

HP OpenVMS

OpenVMS Cluster 構成ガイド

AA-RNJ9C-TE.2

2010年10月

OpenVMS Cluster の可用性，スケーラビリティ，システム管理の利便性は，構成，アプリケーション，およびオペレーティング環境に大きく依存しています。本書では，これらの利便性を最大限に生かせるような提案とガイドラインを提供します。

改訂 / 更新情報:

本書は『OpenVMS Cluster 構成ガイド』V8.2 の改訂版です。

ソフトウェア・バージョン:

OpenVMS Integrity V8.4
OpenVMS Alpha V8.4

日本ヒューレット・パカード株式会社

© 2010 Hewlett-Packard Development Company, L.P.

本書の著作権は Hewlett-Packard Development Company, L.P. が保有しており、本書中の解説および図、表は Hewlett-Packard Development Company, L.P. の文書による許可なしに、その全体または一部を、いかなる場合にも再版あるいは複製することを禁じます。

また、本書に記載されている事項は、予告なく変更されることがありますので、あらかじめご承知おきください。万一、本書の記述に誤りがあった場合でも、日本ヒューレット・パカードは一切その責任を負いかねます。

本書で解説するソフトウェア (対象ソフトウェア) は、所定のライセンス契約が締結された場合に限り、その使用あるいは複製が許可されます。

日本ヒューレット・パカードは、弊社または弊社の指定する会社から納入された機器以外の機器で対象ソフトウェアを使用した場合、その性能あるいは信頼性について一切責任を負いかねます。

Intel®、および Itanium®は、米国およびその他の国における Intel Corporation またはその子会社の商標または登録商標です。

Microsoft®および Windows®は、米国における Microsoft Corporation の登録商標です。

OSF, OSF/1、および Motif は、米国およびその他の国における The Open Group の商標です。

その他のすべての商標および登録商標は、それぞれの所有者が保有しています。

原典：『Guidelines for OpenVMS Cluster Configurations』

© 2010 Hewlett-Packard Development Company, L.P.

本書は、日本語 VAX DOCUMENT V 2.1を用いて作成しています。

目次

まえがき	xvii
1 OpenVMS Cluster システム構成の概要	
1.1 OpenVMS Cluster の構成	1-1
1.2 ハードウェアの構成要素	1-2
1.3 ソフトウェア構成要素	1-3
1.3.1 OpenVMS オペレーティング・システム構成要素	1-4
1.3.2 ネットワーキング構成要素	1-6
1.3.3 ストレージ拡張ソフトウェア	1-6
1.3.4 システム管理ソフトウェア	1-7
1.4 OpenVMS Cluster システムの構成	1-7
1.4.1 一般的な構成規則	1-7
2 ビジネス要件とアプリケーション要件の決定	
2.1 ビジネス要件の決定	2-1
2.1.1 予算	2-1
2.1.2 可用性	2-2
2.1.3 スケーラビリティと将来の成長率	2-2
2.1.4 物理的な位置の要件	2-3
2.1.5 セキュリティ	2-3
2.2 アプリケーション要件の決定	2-3
2.2.1 メモリの追加	2-4
2.2.2 プロセッサ・リソース, メモリ・リソース, I/O リソースの分散	2-4
2.2.3 システム管理ツールとユーティリティ	2-5
3 OpenVMS Cluster システムの選択	
3.1 Integrity サーバと Alpha システム	3-1
3.2 システムのタイプ	3-1
3.3 システムの選択	3-2
3.4 可用性について	3-2
3.5 システム仕様	3-2

4	OpenVMS Cluster インターコネクットの選択	
4.1	特徴	4-2
4.2	インターコネクト・タイプの比較	4-2
4.3	複数のインターコネクト	4-3
4.4	混成インターコネクト	4-4
4.5	サポートするインターコネクト	4-4
4.6	Fibre Channel インターコネクト	4-4
4.6.1	利点	4-5
4.6.2	スループット	4-5
4.7	MEMORY CHANNEL インターコネクト (Alpha のみ)	4-5
4.7.1	利点	4-6
4.7.2	スループット	4-7
4.7.3	サポートされているアダプタ	4-7
4.8	SCSI インターコネクト	4-7
4.8.1	OpenVMS Alpha の構成	4-8
4.8.2	2 ノード OpenVMS Integrity サーバの共有 SCSI 構成	4-8
4.8.3	利点	4-10
4.8.4	スループット	4-11
4.8.5	SCSI インターコネクト距離	4-11
4.8.6	サポートされているアダプタ, バス・タイプ, コンピュータ	4-12
4.9	SAS インターコネクト (Integrity のみ)	4-13
4.9.1	利点	4-13
4.9.2	スループット	4-14
4.9.3	サポートされるアダプタ, バス, タイプ, コンピュータ	4-14
4.10	LAN インターコネクト	4-14
4.10.1	複数の LAN アダプタ	4-15
4.10.1.1	複数の LAN パスの負荷分散	4-15
4.10.1.2	LAN パスの可用性の向上	4-16
4.10.2	LAN ベースのクラスタの構成上のガイドライン	4-16
4.10.3	Ethernet (10/100) および Gigabit (1/10) Ethernet の利点	4-17
4.10.4	Ethernet (10/100) および Gigabit (1/10) Ethernet のスループット	4-17
4.10.5	10 Gigabit Ethernet クラスタの構成上のガイドライン	4-18
4.11	Cluster over IP	4-19
4.11.1	構成上のガイドライン	4-20
4.11.2	IP の可用性	4-21
4.11.3	IP の利点	4-21
4.11.4	IP の性能	4-21
5	OpenVMS Cluster ストレージ・サブシステムの選択	
5.1	ストレージ製品の選択肢	5-1
5.1.1	デバイスの選択基準	5-2
5.1.2	インターコネクトとストレージ・デバイスの選択	5-2
5.1.3	設置面積とストレージ・デバイスの選択	5-3
5.2	記憶容量の必要量の決定	5-3
5.2.1	記憶容量の必要量の見積り	5-3
5.2.2	その他のディスク容量の必要量	5-4
5.3	ディスク・パフォーマンス・オブティマイザの決定	5-5

5.3.1	パフォーマンス・オブティマイザ	5-5
5.4	ディスクの可用性の必要要件の決定	5-7
5.4.1	可用性の必要要件	5-7
5.4.2	デバイスとデータの可用性オブティマイザ	5-7
5.5	SAS 方式のストレージ	5-8
5.5.1	ストレージ・デバイス	5-8
5.6	SCSI 方式のストレージ	5-9
5.6.1	サポートされているデバイス	5-9
5.7	Fibre Channel 方式のストレージ	5-9
5.7.1	ストレージ・デバイス	5-9
5.8	ホスト方式のストレージ	5-10
5.8.1	内部バス	5-10
5.8.2	ローカル・アダプタ	5-10
6	SCSI と Fibre Channel ストレージに対するマルチパスの構成	
6.1	マルチパス SCSI サポートの概要	6-2
6.1.1	直接 SCSI から直接 SCSI へのフェールオーバ	6-3
6.1.2	直接 SCSI から MSCP サービス対象へのフェールオーバ (ディスクのみ)	6-5
6.1.3	両方の種類のマルチパス・フェールオーバを組み合わせた構成	6-6
6.2	構成要件と制限事項	6-8
6.3	HSxフェールオーバ・モード	6-10
6.3.1	透過的フェールオーバ・モード	6-11
6.3.2	マルチパス・フェールオーバ・モード (ディスクのみ)	6-12
6.3.3	マルチパス・モードのコントローラ用のポート・アドレス指定	6-14
6.4	パラレル SCSI マルチパス構成 (ディスクのみ)	6-15
6.4.1	透過的フェールオーバ	6-15
6.4.2	マルチパス・フェールオーバと複数のパス	6-16
6.4.3	マルチポート・ストレージ・コントローラによる構成	6-17
6.5	パラレル SCSI マルチパス構成のディスク・デバイス命名	6-20
6.5.1	ノード割り当てクラスの概要	6-21
6.5.2	ポート割り当てクラスの概要	6-21
6.5.3	HSZ 割り当てクラスを使用したデバイス名	6-22
6.6	Fibre Channel マルチパス構成	6-25
6.7	マルチパス構成の実装	6-28
6.7.1	有効なマルチパス構成	6-29
6.7.2	無効なマルチパス構成	6-31
6.7.3	マルチパス・システム・パラメータ	6-31
6.7.4	パス識別子	6-33
6.7.5	パスの表示	6-34
6.7.5.1	SHOW DEVICE/FULL によるパスの表示	6-34
6.7.5.2	SHOW DEVICE/MULTIPATH_SET によるパスの表示	6-36
6.7.6	パス・ポーリング	6-37
6.7.7	現在のパスへの手動切り替え	6-38
6.7.8	OpenVMS によるパス選択	6-39
6.7.9	OpenVMS によるマルチパス・フェールオーバの方法	6-42
6.7.10	直接パスへの自動フェールバック (ディスクのみ)	6-43
6.7.11	パス切り替え候補としてのパスの有効化または無効化	6-45

6.7.12	パフォーマンスについての考慮	6-45
6.7.13	コンソールについての考慮	6-47
7	OpenVMS Cluster ストレージ・インターコネクトとしての Fibre Channel の構成	
7.1	OpenVMS Fibre Channel サポートの概要	7-2
7.2	Fibre Channel 構成のサポート	7-5
7.2.1	Fibre Channel 修正キット	7-8
7.2.2	混成バージョンと混成アーキテクチャのクラスタ・サポート	7-8
7.3	構成例	7-9
7.3.1	デュアル・ポート・ストレージを備えたシングル・ホスト	7-9
7.3.2	1つのデュアル・ポート・ストレージ・コントローラを備えた複数のホスト	7-10
7.3.3	ストレージ・コントローラ冗長性を備えた複数のホスト	7-11
7.3.4	複数の独立したスイッチを備えた複数のホスト	7-12
7.3.5	デュアル・ファブリックを備えた複数のホスト	7-13
7.3.6	より大規模なファブリックを備えた複数のホスト	7-14
7.4	Fibre Channel アドレス, WWID, デバイス名	7-15
7.4.1	Fibre Channel アドレスと WWID	7-15
7.4.2	Fibre Channel デバイス用の OpenVMS 名	7-18
7.4.2.1	Fibre Channel ストレージ・アダプタ名	7-18
7.4.2.2	Fibre Channel パス名	7-19
7.4.2.3	Fibre Channel ディスク・デバイスの識別	7-19
7.5	Fibre Channel テープのサポート	7-22
7.5.1	最小ハードウェア構成	7-22
7.5.2	Fibre Channel テープ・デバイスの命名	7-23
7.5.2.1	テープ・デバイスおよび媒体チェンジャ・デバイスの名前	7-24
7.5.2.2	ワールドワイド識別子 (WWID) の使用	7-25
7.5.2.3	ファイルベースのデバイスの命名	7-26
7.5.3	Fibre Channel テープ・デバイスの管理サポート	7-27
7.5.4	Fibre Channel テープ・デバイスの構成	7-28
7.5.4.1	基本構成手順: 要約	7-29
7.5.4.2	基本構成手順: 詳細	7-30
7.5.4.3	ユーザ指定デバイス名の作成	7-33
7.5.5	既存の Fibre Channel テープ・デバイス名の変更	7-37
7.5.6	Fibre Channel での物理テープ・デバイスの移動	7-37
7.5.7	Fibre Channel 上の NSR のスワップアウト	7-38
7.5.8	Fibre Channel テープ・デバイスへのサービスの提供	7-38
7.5.9	Fibre Channel テープ・デバイスの交換	7-38
7.5.10	Fibre Channel テープ・デバイスの物理位置の決定	7-39
7.5.11	スタンドアロン環境での Fibre Channel テープ・デバイスへのアクセス	7-40
7.5.12	マルチパス・テープのサポート	7-40
7.6	AlphaServer コンソールによる FC の構成 (Alpha のみ)	7-41
7.6.1	コンソールによる FC 構成の確認	7-41
7.6.2	FC ディスクのブートおよびダンプ用のセットアップ	7-44
7.7	OpenVMS Integrity システムでの Fibre Channel ストレージ・デバイスからのブート	7-48
7.7.1	ブート可能ファームウェアのインストール	7-48

7.7.2	ファームウェア・バージョンのチェック	7-50
7.7.3	FC 上のブート・デバイス・バスの構成	7-50
7.8	OpenVMS で使用するためのストレージ・コントローラのセットアップ (Alpha のみ)	7-52
7.8.1	デバイス識別子を CCL 用に設定	7-53
7.8.2	デバイス識別子をディスク・デバイス用に設定	7-54
7.9	共用 FC システム・ディスクを備えたクラスタの作成	7-54
7.9.1	共有 FC ディスクでブートするためのクラスタ・ノードの追加構成 (Integrity のみ)	7-59
7.9.2	オンライン再構成	7-60
7.9.3	HSG ホスト接続テーブルと構成されないデバイス	7-60
7.10	I/O パフォーマンス向上のための割り込みコアレス機能 (Alpha のみ)	7-61
7.11	Fast Path の使用	7-62
7.12	FIBRE_SCAN ユーティリティによるデバイス情報の表示	7-63
8	可用性を目的とした OpenVMS Cluster の構成	
8.1	可用性の要件	8-1
8.2	OpenVMS Cluster による可用性の提供方法	8-1
8.2.1	ストレージ領域までの共用アクセス	8-2
8.2.2	構成要素の冗長性	8-2
8.2.3	フェールオーバの仕組み	8-2
8.2.4	関連ソフトウェア製品	8-4
8.3	高い可用性を備えた OpenVMS Cluster 構成の実現手法	8-4
8.3.1	可用性の強化手法	8-4
8.4	高い可用性を備えた OpenVMS Cluster の保守	8-5
8.4.1	可用性の保守の手法	8-5
8.5	LAN OpenVMS Cluster における可用性	8-6
8.5.1	構成要素	8-8
8.5.2	長所	8-8
8.5.3	短所	8-9
8.5.4	可用性の主な強化手法	8-9
8.6	マルチ LAN の構成	8-9
8.6.1	MOP サーバの選択	8-10
8.6.2	2 つの LAN セグメントの構成	8-10
8.6.3	3 つの LAN セグメントの構成	8-12
8.7	Cluster over IP における可用性	8-13
8.7.1	構成要素	8-14
8.7.2	長所	8-14
8.7.3	可用性と性能の主な強化手法	8-14
8.8	MEMORY CHANNEL OpenVMS Cluster における可用性	8-15
8.8.1	構成要素	8-15
8.8.2	長所	8-16
8.8.3	短所	8-16
8.8.4	可用性の主な強化手法	8-17
8.9	サテライトを持つ OpenVMS Cluster における可用性	8-17
8.9.1	構成要素	8-19
8.9.2	長所	8-19

8.9.3	短所	8-19
8.9.4	可用性の主な強化手法	8-19
8.10	マルチサイト OpenVMS Cluster システム	8-20
8.10.1	構成要素	8-20
8.10.2	長所	8-21
8.11	ディザスタ・トレラントな OpenVMS Cluster 構成	8-22
9	スケーラビリティを目的とした OpenVMS Cluster の構成	
9.1	スケーラビリティとは?	9-1
9.1.1	スケーラブルな次元	9-2
9.2	スケーラビリティの高い OpenVMS Cluster の構成手法	9-4
9.2.1	スケーラビリティの手法	9-4
9.3	SCSI OpenVMS Cluster におけるスケーラビリティ (Alpha のみ)	9-5
9.3.1	3 ノード Fast Wide SCSI Cluster	9-6
9.3.2	4 ノード Ultra SCSI ハブ構成	9-7
9.4	サテライトを備えた OpenVMS Cluster のスケーラビリティ (Alpha, VAX, およ び Integrity)	9-9
9.4.1	6 サテライト OpenVMS Cluster	9-9
9.4.2	2 ブート・ノードを備えた 6 サテライト OpenVMS Cluster	9-10
9.4.3	2 LAN セグメントを備えた 12 サテライト LAN OpenVMS Cluster	9-11
9.4.4	サイト間リンクを備えた 45 サテライト OpenVMS Cluster	9-13
9.4.5	ハイパワー・ワークステーション OpenVMS Cluster (1995 年のテク ノロジー)	9-14
9.4.6	ハイパワー・ワークステーション OpenVMS Cluster (2004 年のテク ノロジー)	9-16
9.4.7	サテライトを備えた OpenVMS Cluster のガイドライン	9-17
9.4.8	拡張 LAN 構成のガイドライン	9-17
9.4.9	OpenVMS Cluster のシステム・パラメータ	9-19
9.5	Cluster over IP におけるスケーラビリティ	9-20
9.5.1	複数ノードの IP ベースクラスタ・システム	9-20
9.5.2	IP ベース・クラスタの構成上のガイドライン	9-21
9.6	I/O のスケーリング	9-21
9.6.1	MSCP サービス対象のストレージへのアクセス	9-22
9.6.2	ディスク・テクノロジー	9-23
9.6.3	読み込み/書き込みの比率	9-23
9.6.4	I/O サイズ	9-24
9.6.5	キャッシュ	9-24
9.6.6	"ホット"・ファイルの管理	9-25
9.6.7	ボリューム・シャドウイング	9-25
10	OpenVMS Cluster システム管理の手法	
10.1	シンプルな構成と複雑な構成	10-1
10.2	システム・ディスクの手法	10-2
10.2.1	シングル・システム・ディスク	10-2
10.2.2	マルチ・システム・ディスク	10-3
10.2.3	マルチ・システム・ディスク OpenVMS Cluster	10-4
10.2.4	OpenVMS Cluster システムの分割	10-6

10.2.5	まとめ：シングル・システム・ディスク対マルチ・システム・ディスク	10-7
10.3	OpenVMS Cluster 環境の手法	10-7
10.3.1	共通環境	10-8
10.3.2	別々の共通ディスクに環境ファイルを配置	10-8
10.3.3	マルチ環境	10-9
10.4	その他のマルチ環境の手法	10-10
10.4.1	複数の SYSUAF.DAT ファイルの使用	10-10
10.4.2	複数のキュー・マネージャの使用	10-10
10.5	クォーラムの手法	10-11
10.5.1	クォーラム手法のオプション	10-11
10.6	状態遷移の手法	10-12
10.6.1	状態遷移の処理	10-13
10.7	マルチ・バージョンの移行サポートと保証サポート	10-14

A OpenVMS Cluster インターコネクトとしての SCSI

A.1	この付録で使用する表記法	A-1
A.1.1	SCSI ANSI 標準	A-2
A.1.2	図中で使用する記号	A-2
A.2	SCSI ストレージのアクセス	A-2
A.2.1	OpenVMS Cluster システムにおけるシングル・ホスト SCSI アクセス	A-3
A.2.2	OpenVMS Cluster システムにおけるマルチホスト SCSI アクセス	A-3
A.3	構成上の要件とハードウェア・サポート	A-4
A.3.1	構成上の要件	A-4
A.3.2	ハードウェアのサポート	A-6
A.4	SCSI インターコネクトの概念	A-7
A.4.1	デバイス数	A-7
A.4.2	パフォーマンス	A-8
A.4.3	接続距離	A-9
A.4.4	ケーブル配線と終端	A-10
A.5	SCSI OpenVMS Cluster ハードウェア構成	A-11
A.5.1	アドオン SCSI アダプタを使用するシステム	A-12
A.5.1.1	アドオン SCSI アダプタによる基本的なシステムの構築	A-12
A.5.1.2	エンクロージャを追加または物理的距離を延長したシステム、または、HSZ コントローラ付きのシステム	A-14
A.5.1.3	ディファレンシャル・ホスト・アダプタを使用するシステムの構築	A-18
A.6	インストール	A-21
A.6.1	手順 1: SCSI 接地要件の確認	A-22
A.6.2	手順 2: SCSI ノード ID の構成	A-22
A.6.2.1	マルチホスト SCSI バスにおけるデバイス ID の構成	A-23
A.6.2.2	シングル・ホスト SCSI バスにおけるデバイス ID の構成	A-24
A.6.3	手順 3: 電源入力と SCSI デバイスの検査	A-25
A.6.4	手順 4: SCSI コンソール・パラメータの表示と設定	A-26
A.6.5	手順 5: OpenVMS オペレーティング・システムのインストール	A-29
A.6.6	手順 6: その他のシステムの構成	A-29
A.7	補足情報	A-29

A.7.1	OpenVMS Cluster 構成コマンド・プロシージャの実行.....	A-29
A.7.2	マルチホスト SCSI 環境におけるエラー・レポートと OPCOM メッセージ.....	A-32
A.7.2.1	SCSI バスのリセット.....	A-32
A.7.2.2	SCSI タイムアウト.....	A-33
A.7.2.3	マウント検査.....	A-33
A.7.2.4	シャドウ・ボリューム処理.....	A-34
A.7.2.5	マルチホスト SCSI 環境で表示される OPCOM メッセージ.....	A-34
A.7.2.6	エラー・ログの基本.....	A-34
A.7.2.7	マルチホスト SCSI 環境におけるエラー・ログ・エントリ.....	A-35
A.7.3	制限事項と既知の問題.....	A-36
A.7.4	トラブルシューティング.....	A-39
A.7.4.1	終端問題.....	A-39
A.7.4.2	正しくない構成が原因のブート障害やマウント障害.....	A-39
A.7.4.2.1	ブートストラップ・プロセス間のバグチェック.....	A-39
A.7.4.2.2	デバイス構成上の障害.....	A-40
A.7.4.2.3	マウント障害.....	A-40
A.7.4.3	接地.....	A-41
A.7.4.4	インターコネクトの長さ.....	A-41
A.7.5	SCSI アービトレーション上の注意.....	A-41
A.7.5.1	マルチディスク環境におけるアービトレーション問題.....	A-42
A.7.5.2	アービトレーション問題の対策.....	A-42
A.7.5.3	アービトレーションとバス・アイソレータ.....	A-43
A.7.6	OpenVMS Cluster システムが実行中の SCSI デバイスの取り外しや取り付け.....	A-44
A.7.6.1	ホット・プラグを説明するための用語.....	A-44
A.7.6.2	ホット・プラグの規則.....	A-45
A.7.6.3	デバイスまたはセグメントを確実に非アクティブにするための手順.....	A-48
A.7.6.4	StorageWorks SBB ディスクのホット・プラグの手順.....	A-49
A.7.6.5	HSZxxのホット・プラグ手順.....	A-49
A.7.6.6	ホスト・アダプタのホット・プラグ手順.....	A-50
A.7.6.7	DWZZxコントローラのホット・プラグ手順.....	A-51
A.7.7	マルチホスト SCSI OpenVMS Cluster システムで使用されるデバイスの OpenVMS 要件.....	A-53
A.7.8	接地要件.....	A-54

B MEMORY CHANNEL 技術概要

B.1	製品概要.....	B-1
B.1.1	MEMORY CHANNEL の特長.....	B-2
B.1.2	MEMORY CHANNEL バージョン 2.0 の特長.....	B-2
B.1.3	ハードウェアの構成要素.....	B-3
B.1.4	高可用構成のためのバックアップ・インターコネクト.....	B-4
B.1.5	ソフトウェアの要件.....	B-5
B.1.5.1	メモリの要件.....	B-5
B.1.5.2	大容量メモリ・システムにおける NPAGEVIR パラメータの使用.....	B-5
B.1.6	構成.....	B-5
B.1.6.1	構成サポート.....	B-8
B.2	技術概要.....	B-10
B.2.1	従来のネットワークや SMP との比較.....	B-10

B.2.2	OpenVMS Cluster アーキテクチャの MEMORY CHANNEL	B-12
B.2.3	MEMORY CHANNEL アドレス指定	B-12
B.2.4	MEMORY CHANNEL の実装	B-15

C マルチサイト OpenVMS Cluster

C.1	マルチサイト OpenVMS Cluster システムとは?	C-1
C.1.1	ATM, DS3, FDDI, [D]WDM サイト間リンク	C-2
C.1.2	マルチサイト OpenVMS Cluster システムの長所	C-3
C.1.3	一般的な構成ガイドライン	C-4
C.2	Cluster over IP によるマルチサイト OpenVMS Cluster システムの構成	C-4
C.3	FDDI によるマルチサイト OpenVMS Cluster システムの構成	C-5
C.4	WAN サービスによるマルチサイト OpenVMS Cluster システムの構成	C-6
C.4.1	ATM 通信サービス	C-6
C.4.2	DS3 通信サービス (T3 通信サービス)	C-7
C.4.3	FDDI-to-WAN ブリッジ	C-7
C.4.4	OpenVMS Cluster システムにおける ATM と DS3 の構成ガイドライン	C-8
	C.4.4.1 要件	C-9
	C.4.4.2 推奨値	C-9
C.4.5	可用性の考慮	C-12
C.4.6	仕様	C-12
C.5	マルチサイトにおける OpenVMS Cluster システムの管理	C-15
C.5.1	手法とツール	C-15
C.5.2	パフォーマンスの監視	C-16

索引

例

6-1	パスの手動切り替え時のメッセージ	6-38
6-2	他のノードがパス切り替えを検出したときに表示されるメッセージ	6-39
7-1	wwidmgr -show wwid の使用	7-42
7-2	wwidmgr -show wwid -full の使用	7-43
7-3	widmgr -quickset の使用	7-45
7-4	FC システム・ディスクによるブート・シーケンス	7-47
7-5	スタンドアロン FC ノードにおけるクラスタリングの有効化	7-55
7-6	共用 FC システム・ディスクがあるクラスタへノードを追加する	7-57
A-1	SHOW DEVICE コマンドの出力例	A-25
A-2	SCSI クラスタにノードを追加	A-31



1	OpenVMS Cluster システムの構成要素と機能	xviii
1-1	ハードウェアとオペレーティング・システム構成要素	1-5
4-1	2 ノードの OpenVMS Integrity クラスタ・システム	4-10
4-2	ポイント・ツー・ポイント 10 Gigabit Ethernet OpenVMS Cluster	4-18
4-3	スイッチ型 10 Gigabit Ethernet OpenVMS Cluster	4-18
6-1	マルチバス・フェールオーバの構成	6-4
6-2	1 つのインターコネクトによる直接 SCSI から MSCP のサービス対象構成	6-5
6-3	インターコネクトが 2 つの直接 SCSI から MSCP サービス対象構成	6-7
6-4	透過モードにおけるストレージ・サブシステム	6-11
6-5	マルチバス・モードのストレージ・サブシステム	6-13
6-6	マルチバス・モードの平行 SCSI コントローラのポート・アドレス指定	6-14
6-7	マルチバス・モードの Fibre Channel コントローラのポート・アドレス指定	6-15
6-8	透過的フェールオーバ付きの平行 SCSI 構成	6-16
6-9	マルチバス・フェールオーバと複数のバスによる平行 SCSI 構成	6-17
6-10	透過モードによるシングル・インターコネクトのマルチポート・平行 SCSI 構成	6-18
6-11	透過モードによる複数バスのマルチポート・平行 SCSI 構成	6-19
6-12	マルチバス・モードによる複数バスのマルチポート・平行 SCSI 構成	6-20
6-13	ノード割り当てクラスで命名したデバイス	6-21
6-14	ポート割り当てクラスを使用したデバイス名	6-22
6-15	HSZ 割り当てクラスを使用したデバイス名	6-23
6-16	デュアル・ポート・ストレージ・コントローラが 2 つとデュアル・ポート MDR が 1 つ、およびバスが 2 本のシングル・ホスト	6-25
6-17	デュアル・ポート・ストレージ・コントローラが 2 つとデュアル・ポート MDR が 1 つ、およびバスが 4 本のシングル・ホスト	6-27
6-18	デュアル・ポート・ストレージ・コントローラ 2 つと、デュアル・ポート MDR 1 つと、バスが 4 本の 2 つのホスト	6-28
6-19	共用バスと共用ストレージ・コントローラを持つ 2 つのホスト	6-29
6-20	共用、マルチポート・ストレージ・コントローラを持つ 2 つのホスト	6-30
6-21	無効なマルチバス構成	6-31
6-22	Fibre Channel バスの命名	6-33
6-23	複数の直接バスによる構成	6-34
7-1	スイッチ・トポロジ (論理図)	7-3
7-2	スイッチ・トポロジ (実体図)	7-4
7-3	MSA 1000 を使用したアービトラート型トポロジ	7-5
7-4	1 つのデュアル・ポート・ストレージ・コントローラを備えたシングル・ホスト	7-9
7-5	1 つのデュアル・ポート・ストレージ・コントローラを備えた複数のホスト	7-10
7-6	ストレージ・コントローラ冗長性を備えた複数のホスト	7-11
7-7	複数の独立したスイッチを備えた複数のホスト	7-12
7-8	デュアル・ファブリックを備えた複数のホスト	7-13
7-9	より大規模なデュアル・ファブリックを備えた複数のホスト	7-14
7-10	4 つのファブリックを備えた複数のホスト	7-15
7-11	Fibre Channel ホスト・アドレスとポート・アドレス	7-16

7-12	Fibre Channel ホストとポートの WWID とアドレス	7-17
7-13	Fibre Channel イニシエータ名とターゲット名	7-19
7-14	Fibre Channel ディスク・デバイスの命名	7-20
8-1	LAN OpenVMS Cluster システム	8-7
8-2	2 LAN セグメントの OpenVMS Cluster 構成	8-11
8-3	3 LAN セグメントの OpenVMS Cluster 構成	8-12
8-4	ロジカル LAN フェールオーバー IP OpenVMS Cluster システム	8-13
8-5	MEMORY CHANNEL クラスタ	8-15
8-6	サテライトを持つ OpenVMS Cluster	8-18
8-7	WAN リンクで結合されたマルチサイト OpenVMS Cluster 構成	8-20
9-1	OpenVMS Cluster 拡張次元	9-2
9-2	3 ノード Fast Wide SCSI Cluster	9-6
9-3	4 ノード Ultra SCSI ハブ構成	9-8
9-4	6 サテライト LAN OpenVMS Cluster	9-10
9-5	2 ブート・ノードを備えた 6 サテライト LAN OpenVMS Cluster	9-11
9-6	2 LAN セグメントを備えた 12 サテライト OpenVMS Cluster	9-12
9-7	サイト間リンクを備えた 45 サテライト OpenVMS Cluster	9-13
9-8	ハイパワー・ワークステーション・サーバの構成 (1995 年)	9-15
9-9	ハイパワー・ワークステーション・サーバの構成 (2004 年)	9-16
9-10	複数ノードの IP ベース・クラスタ・システム	9-20
9-11	直接アクセスと MSCP サービスによるアクセスの比較	9-22
9-12	ホット・ファイルの分散	9-25
10-1	シングル・システム・ディスクのシンプルな LAN OpenVMS Cluster	10-2
10-2	共通環境におけるマルチ・システム・ディスク	10-5
10-3	マルチ環境 OpenVMS Cluster	10-9
A-1	図中で使用する記号の意味	A-2
A-2	共用 SCSI アクセスによる可用性が高いサーバ	A-4
A-3	最大スタブ長	A-11
A-4	概念図: 基本的な SCSI システム	A-13
A-5	構成例: AlphaServer 1000, KZPAA アダプタ, BA350 エンクロージャ使用の基本的な SCSI システム	A-14
A-6	概念図: DWZZA を利用した物理的な距離の延長またはエンクロージャの追加	A-15
A-7	構成例: DWZZA を利用した物理的な距離の延長またはエンクロージャの追加	A-16
A-8	構成例: SCSI バス上の 3 ホスト	A-17
A-9	構成例: ディファレンシャル・ホスト・アダプタ (KZPSA) を使用した SCSI システム	A-19
A-10	概念図: SCSI ハブを使用した SCSI システム	A-20
A-11	構成例: SCSI ハブ構成における SCSI システム	A-21
A-12	SCSI アクセスの割り当てクラスの設定	A-23
A-13	SCSI バス・トポロジ	A-46
A-14	バス・アイソレータのホット・プラグ	A-47
B-1	MEMORY CHANNEL ハードウェアの構成要素	B-3
B-2	4 ノード MEMORY CHANNEL クラスタ	B-4
B-3	仮想ハブ MEMORY CHANNEL クラスタ	B-4
B-4	MEMORY CHANNEL ベース・クラスタと SCSI ベース・クラスタ	B-6

B-5	MEMORY CHANNEL の CI 方式と SCSI 方式のクラスタ	B-7
B-6	MEMORY CHANNEL DSSI 方式のクラスタ	B-8
B-7	OpenVMS Cluster アーキテクチャと MEMORY CHANNEL	B-12
B-8	物理メモリと I/O アドレス領域	B-13
B-9	MEMORY CHANNEL バス・アーキテクチャ	B-14
C-1	フィラデルフィアとワシントン間のサイト間リンク	C-2
C-2	リモート・サテライトを備えたマルチサイト OpenVMS Cluster 構成	C-4
C-3	Cluster over IP によるマルチサイト OpenVMS Cluster 構成	C-5
C-4	ATM/SONET OC-3 サービス	C-7
C-5	DS3 Service	C-7
C-6	DS3 で結合したマルチサイト OpenVMS Cluster 構成	C-8

表

1	本書の構成	xviii
1-1	OpenVMS Cluster システムの利点	1-2
1-2	OpenVMS Cluster システムのハードウェア構成要素	1-2
1-3	オペレーティング・システム構成要素	1-4
1-4	OpenVMS ネットワーキング構成要素	1-6
2-1	可用性の要件	2-2
2-2	アプリケーション要件の決定	2-3
2-3	アプリケーション・タイプ別のリソース要件	2-5
2-4	HP のシステム管理ツール	2-5
4-1	インターコネクトの特徴	4-2
4-2	インターコネクト・タイプの比較	4-3
4-3	最大転送速度, 単位 MB/秒	4-11
4-4	最大 SCSI インターコネクト距離	4-11
4-5	SCSI アダプタ	4-12
4-6	最大転送速度, 単位 MB/秒	4-14
4-7	SAS アダプタ	4-14
4-8	OpenVMS Clustser 用にサポートされている LAN インターコネクト	4-15
5-1	インターコネクト, および接続するストレージ・デバイス	5-2
5-2	ディスク容量の見積り	5-3
5-3	ディスク・パフォーマンス・オプティマイザ	5-5
5-4	ストレージの可用性オプティマイザ	5-7
6-1	マルチパス SCSI と FC 構成の要件	6-8
6-2	マルチパス FC と SCSI 構成の制限事項	6-9
6-3	マルチパス・システム・パラメータ	6-32
6-4	SHOW DEVICE/FULL マルチパス用語	6-36
7-1	Fibre Channel の特徴と OpenVMS のサポート	7-2
7-2	FC PERFORMANCE コマンド修飾子	7-64
8-1	可用性の要件	8-1
8-2	フェールオーバーのメカニズム	8-3
8-3	可用性を強化する製品	8-4
8-4	可用性の強化手法	8-4
8-5	可用性の保守手法	8-6

9-1	OpenVMS Cluster におけるスケーラブルな次元.....	9-2
9-2	スケーラビリティの手法.....	9-4
9-3	拡張 LAN 構成ガイドライン.....	9-18
9-4	OpenVMS Cluster システム・パラメータ.....	9-19
9-5	ディスク・テクノロジーのまとめ.....	9-23
9-6	キャッシュの種類.....	9-24
10-1	マルチ・システム・ディスクの長所.....	10-3
10-2	マルチ・システム・ディスクの使用方法.....	10-6
10-3	シングル・システム・ディスクとマルチ・システム・ディスクの比較.....	10-7
10-4	クォーラムの手法.....	10-12
10-5	OpenVMS Cluster ポーリング・パラメータ.....	10-13
A-1	SCSI マルチホスト OpenVMS Cluster 構成の要件.....	A-5
A-2	SCSI OpenVMS Cluster システムでサポートされているハードウェア.....	A-7
A-3	最大データ転送速度 (MB/s).....	A-9
A-4	SCSI インターコネクットの最大距離.....	A-10
A-5	SCSI OpenVMS Cluster システムのインストール手順.....	A-22
A-6	SCSI 環境パラメータ.....	A-26
A-7	追加ノードのインストール手順.....	A-30
A-8	正しい接地のための確認手順.....	A-55
B-1	MEMORY CHANNEL 構成のサポート.....	B-9
B-2	SMP, MEMORY CHANNEL, 通常のネットワークの比較.....	B-11
B-3	MEMORY CHANNEL ページ属性.....	B-15
C-1	DS3 プロトコル・オプション.....	C-10
C-2	Bellcore と OpenVMS Cluster の要件と目標値の用語.....	C-13
C-3	OpenVMS Cluster DS3 と SONET OC3 エラー・パフォーマンスの要件.....	C-14

まえがき

本書は、ビジネス、アプリケーション、コンピューティングのさまざまなニーズに応じた OpenVMS Cluster 構成を設計するための解説書です。

システム、インターコネクト、ストレージ・デバイス、およびソフトウェアの選択に必要な情報をまとめています。また、上記の構成要素を組み合わせることで、可用性、スケーラビリティ、パフォーマンス、およびシステム管理の利便性を実現するのにも役立ちます。

注意

本書の説明は Integrity サーバ・システムと AlphaServer システムの組み合わせに対してのみ適用されます。VAX システムと AlphaServer システムの組み合わせについては、以前のバージョンのマニュアルを参照してください。

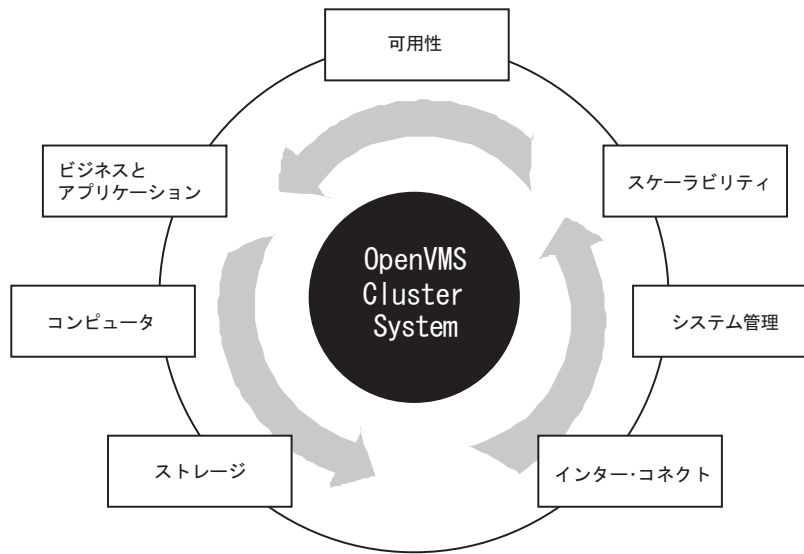
対象読者

本書の対象は、OpenVMS Cluster 製品のユーザや販売代理店、および、OpenVMS クラスタ・システムを構成する管理者であり、コンピュータと OpenVMS Cluster の基本的概念を理解していることが前提です。

本書について

図 1 に示すように OpenVMS Cluster システムには、多くの構成要素と機能がありますが、実質的には 1 つのシステムとして機能します。

図 1 OpenVMS Cluster システムの構成要素と機能



ZK7062AGE

クラスタをフルに活用するには、OpenVMS Cluster 構成の構成要素と機能を理解する必要があります。本書では、表 1 の構成に従い、クラスタの概念について説明します。

表 1 本書の構成

参照先	章タイトル	目的
第 1 章	OpenVMS Cluster システム構成の概要	OpenVMS Cluster のハードウェア、ソフトウェア、および全体的な概念について説明します。
第 2 章	ビジネス要件とアプリケーション要件の決定	ビジネスとアプリケーションのニーズの解析と、そのニーズをどのようにクラスタに適用できるかを説明します。
第 3 章	OpenVMS Cluster システムの選択	コンピュータの要件を紹介し、選択肢の選び方について説明します。
第 4 章	OpenVMS Cluster インターコネクトの選択	クラスタ・インターコネクトを紹介し、選択肢の選び方について説明します。
第 5 章	OpenVMS Cluster ストレージ・サブシステムの選択	ストレージの要件を紹介し、選択肢の選び方について説明します。
第 6 章	SCSI と Fibre Channel ストレージに対するマルチパスの構成	Parallel SCSI インターコネクトや Fibre Channel インターコネクトでストレージまでの複数のパスを構成し、可用性を高める方法について説明します。
第 7 章	OpenVMS Cluster ストレージ・インターコネクトとしての Fibre Channel の構成	ストレージ・インターコネクトとしてファイバ・チャンネルを備えた OpenVMS Cluster の構成方法について説明します。
第 8 章	可用性を目的とした OpenVMS Cluster の構成	クラスタ・システムの可用性の改善方法について説明します。

(次ページに続く)

表 1 (続き) 本書の構成

参照先	章タイトル	目的
第 9 章	スケーラビリティを目的とした OpenVMS Cluster の構成	OpenVMS Cluster システムのさまざまな次元で OpenVMS Cluster システムを拡張する方法と、その拡張に伴う交換要件について説明します。
第 10 章	OpenVMS Cluster システム管理の手法	OpenVMS Cluster システムの管理に関係のある項目とその扱い方について説明します。
付録 A	OpenVMS Cluster インターコネクトとしての SCSI	複数のホストが SCSI デバイスとのアクセスを直接共用できるよう、シングル SCSI バスに複数のホストとストレージを構成する方法について説明します。
付録 B	MEMORY CHANNEL 技術概要	MEMORY CHANNEL インターコネクト使用の理由、タイミング、方法について説明します。
付録 C	マルチサイト OpenVMS Cluster	マルチサイト OpenVMS Cluster システムの長所、構成オプション、要件、および管理について説明します。

関連資料

本書で説明されているトピックについての詳しい情報は、下記のドキュメントを参照してください。

- 『OpenVMS Cluster Software Software Product Description (SPD 29.78.xx)』
- 『OpenVMS Cluster システム』
- 『Volume Shadowing for OpenVMS 説明書』
- 『OpenVMS Alpha パーティショニングおよび Galaxy ガイド』
- 『Availability Manager User's Guide』
- 『TCP/IP Services for OpenVMS Release Notes』 および関連ドキュメント
- 『DECnet for OpenVMS Networking Manual』
- 『Guide to OpenVMS File Applications』
- 『OpenVMS システム・セキュリティ・ガイド』
- 『OpenVMS システム管理者マニュアル』

OpenVMS の製品およびサービスについての情報は、以下の URL の OpenVMS の Web サイトをご覧ください。

<http://www.hp.com/jp/openvms>

または

<http://www.hp.com/go/openvms>

本書で使用する表記法

VMScluster システムは、現在、OpenVMS Cluster システムを指します。本書では、特に指定のないかぎり、OpenVMS Cluster またはクラスタは VMScluster と同意語です。

本書では、DECwindows および DECwindows Motif はすべて DECwindows Motif for OpenVMS ソフトウェアを意味します。

また、本書では次の表記法も使用しています。

表記法	意味
Ctrl/x	Ctrl/x という表記は、Ctrl キーを押しながら別のキーまたはポインティング・デバイス・ボタンを押すことを示します。
PF1 x	PF1 x という表記は、PF1 に定義されたキーを押してから、別のキーまたはポインティング・デバイス・ボタンを押すことを示します。
Return	例の中で、キー名が四角で囲まれている場合には、キーボード上でそのキーを押すことを示します。テキストの中では、キー名は四角で囲まれていません。 HTML 形式のドキュメントでは、キー名は四角ではなく、括弧で囲まれています。
...	例の中の水平方向の反復記号は、次のいずれかを示します。 <ul style="list-style-type: none">• 文中のオプションの引数が省略されている。• 前出の 1 つまたは複数の項目を繰り返すことができる。• パラメータや値などの情報をさらに入力できる。
.	垂直方向の反復記号は、コードの例やコマンド形式の中の項目が省略されていることを示します。このように項目が省略されるのは、その項目が説明している内容にとって重要ではないからです。
()	コマンドの形式の説明において、括弧は、複数のオプションを選択した場合に、選択したオプションを括弧で囲まなければならないことを示しています。
[]	コマンドの形式の説明において、大括弧で囲まれた要素は任意のオプションです。オプションは 1 つまたは複数を選択しても、あるいは 1 つも選択しなくても構いません (ただし、OpenVMS ファイル指定のディレクトリ名の構文や、割り当て文の部分文字列指定の構文の中では、大括弧に囲まれた要素は省略できません)。

表記法	意味
[]	コマンド形式の説明では、括弧内の要素を分けている垂直棒線はオプションを 1 つまたは複数選択するか、または何も選択しないことを意味します。
{ }	コマンドの形式の説明において、中括弧で囲まれた要素は必須オプションです。いずれか 1 つのオプションを指定しなければなりません。
太字	太字のテキストは、新しい用語、引数、属性、要件を示しています。また、変数を示す場合にも使用されています。
<i>italic text</i>	イタリック体のテキストは、重要な情報を示します。また、システム・メッセージ (たとえば内部エラー <i>number</i>)、コマンド・ライン (たとえば <i>PRODUCER=name</i>)、コマンド・パラメータ (たとえば <i>device-name</i>) などの変数を示す場合にも使用されます。
UPPERCASE TEXT	英大文字のテキストは、コマンド、ルーチン名、ファイル名、ファイル保護コード名、システム特権の短縮形を示します。
Monospace type	モノスペース・タイプの文字は、コード例および会話型の画面表示を示します。 C プログラミング言語では、テキスト中のモノスペース・タイプの文字は、キーワード、別々にコンパイルされた外部関数およびファイルの名前、構文の要約、または例に示される変数または識別子への参照などを示します。
-	コマンド形式の記述の最後、コマンド・ライン、コード・ラインにおいて、ハイフンは、要求に対する引数とその後の行に続くことを示します。
数字	特に明記しない限り、本文中の数字はすべて 10 進数です。10 進数以外 (2 進数, 8 進数, 16 進数) は、その旨を明記してあります。

OpenVMS Cluster システム構成の概要

この章では、OpenVMS Cluster の一般的な構成規則とともに、ハードウェア構成要素とソフトウェア構成要素について説明します。

1.1 OpenVMS Cluster の構成

OpenVMS Cluster システムは、OpenVMS システム、ストレージ・サブシステム、インターコネクト、ソフトウェアが 1 つのグループとしてまとまって働く仮想システムです。OpenVMS Cluster システムは、Alpha、VAX、または Intel Itanium プロセッサを搭載した HP Integrity サーバのいずれかの同じアーキテクチャで構成された環境で、OpenVMS が動作するように構成できます。また、OpenVMS Cluster システムは、異なるアーキテクチャを組み合わせたクラスタ環境として実行することもできます。有効な組み合わせは、Alpha と VAX、または Alpha と HP Integrity サーバの 2 種類の組合せです。

OpenVMS Cluster システムでは、各システムが次のことを行います。

- 処理リソース、キュー、データ・ストレージ・デバイスを共有します。
- 個別にブートや障害処理できます。
- OpenVMS オペレーティング・システムを実行します。

また OpenVMS Cluster システムは、1 つのエンティティとして管理されます。

注意

異種混在型のクラスタでは、システム・ディスクおよびシステム・ブート・ブロックごとに 1 つのアーキテクチャのみサポートします。

表 1-1 は、OpenVMS Cluster システムの利点を述べたものです。

OpenVMS Cluster システム構成の概要

1.1 OpenVMS Cluster の構成

表 1-1 OpenVMS Cluster システムの利点

利点	説明
資源の共用	複数のシステムが同じストレージ・デバイスにアクセスできます。これはクラスタ全体でユーザがファイルを共用するためです。アプリケーション、バッチ、およびプリントジョブ処理も複数のシステムに配布できます。共用資源にアクセスするジョブは、どのシステムでも実行できます。
可用性	個々のシステムがシャットダウンしても (ダウンが予定されたものであるかどうかに関係なく)、データやアプリケーションは引き続き利用できます。さまざまな構成により、ディザスタ・トレラントな操作にも対応できるさまざまなレベルの可用性を設定できます。
柔軟性	OpenVMS Cluster コンピューティング環境は、幅広い価格帯とパフォーマンス域にまたがる互換ハードウェアとソフトウェアを提供します。
スケラビリティ	その他のシステムに影響を与えることなく処理資源とストレージ・リソースを追加できます。ハイエンド・マルチプロセッサ・システムから小型ワークステーションまで、各種システムを相互接続でき、ニーズの増大に合わせて簡単に構成できます。パフォーマンスと可用性のレベルは拡張に応じて制御できます。
管理の容易さ	OpenVMS Cluster 管理は、効率的で安全です。OpenVMS クラスタは単独のシステムとして管理されるため、さまざまなタスクも 1 回実行するだけですみます。OpenVMS Cluster では、ユーザ、バッチ、印刷作業の各負荷のバランスが自動的に調整されます。
オープン・システム	IEEE, POSIX, OSF/1, Motif, OSF DCE, ANSI SQL, および TCP/IP の各標準に準拠しているため、OpenVMS Cluster システムはアプリケーションの移植性と相互運用性を提供できます。

1.2 ハードウェアの構成要素

OpenVMS Cluster システムは、システム、インターコネクト、アダプタ、ストレージ・サブシステム、周辺機器など、さまざまなハードウェア構成要素からなります。表 1-2 は、これらの構成要素と例をまとめたものです。サポートする最新のハードウェア一覧については、ソフトウェア仕様書を参照してください。

表 1-2 OpenVMS Cluster システムのハードウェア構成要素

構成要素	説明	例
システム	1 つの処理体として機能する 1 つ以上のプロセッサ、メモリ、入出力 (I/O) アダプタを組み込んだキャビネット。 関連項目：OpenVMS システムの詳細については、第 3 章を参照してください。	OpenVMS Cluster システムには、サポートされている Alpha, VAX, または HP Integrity サーバの各システムを組み込むことができます。

(次ページに続く)

表 1-2 (続き) OpenVMS Cluster システムのハードウェア構成要素

構成要素	説明	例
インターコネクト	<p>OpenVMS Cluster ノード間のハードウェア接続であり、これを通じてノード通信を行います。</p> <p>関連項目： OpenVMS Cluster インターコネクトの詳細については、第 4 章を参照してください。</p>	<p>OpenVMS Cluster システムには、次に示す相互接続を 1 つ以上使用します。</p> <ul style="list-style-type: none"> • Small Computer Systems Interface (SCSI) (ノードからストレージの間のみ) • Serial Attached SCSI (SAS) (ノードからストレージの間のみ) • Fibre Channel (ノードからストレージの間のみ) • Ethernet, Fast Ethernet, Gigabit Ethernet, 10 Gigabit Ethernet <p>サポートするアダプタの一覧については、ソフトウェア仕様書を参照してください。</p>
ストレージ・サブシステム	<p>データを保存するデバイスとデバイスを管理するオプション・コントローラ</p> <p>関連項目： OpenVMS ストレージ・サブシステムの詳細については、第 5 章を参照してください。</p>	<p>ストレージ・サブシステムには、次のものがあります。</p> <ul style="list-style-type: none"> • SCSI あるいは SAS ディスクおよびテープ • ストレージ・アレイ • SCSI あるいは SAS ストレージ・コントローラ • InfoServer システム
アダプタ	<p>OpenVMS Cluster 内のノードをインターコネクトとストレージに接続するデバイス。</p> <p>関連項目： アダプタの詳細については、第 4 章を参照してください。</p>	<p>PCI (Peripheral Component Interconnect) および PCIe (PCI-Express) システムで使用されるアダプタは、次のとおりです。</p> <ul style="list-style-type: none"> • KZPSA (SCSI) • DE435 (Ethernet) • DEGPA (Gigabit Ethernet) • DEGXA (Gigabit Ethernet) • AB287A (10Gigabit Ethernet) • AD385A (10Gigabit Ethernet) • AD339A (Ethernet) (PCIe) • NC360M (Ethernet) (PCIe)

1.3 ソフトウェア構成要素

OpenVMS Cluster システム・ソフトウェアのタイプは、次のように分類できます。

- OpenVMS オペレーティング・システム構成要素
- ネットワーキング構成要素

OpenVMS Cluster システム構成の概要

1.3 ソフトウェア構成要素

- ストレージ拡張ソフトウェア
- システム管理ソフトウェア
- ビジネス・アプリケーション

1.3.1 OpenVMS オペレーティング・システム構成要素

オペレーティング・システムは、ハードウェア構成要素とソフトウェア構成要素、および資源が適切に動作するように管理します。

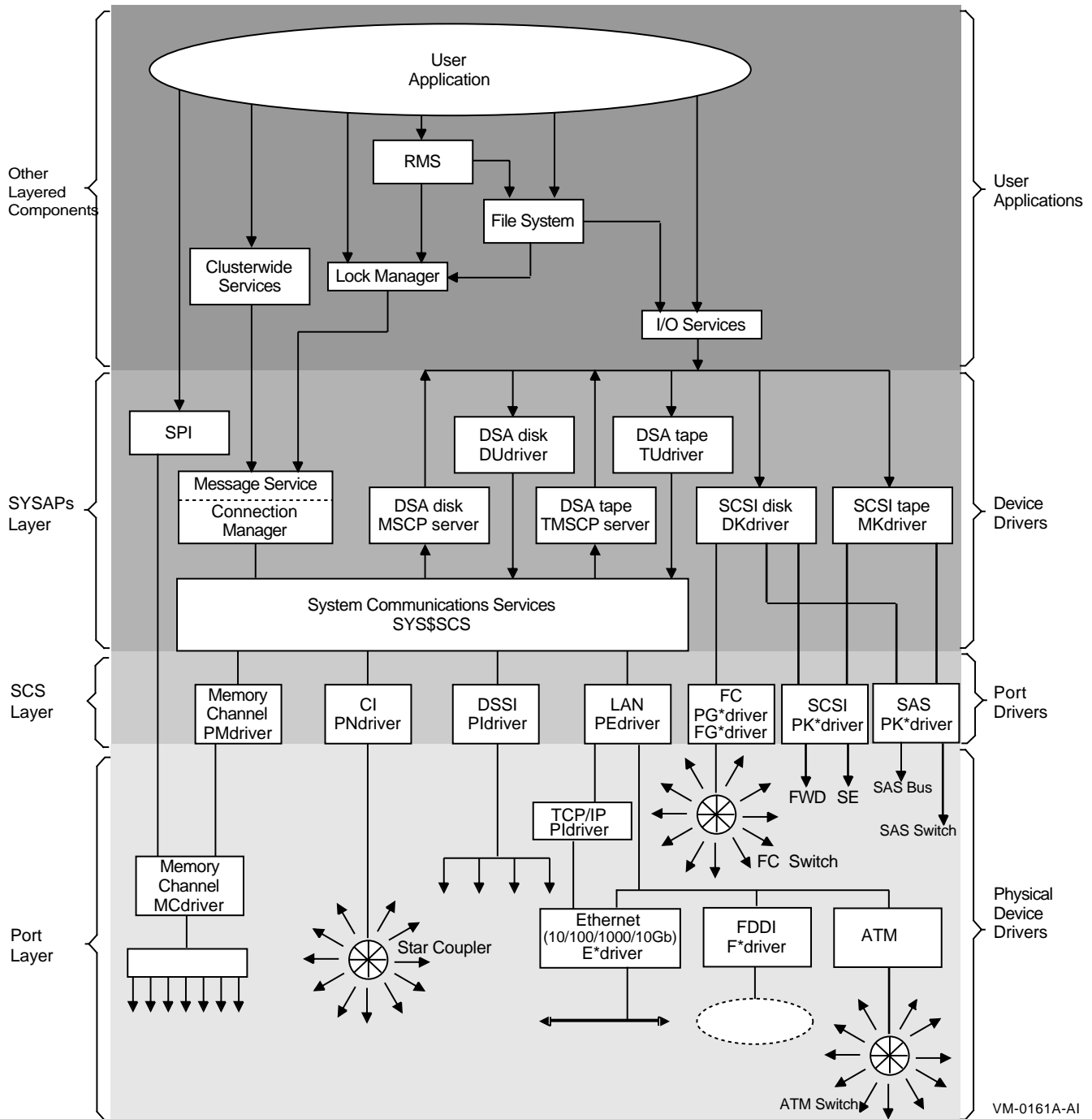
表 1-3 は、OpenVMS Cluster の動作に必要なシステム構成要素です。これらの構成要素は、OpenVMS オペレーティング・システム・ライセンスと OpenVMS Cluster ライセンスがないと使用できません。

表 1-3 オペレーティング・システム構成要素

構成要素	機能
レコード管理サービス (RMS) と OpenVMS ファイル・システム	OpenVMS Cluster 環境のディスクやテープ上の共用読み込みアクセスと書き込みアクセスを提供します。
クラスタ全体のプロセス・サービス	クラスタ全体でプロセスを作成、削除するだけでなく、SHOW SYSTEM や SHOW USERS などの OpenVMS コマンドのクラスタ規模による操作を可能にします。
分散ロック・マネージャ	多数のユーザによる共用資源へのアクセスを同期します。
分散ジョブ・コントローラ	バッチやプリント・キューをクラスタ全体で共用可能にし、資源を最適化します。
接続マネージャ	OpenVMS Cluster メンバのメンバシップとクォーラムを制御します。
SCS (システム通信サービス)	System Communications Architecture (SCA) により、ノード間の OpenVMS Cluster 通信を可能にします。
MSCP サーバ	直接アクセスできるローカル接続ディスクを、OpenVMS Cluster 内の他のシステムにも開放します。
TMSCP サーバ	直接アクセスできるローカル接続テープを、OpenVMS Cluster 内の他のシステムにも開放します。

図 1-1 は、代表的な OpenVMS Cluster システムにおけるハードウェアとオペレーティング・システム構成要素の組み合わせを示したものです。

図 1-1 ハードウェアとオペレーティング・システム構成要素



注意

すべてのインターコネクタが OpenVMS の 3 つのアーキテクチャでサポートされるわけではありません。CI, DSSI, FDDI インターコネクタは, Alpha および VAX システムでサポートされます。Memory Channel および ATM インターコネクタは Alpha システムでのみサポートされます。

1.3.2 ネットワーキング構成要素

表 1-4 は、1 つの OpenVMS Cluster システム・ノードと他の OpenVMS Cluster ノード間の通信と資源を共有するためのオプション・ネットワーキング・ソフトウェアを説明したものです。

表 1-4 OpenVMS ネットワーキング構成要素

オプション・ソフトウェア	機能
DECnet-Plus/DECnet	ネットワーク通信はノード間通信に必要です。
Advanced Server for OpenVMS (Advanced Server) および HP Common Internet File System (CIFS) for OpenVMS (Advanced Server)	PC システムを OpenVMS Cluster システムにリンクするクライアント・サーバ・ネットワーキング・ソフトウェアです。Advanced Server は Alpha システムでのみ利用できます。CIFS は Integrity サーバおよび Alpha システムの両方で利用できます。
HP TCP/IP Services for OpenVMS ソフトウェア	OpenVMS 用の Network File System (NFS) サーバであり、Internet ネットワーキング・プロトコルをサポートします。Cluster over IP ではクラスタ通信に TCP/IP を必要とします。

1.3.3 ストレージ拡張ソフトウェア

オプションのストレージ拡張ソフトウェアは、ストレージ・サブシステムのパフォーマンスや可用性を改善します。

たとえば、次のようなものがあります。

- Volume Shadowing for OpenVMS (RAID レベル 1)
- DECram for OpenVMS (RAM ディスク)
- RAID Software for OpenVMS (RAID レベル 0 アレイ (ディスク・ストライピングをサポート) と RAID レベル 5 アレイ (パリティ付きアクセス・ストライピング))
- Hierarchical Storage Manager (HSM)

OpenVMS ストレージ管理製品に関する最新情報については、次の OpenVMS Web サイトにある OpenVMS 製品のページを参照してください。

<http://www.hp.com/go/openvms>

または

<http://www.hp.com/jp/openvms>

1.3.4 システム管理ソフトウェア

システム管理ソフトウェアは、OpenVMS Cluster システムを管理します。

たとえば、次のようなものがあります。

- HP Archive/Backup System for OpenVMS
- HP Code Management System for OpenVMS
- HP WBEM Services for OpenVMS Integrity

1.4 OpenVMS Cluster システムの構成

OpenVMS Cluster の特徴や利点を活用するには、正しく構成することが重要です。OpenVMS Cluster の理想的な構成では、次の基準を満たすことができます。

- ハードウェア構成要素とソフトウェア構成要素の最善の組み合わせにより、ビジネス要件を満たす。
- ビジネスの最優先分野に戦略的に予算を当て、最大の利益を上げる。
- ビジネス・ニーズの増大や変化に応じて、最新のニーズを満たすとともに投資結果を確保する。

OpenVMS Cluster システムの構成は、さまざまな要因を考慮する必要があるため、慎重に行う必要があります。新しい要因が発生するたびに計画は変更する必要があるかもしれません。設計内容の更新時には、トレードオフに対する利点を評価し、ニーズに最適な決定を下します。

1.4.1 一般的な構成規則

以下の一般規則が、OpenVMS Cluster システムに適用されます。

- Integrity システムと Alpha システムからなる OpenVMS Cluster システムには、96 台 (合計) を超えるシステムを組み込むことはできません。
- Alpha システムと Integrity システムは、同じシステム・ディスクからはブートできません。システム・ディスクはアーキテクチャ固有であり、同じアーキテクチャのシステムでないと共用できません。
- OpenVMS ノードは、他の OpenVMS Cluster ノードと直接通信できるものとします。

共用 (マルチホスト) SCSI バスや共用 (マルチホスト) Fibre Channel インターコネクトを使用する構成は、サポートされている他の OpenVMS Cluster インターコネクトとも構成できるものとします。SCSI バスや Fibre Channel ではノード・ツー・ノード通信ができないからです。

関連項目： SCSI インターコネクットの詳細については、第 4.8 節を参照してください。Fibre Channel インターコネクットの詳細については、第 4.6 節を参照してください。

SAS インターコネクットを使用する構成では、サポートされている他の OpenVMS Cluster インターコネクットを構成する必要もあります。

関連項目： SAS の詳細については、第 4.9 節を参照してください。

- OpenVMS Cluster ノードまたはストレージ・デバイスを 1 つの OpenVMS Cluster システムにまとめて一度に組み込むことはできません。1 つずつ組み込みます。
- DECnet-Plus ソフトウェアは、OpenVMS Cluster 構成には不要です。ただし、DECnet メールボックスによるノード間のプロセス・ツー・プロセス通信が必要な場合、DECnet-Plus が必要です。
- Cluster over IP は TCP/IP Services を必要とします。いずれかのノードでクラスタ通信に IP を必要とする場合、すべてのメンバで IP を有効にする必要があります。

以上の一般的な規則以外に、構成によりさらに詳細なガイドラインが適用されます。本書ではこの後、構成ごとのコンテキストでそれらのガイドラインについて説明します。

ビジネス要件とアプリケーション要件の決定

この章では、OpenVMS Cluster のビジネス要件とアプリケーション要件の決定方法について説明します。

2.1 ビジネス要件の決定

ビジネス要件の種類は、OpenVMS Cluster の構成に関係します。OpenVMS Cluster システムに関する代表的なビジネス要件には以下のものがあります。

- 予算
- 可用性
- スケーラビリティと将来の成長率
- 物理的な位置の要件
- セキュリティ

これらの要件には、スケーラビリティと物理的な位置といったように互いに競合するものがあります。たとえば、OpenVMS Cluster のシステムを拡張したくても設置場所の広さや位置という制約があります。このような状況では、何を最優先し、何を妥協するかを決める必要があります。

2.1.1 予算

ビジネス上の多くの決定事項と同じく、多くの選択肢がコストで決まります。要件の優先順位が決まれば、ビジネス上で最大のニーズがある分野に予算を振り向ける判断ができます。

予算の決定にあたっては、購入コストの他、以下のような初期システム・コストを計画に入れてください。

- サービスと更新
- 電力消費
- 冷却
- システム管理

ビジネス要件とアプリケーション要件の決定

2.1 ビジネス要件の決定

2.1.2 可用性

コンピューティング・システムがどのような可用性であるかを決めます。通常の組織には、表 2-1 に示すように、3 つ (場合によってはそれらは重複します) の大きなカテゴリのいずれかが当てはまります。

表 2-1 可用性の要件

可用性の要件	説明
一般の用途	システムやアプリケーションを利用できずに待機していても、ほとんど、あるいはまったく影響がないビジネスの機能向けです。
24 × 365	年間を通して主要期間または業務時間帯のほとんどにおいて、連続したコンピューティング・サービスを必要とするビジネス機能向けです。最小限のダウン・タイムのみ許容可能です。
耐障害性	可用性の要件が最も厳しく要求されるビジネス機能向けです。このようなビジネスには、地震、洪水、停電などの災害への対策が要求されます。

関連項目：可用性についての説明は、本書の第 8 章を参照してください。

2.1.3 スケーラビリティと将来の成長率

スケーラビリティとは、任意のシステム、ストレージ、インターコネクト・ディメンションにおいて OpenVMS Cluster を拡張できる一方で、当初の構成機器をフルに活用できる能力です。ノード・レベルのスケーラビリティは、ノードのハードウェアとソフトウェアをアップグレードしたり、追加できる能力を指します。OpenVMS Cluster レベルのスケーラビリティとは、処理能力、インターコネクト、ストレージを、ノード全般に渡って強化することで OpenVMS Cluster システム全体のキャパシティを強化できる能力を指します。

HP では、それぞれ異なる処理特性、ストレージ特性、インターコネクト特性を持つフレレンジの Alpha システムおよび Integrity システムを提供しています。適切なレベルにおける投資とは、現在のビジネス要件に対してある程度の余裕を持って対応できるキャパシティを伴った投資を指します。キャパシティの余地は将来の成長率に対応するためですが、現在のニーズに対して余裕がない設計では、OpenVMS Cluster のスケーラビリティに制限が生じるだけでなく、場合によっては縮小しなければならないこともあります。

将来の成長率を考慮して設計すれば、ほとんどの初期投資を活用できるだけでなく、当初の機器を再利用でき、後で無駄なアップグレードを実施しなくて済みます。

関連項目：スケーラビリティの要件の解析については、第 9 章を参照してください。

2.1.4 物理的な位置の要件

物理的な制約は OpenVMS Cluster の構成方法を考える上で重要です。狭いコンピュータ室やオフィス向けのクラスタの設計は、建物全体や数マイルに渡って展開されるクラスタの設計とはまったく異なります。電源や空調の要件も設計を構成する上で影響を及ぼす問題です。

クラスタを設計するときは、将来の物理的な拡張、電源や冷却要件の強化を考慮に入れておきます。

関連項目：複数の拡張ローカル・エリア・ネットワーク (LAN) の構成については、第 8.6 節と第 9.4.8 項を参照してください。

地理的に異なる場所で IP を使用してクラスタを構築する方法については、第 C.2 節を参照してください。

2.1.5 セキュリティ

安全な環境とは、権限がないユーザによるシステム・アクセスを物理的または電子的に制限できる環境を指します。通常のビジネスでは、パフォーマンスの低下がほとんど、あるいはまったくない状態で環境を安全化できます。ただし、セキュリティを最優先すると、利便性、コスト、パフォーマンス面で妥協が必要になることがあります。

関連項目：詳細については、『OpenVMS システム・セキュリティ・ガイド』を参照してください。

2.2 アプリケーション要件の決定

アプリケーションには処理パワー、メモリ、ストレージ、I/O リソースが必要です。アプリケーション要件の決定により、アプリケーションのニーズに合った OpenVMS Cluster システムを設計できます。アプリケーション要件を決定するには、表 2-2 の手順に従ってください。

表 2-2 アプリケーション要件の決定

手順	説明
1	現在実行している、または実行しようとするアプリケーションのリストを作成します。

(次ページに続く)

表 2-2 (続き) アプリケーション要件の決定

手順	説明
2	<p>アプリケーションごとに、プロセッサ要件、メモリ要件、I/O 要件を記述します (詳細については、各アプリケーションのマニュアルを参照してください。)</p> <p>プロセッサ・パワーは、アプリケーションが実行する計算の数に比例します。またノード間やノードとストレージ間のデータ転送に備えて十分なプロセッサ・パワーが必要です。</p> <p>アプリケーションとその他の OpenVMS Cluster 機能に十分なメモリ容量が必要です。メモリ容量に余裕があればすぐれたシステム・パフォーマンスを実現できます。最初はメモリにコストをかけることをお勧めします。</p> <p>I/O パフォーマンス要件は、アプリケーションによって異なります。ノード、インターコネクト、アダプタなどのコンポーネントを選択するときは、各コンポーネント固有の速度をよく検討して最速のコンポーネントを選択し、ボトルネックの発生を防ぎます。</p>
3	<p>すべてのアプリケーションの CPU 要件、メモリ要件、I/O 要件を合算します。この合計値にさらにユーザ要件や周辺機器などの特別要件を加えます。</p>
4	<p>すべてのアプリケーション要件が決まったら、CPU リソース、メモリ・リソース、I/O リソースによる要件が、上記の要件を 20%上回っていることを確認します。</p>

2.2.1 メモリの追加

OpenVMS Cluster でシステムを実行するには、スタンドアロンの場合より約 5%多いメモリが必要になります。この追加メモリは、共用クラスタ・リソース・ベースのサポートに充てられます。共用クラスタ・リソース・ベースは、スタンドアロン構成よりも大きくなります。

メモリを追加すれば、OpenVMS Cluster のノードで、スタンドアロン・システムの場合と同数のユーザやアプリケーションをサポートできます。クラスタ構成の拡張とともに、ノードごとにシステム作業に必要なメモリ容量も増えます。ノードごとのメモリ容量の増加は、クラスタ内で共用されるデータ・レベルとリソース管理の分散状態によって異なるので一定のルールを適用することはできません。ノードが使用頻度の高いリソースのリソース・マネージャの場合、メモリを追加すれば、そのリソースのクラスタ・ユーザのパフォーマンスを強化できます。

関連項目：メモリの要件についてはソフトウェア仕様書を参照してください。

2.2.2 プロセッサ・リソース、メモリ・リソース、I/O リソースの分散

アプリケーションのパフォーマンスは、プロセッサ・リソース、メモリ・リソース、I/O リソースをいかにうまくバランスさせるかによって異なります。アプリケーションによって、これらのリソースの優先順位も異なります。アプリケーション要件をよく検討し、その要件に合わせてこれらのリソースのバランスを図ってください。表 2-3 は、アプリケーション・タイプ別のリソース要件をまとめたものです。

表 2-3 アプリケーション・タイプ別のリソース要件

アプリケーション・タイプ	例	必要要件
一般タイムシェアリング	プログラム開発, ドキュメント作成, オフィス・オートメーション	プロセッサと I/O に集中
データベースの検索と更新およびレポートの表示	トランザクション処理, 資金移動, オンライン発注エントリまたは予約システム	I/O とメモリの集約型
シミュレーション, モデリング, 計算	コンピュータ支援設計, 製造, イメージ処理, グラフィックス・アプリケーション	プロセッサとメモリの集約型

2.2.3 システム管理ツールとユーティリティ

OpenVMS オペレーティング・システムは, OpenVMS Cluster 構成におけるビジネス要件とアプリケーション要件の決定に必要な多くのユーティリティやツールをサポートしています。表 2-4 は, これらの製品の内容をまとめたものです。Disaster Tolerant Cluster Service (DTCS) を除くすべての製品は, OpenVMS オペレーティング・システムに組み込まれ, 各環境に合わせてカスタマイズされます。DTCSの詳細については, 弊社のサービス担当者に問い合わせるか, 次の DTCS の Web サイトを参照してください。

<http://h71000.www7.hp.com/availability/>

表 2-4 HP のシステム管理ツール

ツール	機能
Accounting ユーティリティ	リソースの使用状況を追跡します。
AUTOGEN コマンド・プロシージャ	使用状況に応じてシステム・パラメータの設定を最適化します。
Availability Manager	HP Availability Manager は, 拡張 LAN あるいは WAN 上で 1 つ以上の OpenVMS Integrity システムまたは OpenVMS Alpha システムを監視するためのシステム管理ツールです。情報を集めようとしているノードは同じ拡張 LAN 上にあり, WAN アナライザとともに収集ノードと通信するインタフェースが必要となります。Availability Manager は, 同時に複数の OpenVMS ノードからシステム・データ, プロセス・データを収集し, データを分析しネイティブの Java GUI で出力を表示します。
HP Insight Management Agents for OpenVMS	OpenVMS システム上のデバイスを参照できます。Insight Manager を Microsoft Windows サーバにインストールすると, すべてのプラットフォームを単一の Windows サーバから管理できます。
HP WBEM Services for OpenVMS	WBEM (Web-Based Enterprise Management) を使用すると, 管理アプリケーションによって, 必要な場所で必要なときに, システム情報を取り出し, システム操作を要求できるようになります。これによりユーザは, 複数のプラットフォームおよびオペレーティング・システム間でシステムを一貫性を持って管理することができ, インフラストラクチャを最適化する統合ソリューションによって, 操作効率を大幅に改善することができます。

(次ページに続く)

ビジネス要件とアプリケーション要件の決定

2.2 アプリケーション要件の決定

表 2-4 (続き) HP のシステム管理ツール

ツール	機能
Disaster Tolerant Cluster Services (DTCS)	様々な場所に存在するディザスタ・トレラント・クラスタを構成および管理するための、ツールおよびカスタマイズされたサービスです。
Monitor コーティリティ	基本パフォーマンス・データを提供します。
OpenVMS Data Collector および Performance Analyzer (ECP Data Collector と ECP Performance Analyzer)	これらの ECP 製品を使用するために、追加のコストは必要ありません。ECP Data Collector では、パフォーマンス・データの収集、保管、レポート、およびファイル表示が実行できます。ECP Performance Analyzer は、ボトルネックを特定し、問題の修復方法を示すことで、OpenVMS Cluster システムのパフォーマンス分析とキャパシティ計画を支援します。Data Collector に組み込まれている API により、収集したデータに直接アクセスできます。
Performance Data Collector for OpenVMS (TDC V2)	OpenVMS Version 7.3-2 またはそれ以降が動作している AlphaServer システムのパフォーマンス・データを収集するために使用されます。デフォルトでは、TDC は定期的にデータを収集し、ファイルに格納します。その後、ユーザ・アプリケーションはファイルからデータを取り出すことができます。
Show Cluster コーティリティ	OpenVMS Cluster 構成における処理とパフォーマンスを監視します。
Systems Communications Architecture Control Program (SCACP)	LAN クラスタ通信を監視および管理できます。

Availability Manager , Performance Data Collector for OpenVMS (TDC V2) , および OpenVMS Management Station の詳細については、次の URL を参照してください。

http://h71000.www7.hp.com/openvms/system_management.html

Accounting コーティリティ , AUTOGEN コマンド・プロシージャ , Monitor コーティリティ , Show Cluster コーティリティ , および System Communications Architecture Control Program (SCACP) の詳細については、『OpenVMS システム管理ユーティリティ・リファレンス・マニュアル』を参照してください。

OpenVMS Cluster システムの選択

この章では、ビジネスやアプリケーションの要件に対応できる OpenVMS Cluster 用のシステムを選択について説明します。

3.1 Integrity サーバと Alpha システム

OpenVMS Cluster は、複数の OpenVMS Integrity システムか、OpenVMS Integrity システムと OpenVMS Alpha システムの組み合わせで構築できます。現在サポートしているサーバ・モデルについては OpenVMS のソフトウェア仕様書を参照してください。

- HP OpenVMS Integrity オペレーティング・システム

Intel Itanium アーキテクチャに対応した OpenVMS Integrity は、価格とパフォーマンス、信頼性、および OpenVMS のスケーラビリティの利点を、業界標準の HP Integrity サーバ・システム上で提供します。

- OpenVMS Alpha オペレーティング・システム

64 ビット RISC (縮小命令セット・コンピューティング) アーキテクチャの OpenVMS Alpha は、柔軟性と拡張性を備えた標準 I/O サブシステムにより、価格およびパフォーマンス面で業界をリードしています。

3.2 システムのタイプ

HP Integrity システムは、以下のようなシステム・レンジでコンピューティング環境に対応します。

- エントリ・サーバ
- スタンドアロン・サーバ
- スケーラブル・ブレード
- Integrity VM ゲストシステム

3.3 システムの選択

選択するシステムは、ビジネスやアプリケーションのニーズと予算によって異なります。適切な選択のためには、システムとその特性をよく理解する必要があります。OpenVMS がサポートする Integrity サーバ・システムについては、ソフトウェア仕様書あるいは下記の URL にあるハードウェア・サポート・リストを参照してください。

<http://www.hp.com/go/openvms/>

3.4 可用性について

OpenVMS Cluster システムは、複数のシステムが共通のリソースにアクセスする高度に統合された環境です。このリソース共用により、サービスとデータの可用性が高められます。OpenVMS Cluster システムには透過的かつ自動的なフェールオーバー機能があり、システム管理者やユーザの手をほとんど煩わしません。

関連項目：フェールオーバー機能と可用性の詳細については、第 8 章を参照してください。

3.5 システム仕様

サーバの注文情報および構成情報については、下記の URL の HP の Web サイトを参照してください。ストレージ・デバイス、プリンタ、ネットワーク・アプリケーションのサポートについての詳細な情報も Web サイトを参照してください。

<http://www.hp.com/>

OpenVMS Cluster インターコネクットの選択

インターコネクットは、コンピュータ同士、およびコンピュータとストレージ・サブシステムを接続する物理パスです。OpenVMS Cluster システムは、最も適切で効率的な方法を使用して、メンバが相互におよびストレージと通信できるように、多様なインターコネクット (バスとも呼ぶ) をサポートしています。

OpenVMS Cluster システムでインターコネクット経由の通信を行うためのソフトウェアを SCS (System Communications Services) と呼びます。ノード間 SCS 通信をサポートするインターコネクットをクラスタ・インターコネクットと呼びます。1つのクラスタ内でのノードとストレージ間の接続を可能にするインターコネクットを共用ストレージ・インターコネクットと呼びます。

OpenVMS は、以下のタイプのインターコネクットをサポートします。

- クラスタ・インターコネクット (ノード間専用)
 - Ethernet
 - Fast Ethernet
 - Gigabit Ethernet
 - 10 Gigabit Ethernet (Integrity のみ)
- 共用ストレージ・インターコネクット (ノード・ストレージ間専用)
 - Fibre Channel (Alpha のみ)
 - SCSI (Small Computer Systems Interface) (Integrity あるいは Alpha。Integrity は特定の構成に限定)
 - SAS (Serial Attached SCSI) (Integrity のみ)
- ノード間インターコネクットおよびノード・ストレージ間インターコネクット
 - Ethernet, Fast Ethernet, Gigabit Ethernet
 - 10 Gigabit Ethernet (Integrity のみ)

注意

Cluster over IP は Ethernet , Fast Ethernet , Gigabit Ethernet および 10 Gigabit Ethernet でサポートされます。

注意

CI, DSSI および FDDI インターコネクットは Alpha および VAX システムでサポートされます。Memory Channel および ATM インターコネクットは Alpha システムでのみサポートされます。これらのインターコネクットについての説明は、本書の以前のバージョンを参照してください。

4.1 特徴

この章で説明するインターコネクットにはいくつか共通の特徴があります。表 4-1 に、これらの特徴をまとめました。

表 4-1 インターコネクットの特徴

特徴	説明
スループット	インターコネクットを経由して転送されるデータ量。 インターコネクットによっては、より多くのプロセッサ・オーバーヘッドが必要になります。たとえば、Ethernet インターコネクットと FDDI インターコネクットには、CI や DSSI よりも多くのプロセッサ・オーバーヘッドが必要です。 パケット・サイズが大きければ、小さい場合よりも高いデータ転送速度 (スループット) が可能になります。
ケーブル範囲	3 m から 40 km の範囲で相互接続します。
最大ノード数	1 つのインターコネクットに接続できるノード数は、インターコネクットのタイプによって異なります。OpenVMS Cluster システムの設定時には、このことを忘れないでください。
サポートされるシステムとストレージ	各 OpenVMS Cluster ノードとストレージのサブシステムごとに、内部システム・バスとインターコネクットを結ぶアダプタが必要です。インターコネクット・タイプの選択時には、まずストレージとプロセッサの I/O パフォーマンスを考え、次にアダプタ・パフォーマンスを考えてください。

4.2 インターコネクット・タイプの比較

表 4-2 は、各種インターコネクットの主な仕様です。

表 4-2 インターコネクト・タイプの比較

インターコネクト	最大スループット (Mb/s)	ハードウェア支援データ・リンク ¹	ストレージへの接続	トポロジ	クラスタあたりの最大ノード数	最大長
汎用						
Ethernet Fast Gigabit 10 Gigabit	10/100/1000	No	MSCP 対応	ハブやスイッチに対しリニアまたはラジアル	96 ²	100 m ⁴ / 100 m ⁴ / 550 m ³
共用ストレージ専用						
Fibre Channel	1000	No	直結 ⁵	スイッチに対しラジアル	96 ²	10 km ⁶ /100 km ⁷
SCSI	160	No	直結 ⁵	バスまたはハブに対しラジアル	8-12 ⁸	25 m
SAS	6000	No	直結	ポイント・ツー・ポイント、スイッチに対してラジアル	96 ²	6 m

¹ハードウェア支援データ・リンクを利用すればプロセッサ・オーバーヘッドを節約できます。

²OpenVMS Cluster コンピュータ。

³マルチモード・ファイバ (MMF) を基準とします。ATM, DS3 などの共通キャリアを使用して、このインターコネクトと WAN インタースイッチ・リンク (ISL) 間をブリッジすることによって、より長距離に対応できます。

⁴シールドなしツイスト・ペア・ケーブル (UTP) を基準とします。ATM, DS3 などの共通キャリアを使用して、このインターコネクトと WAN インタースイッチ・リンク (ISL) 間をブリッジすることによって、より長距離に対応できます。

⁵直接接続されている SCSI および Fibre Channel ストレージは、汎用のクラスタ・インターコネクトのいずれかを介して MSCP に対応させることができます。

⁶シングル・モード・ファイバ、ポイント・ツー・ポイント・リンクを基準とします。

⁷シングル・モード・ファイバを使用した、インタースイッチ・リンク (ISL) を基準とする長距離 (最大 100 km) のサポートです。さらに、DRM 構成では、Open Systems Gateway および Wave Division Multiplexor を使用した長距離 ISL を提供します。

⁸最大 3 台の OpenVMS Cluster コンピュータ。DWZZH-05 および適切なアービトレーションにより最大 4 台、最大 15 台のストレージ・デバイス。

4.3 複数のインターコネクト

複数のインターコネクトを使用すると、以下の利点があります。

- フェールオーバ

インターコネクトやアダプタの 1 つにでも障害が発生すると、別のインターコネクトにノード通信が自動的に移行します。

- MSCP サーバ負荷分散

複数の MSCP サーバ構成では、OpenVMS Cluster は負荷分散により最適なパスを自動的に選択します。そのため、1 か所のアダプタで I/O ボトルネックが発生することが少なくなります。構成によっては、1 つのノード間に複数のパスがあると、パスが 1 つだけの場合よりも多くの情報を転送できます。

関連項目：動的 MSCP 負荷分散の例については、第 9.4.3 項を参照してください。

4.4 混成インターコネクット

1 つの OpenVMS Cluster システムで、2 種類以上のタイプのインターコネクットを使用することができます。異なるタイプのインターコネクットでは、それぞれのタイプの利点を組み合わせて OpenVMS Cluster システムを拡張します。

注意

クラスタ内のいずれかのノードでクラスタ通信に IP を必要とする場合は、すべてのメンバ・システムで IP クラスタ通信を有効にする必要があります。

4.5 サポートするインターコネクット

サポートするインターコネクットの最新情報は、OpenVMS Cluster の最新のソフトウェア仕様書を参照してください。

関連項目:各 Integrity サーバおよび AlphaServer システムでサポートするインターコネクットとアダプタについての詳細は、下記の OpenVMS の Web サイトを参照してください。

<http://www.hp.com/go/openvms>

左ナビゲーション・パネルの関連リンクから「HP Integrity servers」を選択してください。その後、確認したい Integrity システムを選択し QuickSpecs を選択します。QuickSpecs には、そのシステムでサポートされるアダプタなど、すべてのオプションが簡単に紹介されています。

Alpha の場合は、左ナビゲーション・パネルから「AlphaSystems」を選択してください。その後、知りたい AlphaServer システムを選択し QuickSpecs を選択します。各 AlphaServer の QuickSpecs には、そのシステムでサポートされるアダプタなど、すべてのオプションが簡単に紹介されています。

4.6 Fibre Channel インターコネクット

Fibre Channel は、PCI ベースの Alpha システム用の高性能 ANSI 標準ネットワーク・ストレージ・インターコネクットです。全二重シリアル・インターコネクットとして、毎秒 100 MB で送受信できます。Fibre Channel は、Fibre Channel スイッチへの複数ノード接続により SCSI ストレージの同時アクセスをサポートします。ノード間通信には、2 番目のタイプのインターコネクットが必要です。

4.6.1 利点

Fibre Channel インターコネクットには、以下の利点があります。

- 2 Gb/s , 4 Gb/s , 8 Gb/s の高速伝送速度 (速度は使用するアダプタに依存)。
- 部門構成から企業構成までサポートするスケーラブルな構成。
- 長距離インターコネクット

Fibre Channel では、1 リンク当たり 500 m のマルチモード・ファイバをサポートしています。Fibre Channel では、長距離インタースイッチ・リンク (ISL) をサポートしています (シングル・モード・ファイバを使用した場合は 1 リンク当たり最大 100 km , FC/ATM リンクの場合は 1 リンク当たり最大 600 km)。

さらに、SANworks の Data Replication Manager (DRM) 構成では、Open Systems Gateway および Wave Division Multiplexor を使用した長距離 ISL を提供します。

- 高可用性
マルチパス・サポートにより、単一点障害 (single point of failure) の発生しない構成が可能です。

4.6.2 スループット

Fibre Channel インターコネクットは、最大 2 Gb/s , 4 Gb/s , あるいは 8 Gb/s で伝送できます (速度は使用するアダプタに依存)。100 MB/s で同時送受信が可能な全二重シリアル・インターコネクットです。

4.7 MEMORY CHANNEL インターコネクット (Alpha のみ)

MEMORY CHANNEL は、PCI ベースの Alpha システム用の高性能クラスター・インターコネクット・テクノロジーです。少ない待ち時間、高い帯域幅、直接メモリ・アクセスにより、MEMORY CHANNEL は OpenVMS Cluster 固有の能力を 1 つの仮想システムへと拡張します。

MEMORY CHANNEL 接続のサポートには、3 つのハードウェア・コンポーネントが必要です。

- PCI MEMORY CHANNEL 間アダプタ
- リンク・ケーブル (延長 3 m または 10 フィート)
- MEMORY CHANNEL ハブのポート (ケーブルで 2 つの PCI アダプタを接続する 2 ノード構成を除く)

MEMORY CHANNEL ハブは、システム間の接続を提供する PC サイズのユニットです。MEMORY CHANNEL ハブは最大 8 つの Alpha ノードをサポートできます。システムは、2 つの MEMORY CHANNEL アダプタで構成してアダプタ障害時のフェールオーバーにすることができます。各アダプタは別々のハブに接続する必要があります。

ノードが 2 つだけのクラスタに MEMORY CHANNEL ハブは必要ありません。2 ノード構成では、モジュール・ジャンパを使用して 1 つの PCI アダプタを仮想ハブに構成します。

4.7.1 利点

MEMORY CHANNEL テクノロジには以下の特徴があります。

- 価格/パフォーマンス比がすぐれている。

MEMORY CHANNEL は、CI 帯域幅の数倍である、最小限の待ち時間による 100 MB/s インターコネクットを提供します。MEMORY CHANNEL アーキテクチャは、業界標準 PCI バスとして設計されています。

- 既存のアプリケーションを変更する必要がない。

MEMORY CHANNEL は既存のクラスタ・ソフトウェアとシームレスに運用でき、既存のアプリケーションを変更する必要がありません。新しい MEMORY CHANNEL ドライバ、PMDRIVER、MCDRIVER は、既存のポート・ドライバと同様に OpenVMS Cluster のレイヤと統合します。クラスタ・ソフトウェアの上位レイヤには影響がありません。

- SCSI クラスタの CI、DSSI、LAN の負荷を解消。

MEMORY CHANNEL にストレージを直接接続することはできませんが、各インターコネクットのパワーを最大限に活用することができます。

MEMORY CHANNEL は、CI や DSSI に代わるものではありませんが、これらのインターコネクットと併用することで、ノード間のトラフィックの負荷が解消されます。すなわち MEMORY CHANNEL をストレージ・トラフィック専用としてクラスタ全体の通信を最適化します。

SCSI インターコネクットと LAN インターコネクットと併用すれば、MEMORY CHANNEL によって LAN からノード間のトラフィックの負荷が取り除かれ、さらに多くの TCP/IP トラフィックや DECnet トラフィックの処理ができるようになります。

- フェール隔離処理を提供する。

システム障害が発生すると、MEMORY CHANNEL ノードは OpenVMS Cluster のノードと同様に対応します。すなわち、障害が発生したノードがクラスタに復帰するまで、障害が発生しなかった他のノードが引き続き処理を続行するわけです。

4.7.2 スループット

MEMORY CHANNEL インターコネクットの最大スループットは 100 MB/s という高い値を備えています。1 つの MEMORY CHANNEL で不十分な場合は、最大 2 つのインターコネクット (および 2 つの MEMORY CHANNEL ハブ) にスループットを振り分けることができます。

4.7.3 サポートされているアダプタ

MEMORY CHANNEL アダプタは PCI バスに接続します。MEMORY CHANNEL アダプタの CCMAA-BA は、以前のアダプタに比べ、パフォーマンスが改善されました。

関連項目：各 AlphaServer システムでサポートされる CCMAA-BA アダプタの詳細については、次の OpenVMS の Web サイトにアクセスしてください。

<http://www.hp.com/go/openvms>

左ナビゲーション・パネルから「AlphaSystems」を選択してください。その後、知りたい AlphaServer システムを選択し QuickSpecs を選択します。各 AlphaServer の QuickSpecs には、そのシステムでサポートされるアダプタなど、すべてのオプションが簡単に紹介されています。

4.8 SCSI インターコネクット

SCSI インターコネクットは、業界標準のインターコネクットであり、1 つ以上のコンピュータ、周辺機器、インターコネクット・コンポーネントをサポートしています。SCSI は、シングル・パス、デジター・チェーンのマルチドロップ・バスです。また、SCSI は、シングル 8 ビットまたは 16 ビットのデータ・パスであり、エラー検出用のバイト・パリティを備えています。長距離通信用には、安価なシングル・エンド・シグナル通知と作動シグナル通知の両方を利用できます。

OpenVMS Cluster では、1 つの SCSI インターコネクット上の複数のコンピュータが同時に SCSI ディスクにアクセスできます。この種の構成をマルチホスト SCSI 接続あるいは共有 SCSI ストレージと呼び、特定のアダプタと限られた構成に制限されます。ノード間通信には、別のタイプのインターコネクットが必要です。

OpenVMS Cluster システムで共有 SCSI ストレージを使用すると、単一の SCSI バスに接続された複数のコンピュータが、SCSI ストレージ・デバイスへのアクセスを共用することが可能になります。この機能により SCSI ストレージへの共有アクセスを使用して、より可用性の高いサーバ環境を構築することが可能になります。

4.8.1 OpenVMS Alpha の構成

SCSI ストレージに対するマルチホスト・アクセスでは、以下のコンポーネントが必要です。

- マルチホスト構成でサポートされている SCSI ホスト・アダプタ (表 4-5 を参照してください。)
- SCSI インターコネクット
- SCSI インターコネクットの端点ごとに 1 つのターミネータ
- マルチホスト構成でサポートされているストレージ・デバイス (RZnn; OpenVMS Cluster SPD [29.78.nn]を参照してください。)

さらに大きな構成規模では、以下のコンポーネントを利用できます。

- ストレージ・コントローラ (HSZnn)
- シングル・エンドをディファレンシャル・シグナル通知に変換し SCSI インターコネクット長を倍増するためのバス・アイソレータ (DWZZA, DWZZB, または DWZZC)。

注意

このサポートは一部のアダプタに限定されます。OpenVMS では、Ultra SCSI アダプタ KZPEA, KZPDC, A6828A, A6829A, および A7173A などの最新の SCSI アダプタに対してはサポートしません。

関連項目： SCSI 構成の接続方法の詳細については、付録 A を参照してください。

4.8.2 2 ノード OpenVMS Integrity サーバの共有 SCSI 構成

2 ノードの OpenVMS Integrity クラスタ・システムでの共有 SCSI ストレージのサポートは OpenVMS Version 8.2-1 で導入されています。これより前のバージョンでは、共有 SCSI ストレージは以前の SCSI ホストベース・アダプタ (HBA) を使用して OpenVMS Alpha システムでのみサポートされていました。

OpenVMS Integrity クラスタ・システムでの共有 SCSI ストレージには以下の制限事項が適用されます。

- 1 つの SCSI バスに接続できる OpenVMS Integrity サーバ・システムは最大 2 つです。
- 各システムに接続できる共有 SCSI バスは最大 4 つです。
- サポートされるシステムは、rx1600 シリーズ, rx2600 シリーズ, および rx4640 システムです。
- サポートされる HBA は A7173A HBA のみです。

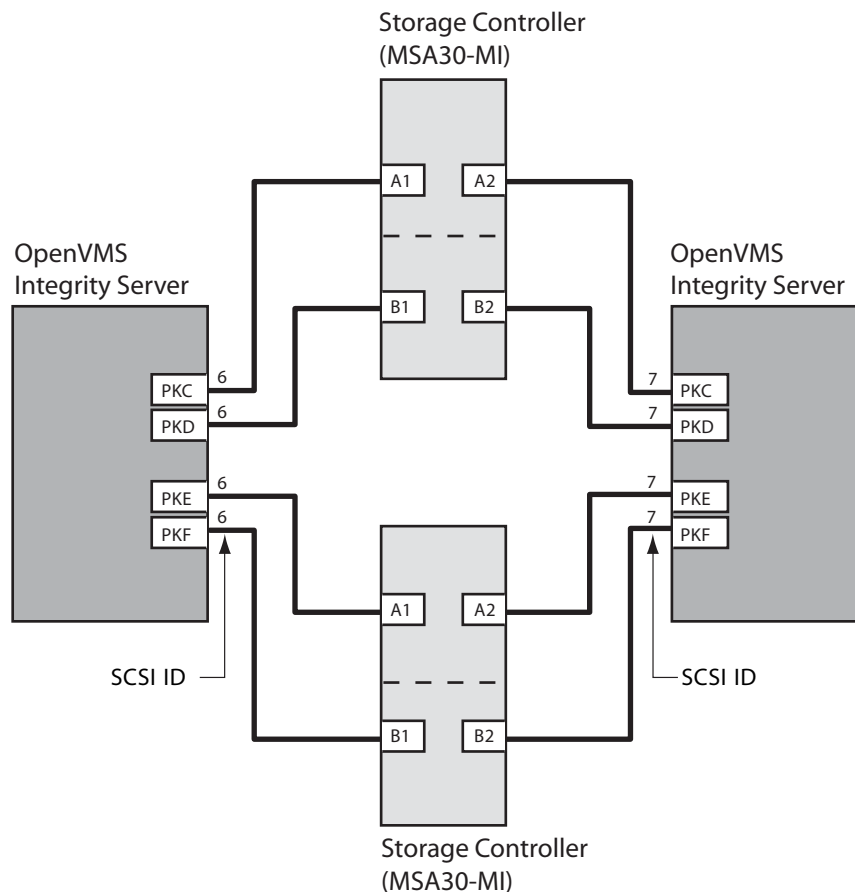
- サポートされる SCSI ストレージは MSA30-MI ストレージ・エンクロージャのみです。
- サポートされる ディスク・タイプは Ultra320 SCSI ディスクのみです。

図 4-1 に 2 ノードの共有 SCSI 構成を图示します。ホスト間の OpenVMS クラスタ通信には 2 つめのインターコネク트의として LAN が必要になります (OpenVMS クラスタ通信は System Communications Architecture (SCA) 通信と呼ばれます)。

この構成には SCSI ID 6 および 7 が必要になります。どちらか一方のシステムでは、出荷時設定のデフォルト ID 7 ではなく、SCSI ID 6 で A7173A アダプタ・ポートに接続されていなければなりません。SCSI ID の変更には、IPF Offline Diagnostics and Utilities CD に含まれている U320 SCSI pscsi.efi ユーティリティを使用します。変更の手順については、次の URL にある『HP A7173A PCI-X Dual Channel Ultra320 SCSI Host Bus Adapter Installation Guide』を参照してください。

<http://docs.hp.com/en/netcom.html>

図 4-1 2 ノードの OpenVMS Integrity クラスタ・システム



VM-1189A-AI

4.8.3 利点

SCSI インターコネクトには、以下の利点があります。

- 最小限のコストによるストレージへの共用直接アクセス
SCSI は、業界標準であり、業界内では幅広く使用されているため、リーズナブルな価格で各社の製品を入手できる。
- スケーラブルな構成により適切な価格で高いパフォーマンスを得ることができる。

以下の選択肢があります。

- SCSI インターコネクトの幅
Narrow (8 ビット) または Wide (16 ビット)。
- 通信モード

最も一般的で、安価なシングル・エンド・シグナル通知か、高い信号の一貫性と長い SCSI インターコネクットに対応できるディファレンシャル・シグナル通知

- 信号速度 (Standard, Fast, Ultra モード)
- SCSI バスを共用するノード (2 ノードまたは 3 ノード)
- 1 ノードに接続できる共用 SCSI バスの数 (最大 6 本)
- ストレージのタイプとサイズ (RZnn または HSZnn)
- コンピュータのタイプとサイズ (AlphaStation または AlphaServer)

4.8.4 スループット

表 4-3 は、SCSI インターコネクットのスループットをまとめたものです。

表 4-3 最大転送速度，単位 MB/秒

モード	Narrow (8-Bit)	Wide (16-Bit)
Standard	5	10
Fast	10	20
Ultra	20	40

4.8.5 SCSI インターコネクット距離

SCSI インターコネクットの最大長は、構成で使用するシグナル通知方法によって決まり、シングル・エンド・シグナル通知の場合はデータ転送速度で決まります。

SCSI インターコネクット用の電気シグナル通知は、シングル・エンドとディファレンシャルの 2 種類があります。どちらの種類も Standard モード、Fast モード、または Ultra モードで動作します。ディファレンシャル・シグナル通知の場合、最大 SCSI ケーブル長は、Standard モードと Fast モードで同じになります。

表 4-4 は、シグナル通知方法のタイプによる SCSI インターコネクット距離の違いをまとめたものです。

表 4-4 最大 SCSI インターコネクット距離

シグナル通知方式	データ転送速度	最大ケーブル長
シングル・エンド	Standard	6 m ¹

¹SCSI 標準では、このインターコネクットに最大長 6 m を指定しています。ただし、データの一貫性を高いレベルで維持するためには、最大長はできれば 4 m までにしてください。

(次ページに続く)

OpenVMS Cluster インターコネクットの選択

4.8 SCSI インターコネクット

表 4-4 (続き) 最大 SCSI インターコネクット距離

シグナル通知方式	データ転送速度	最大ケーブル長
シングル・エンド	Fast	3 m
シングル・エンド	Ultra	20.5 m ²
ディファレンシャル	Standard または Fast	25 m
ディファレンシャル	Ultra	25.5 m ³

²この長さは、デバイスが両端に接続される場合にだけ適用します。デバイスをインターコネクット上に分布する場合は、最低 1 m ずつ離してください。またインターコネクットが 4 m を超えないようにします。

³デバイスは 3 つ以上接続できます。

4.8.6 サポートされているアダプタ、バス・タイプ、コンピュータ

表 4-5 は、SCSI アダプタと、それがサポートする内部バスとコンピュータをまとめたものです。

表 4-5 SCSI アダプタ

アダプタ	内部バス	サポート コンピュータ
埋め込み (NCR-810 based)/KZPAA ¹	PCI	システムのオプション仕様を参照してください。
KZPSA ²	シングル・ホスト構成で PCI	KZPSA をサポートするすべての Alpha コンピュータがサポート ³
KZTSA ²	TURBOchannel	DEC 3000
KZPBA-CB ⁴	PCI	シングル・ホスト構成で KZPBA をサポートするすべての Alpha コンピュータがサポート ³

¹シングル・エンド。

²Fast Wide Differential (FWD)。

³各システム別のハードウェア・マニュアルを参照してください。

⁴Ultra Differential。 Ultra シングル・エンド・アダプタ (KZPBA-CA) は、マルチホスト・システムをサポートしません。

関連項目：各 OpenVMS Integrity あるいは OpenVMS Alpha システムでサポートされる SCSI アダプタについては、次の OpenVMS の Web サイトを参照してください。

<http://www.hp.com/go/openvms>

左ナビゲーション・パネルから「AlphaSystems」あるいは「Integrity Server」を選択してください。その後、システムを選択し QuickSpecs を選択します。QuickSpecs には、そのシステムでサポートされるアダプタなど、すべてのオプションが紹介されています。

4.9 SAS インターコネクット (Integrity のみ)

SAS は、シリアル通信 (一度に 1 ビットずつ) を使用して SCSI ストレージ・デバイスとの間でデータを転送するポイント・ツー・ポイント・アーキテクチャです。SAS は、SAS デバイスとディファレンシャル・シグナリング方式を使用して信頼性の高い高速シリアル通信を実現します。

SAS は、ファイバーチャネルのハイエンド機能 (マルチイニシエータのサポートや全二重通信など) および SATA を利用した物理インタフェース (優れた互換性と投資保護) により、高い性能と信頼性、伝統的な SCSI 技術の使いやすさを兼ね備えています。

SAS デバイス

SAS デバイスには、イニシエータ、ターゲット、エキスパンダの 3 つのタイプがあります。イニシエータ・デバイスは HBA あるいはコントローラです。イニシエータは 1 つ以上のターゲット (SAS ハード・ディスク・ドライブ、SATA ハード・ディスク・ドライブ、あるいは SAS テープ・ドライブ) と接続され SAS ドメインを形成します。エキスパンダは、イニシエータに接続されるターゲットの数を増やし、これにより大規模な SAS ドメインを作成するローコスト、ハイスピードのスイッチです。各 SAS デバイスは、ドメイン内で簡単に識別できるように、製造者によって割り当てられたユニークな名前 (SAS アドレス) を持ちます。

ディファレンシャル・シグナリング

すべての SAS デバイスはポートと呼ばれる接続ポイントを持っています。各 SAS デバイスのポートには、物理層 (PHY) と呼ばれる 1 つあるいは複数のトランシーバ・メカニズムが置かれています。2 つのペア線で構成される物理リンクが、1 つのデバイスのポートにある各 PHY のトランスミッタと別のデバイスのポートにある PHY のレシーバを接続します。SAS インタフェースは、複数の物理リンクを組み合わせるポートごとに 2x、3x、4x、あるいは 8x の接続を作成し、帯域幅を拡張することができます。1 つの PHY を持つポートをナロー (narrow) と呼び、2 から 4 つの PHY を持つポートをワイド (wide) と呼びます。

SAS はディファレンシャル・シグナリングを使用して物理リンク経由でデータを転送するため、パラレル SCSI を高速で使用した場合に見られる静電容量、インダクタンス、ノイズの影響を低減できます。SAS 通信は全二重で、各 PHY は 2 つのペア線を使って情報を同時に送受信できます。

4.9.1 利点

SAS を使用したマルチノード・クラスタはクラスタ化された Fibre Channel ローカルループ・トポロジの代わりとして使用できます。拡張性の高い SAS アーキテクチャにより、単一障害点のない高性能で可用性の高いトポロジが可能です。

OpenVMS Cluster インターコネクットの選択

4.9 SAS インターコネクット (Integrity のみ)

SAS ソリューションは、ローコストの大容量記憶装置 (SATA) あるいはミッション・クリティカル・アプリケーションのための高性能および高信頼性 (SAS) の両方に対応し、構成の柔軟性と簡略性を高めることができます。

4.9.2 スループット

SAS は、パラレル SCSI (すなわち 320 Mb/s) よりも速い速度で動作します。SAS インターコネクットのスループットを表 4-6 に示します。SCSI インターコネクットのスループットは表 4-3 を参照してください。

表 4-6 最大転送速度，単位 MB/秒

モード	MB/秒単位の速度
SAS 1 (3Gig SAS と呼ばれる)	300 MB/s あるいは 3 Gb/s
SAS 2 (6Gig SAS と呼ばれる)	600MB/s あるいは 6Gb/s

4.9.3 サポートされるアダプタ，バス，タイプ，コンピュータ

表 4-7 に、SAS アダプタとそれらがサポートする内部バスおよびコンピュータを示します。

表 4-7 SAS アダプタ

アダプタ	内部バス	サポートするコンピュータ
8p SAS HBA	PCI-X	Integrity サーバ rx3600, rx6600 のコア I/O
HP SC44Ge Host Bus Adapter	PCIe	PCIe バックプレーンの Integrity サーバ (rx2660, rx3600, rx6600)

4.10 LAN インターコネクット

Ethernet (Fast Ethernet, Gigabit Ethernet, および 10 Gigabit Ethernet) は、LAN ベースのインターコネクットです。

これらのインターコネクットには、以下の機能があります。

- OpenVMS Cluster システムとローカル・エリア・ネットワーク (LAN) 内でのシングル・パス接続
- LAN フェールオーバーのサポート
- 複数のアダプタを使用した複数パスのサポート
- 長距離インターコネクット

表 4-2 に示す、各 LAN タイプに固有の最大長に加え、LAN と WAN インタースイッチ・リンク間のブリッジによって長距離を実現することができます。

- ノードの物理的な広範囲にわたる展開
- シングル・インターコネクットでの複数クラスタ (それぞれ最大 96 ノード) のサポート

各 OpenVMS プラットフォーム (Integrity および Alpha) 上で OpenVMS Cluster インターコネクットとしてサポートされている LAN を、表 4-8 に示します。

表 4-8 OpenVMS Cluster 用にサポートされている LAN インターコネクット

LAN のタイプ	プラットフォーム
Ethernet	Integrity, Alpha
Fast Ethernet	Integrity, Alpha
Gigabit Ethernet	Integrity, Alpha
10 Gigabit Ethernet	Integrity のみ

複数の LAN アダプタについて説明した後に、サポートされている各 LAN インターコネクット、Ethernet に固有の情報を示します。

4.10.1 複数の LAN アダプタ

複数の LAN アダプタがサポートされます。LAN アダプタは、さまざまなタイプの LAN を使用する場合でも、同じ LAN タイプに複数の異なるアダプタ・モデルを使用する場合でも対応することができます。

複数の LAN アダプタを使用すると、以下のことが実現できます。

- 複数の LAN パスに負荷を分散してノード間のスループットを上げることができます。
- ノード間の LAN 通信の可用性を改善することができます。

4.10.1.1 複数の LAN パスの負荷分散

複数のノード間の LAN パスが利用できる場合、OpenVMS Cluster ソフトウェアは、以下の基準に基づいて、使用するパスのセットを選択します。この基準は厳密な優先順位で評価されます。

1. パス上でのパケット損失の最新履歴

最近、高いレートでパケットを損失しているパスは、損失過多と呼ばれ、考慮対象から外されます。受け入れ可能な損失履歴を持つチャンネルは、タイトと呼ばれ、使用対象としてさらに検討されます。

2. 優先順位

管理優先順位値は、個別の LAN パスとローカルの LAN デバイスの両方に割り当てることができます。LAN パスの優先順位値は、これらの優先順位の合計値になります。優先順位値が、タイト・パスの最上位優先順位値以下であるタイト LAN パスのみ使用対象としてさらに検討されます。

3. 最大パケット・サイズ

最大パケット・サイズが、すべてのタイト等価優先順位チャンネルの中で最も大きいパケット・サイズの最大サイズと同じタイト等価優先順位チャンネルは、使用対象としてさらに検討されます。

4. 等価待ち時間

前項までの基準を満たした LAN パスは、その待ち時間(コンピューテッド・ネットワーク遅延)が最速のチャンネルの待ち時間と非常に近い場合に使用されます。各 LAN パスの遅延は、そのパスのクラスタ通信トラフィックを使用して測定されます。LAN パスが前項までの基準を満たさないため、クラスタ通信の使用対象から外された場合、そのパスの遅延は数秒間隔で測定され、遅延、またはパケット・ロスのレートが前項までの基準を満たすまで改善されているか確認されます。

パケット伝送は、上記の基準を満たすローカル・アダプタとリモート・アダプタ間のすべての通信パスにラウンド・ロビン方式で分散されます。

4.10.1.2 LAN パスの可用性の向上

LAN は、距離が広い範囲におよぶ場合に適しているため、インターサイト・リンクのスループットを高可用性で強化する場合があります。これを行うためには、それぞれ別のインターサイト LAN リンクに接続されている、複数の LAN アダプタにより重要ノードを構成します。

インターサイト・リンクの故障は、通常、インターサイト・リンクの機械上の破損が原因です。これは、パスの多様性、つまり、複数のインターサイト・リンクのパスを物理的に分離することで回避できます。パスの多様性を利用すると、インターサイト・リンクに影響を与えるような災害によって構成が影響を受ける可能性は確実に少なくなります。

4.10.2 LAN ベースのクラスタの構成上のガイドライン

以下のガイドラインは、LAN ベースのすべての OpenVMS Cluster システムに適用されます。

- OpenVMS Integrity および OpenVMS Alpha システムは、各アーキテクチャ上でサポートされているさまざまな LAN アダプタを組み合わせることで構成できます。サポートされているアダプタについては、最新のソフトウェア仕様書を参照してください。

- OpenVMS Cluster 通信に使用されるすべての LAN バスは、10 Mb/s 以上のスループットと少ない待ち時間で動作する必要があります。あるタイプの LAN のノードを別のタイプの LAN のノードに接続する場合は、ブリッジまたはスイッチにより変換する必要があります。LAN セグメントをブリッジ接続して、拡張 LAN を作成できます。
- 複数の、異なる OpenVMS Cluster システムを単一の拡張 LAN に構成できます。OpenVMS Cluster ソフトウェアは、クラスタ・メンバシップ・チェックを実行して、システムが正しい LAN OpenVMS クラスタに参加することを保証します。

4.10.3 Ethernet (10/100) および Gigabit (1/10) Ethernet の利点

Ethernet (10/100) インターコネク트는、すべての OpenVMS Cluster インターコネク트의の中で最も経済的 です。

Gigabit (1/10) Ethernet インターコネク트는、第 4.10 節の項に示されている利点に加え、以下の利点を提供します。

- 超高速スループット (1 Gb/s, 10 Gb/s)
- クラスタ通信での Jumbo フレーム (フレームあたり 7552 バイト) のサポート

4.10.4 Ethernet (10/100) および Gigabit (1/10) Ethernet のスループット

Ethernet テクノロジによるベースバンドの伝送速度の範囲は以下のとおりです。

- Standard Ethernet で 10 Mb/s
- Fast Ethernet で 100 Mb/s
- Gigabit Ethernet で 1 Gb/s
- 10 Gigabit Ethernet で 10 Gb/s

Ethernet アダプタにはハードウェア・アシストがないので、CI や DSSI の場合よりもプロセッサ・オーバヘッドが高くなります。

多くの Ethernet 接続ノードで OpenVMS Cluster システムを構成するときや、Ethernet が多くの PC やプリンタもサポートする場合、ネットワーク・デザイン全体のキャパシティを考慮してください。Ethernet 上の全体的なネットワーク・トラフィックにより、OpenVMS Cluster 通信に利用できるスループットが低下することがあります。Fast Ethernet と Gigabit Ethernet では、スループットが大幅に改善することがあります。Ethernet アダプタを複数使用すると、全般的なネットワーク・トラフィックの負荷が軽減されてクラスタ・パフォーマンスが改善されることがあります。

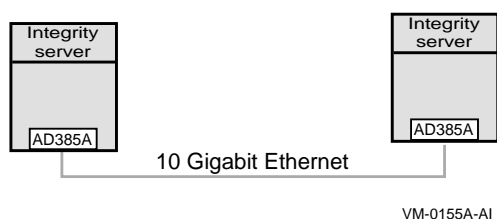
関連項目：LAN 構成におけるガイドラインについては、第 4.10.2 項を参照してください。

4.10.5 10 Gigabit Ethernet クラスタの構成上のガイドライン

10 Gigabit Ethernet クラスタでシステムを構成する場合は、以下のガイドラインに従ってください。

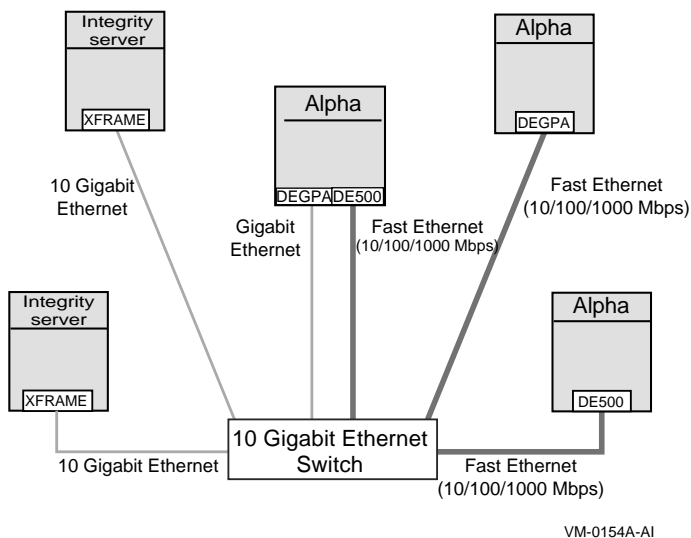
- 2 ノード Gigabit Ethernet クラスタにスイッチは不要です。図 4-2 に示すように、ポイント間で接続できるからです。

図 4-2 ポイント・ツー・ポイント 10 Gigabit Ethernet OpenVMS Cluster



- ほとんどの 10 Gigabit Ethernet スイッチは、Gigabit Ethernet または Gigabit Ethernet と Fast Ethernet を組み合わせて構成できます (100 Mb/s)。
- 各ノードでは、スイッチまでをシングル接続とすることもできますが、図 4-3 に示すように、複数のバスで構成して可用性を高めることもできます。

図 4-3 スイッチ型 10 Gigabit Ethernet OpenVMS Cluster



- OpenVMS バージョン 7.3 から Jumbo フレーム (各 7552 バイト) のサポートが利用できるようになりました。(Jumbo フレームのサポートが導入される前は、クラスタ通信でサポートされるフレーム・サイズは、最大 1518 バイトの標準 Ethernet フレーム・サイズだけでした。)

- DEGPA はブート・デバイスに使用できませんが、サテライトは Gigabit スイッチ上に構成した 10/100 Ethernet ネットワーク・アダプタを経由してブートすることができます。DEGXA Gigabit Ethernet アダプタは、Broadcom 製の BCM5703 チップ (TIGON3) ネットワーク・インタフェース・カード (NIC) です。DEGXA Gigabit Ethernet アダプタの採用により、LAN デバイスとクラスター・インターコネクット・デバイスの両方として既存の Gigabit Ethernet のサポートが継続されます。DEGXA はブート・デバイスとして使用できます。

4.11 Cluster over IP

OpenVMS Cluster ではクラスタ通信にインターネット・プロトコルを使用することもできます。クラスタの基本的なルールでは、クラスタ内のすべてのノードは他のすべてのノードと直接通信できなければなりません。クラスタ内のノードはデータセンタの同じ LAN 内に配置することも、地理的に離れた場所に分散させることもできます。

ノードが同じ LAN 内に存在する場合、クラスタ通信には LAN を使用するのが好まれます。ノードが複数のサイトあるいは複数の LAN に分散している場合、クラスタ通信には IP が好まれます。

レイヤ 2 サービスが利用できない場合あるいはそのコストが高い場合、2 つのサイト間のクラスタ通信に Cluster over IP が使用できます。

注意

2 つのサイト間に拡張 LAN あるいは VLAN を構築して、2 つの異なるサイトのノード間のクラスタ通信に LAN を使用することも可能です。

LAN 経由のクラスタ・プロトコル (SCS, 別名 SCA: System Communication Architecture) はポート・エミュレータ・ドライバ (PEDRIVER) によって提供され、PEDRIVER は図 1-1 に示すようにクラスタ通信に LAN に加えて TCP/IP を使用して SCS 通信を提供します。PEDRIVER は UDP を使用して 2 つのノード間でパケットを伝送します。

Cluster over IP は次の機能を提供します。

- IP マルチキャストと IP ユニキャストにより、IP のみのネットワークで LAN セグメントを越えてノードを検出することが可能です。マルチキャストおよびユニキャスト・ノード検出メカニズムは、クラスタ内の新しいノードを追跡するのに使用されます。ユニキャスト・ノード検出に基いたファイルは、構成ファイルの代わりとして使用できます。

SYS\$SYSTEM:PE\$IP_CONFIG.DAT には、オプションでクラスタのノードの IP マルチキャスト・アドレスおよび IP ユニキャスト・アドレスが含まれます。同じ IP マルチキャスト・ドメイン内のノードの検出には IP マルチキャスト・メッセージが使用され、異なる IP マルチキャスト・ドメインにあるリモート・ノードは、IP ユニキャスト・メッセージング技術を使用してクラスタに参加することができます。

SYS\$SYSTEM:TCPIPCLUSTER.DAT には、クラスタ通信が有効になっている IP インタフェース・アドレス名と IP アドレスが含まれます。また、TCP/IP ルート情報も含まれます。

- クラスタの動作を妨げずに動的にノードの追加および削除が可能です。これにより、クラスタをリブートせずに問題のない透過的な移行が可能になります。また、クラスタをリブートせずに新しいバージョンへのローリング・アップグレードを実行することができます。
- 最小限の待ち時間でクラスタ状態の遷移とフェールオーバのサポートが可能です。PEDRIVER が提供する遅延調査機能と遅延計測機能により、最小の待ち時間でパスを選択することによって IP ネットワークにおける待ち時間を低減することができます。
- 以前のバージョンの LAN ベースの OpenVMS Cluster が稼動するサーバとの相互運用性を提供します。
- 利用可能で正常に動作しているすべてのインタフェース間での動的な負荷バランスを支援します。
- システム管理者は、上位層に対して透過的な正常なインタフェース・セットから障害のあるインタフェースを検出し取り外すことができます。IP 経由のクラスタ通信の管理には Availability Manager が使用できます。
- 会社の方針で非 IP プロトコルの使用が制限されていたり、いずれかのスイッチが通信会社回線で LAN ブリッジが利用できないような環境でも使用できます。

4.11.1 構成上のガイドライン

IP ベースの OpenVMS Cluster には以下のガイドラインが適用されます。

- Cluster over IP を構成する前に、TCP/IP ソフトウェアが構成されていることを確認してください。ネットワークおよび TCP/IP が正しく構成されているかどうか確認するには、PING ユーティリティを使用して、サブネットの外からノードに対して接続試験用のパケットを送信してください。
- クラスタ通信に IP を使用して接続できるノードの最大数は 96 で、これは混成アーキテクチャ構成でもサポートされます。混成バージョンのクラスタ環境であっても、OpenVMS のサポート・マトリックスに違いはありません。
- OpenVMS Version 8.4 以降 Cluster over IP がサポートされ、クラスタ通信には HP OpenVMS TCP/IP Version 5.7 が必要となります。

- Alpha サテライト・ノードおよびそのディスク・サーバは同一 LAN 上になければなりません。
- クラスタ通信に IP を使用する OpenVMS Cluster では、少なくとも 100Mbps のスループットが必要になります。また、100Mbps/1Gbps/10Gbps Ethernet をサポートします。OpenVMS は、最大 500 マイルの距離でクラスタをサポートします (能力的には 60,000 マイルまでサポート可能)。

4.11.2 IP の可用性

ロジカル LAN フェールオーバを使用してクラスタ通信のための IP アドレスを設定することができます。ロジカル LAN フェールオーバ機能は、どのようなリンク障害にも対応する高い可用性を提供します。ロジカル LAN フェールオーバについては、第 8.7 節を参照してください。

4.11.3 IP の利点

Cluster over IP には以下のような利点があります。

- Cluster over IP は、いずれかのスイッチが通信会社回線でレイヤ 2 のサイト間サービスが利用できないようなサイトでも配備できます。
- 会社の方針でいくつかの非 IP プロトコルが制限されることはあっても、IP プロトコルの使用は通常制限されません。
- 専用のハードウェアや人材のためのコストを必要とせず、LAN ブリッジによるマルチサイトの DT (Disaster Tolerant) クラスタを構築できます。
- ローコストで高速な IP サービス

4.11.4 IP の性能

クラスタ通信に IP を使用しても同程度の性能レベルを維持することは重要な課題です。長距離クラスタでは、クラスタ・メンバ・ホスト間での IP スタックの通信に関連する遅延に比べて、サイト間が地理的に離れているために生じる光速の電気通信の遅れが、遅延の主な原因になります。フェールオーバまでの待ち時間とシステムの安定性との間にはトレードオフがあります。サイト間の距離がサポートされている上限に近づくとつれて、通常クラスタ通信 (フェールオーバ無し) を局所化することが、システム性能の最適化に重要です。

OpenVMS Cluster ストレージ・サブシステムの選択

この章では、ストレージ・サブシステムの設計方法について説明します。設計は次の手順で行います。

1. ストレージ製品の選択肢
2. 記憶容量の必要量の見積り
3. ディスク・パフォーマンス・オブティマイザの決定
4. ディスクの可用性の必要要件の決定
5. 項目別の利点とトレードオフの確認:
 - SAS 方式のストレージ
 - SCSI 方式のストレージ
 - Fibre Channel 方式のストレージ
 - ホスト方式のストレージ
 - LAN InfoServer

この章では、これらの手順を詳細に説明します。

5.1 ストレージ製品の選択肢

OpenVMS Cluster のストレージは、SCSI-2 (Small Computer Systems Interface) 標準方式のモジュラ拡張システムである StorageWorks 製品ファミリから選択します。StorageWorks では、次のモジュラ要素から選択して複雑なストレージ・サブシステムを構成できます。

- ディスク、テープ、CD-ROM、半導体ディスク (SSD) などのストレージ・デバイス
- アレイ・コントローラ
- 電源
- パッケージング
- インターコネクト
- ソフトウェア

5.1.1 デバイスの選択基準

ストレージ・デバイスは次の基準で選択してください。

- サポートされているインターコネク
- 容量
- I/O 速度
- 設置面積
- 購入価格，サービス費用，保守費用

5.1.2 インターコネクとストレージ・デバイスの選択

OpenVMS Cluster システムには，メンバ・システムからストレージ・デバイスへのアクセスのために，ストレージ・デバイスを OpenVMS Cluster インターコネクに直接接続できるという利点があります。

OpenVMS Cluster システムでは，次に示すストレージ・デバイスとアダプタを OpenVMS Cluster インターコネクに接続できます。

- LSI 1068 および LSI Logic 1068e (SAS 上)
- HSZ シリーズと RZ シリーズ (SCSI 上)
- HSG コントローラと HSV コントローラ (Fibre Channel 上)
- ローカル・システム・アダプタ

表 5-1 は，特定のインターコネクに接続できるストレージ・デバイスの種類をまとめたものです。

表 5-1 インターコネク，および接続するストレージ・デバイス

ストレージ・インターコネク	ストレージ・デバイス
SCSI	HSZ コントローラ， SCSI ストレージ・デバイス
Fibre Channel	HSG コントローラ， HSV コントローラ， SCSI ストレージ・デバイス
SAS	LSI 1068 および LSI 1068e コントローラ， SCSI ストレージ・デバイス

5.1.3 設置面積とストレージ・デバイスの選択

床面積当たりのコストが高く、ストレージ・デバイスで大きな床面積を占有したくない場合、次の選択肢があります。

- 設置面積が小さくても容量が大きいディスク・ストレージ・アレイを選択する。天井が高い実験室向けには、積み重ね方式のキャビネットが用意されているストレージ・デバイスがあります。
- ディスクのコストよりも容量を優先する。
- 新しいストレージ・アレイやディスクに定期的にアップグレードする習慣をつける。ストレージ・デバイスのテクノロジーの進歩とともに、パフォーマンスと容量が増える一方で物理的なサイズが小さくなっています。
- 電源と空調機用に適切な床面積を検討する。

5.2 記憶容量の必要量の決定

記憶容量は、システム・ファイル、アプリケーション・ファイル、ユーザ・ファイルを保存するために必要なストレージ・デバイスの容量です。記憶容量を算出すれば、OpenVMS Cluster 構成に必要な記憶容量が決まります。

5.2.1 記憶容量の必要量の見積り

オンライン記憶容量の必要量を見積もるには、表 5-2 の説明に従って、OpenVMS Cluster システムのソフトウェアごとに必要なストレージの容量を合算します。

表 5-2 ディスク容量の見積り

ソフトウェア・コンポーネント	説明
OpenVMS オペレーティング・システム	OpenVMS オペレーティング・システムに必要なブロック ¹ 数を見積ります。 関連項目： OpenVMS インストール・マニュアルと SPD (Software Product Description [ソフトウェア仕様書]) を参照してください。
ページ・ファイル、スワップ・ファイル、ダンプ・ファイル	ページ・ファイル、スワップ・ファイル、ダンプ・ファイルに必要なディスク容量を AUTOGEN で決定します。 関連項目： 以上のファイル・サイズの計算や変更については、『OpenVMS システム管理者マニュアル』を参照してください。
サイト固有のユーティリティとデータ	サイト固有のユーティリティ、コマンド・プロシージャ、オンライン・ドキュメント、関連ファイルごとのディスク記憶容量の必要量を見積ります。
アプリケーション・プログラム	アプリケーション製造元のデータを基に、OpenVMS Cluster システムにインストールするアプリケーションごとの容量を見積もります。 関連項目： SPD (Software Product Description) を参考にして、使用するレイヤ製品の通常の操作に必要な容量を見積もります。

¹記憶容量の単位はブロックです。1 ブロックは 512 バイトです。

(次ページに続く)

表 5-2 (続き) ディスク容量の見積り

ソフトウェア・コンポーネント	説明
ユーザ作成プログラム	ユーザ作成プログラムと付属データベースに必要な容量を見積もります。
データベース	各データベースのサイズを見積もります。この情報は、アプリケーション固有のデータベースに関連するドキュメンテーションに必要です。
ユーザ・データ	以下のガイドラインに基づいてユーザ・ディスク容量の必要量を見積もります。 <ul style="list-style-type: none">• 臨時ユーザには、10,000 ブロックから 100,000 ブロックを割り当てます。 臨時ユーザは、電子メールの読み取り、書き込み、削除をします。臨時ユーザがプログラムを持つことはほとんどなく、したがって臨時ユーザがファイルを長期間保存することは滅多にありません。• 標準ユーザには 250,000 ブロックから 1,000,000 ブロックを割り当てます。 標準ユーザは、電子通信、オンラインの情報保存にシステムを幅広く使用し、個人使用目的のプログラムも少数使用します。• 上級ユーザには、1,000,000 ブロックから 3,000,000 ブロックを割り当てます。 上級ユーザには、電子メール用の標準システム用以外に、プログラム開発用、データ・ファイル保存用に大量のストレージ領域が必要になります。上級ユーザには、開発や保守管理の対象になるプロジェクトやプログラムの数により、数十万ブロックのストレージ領域が必要となることもあります。
合計必要量	以上の見積り容量の合計が、OpenVMS Cluster システムの構成に現在必要な、ディスク容量の概算容量になります。

5.2.2 その他のディスク容量の必要量

合計ディスク容量の必要量を決定する前に、オンライン・ストレージやバックアップ・ストレージの将来的な増加量を検討することも可能です。

たとえば、OpenVMS Cluster システムでは、どの程度の割合で新規ファイルが作成されるでしょうか。この値を見積り、表 5-2 で計算した合計ディスク記憶容量に追加すれば、オンライン・ストレージにおける現在と将来のニーズをより正確に計算することができます。

バックアップ・ストレージの必要量を決定するには、古いデータやアーカイブ・データの扱いを決める必要があります。ほとんどのストレージ・サブシステムでは、新しいファイルが積極的に使用される一方で、古いファイルは使用されなくなります。オンライン・ストレージからバックアップ・ストレージに定期的に古いファイルを移動することで、オンライン・ストレージを新しいファイルに空けることができ、オンライン・ストレージの必要量を制御することができます。

バックアップ・ストレージの容量を適切に維持することで、アーカイブ処理の効果が意味を持つとともに、オンライン・ストレージの必要量を節約することができます。

5.3 ディスク・パフォーマンス・オプティマイザの決定

推定ディスク・パフォーマンス作業負荷の見積りと、作業負荷データの解析により、ディスク・パフォーマンスの必要量を決定することができます。

Monitor ユーティリティと DECamsd では、アプリケーションとビジネスのニーズに対して、どのパフォーマンス・オプティマイザが最適であるかを判断できます。

5.3.1 パフォーマンス・オプティマイザ

パフォーマンス・オプティマイザは、アプリケーションとデータ用にストレージのパフォーマンスを強化するためのソフトウェア製品やハードウェア製品です。表 5-3 は、各種パフォーマンス・オプティマイザの働きをまとめたものです。

表 5-3 ディスク・パフォーマンス・オプティマイザ

オプティマイザ	説明
DECram for OpenVMS	システム管理者がメモリ内に論理ディスクを作成し、I/O パフォーマンスを強化するために必要なディスク・デバイス・ドライバ。メモリ内 DECram ディスク上のデータは、ハードウェア・ディスク上のデータよりも高速でアクセスできます。DECram ディスクは、Volume Shadowing for OpenVMS でシャドウ・セットにすることができ、MSCP サーバによるサービスも可能です。 ¹
半導体ディスク (SSD)	一般のシステムでは、I/O 要求の約 80%がオンライン保存データの約 20%の部分に集中します。半導体デバイスは、このデータ・サブセットに必要な高速アクセスを提供します。

¹ MSCP サーバは、直接アクセスが可能なローカル接続ディスクを OpenVMS Cluster 上の他のシステムに開放します。

(次ページに続く)

表 5-3 (続き) ディスク・パフォーマンス・最適化

最適化	説明
ディスク・ストライピング	<p>ディスク・ストライピング (RAID レベル 0) では、アプリケーションがディスク・ドライブ・アレイを並列にアクセスでき、高いスループットを実現できます。ディスク・ストライピングでは、いくつかのディスクが“ストライプ・セット”にグループ化され、アプリケーション・データは“チャンク”に分割され、これによってデータはストライプ・セット内のディスクに総当たり的に均等に分散します。</p> <p>アクセス時間が節約されるので、ディスク・ストライピングでは特に以下のアプリケーションでパフォーマンスが向上します。</p> <ul style="list-style-type: none"> • 大量のデータを並列転送する場合。 • 複数のドライブ間に負荷を分散させる必要がある場合。
	<p>ディスク・ストライピングは 2 種類あります。</p> <ul style="list-style-type: none"> • コントローラ方式のストライピング。HSJ コントローラと HSG コントローラがいくつかのディスクを 1 つのストライプ・セットに結合します。このストライプ・セットは、1 つのボリュームとして OpenVMS を構成します。この種のディスク・ストライピングはハードウェア方式です。 • ホスト方式のストライピング。RAID for OpenVMS を使用して、OpenVMS ホストにストライプ・セットを作成します。OpenVMS ソフトウェアは 1 つの I/O 要求を複数の同時要求に分割し、それをストライプ・セットのディスクに送信します。この種のディスク・ストライピングはソフトウェア方式です。
	<p>注意: Volume Shadowing for OpenVMS ソフトウェアをディスク・ストライピングと組み合わせると、ストライプ・セット・メンバを冗長化できます。コントローラ方式のストライプ・セットをシャドウ化したり、ホスト方式のストライプ・セットをシャドウ化することができます。</p>
拡張ファイル・キャッシュ (XFC)	<p>OpenVMS Alpha には、XFC (extended file cache) を持つホスト方式のキャッシュが用意されています。これは、仮想 I/O キャッシュ (VIOC) と置き換えることも共存させることもできます。XFC はクラスタ規模のファイル・システム・データ・キャッシュで、VIOC では使用できなかった機能をいくつか持っています。パフォーマンスを改善する、先読みキャッシングやキャッシュの自動サイズ変更機能などがこれに含まれます。OpenVMS Integrity も XFC をサポートしていますが、VIOC はサポートしていません。</p>
ディスク・キャッシュ付きコントローラ	<p>ストレージ・テクノロジーによっては、メモリでディスク・キャッシュを実装するものがあります。キャッシュのデータに対応できるアクセスは、シーク・タイムやローテーション待ち時間なしで即座に処理されます。このようなアクセスでは、I/O 応答時間の 2 つの大きな要素が取り除かれます。HSZ コントローラ、HSG コントローラにはキャッシュが組み込まれています。RF ディスクと RZ ディスクにはすべて、ディスク・キャッシュが埋め込みコントローラの一部として組み込まれています。</p>

関連項目：以上のパフォーマンス・最適化で、OpenVMS Cluster における I/O のスケーラビリティが強化される様子については、第 9.6 節を参照してください。

5.4 ディスクの可用性の必要要件の決定

ストレージ・サブシステムでは、可用性はストレージ・デバイスの可用性とデバイスへのパスの可用性によって決まります。

5.4.1 可用性の必要要件

ストレージ・サブシステムの可用性を最適化して強化するには、コストがかかります。可用性コストの解析では、障害発生時に利用できなくなるデータのコストに対してデータ保護に必要なコストを比較検討します。ビジネスの特性によっては、ストレージ・サブシステムの障害による影響は、軽度の影響、中程度の影響、重度の影響に分けることができます。

デバイスとデータの可用性オプションを利用すれば、ストレージ・サブシステムの障害時の影響を緩和したり、場合によっては帳消しにすることができます。

5.4.2 デバイスとデータの可用性オプティマイザ

可用性の必要要件によっては、最も優先順位の高いアプリケーションとデータを対象にして、表 5-4 の説明に従って可用性オプティマイザを選択してください。

表 5-4 ストレージの可用性オプティマイザ

可用性オプティマイザ	説明
冗長アクセス・パス	データまでの冗長アクセス・パスを構成することで、デバイスまでのパスにおけるハードウェア障害を防ぎます。
Volume Shadowing for OpenVMS ソフトウェア	シャドウ・セットを構成する 1 つ以上の物理的に同一なディスクにデータを書き込むことで、仮想ディスクに書き込まれたデータの複製を作成します。複製データを利用して、あるディスクが使用できなくても、ユーザは目的のデータをアクセスできます。また、シャドウ・セットのメンバに障害が発生しても、シャドウ・ソフトウェアにより、そのドライブはシャドウ・セットから取り除かれ、残ったドライブを利用して処理が続行されます。シャドウ処理はアプリケーションには透過的に実行され、メディア、コントローラ、インターコネクトの障害が発生している間もデータの保存や配信が可能です。 シャドウ・セットには、最大で 3 つのメンバを組み込むことができ、シャドウ・メンバ・セットは、OpenVMS Cluster のストレージ・サブシステム内のどこにでも配置できます。 関連項目：ボリューム共有の詳細については、『Volume Shadowing for OpenVMS 説明書』を参照してください。

(次ページに続く)

表 5-4 (続き) ストレージの可用性オプティマイザ

可用性オプティマイザ	説明
システム・ディスクの冗長化	複数のアクセス・パスがあるディスク・ドライブにシステム・ファイルを配置します。システム・ディスクを組み込んだシャドウ・セットを構成すると、OpenVMS Cluster の可用性が増します。また、複数のシステム・ディスクで OpenVMS Cluster システムを構成することもできます。 関連項目：詳細については、第 10.2 節を参照してください。
データベースの冗長化	バッチ・ジョブで夜間に更新されるデータベースの一定のファイルやパーティションの冗長コピーを保存します。シャドウ・セットではディスク全体のコピーが作成されますが、この方法では、別のディスクにバックアップ・コピーを作成するか、ファイルやデータベースを選択してスタンバイ・テープにバックアップ・コピーを作成するだけで済みます。
新しいデバイス	新しいデバイスを選択して障害に備えます。一般に、デバイスが新しいほど信頼性と平均故障時間 (MTBF) がすぐれています。また、コントローラが新しいほど、最新のチップ・テクノロジーで信頼性が高くなります。
包括的なバックアップ計画	周期の短い定期的なバックアップは、データの可用性を保証する最も効果的な方法です。 関連項目：Fibre Channel テープ・サポートの詳細については、第 7.5 節を参照してください。バックアップ計画と OpenVMS Backup の詳細については、『OpenVMS システム管理者マニュアル』を参照してください。バックアップ・ソフトウェアとバックアップ・ソリューションの詳細については、 http://h18006.www1.hp.com/storage/tapestorage.html および http://h71000.www7.hp.com/openvms/storage.html を参照してください。

5.5 SAS 方式のストレージ

SAS は、シリアル通信 (一度に 1 ビットずつ) を使用して SCSI ストレージ・デバイスとの間でデータを伝送するポイントツーポイント (2 地点間) アーキテクチャです。

5.5.1 ストレージ・デバイス

デュアルドメイン SAS は、単一の SAS ドメイン経路で障害が発生した場合の解決策として別のドメインを作成します。追加ドメインは、デュアルドメイン SAS をサポートする HP Smart Array コントローラ上のオープン・ポートを使用します。デュアル・ドメインをサポートする Smart Array 上の 2 つ目のポートはユニークな識別子を生成し、そのドメイン内でサポートします。

以下の SAS コントローラがサポートされます。

- LSI 1068
- LSI Logic 1068e

以下の SMART アレイは SAS バックプレーンを持っていますが、SAS HBA と考えることはできません。

- P400i
- P411
- P700m
- P800

SAS でサポートされる外部コントローラは無く、MSA60/70 などの JABOD か内臓ディスクのみを SAS HBA に接続できます。ただし、P700m は MSA2000SA (SAS 対応の MSA2000) に接続できます。

5.6 SCSI 方式のストレージ

SCSI (Small Computer Systems Interface) バスは、ANSI 業界標準のストレージ・インターコネクタです。SCSI バスには合計で 8 ノードまたは 16 ノード (内 3 ノードを CPU にすることが可能) を接続できます。

5.6.1 サポートされているデバイス

次のデバイスは、1 本のホスト SCSI バスまたは複数の SCSI バスに接続できます。

- RZ シリーズ・ディスク
- HSZ ストレージ・コントローラ

以下のデバイスは、1 本のホスト SCSI バスにだけ接続できます。

- EZ シリーズ・ディスク
- RRD シリーズ CD-ROM
- TZ シリーズ・テープ

5.7 Fibre Channel 方式のストレージ

Fibre Channel インターコネクタは、ANSI 業界標準のストレージ・インターコネクタです。

5.7.1 ストレージ・デバイス

HSG ストレージ・コントローラおよび HSV ストレージ・コントローラは、1 つのホスト Fibre Channel インターコネクタまたは複数の Fibre Channel インターコネクタに接続できます。Fibre Channel ハードウェア・サポートの詳細については、第 7.2 節を参照してください。

5.8 ホスト方式のストレージ

ホスト方式のストレージ・デバイスは、ローカル・アダプタにより OpenVMS Cluster メンバ・システムにローカルに接続できます。ローカルに接続されたこのストレージは、ノードを MSCP サーバとして構成すれば、他の OpenVMS Cluster メンバにも開放できます。

ローカル・アダプタでは、各ディスクを 2 本のアクセス・パスに接続できます (デュアル・ポート)。デュアル・ポーティングにより、ノード間のディスクの自動フェールオーバーが可能になります。

5.8.1 内部バス

ローカルに接続されたストレージ・デバイスは、システムの内部バスに接続します。

- PCI
- PCI-X
- PCI-Express
- EISA
- ISA
- XMI
- SCSI
- TURBOchannel
- Futurebus+

サポートするバスについての詳細は、『HP OpenVMS I/O User's Reference Manual』を参照してください。

5.8.2 ローカル・アダプタ

次に示すのは、ローカル・アダプタとそのバス・タイプです。

- KGPSA (PCI)
- KZPSM (PCI)
- KZPDA (PCI)
- KZPSC (PCI)
- KZPAC (PCI)
- KZESC (EISA)
- KZMSA (XMI)

- PB2HA (EISA)
- PMAZB (TURBOchannel)
- PMAZC (TURBOchannel)
- KDM70 (XMI)
- KDB50 (VAXBI)
- KDA50 (Q-bus)

SCSI と Fibre Channel ストレージに対するマルチパスの構成

この章では、マルチパス SCSI サポートについて説明します。この章の内容は、特に注記のない限り、ディスクとテープに適用されます。この SCSI プロトコルは、パラレル SCSI インターコネクと Fibre Channel インターコネクの両方で使用します。なお、この章で SCSI とは、パラレル SCSI デバイスまたは Fibre Channel (FC) デバイスを意味します。

注意

OpenVMS Alpha のバージョン 7.3-1 から SCSI ディスク・デバイスへのローカル・パスと MSCP サービス対象パスとの間のフェールオーバが導入されました。この機能は、MPDEV_REMOTE システム・パラメータの設定を 1 (デフォルト設定) にすると有効化できます。このタイプのフェールオーバは、テープ・デバイスには適用されません。

他社製のディスク・キャッシュ製品、ディスク・シャドウ製品、または類似の製品には、SCSI マルチパス機能を使用できないものがあります。特に、OpenVMS Alpha SCSI ディスク・クラス・ドライバ (SYS\$DKDRIVER.EXE)、OpenVMS Alpha SCSI テープ・クラス・ドライバ (SYS\$MKDRIVER.EXE、あるいは SCSI 汎用クラス・ドライバ (SYS\$GKDRIVER) のいずれかのドライバ・ディスパッチ・テーブル (DDT) の変更依存する他社製の製品は、SCSI マルチパス機能を使用した場合に正しく動作するために修正が必要になる場合があります。ソフトウェアの製造元でマルチパス機能がサポートされるまで、マルチパスフェールオーバに合わせて構成された SCSI デバイス (HSZ70 コントローラと HSZ80 コントローラにマルチパス・モードで接続された SCSI デバイスなど) では、このようなソフトウェアを使用しないでください。HP は、SCSI マルチパス・フェールオーバと互換性を維持するためにこれらの他社製品を修正するためのドライバ・ディスパッチ・テーブル (DDT) ルーチンを提供します。

マルチパス SCSI 機能を使用する上での主な要件と制限事項については、第 6.2 節を参照してください。

この章では、Fibre Channel とパラレル SCSI インターコネクを模式的に表現しています。いずれもノードとストレージ・サブシステムが接続されている水平線で表します。図 7-1 に示すように、Fibre Channel インターコネクは、物理的にはスイッチから常に放射状に配線されます。パラレル SCSI は、ハブに対して放射状に配線できます。また、デイジー・チェーン・パスにすることもできます。

ストレージ・サブシステムの複数の SCSI ディスクと SCSI バスの表現も、図では単純化しています。1 つ以上の HSZx コントローラ、HSGx コントローラ、または HSVx コントローラが論理ユニットとしてホストに提供する複数のディスクと SCSI バスは、図では 1 つの論理ユニットで表しています。

この章の構成は、以下のとおりです。

- マルチパス SCSI サポートの概要 (第 6.1 節)
- 構成要件と制限事項 (第 6.2 節)
- HSxフェールオーバー・モード (第 6.3 節)
- パラレル SCSI マルチパス構成 (ディスクのみ) (第 6.4 節)
- パラレル SCSI マルチパス・ディスク構成のデバイス命名 (第 6.5 節)
- Fibre Channel マルチパス構成 (第 6.6 節)
- マルチパス構成の実装 (第 6.7 節)

6.1 マルチパス SCSI サポートの概要

マルチパス SCSI 構成では、デバイスまでの 1 本のパスに対し、同じデバイスまでの別のパスによるフェールオーバーを用意します。同じデバイスに複数のパスがあれば I/O 動作におけるそのデバイスの可用性を強化できます。また複数のパスにより全体的なパフォーマンスも向上します。図 6-1 は、マルチパス SCSI 構成を示したものです。1 台のコンピュータから同じ仮想ストレージ・デバイスに 2 本のパスが構成されています。

ディスク・デバイス用のマルチパス SCSI 構成は、パラレル SCSI と Fibre Channel のどちらにおいても図 6-1 に示すように、ストレージ・インターコネクトとして使用できます。テープ・デバイス用のマルチパス SCSI 構成は、Fibre Channel においてのみ、ストレージ・インターコネクトとして使用できます。

1 つのデバイスに複数のパスがあるとき、これをマルチパス・セットといいます。システムがデバイスにパスを構成するとき、名前は同じでパスが異なるデバイスがないか確認します。該当するデバイスがあり、マルチパス・サポートが有効な場合は、システムはマルチパス・セットを形成するか、既存セットに新しいパスを追加します。マルチパス・サポートが無効な場合には、1 つのデバイスに複数のパスは構成されません。

システムはマルチパス・セットを 1 つのデバイスで表します。システムは、デバイスまでの 1 本のパスを“現在の”パスとして選択し、障害が発生するか、システム管理者からシステムに別のパスへの切り替えが要求されるまで、すべての I/O はこのパスで行われます。

マルチパス SCSI サポートには、以下のようなフェールオーバーがあります。

- 直接 SCSI から直接 SCSI (直接接続された SCSI パスから直接接続された SCSI パスへ)
- 直接 SCSI から MSCP サービス対象 (直接接続された SCSI パスから MSCP でサービスされているパスへ) (ディスクのみ)

- MSCP サービス対象から直接 SCSI (MSCP でサービスされているパスから直接接続された SCSI パスへ) (ディスクのみ)

直接 SCSI から直接 SCSI フェールオーバには、マルチポート SCSI デバイスを使用します。直接 SCSI から MSCP サービス対象のフェールオーバには、SCSI パスごとに複数のホストが必要ですが、マルチポート SCSI デバイスは必要ありません。これら 2 種類のフェールオーバは組み合わせることができます。それぞれ単独で使用した場合と組み合わせた場合について以下の項で説明します。

6.1.1 直接 SCSI から直接 SCSI へのフェールオーバ

直接 SCSI から直接 SCSI へのフェールオーバは、マルチポート SCSI デバイスのシステムに使用できます。マルチポート SCSI デバイスの例としては、デュアル HSZ70, HSZ80, HSG80, デュアル MDR, HSV110 があります。マルチポート SCSI デバイスは、同じ物理インターコネクト上の複数のポートで構成できるので、ポートのどれかに障害が発生しても、ホストは別のポートからデバイスにアクセスできます。これを透過的フェールオーバ・モードといい、ディスク・デバイスに対してバージョン 6.2以降の OpenVMS でサポートしています。

OpenVMS バージョン 7.2 では、新しいフェールオーバ・モードのサポートを導入しました。これは、異なる物理インターコネクト上でマルチポート・ディスク・デバイスをそのポートとともに構成できる機能です。これをマルチパス・フェールオーバ・モードといいます。

HSxフェールオーバ・モードは、HSxコンソール・コマンドで選択します。透過モードとマルチパス・モードの詳細については、第 6.3 節を参照してください。

図 6-1 は、マルチパス・フェールオーバの一般的な構成を表しています。

注意

複数の直接 SCSI パスをディスク・デバイスに構成できるのは、接続されているノードのすべてにおいてマルチパス・サポートが有効になっており、HSZ/G がマルチパス・フェールオーバ・モードである場合だけです。

図 6-1 に示す 2 つの論理ディスク・デバイスは、HSxコントローラ・モジュールがシステムに提供する仮想ストレージ・ユニットを表しています。各論理ストレージ・ユニットを、2 つの HSxコントローラ・モジュールの両方で同時に“オンライン”にすることはできません。複数の論理ユニットがあるとき、異なる HSxコントローラに対してはオンラインにできるので、両方の HSxコントローラを同時にアクティブできます。

透過モードでは、他のコントローラの機能停止を HSxコントローラが検出すると論理ユニットはコントローラを切り替えます。

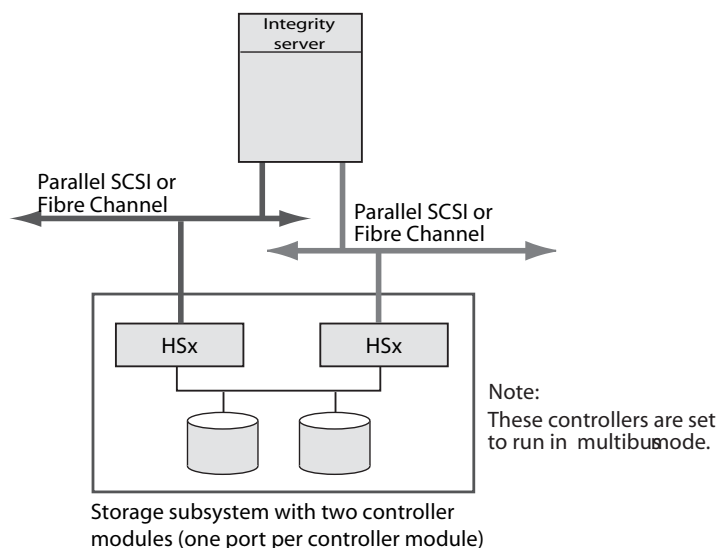
SCSI と Fibre Channel ストレージに対するマルチパスの構成

6.1 マルチパス SCSI サポートの概要

図 6-1 に示すように、マルチパス・モードでは、以下のイベントが発生すると論理ユニットによってコントローラが切り替えられます。

- 他のコントローラの機能停止を HSxコントローラが検出した。
- 現在のパス上の障害発生を OpenVMS マルチパス・ソフトウェアが検出し、コントローラ切り替えコマンドを発行した。
- コントローラ切り替えコマンドを OpenVMS システム管理者が発行した。

図 6-1 マルチパス・フェールオーバの構成



VM-0067A-AI

図 6-1 の以下の点に着目してください。

- ホストに 2 つのアダプタがある。
- 両方のインターコネクトとして、パラレル SCSI (HSZ70 または HSZ80) または Fibre Channel (HSGx または HSVx) のどちらでも使用できるが、組み合わせることはできない。
- マルチパス・フェールオーバ・モードに構成された 2 つの HSx コントローラがストレージ・デバイスのキャビネットに組み込まれている。

マルチパス構成は、透過的フェールオーバに対して以下の利点があります。

- 2 つのホスト・アダプタと 2 つの HSx コントローラ・モジュールの運用により、全体のパフォーマンスが強化される。
- パス上のホスト・アダプタ、インターコネクト、または HSx コントローラ・モジュールに障害が発生してもストレージにアクセスできるため、可用性が高い。

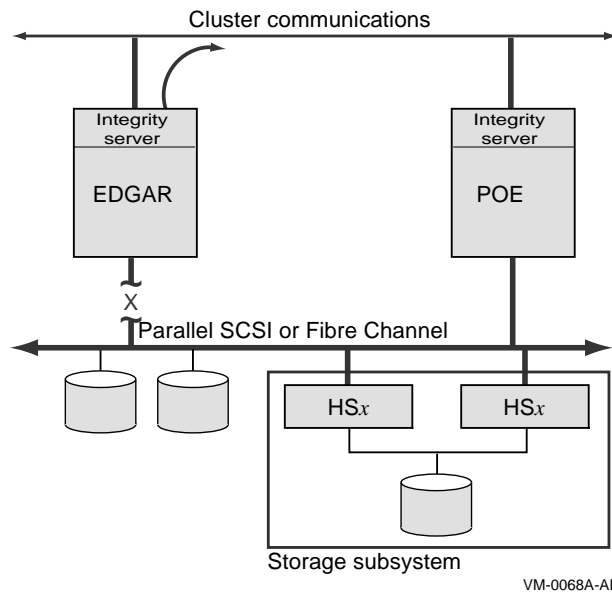
マルチパス・フェールオーバ構成では、SAS およびそのコントローラ (LSI 1068 あるいは LSI 1068e) も使用することができます。

6.1.2 直接 SCSI から MSCP サービス対象へのフェールオーバ (ディスクのみ)

OpenVMS では、SCSI バスを共有する複数のホストをサポートしています。この機能を、マルチホスト SCSI OpenVMS Cluster システムといいます。この構成では、SCSI バスは共有ストレージ・インターコネクタです。クラスタ通信は 2 番目のインターコネクタ (LAN) を通じて行われます。

図 6-2 に示すように、マルチホスト SCSI OpenVMS Cluster システムのマルチパス・サポートでは、直接関連付けられた SCSI から MSCP のサービス対象である SCSI ストレージへ、フェールオーバが可能です。

図 6-2 1つのインターコネクタによる直接 SCSI から MSCP のサービス対象構成



この構成では、次の点に着目してください。

- 共有ストレージ・インターコネクタに 2 つのホストが接続されている。
- 2 つのホストは、クラスタ通信の 2 番目のインターコネクタ (LAN) で接続されている。
- ストレージ・デバイスのポートは 1 つでも複数でもよい。
- ストレージまでのノード Edgar の SCSI 接続に障害が発生すると、SCSI ストレージはクラスタ・インターコネクタで接続された残りのホストにより、MSCP サービスの対象になる。

このようなマルチホスト SCSI OpenVMS Cluster システムでのマルチパス・サポートでも、MSCP サービス対象 SCSI ストレージから直接接続された SCSI ストレージへのフェールオーバーは有効化されます。たとえば、図 6-2 の構成では、以下のようなイベントが発生する可能性があります。

- ノード POE はノード EDGAR を MSCP サーバとして使用して、共用ストレージ・インターコネクタ上のストレージ・デバイスにアクセスしている。
- ノード EDGAR で、共用ストレージへの直接接続に失敗するか、ノード EDGAR が停止するか、またはノード EDGAR はクラスタ・インターコネクタ経由で到達できなくなる。
- ノード POE では、共用ストレージへの直接パスを使用するように切り替えが行われる。

注意

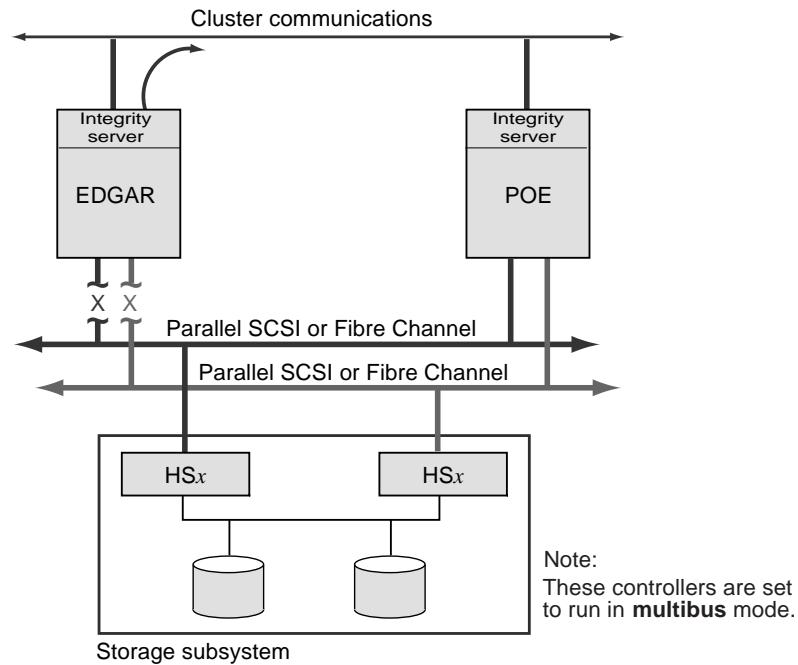
本書では、直接 SCSI から MSCP サービス対象パスへのフェールオーバー機能は、直接パスとサービス対象パスとの間のいずれかの方向におけるフェールオーバー機能を意味します。

直接 SCSI から MSCP サービス対象へのフェールオーバー構成の場合、SAS とそのコントローラ (LSI 1068 あるいは LSI 1068e) も使用できます。

6.1.3 両方の種類のマルチパス・フェールオーバーを組み合わせた構成

図 6-3 に示すように、マルチホスト SCSI OpenVMS クラスタ・システムでは、クラスタを両方の種類のマルチパス・フェールオーバー (直接 SCSI から直接 SCSI と直接 SCSI から MSCP サービス対象の SCSI) に構成することによりディスク・ストレージの可用性を強化できます。

図 6-3 インターコネクトが2つの直接 SCSI から MSCP サービス対象構成



VM-0069A-AI

この構成では、次の点に着目してください。

- 両方のノードが両方のストレージ・インターコネクトに直接接続されている。
- 両方のノードがクラスタ通信の2番目のインターコネクトに接続されている。
- 各 HSxストレージ・コントローラが1本のインターコネクトにだけ接続されている。
- 両方の HSxストレージ・コントローラは同じキャビネットにある。

この構成には、「直接 SCSI フェールオーバー」と「直接から MSCP サービス対象 SCSI へのフェールオーバー」の両方の利点があります。

このタイプの構成では、SAS およびそのコントローラ (LSI 1068 あるいは LSI 1068e) も使用できます。

6.2 構成要件と制限事項

マルチパス SCSI と FC 構成の要件については、表 6-1 を参照してください。

表 6-1 マルチパス SCSI と FC 構成の要件

構成要素	説明
ホスト・アダプタ	パラレル SCSI の場合、KZPBA-CB を使用します。OpenVMS でディスク・マルチパス・フェールオーバをサポートする唯一の SCSI ホスト・アダプタです。Fibre Channel の場合、KGPSA-BC と KGPSA-CA の両方が OpenVMS でマルチパス・フェールオーバをサポートします。
Alpha コンソール・ファームウェア	HSZ70 と HSZ80 を備えたシステムでは、最小更新レベルは、5.3 または 5.4 であり、AlphaServer によって異なります。HSG80 を備えたシステムでは、最小更新レベルは 5.4 です。
コントローラ・ファームウェア	HSZ70 の場合、最小更新レベルは 7.3 です。HSZ80 の場合、8.3 です。HSG80 の場合、8.4 です。MDR の場合、最小更新レベルは 1170 です。
コントローラ・モジュール・モード	マルチパス・モードに設定します (ディスクのみ)。選択は、HSxコンソールで行います。
全接続	<p>マルチパス・モードで HSx に接続されているすべてのホストには、両方の HSx コントローラ・モジュールまでのパスが必要です。異なるコントローラに排他的に接続されているホストは、コントローラ間で論理ユニットをあちらこちらに切り替え、I/O が実行されるのを防ぐためです。</p> <p>これを防ぐには、常にホストからコントローラ・モジュールへ全接続されるようにしておきます。コントローラとのホストの接続に障害が発生すると、以下のどれかの操作が実行され、パスの切り替えが無限に繰り返されるのを防ぎます。</p> <ul style="list-style-type: none">• 接続を迅速に修復する。• 他のホストが特定の接続のコントローラに切り替えられるのを防ぐ。そのためには、部分接続のコントローラの原因になるパスへの切り替えを無効にするか (第 6.7.11 項参照)、部分的に接続されているコントローラをシャットダウンする。• 部分的に接続されたホストを両方のコントローラから切り離す。

(次ページに続く)

表 6-1 (続き) マルチパス SCSI と FC 構成の要件

構成要素	説明
割り当てクラス	<p>パラレル SCSI の場合、有効な HSZ 割り当てクラスが必要です (第 6.5.3 項参照)。SCSI バスが HSZ コントローラだけで構成され、すべてのコントローラが有効な HSZ 割り当てクラスを持つ場合、そのバスの古い SCSI デバイス命名規則に従う必要はありません。つまり、そのアダプタに同じポート割り当てクラスを割り当てたり、同じノード割り当てクラスと同じ OpenVMS アダプタ・デバイス名を割り当てする必要はありません。</p> <p>ただし、非 HSZ デバイスがバスにある場合、あるいは HSZ 割り当てクラスなしの HSZ コントローラがある場合、ノードとポート割り当てクラス代入、およびコントローラ・デバイスの名前については共用 SCSI バスの標準ルールに従う必要があります。</p> <p>AlphaServer 2x00(A) を除き、HSZ 割り当てクラスによるデバイスからのブートは、KZPBA-CB をサポートするすべての AlphaServer でサポートされています。</p> <p>コントローラ割り当てクラスは、FC デバイスには使用しません。</p> <p>注意: Volume Shadowing for OpenVMS を使用している場合は、すべてのディスクに 0 以外の割り当てクラスが必要です。HSG コントローラと HSV コントローラに接続されているすべての FC ディスクには、割り当てクラス値として 1 が自動的に代入されます。この値は、ボリューム・シャドウイング要件を満たす値です。</p>

マルチパス FC と SCSI 構成の制限事項については、表 6-2 を参照してください。

表 6-2 マルチパス FC と SCSI 構成の制限事項

機能	説明
サポートされているデバイス	<p>コントローラ・モジュール HSZ70, HSZ80, HSG60, HSG80, および HSV110 に関連付けられている DKDRIVER ディスク・デバイスをサポートします。MDR に関連付けられている MKDRIVER テープと GKDRIVER テープ・ロボットはサポートされます。テープなどの他のデバイス・タイプや GKDRIVER などの汎用クラス・ドライバはサポートしていません。</p> <p>大きな負荷がかかっている場合、HSZ70 または HSZ80 コントローラでは、ホストによるコントローラの手動または自動切り替えができない場合があります。テストの結果、この問題はめったに起こらないことが判明しています。この問題は、ファームウェア HSOE V7.7 およびそれ以降のバージョンを使用する HSZ70 では修正済みです。将来のリリースでは HSZ80 についても修正される予定です。この問題は HSGx コントローラ、HSVx コントローラ、MDR では発生しません。</p>

(次ページに続く)

表 6-2 (続き) マルチパス FC と SCSI 構成の制限事項

機能	説明										
混成バージョンと混成アーキテクチャ・クラスタ	<p>HSZ, HSG, または HSV にマルチパス・モードで接続されているホストでは, OpenVMS バージョン 7.2 またはそれ以上を実行する必要があります。</p> <p>サービス対象パスへのマルチパス・フェールオーバを使用するには, 共有 SCSI または Fibre Channel デバイスへ直接アクセスしているすべてのシステムで MPDEV_REMOTE を有効にする必要があります。</p>										
SCSI から MSCP へのフェールオーバ MSCP から SCSI へのフェールオーバ	<p>マルチパス・ホストは, 共有 SCSI パス (パラレル SCSI または Fibre Channel) 経由で SCSI ディスク・デバイスに接続する必要があります。これらのホストでは MPDEV_REMOTE システム・パラメータを 1 に設定する必要があります。</p>										
Volume Shadowing for OpenVMS	<p>一定のシステム・パラメータにデフォルト設定を使用すると, マルチパス・サポート用に構成されたシャドウ・セット・メンバが削除されることがあります。このような現象は, Volume Shadowing for OpenVMS を使用しているシャドウ・セット・メンバに見られます。</p> <p>これを防ぐには, Volume Shadowing for OpenVMS でマルチパス・シャドウ・セットを構成するときは, 以下の推奨設定をこれらのパラメータに使用してください。</p>										
	<table border="1"> <thead> <tr> <th>システム・パラメータ</th> <th>推奨設定</th> </tr> </thead> <tbody> <tr> <td>MSCP_CMD_TMO</td> <td>最小値として 60。 ほとんどの構成に 60 が適切ですが, 構成によってはさらに高い値が必要なものもあります。</td> </tr> <tr> <td>SHADOW_MBR_TMO</td> <td>MSCP_CMD_TMO の 3 倍以上</td> </tr> <tr> <td>SHADOW_SYS_TMO</td> <td>MSCP_CMD_TMO の 3 倍以上</td> </tr> <tr> <td>MVTIMEOUT</td> <td>SHADOW_MBR_TMO の 4 倍以上</td> </tr> </tbody> </table>	システム・パラメータ	推奨設定	MSCP_CMD_TMO	最小値として 60。 ほとんどの構成に 60 が適切ですが, 構成によってはさらに高い値が必要なものもあります。	SHADOW_MBR_TMO	MSCP_CMD_TMO の 3 倍以上	SHADOW_SYS_TMO	MSCP_CMD_TMO の 3 倍以上	MVTIMEOUT	SHADOW_MBR_TMO の 4 倍以上
システム・パラメータ	推奨設定										
MSCP_CMD_TMO	最小値として 60。 ほとんどの構成に 60 が適切ですが, 構成によってはさらに高い値が必要なものもあります。										
SHADOW_MBR_TMO	MSCP_CMD_TMO の 3 倍以上										
SHADOW_SYS_TMO	MSCP_CMD_TMO の 3 倍以上										
MVTIMEOUT	SHADOW_MBR_TMO の 4 倍以上										

6.3 HSxフェールオーバ・モード

デュアル冗長構成, 透過的フェールオーバ・モード, マルチパス・フェールオーバ・モードの場合, HSZ70, HSZ80, HSGxは, フェールオーバ操作の2つのモードを実装しています。

注意

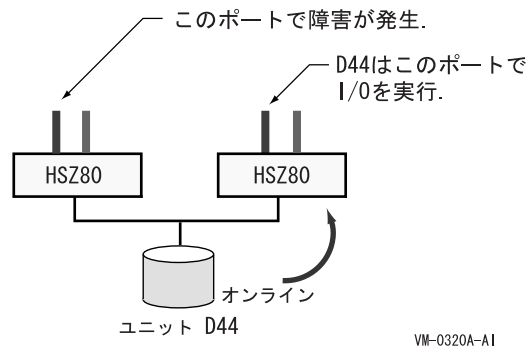
OpenVMS Alpha バージョン 7.3 から, HSGxで透過的フェールオーバ・モードがサポートされるようになりました。

以下の項で説明しますが, HSxフェールオーバ・モードでシステムを正しく機能させるためには, インターコネクト・ハードウェアとホスト・オペレーティング・システム・ソフトウェアの構成との互換性が必要です。

6.3.1 透過的フェールオーバ・モード

透過的フェールオーバ・モードでは、HSxはデュアル・コントローラ・ペアの1つのポート上に各論理ユニットを提供します。さまざまなポートにさまざまな論理ユニットを割り当てることができますが、個々の論理ユニットは、同時に1つのポートからでないとアクセスできません。図 6-4 のように、コントローラ・モジュールの障害をHSZが検出すると、HSZは正常なコントローラ上の対応するポートに論理ユニットを移動します。

図 6-4 透過モードにおけるストレージ・サブシステム



透過モードでは、2つのポートが同じホスト・バスにあることが前提なので、ホストのデバイス・ビューを変更しなくても論理ユニットがポート間を移動できます。このフェールオーバ・モードでバス構成が正しいことを確認するのはシステム・マネージャです。OpenVMSでは、HSZコントローラの透過的フェールオーバをバージョン6.2からサポートしています。

HSZまたはHSGで透過的フェールオーバ・モードを選択するには、構成に従って以下のコマンドのどれかをコンソールに入力します。

```
HSZ> SET FAILOVER COPY=THIS_CONTROLLER
```

または

```
HSZ> SET FAILOVER COPY=OTHER_CONTROLLER
```

透過モードのHSZにおけるコンソールSHOWコマンドの例を以下に示します。

SCSI と Fibre Channel ストレージに対するマルチパスの構成

6.3 HSxフェールオーバ・モード

```
z70_A => SHOW THIS_CONTROLLER
Controller:
  HSZ70 ZG64100160 firmware XB32-0, Hardware CX25
  Configured for dual-redundancy with ZG64100136
  In dual-redundant configuration
  Device Port SCSI address 7
  Time: 02-DEC-1998 09:22:09
Host port:
  SCSI target(s) (0, 2, 3, 4, 5, 6)
  Preferred target(s) (3, 5)
  TRANSFER_RATE_REQUESTED = 20MHZ
  Host Functionality Mode = A
  Allocation class          0
  Command Console LUN is target 0, lun 1
Cache:
  32 megabyte write cache, version 4
  Cache is GOOD
  Battery is GOOD
  No unflushed data in cache
  CACHE_FLUSH_TIMER = DEFAULT (10 seconds)
  NOCACHE_UPS
```

6.3.2 マルチパス・フェールオーバ・モード (ディスクのみ)

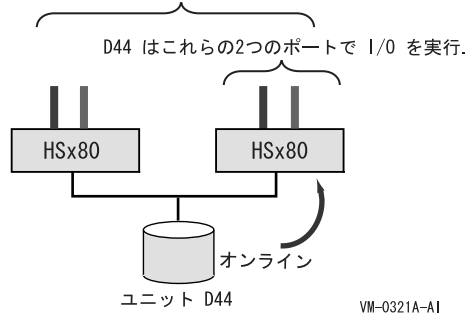
マルチパス・フェールオーバ・モードであるとき、HSxは、デュアル・コントローラ・ペアのポート上にあるホストから SCSI 照会に対しては、必ず応答します。これにより、デバイスに対するすべてのパスを認識できます。デバイスまでのすべてのパスがホストによって認識される場合、次の2つの利点があります。

- 現在のパスに障害を検出するとホストが代替パスを選択できる。これは、HSxコントローラが障害を検出したときに発生するフェールオーバに対する補助機能であり、透過モードで提供される。
- パスが同じホスト・バスでなくてもかまわない。ホストが代替パスを認識すれば、アドレス指定方式の調節により、別のパスを選択できる。そのため、SCSIバスだけで障害の原因になることがなくなる。

論理ユニットはすべてのポートで認識できますが、I/O を実行できるのはオンラインである1つのコントローラのポート上だけです。異なる論理ユニットがオンラインになることはできますが、図 6-5 に示すように、各論理ユニットは一度に1つのコントローラに対してのみオンラインになります。

図 6-5 マルチパス・モードのストレージ・サブシステム

D44 は4つのすべてのポート上の SCSI Inquiry コマンドに対応.



論理ユニットがどのコントローラに対してオンラインであるかを知るには、HSxコンソール・コマンドに次のように入力します。

```
z70_A => SHOW UNIT FULL
      LUN                                     Uses
-----
      D200                                   DISK20300
      Switches:
      RUN                                     NOWRITE_PROTECT       READ_CACHE
      MAXIMUM_CACHED_TRANSFER_SIZE = 32
      ACCESS_ID = ALL
      State:
      ONLINE to the other controller
      PREFERRED_PATH = OTHER_CONTROLLER
      Size: 2050860 blocks
```

ホストが論理ユニットに対して実行する I/O では 1 本のパスだけを使用します。これは、そのパスに障害が発生するまで変更されません。コントローラにポートが 2 つある場合、異なるホストが、論理ユニットがオンラインになっているコントローラの異なるポートを利用して同じ論理ユニットにアクセスできます。

マルチパス・フェールオーバー・モードの HSx は OpenVMS バージョン 7.2 から導入されたマルチパス機能にだけ使用できます。

マルチパス・フェールオーバー・モードを選択するには、構成に応じて、HSx: コンソールに、以下のどれかを入力します。

```
HSZ> SET MULTIBUS_FAILOVER COPY=THIS_CONTROLLER
```

または

```
HSZ> SET MULTIBUS_FAILOVER COPY=OTHER_CONTROLLER
```

SCSI と Fibre Channel ストレージに対するマルチパスの構成

6.3 HSxフェールオーバー・モード

マルチパス・モードの HSxコントローラでコンソール SHOW コマンドを実行する例を以下に示します。

```
z70_B => SHOW THIS_CONTROLLER
Controller:
  HSZ70 ZG64100136 firmware XB32-0, Hardware CX25
  Configured for MULTIBUS_FAILOVER with ZG64100160
  In dual-redundant configuration
  Device Port SCSI address 6
  Time: NOT SET
Host port:
  SCSI target(s) (0, 2, 3, 4, 5, 6)

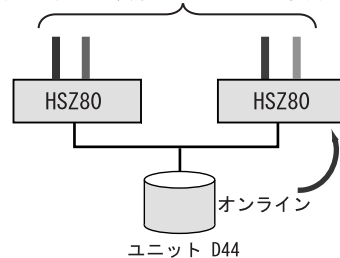
  TRANSFER_RATE_REQUESTED = 20MHZ
  Host Functionality Mode = A
  Allocation class          0
  Command Console LUN is target 0, lun 1
Cache:
  32 megabyte write cache, version 4
  Cache is GOOD
  Battery is GOOD
  No unflushed data in cache
  CACHE_FLUSH_TIMER = DEFAULT (10 seconds)
  NOCACHE_UPS
```

6.3.3 マルチパス・モードのコントローラ用のポート・アドレス指定

パラレル SCSI と FC では、マルチパス・コントローラによるアドレス指定の方法が異なります。パラレル SCSI (HSZ70 と HSZ80) では、すべてのポートが同じ SCSI ターゲット ID に割り当てられます。図 6-6 の HSZ80 の構成はこれを示したものです。

図 6-6 マルチパス・モードのパラレル SCSI コントローラのポート・アドレス指定

D44 は4つのすべてのポート上の SCSI Inquiry コマンドに対応。
4つのポートはすべて同じ SCSI ターゲットIDを持つ。
これらのポートは、別々のSCSI バス上になければならない。



VM-0322A-A1

すべてのポートに同じターゲット ID が割り当てられるのは、ターゲット ID が OpenVMS デバイス名の一部であり (たとえば \$4\$DKC600 の 6 など)、すべてのパス

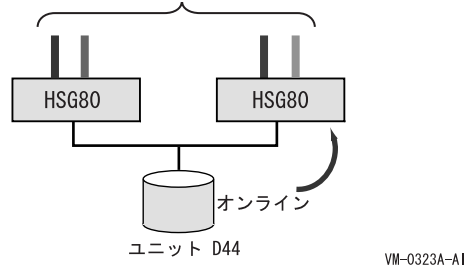
についてデバイス名が同じでなくてはならないためです。したがって、ポートは別々の SCSI バスにないとアドレスが競合します。

図 6-7 に示すように、HSGxまたはHSVxを使用した Fibre Channel 構成では、すべてのポートに固有の FC アドレスと WWID が割り当てられます。MDR の場合も同じです。

図 6-7 マルチパス・モードの Fibre Channel コントローラのポート・アドレス指定

D44 は4つのすべてのポート上の SCSI Inquiry コマンドに対応。

4つのポートはすべて異なるFCアドレス および WWIDを持つ。
ポートは同じFCインターコネクタ上であっても、
異なるFCインターコネクタ上であってもかまわない。



OpenVMS FC デバイス名に使用しないため、HSGxおよびHSVx上のポートの FC アドレスと WWID はすべて異なります。そのため、同じ FC インターコネクタに何本でもポートを接続できます。実際、インターコネクタが 1 本しかなくても、可用性とパフォーマンスの強化のために、マルチパス・モードの HSGxまたは HSVxのポートはすべて接続したほうが良いでしょう。

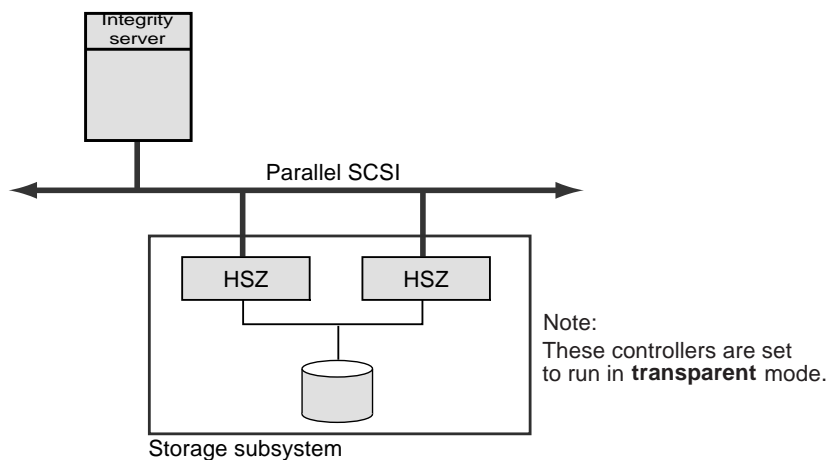
6.4 パラレル SCSI マルチパス構成 (ディスクのみ)

この節の図では、透過的フェールオーバとマルチパス・フェールオーバに構成されたシステムを表示します。HSZ80 など複数のポートを持つコントローラ・モジュールにおける注意事項をまとめています。

6.4.1 透過的フェールオーバ

図 6-8 に示すように、パラレル SCSI 構成の透過的フェールオーバでは、両方のコントローラ・モジュールが同じ SCSI バスにあるものとします。

図 6-8 透過的フェールオーバ付きの平行 SCSI 構成



VM-0070A-AI

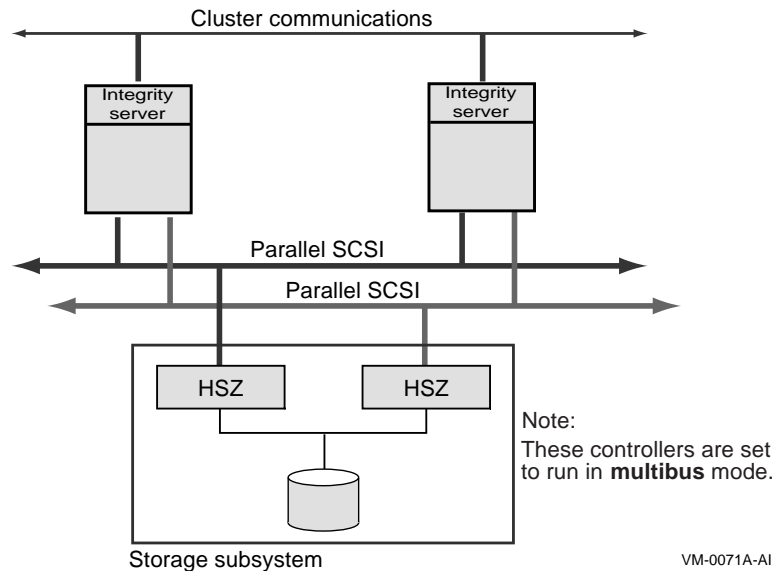
この構成の内容は以下のとおりです。

- 各論理ユニットが認識できるコントローラ・モジュールは、常に1つだけである。他のコントローラ・モジュールが同じ SCSI アドレスで応答することはない、同じアドレスを他の SCSI アドレスにも使用できる。
- 1つの HSZ コントローラ・モジュールが他のコントローラの障害を検出すると、論理ユニットを自らフェールオーバする。障害が発生したコントローラの SCSI アドレスは残ったコントローラが引き受けます。

6.4.2 マルチパス・フェールオーバと複数のパス

ホストからストレージまで複数のパスを持つ平行 SCSI 構成では、透過的フェールオーバによる構成よりも高い可用性とパフォーマンスが得られます。図 6-9 に、この構成を示します。

図 6-9 マルチパス・フェールオーバと複数のパスによる平行 SCSI 構成



この構成では、次の点に着目してください。

- 各論理ユニットは、同じ ID で両方のコントローラに認識され、構成できます。ただし論理ユニットは、一度にどちらか一方のコントローラ、つまりオンラインになっているコントローラにおける読み書き I/O にしか応答できません。
- ID 競合を防ぐため、コントローラ・モジュールには異なる SCSI パスを割り当てます。
- 以下のどちらかのイベントが発生すると、HSZ は論理ユニットを他のコントローラに移動します。
 - HSZ がコントローラ・フェールオーバを検出した。
 - ホストが論理ユニットの SCSI START コマンドを他のコントローラに送信した。

6.4.3 マルチポート・ストレージ・コントローラによる構成

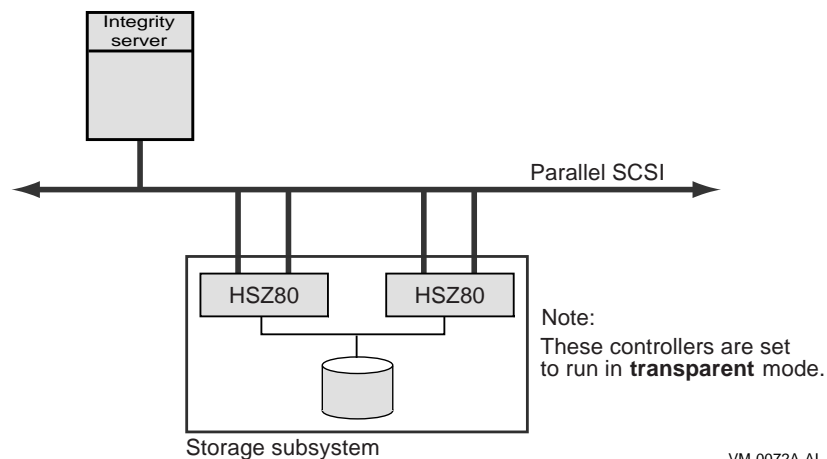
可用性とパフォーマンスを強化するには、HSZ80 などのマルチポート・ストレージ・コントローラを使用します。HSZ80 ストレージ・コントローラは HSZ70 に似ていますが、HSZ80 コントローラごとに 2 つのポートがあるという点で異なります。

この項では、マルチポート・ストレージ・コントローラを使用する 3 つの構成を示します。これらの構成は、可用性の高い順に紹介しています。

SCSI と Fibre Channel ストレージに対するマルチパスの構成 6.4 平行 SCSI マルチパス構成 (ディスクのみ)

図 6-10 は、HSZ80 を透過モードで使用するインターコネク트가 1 本のシングル・ホストです。

図 6-10 透過モードによるシングル・インターコネク트의マルチポート・平行 SCSI 構成

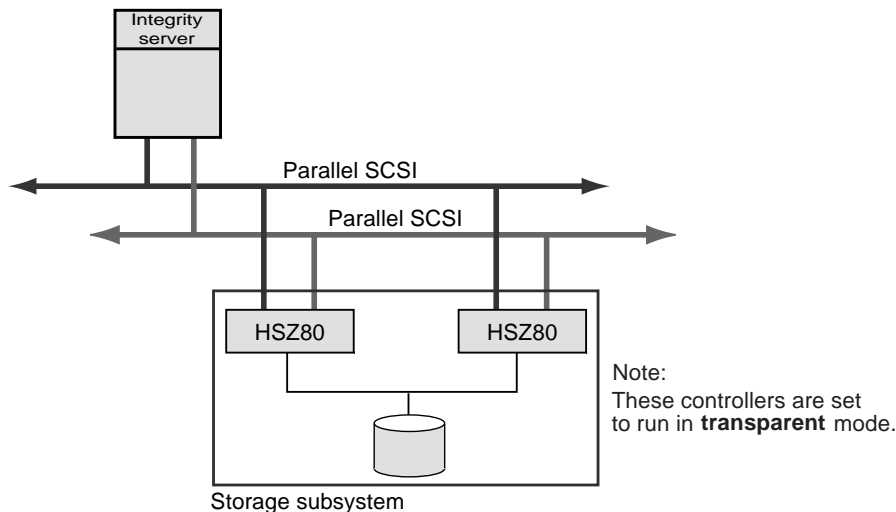


この構成では、次の点に着目してください。

- 各論理ユニットは、ストレージ・コントローラごとに 1 つのポートで認識される。
- ポートに障害が発生すると、HSZ80 は他の HSZ80 の対応するポートにトラフィックをフェールオーバーする。

図 6-11 は、ホストから 2 本のパスを使用し、透過モードで構成したシステムです。

図 6-11 透過モードによる複数パスのマルチポート・平行 SCSI 構成



VM-0088A-AI

この構成の内容は以下のとおりです。

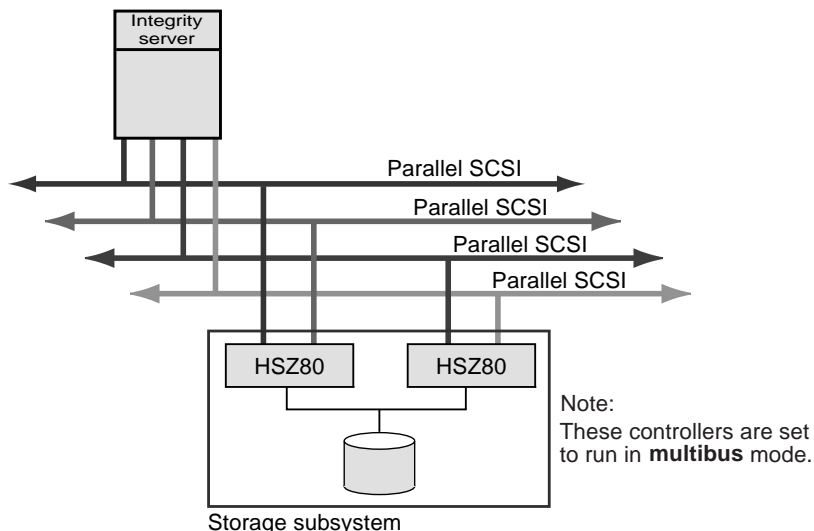
- 物理的に対応するポートは、同じ SCSI バスにあるものとする。
- ストレージ・コントローラごとにバスを最高で 2 本接続できる。

この構成では、バスが 2 本あっても、ホストから各論理ユニットまでのパスは 1 本だけです。コントローラに障害が発生すると、論理ユニットは他のコントローラの対応するポートに移動します。どちらのポートも同じホスト・バスにあるものとします。

この構成では SCSI バスを同時にアクティブにできるため、図 6-10 の構成よりもすぐれたパフォーマンスが得られます。

図 6-12 は、マルチパス・モードでマルチポート HSZ80 ストレージ・コントローラ構成を使用するシステムです。

図 6-12 マルチパス・モードによる複数パスのマルチポート・平行 SCSI 構成



VM-0073A-AI

この構成の内容は以下のとおりです。

- ホストはどのポートでも同じ ID で各論理ユニットを認識する (そのため、ポートはすべてホスト側で構成します)。
- ポートは、すべて異なる SCSI バスとする。
- ホストは、バスを 1 本ずつ使用する。
- 論理ユニットは、“オンライン”になっているコントローラの 2 つのポートに対し、I/O を同時に実行できます。つまり、複数のホストがある場合、ストレージ・デバイスに対して 2 本のバスを同時にアクティブにできます。

6.5 平行 SCSI マルチパス構成のディスク・デバイス命名

SCSI デバイス名は、システムの拡張とともに拡張され、複雑になります。拡張前の SCSI デバイス名はバスとは無縁です。デバイス名は、デバイスをアクセスするためのノード、ホスト・アダプタ、SCSI バス ID、論理ユニット番号 (LUN) を表します。以下のような理由で、ホストが複数あって、バスが複数になる環境に、バス方式の名前は不適當です。

- デバイスに直接アクセスするノードが複数あると、ノード名を使用できない。
- 共有バスのコントローラ名が一致しないと、ホスト・アダプタのコントローラ名は使用できない。
- ノードが複数のアダプタでデバイスに接続されている場合、ホスト・アダプタのコントローラ名は使用できない。

以上の制限事項のうち，最初の2つはノード割り当てクラスとポート割り当てクラスを使用すれば解決できます。3番目の問題には，HSZ コントローラ方式の割り当てクラスを使用する必要があります。これら3つの割り当てクラスについては，以下の項でその概要を説明します。

6.5.1 ノード割り当てクラスの概要

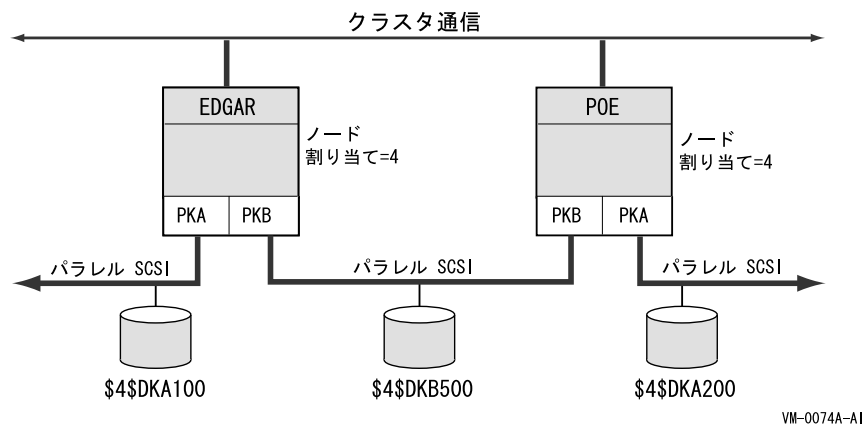
ノード割り当てクラスは，ノード名の代わりにデバイス名に使用します。同じ SCSI デバイスに対して複数のノードが直接的な接続関係にあるとき，ノード割り当てクラスが必要です。

ノード割り当てクラスは，SCSI ストレージ・デバイスとアクセスを共用するすべてのノードが次の要件にあてはまるときに使用できます。

- デバイスに直接パスを1本だけ持っている。
- 共用バスで同じホスト・コントローラ名を使用している。
- 非共用デバイスに対する固有の名前に必要かつ十分な SCSI ID を持っている。

図 6-13 は，ノード割り当てクラスで命名されたデバイスの構成です。

図 6-13 ノード割り当てクラスで命名したデバイス



6.5.2 ポート割り当てクラスの概要

デバイス名におけるポート割り当てクラスは，デバイスのアクセスに使用するホスト・アダプタを表します。デバイス名のノード割り当てクラスの代わりにポート割り当てクラスを使用し，アダプタ・コントローラ名は常に A です。

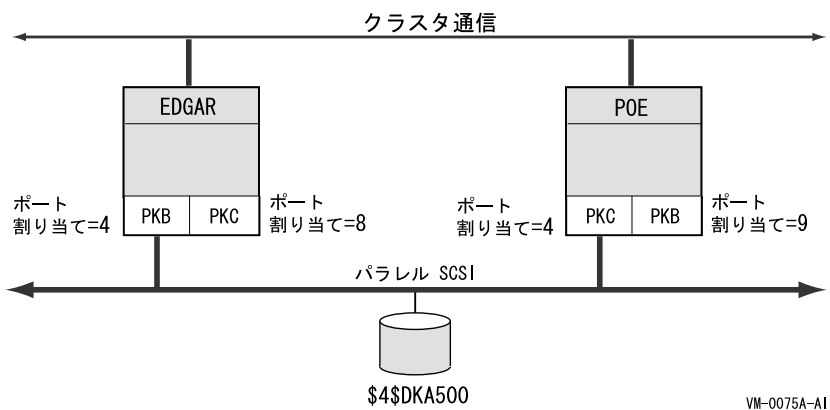
SCSI システムが固有のデバイス名を作成するとき，SCSI ID が不十分な場合や，共用バス上のアダプタのコントローラ名が一致しない場合にポート割り当てクラスを使用できます。ポート割り当てクラスをデバイス名に使用できるのは，SCSI ストレージ

SCSI と Fibre Channel ストレージに対するマルチパスの構成 6.5 平行 SCSI マルチパス構成のディスク・デバイス命名

ジ・デバイスに対するアクセスを共有するすべてのノードに、そのデバイスまでの直接パスが1本しかない場合だけです。

図 6-14 は、ポート割り当てクラスで命名されたデバイスの構成です。

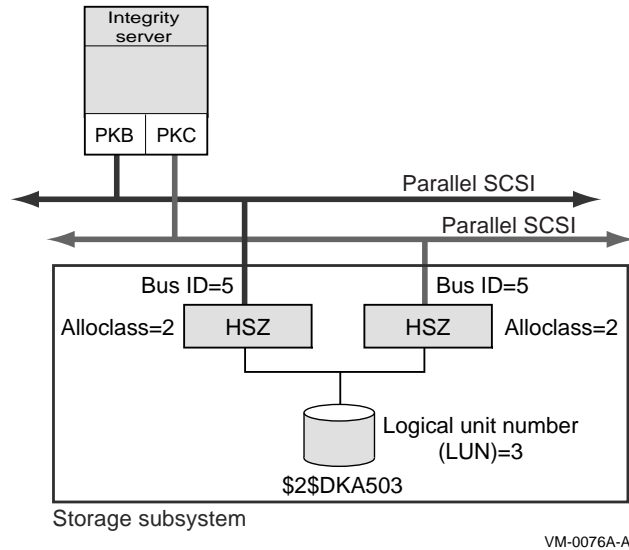
図 6-14 ポート割り当てクラスを使用したデバイス名



6.5.3 HSZ 割り当てクラスを使用したデバイス名

同じストレージ・デバイスに複数のバスが接続されているノードがある場合、図 6-15 に示す新しい HSZ 割り当てクラスを使用します。

図 6-15 HSZ 割り当てクラスを使用したデバイス名



HSZ 割り当てクラスは、割り当てクラス HSC、HSD、HSJ に似ています。HSZ 割り当てクラス番号を使用したデバイス名の形式は、以下のとおりです。

```
$HSZ-allocation-class$ddcu
```

形式の意味は次のとおりです。

- HSZ-allocation-class は、システム管理者によって特定の HSZ ストレージ・コントローラに割り当てられる 1 から 999 までの十進値である。
- dd はデバイス・クラスを表します。ディスクの場合、DK である。
- c は、コントローラを表します。HSZ 割り当てクラスを使用する場合 A である。
- u は、デバイス・ユニット番号を表します。デバイスの SCSI バス ID と論理ユニット番号 (LUN) で決まる。

システム管理者は、構成に応じて以下のコマンドのどちらかを使用して HSZ コントロールから HSZ 割り当てクラスを設定します。

```
HSZ> SET THIS_CONTROLLER ALLOCATION_CLASS = n
```

または

```
HSZ> SET OTHER_CONTROLLER ALLOCATION_CLASS = n
```

ここで n は 1 から 999 の値です。

デュアル冗長構成の 1 つのコントローラ・モジュールに割り当てクラスが設定されると、他のコントローラでも自動的に同じ値に設定されます。

SCSI と Fibre Channel ストレージに対するマルチパスの構成

6.5 パラレル SCSI マルチパス構成のディスク・デバイス命名

以下の例では、割り当てクラスは 199 に設定されます。この例では、この値が両方のコントローラに設定される様子を表しています。

```
z70_B => SET THIS ALLOCATION_CLASS=199
z70_B => SHOW THIS_CONTROLLER
Controller:
    HSZ70 ZG64100136 firmware XB32-0, Hardware CX25
    Configured for MULTIBUS_FAILOVER with ZG64100160
    In dual-redundant configuration
    Device Port SCSI address 6
    Time: NOT SET
Host port:
    SCSI target(s) (0, 2, 3, 4, 5, 6)

    TRANSFER_RATE_REQUESTED = 20MHZ
    Host Functionality Mode = A
    Allocation class          199
    Command Console LUN is target 0, lun 1
Cache:
    32 megabyte write cache, version 4
    Cache is GOOD
    Battery is GOOD
    No unflushed data in cache
    CACHE_FLUSH_TIMER = DEFAULT (10 seconds)
    NOCACHE_UPS
z70_B => SHOW OTHER_CONTROLLER
Controller:
    HSZ70 ZG64100160 firmware XB32-0, Hardware CX25
    Configured for MULTIBUS_FAILOVER with ZG64100136
    In dual-redundant configuration
    Device Port SCSI address 7
    Time: NOT SET
Host port:
    SCSI target(s) (0, 2, 3, 4, 5, 6)

    TRANSFER_RATE_REQUESTED = 20MHZ
    Host Functionality Mode = A
    Allocation class          199
    Command Console LUN is target 0, lun 1
Cache:
    32 megabyte write cache, version 4
    Cache is GOOD
    Battery is GOOD
    No unflushed data in cache
    CACHE_FLUSH_TIMER = DEFAULT (10 seconds)
    NOCACHE_UPS
```

以下の規則は、SCSI デバイス名に HSZ 割り当てクラスを使用する場合に適用します。

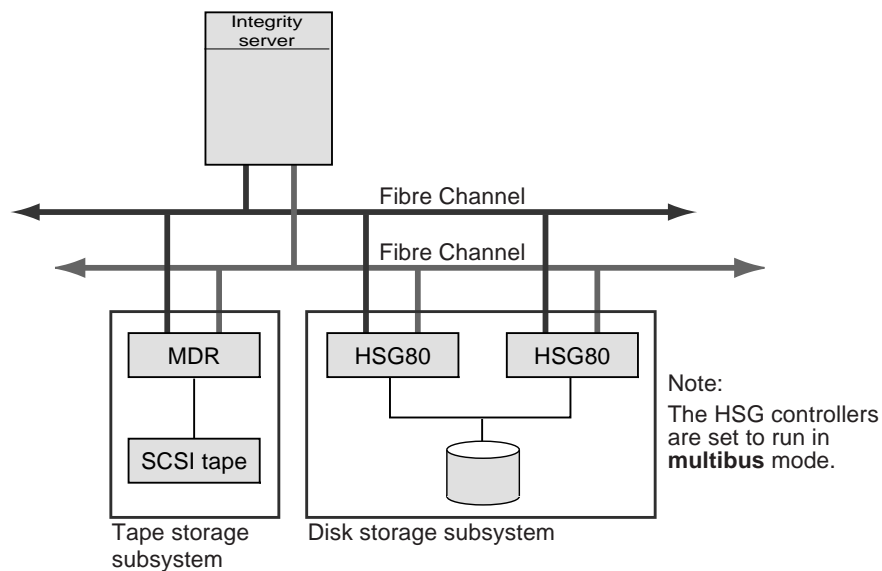
1. マルチパス・モードでは、デバイス名に HSZ 割り当てクラスを使用します (そうしないと、デバイスが構成されません)。
2. 透過モードでは、デバイス名に HSZ 割り当てクラスを使用できますが、必須ではありません。

3. HSZ の両方のコントローラに同じ HSZ 割り当てクラス番号を設定します。これは HSZ ファームウェアが自動的に処理します。
4. クラスタ内では、すべてのタイプの割り当てクラスで HSZ 割り当てクラス番号を固有にします。
5. HSZ 割り当てクラスを持つデバイスを参照するときは、HSZ 割り当てクラスを指定します。たとえば、DKA500 と NODE10\$DKA500 という名前は使用できません。また、\$GETDVI システム・サービスは、このようなデバイスに対しては、HSZ 割り当てクラスを組み入れた完全な名前だけを戻します。

6.6 Fibre Channel マルチパス構成

図 6-16 は、テープ・ストレージ・サブシステムとディスク・ストレージ・サブシステムの両方で構成されるマルチパス構成を示したものです。この図では、ディスク・ストレージ・コントローラがマルチパス・モードで構成されています。

図 6-16 デュアル・ポート・ストレージ・コントローラが 2 つとデュアル・ポート MDR が 1 つ、およびバスが 2 本のシングル・ホスト



VM-0090A-AI

この構成では、次の点に着目してください。

- ホストにアダプタが 2 つあり、それぞれが別々のバスに接続されている。
- 各 HSGx または HSVx のストレージ・コントローラ上の各ポートが別々のインターコネクタに接続されている。

SCSI と Fibre Channel ストレージに対するマルチパスの構成

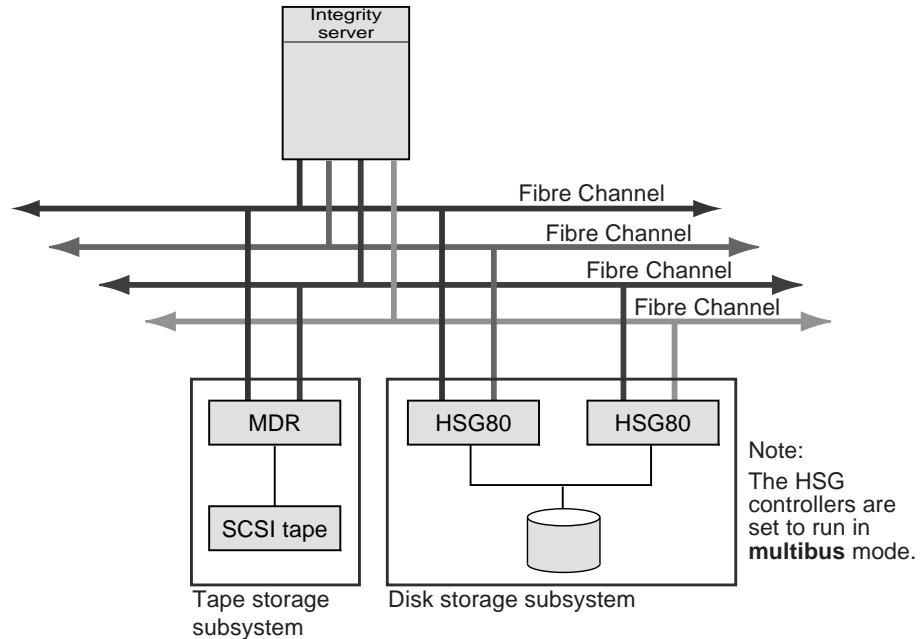
6.6 Fibre Channel マルチパス構成

- Modular Data Router (MDR) または Network Storage Router (NSR) 上の各ポートが別々のインターコネクタに接続されている。
- 両方のストレージ・コントローラが同じディスクにアクセスできる。
- ホストから同じ論理ユニットまで 4 本のパスがある。

各 HSG80 ポートに、専用の Fibre Channel アドレスと Fibre Channel ポート WWID がある点に着目してください。これは、すべてのポートが同じ SCSI アドレスに回答するために異なる SCSI バスに接続されているマルチバス・モードの HSZ80 と異なる点です。独立した FC アドレスにより、デュアル HSG80 の両方のポートを同じ FC で使用できます。

図 6-17 は、図 6-16 と似ていますが、この図では、2 本の Fibre Channel インターコネクトが追加されています。

図 6-17 デュアル・ポート・ストレージ・コントローラが2つとデュアル・ポート MDR が1つ、およびバスが4本のシングル・ホスト



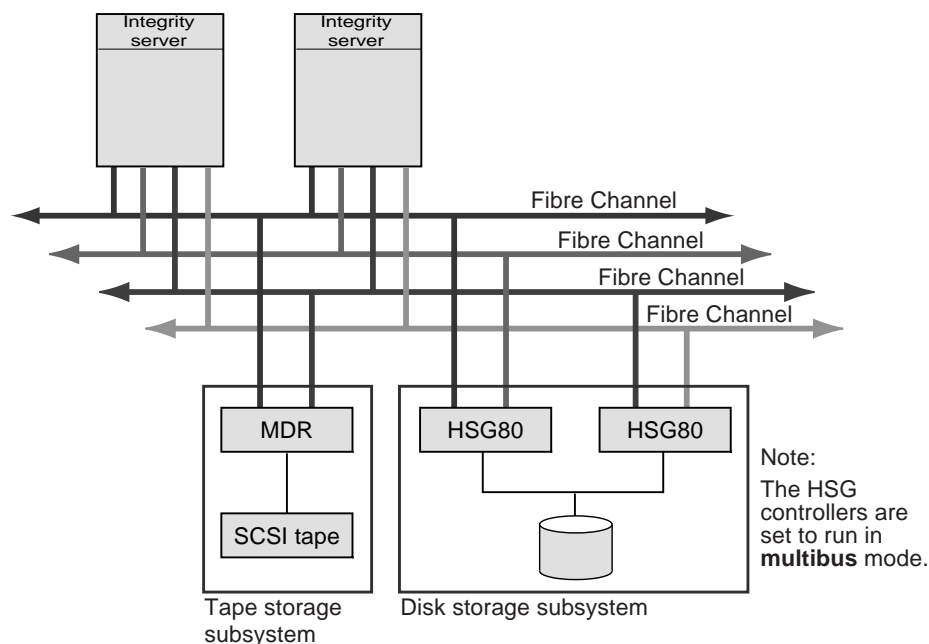
VM-0091A-AI

この構成では、次の点に着目してください。

- ホストにアダプタが4つあり、それぞれが別々のバスに接続されている。
- 各 HSGxまたは HSVxストレージ・コントローラ上の各ポートが別々のインターコネクトに接続されている。
- Modular Data Router (MDR) または Network Storage Router (NSR) 上の各ポートが別々のインターコネクトに接続されている。
- ホストから同じ論理ユニットまで4本のバスがある。

図 6-18 は、前述の 2 つの図を基に作成されています。この図では、シングル・ホストではなく、2 つのホストを使用しています。

図 6-18 デュアル・ポート・ストレージ・コントローラ 2 つと、デュアル・ポート MDR 1 つと、バスが 4 本の 2 つのホスト



VM-0911A-AI

この構成では、次の点に着目してください。

- 各ホストにアダプタが 4 つあり、それぞれが別々のバスに接続されている。
- 各 HSGx または HSVx ストレージ・コントローラ上の各ポートが別々のインターコネクタに接続されている。
- Modular Data Router (MDR) または Network Storage Router (NSR) 上の各ポートが別々のインターコネクタに接続されている。
- 各ホストからディスク・ストレージ・サブシステムの同じ論理ユニットまで 4 本のバスと、テープ・ストレージ・サブシステムまで 2 本のバスがある。

6.7 マルチパス構成の実装

パラレル SCSI, SAS および Fibre Channel インターコネクタは、マルチパス構成をサポートします。これらの構成の実装は似ており、バスを指定するシステム・パラメータとコマンドは同じです。パス識別子の構文は異なります。

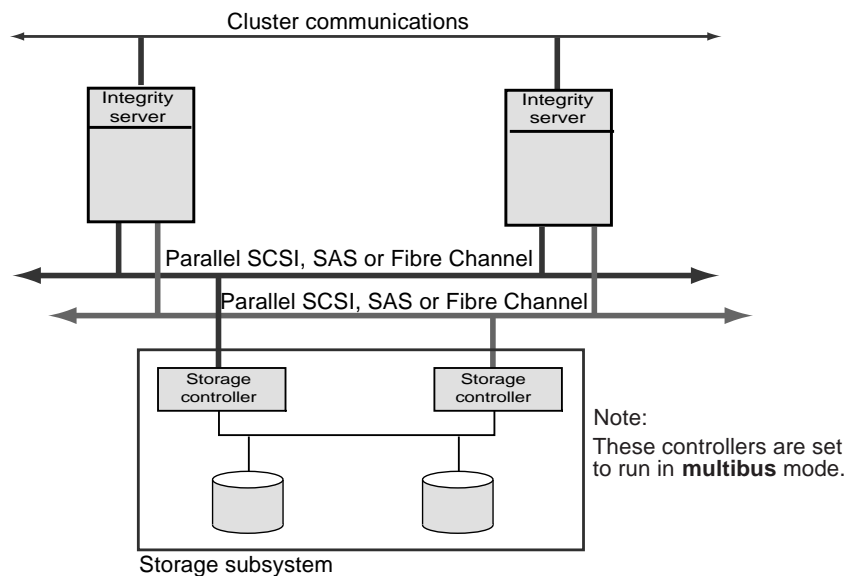
デバイスまでの複数のパスの実装は、次の手順で行います。

1. マルチパス・サポートが必要なデバイスまで、複数の物理パスがあるシステムの構成。
2. HSxコントローラをマルチパス・モードに設定 (ディスクのみ)。
3. オプションで、構成に応じたマルチパス・システム・パラメータとコンソール・パラメータの設定によるマルチパス・サポートの指定。
4. DCL コマンド SET DEVICE/qualifier/PATH=path-identifierによるマルチパス機能の操作の調整。

6.7.1 有効なマルチパス構成

図 6-19 は、有効なマルチパス、マルチホスト構成です。

図 6-19 共有バスと共有ストレージ・コントローラを持つ 2 つのホスト



VM-0080A-AI

この構成では、次の点に着目してください。

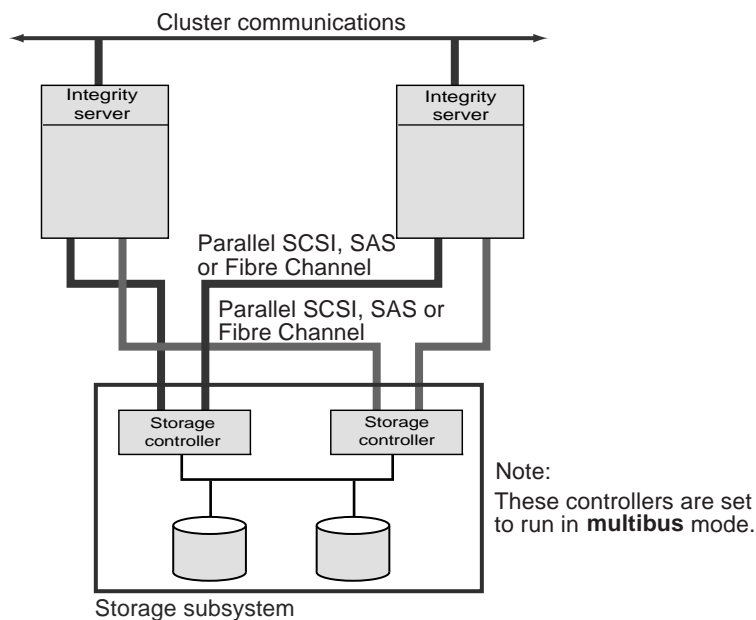
- 両方のホストにアダプタが 2 つある。
- 両方のホストが同じ 2 本のバスに接続されている。
- 両方のホストがストレージを共有している。
- 両方のストレージ・コントローラが 1 本のバスで接続されている。
- 2 つのストレージ・コントローラが同じディスクに接続されている。

SCSI と Fibre Channel ストレージに対するマルチパスの構成 6.7 マルチパス構成の実装

この構成では、各ホストに 2 本の直接バスがあり、各デバイスには 1 本の MSCP サービス対象のバスが接続されています。

図 6-20 は、同じバスで構成されていないシステムの有効なマルチパス構成です。

図 6-20 共有、マルチポート・ストレージ・コントローラを持つ 2 つのホスト



VM-0089A-AI

この構成では、次の点に着目してください。

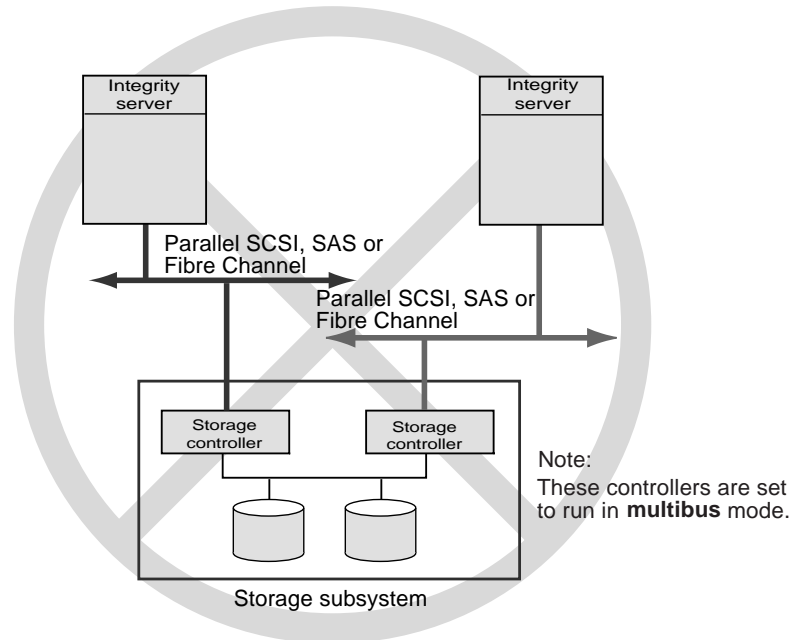
- 各ホストにアダプタが 2 つある。
- 各ホストが 2 本のバスに接続されているが、ホストはバスを共有していない。
- 両方のホストがストレージを共有している。
- 各ストレージ・コントローラに接続が 2 本あり、それぞれがホストに接続されている。
- 2 つのストレージ・コントローラが同じディスクに接続されている。

この構成では、各ホストに直接バスが 2 本あり、1 本はストレージ・コントローラに接続され、もう 1 本は MSCP サービス対象のバスが各デバイスに接続されている。

6.7.2 無効なマルチパス構成

図 6-21 は、無効なマルチパス構成を示しています。クラスタ上の複数のホストを HSZ または HSG に接続するには、同じコントローラ・モジュール (表 6-1 参照) に接続する必要がありますが、この構成では、各ホストが別々のコントローラ・モジュールに接続されています。無効になるのはこのためです。

図 6-21 無効なマルチパス構成



VM-0079A-AI

6.7.3 マルチパス・システム・パラメータ

マルチパス・サポートは表 6-3 に示されるシステム・パラメータで制御されます。(特定のマルチパス・システム・パラメータは、オペレーティング・システム用です。)

表 6-3 マルチパス・システム・パラメータ

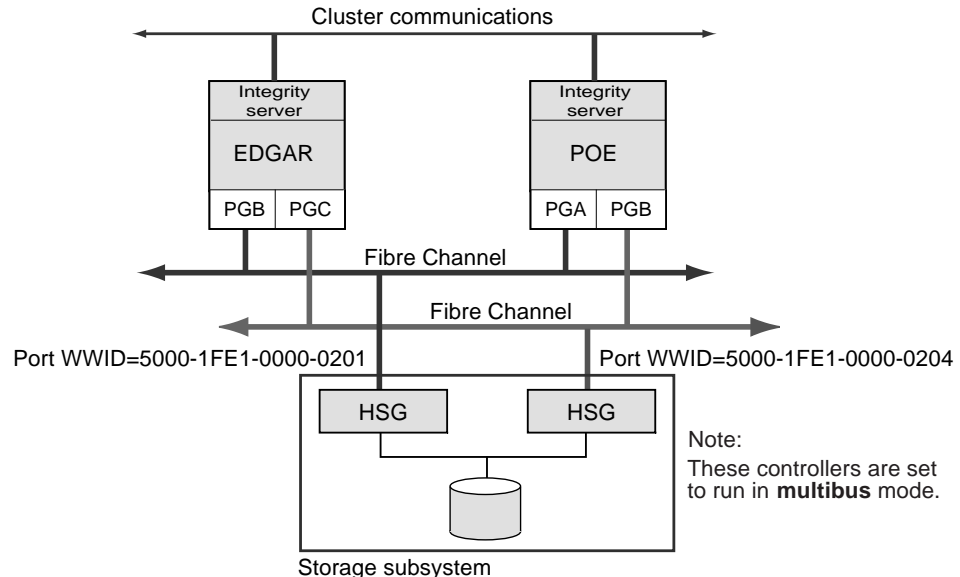
パラメータ	説明
MPDEV_ENABLE	ON (1)のとき、マルチパス・セットを構成します。OFF (0)の場合、新たなマルチパス・セットおよび既存のマルチパス・セットへの新たなパスは無効です。ただし、既存のマルチパス・セットは有効です。デフォルトはON です。 MPDEV_ENABLE が OFF の場合、MPDEV_REMOTE と MPDEV_AFB_INTVL は無効です。
MPDEV_LCRETRIES	論理ユニットがオンラインのコントローラまでの直接パスをシステムが再試行する回数を制御します。この回数を超えると、別のコントローラか、MSCP サービス対象のパスに直接パスを移動します (MSCP パスはディスクにのみ適用されます)。有効範囲は、1 から 256 です。デフォルトは 1 です。
MPDEV_POLLER	ON (1)のとき、マルチパス・セット・メンバまでのパスに対するポーリングが有効になります。ポーリングにより非アクティブ・パス上のエラーを早めに検出することができます。パスが利用できなくなったり、パス・サービスが復活すると、システム管理者には OPCOM メッセージが通知されます。OFF (0)の場合、マルチパス・ポーリングが無効になります。デフォルトは ON です。自動フェールバック機能を使用する場合は、このパラメータを ON に設定する必要があります。
MPDEV_REMOTE (ディスクのみ)	ON (1)のとき、MSCP サービス対象のパスがマルチパス・セットのメンバになります。OFF (0)のとき、SCSI または Fibre Channel デバイスまでのローカル・パスだけで新たなマルチパス・セットが作成されます。MPDEV_REMOTE は、デフォルトで有効になっています。このパラメータを OFF に設定しても、リモートパスがある既存のマルチパス・セットには影響はありません。 サービス対象パスへのマルチパス・フェールオーバーを使用するには、共用 SCSI または Fibre Channel デバイスへ直接アクセスしているすべてのシステムで MPDEV_REMOTE を有効にする必要があります。この機能は、OpenVMS Alpha バージョン 7.3-1 から導入されました。このため、MPDEV_REMOTE を有効にしたすべてのノードで、OpenVMS Alpha バージョン 7.3-1 およびそれ以降を実行する必要があります。MPDEV_ENABLE を OFF (0) に設定すると、マルチパス・セットへの新しいすべてのパスが無効になるため、MPDEV_REMOTE には影響ありません。デフォルトは ON です。
MPDEV_AFB_INTVL (ディスクのみ)	自動フェールバック間隔を秒単位で指定します。自動フェールバック間隔は、同じデバイスで MSCP パスから直接パスまで次のフェールバックを実行するまでの最小経過時間を秒単位で表したものです。 自動フェールバックを有効にするには、MPDEV_POLLER を ON に設定する必要があります。MPDEV_AFB_INTVL を 0 に設定すると、MPDEV_POLLER を無効にしなくても、自動フェールバックを無効にすることができます。デフォルトは 300 秒です。
MPDEV_D1	オペレーティング・システム用です。
MPDEV_D2	オペレーティング・システム用です。
MPDEV_D3	オペレーティング・システム用です。
MPDEV_D4	オペレーティング・システム用です。

6.7.4 パス識別子

以下の項で説明するシステム管理コマンドでは、マルチパス・フェールオーバーの操作を監視、制御できます。これらのコマンドでは、マルチパス・セット内の各パスを別々のものとして指定するパス識別子を渡します。

直接 Fibre Channel パスはローカル・アダプタ名とリモート Fibre Channel ポート WWID、つまり、イニシエータとターゲットで識別します。たとえば、図 6-22 では、左側のホスト・アダプタから左側の HSG ストレージ・コントローラまでのパスのパス識別子は PGB0.5000-1FE1-0000-0201 です。(各 HSG の 2 番目のシリアル・ポートは便宜上省略しています。) ストレージ・コントローラの WWID はコンソールから取得します。

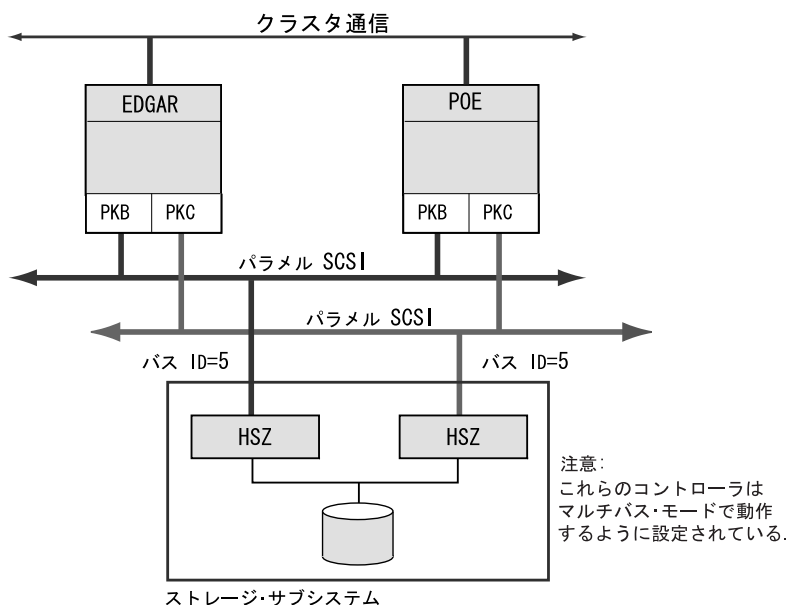
図 6-22 Fibre Channel パスの命名



直接パラレル SCSI パスは、ローカル・ホスト・アダプタ名とリモート SCSI バス ID、つまり、イニシエータとターゲットで識別します。たとえば、図 6-23 では、ディスクまでのノード Edgar の 2 本の直接パスの名前は PKB0.5 と PKC0.5 です。

MSCP サービス対象のパスのパス識別子はMSCPです。

図 6-23 複数の直接バスによる構成



VM-0077A-A1

6.7.5 パスの表示

マルチパス・サポートが有効な場合、以下のどちらかの SHOW DEVICE DCL コマンドにより、デバイスまで複数のパスを表示できます。

```
SHOW DEVICE/FULL device-name
```

```
SHOW DEVICE/MULTIPATH_SET device-name
```

SHOW DEVICE/FULL device-name コマンドでは、デバイスに関する一般的な情報を最初に表示し、次に、デバイスまでのすべてのパスをパス識別子で表示します (第 6.7.4 項参照)。

SHOW DEVICE/MULTIPATH_SET device-name コマンドでは、複数のパスを持つデバイスに関する簡単なマルチパス情報だけを表示します。

マルチパス情報は、マルチパス・デバイスに直結されたノード上にものみ表示されます。

6.7.5.1 SHOW DEVICE/FULL によるパスの表示

以下の例は、SHOW DEVICE/FULL device-name コマンドの出力です。表示の先頭 (device has multiple I/O paths) で複数のパスの使用が明示されており、複数パスの記述が表示の最後まで、I/O paths to device の下まで続いています。また、表示の先頭にある Error count と Operations completed の値は、各パスのカウンターの合計となっています。

\$ SHOW DEVICE/FULL \$1\$DGA23:

Disk \$1\$DGA23: (WILD8), device type HSG80, is online, mounted, file-oriented device, shareable, device has multiple I/O paths, served to cluster via MSCP Server, error logging is enabled.

Error count	3	Operations completed	32814199
Owner process	"	Owner UIC	[SYSTEM]
Owner process ID	00000000	Dev Prot	S:RWPL,O:RWPL,G:R,W
Reference count	9	Default buffer size	512
WWID	01000010:6000-1FE1-0000-0D10-0009-8090-0677-0034		
Total blocks	17769177	Sectors per track	169
Total cylinders	5258	Tracks per cylinder	20
Host name	"WILD8"	Host type, avail	HP AlphaServer GS160 6/731, yes
Alternate host name	"W8GLX1"	Alt. type, avail	HP AlphaServer GS160 6/731, yes
Allocation class	1		
Volume label	"S5SH_V72_SSS"	Relative volume number	0
Cluster size	18	Transaction count	8
Free blocks	12812004	Maximum files allowed	467609
Extend quantity	5	Mount count	8
Mount status	System	Cache name	"_\$1\$DGA8:XQPCACHE"
Extent cache size	64	Maximum blocks in extent cache	1281200
File ID cache size	64	Blocks currently in extent cache	0
Quota cache size	0	Maximum buffers in FCP cache	1594
Volume owner UIC	[1,1]	Vol Prot	S:RWCD,O:RWCD,G:RWCD,W:RWCD

Volume Status: ODS-2, subject to mount verification, file high-water marking, write-back XQP caching enabled, write-through XFC caching enabled.

Volume is also mounted on H2OFRD, FIBRE3, NORLMN, SISK0, BOOLA, FLAM10, W8GLX1.

I/O paths to device	5		
Path PGA0.5000-1FE1-0000-0D12	(WILD8), primary path.		
Error count	2	Operations completed	130666
Path PGA0.5000-1FE1-0000-0D13	(WILD8), current path.		
Error count	1	Operations completed	30879310
Path PGA0.5000-1FE1-0000-0D11	(WILD8).		
Error count	0	Operations completed	130521
Path PGA0.5000-1FE1-0000-0D14	(WILD8).		
Error count	0	Operations completed	130539
Path MSCP (W8GLX1).			
Error count	0	Operations completed	1543163

マルチパス・デバイスのパスごとに、パス識別子、そのパスに関連付けられたホスト名、パス状態、エラー・カウント、操作カウントが表示されます。

このコード例で複数のパスの部分に表示される用語については、表 6-4 を参照してください。

表 6-4 SHOW DEVICE/FULL マルチパス用語

用語	説明
WWID	SCSI 論理ユニットのワールドワイド ID。
Host name (ホスト名)	現在のパスで使用されているシステムの名前。ホスト名は、マルチパス・デバイスへの MSCP パスが存在する場合に表示されます。
Alternate host name (代替ホスト名)	デバイスへアクセスすることができる別のシステムの名前。現在のパスが直接パスの場合は、MSCP パスに現在関連付けられているホストになります。現在のパスが MSCP パスの場合は、ローカル・システムの名前になります。代替ホスト名は、マルチパス・ディスク・デバイスへの MSCP パスが存在する場合に表示されます。
Primary path (プライマリ・パス)	オペレーティング・システムが最初に検出するパスです。
Current path (現在のパス)	I/O に現在使用されているパスです。
User disabled (ユーザ無効)	DCL コマンド SET DEVICE/NOENABLE が、このパスに対して実行済みです。
Polling disabled (ポーリング無効)	DCL コマンド SET DEVICE/NO POLL が、このパスに対して実行済みです。
Not responding (応答なし)	デバイスへのこのパスが、最後にチェックされた時点で使用できなくなっています。通常、マルチパス・ポーリングは、パスが良好な場合 60 秒ごとにチェックし、パスに異常がある場合 30 秒ごとにチェックします。
Unavailable (使用不可)	ソフトウェア・コネクタがこのパスに関連付けられていないためにこのパスは使用できません。

6.7.5.2 SHOW DEVICE/MULTIPATH_SET によるパスの表示

指定デバイスの複数のパスの簡単なリストは、割り当てクラスのすべてのデバイスまたは、すべてのデバイスについて、DCL コマンドで表示できます。

```
SHOW DEVICE/MULTIPATH_SET [device-name]
```

デバイス名は任意です。省略するとマルチパス・セットを構成するすべてのデバイスが表示されます。検出されたマルチパス・デバイスごとに、デバイス名、ホスト名、デバイス状態、エラー・カウント、アクセスできるパス数、パスの総数、現在のパスのパス識別子が表示されます。

アクセスできるパス数は、次の 2 つの理由から、パスの総数よりも少なくなる場合があります。

- システム管理者がコマンド SET DEVICE/PATH=pathname/NOENABLE を使用してパスを無効にした。
- パスが Not Responding (応答なし) として示されている場合、オペレーティング・システムによってパスの総数が減らされる。(この動作は、OpenVMS Alpha バージョン 7.3-1 から導入)

表示されるホスト名は、現在のパスのホスト名です。直接パスの場合、ローカル・システムのホスト名です。MSCP サービス対象のパスの場合、デバイスのアクセス・サービスを担当するリモート・システムのホスト名です。

以下の例は、SHOW DEVICE/MULTIPATH コマンドの出力です。

```
$ SHOW DEVICE/MULTIPATH
Device          Device          Error Paths Current
Name           Status          Count Avl/Tot  path
$1$DGA8:       (H2OFRD) Mounted         3  5/ 5  PGA0.5000-1FE1-0000-0D12
$1$DGA10:      (H2OFRD) ShadowSetMember 1  5/ 5  PGA0.5000-1FE1-0000-0D14
$1$DGA11:      (WILD8)  ShadowSetMember 3  3/ 3  MSCP
$1$DGA23:      (H2OFRD) Mounted         6  5/ 5  PGA0.5000-1FE1-0000-0D13
$1$DGA30:      (H2OFRD) ShadowSetMember 8  5/ 5  PGA0.5000-1FE1-0000-0D13
$1$DGA31:      (WILD8)  ShadowMergeMbr  5  3/ 3  MSCP
$1$DGA33:      (H2OFRD) Online          0  5/ 5  PGA0.5000-1FE1-0000-0D12
$1$DGA40:      (H2OFRD) Mounted         2  5/ 5  PGA0.5000-1FE1-0000-0D13
$1$DGA41:      (H2OFRD) ShadowMergeMbr 8  5/ 5  PGA0.5000-1FE1-0000-0D12
$70$DKA100:    (H2OFRD) Mounted         0  3/ 3  PKD0.1
$70$DKA104:    (H2OFRD) ShadowSetMember 0  3/ 3  PKD0.1
$70$DKA200:    (H2OFRD) ShadowSetMember 0  3/ 3  PKD0.2
$70$DKA300:    (H2OFRD) ShadowSetMember 0  3/ 3  PKC0.3
$80$DKA1104:   (H2OFRD) ShadowSetMember 0  3/ 3  PKD0.11
$80$DKA1200:   (H2OFRD) ShadowSetMember 0  3/ 3  PKD0.12
$80$DKA1204:   (H2OFRD) ShadowSetMember 0  3/ 3  PKC0.12
$80$DKA1207:   (H2OFRD) Mounted         0  3/ 3  PKD0.12
$80$DKA1300:   (H2OFRD) Mounted         0  3/ 3  PKD0.13
$80$DKA1307:   (H2OFRD) ShadowSetMember 0  3/ 3  PKD0.13
$80$DKA1500:   (H2OFRD) Mounted         0  3/ 3  PKD0.15
$80$DKA1502:   (H2OFRD) ShadowSetMember 0  3/ 3  PKD0.15
```

\$70\$DKA などデバイス名の一部を指定すると、デバイスが\$70\$DKA で始める複数のパスとともに表示されます。

6.7.6 パス・ポーリング

SCSI マルチパス・サポートが有効な場合、各ホスト・アダプタから HSZ コントローラまたは HSG コントローラあるいは MDR までのすべての I/O パスを定期的にポーリングして各 I/O パスの状態を判定します。システムがパスの変化を検出すると、以下のようなメッセージをコンソールとオペレータのログに出力します。

```
All multipath devices on path PKB0.5 are either disabled or not reachable.
```

または

```
At least one multipath device on path PKB0.5 is enabled and reachable.
```

パス上のすべてのデバイスを削除すると、パス障害が報告されます。このとき、ホストから HSx コントローラまでのパスは有効な場合もありますが、ポーリングするデバイスがないと有効かどうかはわかりません。

ポーリングは、以下のコマンドでオンまたはオフします。

```
SET DEVICE device/[NO] POLL/PATH=path-identifier
```

長期間サービス対象外にするパスでは、ポーリングをオフにすればシステム・オーバーヘッドを節約できて便利です。

6.7.7 現在のパスへの手動切り替え

デバイスの現在のパスは、SET DEVICE コマンドに/SWITCH 修飾子を指定して手動で切り替えることができます。パスの切り替えは、一般には、複数の HSx コントローラ・モジュール、MDR、およびバス全体の I/O 負荷をバランスさせるときに利用します。

現在のパスの切り替えのコマンド構文は次のとおりです。

```
SET DEVICE device-name/SWITCH/PATH=path-identifier
```

このコマンドには、OPER 特権が必要です。また、デバイスが現在別のプロセスで割り当てられている場合は (特に多いのはテープ・デバイスの場合)、SHARE 特権が必要です。

以下のコマンドでは、デバイス \$2\$DKA502 から MSCP サービス対象のパスまでのパスを切り替えます。

```
$ SET DEVICE $2$DKA502/SWITCH/PATH=MSCP
```

このコマンドを実行するとパスの切り替えが開始し、ただちに DCL プロンプトに戻ります。DCL プロンプトが再表示されて、パスの切り替えの終了するまでに遅延が発生することがあります。

マウントされたデバイスのパスの手動切り替えは、パス切り替えコマンドによって起動されるマウント検証中に行われます。このとき、例 6-1 のような、通常のマウント検証メッセージと、パス切り替えメッセージが表示されます。

例 6-1 パスの手動切り替え時のメッセージ

```
%%%%%%%%%% OPCOM 15-JUN-2001 09:04:23.05 %%%%%%%%%%%  
Device $1$DGA23: (H2OFRD PGA) is offline.  
Mount verification is in progress.  
  
%%%%%%%%%% OPCOM 15-JUN-2001 09:04:25.76 %%%%%%%%%%%  
09:04:25.76 Multipath access to device $1$DGA23: has been manually switched  
from path PGA0.5000-1FE1-0000-0D11 to path PGA0.5000-1FE1-0000-0D14  
  
%%%%%%%%%% OPCOM 15-JUN-2001 09:04:25.79 %%%%%%%%%%%  
Mount verification has completed for device $1$DGA23: (H2OFRD PGA)
```

パス切り替えの終了は、SHOW DEVICE/FULL コマンドか SHOW DEVICE /MULTIPATH コマンドを実行すれば確認できます。

切り替え操作時に、パスの手動切り替えで指定したパスに障害が発生すると、自動パス切り替えが実行されます。この場合、コマンドで指定したパスとは別のパスに切り替えられる場合があります。

パスの手動切り替えで論理ユニットを 1 つの HSG80 コントローラから別のコントローラに切り替えると、クラスタの別のノードにコマンドが影響する場合があります。そのようなノードでは、それぞれの現在のパスにおけるマウント検証が実行され、他の HSG80 コントローラ上のパスへの自動切り替えが発生します。例 6-2 は、このイベントを表すメッセージです。

例 6-2 他のノードがパス切り替えを検出したときに表示されるメッセージ

```
%%%%%%%% OPCOM 15-JUN-2001 09:04:26.48 %%%%%%%%%%
Device $1$DGA23: (WILD8 PGA, H20FRD) is offline.
Mount verification is in progress.

%%%%%%%% OPCOM 15-JUN-2001 09:04:26.91 %%%%%%%%%%
09:04:29.91 Multipath access to device $1$DGA23: has been auto switched from
path PGA0.5000-1FE1-0000-0D12 (WILD8) to path PGA0.5000-1FE1-0000-0D13 (WILD8)

%%%%%%%% OPCOM 15-JUN-2001 09:04:27.12 %%%%%%%%%%
Mount verification has completed for device $1$DGA23: (WILD8 PGA, H20FRD)
```

ノード WILD8 では、各パスが直接パスであるため、WILD8 ノード名はパスごとに表示されます。マウント検証メッセージ in progress と completed のノード名フィールドには、ローカル・パスと代替の MSCP が両方表示されます。この例の WILD8 PGA, H20FRD という名前は、WILD8 でローカルの PGA パスが使用されていることと、ノード H20FRD を経由する MSCP パスが代替パスであることを示しています。

6.7.8 OpenVMS によるパス選択

マルチパス・デバイスへの現在のパスの選択方法は、デバイス・タイプと、パス選択の要因となったイベントによって決まります。

システム起動時の初期構成のためのパス選択

マルチパス・ディスク (DG, DK) またはテープ・デバイス (MG) への新しいパスが構成されると、現在のパスとして、最もデバイス数の少ない直接パスが自動的に選択されます。このときには、オペレータ・メッセージは表示されません。(このタイプのパス選択は、OpenVMS Alpha バージョン 7.3-1 から導入されました。) DG, DK, または MG デバイスでは、システム・ブート後初めてデバイスが使用されるまで、またはパスの手動切り替えが SET DEVICE/SWITCH コマンドで実行されるまで、このタイプのパス選択が行われます。

汎用のマルチパス SCSI デバイス (GG, GK) への新しいパスが構成されると、現在のパスとして、最初に検出されたパス (プライマリ・パス) が自動的に選択されます。GG および GK デバイスの場合は、新しいパスが構成されても、プライマリ・パスが現在のパスとしてそのまま使用されます。通常、GG および GK デバイスは、HSG コントローラ LUN または HSV コントローラ LUN または テープ・メディア・ロボットのコンソール LUN になります。

ディスク・デバイスのマウント時のパス選択

MOUNT コマンドを実行すると、マルチパス・ディスク・デバイスへの現在のパスが変わる場合があります。MOUNT コマンドで実行された I/O により、ディスク・デバイスを別の HSx コントローラにフェールオーバーする必要がある直接パスの検索が実行されます。

DG または DK ディスク・デバイスでの MOUNT コマンドによるパス選択は、次のように実行されます。

1. 現在のパスが直接パスで、このパスのデバイスへのアクセスにコントローラ・フェールオーバーが必要ない場合は、現在のパスが使用されます。
2. 現在のパスが MSCP パスで、そのパスがパスの手動切り替えコマンドで選択された場合は、現在のパスが使用されます。
3. 最もデバイス数の少ない現在のパスを開始地点として、すべての直接パスがチェックされます。直接パスは、他のデバイスに対し、現在のパスとして使用率の高い順に考慮されます。デバイスへのアクセスにコントローラ・フェールオーバーが必要のないパスが見つかり、そのパスが選択されます。選択されたパスが現在のパスでない場合は、パスの自動切り替えが実行され、OPCOM メッセージが発行されます。
4. 最もデバイス数の少ない現在のパスを開始地点として、すべての直接パスが試行されます。直接パスは、他のデバイスに対し、現在のパスとして使用率が高い順に考慮されます。必要に応じて、選択したパスで HSx コントローラへのデバイスのフェールオーバーが試行されます。選択されたパスが現在のパスでない場合は、パスの自動切り替えが実行され、OPCOM メッセージが発行されます。
5. MSCP サービス対象パスが試行されます。MSCP パスが現在のパスでない場合は、パスの自動切り替えが実行され、OPCOM メッセージが発行されます。

MOUNT コマンドでは、作業パスが見つかるまで、このパス選択アルゴリズムが繰り返し行われます。正確な再試行回数は、前の試行での経過時間と、MOUNT コマンドで指定した修飾子の両方で決まります。

OpenVMS Alpha バージョン 7.3-1 から導入されたこのパス選択プロセスには、次の利点があります。

- クラスタ内の他のホストへ影響を最小限に抑えられる。
- クラスタ内の他のノードで手動で設定した静的負荷調整を維持できる。

- HSxコンソール・コマンドを使用して、2つのHSxコントローラ間におけるデバイスの初期のデフォルト分散を設定できる。
- このホストからディスク・デバイスまでの利用可能なパスの使用量を調整できる。
- MSCP サービス対象パスよりも直接パスが優先される。

この選択プロセスでは、2つのHSxコントロール間でデバイスを分散できます。これを行うには、次のようなHSxコンソール・コマンドを使用します。

```
HSG> SET UNIT PREFERRED_PATH=THIS_CONTROLLER  
HSG> SET UNIT PREFERRED_PATH=OTHER_CONTROLLER
```

また、第6.7.7項で説明したパスの手動切り替えのDCLコマンドを使用して、同じHSxコントローラで別のホスト・パス・アダプタまたは別のポートを選択したり、デバイスを強制的に別のHSxコントローラへフェールオーバーすることができます。

テープ・ドライブ・デバイス・マウント時のパス選択

マルチパス・テープ・ドライブのサポートと、このタイプのパス選択は、OpenVMS Alphaバージョン7.3-1から導入されました。MOUNTコマンド実行時のパス選択は、以下の理由から、マルチパス・テープ・ドライブ・デバイスの場合と、ディスク・デバイスの場合とで若干異なります。

- テープ・ドライブは、複数のホストで同時に共有できない。
- ディスクでは、別のホストが実行しているI/Oに割り込む際に問題が発生しますが、テープ・ドライブではそのような問題は発生しない。
- マルチパス・テープ・デバイスへの直接パスとMSCPサービス対象パスとの間では、フェールオーバーは行われません。

MG テープ・ドライブ・デバイスでのMOUNTコマンドによるパス選択は、次のように実行されます。

1. コントローラ・フェールオーバーが必要な場合でも、可能であれば現在のパスが使用されます。
2. 最もデバイス数の少ない現在のパスを開始地点として、直接パスがチェックされます。直接パスは、他のデバイスに対し、現在のパスとして使用率の高い順に考慮されます。デバイスへのアクセスにコントローラ・フェールオーバーが必要のないパスが見つかり、そのパスが選択されます。選択されたパスが現在のパスでない場合は、パスの自動切り替えが実行され、OPCOMメッセージが発行されません。
3. 最もデバイス数の少ない現在のパスを開始時点として、直接パスが再度チェックされます。直接パスは、他のデバイスに対し、現在のパスとして使用率の高い順に考慮されます。選択されたパスが使用可能で、現在のパスでない場合は、パスの自動切り替えが実行され、OPCOMメッセージが発行されます。

6.7.9 OpenVMS によるマルチパス・フェールオーバーの方法

マウント検証の対象となっているデバイスに対する I/O 操作が失敗し、失敗の状態から再試行が可能とされる場合、マウント検証が呼び出されます。デバイスがマルチパス・デバイス、またはマルチパス・デバイスを含むシャドウ・セットである場合、マウント検証中にそのデバイスへの代替パスが自動的に試行されます。これによって、デバイスがある HSxコントローラまたは MDR から別のものへフェールオーバーした状態から透過的に回復することができ、デバイスへのパスの障害から、透過的に回復できます。

マウント検証の対象となるのは、次のデバイスです。

- ホスト方式のボリューム・シャドウイング・セットを含む、Files-11ボリュームとしてマウントされているディスク・デバイス
- ANSI テープ・ボリュームとして使用されているテープ・デバイス
- フォーリン・ボリュームとしてマウントされているテープ・デバイス

フォーリン・マウント・ディスク・ボリュームと汎用 SCSI デバイス (GG および GK) は、マウント検証の対象とならないため、自動マルチパス・フェールオーバーは実行できません。

マウント検証中のパス選択は以下のように実行されます。

1. 現在のパスが直接パスで、このパスのデバイスへのアクセスにコントローラ・フェールオーバーが必要ない場合は、現在のパスが試行されます。
2. 現在のパスが MSCP パスで、そのパスがパスの手動切り替えコマンドで選択された場合は、現在のパスが試行されます (ディスクのみ)。
3. 最もデバイス数の少ない現在のパスを開始地点として、すべての直接パスがチェックされます。直接パスは、他のデバイスに対し、現在のパスとして使用率の高い順に考慮されます。デバイスへのアクセスにコントローラ・フェールオーバーが必要のないパスが見つかり、そのパスが選択されます。
4. 手順 3 は、MPDEV_LCRETRIES システム・パラメータで指定した回数だけ繰り返されます。これは、HSxまたは MDR コントローラ・フェールオーバーが必要のないパスの選択におけるもう 1 つの特徴です。MPDEV_LCRETRIES のデフォルト値は 1 です。
5. 最もデバイス数の少ない現在のパスを開始時点として、すべての直接パスが試行されます。直接パスは、他のデバイスに対し、現在のパスとして使用率の高い順に考慮されます。必要に応じて、選択したパスで、HSxまたは MDR コントローラへのデバイスのフェールオーバーが試行されます。
6. MSCP サービス対象パスが存在する場合は、そのパスが試行されます (ディスクのみ)。

動作するパスが見つかるまで、またはマウント検証がタイムアウトになるまで手順 1 から 6 を繰り返します。動作するパスが見つかり、そのパスが異なる場合は、現在のパスが自動的にその新しいパスに切り替わり、OPCOM メッセージが発行されます。マウント検証が完了すると、失敗した I/O 操作が再開され、新しい I/O の処理が実行されます。このパス選択プロセスでは、以下のような理由から、HSx コントローラ間で不必要にデバイスのフェールオーバーが行われないような仕組みになっています。

- HSx コントローラ・モジュール間のフェールオーバーが発生すると、同期する必要のあるキャッシュ・データの量にもよるが、約 1 秒から 15 秒の遅延が発生する。
- デバイスまでのアクセスを共用する他のノードが、代替パスにより通信を再び確立する必要がある。

このパス選択プロセスでは、MSCP サービス対象パスよりも直接パスが優先して選択されます。これは、MSCP サービス対象パスを使用すると、サーバ・システムに余計な CPU および I/O オーバヘッドがかかるからです。それに比べ、直接パスでは、MSCP サーバ・システムに余計な CPU または I/O オーバヘッドがかかることはありません。このプロセスで MSCP サービス対象パスが選択されるのは、利用可能な直接パスがどれも稼動していない場合だけです。さらに、このパス選択プロセスでは、不必要なコントローラ・フェールオーバーを行わないという制約に従って、利用可能な直接パスの使用量が調整されます。

6.7.10 直接パスへの自動フェールバック (ディスクのみ)

第 6.7.9 項で説明したマルチパス・フェールオーバーは、MSCP サービス対象パスにも適用されます。つまり、現在のパスが MSCP サービス対象パス経由であるときに、そのサービス対象パスで障害が発生した場合は、マウント検証によって、元の使用状態の直接パスへ自動フェールオーバーが行われます。

ただし、MSCP パスで I/O エラーが発生した場合は、直接パスへフェールバックを行う必要があります。以下の順にイベントが発生したとします。

- デバイスに対して直接パスが使用されている。
- すべての直接パスで障害が発生し、デバイスへの次の I/O が失敗する。
- マウント検証によって、MSCP サービス対象パスへの自動切り換えが行われる。
- しばらくすると、デバイスへの直接パスが回復する。

この場合、デバイスに対して直接パスの方が望ましい場合でも、MSCP サービス対象パスが継続して使用されます。これは、MSCP サービス対象パスでエラーが発生しなければ、パス選択プロセスが実行されないからです。

自動フェールバック機能を使用すると、このような状態を解決することができます。マルチパス・ポーリングにより、次のすべての条件に該当することが確認されると、直接パスへフェールバックされます。

- デバイスへの直接パスが応答すること。
- デバイスへの現在のパスが MSCP サービス対象パスであること。
- 現在のパスが、パスの手動切り替えによって選択されていないこと。
- 対象デバイスで、MPDEV_AFB_INTVL で指定した直近の秒数内に自動フェールバックが行われていないこと

マウント検証によって自動フェールバックが実行されると、自動フェールオーバ・プロセスが対象デバイスで行われます。

マルチパス・ポーリングは、主に、未使用パスの状態をテストして、以下のような状況にならないようにするために使用します。

- システムに、デバイスへのパス A と B が存在する。
- システムではパス A が使用されている。
- パス B が機能しなくなった場合、それは検出されない。
- しばらくして、パス A が単独でブレイクする。
- システムでパス B へのフェールオーバが行われるが、パス B がブレイクしていることが検出される。

ポーリングは、障害が発生してから 1 分以内にパス B の障害を検出し、OPCOM メッセージを発行します。システム管理者がこのメッセージに気がつくことにより、直ちに適切な処置を行うことができます。

デバイスは、パス・ポーリングが発行した SCSI INQUIRY コマンドに正常に応答しても、そのパスでのパスの切り替えやマウント検証に失敗する場合があります。システム管理者またはオペレータは、以下の 3 つの方法で自動フェールバックを制御できます。

1. MPDEV_AFB_INTVL システム・パラメータで、特定のデバイスにおける自動フェールバックの最小試行間隔を指定します。
2. MPDEV_AFB_INTVL の値を 0 に設定して、自動フェールバックを無効にします。0 に設定すると、システムのどのデバイスでも自動フェールバックは行われません。
3. デバイスを MSCP サービス対象パスに手動で切り替えて、特定のデバイスにおける自動フェールバックを一時的に無効にします。これは、現在のパスが MSCP サービス対象パスであっても行うことができます。

マウントされたデバイスに直接パスと MSCP サービス対象パスの両方が存在する場合は、パス選択手順、自動フェールオーバ手順、自動フェールバック機能により、そのデバイスへの現在のパスは、通常、直接パスになります。例外として、パスが MSCP サービス対象パスに手動で切り替えられた場合、または使用状態の直接パスが存在しない場合などは、MSCP サービス対象パスが現在のパスになります。

6.7.11 パス切り替え候補としてのパスの有効化または無効化

デフォルトで、すべてのパスがパス切り替えの候補になります。パスは、SET DEVICE コマンドに/[NO]ENABLE 修飾子を指定して、切り替え候補として無効化または再有効化できます。以下のような場合にこの操作をします。

- 特定のパスに障害があることがわかっている場合や、そのパスにフェールオーバーを実行するとクラスタのメンバ内にアクセスを失うメンバが発生する場合。
- 選択したパスがサービス利用中であるとき、そのパスに対する自動切り替えを避けたい場合。

現在のパスは無効にできませんので注意してください。

パスを有効、無効にするためのコマンド構文は以下のとおりです。

```
SET DEVICE device-name/[NO]ENABLE/PATH=path-identifier
```

以下のコマンドでは、デバイス\$2\$DKA502 の MSCP サービス対象のパスが有効になる。

```
$ SET DEVICE $2$DKA502/ENABLE/PATH=MSCP
```

以下のコマンドでは、デバイス\$2\$DKA502 のローカル・パスが無効になる。

```
$ SET DEVICE $2$DKA502/NOENABLE/PATH=PKC0.5
```

パスを無効にするときは注意が必要です。図 6-21 にあるような無効な構成を作成しないようにしてください。

6.7.12 パフォーマンスについての考慮

MSCP パスが現在のパスでない場合、ディスク・マルチパス・セット内に MSCP サービス対象パスが存在していても、安定した状態の I/O パフォーマンスには大きな影響はありません。

マルチパス・セット内に MSCP サービス対象パスが存在する場合は、特定の異常障害でのマウント検証中に、作業パスの検出に時間がかかる場合があります。直接パスが最初に試行されるため、MSCP パスが存在しても回復時間には影響ありません。

ただし、直接パスから MSCP サービス対象パスへ動的に切り替える機能を使用すると、マルチパス・ディスク・ストレージへの直接パスを使用する特定の MSCP サーバ・システムに対する I/O サービス負荷が大幅に増える可能性があります。サービス対象の I/O は、MSCP サーバのその他のほとんどすべての処理より優先されるため、MSCP サービス対象パスへのフェールオーバーは、サーバのキャパシティとサービス対象の I/O 要求の増加率により、その MSCP サーバ上の他のアプリケーションの応答に影響を与える可能性があります。

たとえば、アプリケーションの作業負荷を処理できるだけの十分な CPU および I/O 処理能力がある OpenVMS Cluster 構成において、すべての共用 SCSI ストレージに直接 SCSI パスでアクセスするとします。このような構成の場合は、障害が発生しても、制限された数のデバイスが MSCP サービス対象パスへ強制的に切り替えられるので、機能することができます。しかし、さらに多くの障害が発生すると、MSCP サービス対象パスに対する負荷がクラスタのキャパシティに近づくため、アプリケーションのパフォーマンスが許容できないレベルに低下します。

システム管理者は、MSCP_BUFFER および MSCP_CREDITS システム・パラメータを使用して、MSCP サービスにリソースを割り当てることができます。MSCP サーバに、すべての受信 I/O 要求を処理するだけのリソースがない場合、この MSCP サーバの MSCP パスに存在するデバイスにアクセスしているシステムのパフォーマンスが低下します。

MONITOR MSCP コマンドを使用すると、MSCP サーバでリソースが不足していないかどうか確認できます。Buffer Wait Rate が 0 以外の場合は、その MSCP サーバが、リソースの待機中に一部の I/O を停止する必要があったことを示しています。

これらのパラメータには、適切な推奨値はありません。ただし、OpenVMS Alpha バージョン 7.2-1 から、MSCP_BUFFER のデフォルト値が 128 から 1024 に増やされました。

オンライン・ヘルプの SYSGEN ユーティリティに関するトピックで説明しているように、MSCP_BUFFER には、MSCP サーバのローカル・バッファ領域に割り当てられるページレットの数を指定します。また、MSCP_CREDITS には、1 つのクライアント・システムからアクティブにできる未処理の I/O 要求の数を指定します。たとえば、多くのディスクがいくつかの OpenVMS システムにサービスしているシステムの場合は、MSCP_BUFFER の値を 4000 以上に設定し、MSCP_CREDITS の値を 128 以上に設定します。

システム・パラメータの変更については、『OpenVMS システム管理者マニュアル』を参照してください。

MSCP サービス対象パスへのフェールオーバーに依存している構成を、MSCP サービス対象パスに対する負荷レベルを最悪な状態にして、テストすることをお勧めします。複数サイト SAN を使用する複数サイトのディザスタ・トレラント・クラスタを構成している場合は、SAN をパーティション化し、MSCP サービス対象パスを強制的に使用する原因となる障害について検討してください。対称的なデュアル・サイト構成では、MSCP サービス対象パスからアクセスする SAN ストレージの 50 % をキャパシティとして指定することをお勧めします。

構成のキャパシティをテストするには、パスの手動切り替えを使用して、MSCP サービス対象パスを強制的に使用します。

6.7.13 コンソールについての考慮

ここでは、パラレル SCSI マルチパス・ディスク・デバイスとコンソールを併用する方法を説明します。FC マルチパス・デバイスとコンソールを併用する方法については、第 7.6 節を参照してください。

コンソールでは、一般的な、パス依存の SCSI デバイス名を使用します。たとえば、ディスクのデバイス名形式では、DK の後に、ホスト・アダプタを表す文字、SCSI ターゲット ID、LUN を続けます。

つまり、マルチパス・デバイスでは、アクセス時に使用するすべてのホスト・アダプタ名からなる複数の名前が使用されます。以下のコンソール出力例では、デバイス・コマンド、つまりコンソール・デバイス名が左列に表示されています。中央と右の列は、デバイス・タイプ固有のその他情報です。

ここで、たとえばデバイス dkb100 とデバイス dkc100 が同じデバイスまでの 2 本のパスとします。dkb100 は、アダプタ PKB0 を通るパスの名前であり、dkc100 はアダプタ PKC0 を通るパスの名前です。これは中央の列を見ればわかります。ここでは、情報名に HSZ 割り当てクラスが指定されています。HSZ 割り当てクラスにより、同じ HSZ デバイスまでの“デバイス”がどれかわかります。

注意

コンソールでは、コンソール INIT コマンドを発行するまで HSZ 割り当てクラスの値の変更を認識できません。

```
>>>SHOW DEVICE
dkb0.0.0.12.0          $55$DKB0          HSZ70CCL  XB26
dkb100.1.0.12.0       $55$DKB100       HSZ70     XB26
dkb104.1.0.12.0       $55$DKB104       HSZ70     XB26
dkb1300.13.0.12.0     $55$DKB1300      HSZ70     XB26
dkb1307.13.0.12.0     $55$DKB1307      HSZ70     XB26
dkb1400.14.0.12.0     $55$DKB1400      HSZ70     XB26
dkb1500.15.0.12.0     $55$DKB1500      HSZ70     XB26
dkb200.2.0.12.0       $55$DKB200       HSZ70     XB26
dkb205.2.0.12.0       $55$DKB205       HSZ70     XB26
dkb300.3.0.12.0       $55$DKB300       HSZ70     XB26
dkb400.4.0.12.0       $55$DKB400       HSZ70     XB26
dkc0.0.0.13.0         $55$DKC0          HSZ70CCL  XB26
dkc100.1.0.13.0       $55$DKC100       HSZ70     XB26
dkc104.1.0.13.0       $55$DKC104       HSZ70     XB26
dkc1300.13.0.13.0     $55$DKC1300      HSZ70     XB26
dkc1307.13.0.13.0     $55$DKC1307      HSZ70     XB26
dkc1400.14.0.13.0     $55$DKC1400      HSZ70     XB26
dkc1500.15.0.13.0     $55$DKC1500      HSZ70     XB26
dkc200.2.0.13.0       $55$DKC200       HSZ70     XB26
dkc205.2.0.13.0       $55$DKC205       HSZ70     XB26
dkc300.3.0.13.0       $55$DKC300       HSZ70     XB26
dkc400.4.0.13.0       $55$DKC400       HSZ70     XB26
```

SCSI と Fibre Channel ストレージに対するマルチパスの構成 6.7 マルチパス構成の実装

dva0.0.0.1000.0	DVA0	
ewa0.0.0.11.0	EWA0	08-00-2B-E4-CF-0B
pka0.7.0.6.0	PKA0	SCSI Bus ID 7
pkb0.7.0.12.0	PKB0	SCSI Bus ID 7 5.54
pkc0.7.0.13.0	PKC0	SCSI Bus ID 7 5.54

コンソールでは、現在のパスで I/O に障害が発生すると、デバイスまでの代替パスに自動的に切り替えます。多くのコンソール・コマンドでは、コンソールがアクセスするデバイスのリストを優先順に指定できます。デバイスの複数のパスに対応するマルチパス構成では、デバイスの複数のパスに対応するコンソール・デバイス名のリストを指定できます。たとえば、以下のようなブート・コマンドでは、最初に DKB100 パスでマルチパス・デバイスがブートされ、それが失敗すると、DKC100 パスでブートされます。

```
BOOT DKB100, DKC100
```

OpenVMS Cluster ストレージ・インターコネクトとしての Fibre Channel の構成

OpenVMS の大きな利点は、ネットワーク構成と OpenVMS Cluster System 構成用のインターコネクトとプロトコルを幅広くサポートしていることです。この章では、シングル・システム用のストレージ・インターコネクトと、マルチホスト OpenVMS Cluster システム用の共用ストレージ・インターコネクトとしての Fibre Channel 用 OpenVMS サポートについて説明します。明記された一部の例外を除き、この章の内容は、OpenVMS Alpha システムと OpenVMS Integrity システムに等しく適用されます。

この章の構成は、以下のとおりです。

- OpenVMS Fibre Channel サポートの概要 (第 7.1 節)
- Fibre Channel 構成のサポート (第 7.2 節)
- 構成例 (第 7.3 節)
- Fibre Channel アドレス、WWID、デバイス名 (第 7.4 節)
- Fibre Channel テープのサポート (第 7.5 節)
- AlphaServer コンソールによる Fibre Channel の構成 (第 7.6 節)
- OpenVMS Integrity システムでの Fibre Channel ストレージ・デバイスからのブート (第 7.7 節)
- OpenVMS で使用するためのストレージ・コントローラのセットアップ (第 7.8 節)
- 共用 Fibre Channel システム・ディスクによるクラスタの作成 (第 7.9 節)
- I/O パフォーマンス向上のための割り込みコアレス機能の使用 (第 7.10 節)
- ご使用の環境における Fast Path の使用方法 (第 7.11 節)
- FIBRE_SCAN によるデバイスの表示 (第 7.12 節)

Fibre Channel 構成のマルチパス・サポートについては、第 6 章を参照してください。

注意

この章では、Fibre Channel インターコネクトを模式的に図示します。ノードとストレージ・サブシステムが接続されている部分は水平線で表します。物理的には、Fibre Channel インターコネクトは、図 7-1 に示すように、スイッチから常に放射状に配線されます。

ストレージ・サブシステムにおける複数の SCSI ディスクと SCSI バスも簡略化されています。1つ以上の HSGxコントローラがホストに対する論理ユニットとしてサービスする複数のディスクと SCSI バスは、図では単独の論理ユニットとして表示されています。

この章では、表現を簡略化するために、階層型の Fibre Channel ストレージ・コントローラを意味する HSG という用語で HSG60 と HSG80 を表しています。ただし、HSG60 と HSG80 の違いを強調する場合は例外です。

7.1 OpenVMS Fibre Channel サポートの概要

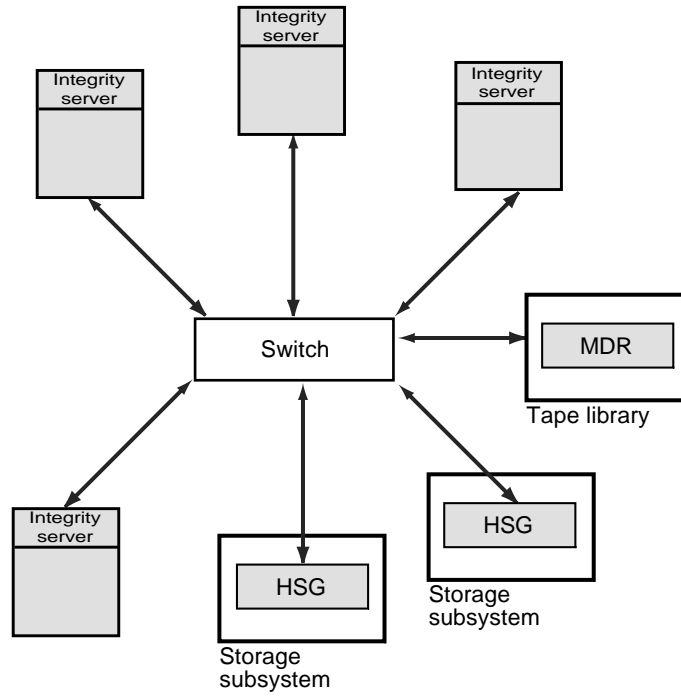
Fibre Channel は、他のインターコネクトに比べさまざまな利点を備えた ANSI 標準ネットワークとストレージ・インターコネクトです。その主な特徴と、これらの特徴に対して OpenVMS が提供するサポートを表 7-1 に示します。

表 7-1 Fibre Channel の特徴と OpenVMS のサポート

特徴	OpenVMS のサポート
高速伝送	OpenVMS は、2 Gb/s、全二重、シリアル・インターコネクトをサポートしています (毎秒 200 MB のデータを同時に送受信可能)。
媒体の選択	OpenVMS は、光ファイバ媒体をサポートしています。
長距離インターコネクト	OpenVMS は、リンク当たり 500 m のマルチモード光ファイバ媒体とリンク当たり最大 100 km のシングル・モード光ファイバ媒体 (インタースイッチ・リンク[ISL]) をサポートしています。
複数のプロトコル	OpenVMS は、SCSI-3 をサポートしています。将来的には、IP をサポート予定。
各種トポロジ	OpenVMS は、スイッチ FC (高いスケーラビリティの複数同時通信) と複数のスイッチ (ファブリック) をサポートしています。また、StorageWorks Modular Storage Array (MSA) 1000 ストレージ・システムでアービトラート型ループのサポートが予定されています。正式なサポートのアナウンスは、次の URL の OpenVMS の Web ページで行われます。 http://www.hp.com/go/openvms

図 7-1 はスイッチ・トポロジを論理的に表した図です。FC ノードは、Alpha ホストかストレージ・サブシステムのどちらかです。ノードからスイッチまでの各リンクは、専用の FC 接続です。スイッチは、ノード・ペア間で保存/送信パケットの配信をします。非結合ノード・ペア間の同時通信はスイッチでサポートしています。

図 7-1 スイッチ・トポロジ (論理図)



VM-0094A-AI

図 7-2 は、Fibre Channel スイッチ・トポロジの実体図を表しています。図 7-2 は解りやすくするために簡素化しています。一般的な構成では第 7.3.4 項に示すように、可用性を高めるために Fibre Channel インターコネクトを多重化しています。

OpenVMS Cluster ストレージ・インターコネクタとしての Fibre Channel の構成
7.1 OpenVMS Fibre Channel サポートの概要

図 7-2 スイッチ・トポロジ (実体図)

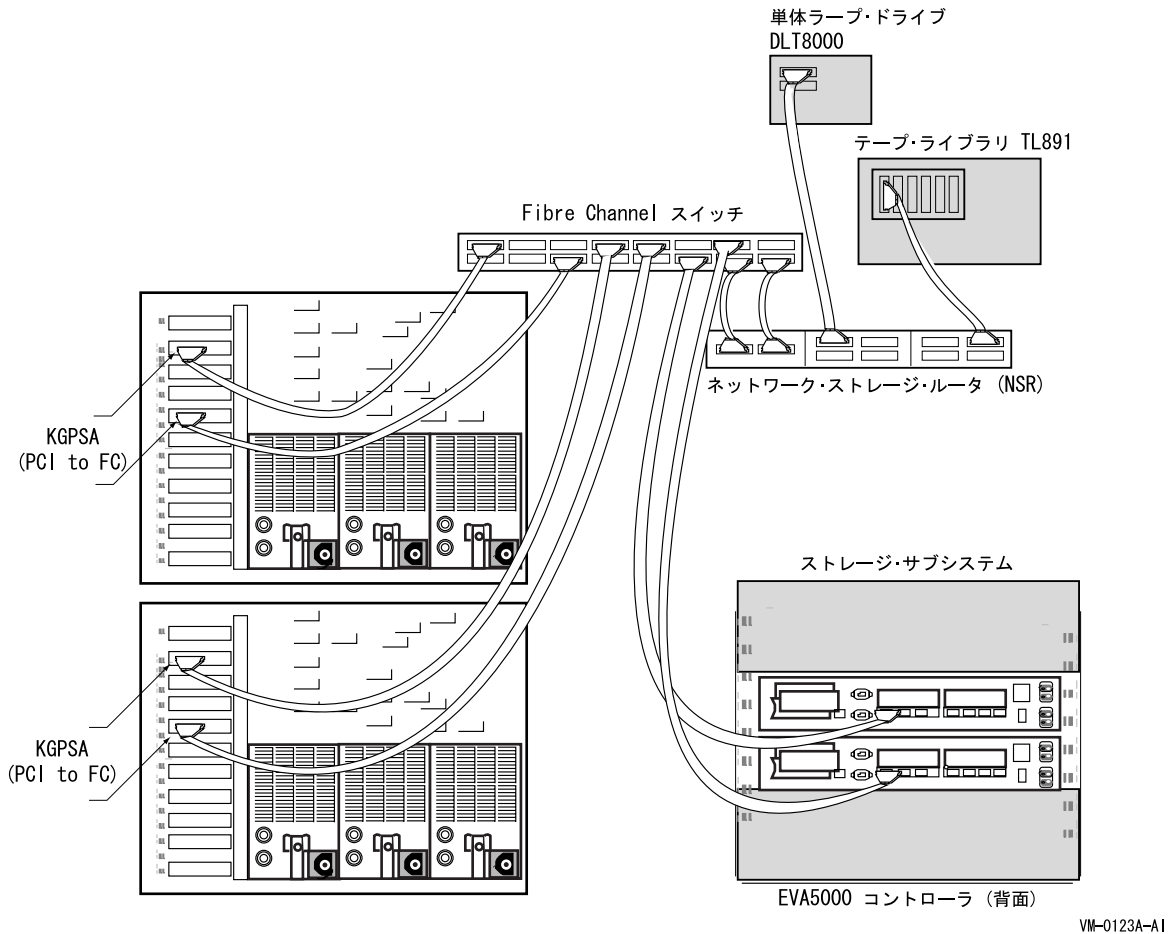


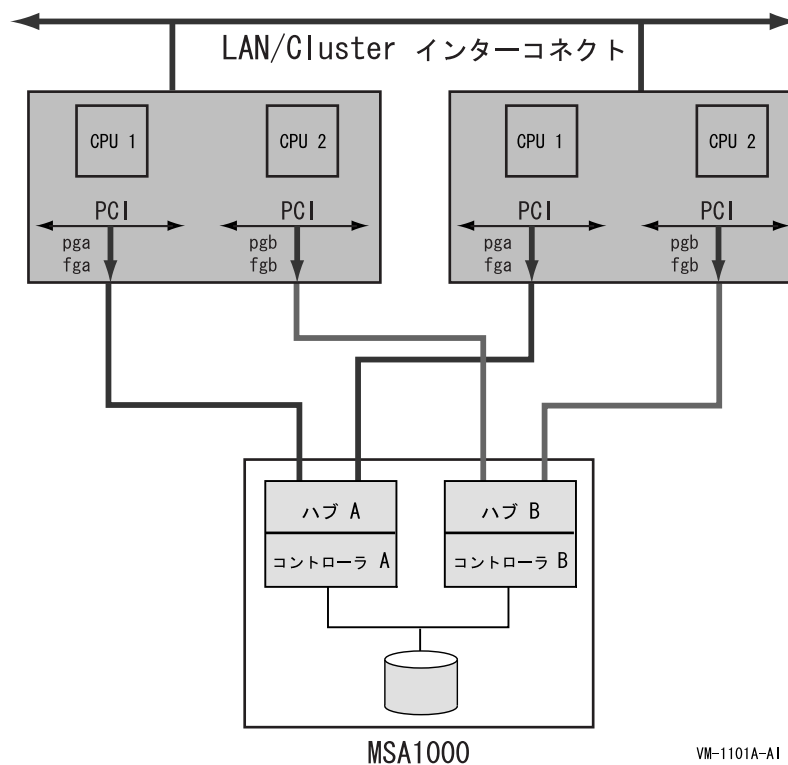
図 7-3 に示すのはアービトレート型ループです。2つのホストとが、デュアル・ポート StorageWorks MSA 1000 ストレージ・システムに接続されています。OpenVMS は、このストレージ・システムでのみアービトレート型ループ・トポロジをサポートします。

注意

このトポロジのサポートは次の URL の OpenVMS の Web ページでアナウンスされる予定です。

<http://www.hp.com/go/openvms>

図 7-3 MSA 1000 を使用したアービトレート型トポロジ



7.2 Fibre Channel 構成のサポート

OpenVMS は、ソフトウェア仕様書にリストされている Fibre Channel デバイスをサポートしています。OpenVMS でサポートされている Fibre Channel ファブリック・コンポーネントについては、最新バージョンの『HP StorageWorks SAN Design Reference Guide』(AA-RMPNM-TE) を参照してください。

Fibre Channel ハードウェア名の規則では、Fibre Channel 固有のハードウェアを G という文字で表します。他の Fibre Channel 機器による Fibre Channel 構成はサポートしていません。オペレーティング・システムとファームウェアの必要最小バージョンを確認するには、リリース・ノートを参照してください。

すべての OpenVMS Fibre Channel 構成で、実行している OpenVMS バージョン用の最新のアップデート・キットを使用することをお勧めします。

これらのキットのルート名は FIBRE_SCSI です。FIBRE_SCSI は以前は FIBRECHAN と呼ばれていました。このキットは、次の URL から入手することができます。

<http://welcome.hp.com/country/us/eng/support.html>

OpenVMS は、最新の『HP StorageWorks SAN Design Reference Guide』および Data Replication Manager (DRM) のユーザ・マニュアルで説明されている Fibre Channel SAN 構成をサポートします。これには、以下のサポートが含まれていません。

- マルチスイッチ FC ファブリック
- 最大距離 500 m のマルチモード・ファイバのサポート、およびシングル・モード・ファイバを使用した最大距離 100 km のインタースイッチ・リンク (ISL) のサポート。さらに、DRM 構成は、Open Systems Gateway と Wave Division Multiplexors を使用した長距離 ISL を提供します。
- ファブリックおよび HSG ストレージの OpenVMS 以外のシステムとの共有

StorageWorks のマニュアルは次の URL の OpenVMS の Web サイトから入手できます。

<http://www.hp.com/go/openvms>

HP Storage を選択し(左のナビゲーション・バーの関連リンクから)、次にストレージ製品を検索してください。その後、その製品のドキュメントにアクセスできます。

StorageWorks のマニュアルで説明されている構成内では、OpenVMS は以下の機能と制限があります。

- すべての OpenVMS ディスク機能(システム・ディスク、ダンプ・ディスク、シャドウ・セット・メンバ、クォーラム・ディスク、および MSCP サービス対象のディスク)がサポートされます。各仮想ディスクには、クラスタ内で一意の識別子を割り当てる必要があります。
- OpenVMS は、StorageWorks マニュアルで指定されている数のホスト、スイッチ、およびストレージ・コントローラをサポートします。一般的に、ホストおよびストレージ・コントローラの数、利用可能なファブリック接続の数によってのみ制限されます。

- プラットフォームあたりの Fibre Channel ホスト・バス・アダプタの数は、プラットフォーム・タイプによって決まります。現在、最大のプラットフォームでは、そのプラットフォームで実行されている OpenVMS インスタンスの数に関係なく、最大で 26 のアダプタをサポートします。
- OpenVMS では、HSG を SCSI-3 モードで稼動する必要があります。また、HSG がデュアル冗長構成の場合、HSG はマルチバス・フェールオーバ・モードにする必要があります。HSG は、これらのモードで稼動する他のシステムとだけ共有できます。
- OpenVMS Fibre Channel ホスト・バス・アダプタは FC スイッチに直接接続する必要があります。ホスト・バス・アダプタの、Fibre Channel ループへの接続、または別の Fibre Channel エンド・ノードへのポイント・ツー・ポイント接続はサポートされません。
- KGPSA-BC または KGPSA-CA のどちらも、S3 Trio 64V+ ビデオ・カード (PB2GA-JC/JD) と同じ PCI バスには接続できません。AlphaServer 800 で、KGPSA をインストールした場合は、必須 S3 Trio を無効にする必要があります。
- ファブリックのホストは、シングル・クラスタとして構成できます。または複数クラスタと非クラスタ・ノードのいずれか、あるいはその両方として構成することもできます。各クラスタおよび各非クラスタ・システムからそれぞれのストレージ・デバイスに排他的アクセスができるようにする必要があります。HSG/HSV Selective Storage Presentation (SSP) または FC スイッチ・ゾーニング、あるいはその両方を使用すると、各 HSG/HSV ストレージ・デバイスは、1 つのクラスタまたは 1 つの非クラスタ・システムからのみアクセスできます。
- HSG は、制限された数の接続をサポートします。接続は、HSG の特定のポートと特定のホスト・バス・アダプタとの通信が非揮発性レコードとして記録されます。HSG CLI コマンド SHOW CONNECTIONS を参照してください。HSG ACS V8.6 は、最大で 96 の接続をサポートし、HSG ACS V8.5 では、最大 64 の接続が許可され、HSG ACS V8.4 では最大 32 の接続が許可されます。接続制限は、シングルおよびデュアル冗長性コントローラのどちらでも同じです。

FC ファブリックが大きく、アクティブな接続の数が HSG の制限を越えている場合は、接続数を減らすために、ファブリックを再構成するか、FC スイッチ・ゾーニングを使用していくつかの HSG ポートからアダプタのいくつかを「隠す」必要があります。

HSG は、ホスト・バス・アダプタを切断しても、接続テーブルから接続情報を削除しません。このため、ユーザは CLI コマンドを使用して接続情報を明示的に削除し、テーブルが一杯にならないようにする必要があります。

この構成サポートは、このマニュアルの改訂時点で有効なものです。

すでに説明した構成に加え、OpenVMS は SANworks Data Replication Manager もサポートしています。これは長距離で Fibre Channel を使用するためのリモート・データ保管ソリューションです。詳細は、次の Compaq StorageWorks の Web サイトを参照してください。

<http://www.compaq.com/products/storageworks>

7.2.1 Fibre Channel 修正キット

新しい Fibre Channel ハードウェアとより大規模な構成の評価は継続的に行われています。新しいハードウェアとより大規模な構成のために、OpenVMS の Fibre Channel サポートの拡張が必要になる場合があります。OpenVMS のリリース間での Fibre Channel ソフトウェアの拡張と修正は、HP のサポート Web サイトから入手可能な修正キットを使って行うことができます。

各キットの最新バージョンは、常に弊社のサポートの Web サイトにあります。定期的にこの Web サイトをチェックすることをお勧めします。

また、次の URL の Fibre Channel の Web サイトもチェックすることをお勧めします。

<http://h71000.www7.hp.com/openvms/fibre/>

この Fibre Channel Web サイトは、重要な情報、新しいスライド・プレゼンテーションを提供するために定期的に改訂されます。

7.2.2 混成バージョンと混成アーキテクチャのクラスタ・サポート

共用 Fibre Channel OpenVMS Cluster ストレージは、混成バージョンと混成アーキテクチャの両 OpenVMS Cluster システムでサポートされています。混成バージョンのサポートについては、ソフトウェア仕様書を参照してください。混成アーキテクチャのサポートとは、OpenVMS Alpha システムと OpenVMS Integrity システムを組み合わせることで、

混成アーキテクチャ・クラスタでは、それぞれのアーキテクチャごとに最低 1 つのシステム・ディスクが必要になります。

この場合、以下の構成要件を適用します。

- 同じストレージ・デバイスに対する共用アクセス用に構成されたホストは、すべて同じ OpenVMS Cluster にあるものとする。
- クラスタ内のすべてのホストに、LAN、IP ネットワーク、または MEMORY CHANNEL など共通クラスタ通信インターコネクトが必要である。
- FC との直接接続を備えるすべてのホストでは、サポートされているバージョンの OpenVMS Integrity あるいは OpenVMS Alpha が動作している必要がある。
- すべてのホストに、インストールされている混成バージョン・クラスタ用の修正キットが必要である (リリース・ノートを参照)。

- DECEvent をエラー・トレースに使用する場合は、バージョン 2.9 以降が必要である。それ以前のバージョンの DECEvent は Fibre Channel をサポートしていません。

7.3 構成例

この節では、Fibre Channel の構成例を示します。

注意

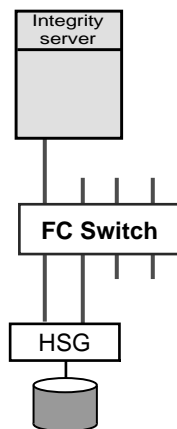
これらの構成は、HSG ストレージ・コントローラと HSV ストレージ・コントローラで有効です。ただし、第 7.3.1 項と第 7.3.2 項については、HSG ストレージ・コントローラにのみ適用されます。

構成は、最初に示す最小の構成に、冗長構成要素を加えて、可用性、パフォーマンス、およびスケーラビリティのレベルを上げていく方法で、交互に構築していきます。

7.3.1 デュアル・ポート・ストレージを備えたシングル・ホスト

図 7-4 は、ストレージ・インターコネクトに Fibre Channel を使用したシングル・システムです。

図 7-4 1つのデュアル・ポート・ストレージ・コントローラを備えたシングル・ホスト



VM-0482A-AI

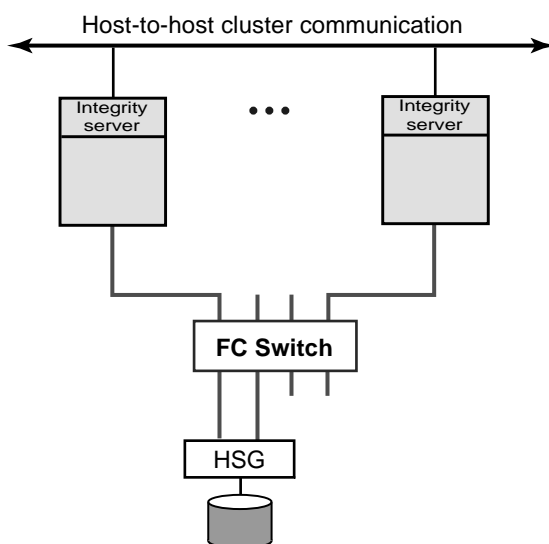
この構成の以下の点に着目してください。

- HSG ストレージ・コントローラまたは HSV ストレージ・コントローラのデュアル・ポートにより、ストレージ・サブシステムの可用性とパフォーマンスが向上します。
- スイッチの予備ポートにより、システムを拡張できます。
- パフォーマンスを最大にするために、論理ユニットを 2 つの HSG ポートまたは HSV ポートに分散できます。
- スイッチと HSG または HSV は単一点障害 (Single point of failure (単一機器の障害がシステム全体の障害になる)) になり得ます。より高い可用性を実現するためには、Volume Shadowing for OpenVMS を使用して、別の Fibre Channel スイッチと HSG コントローラまたは HSV コントローラにデータを複製します。

7.3.2 1 つのデュアル・ポート・ストレージ・コントローラを備えた複数のホスト

図 7-5 は、デュアル・ポート・ストレージ・サブシステムに接続されている複数のホストです。

図 7-5 1 つのデュアル・ポート・ストレージ・コントローラを備えた複数のホスト



VM-0483A-AI

この構成の以下の点に着目してください。

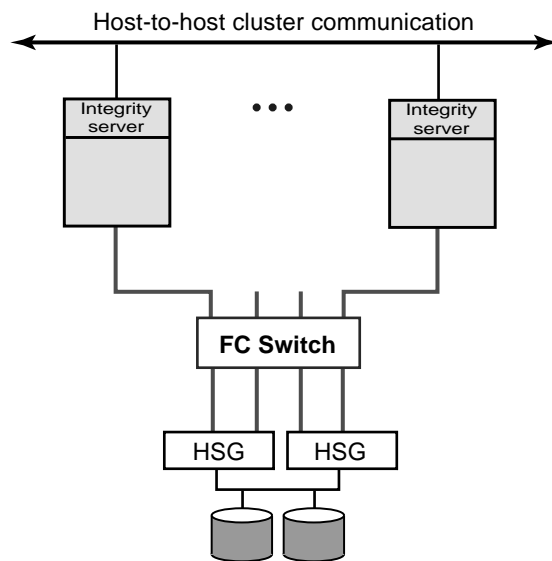
- 複数のホストにより、システム全体の可用性が向上します。
- スイッチの予備ポートにより、システムを拡張できます。

- スイッチと HSG または HSV は単一点障害 (Single point of failure (単一機器の障害がシステム全体の障害になる)) になり得ます。より高い可用性を実現するためには、Volume Shadowing for OpenVMS を使用して、別の Fibre Channel スイッチと HSG コントローラまたは HSV コントローラにデータを複製します。

7.3.3 ストレージ・コントローラ冗長性を備えた複数のホスト

図 7-6 は、2つのデュアル・ポート・ストレージ・コントローラに接続されている複数のホストです。

図 7-6 ストレージ・コントローラ冗長性を備えた複数のホスト



VM-0484A-AI

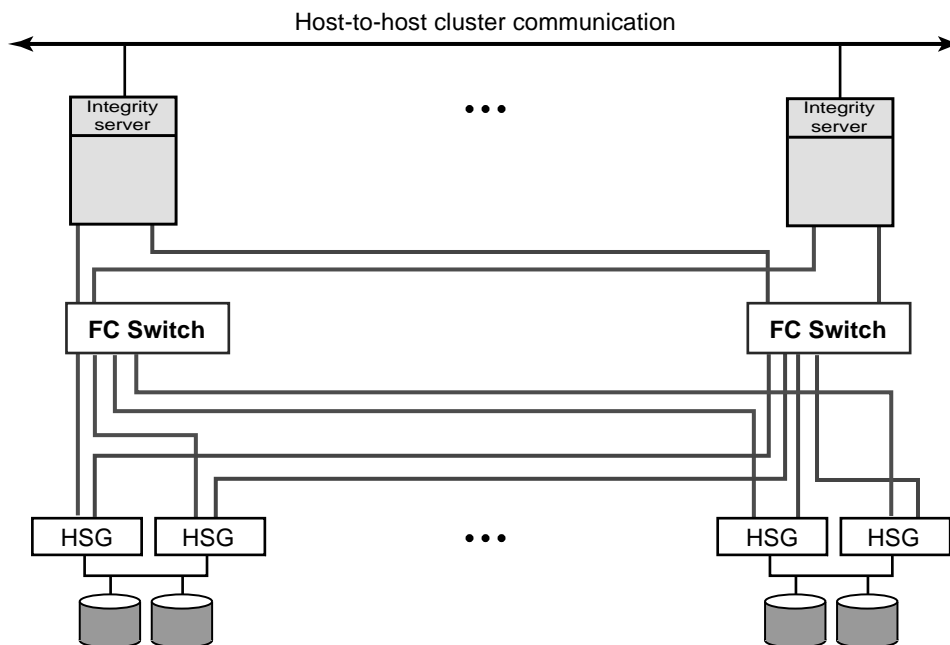
この構成には、以下の利点があります。

- 論理ユニットを4つのHSGポートまたはHSVポートに分散して、より高いパフォーマンスを実現できます。
- 1つのFibre Channel “バス”しかない場合でも、HSGまたはHSVをマルチバス・フェールオーバ・モードで構成できます。
- スイッチは単一点障害 (Single point of failure (単一機器の障害がシステム全体の障害になる)) になり得ます。より高い可用性を実現するためには、Volume Shadowing for OpenVMS を使用して、別の Fibre Channel スイッチと HSG コントローラまたは HSV コントローラにデータを複製します。

7.3.4 複数の独立したスイッチを備えた複数のホスト

図 7-7 は、2つのスイッチに接続されている複数のホストです。このスイッチはそれぞれデュアル・ポート・ストレージ・コントローラのパアに接続されています。

図 7-7 複数の独立したスイッチを備えた複数のホスト



VM-0485A-AI

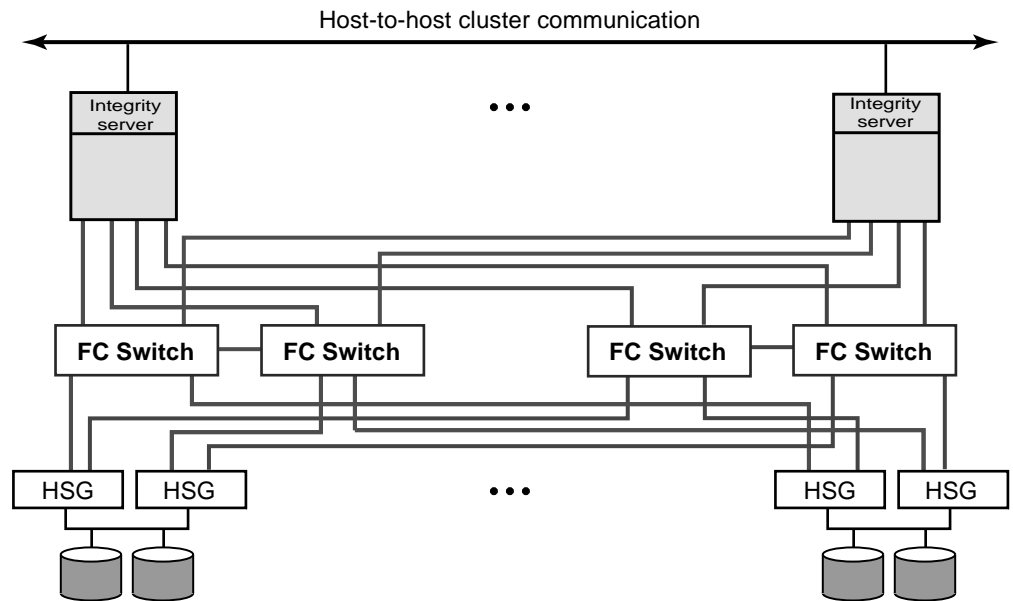
2つのスイッチ構成には、これまでに説明した構成の利点の他に、以下の利点があります。

- 2つのスイッチによって、より高いレベルの可用性が提供されます。単一点障害 (Single point of failure(単一機器の障害がシステム全体の障害になる)) はありません。
- ホスト・バス・アダプタの追加によってパフォーマンスが向上します。
- 各ホストには、ストレージ・サブシステムへの複数の独立パスがあります。2つのスイッチは互いに接続されないため、パスは完全に独立することが保証されます。

7.3.5 デュアル・ファブリックを備えた複数のホスト

図 7-8 は、2つのファブリックに接続されている複数のホストです。各ファブリックは2つのスイッチで構成されています。

図 7-8 デュアル・ファブリックを備えた複数のホスト



VM-0486A-AI

デュアル・ファブリック構成には、これまでに説明した構成の利点の他に、以下の利点があります。

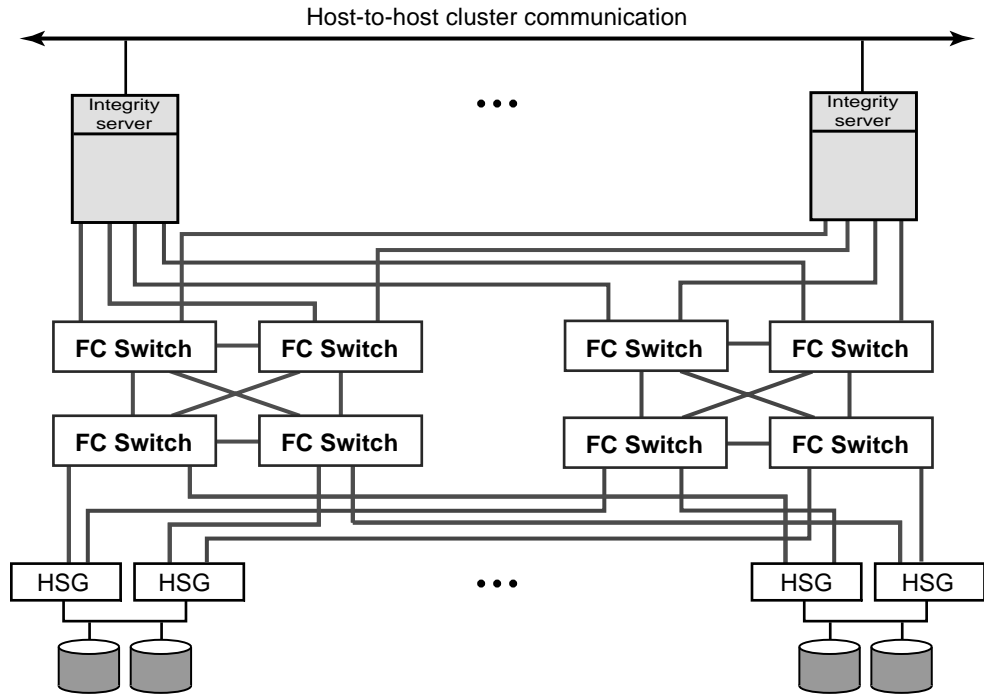
- 各ファブリックにポートを使用して、追加のホストやストレージ・サブシステムを接続することができます。
- 各ホストには、スイッチごとに1つずつ、4つのホスト・バス・アダプタがあります。各ファブリックごとに1つずつ2つのアダプタだけが必須です。冗長アダプタにより、可用性とパフォーマンスを高めることができます。

7.3.6 より大規模なファブリックを備えた複数のホスト

ここに示す構成は、よりレベルの高いパフォーマンスと可用性を実現します。

図 7-9 は、2つのファブリックに接続されている複数のホストです。各ファブリックには、4つのスイッチがあります。

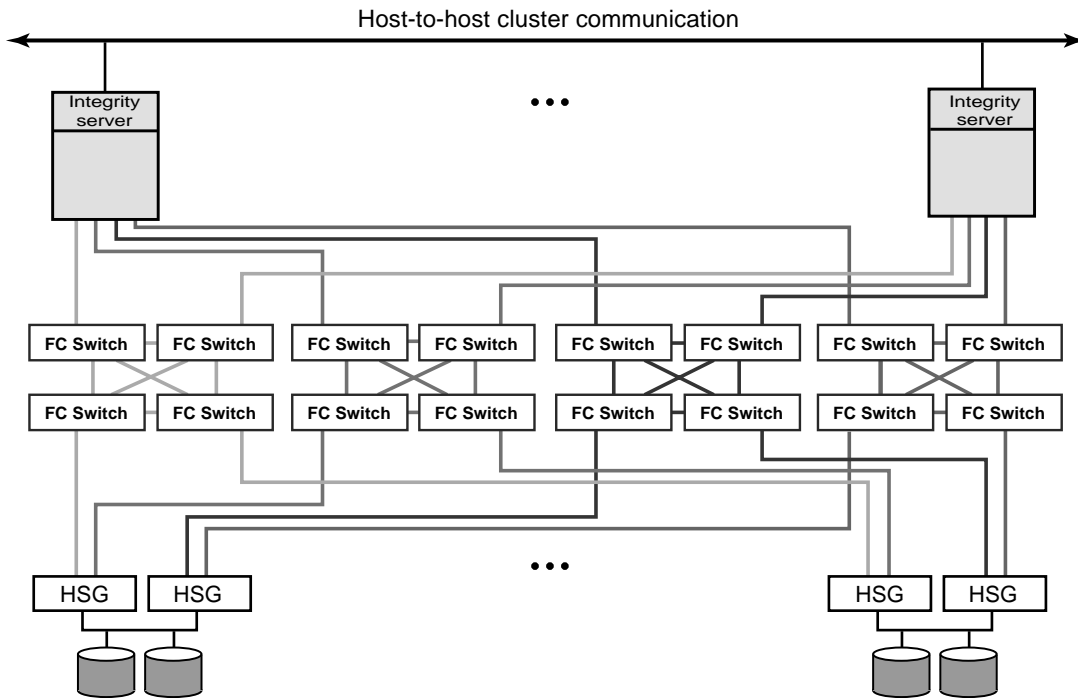
図 7-9 より大規模なデュアル・ファブリックを備えた複数のホスト



VM-0487A-AI

図 7-10 は、4つのファブリックに接続されている複数のホストを示しています。各ファブリックには4つのスイッチがあります。

図 7-10 4つのファブリックを備えた複数のホスト



VM-0488A-AI

7.4 Fibre Channel アドレス，WWID，デバイス名

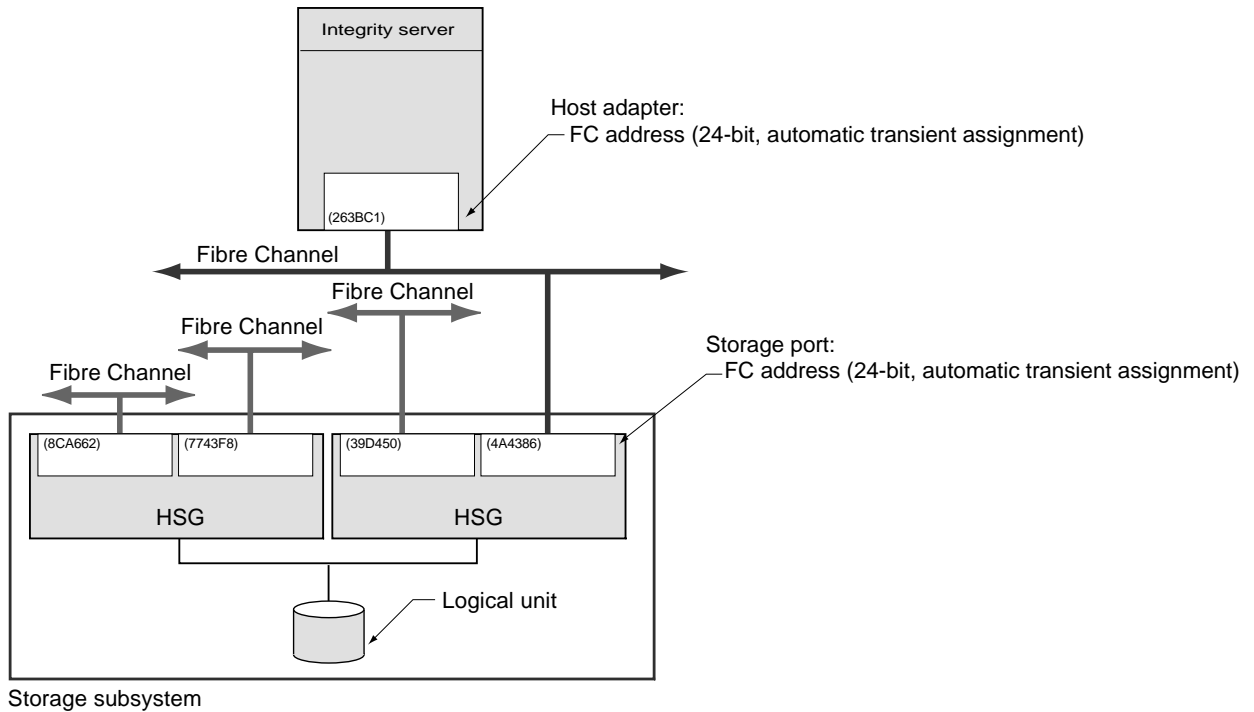
ディスクおよびテープ・ストレージ用の Fibre Channel デバイスには、出荷時にワールドワイド ID (WWID) が割り当てられます。これらの WWID は、システムによる自動 FC アドレス割り当てに使用します。FC WWID とアドレスは、システム管理者が FC 構成でデバイスを識別し検索するときにも利用されます。FC WWID とアドレスは、たとえば Alpha コンソールや HSG コンソールまたは HSV コンソールによって表示されます。そのため、システム管理者はこれらの識別子の意味と OpenVMS デバイス名との関連を理解しておく必要があります。

7.4.1 Fibre Channel アドレスと WWID

一般に、Fibre Channel デバイスは、一時アドレスを持つ構成にします。デバイスのアドレスは、インターコネクトが初期化されるたびに自動的に割り当てられます。デバイスは Fibre Channel が再構成され、再初期化されるたびに新しいアドレスを受け取ります。これは、Fibre Channel デバイスにアドレスジャンパを使用せずに済ませ

るためです。図 7-11 に示すように, ポート当たり 1 つずつ Fibre Channel アドレスが割り当てられます。

図 7-11 Fibre Channel ホスト・アドレスとポート・アドレス

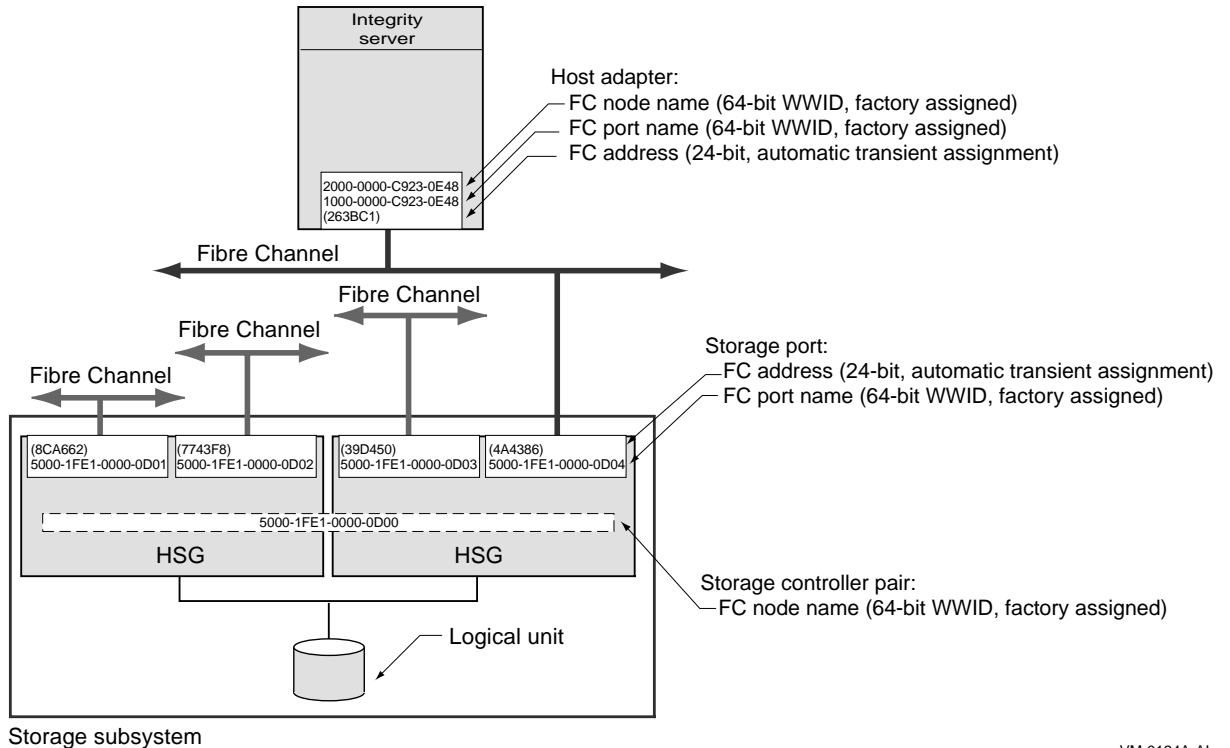


VM-0125A-AI

永続的な識別子を設定するため, 各デバイスの各ポートには WWID が割り当てられます。これは出荷時に設定されます。各 Fibre Channel WWID は一意です。Fibre Channel にはノード WWID も割り当てられ, マルチポート・デバイスを識別します。WWID は, 自動アドレス変更をシステムが検出し, 回復するときに使用します。システム管理者にとっては, 物理デバイスの識別と検索に必要なアドレスです。

図 7-12 は, 出荷時に割り当てられた WWID とその Fibre Channel アドレスを備えた Fibre Channel 構成要素です。

図 7-12 Fibre Channel ホストとポートの WWID とアドレス



VM-0124A-AI

この図の以下の点に着目してください。

- ホスト・アダプタのポート名とノード名はそれぞれ 64 ビットの出荷時に割り当てられた WWID です。
- ホスト・アダプタのアドレスは、自動的に割り当てられた 24 ビットの一時アドレスです。
- 各 HSG ストレージ・ポートまたは HSV ストレージ・ポートには、出荷時に割り当てられた 64 ビットの WWID と、自動的に割り当てられた 24 ビットの一時アドレスがあります。
- HSG コントローラ・ペアまたは HSV コントローラ・ペアは、出荷時に割り当てられた 64 ビットの WWID によるノード名を共有します。

Fibre Channel ホスト・バス・アダプタの FC ノード名と FC ポート名は、SHOW DEVICE/FULL コマンドを使用して表示できます。たとえば、次のように表示されます。

```
$ SHOW DEVICE/FULL FGA0
```

```
Device FGA0:, device type KGPSA Fibre Channel, is online, shareable, error logging is enabled.
```

OpenVMS Cluster ストレージ・インターコネクタとしての Fibre Channel の構成

7.4 Fibre Channel アドレス, WWID, デバイス名

Error count	0	Operations completed	0
Owner process	" "	Owner UIC	[SYSTEM]
Owner process ID	00000000	Dev Prot	S:RWPL,0:RWPL,G,W
Reference count	0	Default buffer size	0
FC Port Name	1000-0000-C923-0E48	FC Node Name	2000-0000-C923-0E48

7.4.2 Fibre Channel デバイス用の OpenVMS 名

各 Fibre Channel ストレージ・アダプタと、ストレージ・アダプタからストレージ・サブシステムまでの各パス、およびストレージ・デバイスごとに OpenVMS 名が割り当てられます。ここで説明する内容は、ディスク・デバイスとテープ・デバイスの両方に適用されます。ただし、第 7.4.2.3 項では、ディスク・デバイスについてのみ説明します。テープ・デバイス名については、第 7.5 節で説明します。

7.4.2.1 Fibre Channel ストレージ・アダプタ名

Fibre Channel ストレージ・アダプタ名の形式は、FGx0 です。OpenVMS によって自動的に割り当てられます。

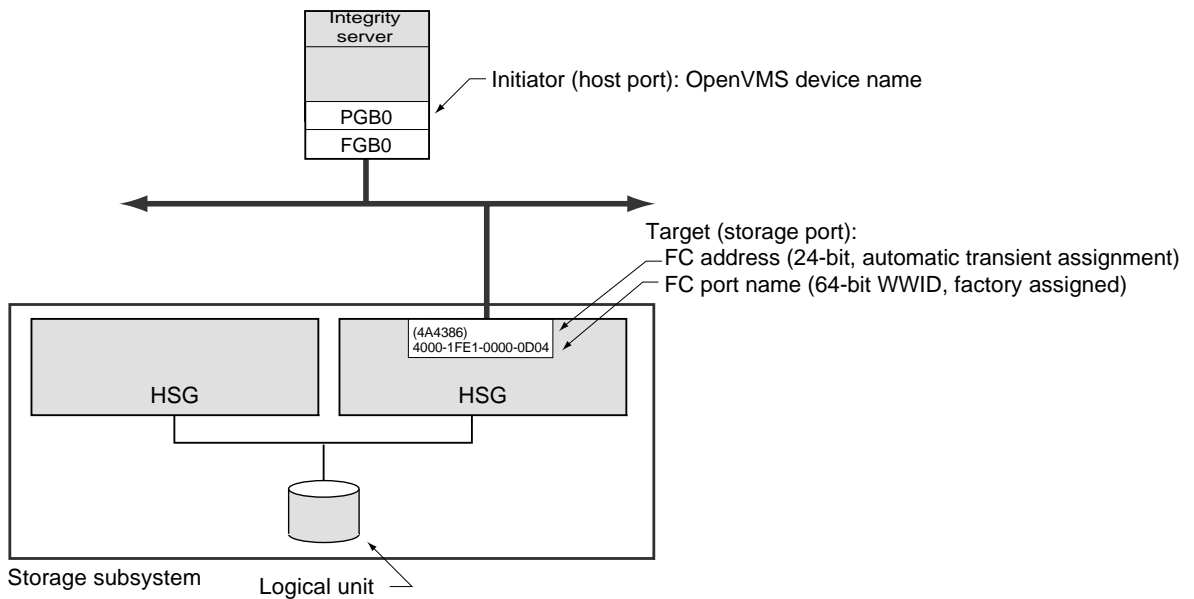
- FG は Fibre Channel を表します。
- x は、スイッチ名を表します。A から Z を使用します。
- 0 は定数です。

命名規則による制限は、システム当たり 26 アダプタです。この命名規則は、将来のリリースでより多くのアダプタをサポートできるよう変更される予定です。

Fibre Channel アダプタは、SCSI や LAN など、複数のプロトコルで実行できます。各プロトコルはアダプタに関連付けられた擬似デバイスです。イニシャル実装では、SCSI プロトコルのみサポートしています。SCSI 擬似デバイス名は、PGx0 です。x は、対応する FGx0 アダプタと同じ装置名を表します。

以上の名前を図 7-13 に示します。

図 7-13 Fibre Channel イニシエータ名とターゲット名



VM-0083A-AI

7.4.2.2 Fibre Channel パス名

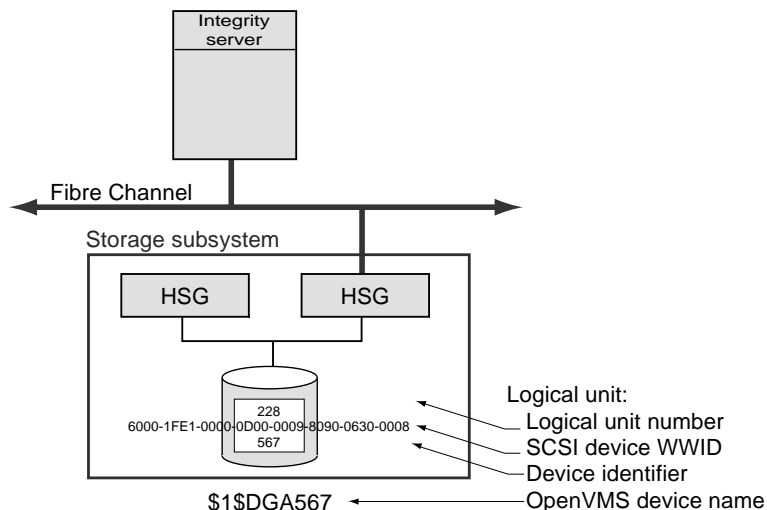
マルチパス SCSI サポートを導入したことにより、第 6 章にあるように、ホストからストレージ・サブシステムまでのパスを指定するようになりました。そのため、小数点(.)で SCSI 擬似デバイス名とアクセス先のストレージ・サブシステム・ポートの WWID を結合しています。たとえば、図 7-13 に示す Fibre Channel パスの名前は PGB0.4000-1FE1-0000-0D04 となります。

Fibre Channel パス名の表示と使用方法の詳細については、第 6 章を参照してください。

7.4.2.3 Fibre Channel ディスク・デバイスの識別

図 7-14 には、各 FC ディスク・デバイスに関連付けられた 4 つの識別子が示されています。

図 7-14 Fibre Channel ディスク・デバイスの命名



VM-0084A-AI

論理ユニット番号 (LUN) は、ストレージ・サブシステム内の特定のデバイスのアドレスとしてシステムが使用します。この番号は HSG コンソールまたは HSV コンソールでシステム管理者が設定し、表示します。OpenVMS SDA ユーティリティで表示することもできます。

各 Fibre Channel ディスク・デバイスには、デバイスの一意の名前を永続的に表す WWID も与えられます。HSG デバイスまたは HSV デバイス WWID は 128 ビットです。この識別子の半分は、論理ストレージ・デバイスを作成した HSG または HSV の WWID であり、残り半分は特定の論理デバイスです。デバイス WWID は、SHOW DEVICE/FULL コマンド、HSG コンソールまたは HSV コンソールと AlphaServer コンソールが表示します。

ストレージ・デバイスに関連付けられた第 3 の識別子は、ユーザが割り当てるデバイス識別子です。デバイス識別子の属性は次のとおりです。

- HSG コンソールまたは HSV コンソールでユーザが割り当てる。
- ユーザが割り当てる識別子は、クラスタ内で一意の識別子とする。
- デバイスとともに移動する。
- MSCP サービス対象のデバイスを除き、0 から 32767 の範囲の 10 進数とする。

FC ディスク・デバイスが MSCP サービス対象の場合、デバイス識別子に指定できる値の上限は 32767 になります。

図 7-14 では、デバイス識別子の値は 567 となっています。この値は、OpenVMS ではデバイス名として作成されるので、クラスタ内で一意であることが必要です。(デバイス識別子を論理ユニット番号 (LUN) と同じ値にしておけば便利です。ただし、この場合も、クラスタ内で一意のデバイス識別子とします。)

Fibre Channel ストレージ・ディスク・デバイス名は、オペレーティング・システムにより、定数 \$1\$DGA とデバイス識別子 (nnnnn) で構成されます。Fibre Channel ディスク・デバイス名では、割り当てクラス値 1 を使用しますが、Fibre Channel テープ・デバイス名では割り当てクラス値 2 を使用します (第 7.5.2.1 項を参照)。この名前の唯一の変数部分はデバイス識別子であり、HSG コンソールまたは HSV コンソールで割り当てます。図 7-14 は、ホストに \$1\$DGA567 として認識されるストレージ・デバイスです。

注意

デバイス識別子 0 は、HSV でサポートされていません。

次の例は、このデバイスの SHOW DEVICE/FULL 出力を表しています。

```
$ SHOW DEVICE/FULL $1$DGA567:
```

```
Disk $1$DGA567: (WILD8), device type HSG80, is online, mounted, file-oriented
device, shareable, device has multiple I/O paths, served to cluster via MSCP
Server, error logging is enabled.
```

Error count	14	Operations completed	6896599
Owner process	"	Owner UIC	[SYSTEM]
Owner process ID	00000000	Dev Prot	S:RWPL,O:RWPL,G:R,W
Reference count	9	Default buffer size	512
WWID	01000010:6000-1FE1-0000-0D00-0009-8090-0630-0008		
Total blocks	17769177	Sectors per track	169
Total cylinders	5258	Tracks per cylinder	20
Host name	"WILD8"	Host type, avail	Compaq AlphaServer GS160 6/731, yes
Alternate host name	"H20FRD"	Alt. type, avail	AlphaServer 1200 5/533 4MB, yes
Allocation class	1		
Volume label	"S5SH_V72_SSS"	Relative volume number	0
Cluster size	18	Transaction count	9
Free blocks	12811860	Maximum files allowed	467609
Extend quantity	5	Mount count	6
Mount status	System	Cache name	"_\$1\$DGA8:XQPCACHE"
Extent cache size	64	Maximum blocks in extent cache	1281186
File ID cache size	64	Blocks currently in extent cache	1260738
Quota cache size	0	Maximum buffers in FCP cache	1594
Volume owner UIC	[1,1]	Vol Prot	S:RWCD,O:RWCD,G:RWCD,W:RWCD

```
Volume Status: ODS-2, subject to mount verification, file high-water marking,
write-back XQP caching enabled, write-through XFC caching enabled.
```

```
Volume is also mounted on H20FRD, FIBRE3, NORLMN, BOOLA, FLAM10.
```

OpenVMS Cluster ストレージ・インターコネクトとしての Fibre Channel の構成

7.4 Fibre Channel アドレス, WWID, デバイス名

I/O paths to device	5		
Path PGA0.5000-1FE1-0000-0D02	(WILD8), primary path.		
Error count	0	Operations completed	14498
Path PGA0.5000-1FE1-0000-0D03	(WILD8), current path.		
Error count	14	Operations completed	6532610
Path PGA0.5000-1FE1-0000-0D01	(WILD8).		
Error count	0	Operations completed	14481
Path PGA0.5000-1FE1-0000-0D04	(WILD8).		
Error count	0	Operations completed	14481
Path MSCP (H20FRD).			
Error count	0	Operations completed	320530

7.5 Fibre Channel テープのサポート

この節では、Fibre Channel テープ機能を利用するために必要な構成要件とユーザ・コマンドについて説明します。Fibre Channel テープ機能は、共用 Fibre Channel ストレージを使用する OpenVMS Cluster システムでの SCSI テープおよび SCSI テープ・ライブラリのサポートを意味します。SCSI テープおよびライブラリは、Fibre Channel と SCSI 間のブリッジによって Fibre Channel に接続されます。現在は、2 つのブリッジ、Modular Data Router (MDR) と Network Storage Router (NSR) が利用できます。

7.5.1 最小ハードウェア構成

最小の Fibre Channel テープのハードウェア構成を以下に示します。

- サポートされている FC HBA 付き Alpha システムまたは Integrity システム
- Fibre-to-SCSI ブリッジ:
 - Network Storage Router (NSR)

NSR も、Alpha システムに直接接続するのではなく、スイッチに接続する必要があります。

OpenVMS では、NSR をインデックス・モードにすることをお勧めします。

コントローラ LUN が LUN 0 にマップされるように、インデックス・マップはターゲット/バス・プライオリティで設定する必要があります。ID が競合しないように注意してください (『hp StorageWorks network storage router M2402』ユーザ・ガイド (269782-003) を参照)。

- Modular Data Router (MDR), 最小ファームウェア・リビジョン 1170

MDR は、Alpha システムに直接接続するのではなく、スイッチに接続する必要があります。また、MDR は、SCSI Command Controller (SCC) モードにする必要があります。通常、このモードがデフォルトです。MDR が SCC モードでない場合は、MDR コンソールで `SetSCCmode On` コマンドを使用してください。

テープ・デバイスおよびテープ・ライブラリ・ロボットは、SCSI ターゲット ID 7 に設定しないようにしてください。この ID は、MDR で使用するために予約されています。

- Fibre Channel スイッチ
- 以下に示すようなテープ・ライブラリ
 - MSL5000 シリーズ
 - ESL9000 シリーズ
 - TL891
 - TL895
- 以下に示すような個別のテープ
 - SDLT 160/320
 - SDLT 110/220
 - HP Ultrium 460
 - HP Ultrium 448c
 - DLT8000
 - TZ89
- SAS テープ・ブレード - C クラス Integrity BladeSystem 用の HP StorageWorks Ultrium 488c および Ultrium 920c

注意

テープは、HSGxxストレージ・サブシステム内でも FCTC-II (Fibre Channel Tape Controller II) を使用する場合でもサポートされません。

テープ・ライブラリ・ロボットは、媒体チェンジャ・デバイスの 1 つの例です。この用語は、この節を通して使用されます。

7.5.2 Fibre Channel テープ・デバイスの命名

ここでは、Fibre Channel テープ・デバイスの命名に関する詳細な背景情報を示します。

テープ・デバイスおよび媒体チェンジャ・デバイスは、第 7.5.3 項で説明する SYSMAN の IO FIND および IO AUTOCONFIGURE のコマンドを使用して自動的に命名および構成されます。Fibre Channel にテープを構成するシステム管理者は、この節でテープの構成手順を参照してください。

7.5.2.1 テープ・デバイスおよび媒体チェンジャ・デバイスの名前

Fibre Channel のテープおよび媒体チェンジャは、Fibre Channel ディスクの命名と同様のスキームを使用して命名されます。

パラレル SCSI では、直接接続されているテープのデバイス名は、デバイスの物理位置を表します。たとえば、MKB301 は、バス B、SCSI ターゲット ID 3、および LUN 1 に存在することを示しています。このような命名スキームは、Fibre Channel 構成には適しません。Fibre Channel 構成では、ターゲットとノードの数が非常に大きくなる場合があります。

Fibre Channel テープ名は、 $S2\$MGA_n$ という形式で表されます。コントローラを表す文字は常に A で、接頭辞は $S2\$$ です。デバイス・ニームニックは、テープを表す場合は MG で、媒体チェンジャを表す場合は GG です。デバイス・ユニット n は、OpenVMS で自動的に生成されます。

名前作成アルゴリズムは、0 から始まる最初の未使用ユニット番号を選択します。Fibre Channel で最初に検出されたテープは $S2\$MGA_0$ 、次のテープは $S2\$MGA_1$ というように順番に命名されていきます。同様に、Fibre Channel で最初に認識された媒体チェンジャの名前は $S2\$GGA_0$ になります。パラレル SCSI バスのテープおよび媒体チェンジャの命名方法は変わりません。

割り当てクラス 2 の使用については注意してください。割り当てクラス 1 は、HSG Fibre Channel ディスク ($S1\$DGA_{nnnn}$) と HSG コンソール・コマンド LUN ($S1\GGA_{nnnn}) のように、ユーザ定義識別子 (UDID) で名前が識別されたデバイスがすでに使用しています。

割り当てクラス 2 は、`SYSS$DEVICES.DAT` ファイルから名前を取得したデバイスが使用します。この名前は、次の項で説明するワールドワイド識別子 (WWID) キーを基準にします。また、GG は、媒体チェンジャと HSG Command Console LUN (CCL) の両方に使用されるニームニックですが、媒体チェンジャは常に割り当てクラス 2 であり、HSG CCL は割り当てクラス 1 である点にも注意してください。

1 つの OpenVMS Cluster システム内では、テープと媒体チェンジャの名前の整合性が自動的に維持されます。テープ・デバイスがクラスタ内のノードで命名されると、クラスタ内の他のすべてのノードが、自動的にそのデバイスと同じ名前を選択します。これにより最初の未使用ユニット番号を選択するアルゴリズムが無効になる場合があります。選択されたデバイス名は、これ以後にクラスタでリブート操作をしても変わりません。

複数の非クラスタ Integrity システムが SAN に存在し、Fibre Channel の同じテープ・デバイスにアクセスする必要がある場合は、上位レベルのアプリケーションで、一貫した命名および同期アクセスを行う必要があります。

7.5.2.2 ワールドワイド識別子 (WWID) の使用

各 Fibre Channel テープ・デバイス名に対し、OpenVMS はその名前に関連付けられている物理デバイスを一意に識別する必要があります。

パラレル SCSI では、直接接続されているデバイスは、物理パス (ポート/ターゲット/LUN) によって一意に識別されます。Fibre Channel ディスクは、ユーザ定義識別子 (UDID) で一意に識別されます。これらの方法は、Fibre Channel テープおよび媒体チェンジャでは非スケーラブルで利用できません。

このため、特定の Fibre Channel テープまたは媒体チェンジャ・デバイスの識別子には、ワールドワイド識別子 (WWID) を使用します。WWID は、デバイス・ファームウェアに常駐し、Fibre Channel 標準によって一意であることが要求されています。

WWID には以下のようないくつかの形式があります。

- IEEE 登録済み WWID (64 ビット・バイナリ)
- ベンダ ID と製品 ID とシリアル番号 (ASCII)

WWID 全体は、バイナリの WWID ヘッダが前に付いた WWID データで構成されます。WWID ヘッダは、WWID データの長さとタイプを示すロングワードです。

通常、デバイスが IEEE WWID をレポートすると、OpenVMS はこの WWID を、そのデバイスの一意に識別される WWID として選択します。デバイスがそのような WWID をレポートしない場合は、ASCII WWID が使用されます。デバイスが IEEE WWID またはシリアル番号情報のどちらもレポートしない場合、OpenVMS はデバイスを構成しません。デバイスの検出プロセス中に、OpenVMS は以下のメッセージを示してデバイスを拒否します。

```
%SYSMAN-E-NOWWID, error for device Product-ID, no valid WWID found.
```

WWID は、バイナリと ASCII データを組み合わせて構成されています。これらのフォーマットは表示可能で、コンソール WWIDMGR ユーティリティで定義されたフォーマットと整合性を図るために使用されます。詳細は、『WWIDMGR Users' Manual』を参照してください。(『Wwidmgr Users' Manual』は、Alpha Systems Firmware Update CD-ROM の [.DOC]ディレクトリにあります。)

WWID ヘッダの後のデータが純粋な ASCII データである場合は、それらのデータを二重引用符で囲む必要があります。

64 ビットの IEEE WWID の表示可能なフォーマットは、ASCII 形式で表される 8 桁の 16 進数 (WWID ヘッダ)、コロン (:) および IEEE WWID データで構成されます。たとえば、次のようになります。

```
0C000008:0800-4606-8010-CD3C
```

ASCII WWID の表示可能なフォーマットは、8 桁の WWID ヘッダ、コロン (:) および 8 バイトのベンダ ID と 16 バイトの製品 ID とシリアル番号を連結したもので構成されます。たとえば、次のようになります。

```
04100022:"COMPAQ DLT8000 JF71209240"
```

注意

ASCII WWID のシリアル番号には、プリント不可能な文字が含まれる場合があります。表示可能なフォーマットでは、そのような文字は \nn として表されます。この nn は、16 進数値を表す 2 桁の ASCII 文字です。たとえば、ヌルは \00 として表されます。

7.5.2.3 ファイルベースのデバイスの命名

Fibre Channel テープ・デバイスと媒体チェンジャ・デバイスは、SYSS\$SYSTEM:SYS\$DEVICES.DAT ファイルの情報に従って構成されます。このファイルは、デバイス当たり 2 つの連続するレコードで構成される ASCII ファイルです。この 2 つのレコードは、次の形式で表されます。

```
[Device $2$devnam]  
WWID = displayable_identifier
```

自動構成中に、Fibre Channel はプローブされ、すべてのデバイスの WWID がフェッチされます。フェッチされた WWID が、SYS\$DEVICES.DAT ファイルのメモリ常駐コピーのエントリと一致すると、デバイスは、その WWID と一対になっているデバイス名を使用して構成されます。

注意

SYS\$DEVICES.DAT ファイルは、ポート割り当てクラス (PAC) 情報にも使用されます。PAC と Fibre Channel テープはブート時にファイル・ベースのデバイス情報にアクセスする必要があるという共通性以外に関連がありません。Fibre Channel テープの命名は同じファイルの 2 回目の使用となります。

デフォルトでは、SYS\$DEVICES.DAT ファイルは、クラスタの共通ディレクトリ SYS\$COMMON:[SYSEXE] に作成されます。

たとえば、SYS\$DEVICES.DAT の以下の部分では、\$2\$MGA300 および \$2\$MGA23 という名前のデバイスが最終的に構成されます。

```
!  
[Device $2$MGA300]  
WWID = 04100022:"COMPAQ DLT8000 JF71209240"  
!  
[Device $2$mga23]  
WWID = 04100022:"DEC TZ89 (C) DECJL01164302"
```

このファイルには、通常 OpenVMS ユーティリティだけが読み込みと書き込みを行います。まれにこのファイルを編集しなければならない場合もあります。第 7.5.5 項で説明するようにデバイスのユニット番号のみ変更できます。ファイルに適用される内部の構文規則を以下にまとめます。

- コメント行 (! で開始される行) およびブランク行を使用できます。
- [Device と、\$2\$xxx] で表されるデバイス名は、空白 (またはなし) で区切ることができます。
- \$2\$接頭辞の指定に失敗すると、コンソールに警告が示されます。

同様に、WWID = を含む行では、等号の両側に空白 (またはなし) を使用できます。すべての行は左揃えにし、512 文字未満にする必要があります。

このファイルの解析では大文字/小文字は区別されません。ただし、1 つだけ重要な例外があります。二重引用符で囲まれた文字はすべて文字通りの意味で解釈されるため、スペースや小文字などの文字は区別されます。二重引用符で囲まれた ASCII データの場合は、コロンと二重引用符の間にスペースを入れることはできません。

また、複数の WWID = 行が 1 行の [Device devnam] に続く場合は、最後の WWID = 値が優先されます。ただし、通常は、[Device devnam] 行ごとに 1 行の WWID = しかありません。

同様に、2 行以上の [Device devnam] で同じデバイス名を指定し、WWID が異なる場合は、ファイルで指定された最後のデバイス名と WWID だけが使用されます。

このファイルはブート時に読み込まれます。また、このファイルには、SYSMAN の IO FIND_WWID コマンドを使って読み込みまたは書き込みを行います。システム固有の追加の SYS\$DEVICES.DAT ファイルがある場合は、システムごとに SYSMAN の IO FIND_WWID を実行することでそのテープの命名レコードが自動的に取りこまれます。SYSMAN の IO FIND_WWID コマンドについては次の項で詳しく説明します。SYS\$DEVICES.DAT ファイルは、以下に示す SYSMAN IO CREATE_WWID および REPLACE_WWID コマンドで変更されている場合があります。

7.5.3 Fibre Channel テープ・デバイスの管理サポート

Fibre Channel テープ・デバイスを管理するために、以下のシステム管理ユーティリティ (SYSMAN) コマンドが用意されています。

- IO FIND_WWID

Fibre Channel 上のすべてのポートをプローブし、検出されていないすべてのテープと媒体チェンジャを認識して、各デバイスに名前を付けます。デバイスとそれらのデバイスに割り当てたデバイス名のリストを表示し、この情報を SYS\$SYSTEM:SYS\$DEVICES.DAT ファイルに自動的に記録します。また、関連のローカルおよびクラスタ全体のメモリ構造も更新します。これはクラスタ全体に実行されます。

このコマンドは、SYSMAN コマンドの IO AUTOCONFIGURE を実行する前に使用します。

CMKRNL 権限が必要です。

- IO LIST_WWID

まだ Fibre Channel に構成されていないすべてのテープ・デバイスの WWID をリストします。SYSMAN コマンドの IO CREATE_WWID を実行する前にこのコマンドを使用してください。

CMKRNL 権限が必要です。

- IO CREATE_WWID

SYSMAN IO LIST_WWID の表示からユーザが特定の (それまで使用されていない) デバイス名を特定の (それまで使用されていない) WWID に割り当てることを可能にします。これはクラスタ全体に実行されます。このコマンドの後にクラスタ全体に対して SYSMAN IO AUTO コマンドを実行して、実際にデバイスを構成します。

このコマンドは、検出された WWID に対してシステム生成のデバイス名を割り当てる SYSMAN IO FIND_WWID コマンドの代りとして使用できます。IO CREATE コマンドは、WWID 相関関係の再定義の手段として IO FIND コマンドの後に実行すべきではありません。IO CREATE_WWID で指定されたデバイスと WWID 文字列は、そのクラスタ以外で使用すべきではありません。

CMKRNL 権限が必要です。

- IO REPLACE_WWID

テープ・デバイスを、同じ FC LUN 位置にある別のテープ・デバイスに物理的に交換する必要がある場合に、該当のファイルおよびメモリ・データ構造を更新します。

CMKRNL 権限が必要です。

Fibre Channel テープ・デバイスについては、以下の DCL サポートを利用できません。

- SHOW DEVICE/FULL コマンドを実行すると、Fibre Channel テープ・デバイスの WWID が表示されます。
- F\$GETDVI レキシカル関数は、キーワード WWID をサポートします。このキーワードは、Fibre Channel テープ・デバイスの WWID を返します。

7.5.4 Fibre Channel テープ・デバイスの構成

ここでは、Fibre Channel で新しいテープまたは媒体チェンジャを構成するために必要な手順を示します。

7.5.4.1 基本構成手順: 要約

クラスタで新しい Fibre Channel テープ・デバイスを構成するための基本手順を以下に示します。

1. 新しいテープ・デバイス (複数可) に電源を投入します。
2. MDR を使用している場合は、MDR の電源を切ってから電源を入れなおし、MDR マッピング情報を更新します。

NSR を使用している場合は、Visual Manager を使用して、次のようにマッピング情報を更新してください。

- 「Mapping」サブメニューをクリックします。ユーザ名は root、パスワードは password です。
- 「Select Map」ボックスが「Indexed」モードを示していることを確認します。
- 「Select Map」ボックスの隣の「Edit/View」ボックスをクリックします。
- これにより空のインデックス・マップが現れます。「Priority」オプションの下で「Target/Bus」プライオリティを選択し、「Fill Map」をクリックします。「Target/Bus」プライオリティにより、コントローラ LUN (Active Fabric (AF) LUN と呼ばれる) が LUN 0 にマップされることに注意してください。
- 新しいマップが現れます。ウィンドウをクローズし、「Mapping」サブメニューに戻ります。
- NSR に FC ポート 0 以外の FC ポートがある場合、「FC Port 1」をクリックし、FC ポート 1 のマッピング処理を繰り返します。それ以外に FC ポートがある場合は、さらに処理を繰り返します。
- 「Reboot」をクリックし、マップの変更を有効にします。

Visual Manager についての詳細は、hp StorageWorks Network Storage Router M2402 のユーザ・ガイドを参照してください。

3. SYSMAN を実行して、デバイス名を割り当て、デバイスを構成します。

```
$ MC SYSMAN
SYSMAN> SET ENVIRONMENT/CLUSTER      ! Execute on all nodes
SYSMAN> IO FIND_WWID                  ! Assign names
SYSMAN> IO AUTOCONFIGURE/LOG          ! Configure devices
SYSMAN> EXIT
```

これらの手順は、最初の構成時に一度だけ実行する必要があります。以降は、システム・リブート時に自動的にデバイスが構成されます。

7.5.4.2 基本構成手順: 詳細

Fibre Channel でテープ・デバイスを構成する前に、デバイスのワールドワイド識別子 (WWID) を検出し、テキスト・ファイル `SYSS$SYSTEM:SYS$DEVICES.DAT` にデバイス名とともに保存する必要があります。これは通常、`SYSMAN` コマンド `IO FIND_WWID` を実行して行うことができます。ただし、`FIND_WWID` によって割り当てられるシステム生成名を使用するのではなく、テープ・デバイスに対してユーザ独自のデバイス名を選択したい場合は、`IO FIND_WWID` の代わりに `IO CREATE_WWID` コマンドを実行します。`IO CREATE_WWID` については次の項で説明します。

`IO FIND_WWID` コマンドは、Fibre Channel 上のすべてのポートをプローブし、MDR または NSR に接続されているすべてのテープと媒体チェンジャを検索します。前に実行した `IO FIND_WWID` コマンドで検出されなかったテープおよび媒体チェンジャについて、`IO FIND_WWID` はデバイス名を割り当て、そのデバイスの WWID を検索し、`SYSS$SYSTEM:SYS$DEVICES.DAT` ファイルにデバイス名と WWID データを保存して、メモリ構造を更新します。

`IO FIND_WWID` の第一の目的は、`SYS$DEVICES.DAT` ファイルを設定することであるため、`IO FIND_WWID` コマンドは、デバイスごとに一度だけ実行する必要があります。`IO FIND_WWID` は、アプリケーションで使用する `2MGAnnnn: デバイス` を構成しません。

このファイルに一度情報を保存すると、次に `IO AUTOCONFIGURE` コマンドを使用したときに、そのファイルのメモリ常駐コピーが読み込まれ、テープと媒体チェンジャ・デバイスが自動的に構成され、必要に応じてデバイス・ドライバがロードまたは接続されます。各システムのリブート時に、`SYS$DEVICES.DAT` ファイルがメモリに読み込まれます。このアクションによって、Fibre Channel でのテープおよび媒体チェンジャの自動構成が開始されます。

`IO FIND_WWID` コマンドを初めて実行したときは、システムの既存のテープ・デバイスと媒体チェンジャ・デバイスがすべて認識されます。後からシステムに Fibre Channel テープ・デバイスを追加した場合は、最初に MDR の電源を切って入れなおし、内部マッピング情報を更新してから、`IO FIND_WWID` コマンドをもう一度実行して、新しいデバイス情報を `SYS$DEVICES.DAT` ファイルに追加する必要があります。NSR では、インデックス・マップを編集してマッピング情報を更新します。

OpenVMS Cluster 環境では、新しい Fibre Channel テープ・デバイスを追加した場合、システムごとにメモリ内のさまざまなデータ構造を更新する必要があります。これを行う場合は、クラスタ内の各 Alpha ノードで `SYSMAN` の `IO FIND_WWID` コマンドを実行することをお勧めします。あるいは、1 つのノードで `IO FIND_WWID` を実行してから、同じシステム・ディスクを共有する他のノードをリブートすることもできます。`SYS$DEVICES.DAT` ファイルはブート時に読み込まれ、その結果、メモリ構造が正常に初期化されるからです。

クラスタ内に複数のシステム・ディスクがある場合は、必ず SYSS\$DEVICES.DAT ファイル内のすべてのコピーの整合性を保つようにします。整合性を保つには、IO FIND_WWID コマンドを実行することをお勧めします。あるいは、IO FIND_WWID を実行して 1 つの SYSS\$DEVICES.DAT ファイルだけを更新し、該当のデバイス名と WWID レコードを元のファイルから残りの SYSS\$DEVICES.DAT ファイルへコピー・アンド・ペーストして、手動で編集することもできます。ただし、2 番目の方法を取る場合、SYSS\$DEVICES.DAT のメモリ常駐コピーをアップデートするには、残りのノードをリブートしなければなりません。

元のファイル全体を別のシステム・ディスクへコピーする方法はお勧めできません。SYSS\$DEVICES.DAT ファイルは、ポート割り当てクラス (PAC) の定義にも使用するため、PAC エントリが誤ってターゲット・システムへ転写される場合があります。

以下に、シングル・ノードで TL891 テープ・ライブラリを使用する構成例を示します。

最初に、SYSMAN コマンド IO FIND_WWID で、検出されていないすべてのテープ・デバイスと、そのデバイス名のリストを表示します。

```
$ MCR SYSMAN IO FIND_WWID
%SYSMAN-I-OUTPUT, command execution on node SAMPLE
On port _SAMPLE$PGA0:, the following tape WWIDs and their proposed device
names have been found but not yet configured:
    [Device $2$GGA0]
    WWID=04100024:"DEC    TL800    (C) DEC3G9CCR82A017"
    [Device $2$MGA0]
    WWID=04100022:"DEC    TZ89     (C) DECCX939S2777"
    [Device $2$MGA1]
    WWID=04100022:"DEC    TZ89     (C) DECCX942S6295"
```

WWID 全体は、等号の右側に示されているものすべてで構成されています。各 WWID は一意ですが、ヘッダ部分は WWID データの基本タイプと長さだけを示すため、一意でない場合があります。

IO FIND_WWID コマンドによって、新しいテープ・デバイスに関する情報を SYSS\$SYSTEM:SYSS\$DEVICES.DAT に記録します。

OpenVMS Cluster ストレージ・インターコネクトとしての Fibre Channel の構成 7.5 Fibre Channel テープのサポート

```
$ TYPE SYS$SYSTEM:SYS$DEVICES.DAT
!
! Updated 23-OCT-2000 14:17:41.85: DEC TL800
!
[Device $2$GGA0]
WWID=04100024:"DEC      TL800      (C) DEC3G9CCR82A017"
!
!
! Updated 23-OCT-2000 14:17:41.93: DEC TZ89
!
[Device $2$MGA0]
WWID=04100022:"DEC      TZ89      (C) DECCX939S2777"
!
!
! Updated 23-OCT-2000 14:17:42.01: DEC TZ89
!
[Device $2$MGA1]
WWID=04100022:"DEC      TZ89      (C) DECCX942S6295"
!
```

次に、**SYSMAN** コマンド **IO AUTOCONFIGURE** でテープ・デバイスを構成します。

```
$ MCR SYSMAN IO AUTOCONFIGURE/LOG

%SYSMAN-I-OUTPUT, command execution on node SAMPLE
%IOGEN-I-PREFIX, searching for ICBM with prefix SYS$
%IOGEN-I-PREFIX, searching for ICBM with prefix DECW$
%IOGEN-I-SCSI POLL, scanning for devices through SCSI port PKA0
%IOGEN-I-SCSI POLL, scanning for devices through SCSI port PKB0
%IOGEN-I-FIBRE POLL, scanning for devices through FIBRE port PGA0
%IOGEN-I-CONFIGURED, configured device GGA0
%IOGEN-I-CONFIGURED, configured device MGA0
%IOGEN-I-CONFIGURED, configured device MGA1
```

最後に、**SHOW DEVICE/FULL** コマンドで、テープ・デバイスの **WWID** を表示します。

```
$ SHOW DEVICE/FULL $2$MG

Magtape $2$MGA0: (SAMPLE), device type TZ89, is online, file-oriented device,
available to cluster, error logging is enabled, controller supports
compaction (compaction disabled), device supports fastskip.

Error count          0      Operations completed          0
Owner process        ""      Owner UIC                    [SYSTEM]
Owner process ID     00000000  Dev Prot                     S:RWPL,O:RWPL,G:R,W
Reference count      0      Default buffer size          2048
WWID 04100022:"DEC  TZ89      (C) DECCX939S2777"
Density              default  Format                        Normal-11
Allocation class     2

Volume status: no-unload on dismount, position lost, odd parity.
```



```
Magtape $2$MGA1: (SAMPLE), device type TZ89, is online, file-oriented device,  
available to cluster, error logging is enabled, controller supports  
compaction (compaction disabled), device supports fastskip.
```

```
Error count          0      Operations completed          0  
Owner process        ""      Owner UIC                      [SYSTEM]  
Owner process ID     00000000  Dev Prot          S:RWPL,O:RWPL,G:R,W  
Reference count      0      Default buffer size          2048  
WWID 04100022:"DEC  TZ89      (C) DECCX942S6295"  
Density              default  Format                      Normal-11  
Allocation class     2
```

```
Volume status: no-unload on dismount, position lost, odd parity.
```

FSGETDVI レキシカル関数も、表示可能な WWID を索します。

```
$ write sys$output f$getdvi("$2$MGA0","WWID")  
04100022:"DEC  TZ89      (C) DECCX939S2777"
```

デバイスが命名および構成されると、INITIALIZE、MOUNT、BACKUP、および COPY などの DCL コマンドを使用してパラレル SCSI テープを使用するのと同じ方法でデバイスを使用できます。Fibre Channel テープの製品固有のサポートについての詳細は、個別のテープ・レイヤ製品のインストール・ガイドを参照してください。

Fibre Channel で媒体チェンジャが自動構成されている間、パラレル SCSI の媒体チェンジャでは、継続して IO CONNECT コマンドでデバイス・ドライバをロードすることを要求します。Fibre Channel 媒体チェンジャはパラレル SCSI と異なり、デバイス名にデバイスの物理位置が含まれていないため、SYSMAN IO CONNECT コマンドを使って Fibre Channel 媒体チェンジャを手動で接続することはできません。

7.5.4.3 ユーザ指定デバイス名の作成

IO FIND_WWID によって生成される省略時の名前を使用する代りにテープ・デバイスに別の名前を使用したい場合は、IO CREATE_WWID コマンドを使用します。次に例を示します。

```
SYSMAN > IO CREATE_WWID $2$MGA3/WWID=04100022:"DEC  TZ89      (C) DECCX939S2341"
```

選択した名前は、メディア・チェンジャーに対しては\$2\$GGAn、テープに対しては\$2\$MGAnの形式でなければなりません。また、nは32767以下でなければなりません。この名前は、クラスタ内の別のデバイスには使用できません。WWIDは、IO LIST_WWIDの出力表示からカット・アンド・ペーストするようにしてください。IO CREATE_WWID コマンドは新しいデバイスのネーミングのために使用してください。既存のデバイス名の変更に使用すべきではありません(既存のデバイス名の変更については第7.5.5項で説明します)。

次の構成例では、IO CREATE_WWID を使用して ESL ライブラリ内の 2 つのテープとメディア・チェンジャーに対してユーザ指定のデバイス名を作成しています。各コマンドは、ノード SYSTM1 および SYSTM2 からなる 2 ノードのクラスタでクラスタ全体に対して実行されています。各ノードは 2 つの Fibre Channel ホスト・バス・

OpenVMS Cluster ストレージ・インターコネクトとしての Fibre Channel の構成

7.5 Fibre Channel テープのサポート

アダプタ PGA0 および PGB0 を持っており、テープへの複数のパスが構成されています。

最初に、SYSMAN コマンド IO LIST_WWID が、それまで検出されていないすべてのデーブ・デバイスを表示します。

```
System1> mcr sysman
SYSMAN> set env/clus
%SYSMAN-I-ENV, current command environment:
      Clusterwide on local cluster
      Username SYSTEM      will be used on nonlocal nodes

SYSMAN> io list_wwid

%SYSMAN-I-OUTPUT, command execution on node SYSTM2
On port _SYSTM2$PGA0:, the following tape WWIDs are not yet configured:

Target 8, LUN 1, HP      ESL9000 Series
WWID=0C000008:0050-8412-9DA1-0026

Target 8, LUN 2, COMPAQ  SDLT320
WWID=02000008:500E-09E0-0009-84D1

Target 8, LUN 3, COMPAQ  SDLT320
WWID=02000008:500E-09E0-0009-4E4E

On port _SYSTM2$PGB0:, the following tape WWIDs are not yet configured:

Target 6, LUN 1, HP      ESL9000 Series
WWID=0C000008:0050-8412-9DA1-0026

Target 6, LUN 2, COMPAQ  SDLT320
WWID=02000008:500E-09E0-0009-84D1

Target 6, LUN 3, COMPAQ  SDLT320
WWID=02000008:500E-09E0-0009-4E4E

%SYSMAN-I-OUTPUT, command execution on node SYSTM1
On port _SYSTM1$PGA0:, the following tape WWIDs are not yet configured:

Target 6, LUN 1, HP      ESL9000 Series
WWID=0C000008:0050-8412-9DA1-0026

Target 6, LUN 2, COMPAQ  SDLT320
WWID=02000008:500E-09E0-0009-84D1

Target 6, LUN 3, COMPAQ  SDLT320
WWID=02000008:500E-09E0-0009-4E4E

On port _SYSTM1$PGB0:, the following tape WWIDs are not yet configured:

Target 5, LUN 1, HP      ESL9000 Series
WWID=0C000008:0050-8412-9DA1-0026

Target 5, LUN 2, COMPAQ  SDLT320
WWID=02000008:500E-09E0-0009-84D1

Target 5, LUN 3, COMPAQ  SDLT320
WWID=02000008:500E-09E0-0009-4E4E
```

OpenVMS Cluster ストレージ・インターコネクタとしての Fibre Channel の構成 7.5 Fibre Channel テープのサポート

```
%SYSMAN-I-NODERR, error returned from node SYSTM1
-SYSTEM-W-NOMORENODE, no more nodes
SYSMAN>
```

コマンドはすべて既存ノードで完了しているため、この NOMORENODE エラーは正常な動作です。次に同じ SYSMAN セッションで IO CREATE_WWID を実行して、デバイス名 \$2\$GGA40、\$2\$MGA40、\$2\$MGA41 を選択します。

```
SYSMAN> io create_wwid $2$GGA40/WWID=0C000008:0050-8412-9DA1-0026
%SYSMAN-I-NODERR, error returned from node SYSTM1
-SYSTEM-W-NOMORENODE, no more nodes
SYSMAN> io create_wwid $2$mga40/WWID=02000008:500E-09E0-0009-84D1
%SYSMAN-I-NODERR, error returned from node SYSTM1
-SYSTEM-W-NOMORENODE, no more nodes
SYSMAN> io create_wwid $2$mga41/WWID=02000008:500E-09E0-0009-4E4E
%SYSMAN-I-NODERR, error returned from node SYSTM1
-SYSTEM-W-NOMORENODE, no more nodes
SYSMAN>
```

IO AUTOCONFIGURE を実行してデバイスを構成します。PGA バスおよび PGB バスの両方が各ノードに対して構成されます。

```
SYSMAN> io auto/lo

%SYSMAN-I-OUTPUT, command execution on node SYSTM2
%IOGEN-I-PREFIX, searching for ICBM with prefix SYS$
%IOGEN-I-PREFIX, searching for ICBM with prefix DECW$
%IOGEN-I-FIBREPOLL, scanning for devices through FIBRE port PGA0
%IOGEN-I-CONFIGURED, configured device GGA40
%IOGEN-I-CONFIGURED, configured device MGA40
%IOGEN-I-CONFIGURED, configured device MGA41
%IOGEN-I-FIBREPOLL, scanning for devices through FIBRE port PGB0
%IOGEN-I-CONFIGURED, configured device GGA40
%IOGEN-I-CONFIGURED, configured device MGA40
%IOGEN-I-CONFIGURED, configured device MGA41

%SYSMAN-I-OUTPUT, command execution on node SYSTM1
%IOGEN-I-PREFIX, searching for ICBM with prefix SYS$
%IOGEN-I-PREFIX, searching for ICBM with prefix DECW$
%IOGEN-I-FIBREPOLL, scanning for devices through FIBRE port PGA0
%IOGEN-I-CONFIGURED, configured device GGA40
%IOGEN-I-CONFIGURED, configured device MGA40
%IOGEN-I-CONFIGURED, configured device MGA41
%IOGEN-I-FIBREPOLL, scanning for devices through FIBRE port PGB0
%IOGEN-I-CONFIGURED, configured device GGA40
%IOGEN-I-CONFIGURED, configured device MGA40
%IOGEN-I-CONFIGURED, configured device MGA41
%SYSMAN-I-NODERR, error returned from node SYSTM1
-SYSTEM-W-NOMORENODE, no more nodes
SYSMAN> exit
Sysm1>
Sysm1> SHOW DEVICE/FULL $2$GG
```

OpenVMS Cluster ストレージ・インターコネクトとしての Fibre Channel の構成
 7.5 Fibre Channel テープのサポート

Device \$2\$GGA40:, device type Generic SCSI device, is online, shareable, device has multiple I/O paths.

Error count	0	Operations completed	0
Owner process	"	Owner UIC	[SYSTEM]
Owner process ID	00000000	Dev Prot	S:RWPL,O:RWPL,G:RWPL,W:RWPL
Reference count	0	Default buffer size	0
WWID	0C000008:0050-8412-9DA1-0026		

I/O paths to device 2
 Path PGA0.1000-00E0-0242-86ED (SYSTM1), primary path, current path.
 Error count 0 Operations completed 0
 Path PGB0.1000-00E0-0222-86ED (SYSTM1).
 Error count 0 Operations completed 0

Systml> sho dev/fu \$2\$MG

Magtape \$2\$MGA40: (SYSTM1), device type COMPAQ SDLT320, is online, file-oriented device, available to cluster, device has multiple I/O paths, error logging is enabled, device supports fastskip (per_io).

Error count	0	Operations completed	2
Owner process	"	Owner UIC	[SYSTEM]
Owner process ID	00000000	Dev Prot	S:RWPL,O:RWPL,G:R,W
Reference count	0	Default buffer size	2048
WWID	02000008:500E-09E0-0009-84D1		
Density	default	Format	Normal-11
Host name	"SYSTM1"	Host type, avail	AlphaServer DS10 466 MHz, yes
Alternate host name	"SYSTM2"	Alt. type, avail	AlphaServer DS10 466 MHz, no
Allocation class	2		

Volume status: no-unload on dismount, position lost, odd parity.

I/O paths to device 2
 Path PGA0.1000-00E0-0242-86ED (SYSTM1), primary path, current path.
 Error count 0 Operations completed 1
 Path PGB0.1000-00E0-0222-86ED (SYSTM1).
 Error count 0 Operations completed 1

Magtape \$2\$MGA41: (SYSTM1), device type COMPAQ SDLT320, is online, file-oriented device, available to cluster, device has multiple I/O paths, error logging is enabled, device supports fastskip (per_io).

Error count	0	Operations completed	0
Owner process	"	Owner UIC	[SYSTEM]
Owner process ID	00000000	Dev Prot	S:RWPL,O:RWPL,G:R,W
Reference count	0	Default buffer size	2048
WWID	02000008:500E-09E0-0009-4E4E		
Density	default	Format	Normal-11
Host name	"SYSTM1"	Host type, avail	AlphaServer DS10 466 MHz, yes
Alternate host name	"SYSTM2"	Alt. type, avail	AlphaServer DS10 466 MHz, no
Allocation class	2		

Volume status: no-unload on dismount, position lost, odd parity.

```
I/O paths to device          2
Path PGA0.1000-00E0-0242-86ED (SYSTM1), primary path, current path.
  Error count                  0    Operations completed          0
Path PGB0.1000-00E0-0222-86ED (SYSTM1).
  Error count                  0    Operations completed          0
Sysctl>
```

7.5.5 既存の Fibre Channel テープ・デバイス名の変更

SYSSYSTEM:SYS\$DEVICES.DAT はテキスト・ファイルであるため編集することができますが、変更できるのは Fibre Channel テープまたは媒体チェンジャー・デバイスのユニット番号だけです。ただし、前述したように、Fibre Channel テープと媒体チェンジャー・デバイスの情報はクラスタ規模のデータ構造、特にクラスタ規模の論理名を使用して、OpenVMS により内部的に保存されます。これらのデータ構造をクリーン・アップするには、クラスタを完全にシャットダウンする必要があります。ローリング・リポート (他のノードのリポート中に最低 1 つのノードを起動したままにしておく方法) は、構造のクリーン・アップには使用できません。

既存のデバイス名を変更するための手順を以下に示します。

1. SYSSYSTEM:SYS\$DEVICES.DAT ファイルを編集して、選択したデバイスのユニット番号を変更します。基本の \$2\$SMGAnnnn または \$2\$SGAnnnn フォーマットで、nnnn の部分だけを編集できます。nnnn の上限値は 32767 で、この値には 10 進数を指定する必要があります。必ず、該当のデバイス・タイプの別のデバイスで使用されていないユニット番号を選択してください。
たとえば、\$2\$MGA26 が別のテープの WWID に関連付けられている場合は、26 以外のユニット番号を選択します。\$2\$GGA4 が別の媒体チェンジャーの WWID にすでに関連付けられている場合は、4 以外のユニット番号を選択します。
2. クラスタ内に複数のシステム・ディスクがある場合は、各 SYS\$DEVICES.DAT ファイルを同じ方法で編集します。
3. クラスタ全体をシャットダウンして、既存のクラスタ・データ構造をクリーン・アップします。
4. クラスタをリポートします。新しいデバイス名が自動的に適用されます。

7.5.6 Fibre Channel での物理テープ・デバイスの移動

テープ・デバイスまたはメディア・チェンジャーを名前を変更せずに移動する場合、リポートする必要はありません。ただし、NSR あるいは MDR が新しい場所でデバイスに FC LUN を割り当てていることを確認し、SYSMAN IO AUTOCONFIGURE を実行してそのデバイスに新しい物理パスを構成してください。メディア・チェンジャーのみの場合、SET DEVICE/SWITCH/PATH=new_path コマンドを使用して手動でそのチェンジャーを新しいパスへ切替える必要もあります。以前のパスは SHOW

DEV/FULL で依然表示されますが、これらのパスは使用されていません。システムをリブートすると古いパスは表示されなくなります。

7.5.7 Fibre Channel 上の NSR のスワップアウト

Alpha あるいは Integrity システムをリブートせずに NSR をスワップアウトすることができます。

新しい NSR を接続した後、Visual Manager の「Mapping」サブメニューを使用して、NSR の各 Fibre Channel ポートでインデックス・マップを有効にし、NSR をリブートすることができます。新しい NSR をマップする別の方法としては、NSR の FTP ユーティリティを使用して以前の NSR から.cfg ファイルをコピーする方法があります。

インデックス・マップを有効にしたら、SYSMAN IO AUTOCONFIGURE を実行してテープへの新しい物理パスを構成してください。メディア・チェンジャーのみの場合、SET DEVICE/SWITCH/PATH=new_path コマンドを使用して手動でそのチェンジャーを新しいパスへ切替える必要もあります。以前のパスは SHOW DEV/FULL で依然表示されますが、これらのパスは使用されていません。システムをリブートすると古いパスは表示されなくなります。

7.5.8 Fibre Channel テープ・デバイスへのサービスの提供

一般的に、OpenVMS Cluster 内の OpenVMS Alpha または OpenVMS Integrity ノードが NSR あるいは MDR と同じ Fibre Channel ファブリックに接続されている場合は、それらのノードすべてに Fibre Channel テープ・デバイスへの直接パスがあります。

媒体チェンジャーは、Fibre Channel とパラレル SCSI のどちらに接続されている場合でも、TMSCP のサービス対象にすることはできません。

7.5.9 Fibre Channel テープ・デバイスの交換

テープ・ドライブを、MDR または NSR 内の同じ FC LUN 位置にある別のテープ・ドライブと物理的に交換する必要がある場合は、該当のデータ構造を IO REPLACE_WWID コマンドを使用して更新します。

たとえば、欠陥のあるテープ・ドライブを、クラスタをリブートせずに新しいドライブと交換する必要があり、このドライブは、その位置にあった前のテープのデバイス名を保持する必要があるとします。

交換用デバイスには、元のデバイスと同じ SCSI ターゲット ID であることが必要です。デバイスを使うすべての処理を終了させてから、次のコマンドを入力して、新し

いテープ・デバイスの WWID で、必要なすべてのファイルとメモリ・データ構造を更新します。

```
$ MCR SYSMAN IO REPLACE_WWID $2$MGA1
```

クラスタ環境内の各 Alpha ノードでこのコマンドを実行します。これは、次のコマンドを使用して実行します。

```
$ MCR SYSMAN  
SYSMAN> SET ENVIRONMENT/CLUSTER  
SYSMAN> IO REPLACE_WWID $2$MGA1
```

場合によっては、デバイス名\$2\$MGA1 は SHOW DEVICE の表示にもはや存在しないためこのコマンドは失敗します。これは、ドライブが誤動作した後、システムが何回かリブートされている場合に発生します。このような場合は、次の例に示すように、デバイス名と WWID の両方を指定する必要があります。

WWID は、置き換えるデバイスと同じ Port/Target/LUN の場所に存在する新しいデバイスの WWID でなければなりません (特定の Port/Target/LUN 位置の WWID の値を調べるには、SYSMAN IO LIST_WWID コマンドを使用します)。

```
$ MCR SYSMAN  
SYSMAN> SET ENVIRONMENT/CLUSTER  
SYSMAN> IO REPLACE_WWID $2$MGA1/WWID=02000008:500E-09E0-0009-4E44
```

注意

このコマンドは、デバイス名の変更あるいは WWID 相関関係の再定義には使用すべきではありません。特定の WWID は、クラスタ内の他のデバイス名と関連付けるべきではありません。

7.5.10 Fibre Channel テープ・デバイスの物理位置の決定

Fibre Channel テープ・デバイスの名前を指定すると、Fibre Channel テープ・デバイスの位置を確認するのに役立ちます。これを行うには、以下の手順を実行します。

1. システム管理者のアカウントで、ANALYZE/SYSTEM を実行します。
2. SDA プロンプトで、CLUE SCSI/SUMMARY と入力します。
3. 「DEVICE」というラベルの列で、デバイスの名前 (たとえば MGA3) を検索します。
4. 「SCSI-LUN」というラベルの列の該当する値に注目してください。この SCSI LUN 値は、FC LUN として MDR または NSR で使用されている値と同じです。また、デバイスの「Port」および「SCSI-ID」というラベルの列にも注目してください。その同じポートと SCSI-ID に関連付けられているデバイスはすべて同じ MDR または NSR の同じ Fibre Channel ポートに接続されています。

5. NSR の場合、Visual Manager の「Mapping」サブメニューを入力して、「Select Map」ボックスの隣の「Edit/View」をクリックし、現在のポートのインデックス・マップを表示します。

「FC LUN」というラベルの列で、手順 4 で説明した値を探します。「FC LUN」の値が見つかったら、同じ行の「SCSI Bus」、「SCSI Target ID」、および「SCSI LU」の対応する値を確認します。この B:T:L 情報には、MDR 内のデバイスの物理位置が記述されています。Visual Manager のメイン・メニューで「Report」をクリックすることにより、他のデバイス情報が参照できます。

MDR が使用されている場合、MDR コンソールの AMC プロンプトで ShowFcScsiMap と入力することにより、同様の情報を参照することができます。

7.5.11 スタンドアロン環境での Fibre Channel テープ・デバイスへのアクセス

Fibre Channel テープ・デバイスは、CDROM 配布キットからブートする際に構成できます。構成手順は、第 7.5.4 項で説明した手順と同じです。特に、テープ・デバイスを使用する前に SYSMAN の IO FIND_WWID および IO AUTOCONFIGURATION コマンドを使用して構成しておく必要があります。

SYSSDEVICES.DAT ファイルは、スタンドアロン環境では作成されません。このため、関連するすべての命名情報はメモリ・データ構造に保存されます。CDROM をブートするたびに、IO FIND_WWID と IO AUTOCONFIGURE コマンドを実行し、テープ・デバイスの命名と構成を行う必要があります。

CDROM ブート環境における Fibre Channel テープ・デバイスの名前は、リポート中には保持されず、読み込み/書き込みシステム・ディスクからのブート時に割り当てられる名前と異なる場合があります。

7.5.12 マルチパス・テープのサポート

NSR あるいは MDR により SCSI テープ・デバイスが Fibre Channel に接続されている Fibre Channel 構成では、Alpha あるいは Integrity ホストから SCSI テープまでの間に複数のパスを存在させることができます。たとえば、4 つの KGPSA アダプタを持つ AlphaServer ホストの場合、Fibre Channel の 1 つのテープまで異なる 4 つのパスを存在させることができます。また、NSR 自体がデュアル・ポートの場合は、NSR へ 2 つのパスを使用できます。つまり、4 つの KGPSA がデュアル・ポート NSR へつながれている AlphaServer システムには、AlphaServer システムから特定のテープ・ドライブへ異なる 8 つのパスが存在します。

OpenVMS Alpha では、Alpha あるいは Integrity システムから SCSI テープまでのあらゆるパスを構成し利用することができます。DCL コマンド SET DEVICE /SWITCH を使用すると、特定のパスを指定できます。また、接続が切断された場合は、自動フェールオーバーが発生します。

注意

テープ・デバイスでは、直接パスと MSCP サービス対象パスとの間のマルチパス・フェールオーバがサポートされていません (SCSI ディスクや Fibre Channel ディスクでの直接パスと MSCP サービス対象パスとの間のマルチパス・フェールオーバとは異なります)。

ただし、マルチパス・セットの TMSCP クライアントについてはサポートされています。この場合は、サービス対象のマルチパス・セットのすべてのメンバを Fibre Channel に直接接続する必要があります。マルチパス・セットの 1 つのメンバで障害が発生すると、別のメンバが、デバイスへのローカル・パスを提供し、クライアントはそのパスを使用します。

7.6 AlphaServer コンソールによる FC の構成 (Alpha のみ)

AlphaServer コンソールは、FC インターコネクトの状態を確認するときに利用します。これにより、インターコネクトが正しくセットアップされているかをブート前に確認できます。FC ディスク・デバイスでブートやダンプするときには、FC ディスク・デバイスのセットアップのためにコンソールから追加作業が必要です。次にこの作業内容を説明します。

7.6.1 コンソールによる FC 構成の確認

コンソールの SHOW コマンドは、コンソールから前回システムの I/O アダプタを調べたときに検出されたデバイスに関する情報を表示できます。他のインターコネクトと異なり、FC ディスク・デバイスは、SHOW DEVICE 出力に自動的に組み込まれません。FC デバイスが WWID で識別され、WWID は SHOW DEVICE の出力に組み込むには大きすぎるためです。代わりに、コンソールには WWID を操作するコマンドが用意されています。これを `wwidmgr` コマンドといいます。このコマンドでは、FC デバイスに関する情報を表示したり、ブートやダンプに使用する FC デバイスのデバイス名を定義することができます。

`wwidmgr` コマンドの使用時には以下の点に着目してください。

- `wwidmgr` コマンドを使用するとき、システムが AlphaServer model 8x00, 4x00, または 1200 の場合、まず診断モードを起動します。他のどのプラットフォームでも、`wwidmgr` コマンドはいつでも発行できます。
- `wwidmgr` コマンドによる変更内容は、次のシステム初期化で初めて有効になります。`wwidmgr` コマンドを実行した後に、`initialize` コマンドを実行してください。

`wwidmgr` コマンドの詳細については、『`Wwidmgr Users' Manual`』を参照してください。(『`Wwidmgr Users' Manual`』は、Alpha Systems Firmware Update CD-ROM の[.DOC]ディレクトリにあります。)

以下の例は、AlphaServer 4100 システムで作成したのですが、`wwidmgr` コマンドの典型的な用例を示しています。他の環境では、さらに追加作業が必要であり、他のシステムにおける出力は少し異なる場合があります。

例 7-1 については、以下の点に着目してください。

- `wwidmgr -show wwid` コマンドでは、システム上の FC デバイスの要約情報を表示できます。このコマンドでは、デバイス接続に関する情報は表示されません。
- ここには、FC アダプタが 2 つとディスクが 5 つ表示されています。(どのディスクも接続先とは無関係にリストの最後に掲載されています。) この例で、各ディスクには HSG80 コンソールでデバイス識別子が割り当てられています。コンソールでは、この識別子をユーザ割り当てデバイス識別子 (UDID) として認識します。

例 7-1 `wwidmgr -show wwid` の使用

```
P00>>>SET MODE DIAG
Console is in diagnostic mode
P00>>>wwidmgr -show wwid
polling kgpsaa0 (KGPSA-B) slot 2 , bus 0 PCI , hose 1
kgpsaa0.0.0.2.1          PGA0          WWN 1000-0000-c920-a7db
polling kgpsaa1 (KGPSA-B) slot 3 , bus 0 PCI , hose 1
kgpsaa1.0.0.3.1          PGB0          WWN 1000-0000-c920-a694
[0] UDID:10 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0016 (ev:none)
[1] UDID:50 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0026 (ev:none)
[2] UDID:51 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0027 (ev:none)
[3] UDID:60 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0021 (ev:none)
[4] UDID:61 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0022 (ev:none)
```

例 7-2 は、FC デバイスに関する情報と接続情報を `wwidmgr show wwid -full` コマンドで表示したものです。この表示は 2 つに分かれています。

- 最初の部分では、アダプタから FC ポートまでのパスをリストします。アダプタは KGPSAA などのコンソール・デバイス名で識別します。FC ポートは、5000-1FE1-0000-0D14 など、それぞれの WWID で識別します。パスに FC ディスクがあると、そのパスの後にリストされます。FC ディスクはそれぞれの現在のコンソール・デバイス名で識別します。このデバイス名は WWID の後に付けます。
- 表示リストの 2 番目の部分には、すべての FC ディスクと、必要なパスがリストされます。この部分で先頭が [0] UDID:10... の個所に、それぞれアダプタ KGPSAA と KGPSAB を通る 2 本のパスを持つパスが、各ディスクまで 4 本あるのがわかります。アダプタを通るパスは、いずれも HSG または HSV 上で別々のポートに達します。タイトルが Con の列は、FC ディスク装置がこのパスで使用している HSG コントローラまたは HSV コントローラに現在オンラインになっているかどうかを示します。

OpenVMS Cluster ストレージ・インターコネクタとしての Fibre Channel の構成
7.6 AlphaServer コンソールによる FC の構成 (Alpha のみ)

例 7-2 wwidmgr -show wwid -full の使用

```
P00>>>wwidmgr -show wwid -full
kgpsaa0.0.0.2.1
- Port: 1000-0000-c920-a7db

kgpsaa0.0.0.2.1
- Port: 2007-0060-6900-075b

kgpsaa0.0.0.2.1
- Port: 20fc-0060-6900-075b

kgpsaa0.0.0.2.1
- Port: 5000-1fe1-0000-0d14
- dga12274.13.0.2.1 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0016
- dga15346.13.0.2.1 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0026
- dga31539.13.0.2.1 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0027
- dga31155.13.0.2.1 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0021
- dga30963.13.0.2.1 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0022

kgpsaa0.0.0.2.1
- Port: 5000-1fe1-0000-0d11
- dga12274.14.0.2.1 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0016
- dga15346.14.0.2.1 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0026
- dga31539.14.0.2.1 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0027
- dga31155.14.0.2.1 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0021
- dga30963.14.0.2.1 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0022

kgpsab0.0.0.3.1
- Port: 1000-0000-c920-a694

kgpsab0.0.0.3.1
- Port: 2007-0060-6900-09b8

kgpsab0.0.0.3.1
- Port: 20fc-0060-6900-09b8

kgpsab0.0.0.3.1
- Port: 5000-1fe1-0000-0d13
- dgb12274.13.0.3.1 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0016
- dgb15346.13.0.3.1 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0026
- dgb31539.13.0.3.1 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0027
- dgb31155.13.0.3.1 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0021
- dgb30963.13.0.3.1 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0022

kgpsab0.0.0.3.1
- Port: 5000-1fe1-0000-0d12
- dgb12274.14.0.3.1 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0016
- dgb15346.14.0.3.1 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0026
- dgb31539.14.0.3.1 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0027
- dgb31155.14.0.3.1 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0021
- dgb30963.14.0.3.1 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0022
```

(次ページに続く)

例 7-2 (続き) `wwidmgr -show wwid -full` の使用

```
[0] UDID:10 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0016 (ev:none)
- current_unit:12274 current_col: 0 default_unit:12274
  via adapter      via fc_nport      Con   DID   Lun
-   kgpsaa0.0.0.2.1 5000-1fe1-0000-0d14 Yes  210013 10
-   kgpsaa0.0.0.2.1 5000-1fe1-0000-0d11 No   210213 10
-   kgpsab0.0.0.3.1 5000-1fe1-0000-0d13 Yes  210013 10
-   kgpsab0.0.0.3.1 5000-1fe1-0000-0d12 No   210213 10

[1] UDID:50 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0026 (ev:none)
- current_unit:15346 current_col: 0 default_unit:15346
  via adapter      via fc_nport      Con   DID   Lun
-   kgpsaa0.0.0.2.1 5000-1fe1-0000-0d14 Yes  210013 50
-   kgpsaa0.0.0.2.1 5000-1fe1-0000-0d11 No   210213 50
-   kgpsab0.0.0.3.1 5000-1fe1-0000-0d13 Yes  210013 50
-   kgpsab0.0.0.3.1 5000-1fe1-0000-0d12 No   210213 50

[2] UDID:51 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0027 (ev:none)
- current_unit:31539 current_col: 0 default_unit:31539
  via adapter      via fc_nport      Con   DID   Lun
-   kgpsaa0.0.0.2.1 5000-1fe1-0000-0d14 Yes  210013 51
-   kgpsaa0.0.0.2.1 5000-1fe1-0000-0d11 No   210213 51
-   kgpsab0.0.0.3.1 5000-1fe1-0000-0d13 Yes  210013 51
-   kgpsab0.0.0.3.1 5000-1fe1-0000-0d12 No   210213 51

[3] UDID:60 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0021 (ev:none)
- current_unit:31155 current_col: 0 default_unit:31155
  via adapter      via fc_nport      Con   DID   Lun
-   kgpsaa0.0.0.2.1 5000-1fe1-0000-0d14 Yes  210013 60
-   kgpsaa0.0.0.2.1 5000-1fe1-0000-0d11 No   210213 60
-   kgpsab0.0.0.3.1 5000-1fe1-0000-0d13 Yes  210013 60
-   kgpsab0.0.0.3.1 5000-1fe1-0000-0d12 No   210213 60

[4] UDID:61 WWID:01000010:6000-1fe1-0000-0d10-0009-8090-0677-0022 (ev:none)
- current_unit:30963 current_col: 0 default_unit:30963
  via adapter      via fc_nport      Con   DID   Lun
-   kgpsaa0.0.0.2.1 5000-1fe1-0000-0d14 Yes  210013 61
-   kgpsaa0.0.0.2.1 5000-1fe1-0000-0d11 No   210213 61
-   kgpsab0.0.0.3.1 5000-1fe1-0000-0d13 Yes  210013 61
-   kgpsab0.0.0.3.1 5000-1fe1-0000-0d12 No   210213 61
```

7.6.2 FC ディスクのブートおよびダンプ用のセットアップ

ブートやダンプに使用する個々のデバイスのセットアップには、`wwidmgr` コマンドを使用します。デバイスをセットアップすると、コンソールからデバイスにアクセスするために必要な情報は不揮発性メモリに保存されます。システム構成が変更され、不揮発性情報が無効になった場合、`wwidmgr` コマンドを再実行する必要があります。

コンソールに用意されているコマンドは `wwidmgr -quickset` というシンプルなコマンドです。このコマンドは、以下のどちらの場合にも使用できます。

- デバイスを 1 つだけセットアップする場合。

- セットアップするすべてのデバイスが HSG または HSV 上の同じポートを介してアクセスする場合。

上記のどちらも自分のケースに当てはまらない場合は、『Wwidmgr Users' Manual』を参照してください。

例 7-3 は、wwidmgr-quickset コマンドの実行例です。以下の点に着目してください。

- コマンド `wwidmgr -quickset -udid 10` は、HSG デバイス識別子または HSV デバイス識別子が 10 の FC ディスクをセットアップします。
- コンソール・デバイス名は、パスに依存します。FC ディスクにアクセスするパスには、それぞれ異なる名前が割り当てられます。この例では、ホストから FC ディスクまでの 4 本のパスに応じて `wwidmgr -quickset` コマンドにより 4 つのコンソール・デバイス名が割り当てられています。
 - `dga10.1001.0.2.1`
 - `dga10.1002.0.2.1`
 - `dgb10.1003.0.3.1`
 - `dgb10.1004.0.3.1`
- 2 番目のコマンド `wwidmgr -quickset -udid 50` は、HSG 識別子または HSV 識別子が 50 の FC ディスクをセットアップします。
- `wwidmgr` コマンドによる変更内容は、次のシステム初期化で初めて有効になります。したがって、次の操作では、`initialize` コマンドを実行します。
- 初期化が済んだら、コンソール `show device` コマンドにより各 FC アダプタを表示します。コマンドには、アダプタから定義済みの各 FC ディスクまでのパスを指定します。各 FC ディスクの、パスなしの OpenVMS デバイス名が 2 列目に表示されます。

例 7-3 wwidmgr -quickset の使用

```
P00>>>wwidmgr -quickset -udid 10
```

```
Disk assignment and reachability after next initialization:
```

```
6000-1fe1-0000-0d10-0009-8090-0677-0016
via adapter:          via fc nport:      connected:
dga10.1001.0.2.1      kgpsaa0.0.0.2.1    5000-1fe1-0000-0d14    Yes
dga10.1002.0.2.1      kgpsaa0.0.0.2.1    5000-1fe1-0000-0d11    No
dgb10.1003.0.3.1      kgpsab0.0.0.3.1    5000-1fe1-0000-0d13    Yes
dgb10.1004.0.3.1      kgpsab0.0.0.3.1    5000-1fe1-0000-0d12    No
P00>>>wwidmgr -quickset -udid 50
```

```
Disk assignment and reachability after next initialization:
```

(次ページに続く)

OpenVMS Cluster ストレージ・インターコネクタとしての Fibre Channel の構成 7.6 AlphaServer コンソールによる FC の構成 (Alpha のみ)

例 7-3 (続き) widmgr -quickset の使用

```
6000-1fe1-0000-0d10-0009-8090-0677-0016
      via adapter:          via fc nport:      connected:
dga10.1001.0.2.1      kgpsaa0.0.0.2.1      5000-1fe1-0000-0d14      Yes
dga10.1002.0.2.1      kgpsaa0.0.0.2.1      5000-1fe1-0000-0d11      No
dgb10.1003.0.3.1      kgpsab0.0.0.3.1      5000-1fe1-0000-0d13      Yes
dgb10.1004.0.3.1      kgpsab0.0.0.3.1      5000-1fe1-0000-0d12      No

6000-1fe1-0000-0d10-0009-8090-0677-0026
      via adapter:          via fc nport:      connected:
dga50.1001.0.2.1      kgpsaa0.0.0.2.1      5000-1fe1-0000-0d14      Yes
dga50.1002.0.2.1      kgpsaa0.0.0.2.1      5000-1fe1-0000-0d11      No
dgb50.1003.0.3.1      kgpsab0.0.0.3.1      5000-1fe1-0000-0d13      Yes
dgb50.1004.0.3.1      kgpsab0.0.0.3.1      5000-1fe1-0000-0d12      No
P00>>>initialize
Initializing...
P00>>>show device
polling ncr0 (NCR 53C810) slot 1 , bus 0 PCI , hose 1   SCSI Bus ID 7
dka500.5.0.1.1      DKA500          RRD45  1645
polling kgpsaa0 (KGPSA-B) slot 2 , bus 0 PCI , hose 1
kgpsaa0.0.0.2.1      PGA0          WWN 1000-0000-c920-a7db
dga10.1001.0.2.1      $1$DGA10      HSG80  R024
dga50.1001.0.2.1      $1$DGA50      HSG80  R024
dga10.1002.0.2.1      $1$DGA10      HSG80  R024
dga50.1002.0.2.1      $1$DGA50      HSG80  R024
polling kgpsaa1 (KGPSA-B) slot 3 , bus 0 PCI , hose 1
kgpsaa1.0.0.3.1      PGB0          WWN 1000-0000-c920-a694
dgb10.1003.0.3.1      $1$DGA10      HSG80  R024
dgb50.1003.0.3.1      $1$DGA50      HSG80  R024
dgb10.1004.0.3.1      $1$DGA10      HSG80  R024
dgb50.1004.0.3.1      $1$DGA50      HSG80  R024
polling isp0 (QLogic ISP1020) slot 4 , bus 0 PCI , hose 1   SCSI Bus ID 15
dkb0.0.0.4.1      DKB0          RZ1CB-CS  0844
dkb100.1.0.4.1      DKB100       RZ1CB-CS  0844
polling floppy0 (FLOPPY) PCEB - XBUS hose 0
dva0.0.0.1000.0    DVA0          RX23
polling ncr1 (NCR 53C810) slot 4 , bus 0 PCI , hose 0   SCSI Bus ID 7
dkc0.0.0.4.0      DKC0          RZ29B  0007
polling tulip0 (DECchip 21040-AA) slot 3 , bus 0 PCI , hose 0
ewa0.0.0.3.0      00-00-F8-21-09-74 Auto-Sensing
```

例 7-4 は、FC システム・ディスクからブートするときのシーケンスです。以下の点に着目してください。

- ブート・デバイスは\$1\$DGA50です。ここで、ユーザは、デバイスまでの 4 本すべてのパスをbootdef_devストリングに指定しています。これにより、どれかのパスに障害が発生してもシステムを確実にブートできます。
- ブート・コマンド・ストリングの最初のパスdga50.1002.0.2.1は、現在は接続されていません(つまり、このディスクは、そのパスのHSG80に対してオンラインになっていません)。コンソールにこのことが表示され 2, 3 回再試行し、

bootdef_dev スtringの次のパスに移動します。このパスは現在接続されているのでブートが成功します。

- ブート後， OpenVMS SHOW DEVICE コマンドは， コンソール wwidmgr - quickset コマンドによってセットアップされた 2 つの FC ディスクだけでなく， wwidmgr -show wwid コマンドによって表示された 5 つすべての FC デバイスが OpenVMS によって構成されたことを確認します。 OpenVMS SHOW DEV /MULTIPATH コマンドは， OpenVMS により各ディスクまでの 4 本のパスのすべてが構成されたことを確認します。

例 7-4 FC システム・ディスクによるブート・シーケンス

```
P00>>>set bootdef_dev dga50.1002.0.2.1,dga50.1001.0.2.1,dgb50.1003.0.3.1,
dgb50.1004.0.3.1
P00>>>b
(boot dga50.1002.0.2.1 -flags 0,0)
dga50.1002.0.2.1 is not connected
dga50.1002.0.2.1 is not connected
dga50.1002.0.2.1 is not connected
dga50.1002.0.2.1 is not connected
failed to open dga50.1002.0.2.1
(boot dga50.1001.0.2.1 -flags 0,0)
block 0 of dga50.1001.0.2.1 is a valid boot block
reading 919 blocks from dga50.1001.0.2.1
bootstrap code read in
Building FRU table
base = 200000 , image_start = 0 , image_bytes = 72e00
initializing HWRPB at 2000
initializing page table at 1f2000
initializing machine state
setting affinity to the primary CPU
jumping to bootstrap code

OpenVMS (TM) Alpha Operating System , Version V7.2
...
$ SHOW DEVICE

Device          Device      Error   Volume      Free  Trans Mnt
Name            Status     Count   Label        Blocks Count Cnt
$1$DGA10:      (FCNOD1) Online      0
$1$DGA50:      (FCNOD1) Mounted    0 V72_SSB     4734189  303  1
$1$DGA51:      (FCNOD1) Online      0
$1$DGA60:      (FCNOD1) Online      0
$1$DGA61:      (FCNOD1) Online      0

$ SHOW LOGICAL SYS$SYSDEVICE
"SYS$SYSDEVICE" = "$1$DGA50:" (LNM$SYSTEM_TABLE)

$ SHO DEV/MULTI
```

(次ページに続く)

例 7-4 (続き) FC システム・ディスクによるブート・シーケンス

Device Name	Device Status	Error Count	Paths	Current path
\$1\$DGA10:	(FCNOD1) Online	0	4/ 4	PGB0.5000-1FE1-0000-0D11
\$1\$DGA50:	(FCNOD1) Mounted	0	4/ 4	PGA0.5000-1FE1-0000-0D12
\$1\$DGA51:	(FCNOD1) Online	0	4/ 4	PGA0.5000-1FE1-0000-0D13
\$1\$DGA60:	(FCNOD1) Online	0	4/ 4	PGB0.5000-1FE1-0000-0D14
\$1\$DGA61:	(FCNOD1) Online	0	4/ 4	PGB0.5000-1FE1-0000-0D11

Device Name	Device Status	Error Count	Paths	Current path
\$1\$GGA42:	Online	0	4/ 4	PGB0.5000-1FE1-0000-0D11

7.7 OpenVMS Integrity システムでの Fibre Channel ストレージ・デバイスからのブート

この節では、OpenVMS Integrity システムで Fibre Channel (FC) ストレージ・デバイスをブートする方法について説明しています。FC ストレージは、OpenVMS システムでサポートするすべてのストレージ・アレイでサポートされます。

OpenVMS Integrity Version 8.2 は、PCI-X デュアル・チャネルの 2-Gb Fibre Channel ホストベース・アダプタ (HBA) である HP A6826A をサポートします。A6826A HBA は次のソフトウェアとファームウェアをサポートします。

- EFI ドライバ Version 1.40
- RISC ファームウェア Version 3.03.001

Fibre Channel デバイスのブートでは、ポイント・ツー・ポイント・トポロジをサポートします。FC arbitrated ループ・トポロジをサポートする予定はありません。

7.7.1 ブート可能ファームウェアのインストール

OpenVMS Integrity システムの FC デバイスからブートするためには、FC HBA のフラッシュ・メモリに格納されている EFI のブート可能ファームウェアをアップデートする必要があります。

FC HBA のフラッシュ・メモリの書き込みでは、次のコンポーネントのファームウェアをアップデートします。

- EFI ドライバのファームウェア
- RISC ファームウェア
- HBA 上の FLASH ROM に常駐している NVRAM

IPF Offline Diagnostics and Utilities CD にある efiutil.efi ユーティリティを使用して、ファームウェアをアップデートします。

これらのファームウェアのアップデート手順は以下のとおりです。

1. 「IPF Offline Diagnostics and Utilities IPF」 CD を挿入します。このイメージは、以下のいずれかの方法で入手できます。

- HP Software Depot サイトのメイン・ページから無料のイメージを申し込んでください：

<http://www.software.hp.com>

Search バーで“ipf offline”と入力してください。

- HP Support and Drivers の Web サイトから.iso マスター・イメージ・ファイルをダウンロードして、CD を作成してください。

<http://welcome.hp.com/country/us/en/support.html>

Support your Product バーで、ご使用のサーバ製品を選択してください。

2. システムで検出されたすべてのアダプタのファームウェアをバッチ・モードで書き込むには、EFI Boot Manager メニューの Boot Options list から EFI Shell を選択してください。

EFI コンソールで次のコマンドを入力してください (fs0:は CD-ROM のブート可能パーティションです):

- a. fs0:\efi\hp\tools\io_cards\fc2p2g\ efiutil all info

このコマンドは、システム上のすべてのアダプタの現在の EFI ドライバと RISC ファームウェア・バージョンを出力します。

- b. fs0:\efi\hp\tools\io_cards\fc2p2g\ efiutil all efi_write

このコマンドは EFI ドライバをアップデートします。

- c. fs0:\efi\hp\tools\io_cards\fc2p2g\ efiutil all risc_fw_write

このコマンドは RISC ファームウェアをアップデートします。

- d. fs0:\efi\hp\tools\io_cards\fc2p2g\ efiutil all nvram_write

このコマンドは NVRAM をアップデートします。

- e. fs0:\> reset

このコマンドはシステムをリセットします。

3. 別の方法として、次に示すように、アダプタ ID と ROM に書き込むファームウェアのファイル名を指定して、個々のアダプタに別々に書き込むこともできます。

- a. Boot Options リストから対応する DVD-ROM のエントリをブートします。あるいは、“Boot Option Maintenance Menu”を選択して CD メディアを指定し、“Boot from a File”を選択した後、“Removable Media Boot”を選択します。

- b. CD メイン・メニューから“View I/O Cards FW Update and Configuration Utilities, and MCA Menu”を選択した後，“2Gb Fibre Channel HBA Utility”を選択します。これにより，efiutil CLI ユーティリティが起動され，システムで検出された Fibre Channel アダプタの一覧が表示されます。
- c. インデックス番号を指定して Fibre Channel アダプタを選択します。EFI ドライバ，RISC ファームウェア・ドライバ，NVRAM をアップデートします。すべてのアダプタがアップデートされるまで，この手順を繰り返します。次に例を示します。

```
efiutil.efi> adapter
Adapter index number [0]?
efiutil.efi> efi_write
efiutil.efi> risc_fw_write
efiutil.efi> nvram_write
```
- d. Quit と入力して efiutil CLI を終了します。この結果，“I/O Cards Firmware and Configuration Menu”に移ります。q と入力して Main Menu へ戻ります。Main Menu から X を選択して終了し，システムをリブートします。

7.7.2 ファームウェア・バージョンのチェック

インストールされているファームウェアのバージョンを確認する方法は，システム初期化時のコンソールからチェックする方法と，efiutil ユーティリティを使用する方法の 2 通りがあります。

- ファームウェアのバージョンは，次の例のように，システム初期化時に表示されるブート・コンソール・メッセージ内に示されます。

```
HP 2 Port 2Gb Fibre Channel Adapter (driver 1.40, firmware 3.03.001)
```

- ファームウェアのバージョン番号は，efiutil info コマンドでも表示されます。

```
fs0:\efi\hp\tools\io_cards\fc2p2g\efiutil info
Fibre Channel Card Efi Utility 1.20 (1/30/2003)
2 Fibre Channel Adapters found:
Adapter      Path                               WWN          Driver (Firmware)
A0 Acpi(000222F0,200)/Pci(1|0) 50060B00001CF2DC 1.40 (3.03.001)
A1 Acpi(000222F0,200)/Pci(1|1) 50060B00001CF2DE 1.40 (3.03.001)
```

7.7.3 FC 上のブート・デバイス・パスの構成

Fibre Channel ストレージ・デバイスでブートするための構成には，HP OpenVMS のインストレーションの後に OpenVMS I64 Boot Manager (BOOT_OPTIONS.COM) を使用することをお勧めします。以下に手順を示します。

1. OpenVMS Installation Menu からオプション 7 “Execute DCL commands and procedures”を選択して DCL プロンプトを表示させます。

2. DCL プロンプトで次のコマンドを入力して OpenVMS I64 Boot Manager ユーティリティを起動します。

```
$$$ @SYS$MANAGER:BOOT_OPTIONS
```

3. ユーティリティが起動されたら、メイン・メニューが表示されます。ブート・オプションとしてシステム・ディスクを追加するには、次の例のようにプロンプトで 1 を入力します。

```
OpenVMS I64 Boot Manager Boot Options List Management Utility
```

- (1) ADD an entry to the Boot Options list
- (2) DISPLAY the Boot Options list
- (3) REMOVE an entry from the Boot Options list
- (4) MOVE the position of an entry in the Boot Options list
- (5) VALIDATE boot options and fix them as necessary
- (6) Modify Boot Options TIMEOUT setting

- (B) Set to operate on the Boot Device Options list
- (D) Set to operate on the Dump Device Options list
- (G) Set to operate on the Debug Device Options list

- (E) EXIT from the Boot Manager utility

```
You can also enter Ctrl-Y at any time to abort this utility
```

```
Enter your choice: 1
```

注意

このユーティリティの使用中は、必要な回数だけ“^”文字をタイプすることにより、それまでに行った応答に変更することができます。DCL プロンプトに戻る場合は Ctrl/Y と入力してください。

4. ユーティリティはデバイス名の入力プロンプトを表示します。このインストレーションで使用しているシステム・ディスク・デバイスを入力してください。次の例では、マルチパス Fibre Channel デバイス \$1\$DGA1 です。

```
Enter the device name (enter "?" for a list of devices): $1$DGA1:
```

5. ユーティリティは、EFI ブート・オプション・リストにエントリを入れたい場所に関する確認プロンプトを表示します。次の例では 1 を入力します。

```
Enter the desired position number (1,2,3,,,) of the entry.  
To display the Boot Options list, enter "?" and press Return.  
Position [1]: 1
```

6. ユーティリティは、OpenVMS ブート・フラグに関するプロンプトを表示しません。デフォルトではフラグは設定されません。OpenVMS フラグ (0, 1 など) をタイプしリターン・キーを押すか、次の例のようにフラグを指定しないでリターン・キーを押します。

```
Enter the value for VMS_FLAGS in the form n,n.  
VMS_FLAGS [NONE]:
```

7. ユーティリティは、ブート・オプション・エントリに記述される説明の入力プロンプトを表示します。デフォルトでは、デバイス名が使用されます。次の例のようにさらに詳しい情報を入力することもできます。

```
Enter a short description (do not include quotation marks).
Description ["$1$DGA1"]: $1$DGA1 OpenVMS V8.2 System
efi$bcfg: $1$dgal (Boot0001) Option successfully added
efi$bcfg: $1$dgal (Boot0002) Option successfully added
efi$bcfg: $1$dgal (Boot0003) Option successfully added
```

8. ブート・オプションが正しく追加できたら、プロンプトで E と入力してユーティリティを終了します。

```
Enter your choice: E
```

9. DCL プロンプトからログアウトして Integrity システムをシャットダウンしてください。

このユーティリティの詳細については、『OpenVMS システム管理者マニュアル (上巻)』を参照してください。

7.8 OpenVMS で使用するためのストレージ・コントローラのセットアップ (Alpha のみ)

HP ストレージ・アレイ・コントローラと、そのコントローラを OpenVMS で使用できるように構成するための情報が記載してあるマニュアルは以下のとおりです。

- **HSG60/80**

『HSG80 ACS Solution Software Version 8.6 for OpenVMS Installation and Configuration Guide』 (AA-RH4BD-TE)

このマニュアルは、次の FTP サイトで入手できます。

<ftp://ftp.compaq.com/pub/products/storageworks/techdoc/raidstorage/AA-RH4BD-TE.pdf>

- **Enterprise Virtual Array**

『OpenVMS Kit V2.0 for Enterprise Virtual Array Installation and Configuration Guide』 (AA-RR03B-TE)

このマニュアルは、次の FTP サイトで入手できます。

<ftp://ftp.compaq.com/pub/products/storageworks/techdoc/enterprise/AA-RR03B-TE.pdf>

- **HP StorageWorks Modular Smart Array 1000**

MSA1000 のドキュメントは次の場所にあります。

<ftp://ftp.compaq.com/pub/products/storageworks/techdoc/msa1000/>

- HP StorageWorks XP Disk Array

HP StorageWorks XP Arrays の製品情報は次の場所にあります。

<http://h18006.www1.hp.com/storage/xparrays.html>

7.8.1 デバイス識別子を CCL 用に設定

OpenVMS では、HSG および HSV のコマンド・コンソール LUN (CCL) に対して一意のデバイス識別子を定義する必要はありませんが、一部の管理ツールでは必要な場合があります。この識別子を定義すると、SHOW DEVICE コマンドを使用して表示可能な CCL デバイスを作成できるため、OpenVMS では、常に一意のデバイス識別子を定義するように指示されます。対象のデバイスが OpenVMS で直接制御できない場合でも、SHOW DEVICE/FULL コマンドを使用して、ストレージ・コントローラへの複数のパスを表示したり、以下の例で\$1\$GGA3 に示すように、障害が発生したパスを診断できます。以下の例では、2 つのパスのうち 1 つに障害が発生しています。

```
Paver> sh dev gg /mul
Device                               Device                               Error                               Current
Name                                 Status                              Count  Paths    path
$1$GGA1:                             Online                              0    2/ 2    PGA0.5000-1FE1-0011-AF08
$1$GGA3:                             Online                              0    1/ 2    PGA0.5000-1FE1-0011-B158
$1$GGA4:                             Online                              0    2/ 2    PGA0.5000-1FE1-0015-2C58
$1$GGA5:                             Online                              0    2/ 2    PGA0.5000-1FE1-0015-22A8
$1$GGA6:                             Online                              0    2/ 2    PGA0.5000-1FE1-0015-2D18
$1$GGA7:                             Online                              0    2/ 2    PGA0.5000-1FE1-0015-2D08
$1$GGA9:                             Online                              0    2/ 2    PGA0.5000-1FE1-0007-04E3

Paver> show dev /full $1$gga9:

Device $1$GGA9:, device type Generic SCSI device, is online, shareable, device
has multiple I/O paths.

Error count                0    Operations completed                0
Owner process              ""    Owner UIC                          [SYSTEM]
Owner process ID          00000000    Dev Prot    S:RWPL,O:RWPL,G:RWPL,W:RWPL
Reference count           0    Default buffer size                0
WWID    02000008:5000-1FE1-0007-04E0

I/O paths to device                2
Path PGA0.5000-1FE1-0007-04E3    (PAVER), primary path, current path.
Error count                0    Operations completed                0
Path PGA0.5000-1FE1-0007-04E1    (PAVER).
Error count                0    Operations completed                0
```

7.8.2 デバイス識別子をディスク・デバイス用に設定

完全なデバイス名を作成するには、ディスク用のデバイス識別子を文字列\$1\$DGA の後に付けます。すべてのディスクにクラスタ内で一意のデバイス識別子を割り当てる必要があります。デバイス識別子として 0 ~ 32767 の数値を指定できますが、HSV ではデバイス識別子 0 は無効です。

7.9 共用 FC システム・ディスクを備えたクラスタの作成

OpenVMS Cluster システムにノードを構成するには、`CLUSTER_CONFIG.COM`(または `CLUSTER_CONFIG_LAN.COM`) コマンド・プロシージャを実行します。(ほとんどのプロンプトで詳細な情報が提供されるフル・バージョンと、簡易バージョンの 2 通りが用意されています。)

割り当てクラス・パラメータが FC に適用されない点を除き、共用 Fibre Channel (FC) バスは `CLUSTER_CONFIG` の目的に従って共用 SCSI バスと類似した扱いを受けます。FC ストレージ・デバイスを組み込んだ構成にパラレル SCSI ストレージ・デバイスがあれば、ノード割り当てクラス値とポート割り当てクラス値の設定規則は有効です。

新しい OpenVMS Cluster システムを構成するには、まずシングル・システムかスタンドアロン・システムでクラスタリングを有効にします。これで、そのクラスタにノードを追加できます。

例 7-5 は、`CLUSTER_CONFIG_LAN.COM` の簡易バージョンを `FCNOD1` というスタンドアロン・ノードでクラスタリングを有効にする方法を示したものです。このプロシージャの最後に、`FCNOD1` はリブートして 1 ノード・クラスタを構成します。

例 7-6 は、`FCNOD1` で `CLUSTER_CONFIG_LAN.COM` の簡易バージョンを実行して、`FCNOD2` という第 2 のノードを追加する方法を示しています。このプロシージャの最後に、クラスタは、`FCNOD2` が `FCNOD1` と同じ FC システム・ディスクからブートできる構成になります。

以下の手順は、両方の例に共通です。

1. ADD のデフォルト・オプション[1]を選択します。
2. `CLUSTER_CONFIG_LAN.COM` から共用 SCSI バスを使用するかどうかを確認するプロンプトが表示されたら、YES と応えます。この場合、SCSI には、パラレル SCSI だけでなく FC も含まれます。

FC があっても、割り当てクラス・パラメータには関係ありません。

3. ノードをサテライトにするかどうかを確認するプロンプトが表示されたら、NO と応えます。

例 7-5 スタンドアロン FC ノードにおけるクラスタリングの有効化

```
$ @CLUSTER_CONFIG_LAN BRIEF
```

```
Cluster Configuration Procedure  
Executing on an Alpha System
```

```
DECnet Phase IV is installed on this node.
```

```
The LAN, not DECnet, will be used for MOP downline loading.
```

```
This Alpha node is not currently a cluster member
```

```
MAIN MENU
```

1. ADD FCNOD1 to existing cluster, or form a new cluster.
2. MAKE a directory structure for a new root on a system disk.
3. DELETE a root from a system disk.
4. EXIT from this procedure.

```
Enter choice [1]: 1
```

```
Is the node to be a clustered node with a shared SCSI or Fibre Channel bus (Y/N)? Y
```

```
Note:
```

```
Every cluster node must have a direct connection to every other  
node in the cluster. Since FCNOD1 will be a clustered node with  
a shared SCSI or FC bus, and Memory Channel, CI, and DSSI are not  
present, the LAN will be used for cluster communication.
```

```
Enter this cluster's group number: 511
```

```
Enter this cluster's password:
```

```
Re-enter this cluster's password for verification:
```

```
Will FCNOD1 be a boot server [Y]? Y
```

```
Verifying LAN adapters in LANACP database...
```

```
Updating LANACP LAN server process volatile and permanent databases...
```

```
Note: The LANACP LAN server process will be used by FCNOD1 for boot  
serving satellites. The following LAN devices have been found:
```

```
Verifying LAN adapters in LANACP database...
```

LAN TYPE	ADAPTER NAME	SERVICE STATUS
===== Ethernet	===== EWA0	===== ENABLED

```
CAUTION: If you do not define port allocation classes later in this  
procedure for shared SCSI bus, all nodes sharing a SCSI bus  
must have the same non-zero ALLOCLASS value. If multiple  
nodes connect to a shared SCSI bus without the same allocation  
class for the bus, system booting will halt due to the error or  
IO AUTOCONFIGURE after boot will keep the bus offline.
```

(次ページに続く)

OpenVMS Cluster ストレージ・インターコネクトとしての Fibre Channel の構成 7.9 共用 FC システム・ディスクを備えたクラスタの作成

例 7-5 (続き) スタンドアロン FC ノードにおけるクラスタリングの有効化

Enter a value for FCNOD1's ALLOCLASS parameter [0]: 5

Does this cluster contain a quorum disk [N]? N

Each shared SCSI bus must have a positive allocation class value. A shared bus uses a PK adapter. A private bus may use: PK, DR, DV.

When adding a node with SCSI-based cluster communications, the shared SCSI port allocation classes may be established in SYS\$DEVICES.DAT. Otherwise, the system's disk allocation class will apply.

A private SCSI bus need not have an entry in SYS\$DEVICES.DAT. If it has an entry, its entry may assign any legitimate port allocation class value:

- n where n = a positive integer, 1 to 32767 inclusive
- 0 no port allocation class and disk allocation class does not apply
- 1 system's disk allocation class applies (system parameter ALLOCLASS)

When modifying port allocation classes, SYS\$DEVICES.DAT must be updated for all affected nodes, and then all affected nodes must be rebooted. The following dialog will update SYS\$DEVICES.DAT on FCNOD1.

There are currently no entries in SYS\$DEVICES.DAT for FCNOD1. After the next boot, any SCSI controller on FCNOD1 will use FCNOD1's disk allocation class.

Assign port allocation class to which adapter [RETURN for none]: PKA

Port allocation class for PKA0: 10

```
Port Alloclass 10 Adapter FCNOD1$PKA
```

Assign port allocation class to which adapter [RETURN for none]: PKB

Port allocation class for PKB0: 20

```
Port Alloclass 10 Adapter FCNOD1$PKA
```

```
Port Alloclass 20 Adapter FCNOD1$PKB
```

WARNING: FCNOD1 will be a voting cluster member. EXPECTED_VOTES for this and every other cluster member should be adjusted at a convenient time before a reboot. For complete instructions, check the section on configuring a cluster in the "OpenVMS Cluster Systems" manual.

Execute AUTOGEN to compute the SYSGEN parameters for your configuration and reboot FCNOD1 with the new parameters. This is necessary before FCNOD1 can become a cluster member.

Do you want to run AUTOGEN now [Y]? Y

Running AUTOGEN -- Please wait.

The system is shutting down to allow the system to boot with the generated site-specific parameters and installed images.

The system will automatically reboot after the shutdown and the upgrade will be complete.

例 7-6 共用 FC システム・ディスクがあるクラスタへノードを追加する

```
$ @CLUSTER_CONFIG_LAN BRIEF
```

```
Cluster Configuration Procedure  
Executing on an Alpha System
```

```
DECnet Phase IV is installed on this node.
```

```
The LAN , not DECnet , will be used for MOP downline loading.  
FCNOD1 is an Alpha system and currently a member of a cluster  
so the following functions can be performed:
```

MAIN MENU

1. ADD an Alpha node to the cluster.
2. REMOVE a node from the cluster.
3. CHANGE a cluster member's characteristics.
4. CREATE a duplicate system disk for FCNOD1.
5. MAKE a directory structure for a new root on a system disk.
6. DELETE a root from a system disk.
7. EXIT from this procedure.

```
Enter choice [1]: 1
```

```
This ADD function will add a new Alpha node to the cluster.
```

```
WARNING: If the node being added is a voting member , EXPECTED_VOTES for  
every cluster member must be adjusted. For complete instructions  
check the section on configuring a cluster in the "OpenVMS Cluster  
Systems" manual.
```

```
CAUTION: If this cluster is running with multiple system disks and  
common system files will be used , please , do not proceed  
unless appropriate logical names are defined for cluster  
common files in SYLOGICALS.COM. For instructions , refer to  
the "OpenVMS Cluster Systems" manual.
```

```
Is the node to be a clustered node with a shared SCSI or Fibre Channel bus (Y/N)? Y
```

```
Will the node be a satellite [Y]? N
```

```
What is the node's SCS node name? FCNOD2
```

```
What is the node's SCSYSTEMID number? 19.111
```

```
NOTE: 19.111 equates to an SCSYSTEMID of 19567
```

```
Will FCNOD2 be a boot server [Y]? Y
```

```
What is the device name for FCNOD2's system root
```

```
[default DISK$V72_SSB:]? Y
```

```
What is the name of FCNOD2's system root [SYS10]?
```

```
Creating directory tree SYS10 ...
```

```
System root SYS10 created
```

```
CAUTION: If you do not define port allocation classes later in this  
procedure for shared SCSI bus , all nodes sharing a SCSI bus  
must have the same non-zero ALLOCLASS value. If multiple  
nodes connect to a shared SCSI bus without the same allocation  
class for the bus , system booting will halt due to the error or  
IO AUTOCONFIGURE after boot will keep the bus offline.
```

(次ページに続く)

OpenVMS Cluster ストレージ・インターコネクトとしての Fibre Channel の構成 7.9 共用 FC システム・ディスクを備えたクラスタの作成

例 7-6 (続き) 共用 FC システム・ディスクがあるクラスタへノードを追加する

```
Enter a value for FCNOD2's ALLOCLASS parameter [5]:
Does this cluster contain a quorum disk [N]? N
Size of pagefile for FCNOD2 [RETURN for AUTOGEN sizing]?

    A temporary pagefile will be created until resizing by AUTOGEN. The
    default size below is arbitrary and may or may not be appropriate.

Size of temporary pagefile [10000]?
Size of swap file for FCNOD2 [RETURN for AUTOGEN sizing]?

    A temporary swap file will be created until resizing by AUTOGEN. The
    default size below is arbitrary and may or may not be appropriate.

Size of temporary swap file [8000]?
Each shared SCSI bus must have a positive allocation class value. A shared
bus uses a PK adapter. A private bus may use: PK, DR, DV.

When adding a node with SCSI-based cluster communications, the shared
SCSI port allocation classes may be established in SYS$DEVICES.DAT.
Otherwise, the system's disk allocation class will apply.

A private SCSI bus need not have an entry in SYS$DEVICES.DAT. If it has an
entry, its entry may assign any legitimate port allocation class value:

    n   where n = a positive integer, 1 to 32767 inclusive
    0   no port allocation class and disk allocation class does not apply
    -1  system's disk allocation class applies (system parameter ALLOCLASS)

When modifying port allocation classes, SYS$DEVICES.DAT must be updated
for all affected nodes, and then all affected nodes must be rebooted.
The following dialog will update SYS$DEVICES.DAT on FCNOD2.

Enter [RETURN] to continue:

$20$DKA400:<VMS$COMMON.SYSEXEC>SYS$DEVICES.DAT;1 contains port
allocation classes for FCNOD2. After the next boot, any SCSI
controller not assigned in SYS$DEVICES.DAT will use FCNOD2's
disk allocation class.

Assign port allocation class to which adapter [RETURN for none]: PKA
Port allocation class for PKA0: 11

    Port Alloclass   11   Adapter FCNOD2$PKA

Assign port allocation class to which adapter [RETURN for none]: PKB
Port allocation class for PKB0: 20

    Port Alloclass   11   Adapter FCNOD2$PKA
    Port Alloclass   20   Adapter FCNOD2$PKB

Assign port allocation class to which adapter [RETURN for none]:

    WARNING: FCNOD2 must be rebooted to make port allocation class
    specifications in SYS$DEVICES.DAT take effect.
Will a disk local only to FCNOD2 (and not accessible at this time to FCNOD1)
be used for paging and swapping (Y/N)? N
```

(次ページに続く)

例 7-6 (続き) 共有 FC システム・ディスクがあるクラスタへノードを追加する

If you specify a device other than DISK\$V72_SSB: for FCNOD2's page and swap files, this procedure will create PAGEFILE_FCNOD2.SYS and SWAPFILE_FCNOD2.SYS in the [SYSEXE] directory on the device you specify.

```
What is the device name for the page and swap files [DISK$V72_SSB:]?
%SYSGEN-I-CREATED, $20$DKA400: [SYS10.SYSEXE]PAGEFILE.SYS;1 created
%SYSGEN-I-CREATED, $20$DKA400: [SYS10.SYSEXE]SWAPFILE.SYS;1 created
The configuration procedure has completed successfully.
```

FCNOD2 has been configured to join the cluster.

The first time FCNOD2 boots, NETCONFIG.COM and AUTOGEN.COM will run automatically.

7.9.1 共有 FC ディスクでブートするためのクラスタ・ノードの追加構成 (Integrity のみ)

OpenVMS Cluster システムで共有 FC ディスクでブートするための追加ノードを構成するためには、OpenVMS I64 Boot Manager (BOOT_OPTIONS.COM) を実行する必要があります。

シングル・システムあるいはスタンドアロン・システムでクラスタ機能を有効に設定した後、次の手順で、共有 FC ディスクでブートする Integrity ノードを追加することができます。

1. ターゲット・ノードで HP OpenVMS Installation Disk をブートします。
2. OpenVMS Installation Menu から、オプション 7 “Execute DCL commands and procedures”を選択します。
3. 第 7.7.3 項の説明に従います。OpenVMS ブート・フラグを入力するよう求められたら、正しいシステム・ルートを設定していることを確認してください。

注意

OpenVMS I64 Boot Manager (BOOT_OPTIONS.COM) ユーティリティは共有 FC ディスクのマウントを必要とします。クラスタに共有 FC ディスクがマウントされていない場合は、このユーティリティが/NOWRITE オプションでマウントしようとしています。共有 FC ディスクがすでにクラスタにマウントされている場合は、ユーザの介入が必要です。このユーティリティについての詳細は、『OpenVMS システム管理者マニュアル (上巻)』を参照してください。

7.9.2 オンライン再構成

ホストで OpenVMS を実行中に FC インターコネクトは再構成できます。具体的には、次の機能があります。

- FC スイッチと HSG の追加、移動、削除
- HSG 仮想ディスク装置の追加、移動、削除
- HSG 仮想ディスク装置のデバイス識別子や LUN 値の変更
- FC ケーブルの切断と再接続。再接続の対象には、同じアダプタ、異なるアダプタ、スイッチ・ポート、HSG ポートがあります。

ほとんどの FC 再構成は、OpenVMS で自動的に認識できません。FC 再構成を安全に実行し、その初期データ構造を新しい状態に合わせて OpenVMS で確実に調整するには、次のように操作します。

1. 再構成に必要なすべてのディスクのマウントを解除します。
2. 再構成を実行します。
3. Fibre Channel に接続されている各ホストに、以下のコマンドを入力します。

```
SYSMAN> IO SCSI_PATH_VERIFY  
SYSMAN> IO AUTOCONFIGURE
```

SCSI_PATH_VERIFY コマンドの目的は、システムの IO データベース内の各 FC パスを調べて、関連付けられているデバイスが変更されていないか確認することです。デバイスの変更が見つかったら、FC パスの接続は IO データベースで解除されます。これで、IO AUTOCONFIGURE コマンドにより、パスは新しいデバイスに合わせて再構成できます。

注意

現在のリリースでは、SCSI_PATH_VERIFY コマンドは FC ディスク装置にのみ使用できます。HSG コマンド・コンソール LUN (CCL) などの汎用 FC デバイスには使用できません。(汎用 FC デバイスとは、\$1\$GGAnnnnnn のような名前を持つデバイスです。したがって、OpenVMS により、HSG の CCL を特定のデバイス識別子で構成したら、そのデバイス識別子は変更しないでください。)

7.9.3 HSG ホスト接続テーブルと構成されないデバイス

Fibre Channel ホスト・バス・アダプタが (Fibre Channel スイッチ経由で) HSG コントローラに接続された場合、その HSG コントローラは HSG 接続テーブルにエントリを作成します。各ホスト・バス・アダプタ、およびそのアダプタが接続される各 HSG ポートに対しては個々の接続が存在します (詳細は、HSG CLI コマンドの SHOW CONNECTIONS を参照)。

HSG 接続が存在する場合、『HSG Array Controller ACS Configuration and CLI Reference Guide』に説明されているコマンドを使用してパラメータを修正することができます。接続は変更可能であるため、ホスト・バス・アダプタの接続を外した際に HSG はテーブルから接続情報を削除しません。その代わりに、ユーザが接続を行った時に CLI コマンドで明示的に接続を削除する必要があります。

HSG コントローラは、限定された数の接続をサポートします。ACS V8.5 は最大 64 の接続が可能で、ACS V8.4 は最大 32 の接続が可能です。接続制限は、シングル冗長コントローラ、デュアル冗長コントローラのどちらでも同じです。接続数が最大数に達すると、それ以上新しい接続は行われません。この状況が発生すると、OpenVMS は、その HSG 上でディスク・デバイス、あるいはディスク・デバイスへのバスを構成しません。

この問題に対する解決策は、必要のない古い接続を削除することです。ただし、Fibre Channel ファブリックが大きく、アクティブな接続の数が HSG の限界を超えている場合は、接続数を減らすためにファブリックを再構成するか、あるいは FC スイッチ・ゾーニングを使用していくつかのアダプタをいくつかの HSG ポートから隠す必要があります。

7.10 I/O パフォーマンス向上のための割り込みコアレス機能 (Alpha のみ)

OpenVMS Alpha バージョン 7.3-1 から、KGPSA ホスト・アダプタで割り込みコアレス機能がサポートされるようになりました。このサポートはデフォルトでオフになっています。割り込みコアレス機能を使用すると、ホストから確認できる割り込み数をアダプタで削減できるため、I/O 作業負荷量の多い環境でパフォーマンスを向上させることができます。この機能は KGPSA ファームウェアに実装されています。

割り込みコアレス機能の現在の設定を参照および変更するには、Fibre Channel Control Program (FCSCP) を使用します。FCSCP を使用するには CMKRNL 特権が必要です。

FCSCP で応答回数と遅延時間 (ミリ秒単位) を指定すると、アダプタでは、指定した応答回数になるまで、または指定した時間が経過するまで (先に指定した値になった方)、割り込みを遅らせます。

割り込みコアレス機能によって、同期 I/O を実行しているアプリケーションのパフォーマンスが低下する場合があります。特定の KGPSA で他の I/O が処理されていない場合は、割り込みコアレス機能を有効にすると、単一の書き込みに対し、平均して 900 マイクロ秒遅延します (指定した応答間隔によっては、さらに長くなる場合があります)。

割り込みコアレス機能は KGPSA ごとに設定します。割り込みコアレス機能を有効にするには、特定の KGPSA で 1 秒につき 2000 以上の I/O が処理されている必要があります。

このコマンドの形式は次のとおりです。

```
RUN SYS$ETC:FC$CP FGx enable-value [delay][response-count]
```

この形式では、次のように定義されます。

- FGxのxには A ~ Z を指定可能。
- enable-valueは、ビット・マスクを表す。ビット 1 は応答コアレス機能を制御し、ビット 0 は割り込みコアレス機能を制御する。指定できる 10 進数値は次のとおり。

1=割り込みコアレス機能

2=応答コアレス機能

3=割り込みコアレス機能と応答コアレス機能

- delay (ミリ秒単位) には 0 ~ 255 の 10 進数値を指定可能。
- response-countには 0 ~ 63 の 10 進数値を指定可能。
- 負の値は、パラメータに変更がないことを表す。
- 戻り値は、変更後の現在の値を表す。

OpenVMS では、FC\$CP コマンドを次のように設定することをお勧めします。

```
$ RUN SYS$ETC:FC$CP FGx 2 1 8
```

7.11 Fast Path の使用

Fast Path のサポートは OpenVMS Alpha Version 7.3 で Fibre Channel に対して導入されており、デフォルトで有効になっています。シンメトリック・マルチプロセッサ・システム (SMP) で使用するように設計されています。Fast Path が有効な場合、I/O 完了処理はプライマリ CPU だけでなく SMP システムのすべてのプロセッサで発生することが可能になります。Fast Path は、SMP システムにおける潜在的な I/O スループットを実質的に向上させ、プライマリ CPU が飽和状態になるのを防ぎます。

Fast Path はまた、さらなる最適化、リソースの事前割り当て、メインライン・コードに対する最適化されたコード・パスを提供する次のような機能もサポートします。

- SCS/IOLOCK8 スピンロックのコンテンションの低減。LAN ドライバは、可能であれば LAN ポート固有のスピンロックを使用して同期を行います。

- プライマリ CPU のオフロード。I/O 処理がセカンダリ CPU で開始し完了できるように LAN ドライバをセカンダリ CPU に割り当てます。この機能によりプライマリ CPU を解放し、プロセッサ間のキャッシュのコンテンションを低減できます。

Fast Path システム・サービスを使用して、Fast Path をプログラムで管理することができます。また、DCL コマンドを使用して Fast Path を管理することも、システム・パラメータ FAST_PATH および FAST_PATH_PORTS を使用して管理することもできます。Fast Path の使用方法の詳細は、『HP OpenVMS I/O User's Reference Guide』を参照してください。

7.12 FIBRE_SCAN ユーティリティによるデバイス情報の表示

FIBRE_SCAN.EXE ユーティリティは、システムの Fibre Channel に接続されたすべてのストレージ・デバイス (構成済みデバイスおよび未構成デバイスの両方) についての情報を表示します。表示される情報は、Fibre Channel ターゲット値および LUN 値、ベンダー ID および製品 ID、デバイス・タイプ、ポート WWID およびデバイス WWID、シリアル番号、ファームウェア・リビジョン・レベル、ポート・ログイン状態などのデータです。プログラムは主にディスクとテープ・デバイスについて表示し、コントローラやその他の汎用 (\$n\$GGAn) デバイスについても情報を表示します。

FIBRE_SCAN は次の 2 つのモードで起動できます。

```
$ MCR SYS$SETC:FIBRE_SCAN      ! Scans all ports on the Fibre Channel.  
$ MCR SYS$SETC:FIBRE_SCAN PGx ! Scans only port x on the Fibre Channel.
```

FIBRE_SCAN は CMKRNL および LOG_IO 特権を必要とします。

FIBRE_SCAN の出力をファイルに落すには、FIBRE_SCAN を実行する前に次のようなコマンドを実行してください。

```
$ DEFINE/USER SYS$OUTPUT xxx.log
```

FIBRE_SCAN は表示専用のユーティリティで、デバイス・ドライバのローディング機能や Fibre Channel 上のデバイスの構成機能はありません。デバイスの構成には SYSMAN IO AUTOCONFIGURE コマンドを使用してください。

OpenVMS Version 8.2-1 で System Dump Analyzer (SDA) コマンド FC PERFORMANCE のサポートが追加されています。FC PERFORMANCE は、DGA デバイスの I/O 性能特性を表示するのに使用されます。

FC PERFORMANCE は、OpenVMS Alpha Version 8.2 および OpenVMS Integrity Version 8.2 でも利用できます。さらに、OpenVMS Alpha Version 7.3-2 用の FIBRE_SCSI_V400 Fibre Channel パッチキットでも提供されています。これらのサ

ポートされるバージョンの Fibre Channel ドライバは、構成されているディスク毎に性能データを保持します。

この SDA コマンドは、指定した DGA デバイスあるいはシステムに構成されているすべての DGA デバイスの I/O 性能特性を表示することができます。デバイス名の指定を省略すると、データがゼロでないすべての DGA ディスクの I/O 性能特性が表示されます。そのデバイスが現在マウントされているかどうかは関係ありません。FC PERFORMANCE 配列は、LUN ごとの I/O 遅延、I/O サイズ、I/O 方向 (read/write) のカウント数を保持します。

I/O 遅延は、要求をホスト FC アダプタに送信してから中断が完了するまでの時間が計測されます。最新のディスク・ドライブでは、平均して 5-10 ms 程度の遅延が発生します。ディスク・コントローラ (HSG/EVA/MSA/XP) およびディスク・ドライブに置かれたキャッシュは、時折より低いアクセス時間を示す場合があります。

デフォルトでは、FC PERFORMANCE コマンドは /SYSTIME 修飾子を使用して遅延を計測します。/SYSTIME 修飾子は、ミリ秒単位で更新される EXESGQ_ SYSTIME を使用します。1 ms 以内に I/O が完了した場合は、時間ゼロで完了したように表示されます。/SYSTIME を使用した場合、1 ms 以内に完了した I/O 操作は <2us のカラムに表示されます。ここで us はミリ秒を意味します。

より正確な値が必要な場合は、/RSCC 修飾子を指定して System Cycle Counter タイマを使用することができます。このタイマについてを含め、コマンド修飾子については表 7-2 を参照してください。

表 7-2 FC PERFORMANCE コマンド修飾子

修飾子	説明
[device-name]	性能特性を表示するデバイス。DGA デバイス名を 1 つだけ指定できます。デバイス名の指定を省略すると、データがゼロのものを除き、システムに構成されている各 DGA デバイスの性能データが表示されます。現在マウントされていない DGA デバイスの性能データも表示されます。
[/RSCC /SYSTIME]	2 つの時間修飾子を使用できます。/RSCC 修飾子は PAL コール Read System Cycle Counter を使用し、CPU サイクル・カウンタを取得します。この修飾子を使用するとより正確度の高い値が取得できますが、1 つの I/O 操作ごとに 2 つの PAL コールを使用することになります。1 ms 未満の粒度が必要な場合はこの修飾子を使用します。/SYSTIME 修飾子は、ミリ秒単位で更新される OpenVMS システム時間を使用します。デフォルトは /SYSTIME です。
[/COMPRESS]	ゼロだけのカラムを画面に表示しません。
[/CSV]	値をコンマで区切った CSV 形式でファイルに出力します。CSV は Microsoft® Excel で読み込んだりグラフ表示することができます。

(次ページに続く)

表 7-2 (続き) FC PERFORMANCE コマンド修飾子

修飾子	説明
[/CLEAR]	性能配列をクリアします。この修飾子は、システム I/O 性能のテストを行う場合に便利です。新たにテストを開始する前に性能を計測しているデバイスの性能配列をクリアし、テストが完了した後にそのテスト結果を取り出すことができます。

以下に示すのは、SDA コマンド FC PERFORMANCE に /COMPRESS 修飾子を指定して取得した \$1\$DGA4321 デバイスの read および write 出力の例です。

この出力例の LBC (ヘッダの最初のカラム) は、論理ブロック数を示しています。LBC カラムの各列の値は 2 の累乗になっています。個々の列には、LBC の次の列に至るまでのすべてのカウント数が含まれています。たとえば LBC 8 の列には、LBC 8 から LCB 15 までのすべての I/O のカウント数が含まれています。256 よりも大きな LBC の I/O はこのマトリックには表示されませんが、マトリックの上に表示される "total blocks" の値には含まれています。

各カラムのカウント数は、カラムヘッダに示されている時間までに完了した I/O の数を示しています。たとえば <2ms カラムは、1 ms よりも長く 2 ms よりも短い時間で完了した I/O を示しています。同様に、<4ms カラムの I/O は、2 ms よりも長く 4 ms よりも短い時間で完了した I/O を示しています。<2us カラムには例外があり、/SYSTIME 修飾子を指定した場合は、1 ミリ秒より短時間で完了したすべての I/O は <2us に含まれます。

<2us、<2ms、<4ms などのカラムヘッダが表示されていますが、/COMPRESS 修飾子が指定されている場合、カラムヘッダに対応する値が全くなければそのカラムは表示されません。デフォルトの /SYSTIME 修飾子の代わりに /RSCC 修飾子が指定されている場合、<4us、<8us、<16us および <256us に対して追加のヘッダが表示されません。

```
SDA> FC PERFORMANCE $1$DGA4321/COMPRESS
Fibre Channel Disk Performance Data
-----
$1$dga4321 (write)
Using EXE$GQ_SYSTIME to calculate the I/O time
    accumulated write time = 2907297312us
    writes = 266709
    total blocks = 1432966

I/O rate is less than 1 mb/sec

LBC  <2us  <2ms  <4ms  <8ms  <16ms  <32ms  <64ms  <128ms  <256ms  <512ms  <1s
===  =====
1   46106 20630 12396 13605 13856 15334 14675 8101   777    8    -   145488
```

OpenVMS Cluster ストレージ・インターコネクタとしての Fibre Channel の構成
 7.12 FIBRE_SCAN ユーティリティによるデバイス情報の表示

2	52	21	8	9	5	5	6	1	2	-	-	109
4	40310	13166	3241	3545	3423	3116	2351	977	88	-	-	70217
8	2213	1355	360	264	205	225	164	82	5	-	-	4873
16	16202	6897	3283	3553	3184	2863	2323	1012	108	-	1	39426
32	678	310	36	39	47	44	33	27	6	-	-	1220
64	105	97	18	26	41	43	42	24	7	-	-	403
128	592	3642	555	60	43	31	23	9	2	-	-	4957
256	-	9	7	-	-	-	-	-	-	-	-	16

106258 46127 19904 21101 20804 21661 19617 10233 995 8 1 266709

Fibre Channel Disk Performance Data

\$1\$dga4321 (read)

Using EXE\$GQ_SYSTIME to calculate the I/O time

accumulated read time = 1241806687us

reads = 358490

total blocks = 1110830

I/O rate is less than 1 mb/sec

LBC <2us <2ms <4ms <8ms <16ms <32ms <64ms <128ms <256ms <512ms <2s
 === =====

1	46620	12755	6587	7767	3758	2643	1133	198	5	-	-	81466
2	574	134	66	158	82	20	21	4	1	-	-	1060
4	162060	35896	20059	18677	15851	11298	5527	1300	25	2	1	270696
8	355	79	46	97	59	36	28	10	-	-	-	710
16	241	103	32	150	77	24	13	1	-	-	-	641
32	916	355	76	302	316	61	25	10	-	-	-	2061
64	725	380	64	248	140	17	10	3	-	-	-	1587
128	13	22	13	36	21	6	-	-	-	-	-	111
256	10	41	28	15	49	13	2	-	-	-	-	158

211514 49765 26971 27450 20353 14118 6759 1526 31 2 1 358490

SDA>

可用性を目的とした OpenVMS Cluster の構成

可用性 (Availability) とは、コンピューティング・システムがアプリケーション・サービスを提供する時間の比率です。OpenVMS Cluster のさまざまな特長を活かせば、ディザスタ・トレランスをはじめ、さまざまなレベルの可用性に OpenVMS Cluster システムを構成できます。

この章では、高い可用性を備えた OpenVMS Cluster システムを構築するための方法や最適な構成例を紹介します。これらの方法や構成例は、可用性の要件ごとの設定やトレードオフの選択に役立ててください。

8.1 可用性の要件

OpenVMS Cluster システムは要求に応じてさまざまなレベルの可用性に構成できます。表 8-1 は、そのような可用性レベルの主な分類を示したものです。

表 8-1 可用性の要件

可用性の要件	説明
一般の用途	システムやアプリケーションが利用できなくてもほとんど、あるいはまったく影響なしで待機できるビジネス向け。
24 x 365	重要時間帯のみならず年間を通してほとんどの業務時間帯は、途切れることのないコンピューティング・サービスが必要なビジネス向け。最小限のダウン・タイムはあっても構わない。
ディザスタ・トレラント	可用性の要件が厳しいビジネス向け。このようなビジネスでは、地震、洪水、停電などに対する対策が必要となる。

8.2 OpenVMS Cluster による可用性の提供方法

OpenVMS Cluster システムでは、以下の方式により可用性を強化しています。

- 複数のシステムでリソース・アクセスを共用できる高度な統合環境
- 主要ハードウェア構成要素の冗長性
- ハードウェア構成要素間のフェールオーバーのソフトウェア・サポート
- 高い可用性をサポートするソフトウェア製品

8.2.1 ストレージ領域までの共用アクセス

OpenVMS Cluster 環境では、複数のシステム上のユーザとアプリケーションが、ストレージ・デバイスとファイルを透過的に共用できます。システムのどれかをシャット・ダウンしてもユーザは共用ファイルとデバイスに引き続きアクセスできます。ストレージ・デバイスの共用方法には、以下の 2 通りがあります。

- 直接アクセス

ディスク・ストレージ・サブシステムとテープ・ストレージ・サブシステムを、ノードではなくインターコネクต์に接続します。これにより、インターコネクต์に接続されたすべてのノードがストレージ・システムまでのアクセスを共用できます。システムにシャットダウンや障害が発生しても他のシステムはストレージ領域にアクセスできます。

- 他ノードのサービスによるアクセス

ノードに接続されたストレージ・デバイスは、OpenVMS Cluster 内の他のノードのサービスを受けることができます。MSCP サーバ・ソフトウェアと TMSCP サーバ・ソフトウェアを利用すれば、すべての OpenVMS Cluster メンバがローカル・デバイスにアクセスできます。ただし、サービス元のノードにシャットダウンや障害が発生すると、他のノードにおけるストレージのアクセス機能に影響が出ます。

8.2.2 構成要素の冗長性

OpenVMS Cluster システムでは、以下の構成要素をはじめ、多くの構成要素に冗長性を設定できます。

- システム
- インターコネクต์
- アダプタ
- ストレージ・デバイスとデータ

冗長構成要素を利用すれば、構成要素のどれかに障害が発生しても、ユーザやアプリケーションは別の構成要素を利用できます。

8.2.3 フェールオーバの仕組み

OpenVMS Cluster システムには、OpenVMS Cluster のどこかで障害が発生しても、それを回復できるフェールオーバのメカニズムがあります。表 8-2 は、これらのメカニズムとそれによる回復レベルをまとめたものです。

表 8-2 フェールオーバーのメカニズム

メカニズム	フェールオーバー時の動作	回復方法
DECnet-Plus クラスタ・エイリアス	ノードに障害が発生すると、OpenVMS Cluster ソフトウェアは新しい着信接続を他のノードに分散します。	手動。障害の発生したノードにログインするユーザは、残りのノードに再接続できます。 この方式に合わせて構成したアプリケーションでは自動処理。その場合、アプリケーションはクラスタ・エイリアス・ノード名で再接続し、その接続は残りのノードのどれかに直結されます。
I/O パス	ストレージ・デバイスまでの冗長パスにより、パスのどれかに障害が発生しても OpenVMS Cluster ソフトウェアは、機能している作業パスがあれば、そこにフェールオーバーします。	他の作業パスがあれば透過的
インターコネク	冗長または複合インターコネクにより、OpenVMS Cluster ソフトウェアは最速の作業パスを利用して他の OpenVMS Cluster メンバに接続します。インターコネク・パスのどれかに障害が発生しても、機能している作業パスがあれば、そこにフェールオーバーします。	透過的
ブート・サーバとディスク・サーバ	少なくとも 2 個のノードをブート・サーバとディスク・サーバに構成すると、サーバのどれかがシャット・ダウンしたり障害を起しても、引き続きサテライトからブートしたり、ディスクにアクセスできます。 ブート・サーバに障害が発生しても、MSCP サービス対象のディスクにアクセスできる代替パスがあれば影響はありません。	自動
ターミナル・サーバと LAT ソフトウェア	ターミナルとプリンタをターミナル・サーバに接続します。ノードに障害が発生すると、LAT ソフトウェアは残ったノードのどれかに自動的に接続します。また、ユーザ・プロセスが LAT ターミナル・セッションから切り離されても、LAT セッションに再接続するとき、切り離されたセッションに LAT ソフトウェアが自動的に再接続します。	手動。ターミナル・ユーザが、障害の発生したノードにログインするには、残ったノードにログインしてアプリケーションを再起動する必要があります。
汎用バッチとプリント・キュー	汎用キューをセットアップすれば、1 つ以上のノード上の (処理が発生した) 実行キューにジョブをフィードできます。ノードのどれかに障害が発生すると、汎用キューは、残ったノード上の実行キューにジョブを発行します。また、/RESTART 修飾子で発行されたバッチ・ジョブは、残ったノードのどれかで自動的に再起動します。	デイスバッチを待機中のジョブには透過的 障害の発生したノードで実行中のジョブには自動的、または手動
自動起動バッチとプリント・キュー	フェールオーバー・リストで、実行キューを自動起動キューとしてセットアップすれば、可用性をフルに活用できます。ノードに障害が発生すると、自動起動キューとそのジョブがフェールオーバー・リストの次の論理ノードに自動的にフェールオーバーし、別のノードで処理を続行します。自動起動キューは、ターミナル・サーバに接続されたプリンタに送信されたプリント・キューに特に適しています。	透過的

関連項目：クラスタ・エイリアス、汎用キュー、自動起動キューについては、『OpenVMS Cluster システム』を参照してください。

8.2.4 関連ソフトウェア製品

表 8-3 は、HP が可用性を強化するために提供しているさまざまな関連 OpenVMS Cluster ソフトウェア製品です。

表 8-3 可用性を強化する製品

製品	説明
Availability Manager	複数ノードのデータを同時に収集、解析し、すべての出力を一元された DECwindows 表示に送信します。解析時に可用性の問題を検出し、対処方法を提示します。
RTR	分散環境においてスケラビリティと位置透過性を備えた連続的で耐障害性のあるトランザクション転送サービスを提供します。実行中のトランザクションは 2 フェーズのコミット・プロトコルにより保証され、データベースは全体に分配することが可能であるとともに、性能の改善のために分割することができます。
Volume Shadowing for OpenVMS	OpenVMS Cluster システム上の任意のディスクを、その OpenVMS Cluster システム内にある他の同じサイズ (同じ物理ブロック数) のディスクとで冗長構成にします。

8.3 高い可用性を備えた OpenVMS Cluster 構成の実現手法

選択するハードウェアとその構成方法は、OpenVMS Cluster システムの可用性に重要な影響があります。この節では、可用性強化を目的とした OpenVMS Cluster 構成設計の基本手法を説明します。

8.3.1 可用性の強化手法

表 8-4 は、高可用性 OpenVMS Cluster の構成方法をまとめたものです。これらの方法は、重要度の高い順に並んでいます。以下の方法のうち多くは、この章で紹介する最適構成の例で取り上げています。

表 8-4 可用性の強化手法

強化手法	説明
単一点障害 (Single point of failure (単一機器の障害がシステム全体の障害になる)) の要因を排除	構成要素に冗長性を与え、ある構成要素に障害が発生しても他の構成要素で対応します。
システム・ディスクのシャドウ化	システム・ディスクは、ノード操作にきわめて重要です。Volume Shadowing for OpenVMS で、システム・ディスクを冗長化できます。
基本データ・ディスクのシャドウ化	Volume Shadowing for OpenVMS でデータ・ディスクを冗長化してデータの可用性を強化します。

(次ページに続く)

表 8-4 (続き) 可用性の強化手法

強化手法	説明
ストレージ領域までの共用直接アクセスを提供	可能な場合、すべてのノードにストレージ領域の共用直接アクセスを与えます。これにより、ストレージ領域へのアクセス時に MSCP サーバ・ノードに対する依存度を緩和できます。
環境上のリスクを最小化	以下の手順により、環境上の問題のリスクを最小限に抑えます。 <ul style="list-style-type: none">一時的な停電に備え、通常電源を代行するシステムを発電機が無停電電力システム (UPS) に用意する。1 つのユニットに障害が発生してもシステム機器の使用に支障が発生しないよう、補助エアコン装置を設置する。
最低 3 ノードを構成	OpenVMS Cluster ノードには、操作を続行するための定足数が定められています。最適構成では、最低 3 ノードを使用し、どれか 1 ノードが使用できなくなっても、残り 2 ノードで定足数を満たして処理を続行します。 関連項目：定足数の実現手法の詳細については、第 10.5 節と『OpenVMS Cluster システム』を参照してください。
予備キャパシティを構成	構成要素ごとに、キャパシティの処理に必要な数に加え、少なくとも 1 つは予備の構成要素を構成します。構成要素の使用パーセンテージは、キャパシティの 80%未満になるようにします。重要な構成要素については、リソースの使用パーセンテージがキャパシティの 80%を十分に下まわるように構成します。これは構成要素のどれかに障害が発生しても残りの構成要素に余裕をもって作業負荷を分散するためです。
予備構成要素をスタンバイ	構成要素ごとに、予備構成要素を 1 つまたは 2 つ用意しておき、構成要素の障害発生に備えます。予備構成要素は定期的にチェックして正しく機能するか確認します。予備構成要素を 1 つまたは 2 つを超えて用意しても複雑になるだけでなく、必要なときに予備構成要素が正しく機能しなくなる可能性も高くなります。
共通環境ノードの使用	似たようなサイズ、パフォーマンスのノードを構成してフェールオーバー時のキャパシティの過剰負荷を避けます。サイズの大きなノードに障害が発生すると、サイズの小さいノードでは転化された負荷に対応できないことがあります。それによるボトルネックで OpenVMS Cluster のパフォーマンスが低下することがあります。
信頼性の高いハードウェアを使用	ハードウェア・デバイスの障害の可能性を考慮します。MTBF (平均故障間隔) に関する製品の記述を確認してください。一般に、テクノロジーが新しいほど信頼性が高くなります。

8.4 高い可用性を備えた OpenVMS Cluster の保守

可用性の強化プロセスは現在進行中です。OpenVMS Cluster システムの管理方法は、その構成方法と同様に重要です。この節では、OpenVMS Cluster 構成の可用性の保守の方法を紹介します。

8.4.1 可用性の保守の手法

初期構成のセットアップ後、表 8-5 の方法に従って、OpenVMS Cluster システムの可用性を保守します。

可用性を目的とした OpenVMS Cluster の構成

8.4 高い可用性を備えた OpenVMS Cluster の保守

表 8-5 可用性の保守手法

方法	説明
フェールオーバー方法の計画	<p>OpenVMS Cluster システムでは、ハードウェア構成要素間のフェールオーバーをソフトウェア・サポートします。どのようなフェールオーバー機能があり、またニーズに応じたカスタマイズができるのはどれなのか、よく確認してください。障害時に回復すべき構成要素はどれか、また、フェールオーバーで生じた余分な作業負荷に構成要素が対応できるか確認します。</p> <p>関連項目：表 8-2 は、OpenVMS Cluster フェールオーバー機能と、それによる回復レベルをまとめたものです。</p>
分散アプリケーションのコーディング	<p>OpenVMS Cluster システムの複数のノードでアプリケーションを同時に実行できるようコーディングします。ノードに障害が発生しても、OpenVMS Cluster システムの残ったメンバは利用でき、ディスク、テープ、プリンタ、その他必要な周辺機器のアクセスを維持します。</p>
変更を最小限にする	<p>実行中のノードに対するハードウェアやソフトウェアの変更は、実装する前に、その必要性をよく検討してください。変更する場合は、生産環境で実施する前に、重要度の低い環境でテストします。</p>
サイズと複雑さを抑制	<p>冗長性を実現できたら、構成の要素数と複雑さを抑制してください。構成が簡潔であれば、ハードウェアやソフトウェアによるエラーだけでなく、ユーザやオペレータによるエラーも減ります。</p>
全ノードに同一ポーリング・タイムを設定	<p>OpenVMS Cluster システムの維持管理に使用するポーリング・タイムは、決められたシステム・パラメータで制御します。これらのシステム・パラメータの値は、すべての OpenVMS Cluster メンバ・ノードで同じ値に設定してください。</p> <p>関連項目：これらのシステム・パラメータの詳細については、『OpenVMS Cluster システム』を参照してください。</p>
管理のための事前対策	<p>システム管理者の熟練度は高いほど適しています。アクセス権限はそれを必要とするユーザやオペレータだけに設定してください。OpenVMS Cluster システムの管理とセキュリティ保持のポリシーは厳密に定義してください。</p>
AUTOGEN の有効活用	<p>通常の AUTOGEN フィードバックでは、システム・パラメータの設定に影響を与えるリソースの使用法が解析できます。</p>
単独サーバやディスクに対する依存度の抑制	<p>複数のシステムやディスクにデータを分散すれば、システムやディスクによる単一点障害を防げます。</p>
バックアップ・ストラテジの実装	<p>定期的に、そして頻繁にバックアップを実行すれば、障害後にデータを確実に回復できます。この表にある方法のどれも確実なバックアップにはなりません。</p>

8.5 LAN OpenVMS Cluster における可用性

図 8-1 は、スモール・キャパシティで可用性の高い LAN OpenVMS Cluster システムの最適な構成です。図 8-1 の後の項で、以下の構成を検討します。

- 構成要素の解析
- 長所と欠点
- 可用性の主な強化手法の実装

8.5.1 構成要素

図 8-1 に示す LAN OpenVMS Cluster 構成の構成要素は以下のとおりです。

構成要素	説明
1	<p>2 本の Ethernet インターコネクト。ネットワーク・キャパシティを強化するには、Gigabit Ethernet あるいは 10 Gigabit Ethernet を使用します。</p> <p>解説：冗長化のためには、最低 2 本の LAN インターコネクトを使用し、全ノードを全 LAN インターコネクトに接続します。</p> <p>インターコネクトが 1 本だけでは、単一点障害 (Single point of failure (単一機器の障害がシステム全体の障害になる)) の要因になります。</p>
2	<p>3 個から 8 個の Ethernet 対応の OpenVMS ノード</p> <p>各ノードには、別のディスクに依存しないよう専用のシステム・ディスクが備えられます。</p> <p>解説：少なくとも 3 個のノードを使用し、定足数を維持します。8 つのシステム・ディスクを管理するのは大変なので、ノード数は 8 個未満にしてください。</p> <p>代替法 1: サテライト・ノードが必要な場合は、1 個または 2 個のノードをブート・サーバとして構成します。ただし、サテライト・ノードの可用性は、サーバ・ノードの可用性に依存します。</p> <p>代替法 2: 第 8.9 節にあるように、8 ノードを超える場合、LAN OpenVMS Cluster 構成を使用します。</p>
3	<p>システム・ディスク</p> <p>ブート順序の制約上、システム・ディスクは、通常は LAN OpenVMS Cluster ではシャドウイングしません。</p> <p>代替法 1: 2 つのローカル・コントローラ間でシステム・ディスクをシャドウイングします。</p> <p>代替法 2: 2 個のノード間でシステム・ディスクをシャドウイングします。2 番目のノードでは、ディスクは非システム・ディスクとしてマウントされます。</p> <p>関連項目：ブート順序とサテライトの依存内容については、第 10.2.4 項を参照してください。</p>
4	<p>重要データ・ディスク</p> <p>すべての重要データ・ディスクには、ボリューム・シャドウイングでコピーを複数作成しておきます。シャドウ・セット・メンバは、少なくとも 2 個のノードに配置して、単一点障害の要因にならないようにします。</p>

8.5.2 長所

この構成には、以下のような長所があります。

- この章で紹介するどの例よりもコストがかからない。
- サイズとパフォーマンスを拡張できる、いくらかの余裕がある。
- LAN インターコネクトにおけるノードの選択肢が最も広い。

8.5.3 短所

この構成には、以下のような短所があります。

- ストレージ領域までの共用の直接アクセスがない。ノードから共用ストレージ領域にアクセスするには、MSCP サーバが必要である。
- LAN ノード間でディスク・シャドウイングをすると、ノード・ブート時にコピーがシャドウされる。
- ブート順序に依存するため、システム・ディスクをシャドウイングしても実利的ではない。

8.5.4 可用性の主な強化手法

図 8-1 の構成では、以下の手法により可用性を強化しています。

- 単一点障害の要因を構成から排除している。
- ボリューム・シャドウイングにより、重要なデータ・ディスクの複数のコピーが別々のノードに用意される。
- 定足数を満たすため、最低 3 ノードを使用しており、ノードのどれかに障害が発生しても、OpenVMS Cluster は処理を続行できる。
- 各ノードに専用のシステム・ディスクを割り当てている。サテライトに依存しない。

8.6 マルチ LAN の構成

高い可用性を備えたマルチ LAN クラスタを構成するには、以下のガイドラインに従ってください。

- LAN セグメントをブリッジし、シングル拡張 LAN を構成する。
- フェールオーバー・サポートとして冗長 LAN セグメント・ブリッジを設定する。
- LAN と MOP のマルチキャスト・メッセージを伝達できる LAN ブリッジを構成する。
- Local Area OpenVMS Cluster Network Failure Analysis Program を利用してネットワークの可用性を監視し、管理する（詳細については、『OpenVMS Cluster システム』を参照）。
- SCS レイヤと NISCA 伝送プロトコルに関するパフォーマンス問題の診断については、『OpenVMS Cluster システム』のトラブルシューティングを参照してください。
- LAN の平均利用率を 50%未満とする。

可用性を目的とした OpenVMS Cluster の構成

8.6 マルチ LAN の構成

関連項目：拡張 LAN (ELAN) の詳細については、第 9.4.8 項を参照してください。
Alpha サテライトと Integrity サテライトの違いについては『OpenVMS Cluster システム』を参照してください。

8.6.1 MOP サーバの選択

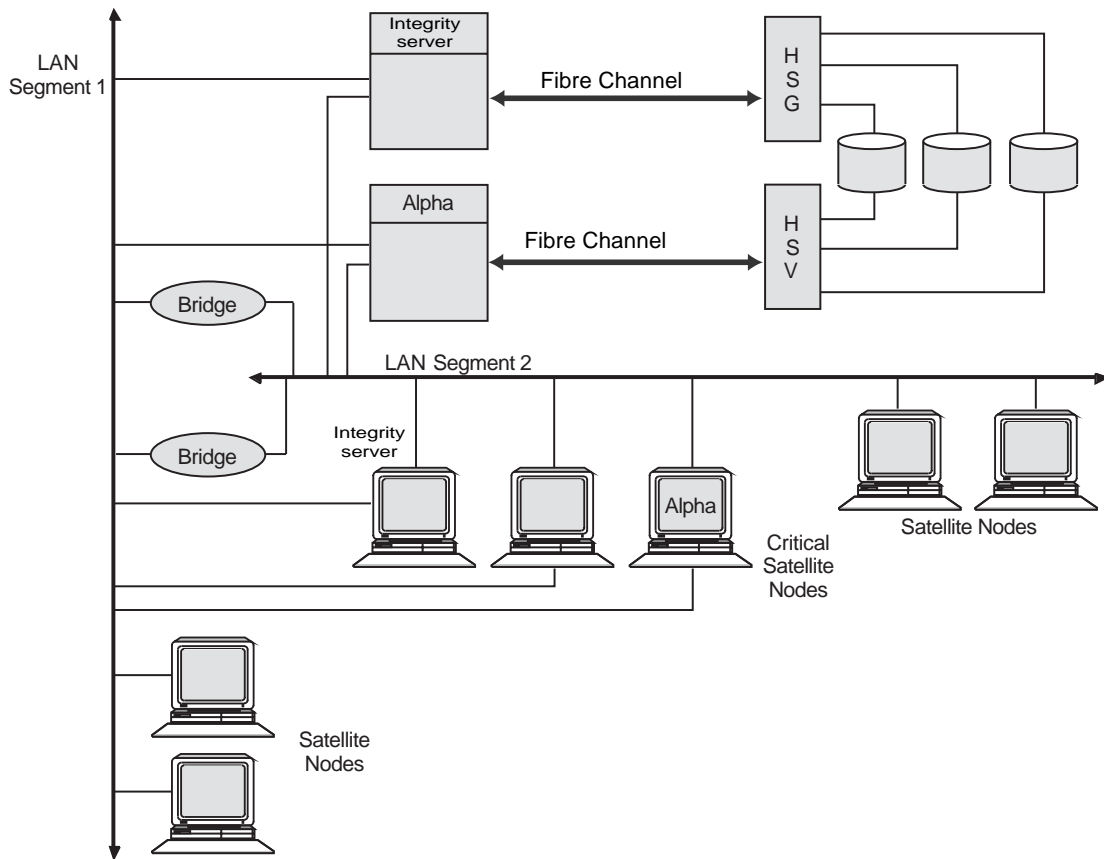
マルチ LAN アダプタとマルチ LAN セグメントを併用するときは、MOP サービスを提供する LAN セグメントまでの接続を分散してください。分散することで、ネットワーク構成要素に障害が発生しても、MOP サーバからサテライトにダウンライン・ロードをすることができます。

サテライトをブートするためのダウンライン・ロードをサポートするには、Integrity ノードと Alpha ノードに十分なキャパシティの MOP サーバが必要です。ネットワーク上で、各 MOP サーバに適した LAN 接続 (必要に応じて Alpha または VAX) を選択すれば、ネットワーク障害が発生しても MOP サービスを続行できます。

8.6.2 2 つの LAN セグメントの構成

図 8-2 は、異なる 2 つの LAN セグメントに接続した OpenVMS Cluster システムの構成例です。この構成には、Alpha ノードと Integrity ノード、サテライト、2 個のブリッジがあります。

図 8-2 2 LAN セグメントの OpenVMS Cluster 構成



VM-3828A-GE

この図では、以下の点に着目してください。

- 重要なノードはマルチ LAN セグメントに接続すると、セグメントやアダプタに障害が発生したときの可用性を強化できます。こうして追加されたネットワーク接続により増加したネットワーク帯域幅の一部は、ディスク・サーバとテープ・サーバに振り向けることができます。LAN アダプタのどれかに障害が発生しても、重要なサテライトは他の LAN アダプタでブートできます。
- 重要度の低いサテライトをシングル LAN セグメントに接続すれば、LAN セグメント間にシステムを均等に分散することでネットワーク負荷のつりあいをとることができます。このようなシステムでは、ブリッジのどれかを通じて他の LAN セグメント上のサテライトと通信します。
- DECnet と MOP サービスでは、LAN アドレスの重複を避けるために、LAN アダプタは 1 つしか使用できません。

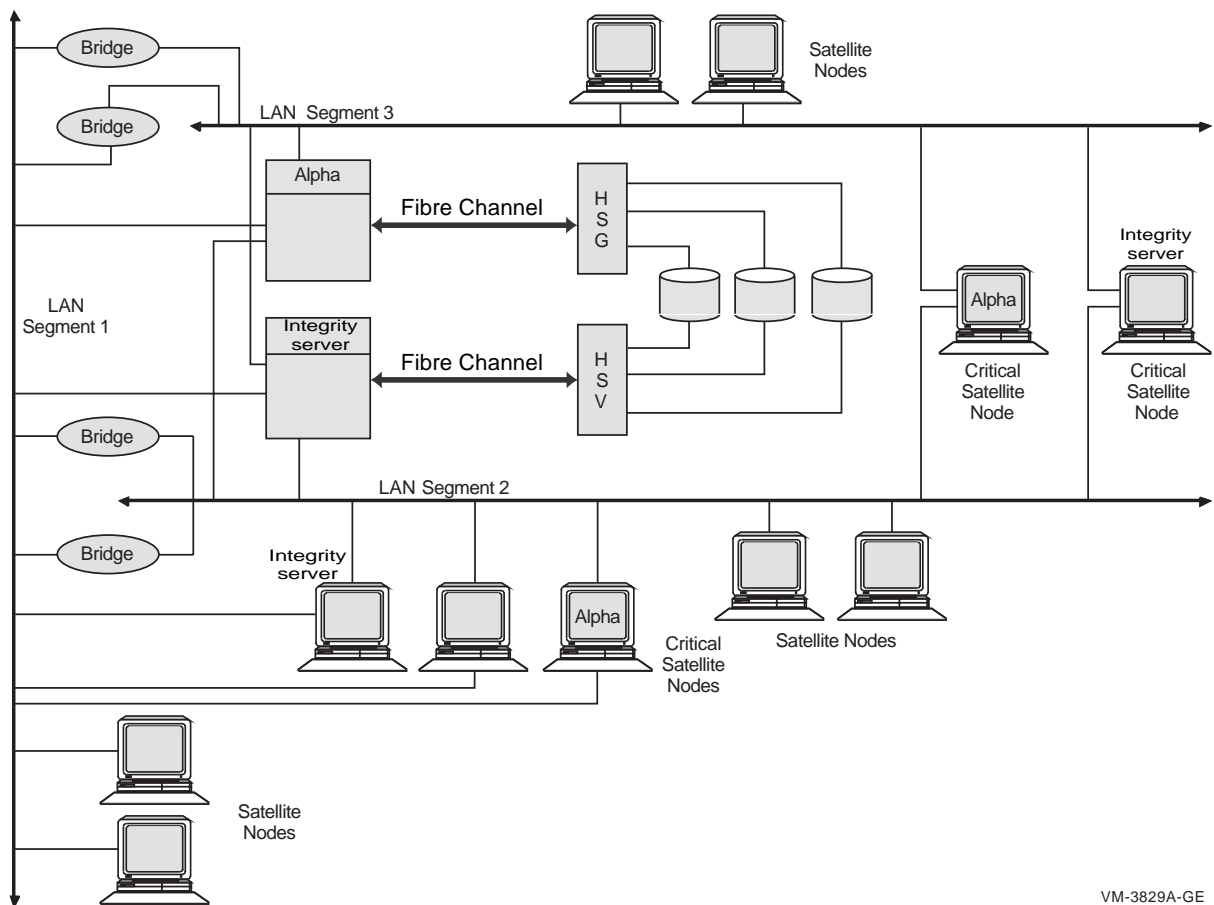
可用性を目的とした OpenVMS Cluster の構成
8.6 マルチ LAN の構成

- LAN 障害でサテライトのブートが不可能になるのを防ぐため、MOP サービスを提供する LAN アダプタ (Alpha) は、LAN セグメント間に分散します。
- 冗長 LAN ブリッジを使用すると、ブリッジによる単一点障害を防げます。

8.6.3 3つのLANセグメントの構成

図 8-3 は、異なる3つのLANセグメントに接続した OpenVMS Cluster システムの構成例です。この構成には、Alpha ノードと Integrity ノード、サテライト、複数のブリッジがあります。

図 8-3 3 LAN セグメントの OpenVMS Cluster 構成



VM-3829A-GE

この図では、以下の点に着目してください。

- ディスク・サーバとテープ・サーバを2つまたは3つのLANセグメントに結合して、可用性とI/Oスループットを強化しています。

- 重要なサテライトをマルチ LAN セグメントに接続したことも可用性を強化しています。ネットワーク構成要素のどれかに障害が発生すると、これらのサテライトは、他の LAN アダプタを使用してブートし、重要なディスク・サーバとのアクセスを維持します。
- LAN セグメント内に重要度の低いサテライトを分散することで、ネットワーク負荷のバランスを維持しています。
- 各 LAN セグメントには、MOP サーバ (Alpha) が提供されています。

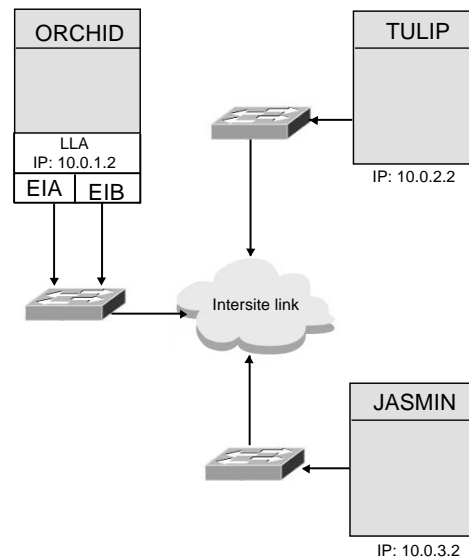
関連項目： LAN におけるブート順序とサテライト依存の詳細については、第 10.2.4 項を参照してください。LAN ブリッジ・フェールオーバの詳細については、『OpenVMS Cluster システム』を参照してください。

8.7 Cluster over IP における可用性

図 8-4 は、中程度のキャパシティで、高い可用性を備えたロジカル LAN フェールオーバ IP OpenVMS Cluster システムです。図の後の項で、以下の構成を検討します。

- 構成要素の解析
- 長所と短所
- 可用性の主な強化手法の実装

図 8-4 ロジカル LAN フェールオーバ IP OpenVMS Cluster システム



VM-drawing5

8.7.1 構成要素

図 8-4 の IP OpenVMS Cluster 構成には以下の構成要素があります。

パート	説明
1	<p>EIA と EIB の 2 つの IP インタフェースでノードに接続されている</p> <p>解説:可用性のため両方のインタフェースを使用してノードに接続します。EIA と EIB がノードに接続する 2 つの LAN インタフェースです。この 2 つの LAN インタフェースはロジカル LAN フェールオーバ・セットを作成するのに使用できます。以下のコマンドを実行して、ロジカル LAN フェールオーバ・セットを作成します。</p> <pre>\$ MC LANCP LANCP> DEFINE DEVICE LLB/ENABLE/FAILOVER=(EIA0, EIB0)</pre> <p>ロジカル LAN フェールオーバ・デバイスでは IP アドレスを構成でき、クラスタ通信に使用できます。</p>
2	<p>サイト間リンク</p> <p>解説:同一ベンダーあるいは 2 つの異なるベンダーのどちらであっても、複数のサイト間リンクを取得してサイト間の高い可用性を確保できます。</p>

8.7.2 長所

図 8-4 の構成には以下のような長所があります。

- 複数のサイト間リンクにより、サイト間で高い可用性が期待できる。
- コンポーネントの 1 つで障害が発生しても他のノードで処理を引き継ぐことができるよう LLA を使用する。他のノードが利用できるため LAN アダプタで障害が発生してもユーザに意識させない。

8.7.3 可用性と性能の主な強化手法

図 8-4 の構成では以下のような手法で可用性を強化しています。

- ロジカル LAN フェールオーバを可能にするデバイスを定義することで、ローカル LAN カードで障害が発生してもノードはそのまま動作し続け、ロジカル LAN フェールオーバ・セットで構成されている他のインタフェースに切り替えられます。
- すべてのノードが他のノードからアクセスを受けることができます。ロジカル LAN フェールオーバ・セットを構成するには次のコマンドを実行します。

```
$ MC LANCP  
LANCP> DEFINE DEVICE LLB/ENABLE/FAILOVER=(EIA0, EIB0)
```

『OpenVMS Cluster システム』の第 8 章で、JASMIN および TULIP の 2 つのノードを持つクラスタにノード ORCHID を追加する方法について説明しています。この例ではロジカル LAN フェールオーバ・セットは ORCHID で作成および構成されます。ローカルの LAN カードが故障しても ORCHID は動作し続け、口

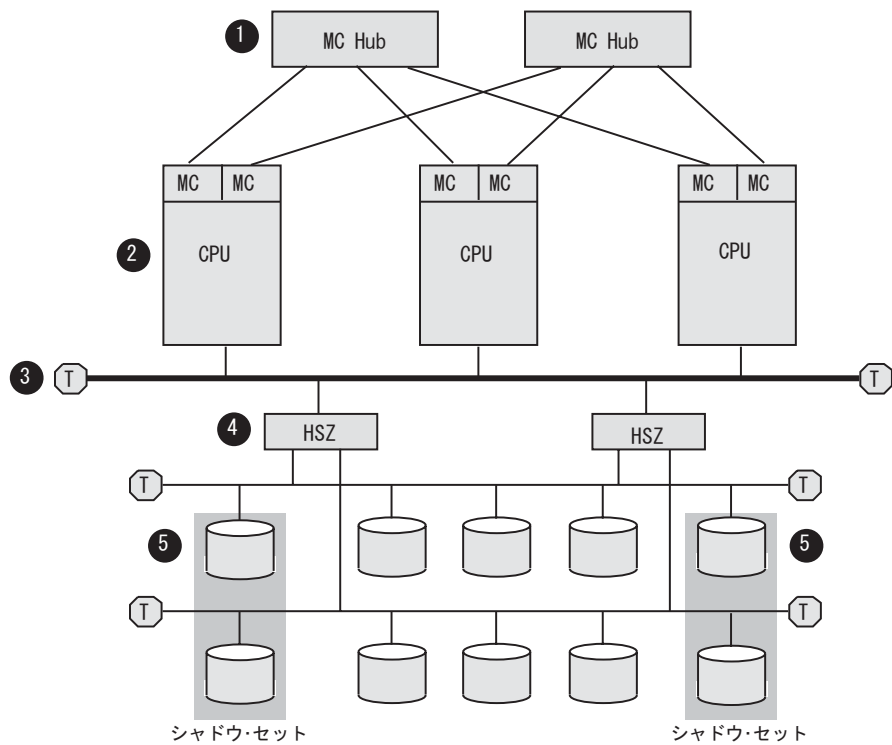
ジカル LAN フェールオーバー・セットに構成されている他のインタフェースに切り替えられます。

8.8 MEMORY CHANNEL OpenVMS Cluster における可用性

図 8-5 は、高い可用性を備えた MEMORY CHANNEL (MC) クラスタ構成です。
図 8-5 の後の項で、以下の構成を検討します。

- 構成要素の解析
- 長所と短所
- 可用性の主な強化手法の実装

図 8-5 MEMORY CHANNEL クラスタ



ZK8709AGE

8.8.1 構成要素

図 8-5 に示す MEMORY CHANNEL 構成には、以下の構成要素があります。

可用性を目的とした OpenVMS Cluster の構成

8.8 MEMORY CHANNEL OpenVMS Cluster における可用性

パート	説明
1	2 個の MEMORY CHANNEL ハブ 解説：2 個のハブとノードとの複数の接続により，単一点障害の要因を排除しています。
2	3 個から 8 個の MEMORY CHANNEL ノード 解説：定足数を満たすには，3 ノードを推奨します。1 本の MEMORY CHANNEL インターコネクで，最高 8 個の OpenVMS Alpha ノードをサポートします。 代替法：2 ノード構成では，ノードに障害が発生したときの定足数を満たすために，クォーラム・ディスクが必要です。
3	FWD (Fast-wide differential) SCSI バス 解説：データ転送速度を強化し (20 百万/秒)，2 つの HSZ コントローラをサポートするため FWD SCSI バスを使用します。
4	2 つの HSZ コントローラ 解説：2 つの HSZ コントローラでコントローラ障害時の冗長性を確保しています。2 つのコントローラにより，2 本のシングル・エンド SCSI バスの接続が実装され，さらに多くのストレージ領域を利用できます。
5	重要なシステム・ディスクとデータ・ディスク 解説：重要なディスクはシャドウ化し，シャドウ・セット・メンバは異なる SCSI バスに配置して，単一点障害の要因を排除します。

8.8.2 長所

この構成には，以下のような長所があります。

- 全ノードが全ストレージ領域に直接アクセスできる。
- SCSI ストレージ領域では，経済的でパフォーマンスの高い商品ハードウェアを提供できる。
- MEMORY CHANNEL インターコネクが，経済的で高いパフォーマンスのノード間通信を提供する。SCSI インターコネクは，経済的な商品ストレージ領域通信の提供により，MEMORY CHANNEL を補完する。

8.8.3 短所

この構成には，以下のような短所があります。

- FWD (Fast-wide differential) SCSI バスが，単一点障害の要因になる。これを防ぐには，1 本に障害が発生しても他のバスにフェールオーバーできるように第 2 の FWD (Fast-wide differential) SCSI バスを追加する。この機能を使用するには，システムで OpenVMS バージョン 7.2 またはそれ以上を実行しておき，マルチバス・サポートを有効にしておく。

8.8.4 可用性の主な強化手法

図 8-5 の構成では、以下の手法により可用性を強化しています。

- 冗長 MEMORY CHANNEL ハブと HSZ コントローラで、ハブが単一点障害の要因になったり、コントローラ障害が発生するのを防ぐ。
- ボリューム・シャドウイングにより、重要なデータ・ディスクの複数のコピーが、別々の HSZ コントローラに用意される。
- 全ノードが全ストレージ領域に対する共用の直接アクセスを備える。
- 定足数を満たすため、最低 3 ノードを使用しており、ノードのどれかに障害が発生しても、OpenVMS Cluster は処理を続行できる。

8.9 サテライトを持つ OpenVMS Cluster における可用性

サテライトとは、システム・ディスクや他の OpenVMS Cluster ストレージ領域に直接アクセスできないシステムです。サテライトは通常、ワークステーションですが、クラスタ内の他のノードによっては、ストレージ領域のサービスを受ける OpenVMS Cluster ノードもサテライトとして使用できます。

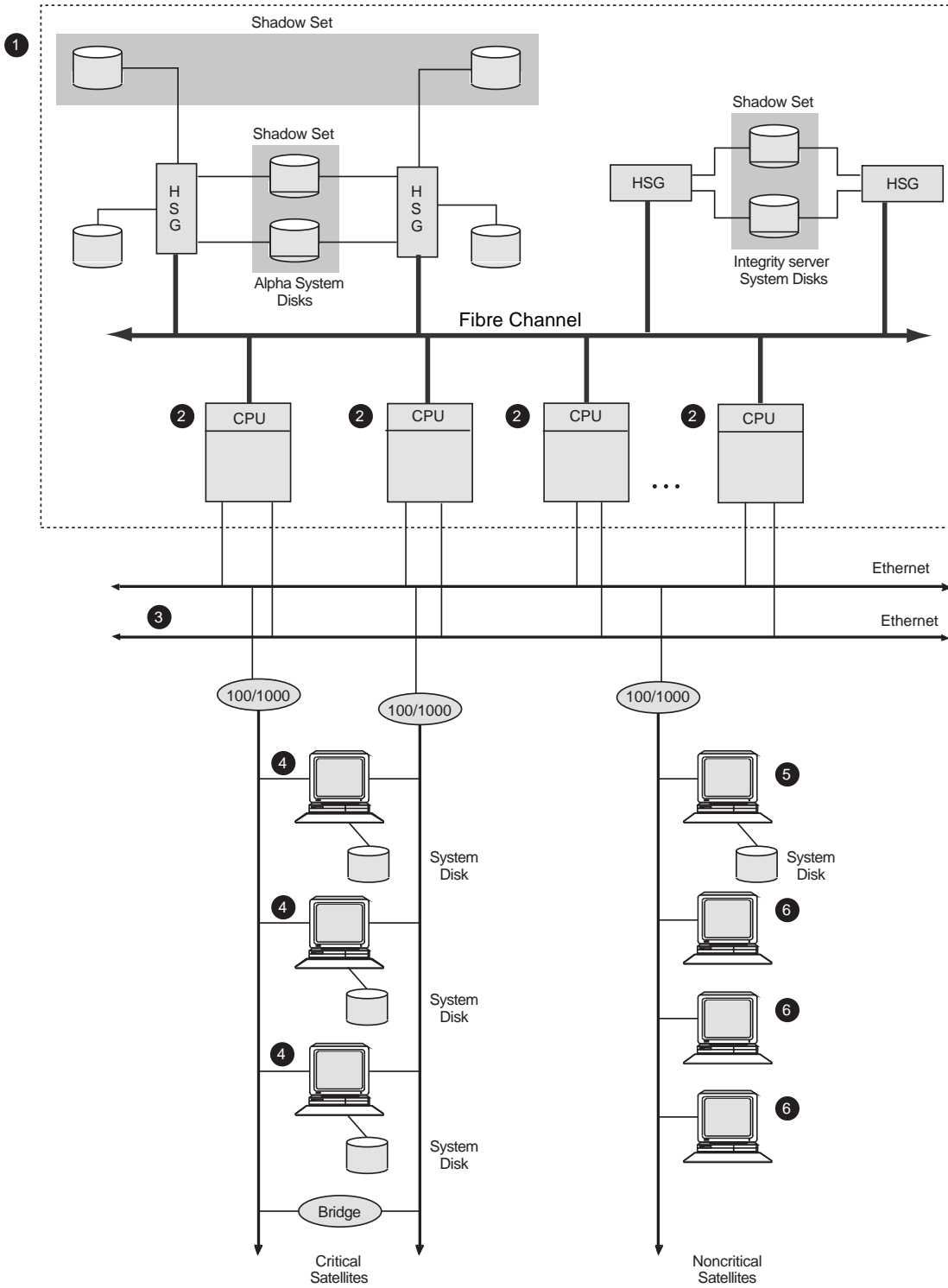
サテライト・ノードの可用性はサーバ・ノードによって大きく異なるので、この章の始めに紹介した構成例ではサテライト・ノードを使用していません。ただし、サテライト/サーバ構成には大きな長所があり、サテライト・ノードを構成に組み込むことで多少の可用性を犠牲にする場合もあります。

図 8-6 は、サテライトを持つ OpenVMS Cluster システムの最適な構成です。図 8-6 の後の項で、以下の構成を検討します。

- 構成要素の解析
- 長所と短所
- 可用性の主な強化手法の実装

可用性を目的とした OpenVMS Cluster の構成
 8.9 サテライトを持つ OpenVMS Cluster における可用性

図 8-6 サテライトを持つ OpenVMS Cluster



ZK6640AGE

8.9.1 構成要素

図 8-6 に示すこのサテライト/サーバ構成には、以下の構成要素があります。

パート	説明
1	ベース構成 ベース構成は、サテライトのサーバ機能を実行します。
2	3 個から 16 個の OpenVMS サーバ・ノード 解説：定足数を満たすには、少なくとも 3 ノードを推奨します。16 ノードを超えると管理がたいへん煩雑になります。
3	ベース・サーバ・ノードとサテライト間の 2 つの Ethernet セグメント 解説：高い可用性を提供します。
4	2 つの Ethernet アダプタで重要な各サテライトを接続している 2 つの Ethernet セグメント。 これら重要なサテライトには、それぞれ専用のシステム・ディスクがあります。 解説：専用のブート・ディスクを備えることで、重要なサテライトの可用性を強化しています。
5	重要度の低いサテライトの場合、ブート・サーバは Ethernet セグメントに配置します。 解説：重要度の低いサテライトに、専用のブート・ディスクは不要です。
6	サテライト数は 1 セグメント当たり 15 に制限。 解説：1 セグメント上のサテライト数が 15 を超えると I/O の輻輳が発生することがあります。

8.9.2 長所

この構成には、以下のような長所があります。

- 1 つの OpenVMS Cluster で多数のノードにサービスを提供できる。
- 多数のノードを長距離間に展開できる。

8.9.3 短所

この構成には、以下のような短所があります。

- シングル LAN アダプタのサテライトが原因の単一点障害により、アダプタに障害が発生するとクラスタ遷移が引き起される。
- 高い可用性を備えたサテライトを実現するには、LAN 接続にコストがかかる。

8.9.4 可用性の主な強化手法

図 8-6 の構成では、以下の手法により可用性を強化しています。

- 単一点障害の要因を構成から排除している。
- すべての共用ストレージ領域がベース構成から MSCP サービスを受ける。多数のノードのサービスが可能な構成になっている。

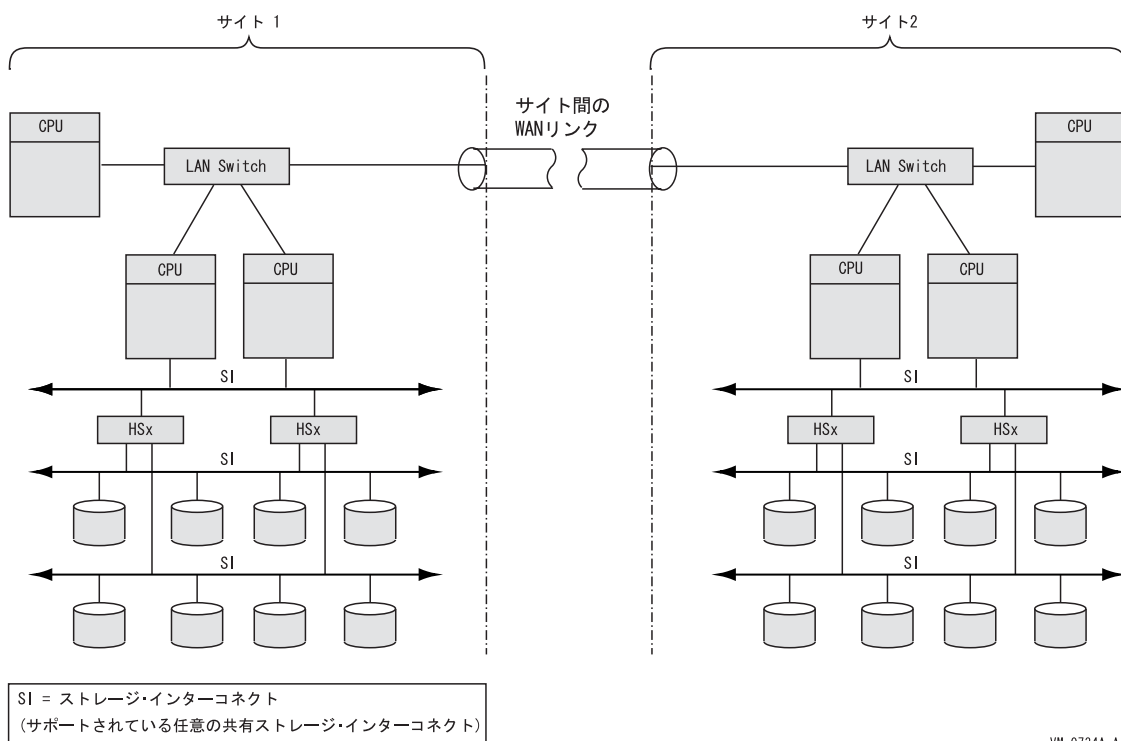
8.10 マルチサイト OpenVMS Cluster システム

マルチサイト OpenVMS Cluster 構成では、遠距離を隔てているサイトのノードに対応します。使用するテクノロジーにもよりますが、サイト間の距離は、最大で約 240 km になります。サイト間は、FDDI および DS3 で結合され、大きな 1 つのクラスターを形成します。ほとんどの電話会社のサービス網から利用できる DS3 サービスでは、マルチサイト・クラスター用の長距離ポイント間通信が可能です。

図 8-7 は、マルチサイト OpenVMS Cluster システムの代表的な構成例です。図 8-7 の後に、以下の構成を検討します。

- 構成要素の解析
- 長所

図 8-7 WAN リンクで結合されたマルチサイト OpenVMS Cluster 構成



8.10.1 構成要素

図 8-7 で、可能な構成の組み合わせをすべて表しているわけではありませんが、マルチサイト OpenVMS Cluster には次の構成要素を組み込むことができます。

- DECconcentrator または GIGAswitch クロスバー・スイッチに接続されたインターサイト・リンク (FDDI または DS3) を持つ 2 個のデータ・センター。

- 2 個のサイトが共用するアプリケーションと互換性を持つインターサイト・リンク・パフォーマンス。
- 最大 96 個の Integrity ノードと Alpha ノード (組み合わせ合計)。一般に、OpenVMS LAN クラスタと拡張 LAN (ELAN) クラスタに適用する規則は、マルチサイト・クラスタにも適用します。

関連項目： LAN 構成のガイドラインについては、第 4.10.2 項を参照してください。ELAN 構成のガイドラインについては、第 9.4.8 項を参照してください。

8.10.2 長所

マルチサイト OpenVMS Cluster システムには、以下の利点があります。

- セカンダリ・サイトに少数のシステムをリモートに配置して、一元システム管理やプライマリ・サイトの他のリソースを活かすことができる。たとえば、メイン・オフィス・データ・センターを倉庫や小規模な製造サイトにリンクすることができる。また、その倉庫や小規模な製造サイトには、サイト固有のデバイスが直接関連付けられた少数のローカル・ノードを配置できる。あるいは、都市全体に広がるオフィス・パークにプライマリ・ビジネス・サイトからエンジニアリング・ワークステーションをインストールできる。
- ハイ・キャパシティ・コンピュータ、テープ・ライブラリ、ディスク・アーカイブ、フォトタイプセッタなどのデバイスを複数のサイトで簡単に共用できる。
- クラスタ内にある任意のサイトのアーカイブ媒体に対してバックアップが可能。一般的な例としては、マルチサイト OpenVMS Cluster のすべてのサイトのデータを、1 個のサイトでディスクやテープにバックアップをとる方法がある。リモート・サイトのデータのバックアップは透過的に行われる (つまり、リモート・サイトでは仲介処理が不要)。
- 一般に、マルチサイト OpenVMS Cluster は、LAN OpenVMS Cluster の可用性の長所をすべて備えています。それだけでなく、複数の遠距離サイトを結合することで、マルチサイト OpenVMS Cluster 構成におけるシステムやシステム要素の可用性をさまざまな方法で強化できます。

– 論理ボリューム/データ可用性

ボリューム・シャドウイングまたは RAID (Redundant arrays of independent disks) により、両サイトにメンバがある論理ボリュームを作成できます。サイトのどちらかが使用できなくなっても、反対のサイトからデータを利用できます。

– サイト・フェールオーバ

VOTES システム・パラメータを調整しておけば、他のサイトに障害が発生したり、他のサイトとの通信ができなくなっても自動的に継続する優先サイトを選択できます。

関連項目：マルチサイト・クラスタのその他の説明については、『OpenVMS Cluster システム』を参照してください。

8.11 ディザスタ・トレラントな OpenVMS Cluster 構成

ディザスタ・トレラントな OpenVMS Cluster 構成では、Volume Shadowing for OpenVMS、高速ネットワーク、特別管理ソフトウェアを使用します。

ディザスタ・トレラントな OpenVMS Cluster 構成では、2 地点のサイトにあるシステムを 1 つの管理可能な OpenVMS Cluster システムに結合できます。前項で説明したマルチサイト・クラスタと同じく、これらの物理的に離れているデータ・センターは、FDDI で結合するか、T3、E3 のいずれかと FDDI の組み合わせにより結合します。

OpenVMS ディザスタ・トレラントな製品を、従来は BRS (Business Recovery Server) と称していましたが、BRS は、システム管理とソフトウェア・サービスのパッケージである Disaster Tolerant Cluster Services に包含されることになりました。Disaster Tolerant Cluster Services の詳細については、弊社のサービス担当者にお問い合わせください。

スケーラビリティを目的とした OpenVMS Cluster の構成

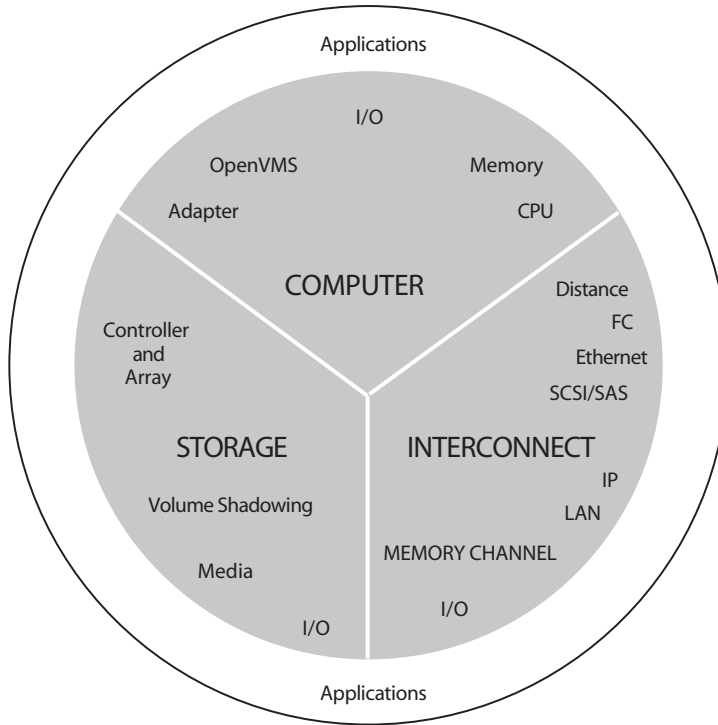
この章では、さまざまな OpenVMS Cluster でスケーラビリティを最大限に活かす方法を説明します。

9.1 スケーラビリティとは？

スケーラビリティは、初期の構成機器をフルに活用しながら、任意のシステム、ストレージ、インターコネクトの各次元で、OpenVMS Cluster システムを拡張できる能力をいいます。図 9-1 にも示すように OpenVMS Cluster システムの拡張はさまざまな次元に見られます。また、各次元ではユーザ・アプリケーションの拡張も可能です。

スケーラビリティを目的とした OpenVMS Cluster の構成
 9.1 スケーラビリティとは？

図 9-1 OpenVMS Cluster 拡張次元



VM-0218A-AI

9.1.1 スケーラブルな次元

表 9-1 は， OpenVMS Cluster におけるシステム，ストレージ，インターコネクットの拡張次元を図示したものです。

表 9-1 OpenVMS Cluster におけるスケーラブルな次元

対象次元	拡張要因
システム	
CPU	システム内に SMP を実装。 クラスタにシステムを追加。 クラスタに各種プロセッサ・サイズを格納。 クラスタに大きなシステムを追加。 Alpha システムから Integrity システムに移行。
メモリ	システムにメモリを追加。
I/O	インターコネクとアダプタをシステムに追加。 MEMORY CHANNEL をクラスタに追加して I/O インターコネクのロードを抑制。

(次ページに続く)

表 9-1 (続き) OpenVMS Cluster におけるスケーラブルな次元

対象次元	拡張要因
システム	
OpenVMS	システム・パラメータの調整。 OpenVMS Integrity へ移行。
アダプタ	ストレージ・アダプタをシステムに追加。 LAN アダプタをシステムに追加。
ストレージ	
媒体	ディスクをクラスタに追加。 テープと CD-ROM をクラスタに追加。
ボリューム・シャドウイング	ディスクのシャドウイングによる可用性の強化。 各コントローラのディスクのシャドウイング。 各システムのディスクのシャドウイング。
I/O	半導体ディスクまたは DECram ディスクをクラスタに追加。 キャッシュ付きディスクとコントローラをクラスタに追加。 RAID ディスクをクラスタに追加。
コントローラとアレイ	ディスクとテープをシステムからコントローラに移動。 ディスクとテープをアレイに組み合わせ。 コントローラとアレイをクラスタに追加。
インターコネクト	
LAN	Ethernet セグメントの追加。 Ethernet から Fast Etherne , Gigabit Ethernet , 10 Gigabit Ethernet へのアップグレード。 冗長セグメントの追加とセグメントのブリッジ化。
Fibre Channel , SCSI , SAS , MEMORY CHANNEL	Fibre Channel , SCSI , SAS , MEMORY CHANNEL インターコネクトをクラスタに追加 , または冗長インターコネクトをクラスタに追加。
I/O	キャパシティを目的とした高速インターコネクトの追加。 キャパシティと可用性を目的とした冗長インターコネクトの追加。
距離	室内または建物内でクラスタを展開。 都市内または数軒のビル間でクラスタを展開。 2 箇所のサイト間でクラスタを展開 (距離 40 km)。

表 9-1 に掲載した構成要素を、ユーザが選択した任意の方法で追加する機能は、OpenVMS Cluster における重要な機能に位置づけられます。この章や、製品のマニュアルの説明やガイドラインに従って、ハードウェアとソフトウェアはさまざまな方法の組み合わせで追加することができます。特定の次元における OpenVMS Cluster の拡張をする場合は、他の次元との関係で長所や短所を検討してください。表 9-2 は、OpenVMS Cluster のスケーラビリティを強化するための手法をまとめたものです。これらのスケーラビリティ手法を理解すれば、高いレベルのパフォーマンスと可用性を維持しながら OpenVMS Cluster の拡張に対応することができます。

9.2 スケーラビリティの高い OpenVMS Cluster の構成手法

選択するハードウェアと、その構成に適用する手法は、OpenVMS Cluster のスケーラビリティと重要な関係があります。この項では、スケーラビリティの高い OpenVMS Cluster の構成方法を説明します。

9.2.1 スケーラビリティの手法

表 9-2 は、スケーラビリティについて重要度の高い順に手法を並べています。この章では、この手法の実装方法を多くの図を使って説明します。

表 9-2 スケーラビリティの手法

手法	説明
キャパシティの計画	キャパシティの 80% (ほぼパフォーマンスとしては飽和状態) を超えてシステムを実行している状態では、それ以上の拡張性は望めません。 ビジネスやアプリケーションが拡大するかどうかを見極める必要があります。プロセッサ、メモリ、I/O に将来、どのような仕様が必要になるかを見積もることも大切です。
全ストレージへの共用の直接アクセス	コンピュータと I/O のパフォーマンスのスケーラビリティは、全システムから全ストレージまで共用の直接アクセスが設定されているかどうかで大きく異なります。 以下に示す FC と LAN OpenVMS Cluster の図では、MSCP のオーバーヘッドなしでストレージへの共用直接アクセスを実装したさまざまな例を示します。 関連項目：MSCP のオーバーヘッドの詳細については、第 9.6.1 項を参照してください。
ノード・カウントの範囲を、3 から 16 に制限	OpenVMS Cluster が小さければパフォーマンスの管理や調整がしやすく、大きな OpenVMS Cluster よりも OpenVMS Cluster 通信のオーバーヘッドが少なく済みます。ノード・カウントを抑制するには、強力なプロセッサにアップグレードし、OpenVMS SMP の能力を活かします。 サーバに負荷がかかり過ぎてコンピューティングのボトルネックになっている場合、アプリケーションをノード間に分散できないか検討します。可能であれば、ノードを追加します。不可能な場合は、プロセッサ (SMP) を追加します。
システム・ボトルネックを取り除く	OpenVMS Cluster 機能のキャパシティを最大化するにあたっては、その機能を実装するハードウェアとソフトウェアの構成要素を検討します。ボトルネックになっている構成要素があると、他の構成要素のポテンシャルをフルに活せなくなります。ボトルネックを探し出し、その要因を削減すれば、OpenVMS Cluster のキャパシティを強化できます。
MSCP サーバを使用可能にする	MSCP サーバを利用すれば、OpenVMS Cluster にサテライトを追加でき、すべてのストレージへのアクセスをノード間で共用できます。また、MSCP サーバには、インターコネクに障害が発生すると共用ストレージに対するフェールオーバを実行する機能があります。
相互依存の緩和と単純な構成	システム・ディスクが 1 つだけの OpenVMS Cluster システムは、そのディスクだけに OpenVMS Cluster の実行を依存します。ディスクや、そのディスクのサービスをしているノード、あるいはノード間のインターコネクに障害が発生すると、OpenVMS Cluster システム全体が使用できなくなります。

(次ページに続く)

表 9-2 (続き) スケーラビリティの手法

手法	説明
サービス用リソースの確保	容量の小さいディスク・サーバで多くのディスクをサテライトに提供する場合、OpenVMS Cluster 全体のキャパシティが制約されます。サーバが過剰負荷になるとボトルネックになり、フェールオーバーによる回復操作を効果的に処理できなくなるので注意してください。
リソースとコンシューマを近づける	サーバ(リソース)とサテライト(コンシューマ)の間を近づけます。OpenVMS Cluster のノード数を追加する場合は、まず分割することを検討してください。詳細については、第 10.2.4 項を参照してください。
適切なシステム・パラメータの設定	OpenVMS Cluster の拡張度合いがめざましい場合、重要なシステム・パラメータが現状に合わなくなることがあります。このようなときは、重要なシステム・パラメータを自動的に計算し、ページ、スワップ、ダンプの各ファイルのサイズ調整ができる AUTOGEN を実行してください。

9.3 SCSI OpenVMS Cluster におけるスケーラビリティ (Alpha のみ)

SCSI 方式の OpenVMS Cluster では、市販のストレージ・レベルのストレージを OpenVMS Cluster に直接使用できます。OpenVMS Cluster で SCSI インターコネクタを使用することで、距離、価格、パフォーマンスの各キャパシティの選択肢が広がります。この SCSI クラスタリング機能は、ローエンドの手ごろなクラスタ・ソリューションを構成すれば、理想的な出発点になります。SCSI クラスタは、デスクトップからデスクサイド、あるいは部門別の大規模な構成まで適用できます。

SCSI インターコネクタの使用にあたっては、以下の一般的な制約に注意してください。

- SCSI インターコネクタでは、ストレージ・トラフィックしか処理しないので、ノード間トラフィックでは、もう 1 本のインターコネクタとペアで使用する必要があります。この節で紹介している図では、MEMORY CHANNEL を代替インターコネクタに使用しています。ただし、CI、DSSI、Ethernet、FDDI も使用できます。
- SCSI ケーブルの長さの合計には、システムの内部ケーブル長も含まれます。たとえば、AlphaServer 1000 ラックマウントでは、内部アダプタを外部コネクタに接続するための内部ケーブルに 1.6 m 必要です。2 m SCSI ケーブルで接続した 2 つの AlphaServer 1000 の場合、それぞれ内部配線に 1.6 m 必要なので、SCSI バスの合計の長さは 5.2 m になります。

関連項目：内部 SCSI ケーブルの長さや、SCSI デバイスのクラスタリングの詳細については、付録 A を参照してください。

- この機能を利用できるのは、古い SCSI アダプタだけです。KZPEA、A6828A、A6829A、および A7173A などの最新の SCSI アダプタは、この機能をサポートしていません。

スケーラビリティを目的とした OpenVMS Cluster の構成
9.3 SCSI OpenVMS Cluster におけるスケーラビリティ (Alpha のみ)

この節の図では、比較的小容量のストレージを備えた 2 ノード SCSI 構成からストレージ容量を最大とし、スケーラビリティにも余裕のある 4 ノード SCSI ハブまで段階を追って紹介します。

9.3.1 3 ノード Fast Wide SCSI Cluster

図 9-2 では、25-m の Fast Wide (FWD) SCSI インターコネクトで 3 ノードが接続されています。HSZ コントローラ内に複数のストレージ・シェルフがあり、図の上部の BA356 にはさらに多くのストレージが格納されています。

図 9-2 3 ノード Fast Wide SCSI Cluster

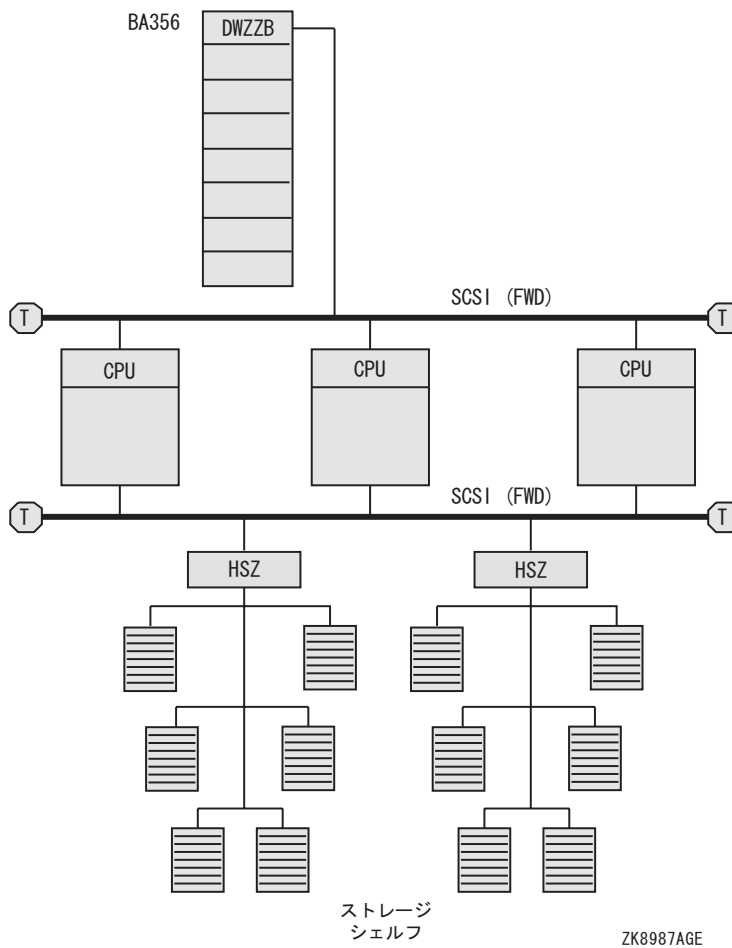


図 9-2 に示すこの構成の長所と短所は次のとおりです。

長所

- 2 ノード Fast Wide SCSI クラスタと HSZ ストレージを備えた 2 ノード Fast Wide SCSI クラスタの構成の長所を備えている。
 - バスの許容長さ (25 m) とスケーラビリティに余裕がある。
 - HSZ にキャッシュを組み込んでいる。RAID 0, 1, 5 テクノロジをサポートしている。HSZ に複数のストレージ・シェルフがある。
 - FWD バスによりスループット 20 MB/s が可能。
 - BA356 キャビネットでは, Narrow (8 ビット) または Wide (16 ビット) SCSI バスを使用できます。

短所

- これまでの図で示した構成よりもコストがかかる。

9.3.2 4 ノード Ultra SCSI ハブ構成

図 9-3 では, SCSI ハブで 4 ノードが接続されています。SCSI ハブの電源と冷却は, BA356 などのストレージ・キャビネットから供給します。SCSI ハブは, ストレージ・キャビネットの SCSI バスに接続されていません。

スケーラビリティを目的とした OpenVMS Cluster の構成
9.3 SCSI OpenVMS Cluster におけるスケーラビリティ (Alpha のみ)

図 9-3 4 ノード Ultra SCSI ハブ構成

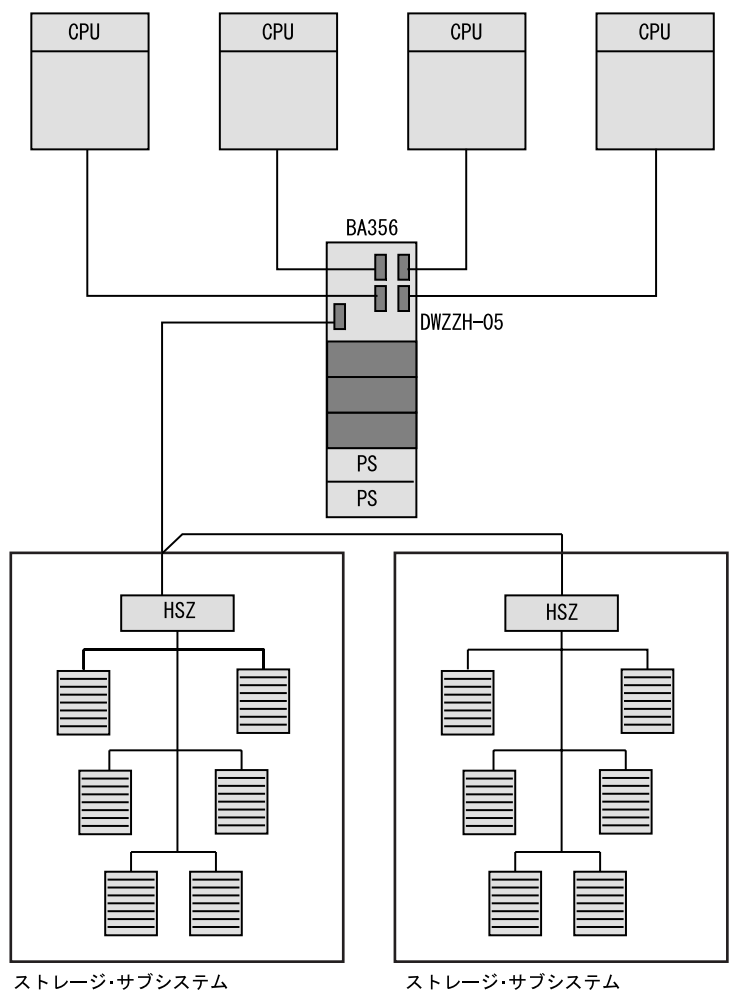


図 9-3 に示すこの構成の長所と短所は次のとおりです。

長所

- バス長さとスケーラビリティが優れている。
- SCSI ハブでは、SCSI バス上のアービトレーションが公正に行われ、システムの挙動が均等化され予測しやすくなる。アービトレーションが公正でないと 4 つの CPU を活用できない。
- ハブのストレージ・ポートに、デュアル HSZ コントローラを最大 2 つまでデジター・チェーン接続できる。
- BA356 に電源が 2 つある (1 つはバックアップ電源)。
- HSZ のキャッシュにより、RAID 0、1、5 テクノロジもサポートされる。
- Ultra SCSI バスによりスループット 40 MB/s が可能。

短所

- CPU や HSZ から別の CPU までの SCSI インターコネクトをデジター・チェーン化してこの構成に CPU 追加することができない。
- 2 ノード Fast Wide SCSI クラスタおよび HSZ ストレージを備えた 2 ノード Fast Wide SCSI クラスタの構成よりはるかにコストがかかる。
- HSZ ストレージしか接続できない。SCSI ハブに直接ディスク・ドライブを接続したストレージ・シェルフは接続できない。

9.4 サテライトを備えた OpenVMS Cluster のスケーラビリティ (Alpha, VAX, および Integrity)

OpenVMS Cluster におけるサテライト数と MSCP でサービスを受けるストレージの容量により、サーバ数とキャパシティが決まります。サテライトは、システム・ディスクや他の OpenVMS Cluster ストレージに直接アクセスしないシステムです。サテライトは、通常はワークステーションですが、OpenVMS Cluster 内の他のノードのサービスを受ける OpenVMS Cluster ノードでもサテライトとして使用できます。

Ethernet LAN セグメント 1 つにつき、10 から 20 のサテライト・ノードを接続できます。図 9-4、図 9-5、図 9-6、図 9-7 は、6 サテライト LAN から 45 サテライト LAN まで段階的に拡張した構成例です。

9.4.1 6 サテライト OpenVMS Cluster

図 9-4 では、6 つのサテライトとブート・サーバが Ethernet で接続されています。

スケーラビリティを目的とした OpenVMS Cluster の構成
9.4 サテライトを備えた OpenVMS Cluster のスケーラビリティ (Alpha , VAX , および Integrity)

図 9-4 6 サテライト LAN OpenVMS Cluster

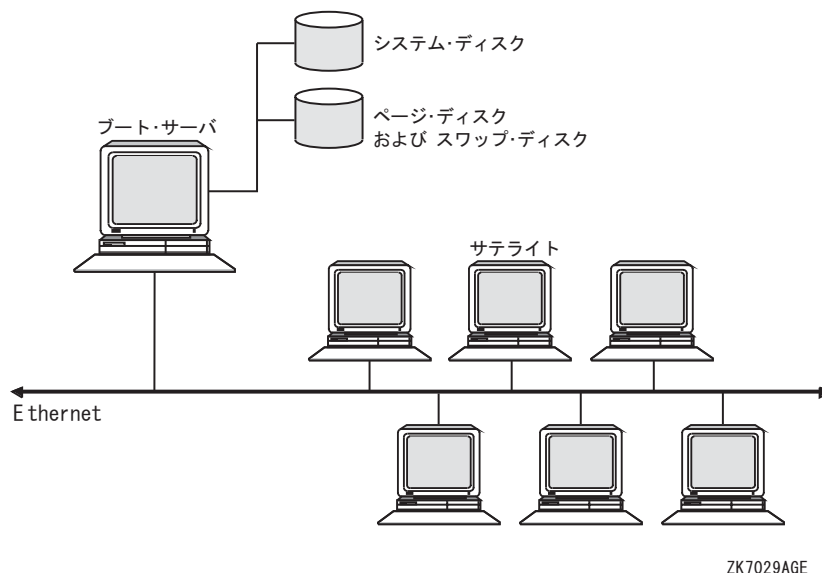


図 9-4 に示すこの構成の長所と短所は次のとおりです。

長所

- MSCP サーバを使用可能にして、サテライトの追加とアクセスできるストレージの拡張に対応できる。
- システム・ディスクが1つなので、システム管理が比較的簡単である。

関連項目：システム・ディスクの管理方法については、第 10.2 節を参照してください。

短所

- Ethernet がボトルネックや単一点障害の要因になる可能性がある。

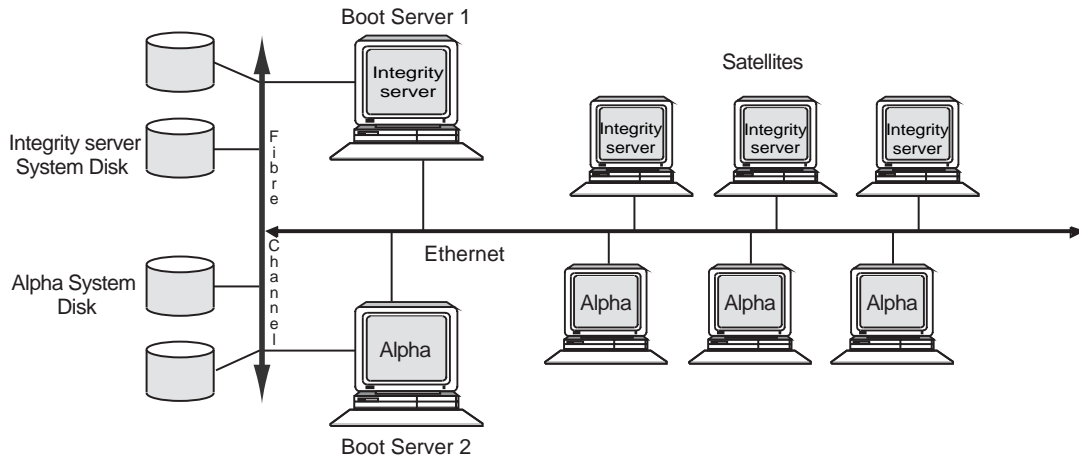
図 9-4 に示すブート・サーバがボトルネックになる場合、図 9-5 に示すような構成が必要になることもあります。

9.4.2 2 ブート・ノードを備えた 6 サテライト OpenVMS Cluster

図 9-5 は、Ethernet で接続された 6 つのサテライトと 2 つのブート・サーバです。ブート・サーバ 1 とブート・サーバ 2 は MSCP サーバの動的負荷バランス調整を実行します。各ノードは、ノード間の負荷を調整、共用します。どれかのノードの機能が停止すると、他のノードがその処理を引き受けます。MSCP 動的負荷調整には、ストレージまでの共用アクセスが必要です。

スケーラビリティを目的とした OpenVMS Cluster の構成
9.4 サテライトを備えた OpenVMS Cluster のスケーラビリティ (Alpha , VAX , および Integrity)

図 9-5 2 ブート・ノードを備えた 6 サテライト LAN OpenVMS Cluster



VM-7030A-GE

図 9-5 に示すこの構成の長所と短所は次のとおりです。

長所

- MSCP サーバを使用可能にして、サテライトの追加とアクセスできるストレージの拡張に対応できる。
- 2 ブート・サーバが MSCP 動的負荷バランス調整を実行する。

短所

- Ethernet がボトルネックや単一点障害の要因になる可能性がある。

図 9-5 に示す LAN が OpenVMS Cluster ボトルネックになる場合、この構成は図 9-6 のようになります。

9.4.3 2 LAN セグメントを備えた 12 サテライト LAN OpenVMS Cluster

図 9-6 は、2 つの Ethernet セグメントで接続された 12 サテライトと 2 ブート・サーバです。これら 2 つの Ethernet セグメントは、LAN ブリッジでも結合されています。サテライトごとにストレージまでのデュアル・パスがあるので、この構成では、MSCP 動的負荷バランス調整も可能です。

スケーラビリティを目的とした OpenVMS Cluster の構成
9.4 サテライトを備えた OpenVMS Cluster のスケーラビリティ (Alpha , VAX , および Integrity)

図 9-6 2 LAN セグメントを備えた 12 サテライト OpenVMS Cluster

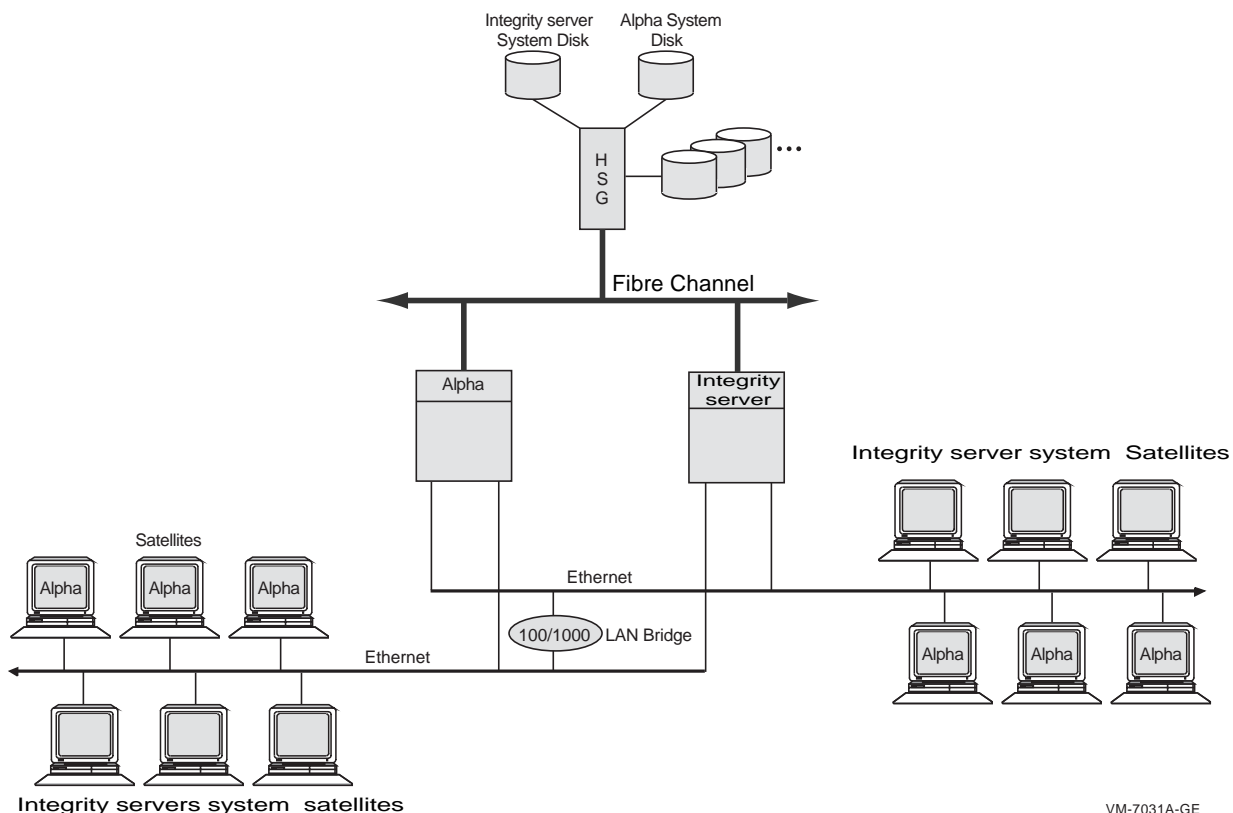


図 9-6 に示すこの構成の長所と短所は次のとおりです。

長所

- MSCP サーバを使用可能にして、サテライトの追加とアクセスできるストレージの拡張に対応できる。
- 2 ブート・サーバで MSCP 動的負荷バランス調整が可能である。

Ethernet LAN のサテライトの観点から見れば、Alpha ノードや Integrity ノードまでのデュアル・パスは MSCP 負荷調整で有利です。

- 2 LAN セグメントにより、LAN キャパシティが増加する。

短所

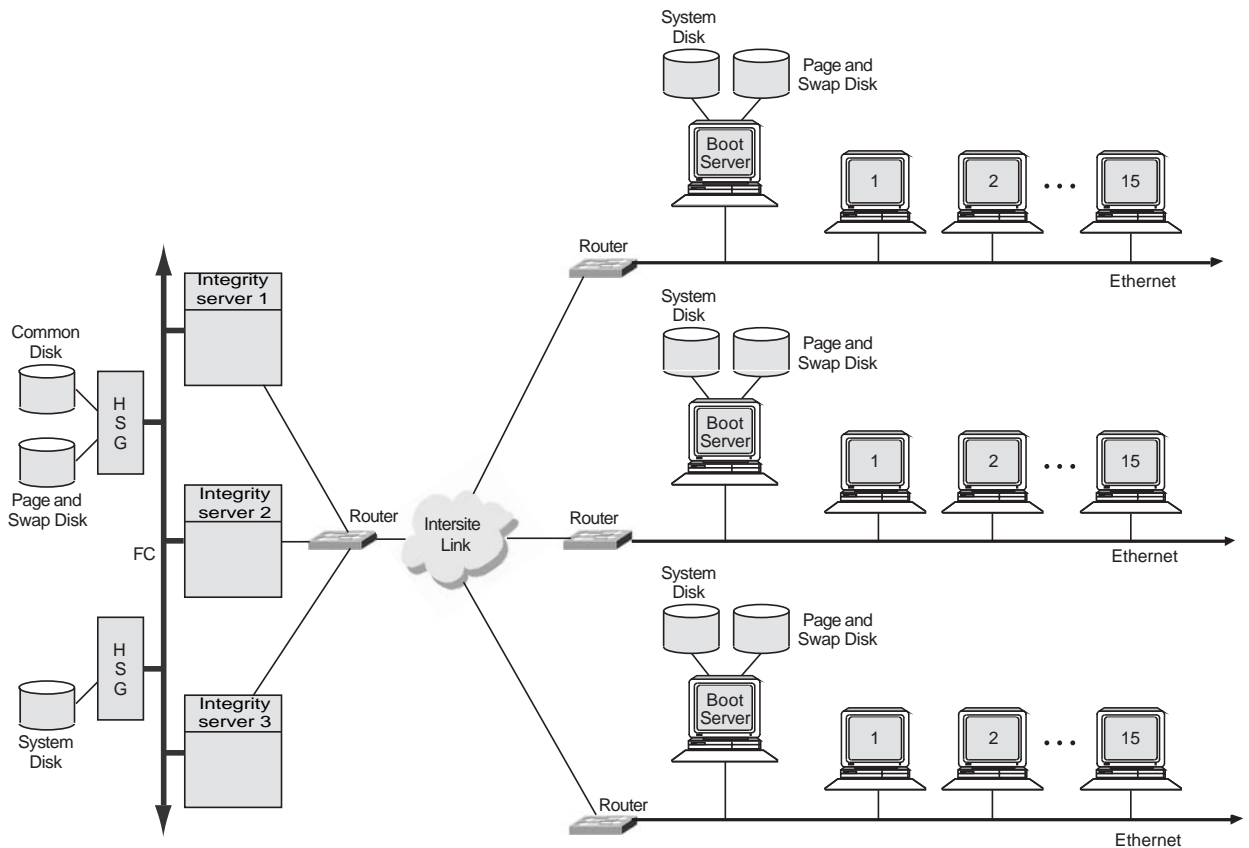
- この OpenVMS Cluster 構成は、サポートできるサテライト数が制限される。
- 1 つしかない HSG コントローラが、ボトルネックと単一点障害の要因になる可能性がある。

図 9-6 に示す OpenVMS Cluster を現在の制限以上に拡張する場合、この構成は図 9-7 に示すような内容になります。

9.4.4 サイト間リンクを備えた 45 サテライト OpenVMS Cluster

図 9-7 は、45 つのサテライト・ノードを接続した大規模な 51 ノード OpenVMS Cluster です。3 つのブート・サーバ、Integrity 1, Integrity 2, Integrity 3 は、3 つのディスク、つまり共通ディスク、ページ・スワップ・ディスク、システム・ディスクを共有します。サイト間リンクはルータに接続されており、3 つの LAN セグメントが接続されています。各セグメントには、ワークステーション・サテライトが 15 個の他、専用のブート・ノードがあります。

図 9-7 サイト間リンクを備えた 45 サテライト OpenVMS Cluster



ZK7034AGE

図 9-7 に示すこの構成の長所と短所は次のとおりです。

長所

- ブート時間を短縮できる。ノード・カウントが多い OpenVMS Cluster では特に効果的。

関連項目： 図 9-7 に示すような OpenVMS Cluster のブートについては、第 10.2.4 項を参照してください。

- MSCP サーバを使用可能にして、サテライトからアクセスできるストレージを拡張できる。
- 各ブート・サーバには、専用のページ・スワップ・ディスクがあるので、システム・ディスクにおける I/O 処理を短縮できる。
- OpenVMS Cluster 全体の環境ファイルは、すべて共通ディスクにある。そのため、サテライト・ブート・サーバはサテライトまでのルート情報だけを処理すればよい。

関連項目：共通ディスクとページ・ディスクおよびスワップ・ディスクの詳細については、第 10.2 節を参照してください。

短所

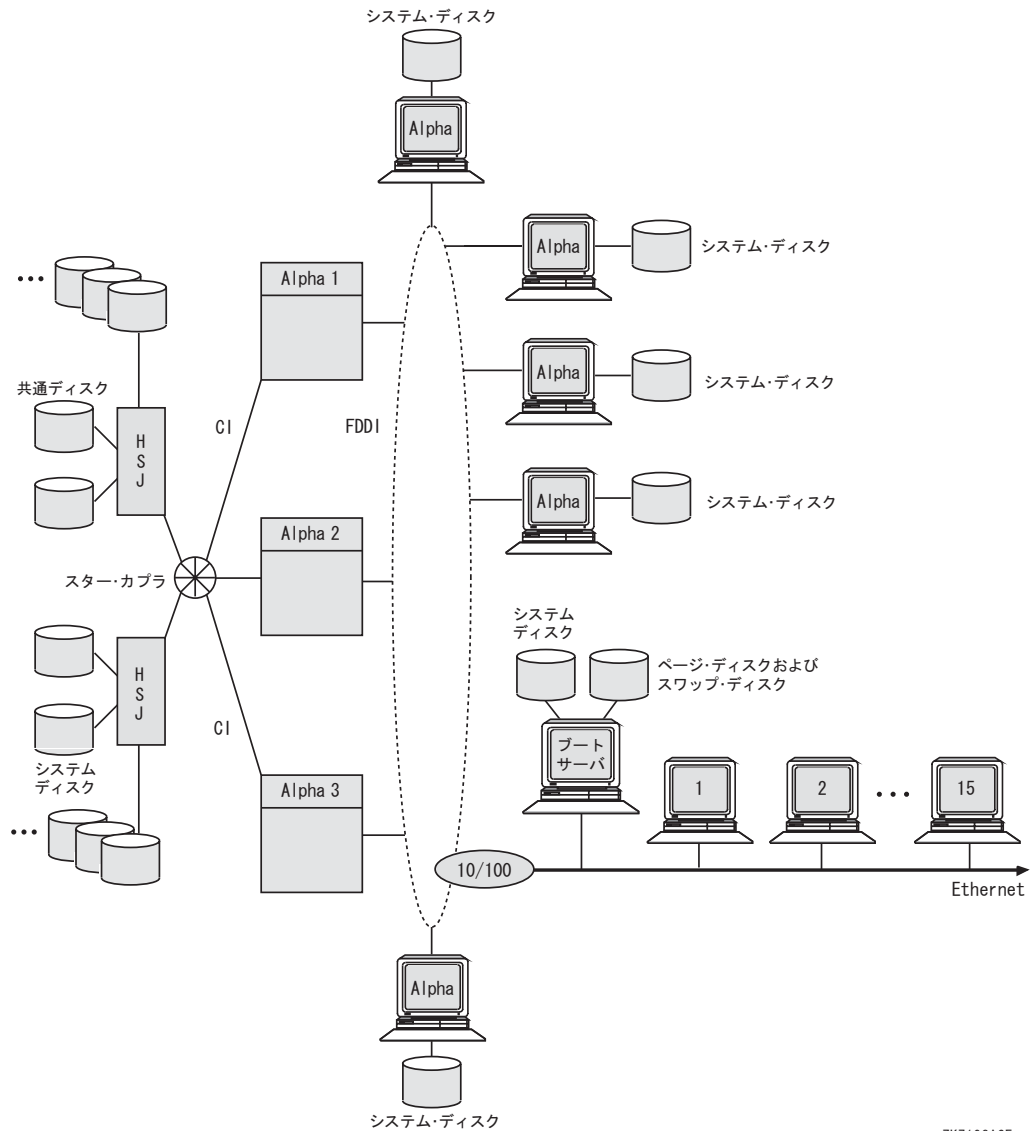
- Ethernet LAN セグメント上のサテライト・ブート・サーバは、同じセグメントのサテライトでないとブートできない。

9.4.5 ハイパワー・ワークステーション OpenVMS Cluster (1995 年のテクノロジー)

図 9-8 は、FDDI リング上で高いパフォーマンスと可用性を実現できる OpenVMS Cluster 構成です。

スケーラビリティを目的とした OpenVMS Cluster の構成
9.4 サテライトを備えた OpenVMS Cluster のスケーラビリティ (Alpha , VAX , および Integrity)

図 9-8 ハイパワー・ワークステーション・サーバの構成 (1995 年)



ZK7198AGE

図 9-8 には、専用のシステム・ディスクを持った Alpha ワークステーションが数台あります。これらは、FDDI リングに接続されています。Alpha ワークステーションを FDDI に接続すると各ワークステーションが自身のシステム・ディスクに直接アクセスできるため、高いパフォーマンスが可能になります。また、FDDI 帯域幅は、Ethernet の場合よりも高くなります。Alpha ワークステーションには FDDI アダプタがあるため、これらのワークステーションを FDDI に接続するのは、ワークステーションの重要度に応じた代替策として有効です。FDDI は Ethernet よりも 10 倍高速であり、Alpha ワークステーションには、FDDI の処理速度を活かせるだけの処理能力があります。(Fast Ethernet は FDDI と同等の速度で、Gigabit Ethernet は Fast Ethernet や FDDI より 10 倍高速です。)

スケーラビリティを目的とした OpenVMS Cluster の構成

9.4 サテライトを備えた OpenVMS Cluster のスケーラビリティ (Alpha , VAX , および Integrity)

9.4.6 ハイパワー・ワークステーション OpenVMS Cluster (2004 年のテクノロジー)

図 9-9 に、LAN に Gigabit Ethernet を使用し、ストレージに Fibre Channel を使用した、ハイパフォーマンスで可用性の高い OpenVMS Cluster 構成を示します。

図 9-9 ハイパワー・ワークステーション・サーバの構成 (2004 年)

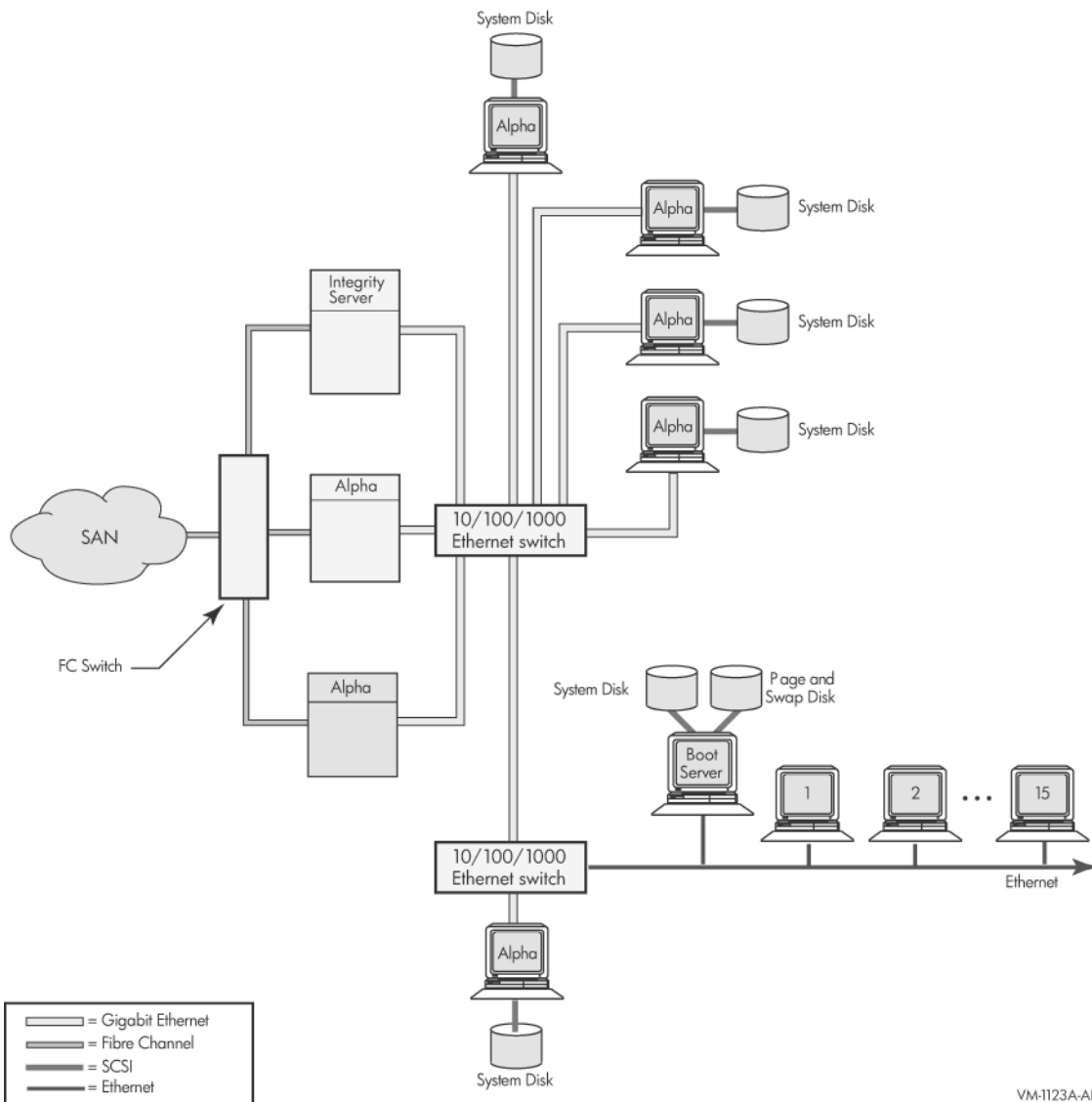


図 9-9 には、専用のシステム・ディスクをそれぞれ持った Alpha ワークステーションが数台あり、これらは Gigabit Ethernet LAN に接続されています。各ワークステーションが自身のシステム・ディスクに直接アクセスできるため、それらを Gigabit Ethernet LAN に接続することで高いパフォーマンスを実現できます。さらに、Gigabit Ethernet の帯域幅は、FDDI よりも 10 倍広く高速です。Alpha ワークステーションには、Gigabit Ethernet の速さを活かせるだけの処理能力があります。

9.4.7 サテライトを備えた OpenVMS Cluster のガイドライン

以下に示すのは、サテライトを持った OpenVMS Cluster をセットアップするときのガイドラインです。

- 大規模 LAN 構成のサテライトには多くのメモリが必要です。これは、ノードごとに他のすべてのノードとの接続を管理するからです。
- ボトルネックを排したネットワークを構成してください(つまり、ネットワーク群やサーバ接続に、十分な帯域幅を割り当ててください)。
- 図 9-5 と図 9-6 に示すように MSCP 動的負荷バランス調整によりリソースを最大限に活用してください。
- 優れたパフォーマンスを維持するため、MSCP サービスが必要なノード数は最小限にしてください。

関連項目：MSCP オーバヘッドの詳細については、第 9.6.1 項を参照してください。

- 時間を節約するため、ブート・シーケンスの効率化を図ってください。特に OpenVMS Cluster が大規模な場合や複数のセグメントがある場合にこれは必要です。LAN とシステム・ディスクの処理を削減する方法と、シーケンス内で別々のノード・グループをブートする方法の詳細については、第 10.2.4 項を参照してください。
- ホスト 1 台に対して複数の LAN アダプタを使用し、別々の LAN バスに接続してください。これは、ノード間の同時双方向通信を可能にし、複数のノードを宛先とするトラフィックを可用 LAN 間に分散させるためです。さらに、複数の LAN アダプタにより、フェールオーバーの機能が強化されます。

9.4.8 拡張 LAN 構成のガイドライン

LAN セグメント間にブリッジとスイッチを使用すると、拡張 LAN を形成できます。シングル LAN に比べたとき、これにより、可用性、接続距離、全帯域幅の強化が可能です。ただし、拡張 LAN では、遅延が増加し、パスによっては帯域幅が低下する因子もあります。パケット損失、キューイング遅延、パケット・サイズのような要因もネットワークのパフォーマンスに影響を与えることがあります。表 9-3 は、以上の因子を扱う場合に、十分な LAN パフォーマンスを確保するためのガイドラインをまとめたものです。

表 9-3 拡張 LAN 構成ガイドライン

係数	ガイドライン
伝播遅延	<p>パケットが LAN を移動する時間は、移動距離と、スイッチまたはブリッジによってリンク間で中継される回数によって異なります。応答時間を重視する場合は、これらの因子を制御する必要があります。</p> <p>ハイ・パフォーマンス・アプリケーションでは、ノード間のスイッチ数を 2 つまでにしてください。ハイ・パフォーマンスが求められない場合、ノード間で使用できるスイッチまたはブリッジの数は 7 つまでとします。</p>
キューイング遅延	<p>スイッチまたはブリッジとホスト・アダプタにおいて、即時着信率がサービス率を超えるとキューイングが発生します。キューイングは以下の方法で制御できます。</p> <ul style="list-style-type: none"> 通信頻度の高いノード間は、スイッチまたはブリッジの数を削減する。 ハイ・パフォーマンスなスイッチまたはブリッジとアダプタだけを使用する。 LAN におけるトラフィック・バーストを削減する。たとえば、小規模な I/O は結合し、小さいパケットのバーストの代わりに 1 つのパケットにまとめられる。 高速プロセッサと高速 LAN を利用し、スイッチまたはブリッジでトラフィックを分離して、LAN セグメントとホスト・プロセッサの利用レベルを下げる。
パケット損失	<p>LAN が配信しなかったパケットは、再送信が必要であり、これはシステムおよびネットワーク・リソースの余分な消費や遅延の増加、あるいは帯域幅の低下をもたらします。ブリッジとアダプタに輻輳が生じると、パケットが廃棄されます。このパケット損失を少なくするには、先に説明したキューイングを制御します。</p> <p>通過時に損傷を受けたパケットも廃棄されます。この問題を制御するには、LAN ハードウェアの構成規則を順守し、電気的な干渉源を排除し、確実にすべてのハードウェアを適正に運用します。</p> <p>再送タイムアウト率では、パケット損失の兆候を知ることができます。ノード間の OpenVMS Cluster トラフィックの場合、送信 1000 回に対しタイムアウト 1 回を限度とします。ハイ・パフォーマンス・アプリケーションに使用される LAN パスでは、さらに大幅に低い率が求められます。OpenVMS Cluster における再送タイムアウトの発生状況をよく監視してください。</p> <p>関連項目：再送タイムアウトの発生状況の監視については、『OpenVMS Cluster システム』を参照してください。</p>
スイッチまたはブリッジの回復遅延	<p>自己診断時間が短いスイッチまたはブリッジを選択し、高速自動再構成ができるようにそれらを調整してください。ネットワークの要件に合わせてスパンニング・ツリーのパラメータを調整することも含みます。</p> <p>関連項目：LAN ブリッジ・フェールオーバーの詳細については、『OpenVMS Cluster システム』を参照してください。</p>
帯域幅	<p>OpenVMS Cluster 通信に使用するすべての LAN パスは、最低 10 Mb/s の公称帯域幅で使用できるものに限定します。LAN セグメントの平均利用率は、任意の 10 秒間に 60%を超えないようにします。</p> <p>Gigabit Ethernet および 10 Gigabit Ethernet 構成の場合は、可能であればジャンボ・フレームを有効にしてください。</p>

(次ページに続く)

スケラビリティを目的とした OpenVMS Cluster の構成
 9.4 サテライトを備えた OpenVMS Cluster のスケラビリティ (Alpha , VAX , および Integrity)

表 9-3 (続き) 拡張 LAN 構成ガイドライン

係数	ガイドライン
トラフィックの分離	<p>相互の通信頻度が高いノード間では、スイッチまたはブリッジを利用してトラフィックを分離しローカライズしてください。たとえば、スイッチまたはブリッジは、他の LAN から OpenVMS Cluster を分離したり、OpenVMS Cluster 内の通信頻度の高いノードを他の OpenVMS Cluster 内のノードから分離するときに使用します。</p> <p>複数のアダプタがある重要なシステム間には、LAN によって独立したバスを提供してください。</p>
パケット・サイズ	<p>LAN バスが、エンド・ツー・エンドで最低 4474 バイトのデータ・フィールドをサポートしているか確認してください。ジャンボフレームを使用した Gigabit Ethernet デバイスに対しては、NISCS_MAX_PKT SZ システム・パラメータを 8192 バイトに設定してください。</p> <p>障害によっては、トラフィックが大きなパケット・サイズをサポートする LAN バスから、より小さいサイズのパケットしかサポートしないバスに切り替わることがあります。この種の障害の自動検出と自動回復は実装可能です。</p>

9.4.9 OpenVMS Cluster のシステム・パラメータ

サテライトとサーバがある OpenVMS Cluster には、OpenVMS Cluster を効率よく管理できるシステム・パラメータが用意されています。表 9-4 は、これらのシステム・パラメータの推奨値をまとめたものです。

表 9-4 OpenVMS Cluster システム・パラメータ

システム・パラメータ	サテライト用の値	サーバ用の値
LOCKDIRWT	0	1 ~ 4。LOCKDIRWT の設定は、ノードがリソース・ディレクトリ・ノードとしてどのようなサービスを行うかに影響します。また、リソース・ツリーのマスタを決定するために使用することもできます。一般的に、1 より大きい設定は、クラスタ・ノードの具体的なワークロードと、各アプリケーションの稼動状況を慎重に調査した後決定するものであり、本書では説明しません。
SHADOW_MAX_COPY	0	4、環境によってはさらに高い値に設定可能
MSCP_LOAD	0	1
NPAGEDYN	スタンドアロン・ノードの場合よりも高い値	サテライト・ノードの場合よりも高い値
PAGEDYN	スタンドアロン・ノードの場合よりも高い値	サテライト・ノードの場合よりも高い値
VOTES	0	1
EXPECTED_VOTES	OpenVMS Cluster ボーツの合計値	OpenVMS Cluster ボーツの合計値
RECNXINTERVL ¹	全ノードに同じ値	全ノードに同じ値

¹ブリッジ時間と LAN 利用率に関係。

関連項目：これらのパラメータの詳細については、『OpenVMS Cluster システム』および『Volume Shadowing for OpenVMS 説明書』を参照してください。

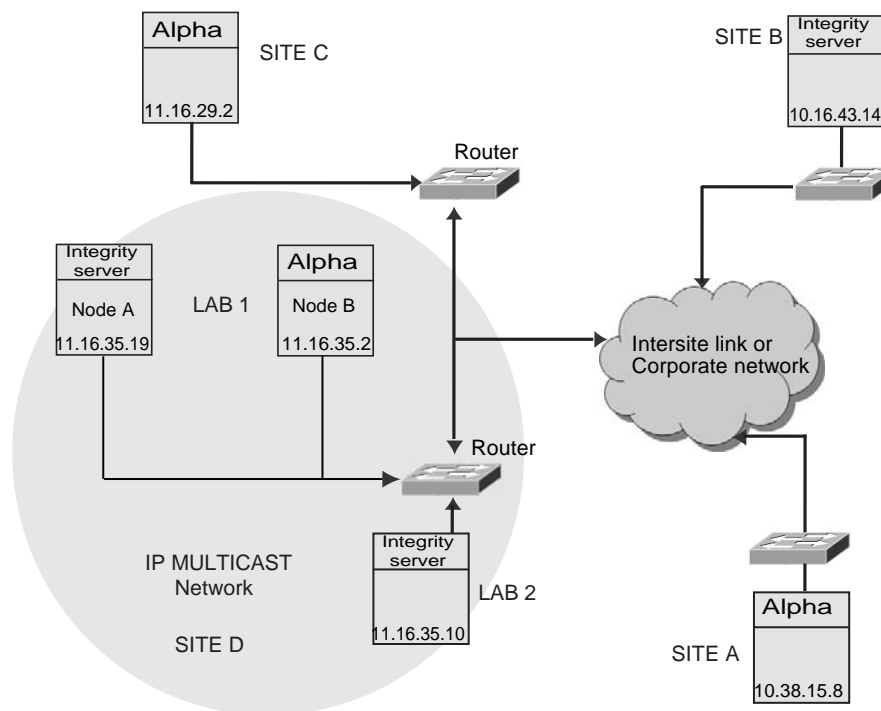
9.5 Cluster over IP におけるスケーラビリティ

Cluster over IP では、地理的に離れた場所にある最大 96 ノードをサポート・ストレージとともに接続することが可能です。IP クラスタ通信に拡張 LAN 構成を使用することができます。LAN スイッチとブリッジはルータに置き換えることができます。ルータは 2 つ以上の論理サブネットの接続に使用でき、ルータの物理インタフェースを 1 対 1 でマッピングする必要はありません。

9.5.1 複数ノードの IP ベースクラスタ・システム

図 9-10 に示すのは、複数のノードが接続された IP ベースのクラスタ・システムです。各ノードは地理的に異なる場所に分散させることができるため、耐障害性と可用性の点で優れています。

図 9-10 複数ノードの IP ベース・クラスタ・システム



長所

- IP 上のクラスタ通信では、10 Gb/s のスループットを提供する 10 Gigabit Ethernet をサポートします。
- 構成が容易です。
- すべてのノードは他のノードとアクセスすることができ、ストレージへの共有直接アクセスが可能です。

9.5.2 IP ベース・クラスタの構成上のガイドライン

以下に示すのは、IP クラスタ通信を使用してクラスタを構成する場合のガイドラインです。

- リモート・ノードの識別には IP ユニキャスト・アドレスが必要です。
- システム管理者が相手先を指定しクラスタ・グループ番号で動的に計算される IP マルチキャスト・アドレスが必要です。クラスタ構成については、『OpenVMS Cluster システム』を参照してください。
- ネットワーク・マスク・アドレスとともにローカル・マシンの IP アドレスが必要となります。
- IP アドレスを設定し SCS に使用されるローカル LAN アダプタが必要になります。

9.6 I/O のスケーリング

OpenVMS Cluster の拡張性にとって、I/O のスケーラビリティは重要な要素です。OpenVMS Cluster に構成要素を追加する場合、追加した構成要素がボトルネックになったり、OpenVMS Cluster 全体のパフォーマンスが低下するのを防ぐため、高い I/O スループットが必要になります。I/O スループットに影響を与える要素は、以下のとおりです。

- ストレージに対する直接アクセスや MSCP サービスによるアクセス
- MSCP_BUFFER および MSCP_CREDITS システム・パラメータの設定
- Files-11などのファイル・システム・テクノロジー
- 磁気ディスク、半導体ディスク、DECramなどのディスク・テクノロジー
- 読み込み/書き込みの比率
- I/O サイズ
- キャッシュとキャッシュ“ヒット”率
- “ホット・ファイル”管理
- RAID ストライピングとホスト方式のストライピング
- ボリューム・シャドウイング

以上の要素は、I/O スケーラビリティそのものに影響する場合と、スケーラビリティの組み合わせに影響する場合があります。以下の項では、以上の要素について説明するとともに、アプリケーションに変更を加えずに I/O スループットとスケーラビリティを最大限に発揮する方法を説明します。

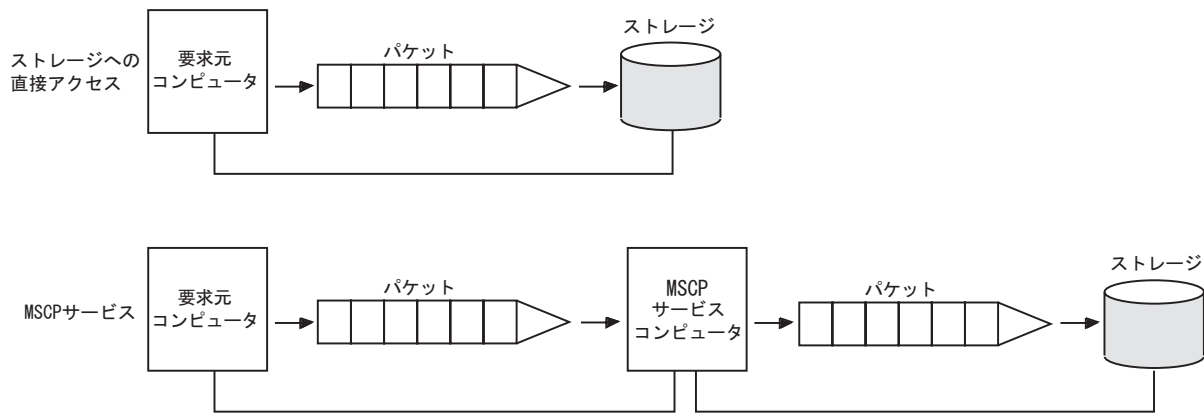
その他、I/O スループットに対しては、インターコネクトやストレージ・サブシステムの種類が影響します。

関連項目：インターコネクトの詳細については第 4 章を，ストレージ・サブシステムの詳細については第 5 章を参照してください。MSCP_BUFFER および MSCP_CREDITS の詳細については、『OpenVMS Cluster システム』を参照してください。

9.6.1 MSCP サービス対象のストレージへのアクセス

MSCP サーバの機能性は，OpenVMS Cluster にとって大きな利点をもたらします。たとえば，ノードとストレージ間の直接通信が可能になります。そして，MSCP サービスによる I/O はオーバーヘッドを引き起こしません。図 9-11 は，サービス・システムによる特別な処理をパケットが要求する仕組みを簡略化したものです。

図 9-11 直接アクセスと MSCP サービスによるアクセスの比較



ZK7032AGE

図 9-11 で，MSCP サービス対象のパケットは，別のシステムによる“中継”を受けてから宛先に到達します。MSCP サービス対象のパケットが宛先ストレージに対応するシステムに達すると，パケットは直接アクセスの場合と同様の処理を受けます。

大量の MSCP サービスを必要とする OpenVMS Cluster では，I/O パフォーマンスやスケーラビリティが低下します。MSCP サービスによる I/O と直接アクセスによる I/O を比較すると，I/O スループットは，全体で約 20%低下します。サテライトがそれぞれのローカル・ストレージを OpenVMS Cluster 全体に提供する方式ではなく，少数の大ノードのストレージを多数のサテライトに提供するように構成してください。

9.6.2 ディスク・テクノロジー

ここ数年、CPU による情報処理能力は、I/O サブシステムによるプロセッサに対するデータ提供能力を越えています。その結果、プロセッサが I/O 処理の終了を待機する時間が増えています。

半導体ディスク (SSD)、DECram、そして RAID レベル 0 は、処理速度と磁気ディスクのアクセス速度のギャップをとりもつテクノロジーです。磁気ディスクのパフォーマンスは、シーク待ち時間と回転待ち時間による制約を受けますが、SSD と DECram はメモリを使用するため、即時アクセスが可能です。

RAID レベル 0 は、1 つのファイルを複数のディスク・ボリュームに分散 (または“ストライピング”) するための技術です。その目的は、アクセス頻度の高いファイルをストライプ・セットにパーティション化し、複数のデバイスに保存することにより、1 つのディスクにボトルネックが発生するのを防ぐことです。この技術では、1 つの I/O を多数のディスクで併行処理することができます。

表 9-5 は、ディスク・テクノロジーとその機能をまとめたものです。

表 9-5 ディスク・テクノロジーのまとめ

ディスク・テクノロジー	特性
磁気ディスク	アクセス時間がかかる。 低費用。 複数のインターコネクタで使用可能。
半導体ディスク	あらゆる I/O サブシステム・デバイス内で最速のアクセス時間。 書き込み頻度が高いファイルの場合に最高のスループット。 複数のインターコネクタで使用可能。
DECram	中小サイズの I/O 要求で最高のスループット。 揮発性ストレージ領域; 一時読み取り専用ファイルに最適。 Alpha システムや VAX システムで利用可能。
RAID レベル 0	HSD コントローラ、HSJ コントローラ、および HSG コントローラで使用可能。

注意: I/O をスケーリングする最も速い方法は、半導体ディスクや DECram に対して共用直接アクセスする場合です。

9.6.3 読み込み/書き込みの比率

I/O をシャドウ・セットに対して行う場合、アプリケーションの読み込み/書き込みの比率が重要です。シャドウ・セットに対する MSCP は、インターコネクタ上でコピーされます。

したがって、100% (100/0) 読み込み処理のアプリケーションでは、シャドウイングにより I/O 処理に複数のパスを利用できるため、ボリューム・シャドウイングの恩恵を受けることができます。読み込み/書き込みの比率が 50/50 のアプリケーションでは、さらにインターコネクタが活用されます。これは、書き込み処理では、I/O を各

シャドウ・メンバに送信する必要があるからです。最も処理速度の遅い I/O を完了するために必要な時間により、遅延が発生する可能性があります。

I/O 読み込み/書き込みの比率は、DCL コマンド MONITOR IO で知ることができます。

9.6.4 I/O サイズ

I/O パケットごとに、プロセッサとメモリにオーバーヘッドが生ずるので、I/O を 1 つのパケットにまとめれば、すべての I/O 処理のオーバーヘッドを節約することができます。アプリケーションで使用できるパケット・サイズが大きければ、スループットが高くなります。パケット・サイズが小さいとオーバーヘッドが大きくなります。

9.6.5 キャッシュ

キャッシュとは、最新のデータや使用頻度の高いデータを、メモリ、コントローラ、またはディスク内のアクセスしやすい場所に保存する技術です。キャッシュは半導体ディスク、DECram、RAID を補完する技術です。特別なコーディングをしなくてもアプリケーションはキャッシュの利点をそのまま利用できます。キャッシュすると、OpenVMS Cluster システム内で構成要素間の I/O 数が減少し、現在の I/O ボトルネックや潜在的な I/O ボトルネックを緩和できます。

表 9-6 は、3 種類のキャッシュ方式をまとめたものです。

表 9-6 キャッシュの種類

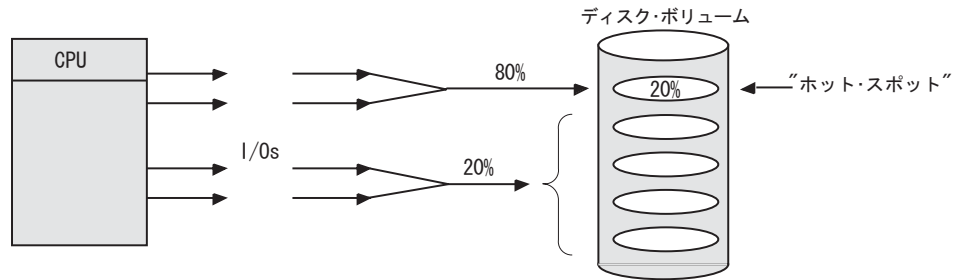
キャッシュの種類	説明
ホスト方式	ホスト・システムのメモリにキャッシュが常駐。ホストからの I/O をサービス。
コントローラ方式	ストレージ・コントローラにキャッシュが常駐。すべてのホストにデータをサービス。
ディスク	ディスクに常駐するキャッシュ。

ホスト方式のディスク・キャッシュの特長は、コントローラ方式やディスク方式のキャッシュとは異なります。ホスト方式のディスク・キャッシュでは、キャッシュそのものをノード間で共用できません。コントローラ方式やディスク方式のキャッシュは共用できます。これは、共用できるコントローラやディスクにキャッシュがあるためです。

9.6.6 "ホット"・ファイルの管理

“ホット”・ファイルは、システム内でほとんどの処理が発生する場所です。多くの環境でホット・ファイルが存在するのは、全 I/O のおよそ 80% がデータの 20% に集中するためです。つまり、図 9-12 に示すように、ディスク・ドライブで領域は均等に分散しているのに、転送されるデータの 80% がディスクの 1 個所に集中するためです。

図 9-12 ホット・ファイルの分散



I/O のスケーラビリティを強化するには、管理が不適切だとボトルネックになりかねないホット・ファイルに着目します。この領域の処理は、I/O、MB 単位の転送データ量、キュー深度で表します。

RAID レベル 0 では、1 つのファイルを複数のディスクに分散してホット・ファイルのバランスを調整します。これにより、パフォーマンスに対するホット・ファイルの影響を緩和できます。

ホット・ファイルによる処理の解析には、以下の DCL コマンドを使用します。

- MONITOR IO コマンド — ホット・ディスクを監視します。
- MONITOR MSCP コマンド — MSCP サーバを監視します。

ホット状態のディスクやサーバを探すには、MONITOR IO コマンドと MONITOR MSCP コマンドを使用します。

9.6.7 ボリューム・シャドウイング

Volume Shadowing for OpenVMS 製品では、複数のディスクにデータのコピーを作成してアプリケーションとエンド・ユーザに対するデータの可用性を強化します。ボリューム・シャドウイングでは、データの冗長性と可用性が高くなるものの、以下の 2 つのレベルで OpenVMS Cluster I/O に影響が生じることがあります。

スケーラビリティを目的とした OpenVMS Cluster の構成

9.6 I/O のスケーリング

要因	影響
物理的な距離	ホスト方式のボリューム・シャドウイングでは、MSCP サーバからサービスを受けるボリュームも含め、OpenVMS Cluster システム内の任意のデバイスのシャドウイングが可能です。この機能では、長距離通信が可能です。MSCP のオーバーヘッドも生じます。たとえば、IP を使用する OpenVMS Cluster システムは、500 マイル離れて配置することができます。Fibre Channel を使用すると、最大 100 km (62 マイル) 離れて設置することができます。距離と MSCP の相関により I/O スループットが低下する場合があります。
読み込み/書き込みの比率	シャドウイングでは、データが複数のボリュームに書き込まれるので、書き込みの激しいアプリケーションでは、シャドウイングをするソフトウェアは、スループットが低下します。これは最も効率よくデータを取り出せるディスク・メンバを選択するからです。

OpenVMS Cluster システム管理の手法

この章では、OpenVMS Cluster をフルに活用するための主なシステム管理の手法について説明します。OpenVMS Cluster の一般的なシステム管理の実践については、『OpenVMS Cluster システム』を参照してください。

この章では、システム・ディスク、クォーラム・ディスク、OpenVMS Cluster 遷移など、システム管理の基本的な概念を理解していることを前提としています。

この章の構成は、以下のとおりです。

- システム・ディスクの手法
- 共通のマルチ環境
- クォーラムの手法
- 状態遷移の手法
- 同じ OpenVMS Cluster に複数の OpenVMS のバージョン
- 1 つの OpenVMS Cluster に複数のプラットフォーム (Integrity システムおよび Alpha システム)

10.1 シンプルな構成と複雑な構成

OpenVMS Cluster ソフトウェアは、大半のシステム管理のタスクが 1 回ですむため、システム管理者の仕事が軽減されます。このことは、OpenVMS Cluster の全機能を必要とせず、シンプルな構成で対応できるビジネス要件において特にはっきりとした効果が出ます。シンプルな構成は、新しくシステム管理者になった人だけでなく、経験豊かなシステム管理者にとっても魅力的であり、ノードが 3 個から 7 個、ユーザが 20 人から 30 人、ストレージが 100 GB 程度の小規模な OpenVMS Cluster に最適です。

複雑な OpenVMS Cluster 構成の場合で、可用性、スケーラビリティ、パフォーマンスを実現するには、高度なシステム管理方法が必要です。

関連項目：複雑な OpenVMS Cluster 構成の例については、図 10-2 を参照してください。

システム管理の選択にあたっては、簡潔性と、複雑な OpenVMS Cluster で要求されるその他の管理タスクとの調和を考慮したシステム管理方法を選択してください。

10.2 システム・ディスクの手法

システム・ディスクには、システム・ファイルと環境ファイルが含まれています。

システム・ファイルは、ランタイム・ライブラリなど主に読み込み専用のイメージとコマンド・プロシージャであり、クラスタ規模でアクセスします。

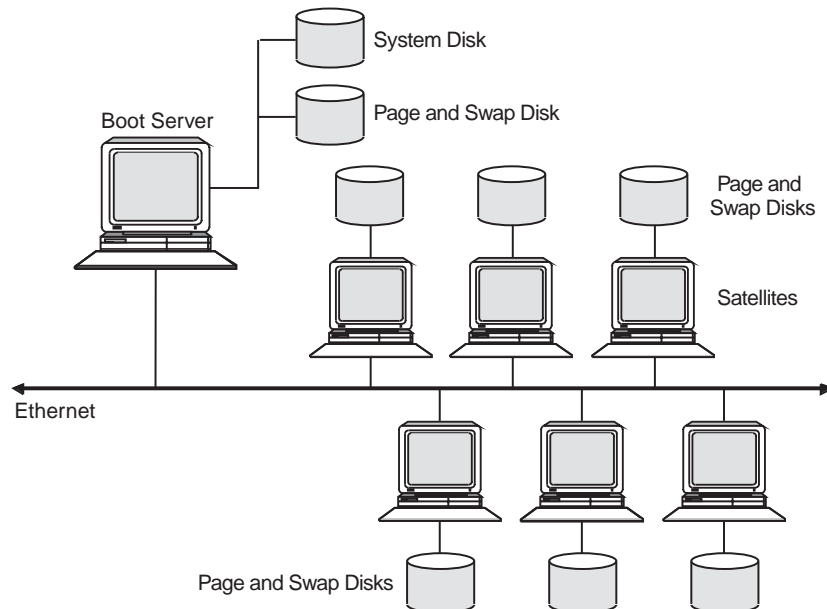
環境ファイルは、個々のユーザに合った環境を作成するためのファイルです。すべての環境ファイルをクラスタ規模でアクセスできるようにすれば共通環境ができます。また、特定の環境ファイルについてアクセスできるユーザやシステムを限定すればマルチ環境になります。

10.2.1 シングル・システム・ディスク

シングル・システム・ディスクと共通環境という簡潔な構成の場合は、システム管理も簡単です。必要なほとんどの作業は1回で済みます。また、システム・ファイルと環境ファイルは同じディスクに常駐します。ページ・ファイルとスワップ・ファイルもシステム・ディスクに常駐します。

図 10-1 は、共通環境を備えたシンプルな OpenVMS Cluster の別の例です。

図 10-1 シングル・システム・ディスクのシンプルな LAN OpenVMS Cluster



ZK7197AGE

図 10-1 では、6 つのサテライトと 1 つのブート・サーバを Ethernet で接続しています。各サテライトには、専用のページ・ディスクとスワップ・ディスクがあり、これでシステム・ディスクの領域を節約するとともに、Ethernet からのページ・ファイルとスワップ・ファイルの I/O 処理を取り除きます。システム・ディスクからページ・ファイルとスワップ・ファイルを取り除いたことで、OpenVMS Cluster のパフォーマンスが強化できます。

シングル・システム・ディスクの構成で、多くの OpenVMS Cluster 要件に対応できますが、マルチ・システム・ディスクにもいくつかの長所があります。

10.2.2 マルチ・システム・ディスク

Integrity システムと Alpha システムの両方を組み込んだ OpenVMS Cluster には、複数のシステム・ディスクすなわち OpenVMS Integrity のシステム・ディスクと Alpha システム・ディスクが必要です。表 10-1 で、(アーキテクチャに関係なく) システム管理者が OpenVMS Cluster で複数のシステム・ディスクを必要とする理由を補足説明します。

表 10-1 マルチ・システム・ディスクの長所

長所	説明
短縮されたブート時間	<p>3 つ以上のシステムを同時にブートしようとする時、シングル・システム・ディスクでは、ボトルネックになる可能性があります。ブート時間は以下の要素によって大きく異なります。</p> <ul style="list-style-type: none"> • LAN の利用率 • システム・ディスクの処理速度 • マウントされているディスクの数 • インストールされているアプリケーションの数 • ブート・ノードとサテライトの距離 • ブート・ノードの処理能力 • システム・ディスクに環境ファイルがあるか • システム・ディスクがシャドウ化されているか <p>書き込み処理の負荷が高い環境ファイル (SYSUAF.DAT など) がシステム・ディスクにない場合、Volume Shadowing for OpenVMS ソフトウェアは、ディスクの読み込みパフォーマンスの強化に有効です。</p>

(次ページに続く)

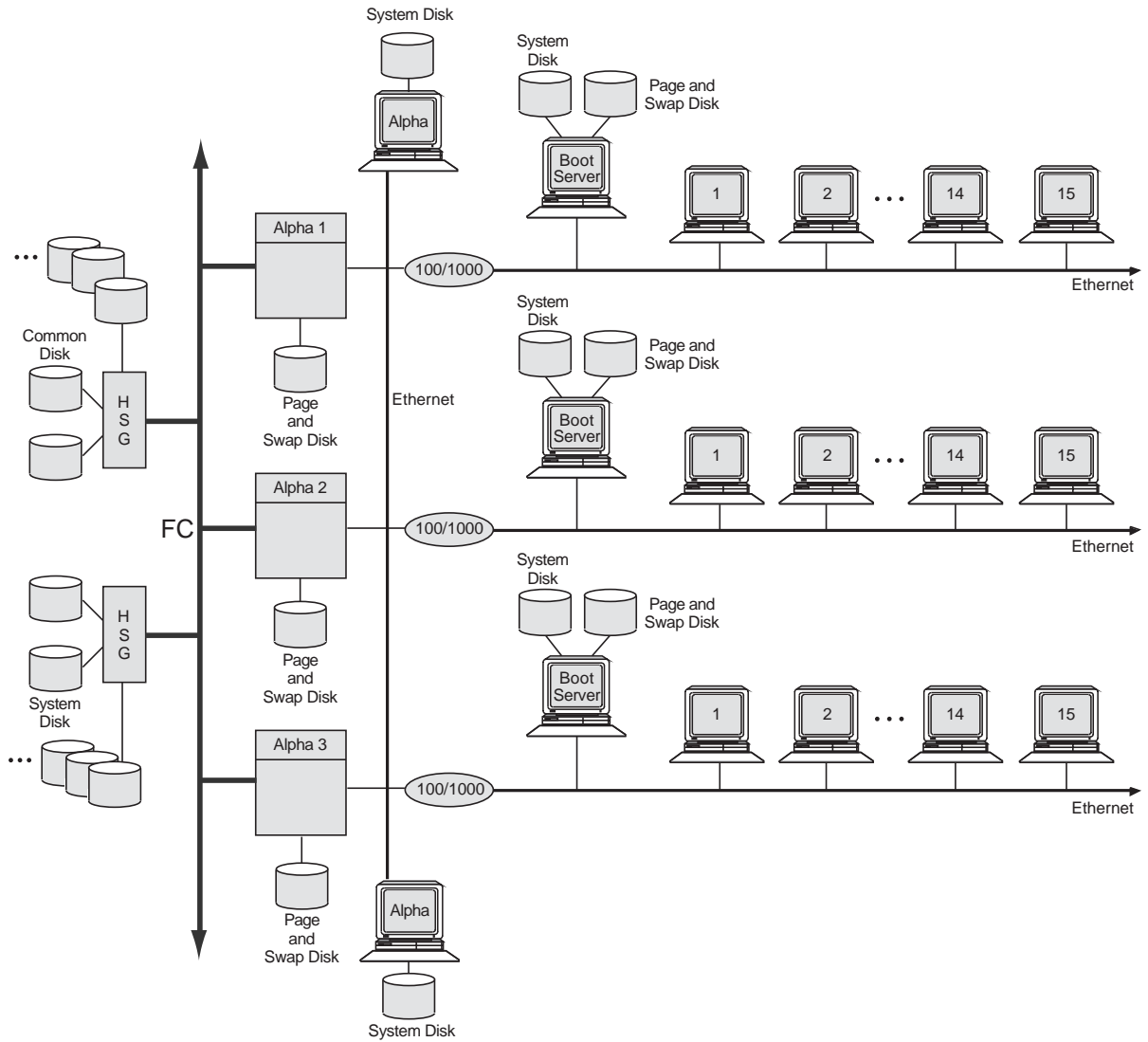
表 10-1 (続き) マルチ・システム・ディスクの長所

長所	説明
強化されたシステムとアプリケーションのパフォーマンス	<p>一定以上の使用頻度の各種アプリケーションを搭載した OpenVMS Cluster では、ノードごとにローカル・システム・ディスクを設けるか、システム・ディスクが担当するシステムを少数に限定するとパフォーマンス面で効果的です。長所は、イメージ起動時間が短く、LAN のサービス対象になるファイルが少ない点です。</p> <p>Alpha ワークステーションは、ローカル・システム・ディスクに適しています。強力な Alpha プロセッサでは、システム・ディスク・アクセス程待たなくてもすむためです。</p> <p>関連項目：詳細については、第 9.4.5 項を参照してください。</p>
LAN の利用率の削減	<p>システム・ディスクによる LAN の利用率が下がれば、LAN によるサービス対象のファイルも減ります。LAN セグメントとそのブート・サーバを、セグメント外の不要なトラフィックから分離すれば、LAN パス接続も少なくて済みます。</p> <p>関連項目：詳細については、第 10.2.4 項を参照してください。</p>
OpenVMS Cluster 可用性の増加	<p>シングル・システム・ディスクは、単一点障害の要因になる可能性があります。ブート・サーバとシステム・ディスクを増加すれば、シングル・リソースにおける OpenVMS Cluster に対する依存度が下がり、可用性が高くなります。</p>

10.2.3 マルチ・システム・ディスク OpenVMS Cluster

図 10-2 のようにシステム・ディスクを配置すると、ブート時間と LAN の利用率を下げることができます。

図 10-2 共通環境におけるマルチ・システム・ディスク



ZK7036AGE

図 10-2 は、マルチ・システム・ディスクの OpenVMS Cluster です。

- Alpha 1, Alpha 2, Alpha 3 に各 1 つずつのページおよびスワップ・ファイル・ディスク
- LAN セグメント上のブート・サーバに各 1 つずつのシステム・ディスク

この構成におけるマルチ・システム・ディスクと、LAN セグメントの分割使用形態は、効率的でタイミングの良いブート・シーケンスを可能にします。

10.2.4 OpenVMS Cluster システムの分割

第 9.4 節に示すワークステーション・サーバの例で、障害後の OpenVMS Cluster のリポートは比較的簡潔です。これはサーバ当たりのサテライト数が少ないためです。ただし、図 10-2 に示す大規模な OpenVMS Cluster 構成のリポートは慎重に計画する必要があります。この項で説明する OpenVMS Cluster の分割とシステム・ディスクの配置により、ブート時間を大幅に節約できます。OpenVMS Cluster を分割すると、LAN セグメントのサテライトの利用率も低下し、サテライト・パフォーマンスが向上します。

この OpenVMS Cluster のディスクには、表 10-2 に示すように決められた機能があります。

表 10-2 マルチ・システム・ディスクの使用法

ディスク	内容	目的
共通ディスク	OpenVMS Cluster 全体のすべての環境ファイル	SYSUAF.DAT, NETPROXY.DAT, QMAN\$MASTER.DAT などの環境ファイルは、ブート時に、サテライトを始めすべてのノードがアクセスできます。これにより、サテライト・ブート・サーバは、システム・ファイルとサテライトまでのルート情報だけを扱うこととすむようになります。 共通環境を作成し、すべてのシステム・ディスクのパフォーマンスを強化する方法については、第 10.3 節を参照してください。
システム・ディスク	Alpha 1, Alpha 2, Alpha 3 のシステム・ルート	サーバ・システムのパフォーマンスの強化。書き込み処理をする環境ファイルをこのシステム・ディスクから取り除き、このディスクをできるだけ読み込み専用にします。ディスクは、スタートアップ時に、SYLOGICALS.COM にクラスタ規模でマウントできます。
サテライト・ブート・サーバのシステム・ディスク	サテライト用のシステム・ファイルまたはルート	Alpha 1, Alpha 2, Alpha 3 に関連付けられたシステム・ディスクを解放し、サテライトのサービス専用とし、各 Ethernet セグメントの全 LAN トラフィックを分割します。
ページ・ディスクとスワップ・ディスク	1 つ以上のシステムのページ・ファイルとスワップ・ファイル	システム・ディスクの I/O 処理を削減し、システム・ディスクの領域をアプリケーションとシステム・ルートに開放します。

図 10-2 に示す構成用のブート・シーケンスでは、共通ディスクのファイルをサテライトがアクセスできるよう、LAN Ethernet セグメントをブートする前に、ノード Alpha 1, Alpha 2, Alpha 3 がすべてブートされているか確認してください。FDDI 構成の場合は、サテライトが Alpha 1, Alpha 2, Alpha 3 のシステム・ディスクからブートしないよう、Ethernet-to-FDDI (10/100) ブリッジで MOP(Maintenance Operations Protocol) のフィルタリングを可能にしておきます。この OpenVMS Cluster をブートする順序は、次のとおりです。

1. Alpha 1, Alpha 2, Alpha 3 をブートする。
2. サテライトブート・サーバをブートする。
3. すべてのサテライトをブートする。

関連項目：拡張 LAN については、第 9.4.7 項を参照してください。

10.2.5 まとめ：シングル・システム・ディスク対マルチ・システム・ディスク

表 10-3 を参考に，OpenVMS Cluster 全体で 1 つのシステム・ディスクとするかマルチ・システム・ディスクを使用するか決定してください。

表 10-3 シングル・システム・ディスクとマルチ・システム・ディスクの比較

シングル・システム・ディスク	マルチ・システム・ディスク
システム・ディスクのファイルにアクセスするまでの待機時間が長くなることもある。	長く待たなくてもシステム・ディスクのファイルにアクセスでき，プロセッサ・パフォーマンスが高速である。
シングル・リソースの競合が多い。	シングル・リソースの競合が少ない。
サテライトのブート時間が長い。	サテライトのブート時間が短い。
管理するシステム・ディスクは 1 つだけである。	複数のシステム・ディスクを管理しなければならない。
システム管理がシンプルである。	システム・パラメータやファイルのクラスタ規模での調整など，システム管理が複雑である。
ハードウェアとソフトウェアのコストが経済的である。	ハードウェアとソフトウェアにより多くのコストがかかる。特にディスクをシャドウ化した場合。
シングル・システム・ディスクの管理には時間と経験があまり要求されないためシステム管理のコストが経済的である。	マルチ・システム・ディスクの管理にはより多くの時間と経験が要求されるためシステム管理により多くのコストが要求される。

10.3 OpenVMS Cluster 環境の手法

処理上のニーズに合わせて，共通環境とマルチ環境のどちらかを選択します。前者では，環境ファイルがクラスタ規模で共用され，後者では，クラスタ規模で共用されるファイルと，一定の OpenVMS Cluster メンバだけがアクセスできるファイルに分かれます。

以下に示すのは，最も頻繁に使用および操作される OpenVMS Cluster 環境ファイルです。

```

SYS$SYSTEM:SYSUAF.DAT
SYS$SYSTEM:NETPROXY.DAT
SYS$SYSTEM:VMSMAIL_PROFILE.DATA
SYS$SYSTEM:NETNODE_REMOTE.DAT
SYS$MANAGER:NETNODE_UPDATE.COM
SYS$SYSTEM:RIGHTSLIST.DAT
SYS$SYSTEM:QMAN$MASTER.DAT

```

関連項目：これらのファイルの管理方法についての詳細は，『OpenVMS Cluster システム』を参照してください。

10.3.1 共通環境

共通 OpenVMS Cluster 環境は、OpenVMS Cluster の全ノードで共通な操作環境です。共通環境は、システム・ファイルを共通バージョンで使用できるため、マルチ環境より簡単に管理できます。環境は、以下のようにセットアップします。

- すべてのノードが同じプログラム、アプリケーション、ユーティリティを実行する。
- すべてのユーザが同じ種類のアカунトを持ち、同じ論理名が定義される。
- すべてのユーザがストレージとキューのアクセスを共用する。
- すべてのユーザが構成のどのノードでもログインでき、他のすべてのユーザと同じ環境で作業できる。

最もシンプルで経済的な環境は、すべての環境ファイルを搭載した 1 つのシステム・ディスクで、OpenVMS Cluster に対応する方法です。この方法の長所は、以下のとおりです。

- ソフトウェア製品のインストールが 1 回ですむ。
- すべての環境ファイルが、同じシステム・ディスクにあり、検索および管理が簡単。
- ブートの依存関係が明快である。

10.3.2 別々の共通ディスクに環境ファイルを配置

すべてのノードが同じシステム・ディスクを共用する OpenVMS Cluster の場合、ほとんどの共通環境ファイルは SYSSSYSTEM ディレクトリに配置します。

ただし、環境ファイルを別のディスクに移動して、OpenVMS Cluster のパフォーマンスを強化することもできます。一般に、環境ファイルの処理はシステム・ディスクの処理の 80% を占めるため、常駐ディスクを分散させればシステム・ディスクの負担が減ります。図 10-2 は、分割された共通ディスクの例です。

SYSUAF.DAT などの環境ファイルを別の共通ディスクに移動すると、SYSUAF.DAT は、SYSSSYSTEM:SYSUAF.DAT のデフォルト・パスでは見つからなくなります。

関連項目：OpenVMS Cluster の全ノードが新しい場所にある SYSUAF.DAT にアクセスできるようにする方法については、『OpenVMS Cluster システム』を参照してください。

10.3.3 マルチ環境

マルチ環境は、ノードによって異なります。個々のノードやノード・サブセットは以下の目的でセットアップします。

- ユーザが実行するタスクの種類や必要とするリソースに応じてマルチ・アクセスを提供する。
- 他のノードでは利用できないリソース・セットを共有する。
- 一般的なタイムシェアリング・ワークは他のプロセッサに実行させ、機能を限定したリソースで特殊機能を実行する。
- ログインしたノードに固有の環境で作業する。

図 10-3 は、マルチ環境の例です。

図 10-3 マルチ環境 OpenVMS Cluster

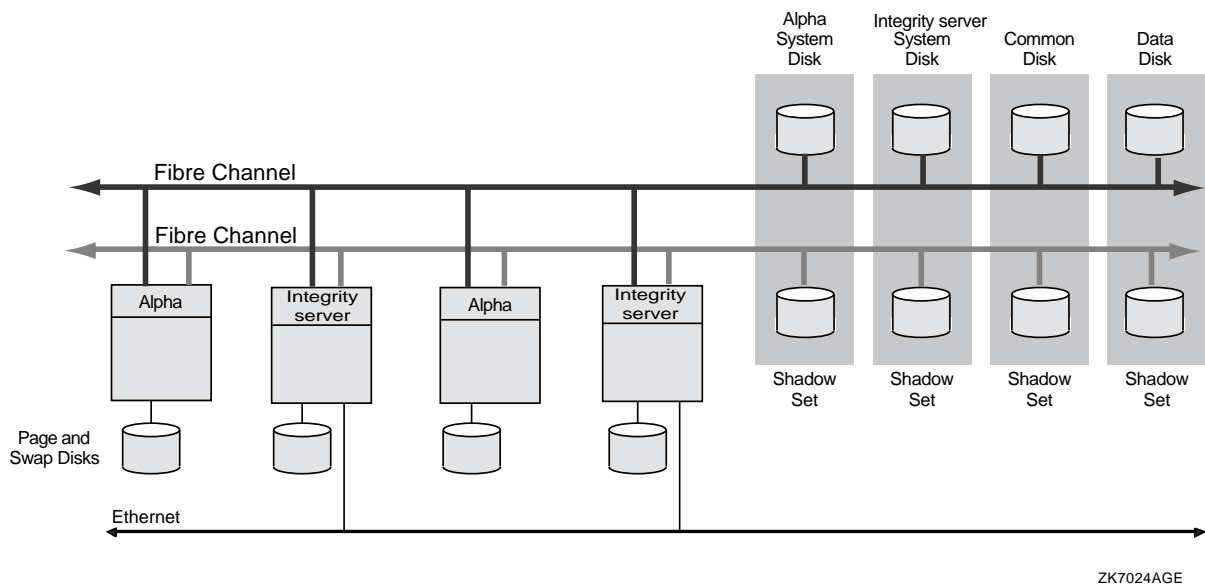


図 10-3 のマルチ環境 OpenVMS Cluster は、2 つのシステム・ディスクからなります。1 つは Integrity ノード用であり、他の 1 つは Alpha ノード用です。共通ディスクには、各ノードやノード・グループ用の環境ファイルがあります。OpenVMS Cluster システム管理者は、一般にシンプルなシングル (共通) 環境を好みますが、すべてのノードがリソースを共有するわけではないマルチ環境の作成時には環境ファイルの多重化が必要です。環境は、それぞれユーザが実行するタスクの種類と使用するリソースに応じて多様化でき、その構成にはさまざまなアプリケーションをインストールできます。

4 個の Fibre Channel ノードのそれぞれが専用のページ・ディスクとスワップ・ディスクを持ち、FC インターコネクト上の Alpha システムと Integrity システム・ディスクにおけるページ処理とスワップ処理の負荷を軽減しています。FC インターコネクト上のディスクはすべてシャドウ化され、障害の発生時にもディスクを保護することができます。

10.4 その他のマルチ環境の手法

この節では、複数の SYSUAF.DAT ファイルや複数のキュー・マネージャの使用方法など、その他のマルチ環境の手法について説明します。

10.4.1 複数の SYSUAF.DAT ファイルの使用

ほとんどの OpenVMS Cluster は、1 つのユーザ権限 (SYSUAF.DAT) ファイルで管理しますが、複数のユーザ権限ファイルを利用すれば、ユーザがアクセスできるシステムを限定することができます。この方法では、すべてのシステムのアクセスを必要とするユーザには複数のパスワードが必要になります。

複数の SYSUAF.DAT ファイルを運用する場合は、セキュリティに気をつけてください。OpenVMS Integrity および OpenVMS Alpha オペレーティング・システムでは、複数のセキュリティ・ドメインをサポートしていません。

関連項目：SYSUAF.DAT エントリなど、シングル・セキュリティ・ドメイン内では同一でなければならないフィールドの一覧については、『OpenVMS Cluster システム』を参照してください。

Alpha システムでは、多くのプロセス割り当て量が必要なため、システム管理者は複数の SYSUAF.DAT ファイルを用意して対応することがあります。しかし、これは最適な解決方法とは言えません。複数の SYSUAF.DAT ファイルを作成するのは、ノードによって環境を変えるためであり、プロセス割り当て量を増やすのが本来の目的ではありません。プロセス割り当て量を増やす方法としては、SYSUAF.DAT ファイルを 1 つとし、SYSUAF.DAT ファイルで指定したプロセス割り当て量を無効にするシステム・パラメータを使用して Alpha システムのリソースを管理する方法をお勧めします。

10.4.2 複数のキュー・マネージャの使用

OpenVMS Cluster におけるバッチ・トランザクションとプリント・トランザクションの数が多すぎて混雑している場合、複数のキュー・マネージャを実装してバッチとプリントの負荷をノード間に分散させることができます。

OpenVMS Cluster には、QMAN\$MASTER.DAT ファイルを 1 つずつ配置します。複数のキュー・マネージャは、複数の*.QMAN\$QUEUES ファイルと*.QMAN\$JOURNAL ファイルで定義します。キュー・マネージャ・ファイルの各ペアを別々のディスクに配置します。QMAN\$MASTER.DAT ファイルにアクセスが集中するようであれば、半導体ディスク (SSD) に常駐させて、OpenVMS Cluster が処理できるバッチ・トランザクションとプリント・トランザクションの数を増やします。たとえば、バッチ・キューとプリント・キューに別々のキュー・マネージャを作成できます。

関連項目：複数のキュー・マネージャの実装例とコマンドについては、『OpenVMS システム管理者マニュアル (上巻)』を参照してください。

10.5 クォーラムの手法

OpenVMS Cluster システムは、クォーラム・アルゴリズムを利用してストレージのアクセスを同期します。クォーラム・アルゴリズムとは、OpenVMS Cluster システムでリソースを共用する方法を“決定する (ポートする)”多数の OpenVMS Cluster メンバが存在するかどうかを確かめる数学的手法です。接続マネージャは、動的な値としてクォーラムを計算し、OpenVMS Cluster メンバの大半が機能しているときにだけ処理を実行します。

クォーラム・ポーツを決定する要素は次のとおりです。

- パラメータ VOTES がゼロを超える値に設定してあるシステム
- クォーラム・ディスクと呼ばれる所定のディスク

各 OpenVMS Cluster システムには、クォーラム・ディスクを 1 つだけ配置できます。クォーラム・ディスクはシャドウ・セットのメンバにはできませんが、システム・ディスクにすることはできます。

接続マネージャは、“クォーラム・ディスク・ウォッチャ”からクォーラム・ディスクの情報を入手します。クォーラム・ディスク・ウォッチャは、クォーラム・ディスクとの間にアクティブな直接接続を持つ、任意のシステムです。

10.5.1 クォーラム手法のオプション

クォーラム・ディスクには、少なくとも直接接続を持つシステムが 2 つ必要です。これによってシステムのどちらかに障害が発生しても、クォーラム・ディスク・ポーツにアクセスできます。

クォーラム手法を検討するにあたり、どのような障害状況のときに OpenVMS Cluster の処理を続行するかを決めておく必要があります。表 10-4 は、その 4 つのオプションを示したものです。

表 10-4 クォーラムの手法

手法のオプション ¹	説明
“予想”最高ノード数が残っていれば、処理を続行する。	すべてのノードにブーツを与え、クォーラム・ディスクは使用しません。この手法では、ノードが 3 個以上必要です。
(3 つ以上のノードの内) ノードが 1 つでも残っていれば、処理を続行する。	この手法では、クォーラム・ディスクが必要です。 全システムのブーツ合計数より 1 つ少ない値までクォーラム・ディスクのブーツ数を増やすと (また、EXPECTED_VOTES システム・パラメータの値も同じだけ増やす)、ノードが 1 つだけのクラスタをクォーラム・ディスク・ウォッチャとして実行できます。この場合、ブーツ・システムの過半数になるまで待たなくても OpenVMS Cluster システムを利用できます。
(2 ノード OpenVMS Cluster) ノードが 1 つでも残っていれば、処理を続行する。	各ノードとクォーラム・ディスクにブーツを与えます。 2 ノード OpenVMS Cluster は、特別な事例です。クォーラム・ディスクを確立すると、2 ノード OpenVMS Cluster の可用性を強化できます。このような構成では、クォーラム・ディスクまたは 1 ノードが故障してもクォーラムを維持し運用を続行することができます。この場合、両方のノードが、両方もクォーラム・ディスク・ウォッチャになるように、(CI, DSSI, SCSI, または Fibre Channel によって) ストレージに直接接続される必要があります。
OpenVMS Cluster の重要ノードが 1 つでも処理を続行する。	一般に、この手法では、サーバにはブーツが与えられ、サテライトには与えられません。3 つ以上のサーバがあり、クォーラム・ディスクがないことが前提です。

¹これらの手法は、互いに排他的です。どれか 1 つだけを選択してください。

関連項目：クォーラム・ディスク管理の詳細については、『OpenVMS Cluster システム』を参照してください。

10.6 状態遷移の手法

OpenVMS Cluster の状態遷移は、OpenVMS Cluster システムに対してシステムを追加したり削除したとき、そして OpenVMS Cluster がクォーラム・ディスクの状態遷移を検出すると発生します。接続マネージャは、これらのイベントを受けて OpenVMS Cluster 内のデータの一貫性を維持するように対応します。

状態遷移が問題になるのは、システムが OpenVMS Cluster システムに対して頻繁に追加あるいは削除され、それによって障害が発生する場合です。

ユーザとアプリケーションに対する状態遷移の間隔と効果は、遷移の理由、構成、使用中のアプリケーションによって決まります。遷移を有効に管理すれば、システム・マネージャは以下の対応ができます。

- 障害の検出と遷移の所要時間
- ボリューム・シャドウイングのコピー、マージ処理など、遷移の影響

10.6.1 状態遷移の処理

以下のガイドラインは、実際の遷移時間を短縮し、遷移後の影響を最小限にするための方法をまとめたものです。

- ノードと OpenVMS Cluster 間の接続が切れないように以下の方法で対応してください。
 - システム全体でインターコネクに冗長性を与える。
 - インターコネク、プロセッサ、アダプタの容量不足だけでなく、ディスクとメモリのリソース欠乏を防ぐ。
 - 無停電電源装置 (UPS) を使用する。
 - 大規模 OpenVMS Cluster におけるワークステーションのシャットオフでクラスタ内の全システムの操作に支障が生じないようにユーザに通知する。
- OpenVMS Cluster のノード数が 2 個でない限り、クォーラム・ディスクを使用しない。
- 可用性を強化できるように、シャドウ・セット・メンバはできるだけ共用バスに常駐するようにする。
- ノード、ディスク、アダプタ、インターコネク、仮想回路の障害を検出する時間は、システム・ポーリング・パラメータで制御します。ポーリング時間を短縮すると、変化に対するクラスタの応答は良くなりますが、一時的な障害に弱くなります。タイマの設定にあたっては、重大障害からの迅速な回復と一時的障害に対する“過敏さ”とのバランスを考えてください。

表 10-5 は、検出時間を短縮する際に設定する OpenVMS Cluster ポーリング・パラメータをまとめたものです。これらのパラメータの値は、各 OpenVMS Cluster メンバと同じ値に設定することをお勧めします。

表 10-5 OpenVMS Cluster ポーリング・パラメータ

パラメータ	説明
QDSKINTERVAL	クォーラム・ディスク・ポーリング間隔を指定します。
RECNXINTERVL	接続マネージャが別のシステムとの通信を復元するときの所要時間を指定します。
TIMVCFAIL	仮想回路障害の検出に必要な時間を指定します。

- アプリケーション回復時間を計画に入れておいてください。アプリケーション・ユーザに対する状態遷移の影響を評価するときは、アプリケーション回復フェーズに、ジャーナル・ファイルの再生、回復ユニットのクリーニング、再度ログインするユーザのことを考慮してください。

関連項目： OpenVMS Cluster 遷移とそのフェーズ、システム・パラメータ、クォーラム管理の詳細については、『OpenVMS Cluster システム』を参照してください。

10.7 マルチ・バージョンの移行サポートと保証サポート

HP は、バージョン混在およびアーキテクチャ混在環境の OpenVMS Cluster システムに対し、移行サポートと保証サポートの 2 つのレベルのサポートを提供します。

保証サポートとは、OpenVMS Cluster 上に異なるバージョンが共存することを認めているものであり、これらの構成を使用している利用者から寄せられた問題にはすべて回答します。保証されている構成については、ソフトウェア仕様書を参照してください。

移行サポートは、カスタマのクラスタ環境で、保証されている OpenVMS Cluster 構成へ最小限の影響で移行するのを支援するためのサポートです。移行サポートでは、新しいバージョンの OpenVMS Alpha や OpenVMS Integrity に段階的に移行していくユーザのために、ともに使用できるバージョンを明確に示します。HP は、これらの構成に対して提出される問題レポートに回答します。ただし、例外的にソリューションの一環として、保証されている構成に移行するようユーザに依頼する場合があります。

混成バージョンのクラスタでは、OpenVMS の旧バージョンに修正キットをインストールする必要があります。

OpenVMS Cluster インターコネクトとしての SCSI

OpenVMS Cluster ストレージ・インターコネクトに接続されたストレージを複数のコンピュータが同時にアクセスできるのは、OpenVMS Cluster システムの長所の 1 つです。さらに、OpenVMS Cluster システムについては、パフォーマンスとストレージのアクセスの高可用性が特長です。

この付録では、OpenVMS Cluster システムにおけるストレージ・インターコネクトとしての SCSI (Small Computer System Interface) のサポート方法を説明します。SCSI インターコネクトを通じて、複数の Alpha コンピュータ (ホストまたはノードとも呼ぶ) から SCSI ディスクに同時にアクセスできます。このような構成を SCSI マルチホスト OpenVMS Cluster と言います。SCSI インターコネクト (SCSI バスとも呼ぶ) は、業界標準インターコネクトであり、1 つ以上のコンピュータ、周辺機器、相互接続の構成要素をサポートします。

注意

OpenVMS Alpha システムでは、最新の SCSI アダプタはサポートしていません。

この付録での説明は、OpenVMS Cluster 環境におけるストレージ・リソースの概念を理解していることが前提となっています。OpenVMS Cluster の概念と構成の要件については、以下の OpenVMS Cluster マニュアルでも説明しています。

- 『OpenVMS Cluster システム』
- 『OpenVMS Cluster Software Software Product Description』 (SPD 29.78.xx)

この付録は、次の 2 つの項目に分かれています。

- 第 A.1 節から第 A.6.6 項までは、SCSI マルチホスト OpenVMS Cluster システムの計画と実装に必要な基本的な手順と概念を説明しています。
- 第 A.7 節は、技術的な内容の詳細と概念を説明しています。

A.1 この付録で使用する表記法

この付録では、ANSI 標準と図で使用する要素を、特定の表記法に従って表記します。

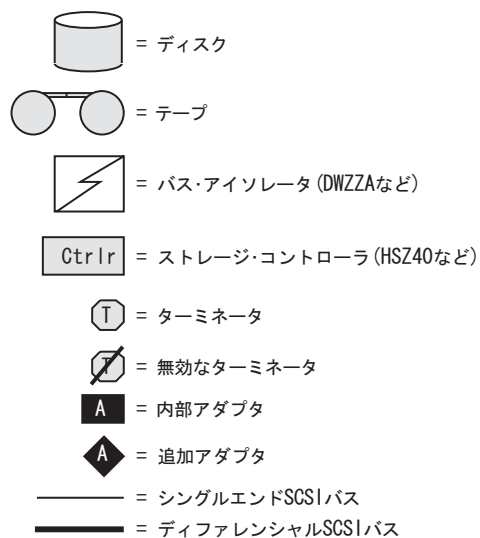
A.1.1 SCSI ANSI 標準

SCSI インターコネクタで構成した OpenVMS Cluster システムでは、標準的な SCSI-2 構成要素または SCSI-3 構成要素を使用します。SCSI-2 構成要素は、American National Standards Institute (ANSI) Standard SCSI-2, X3T9.2, Rev. 10L で定義したアーキテクチャに準拠しています。SCSI-3 構成要素は、SCSI-3 Architecture およびコマンド標準の認定バージョンに準拠しています。説明の便宜上、この付録では、SCSI-2 と SCSI-3 のどちらも SCSI と表記します。

A.1.2 図中で使用する記号

図 A-1 は、この付録の図中で使用する記号です。

図 A-1 図中で使用する記号の意味



ZK7759AGE

A.2 SCSI ストレージのアクセス

OpenVMS Cluster 構成では、以下の方式で、複数の VAX ホストと Alpha ホストを SCSI デバイスに直結できます。

- HSJ コントローラまたは HSC コントローラ付きの CI インターコネクタ
- HSD インターコネクタ付きの Digital Storage System Interconnect (DSSI)
- VAX システムまたは Alpha システムに直結した SCSI アダプタ

SCSI デバイスは、OpenVMS MSCP サーバで間接的にアクセスすることもできます。

以下の項では、SCSI ストレージ・デバイスに対するシングル・ホストとマルチホストのアクセスについて説明します。

A.2.1 OpenVMS Cluster システムにおけるシングル・ホスト SCSI アクセス

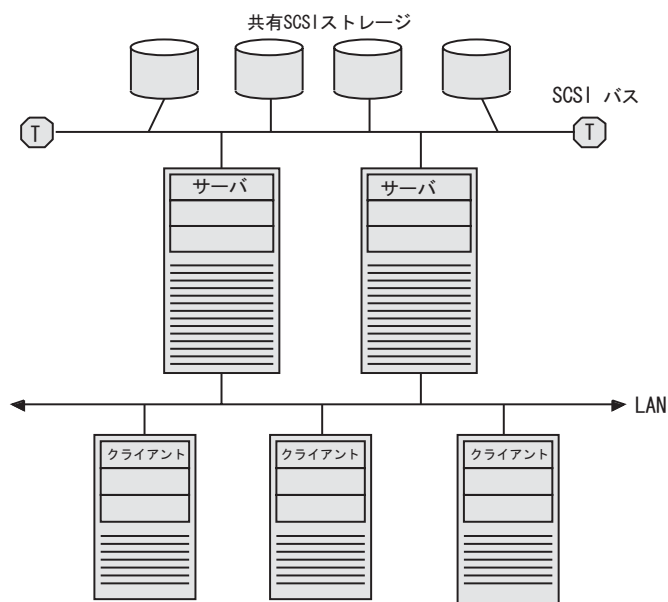
OpenVMS バージョン 6.2 より前のバージョンでは、OpenVMS Cluster システムは、埋め込み SCSI アダプタ、またはオプションの外部 SCSI アダプタ、あるいは専用の RAID (redundant arrays of independent disks) コントローラで単一ホストに接続された SCSI ストレージ・デバイスをサポートしていました。この場合、SCSI バスに接続できるホストは 1 つだけでした。

A.2.2 OpenVMS Cluster システムにおけるマルチホスト SCSI アクセス

OpenVMS Alpha バージョン 6.2 以降、OpenVMS Cluster システム上の複数の Alpha ホストを 1 本の SCSI バスで接続できるようになり、SCSI ストレージ・デバイスのアクセスを共用できるようになりました。SCSI ストレージ・デバイスのアクセスを共用することで高い可用性が実現できます。

図 A-2 は、SCSI インターコネクトで SCSI デバイスのアクセスを共用する OpenVMS Cluster 構成です。ただし、ホスト間 OpenVMS Cluster (System Communications Architecture [SCA]) 通信には、もう 1 本のインターコネクト (ローカル・エリア・ネットワーク[LAN]など) が必要です。

図 A-2 共有 SCSI アクセスによる可用性が高いサーバ



ZK7479AGE

3 ノード OpenVMS Cluster システムを構築するには、ストレージ・インターコネクトに共有 SCSI バスを使用します。また大規模な OpenVMS Cluster 構成に共有 SCSI バスを組み込むこともできます。SCSI バスでクォーラム・ディスクを使用すると、2 ノード構成や3 ノード構成の可用性を強化できます。共有 SCSI ストレージ・デバイス向けには、ホスト方式の RAID (ホスト方式のシャドウイング) と MSCP サーバがサポートされています。

A.3 構成上の要件とハードウェア・サポート

この項では、マルチホスト SCSI OpenVMS Cluster システムにおける構成上の要件とサポートされているハードウェアを紹介します。

A.3.1 構成上の要件

表 A-1 は、SCSI OpenVMS Cluster システムに構成できるソフトウェアとハードウェアの基本構成要素の要件と性能をまとめたものです。

表 A-1 SCSI マルチホスト OpenVMS Cluster 構成の要件

要件	説明
ソフトウェア	<p>SCSI インターコネクト上のストレージまでのアクセスを共有する Alpha ホストは、すべて以下のソフトウェアを実行していなければなりません。</p> <ul style="list-style-type: none"> • OpenVMS Alpha バージョン 6.2 またはそれ以降 • OpenVMS Cluster Software for OpenVMS Alpha バージョン 6.2 またはそれ以降
ハードウェア	<p>表 A-2 は、SCSI OpenVMS Cluster システムにサポートされているハードウェア構成要素をまとめたものです。SCSI OpenVMS Cluster 構成で使用できるその他のハードウェア・デバイスについては、第 A.7.7 項を参照してください。</p>
SCSI のテープ・ドライブ、フロッピー・ドライブ、CD-ROM ドライブ	<p>マルチホスト SCSI インターコネクト上に、SCSI のテープ・ドライブ、フロッピー・ドライブ、CD-ROM ドライブは構成できません。構成で SCSI テープ、フロッピー、CD-ROM ドライブのいずれが必要な場合は、シングル・ホスト SCSI インターコネクト上で構成してください。OpenVMS Cluster 構成の他のホストに対して、SCSI のテープ、フロッピー、CD-ROM ドライブは、MSCP サービスまたは TMSCP サービスが可能です。</p>
1 本の SCSI バス当たりの最高ホスト数	<p>マルチホスト SCSI バスには、最高で 3 つのホストを接続できます。表 A-2 に掲載されているホストは、同じ共有 SCSI インターコネクト上に任意の組み合わせで構成できます。</p>
1 つのホスト当たりの最高 SCSI バス本数	<p>1 つのホストには、最高で 6 本のマルチホスト SCSI バスを接続できます。構成できる非共有 (シングル・ホスト) SCSI バスの本数は、ホスト・バスの空きロット数で決まります。</p>
ホスト間通信	<p>クラスタの全メンバは、DSSI、CI、Ethernet、FDDI、または MEMORY CHANNEL など、ホスト間 (SCA) 通信に使用できるインターコネクトで接続しなければなりません。</p>
ホスト方式の RAID (ホスト方式のシャドウイングなど)	<p>SCSI OpenVMS Cluster 構成でサポート。</p>

(次ページに続く)

表 A-1 (続き) SCSI マルチホスト OpenVMS Cluster 構成の要件

要件	説明
SCSI デバイスの命名	<p>各 SCSI デバイスの名前は、OpenVMS Cluster システム上で一意の名前です。マルチホスト SCSI バスを使用しているシステム上でデバイスを構成するときは、以下の要件に従ってください。</p> <ul style="list-style-type: none">ホストから各 SCSI インターコネクトへの接続では、少なくともアダプタを 1 つ使用する。ポート割り当てクラスを使用する場合を除き、所定の SCSI インターコネクトに接続されているホスト・アダプタは、すべて同じ OpenVMS デバイス名 (PKA0 など) を持つものとする (『OpenVMS Cluster システム』参照)。SCSI インターコネクトに接続された各システムには、ゼロ以外のノード・ディスク割り当てクラス値が必要。これらのノード・ディスク割り当てクラス値は、以下の条件のいずれかに該当する場合異なる場合がある。<ul style="list-style-type: none">SCSI インターコネクトに正のゼロ以外のポート割り当てクラスがある場合ゼロ以外のコントローラ割り当てクラスがある HSZ70 または HSZ80 コントローラが、SCSI インターコネクトに接続されているデバイスだけにアクセスする場合 <p>複数の SCSI インターコネクトがある場合は、すべての SCSI インターコネクトを考慮に入れ、各システムのノード・ディスク割り当てクラスに異なる値を選択できるかどうかを判断する必要があります。また、既存の SCSI インターコネクトに SCSI デバイスを追加する場合は、ノード・ディスク割り当てクラスを異なる値にできるかどうかを再評価する必要があります。このため、弊社では、同じ SCSI インターコネクトに接続されているすべてのシステムで、同じノード・ディスク割り当てクラスを使用することをお勧めします。割り当てクラスの詳細は、『OpenVMS Cluster システム』を参照してください。</p>

A.3.2 ハードウェアのサポート

表 A-2 は、SCSI OpenVMS Cluster システムにサポートされているハードウェア構成要素です。あわせてこれらのハードウェア構成で使用できる最小バージョンも掲載します。つまり、表 A-2 に記載されているバージョンまたはそれ以降のバージョンが必要であるということです。ホストのサポート情報については、次の URL を参照してください。

<http://www.hp.com/country/us/eng/prodserv/servers.html>

このサイトで AlphaServer あるいは AlphaStation システムのドキュメントを参照してください。

ディスク・サポート情報については、次の URL を参照してください。

<http://www.hp.com/country/us/eng/prodserv/storage.html>

SCSI インターコネクト構成と SCSI インターコネクト上のすべてのデバイスは、ANSI 標準 SCSI-2マニュアルまたは SCSI-3 のアーキテクチャ標準とコマンド標準で定義した要件と、この付録で紹介する要件に準拠していなければなりません。SCSI OpenVMS Cluster 構成で使用できるハードウェア・デバイスについては、第 A.7.7 項も参照してください。

表 A-2 SCSI OpenVMS Cluster システムでサポートされているハードウェア

構成要素	サポートされているアイテム	最低ファームウェア (FW) のバージョン ¹
コントローラ	HSZ40-B	2.5 (FW)
	HSZ50	
	HSZ70	
	HSZ80	8.3 (FW)
アダプタ ²	埋め込み (NCR-810 方式)	
	KZPAA (PCI から SCSI へ)	
	KZPSA (PCI から SCSI へ)	A11 (FW)
	KZPBA-CB (PCI から SCSI へ)	5.53 (FW)
	KZTSA (TURBOchannel から SCSI へ)	A10-1 (FW)

¹この欄で定めのない限り、デバイスの最小ファームウェア・バージョンは、実行するオペレーティング・システムのバージョンと同じでなければなりません。SCSI マルチホスト OpenVMS Cluster 構成では、それ以外のファームウェア要件はありません。

²ローカル・ストレージへのシングル・ホスト・アクセス用にこの他のタイプの SCSI アダプタを構成できます。

A.4 SCSI インターコネクトの概念

SCSI 標準では、イニシエータ (通常は、ホスト・システム) と SCSI ターゲット (通常は、周辺機器) 間のインタラクションを管理する規則を定義しています。この標準では、デバイス固有の特性に関係なく、SCSI デバイス (ディスク・ドライブ、テープ・ドライブ、プリンタ、光学媒体装置など) との通信が可能なホスト・サービスを扱います。

以下の項では、SCSI 標準とデフォルトの操作モードについて説明します。また、キャパシティ、パフォーマンス、可用性、通信距離などに関して、デフォルトの SCSI 機能を強化することができるオプションの機構についても説明します。

A.4.1 デバイス数

SCSI バスは、最高で 16 個のデバイスをサポートできる I/O インターコネクトです。Narrow SCSI バスは、最高で 8 基のデバイスをサポートします。Wide SCSI バスは、最高で 16 個のデバイスをサポートします。サポートできるデバイスには、ディスク・ドライブやテープ・ドライブなどのホスト・アダプタ、周辺コントローラ、ディスクリット周辺機器があります。0 から 15 までの固有の ID 番号でアドレス指定します。デバイス ID を割り当てるには、コンソール・コマンドを入力するか、ジャン

パまたはスイッチを設定するか、StorageWorks エンクロージャのロットを選択します。

注意

16 個のデバイスを Wide SCSI バスに接続するには、デバイス自体も Wide アドレス指定をサポートしている必要があります。Narrow デバイスでは、ID 7 を超えるアドレスをサポートしていません。現在、HSZ40 は 7 を超えるアドレスをサポートしていません。Wide アドレス指定をサポートしているホスト・アダプタには、KZTSA、KZPSA、QLogic Wide アダプタ (KZPBA、KZPDA、ITIOP、P1SE、P2SE) があります。マルチホスト SCSI OpenVMS Cluster 構成では、KZPBA-CB のみサポートしています。

先の 8 個のデバイスの制限を超えてデバイスを構成するときは、バスの長さの要件を順守してください (表 A-4 を参照)。

BA356 ボックスに Wide ID を構成するには、BA356 マニュアル『StorageWorks Solutions BA356-SB 16-Bit Shelf User's Guide』(EK-BA356-UG) を参照してください。開始アドレスが 8 の BA356 ボックスに Narrow デバイスは構成しないでください。

SCSI インターコネク特的デバイスを増やすには、一部のデバイスに論理ユニット番号 (LUN) を利用した二次レベルのデバイス・アドレス指定を実装します。各デバイス ID には、最高で 8 つの LUN (0 ~ 7) を利用して、1 つの SCSI デバイスを複数のユニットとしてアドレス指定できます。デバイス ID ごとの LUN の最大値は、8 です。

注意

SCSI インターコネク特的にデバイスを接続するとき、インターコネク特的上の各デバイスには固有のデバイス ID が必要です。固有のデバイス ID を作成するには、デバイスのデフォルト・デバイス ID を変更します。個々のデバイスの ID の設定方法については、該当するデバイスのオーナーズ・ガイドを参照してください。

A.4.2 パフォーマンス

SCSI デバイスのデフォルトの操作モードは、8 ビット非同期モードです。このモードを Narrow モードと呼ぶことがあります。このモードでは、8 ビットのデータがデバイス間で転送されます。各データ転送は、データを受信するデバイスが肯定応答します。デフォルト・モードのパフォーマンスには限度があるので、SCSI 標準では、パフォーマンスを強化するためのオプション機構が定義されています。次に、パフォーマンスを強化する方法を 2 つ示します。

- インターコネク特的上で並列転送されるデータ量を増加します。16 ビット幅と 32 ビット幅のオプションでは、データ転送速度を、それぞれ 2 倍と 4 倍にできま

す。32 ビット・オプションが実装される例は少ないので、この付録では、16 ビット操作についてのみ説明し、16 ビット操作をWideと称します。

- 同期データ転送を使用します。同期モードでは、データを受信するデバイスの肯定応答が後に付いた複数のデータ転送を連続して実行できます。標準では、同期データ転送の低速モード (Standard モードとも呼ぶ) と Fast モードを定義しています。
 - Standard モードでは、最高で每秒 5 百万データを転送できます。
 - Fast モードでは、最高で每秒 1 千万データを転送できます。
 - Ultra モードでは、最高で每秒 2 千万データを転送できます。

SCSI インターコネクトにおける通信は、すべて 2 つのデバイス間で順次発生するため、各デバイス・ペアは使用するオプション機能を互いに調整する必要があります。ほとんどの SCSI デバイスは、これらのオプションを少なくとも 1 つ実装します。

表 A-3 は、8 ビット転送と 16 ビット転送を、Standard、Fast、Ultra の各同期モードで使用した場合のデータ転送速度をまとめたものです。

表 A-3 最大データ転送速度 (MB/s)

モード	Narrow (8 ビット)	Wide (16 ビット)
Standard	5	10
Fast	10	20
Ultra	20	40

A.4.3 接続距離

SCSI インターコネクトで利用できる最大長は、構成で使用するシグナル通知と、データ転送速度で決まります。SCSI インターコネクトの電気シグナル通知には 2 種類あります。

- シングル・エンド・シグナル通知
シングル・エンド方式は最も一般的であり、また経済的です。適度な通信距離を利用できます。
- ディファレンシャル・シグナル通知
この方式では、信号の一貫性が高いので、長距離の SCSI バスが可能です。

表 A-4 は、シグナル通知の方式によって SCSI インターコネクトの距離がどのような影響を受けるかをまとめたものです。

表 A-4 SCSI インターコネクタの最大距離

シグナル通知の方式	データ転送速度	最大ケーブル長
シングル・エンド	Standard	6 m ¹
シングル・エンド	Fast	3 m
シングル・エンド	Ultra	20.5 m ²
ディファレンシャル	Standard または Fast	25 m
ディファレンシャル	Ultra	25.5 m ²

¹SCSI 標準は、このインターコネクタ方式で最長 6 m と指定されています。ただし、データの一貫性を最高に保つため、ケーブル長はなるべく 4 m 以内にしてください。

²詳細については、『StorageWorks Ultra SCSI Configuration Guidelines』(EK-ULTRA-CG) を参照してください。

DWZZA, DWZZB, DWZZC コンバータは、シングル・エンドからディファレンシャルまでを含む各種コンバータであり、シングル・エンドやディファレンシャルの SCSI インターコネクタ・セグメントに接続します。DWZZA は、Narrow (8 ビット) SCSI バス用、DWZZB は Wide (16 ビット) 用の SCSI バスであり、DWZZC は Wide Ultra SCSI バス用です。

ディファレンシャル・セグメントは、以下のような状況に最適です。

- シングル・エンド・インターコネクタの接続距離の上限を超えて敷設したいとき
- シングル・エンド・デバイスとディファレンシャル・デバイス間で通信をしたいとき

DWZZA, DWZZB, DWZZC は、あくまでシグナル・コンバータであり、SCSI デバイス ID を割り当てることはできません。2 つの SCSI デバイス間のバスには、最大で 2 個の DWZZA コンバータまたは 2 つの DWZZB コンバータを定義できます。DWZZC の構成については、『StorageWorks Ultra SCSI Configuration Guidelines』を参照してください。

A.4.4 ケーブル配線と終端

シングル・エンドとディファレンシャルの各 SCSI インターコネクタには、各端点に 1 つずつ、合計 2 つのターミネータが必要です。インターコネクタに指定された最大接続距離は、ターミネータからターミネータまでを計測します。

インターコネクタ・ターミネータの電源は、TERMPWR という SCSI インターコネクタ・ラインから供給します。各 StorageWorks ホスト・アダプタとエンクロージャが、TERMPWR インターコネクタ・ラインに電源を供給し、ホストまたはエンクロージャのどれかに電源が供給されていれば、インターコネクタの終端が機能するようになっています。

ショート・ケーブル(または etch) でインターコネクトに接続されているデバイスをスタブと言います。スタブの長さは、インターコネクトのシグナルの一貫性を維持するために短くなっています。スタブの許容最大長は、インターコネクトが使用するシグナル通知方式によって以下のように決められています。

- シングル・エンド・インターコネクトの場合、最大スタブ長は 0.1 m です。
- ディファレンシャル・インターコネクトの場合、最大スタブ長は 0.2 m です。

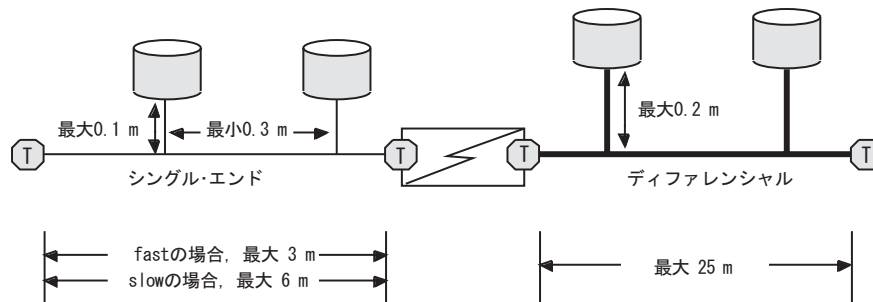
その他、シングル・エンド・インターコネクト上のスタブ間の最短の長さは、0.3 m に定められています。この構成例については、図 A-3 を参照してください。

注意

シングル・エンド・バスとディファレンシャル・バスには、DWZZxコンバータを使用する場合もそれぞれ終端が必要です。

既存のターミネータを超えて SCSI バスを延長するには、そのターミネータを無効にするか、取り外す必要があります。

図 A-3 最大スタブ長



ZK7480AGE

A.5 SCSI OpenVMS Cluster ハードウェア構成

選択するハードウェア構成は、各種要素の組み合わせによって異なります。

- コンピューティングのニーズ
たとえば、可用性の連続性や SCSI OpenVMS Cluster システムからシステムを接続解除または取り外せる性能
- 環境
たとえば、コンピューティング・ファシリティの物理的な属性
- リソース

たとえば、主要機器や空き PCI スロット

構成上の制約については、『OpenVMS Cluster Software Software Product Description』(SPD 29.78.xx)を参照してください。

以下の項では、SCSI 構成の構築方法のガイドラインの他、さまざまなサイトに適した構成方法について説明します。

A.5.1 アドオン SCSI アダプタを使用するシステム

共用 SCSI バス構成では、一般にオプションのアドオン KZPAA、KZPSA、KZPBA、KZTSA の各アダプタを使用します。これらのアダプタを接続しても、SCSI ケーブルの長さに反映されないため、一般に内部アダプタよりも構成が簡単です。また、共用 SCSI バス用のアドオン・アダプタでシステムを構成する際に、内部アダプタで共用できないデバイス(SCSI テープ・ドライブ、フロッピー・ドライブ、CD-ROM ドライブなど)を接続できます。

アドオン・アダプタを使用する場合は、ストレージは BA350、BA353、または HSZ xx StorageWorks エンクロージャで構成します。これらのエンクロージャは、すべてのデータ・ディスク、そして共用の OpenVMS Cluster システムとクォーラム・ディスクに最適です。StorageWorks エンクロージャを使用すれば、ディスクのアクセスを維持したまま個々のシステムをシャット・ダウンできます。

以下の項では、アドオン・アダプタを活かした SCSI OpenVMS Cluster 構成をいくつか紹介します。

A.5.1.1 アドオン SCSI アダプタによる基本的なシステムの構築

図 A-4 は、SCSI アダプタと StorageWorks エンクロージャを使用した基本構成の論理図です。この構成は比較的シンプルであるにもかかわらず、内部バス上で非共用ファイル(ページ・ファイルやスワップ・ファイルなど)とともにテープ、フロッピー、CD-ROM、ディスクを使用できるという利点があります。図 A-5 は、AlphaServer 1000 システムと BA350 エンクロージャを使用した同じタイプの構成です。

BA350 エンクロージャは内部で 0.9 m の SCSI ケーブルを使用します。この構成は、一般に 2 本の 1 m SCSI ケーブルを使用します。(BA353 エンクロージャでも同じ合計ケーブル長で 0.9 m ケーブルを使用します。)合計長さ 2.9 m のケーブルで Fast SCSI モード操作が可能です。

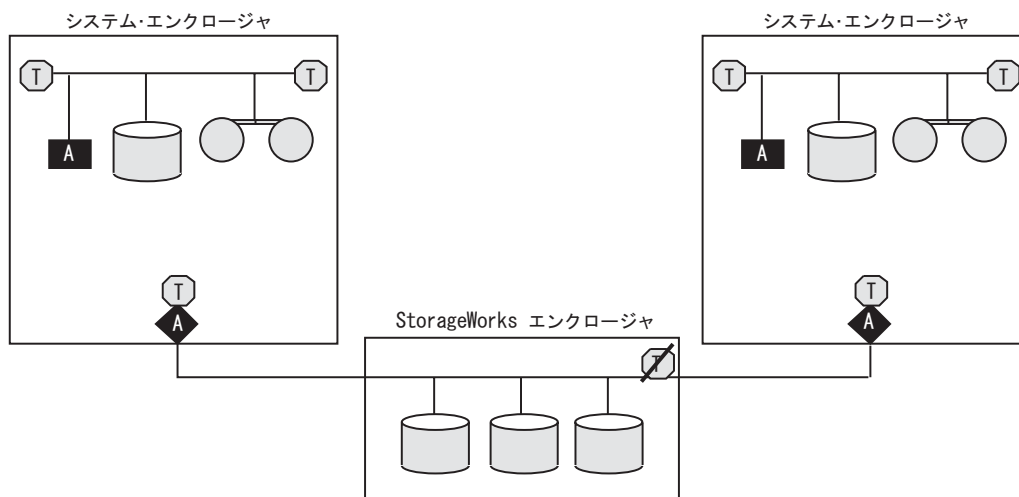
共用 BA350 ストレージ・エンクロージャは、理論的には単一点障害の要因になり得ますが、この基本システムは非常に信頼性が高い SCSI OpenVMS Cluster 構成です。クォーラム・ディスクが BA350 にある場合、OpenVMS Cluster システムとのアクセスを維持しながら、どちらの AlphaStation システムも単独でシャット・ダウンできます。ただし、AlphaStation システムを物理的に取り外すことはできません。終端がない SCSI バスが生じるからです。

OpenVMS Cluster システムの動作を続行しつつシステムの取り外しができるようにするには、第 A.5.1.2 項にあるように、DWZZxコンバータでシステムを構築します。SCSI インターコネクトに障害が発生しても、データのアクセスを維持したい場合は、以下のどちらかの操作をします。

- もう 1 つの BA350 シェルフとともに冗長 SCSI インターコネクトを追加する。
- データをシャドウイングする。

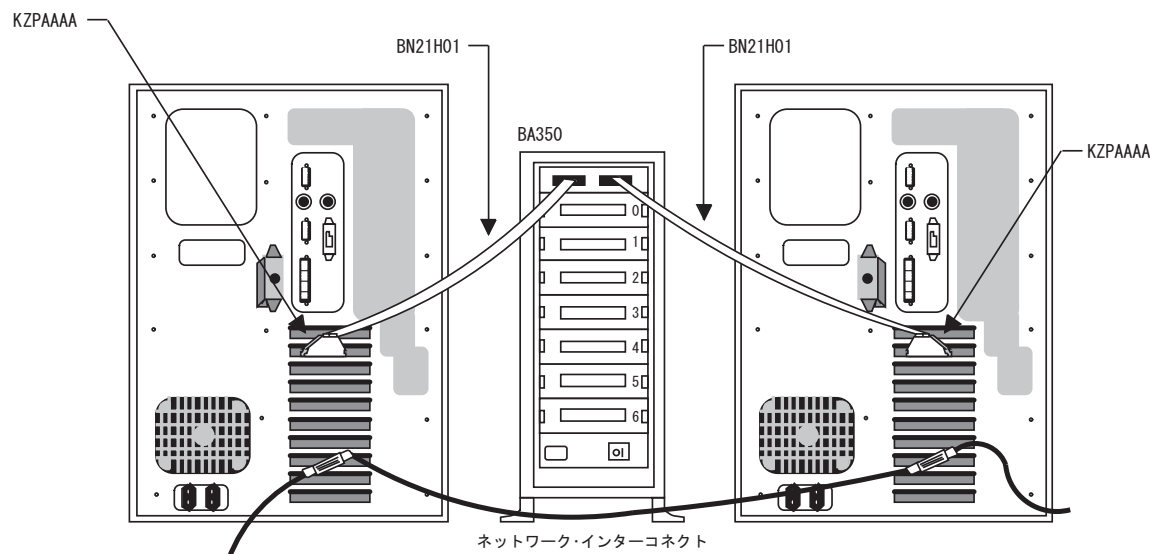
図 A-4 と、この付録のその他の論理構成図では、必要なネットワーク・インターコネクトが省略されています。

図 A-4 概念図: 基本的な SCSI システム



ZK7501AGE

図 A-5 構成例: AlphaServer 1000, KZPAAA アダプタ, BA350 エンクロージャ使用の基本的な SCSI システム



ZK7449AGE

A.5.1.2 エンクロージャを追加または物理的距離を延長したシステム, または, HSZ コントローラ付きのシステム

エンクロージャを追加したり, システム間の物理的な距離を離したい場合, あるいは HSZ コントローラを使用する場合, シングル・エンド・シグナル通知とディファレンシャル・ケーブルの SCSI バスを備えたシステム間に DWZZxコンバータを配置する構成を使用します。

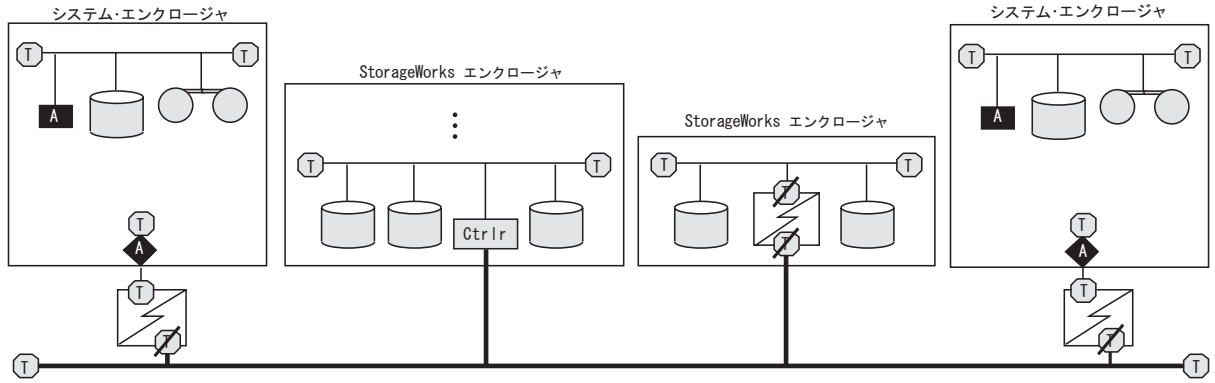
DWZZxコンバータを利用すると, SCSI バスの接続距離の制約が緩和されます。これは DWZZxでは, ディファレンシャル・シグナル通知を使用するバスにシングル・エンド・デバイス接続できるからです。第 A.4.3 項で説明したように, ディファレンシャル・シグナル通知を使用する SCSI バス構成での最大ケーブル長は 25 m です。一方, シングル・エンド構成で Fast モード・データ転送を行うには 3 m が限度です。

DWZZxコンバータは, スタンドアロンのデスクトップ構成要素として, あるいは StorageWorks 互換の構築ブロックとして使用できます。DWZZxコンバータは, 内部 SCSI アダプタやオプションの KZPAA アダプタと併用できます。

HSZ40 は, ディファレンシャル SCSI バスに接続できるハイ・パフォーマンスなディファレンシャル SCSI コントローラであり, 最高で 72 個の SCSI デバイスをサポートします。HSZ40 は, ディファレンシャル・コンバータでシングル・エンドになる DWZZxを組み込んだ共用 SCSI バス上に構成できます。HSZ40 コントローラ上に構成したディスク・デバイスは, RAID セットに組み合わせることができ, これによってパフォーマンスと可用性を強化します。

図 A-6 は、潜在的な物理的距離を広げるために (またはエンクロージャや HSZ40 を追加するために) DWZZA を追加した構成の論理図であり、図 A-7 は、この構成例を具体的に示した図です。

図 A-6 概念図: DWZZA を利用した物理的な距離の延長またはエンクロージャの追加



ZK7482AGE

OpenVMS Cluster インターコネクトとしての SCSI
 A.5 SCSI OpenVMS Cluster ハードウェア構成

図 A-7 構成例: DWZZA を利用した物理的な距離の延長またはエンクロージャの追加

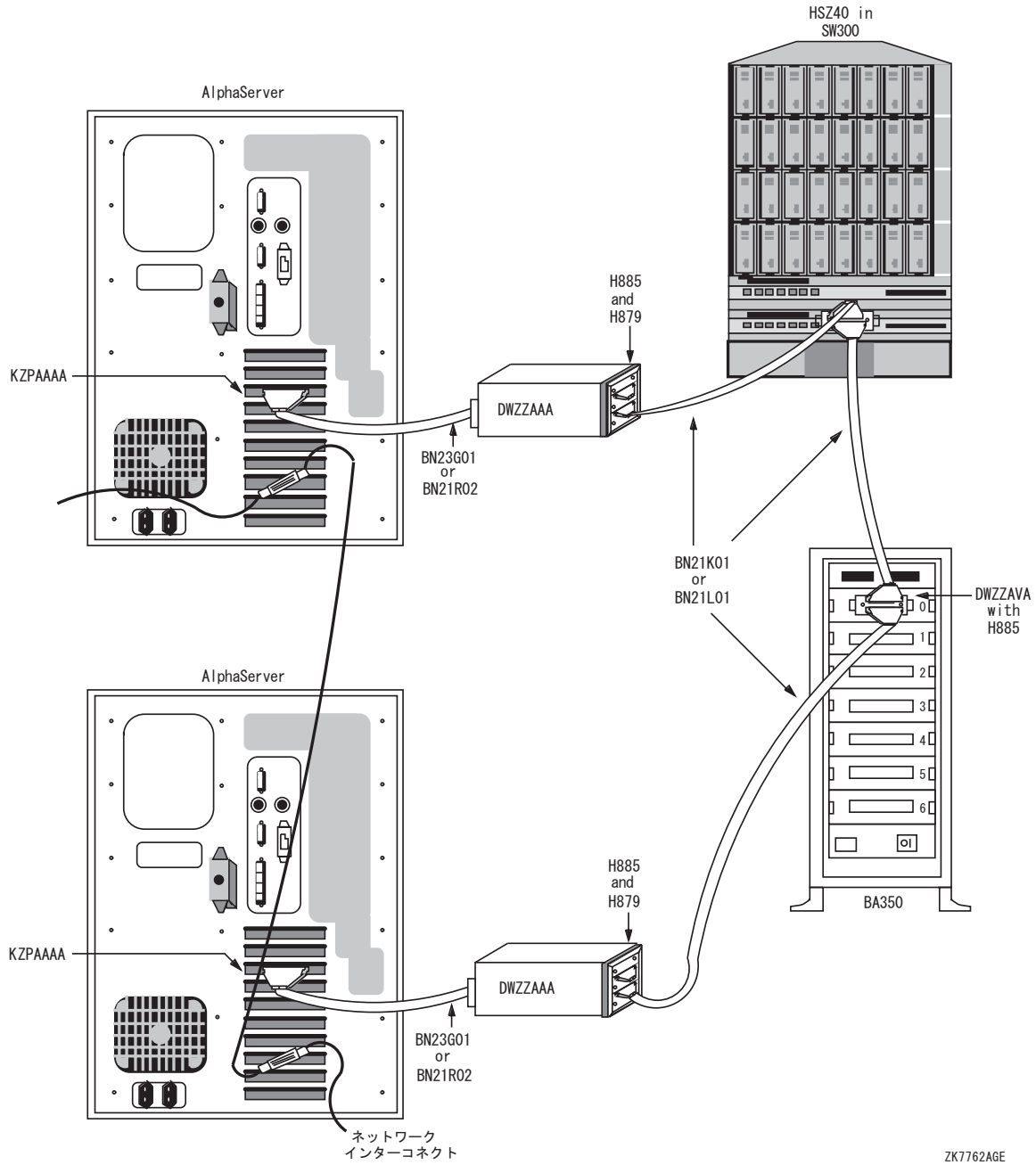
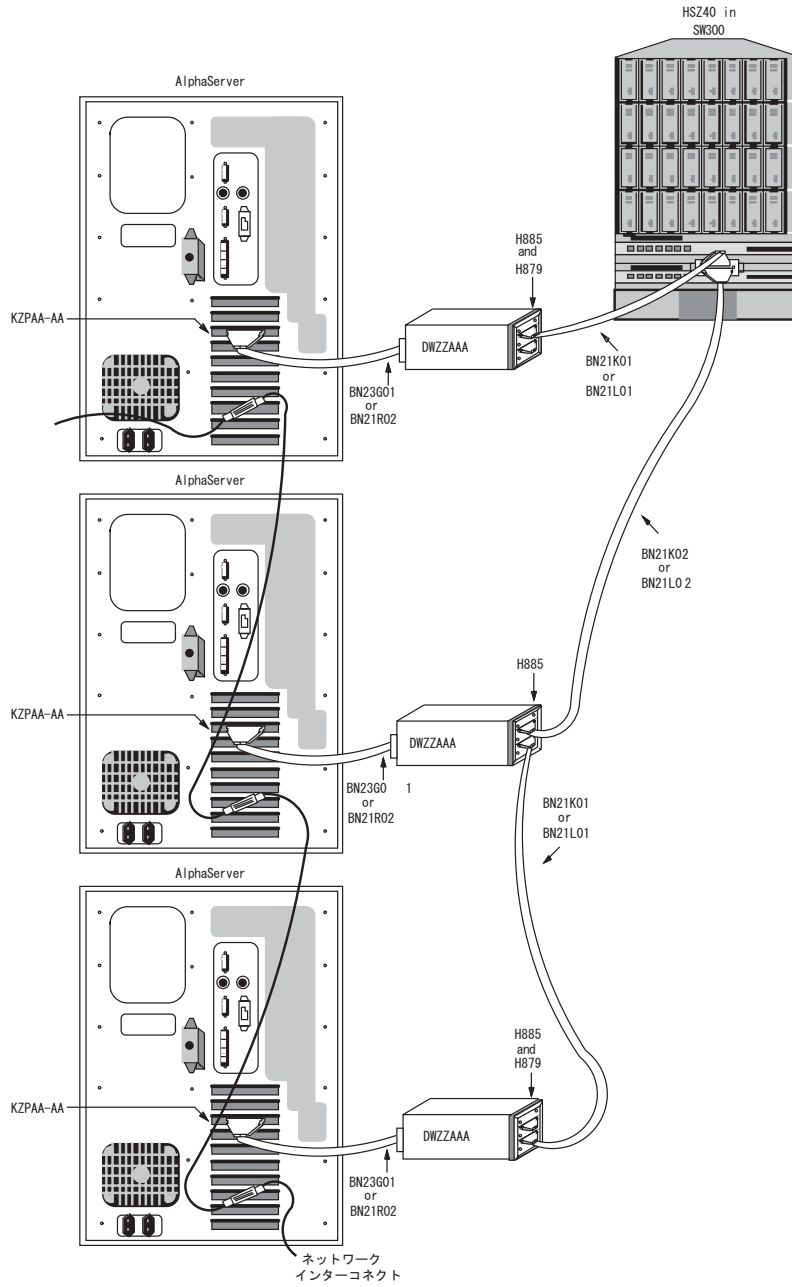


図 A-8 は、3 ホスト SCSI OpenVMS Cluster システムの構成方法を表した図です。

図 A-8 構成例: SCSI バス上の 3 ホスト



ZK7499AGE

A.5.1.3 ディファレンシャル・ホスト・アダプタを使用するシステムの構築

図 A-9 は、同じ SCSI バス上に 2 個の KZPSA アダプタを備えた構成の例です。この構成で、SCSI 終端は、KZPSA から取り外されており、外部ターミネータは“Y”ケーブル上にインストール済みです。そのため、SCSI バスを有効にしたまま SCSI バスから KZPSA アダプタを取り外すことができます。クラスタ内の他のシステムをアクティブにしたまま (メンテナンスや修理目的で) SCSI OpenVMS Cluster 構成から個々のシステムを取り外せる機能があるために、非常に高いレベルの可用性を実現できます。

図 A-9 では、以下の点に注意してください。

- 終端がホスト・アダプタから取り外されている。
- BA356 内部のシングル・エンド・バスは、スロット 0 の DWZZB とパーソナリティ・モジュールの自動ターミネータで終端されている。(外部ケーブルやターミネータはパーソナリティ・モジュールに接続されていない。)
- DWZZB のディファレンシャル終端が取り外されている。

図 A-9 構成例: ディファレンシャル・ホスト・アダプタ (KZPSA) を使用した SCSI システム

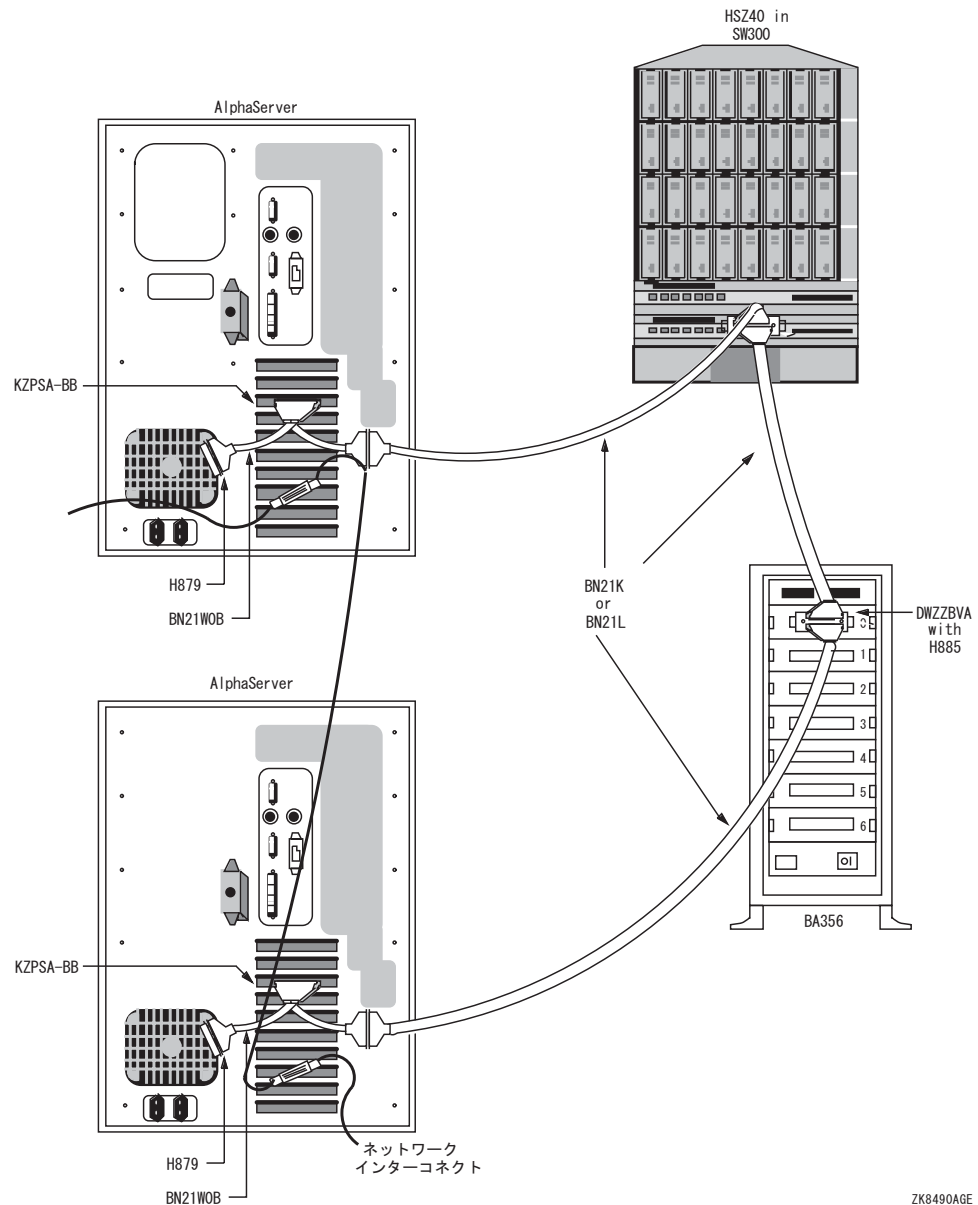


図 A-9 の構成に示すディファレンシャル SCSI バスは、エンクロージャ間でチェーン化され、最長で 25 m となっています。(BA356 バスはディファレンシャル SCSI バスの長さに加えません。ディファレンシャル・バスを構成するのは、BN21W-0B “Y”ケーブルと BN21K/BN21L ケーブルだけです。) このケーブル配線方式が不便な構成や、接続距離が不足する場合は、別のラジアル方式を利用できます。

別のラジアル SCSI ケーブル配線方式とは、SCSI ハブ方式のことです。図 A-10 は、SCSI ハブ構成の論理図です。図 A-11 は、この構成例を具体的に表現したものです。

OpenVMS Cluster インターコネクトとしての SCSI
A.5 SCSI OpenVMS Cluster ハードウェア構成

図 A-10 概念図: SCSI ハブを使用した SCSI システム

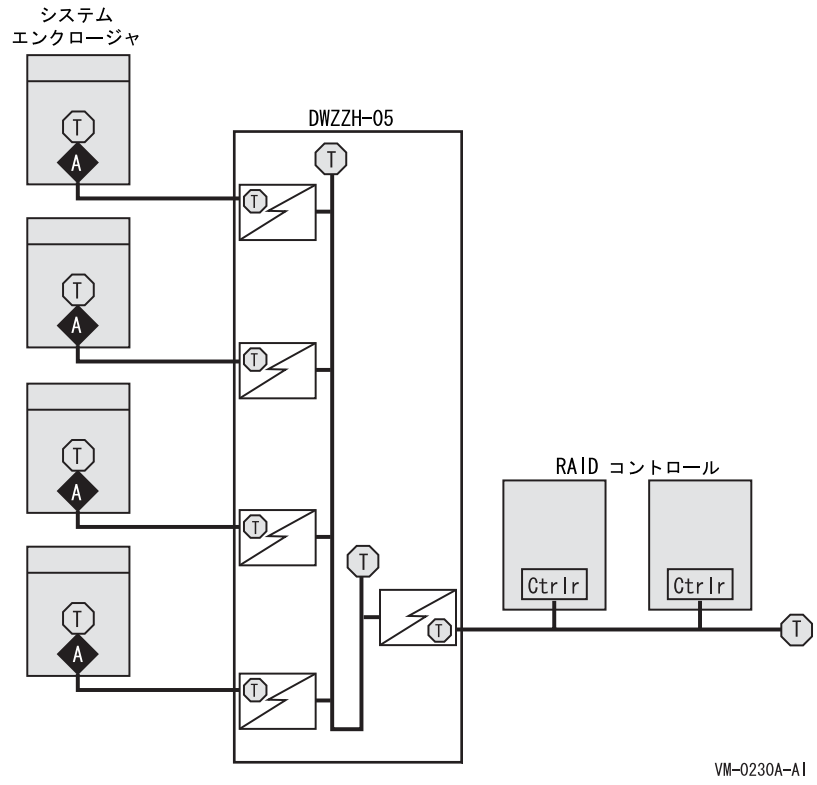
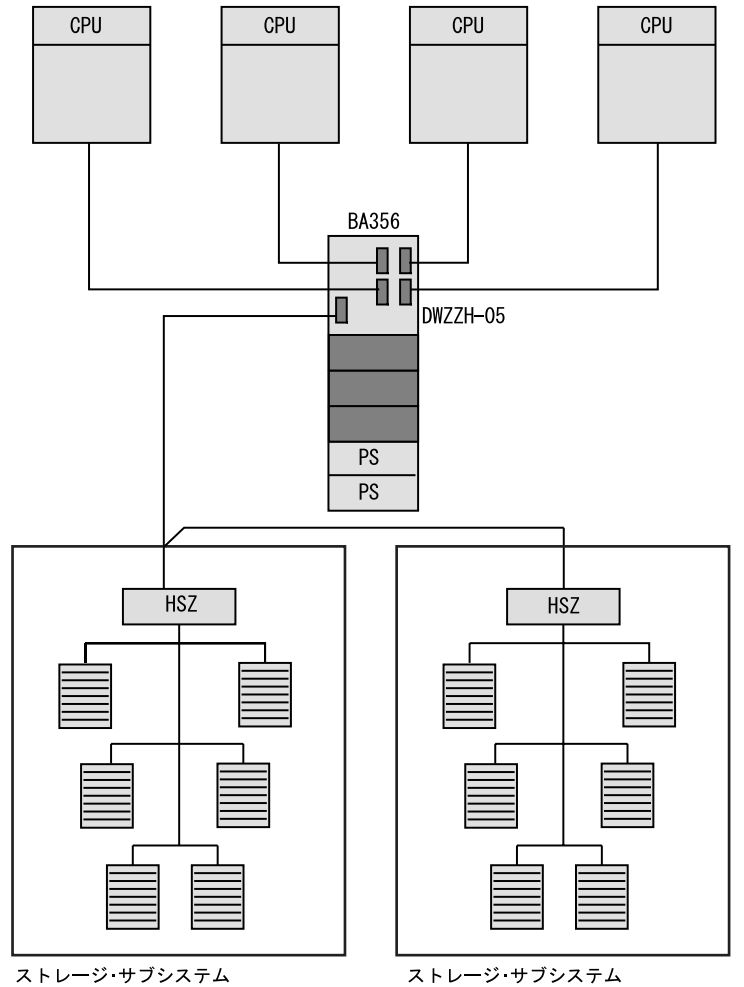


図 A-11 は、SCSI ハブ構成の例です。

図 A-11 構成例: SCSI ハブ構成における SCSI システム



VM-0216A-A1

A.6 インストール

ここでは、SCSI OpenVMS Cluster システムにハードウェアをセットアップし、インストールする手順について説明します。この節では、共用 SCSI バス方式の新しい OpenVMS Cluster システムを作成します。ただし、既存の OpenVMS Cluster 構成に共用 SCSI バスを追加する場合は、『OpenVMS Cluster システム』の手順説明も合わせて参考にし、計画全体を整えてください。

表 A-5 は、SCSI OpenVMS Cluster システムにハードウェアをセットアップし、インストールする手順を説明したものです。

表 A-5 SCSI OpenVMS Cluster システムのインストール手順

手順	説明	参照
1	エンクロージャ間が適切に接地されているか確認する。	第 A.6.1 項と第 A.7.8 項
2	SCSI ホスト ID を構成する。	第 A.6.2 項
3	システムの電源を入れ、デバイスを確認する。	第 A.6.3 項
4	SCSI コンソール・パラメータを設定する。	第 A.6.4 項
5	OpenVMS オペレーティング・システムをインストールする。	第 A.6.5 項
6	追加システムを構成する。	第 A.6.6 項

A.6.1 手順 1: SCSI 接地要件の確認

OpenVMS Cluster システムをインストールする前に、配電システムがローカル要件を満足しているか確認します (電気コードなど)。構成が、共通 SCSI インターコネクタで接続された複数のエンクロージャからなる場合、それらのエンクロージャの接地が適切であるか確認します。安全上と SCSI インターコネクタの正確な機能のためには、正しい接地方法が求められます。

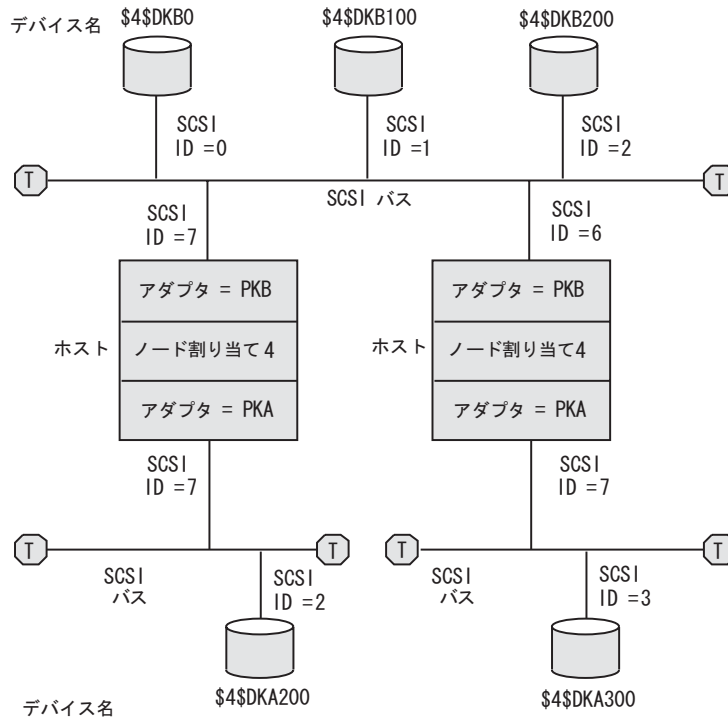
電気工事は有資格者によって行ってください。SCSI システムの接地要件の詳細については、第 A.7.8 項を参照してください。

A.6.2 手順 2: SCSI ノード ID の構成

この項では、SCSI ノードとデバイス ID の構成方法を説明します。マルチホスト SCSI バスとシングル・ホスト SCSI バスでは、SCSI ID は異なる方法で割り当てます。

図 A-12 は 2 つのホストを示しています。それぞれがシングル・ホスト SCSI バスで構成されており、マルチホスト SCSI バスを共用しています。(この図の記号の意味については、図 A-1 を参照してください。)

図 A-12 SCSI アクセスの割り当てクラスの設定



ZK7483AGE

以下の項では、この種のマルチホスト SCSI 構成に ID を割り当てる方法を説明します。このトピックの詳細については、『OpenVMS Cluster システム』を参照してください。

A.6.2.1 マルチホスト SCSI バスにおけるデバイス ID の構成

マルチホスト SCSI バスの構成時には、以下の規則を守ってください。

- マルチホスト・バス上のホスト・アダプタには、それぞれ異なる ID を設定してください。まず ID 7 を割り当て、次に ID 6 というように ID 番号を降順に割り当てていきます。

ホストに 2 本のマルチホスト SCSI バスがある場合、各 SCSI アダプタに別々に ID を割り当てます。同じ ID にアダプタを設定しなければならないという条件はありませんが、同じ ID を使用すると構成の管理が簡単です。(SCSI コンソール・パラメータで、内部アダプタにホスト ID を設定する方法については、第 A.6.4 項を参照してください。)

- マルチホスト SCSI バスに接続されたデバイスやストレージ・コントローラに ID を割り当てる場合、ID 0 (ゼロ) から始め、最速の I/O 応答時間が必要なディスクには最高の ID 番号を割り当てます。

- マルチホスト SCSI バスに接続されたデバイスには、ホストを起点として同じ名前を割り当てます。以下にその方法を示します。
 - 図 A-12 のように、マルチホスト SCSI バスに接続されたすべてのホストが同じノード割り当てクラスに設定されていることを確認し、マルチホスト SCSI バスに接続されたすべてのホスト・アダプタに同じコントローラ名が割り当てられていることを確認します。
 - ポート割り当てクラス (『OpenVMS Cluster システム』参照) または HSZ 割り当てクラス (『OpenVMS Cluster 構成ガイド』参照) を使用します。

A.6.2.2 シングル・ホスト SCSI バスにおけるデバイス ID の構成

デバイス ID の選択は、ノード割り当てクラスを利用するか、ポート割り当てクラスを利用するかで異なります。以下の説明では、ノード割り当てクラスを使用しています。ポート割り当てクラスについては、『OpenVMS Cluster システム』を参照してください。

マルチホスト SCSI 構成では、OpenVMS が生成するデバイス名の形式は \$allocation_class\$DKA300 です。割り当てクラスは、ALLOCLASS システム・パラメータで設定します。OpenVMS は、ブート時に各コントローラに文字を割り当ててコントローラ名 (A, B, C など) とします。単位番号 (0, 100, 200, 300 など) は、SCSI デバイス ID から抽出されます。

マルチホスト SCSI 構成の一部であるシングル・ホスト SCSI バスでデバイスを構成する場合、シングル・ホスト SCSI バスに接続されたディスクに必ず一意のデバイス名を指定してください。そのためには、同じ割り当てクラスを使用するシステムに同じコントローラ名でシングル・ホスト SCSI バスに接続されているデバイスには、異なる ID を割り当てます。バスが共用されていない場合でも、異なるデバイス名を付けます。

たとえば、図 A-12 で、図の下側の 2 つのディスクは、同じ割り当てクラスを使用する 2 つのシステムの SCSI バス A にあります。そのため、異なるデバイス ID (この場合、2 と 3) が割り当てられています。

所定の割り当てクラス、SCSI デバイス・タイプ、コントローラ名 (この例では、\$4\$DKA) で、クラスタには最高で 8 つのデバイスを接続できます。各デバイスごとに SCSI バス ID を割り当てます。8 つすべての ID を使用するには、別のバスのプロセッサと同じ ID の 1 本の SCSI バス上にディスクを構成する必要があります。これによってパフォーマンスにどのような影響が出るかについては、第 A.7.5 項を参照してください。

異なる SCSI バスに、異なる SCSI デバイス・タイプを同じ SCSI ID で構成すれば、SCSI バス ID を“兼用”できます。たとえば、デバイス・タイプ DK とデバイス・タイプ MK の場合、\$4\$DKA100 と \$4\$MKA100 になります。

A.6.3 手順 3: 電源入力と SCSI デバイスの検査

SCSI ケーブルを接続したら、システムの電源を入れます。コンソール SHOW DEVICE コマンドを入力し、すべてのデバイスが SCSI インターコネクトで認識できるか確認します。

SCSI ID の競合がある場合、ディスプレイは現在のデバイスを表示しないか、または、存在しないデバイスを表示することがあります。表示が間違っている場合は、各デバイスの SCSI ID ジャンパ、StorageWorks シェルフによる自動 ID 割り当て、ホスト・アダプタと HSZxx コントローラ ID の設定が間違っていないか確認してください。変更されている場合は、INIT と入力し、SHOW DEVICE を再度実行します。問題が解決しない場合は、SCSI ケーブルの長さや終端を確認してください。

例 A-1 は、コンソール SHOW DEVICE コマンドの出力例です。このシステムには、ホスト SCSI アダプタがプライベート SCSI バス (PKA0) に 1 個、そして追加 SCSI アダプタ (PKB0 と PKC0) がそれぞれ別々の共用 SCSI バスに 1 個ずつあります。

例 A-1 SHOW DEVICE コマンドの出力例

```
>>>SHOW DEVICE
dka0.0.0.6.0          DKA0          RZ26L  442D
dka400.4.0.6.0        DKA400        RRD43  2893
dkb100.1.0.11.0       DKB100        RZ26   392A
dkb200.2.0.11.0       DKB200        RZ26L  442D
dkc400.4.0.12.0       DKC400        HSZ40   V25
dkc401.4.0.12.0       DKC401        HSZ40   V25
dkc500.5.0.12.0       DKC500        HSZ40   V25
dkc501.5.0.12.0       DKC501        HSZ40   V25
dkc506.5.0.12.0       DKC506        HSZ40   V25
dva0.0.0.0.1          DVA0
jkb700.7.0.11.0       JKB700        OpenVMS V62
jkc700.7.0.12.0       JKC700        OpenVMS V62
mka300.3.0.6.0        MKA300        TLZ06  0389
era0.0.0.2.1          ERA0          08-00-2B-3F-3A-B9
pka0.7.0.6.0          PKA0          SCSI Bus ID 7
pkb0.6.0.11.0         PKB0          SCSI Bus ID 6
pkc0.6.0.12.0         PKC0          SCSI Bus ID 6
```

上の例のデバイス名について、以下に説明します。

- DK デバイスは SCSI ディスクです。アダプタ PKA が制御している SCSI バスに接続されているディスクのデバイス名は先頭が DKA で始まります。追加バスにあるディスクの名前は、同様にホスト・アダプタ名を基に割り当てられます (アダプタ PKB の DKB デバイスなど)。

デバイス名の次の文字は、デバイスの SCSI ID です。SCSI ID は、接続先の SCSI バス上でデバイスごとに一意であるか確認してください。

- DK デバイス名の最後の文字は LUN 番号です。この例で、HSZ40 仮想 DK デバイスは、SCSI ID 4、LUN 1 にあります。システムによっては、ゼロ以外の LUN のデバイスが表示されません。
- JK デバイスは、SCSI インターコネクト上の非ディスク装置や非テープ装置です。この例で、JK デバイスは SCSI インターコネクト上で、OpenVMS オペレーティング・システムで実行している他のプロセッサを表します。他のシステムが実行していない場合、JK デバイスは表示されません。この例では、他のプロセッサのアダプタは SCSI ID 7 にあります。
- MK デバイスは SCSI テープです。デバイス MKA300 の A は、プライベート SCSI バスであるアダプタ PKA0 に接続されていることを表しています。
- PK デバイスはローカルな SCSI アダプタです。このようなアダプタの SCSI ID は、右端の列に表示されます。そのバス上の他のデバイスやホスト・アダプタが、使用している ID と異なることを確認してください。

デバイス名の 3 番目の文字 (この例では a) は、各アダプタの名前がそのシステム上で一意になるようにシステムが割り当てる文字です。4 番目の文字は必ずゼロです。

A.6.4 手順 4: SCSI コンソール・パラメータの表示と設定

SCSI OpenVMS Cluster システムを作成するとき、表 A-6 に表示されるコンソール環境パラメータの設定を確認し、必要に応じて、構成上の要件に従って値をリセットします。

表 A-6 では、SCSI コンソール・パラメータを簡単に説明しています。これらのパラメータやその他のシステム・パラメータの設定の詳細については、各システムのマニュアルを参照してください。

注意

ホスト・アダプタの種類によってコンソール環境パラメータは異なります。個々のアダプタについては、それぞれのインストレーション・ガイドおよびユーザズ・ガイドを参照してください。

表 A-6 SCSI 環境パラメータ

パラメータ	説明
<code>bootdef_dev device_name</code>	デフォルトのブート・デバイスをシステムに指定します。
<code>boot_osflags root_number, bootflag</code>	<code>boot_osflags</code> 変数には、システム・ブートストラップのオプションの аспекト (会話型ブートストラップなど) をオペレーティング・システムに伝える情報を格納します。

(次ページに続く)

表 A-6 (続き) SCSI 環境パラメータ

パラメータ	説明
pk*0_disconnect	ターゲットがコマンドで実行中に、ターゲットの接続を SCSI バスから切断します。このパラメータを 1 に設定すると、ターゲットの接続は、コマンドを処理中に SCSI バスから切断されます。このパラメータを 0 に設定すると、ターゲットがコマンドで実行中は SCSI バスの制御を維持します。
pk*0_fast	SCSI アダプタを Fast SCSI モードで実行します。このパラメータを 1 に設定すると、デフォルトの速度が Fast モードになります。パラメータを 0 に設定すると、デフォルトの速度が Standard モードになります。
pk*0_host_id	ホスト・アダプタの SCSI デバイス ID を 0 から 7 の間に設定します。
scsi_poll	システムが一時停止すると、すべての SCSI インターコネクトにおいて、コンソール・ポーリングを有効にします。
control_scsi_term	システムのパルクヘッドの統合 SCSI インターコネクト上のターミナータを有効または無効にします (一部のシステム対象)。

注意

パラメータを変更するには、まず (対応するコンソール SET コマンドで) パラメータを編集します。次にコンソール INIT コマンドを入力するか、Reset ボタンを押して変更を有効にします。

例

ブート・パラメータを設定する前に、以下の例のようにそれらのパラメータの現在の設定を表示します。

1.

```
>>>SHOW *BOOT*
boot_osflags          10,0
boot_reset            OFF
bootdef_dev           dka200.2.0.6.0
>>>
```

boot_osflags パラメータの最初の番号は、システム・ルートを指定します。(この例で最初の番号は、10 です。) boot_reset パラメータは、ブート・プロセスを制御します。デフォルトのブート・デバイスは、OpenVMS オペレーティング・システムの読み込み元デバイスです。その他のブート・デバイスについては、各システムのマニュアルを参照してください。

システムには複数のブート・デバイスを指定できます。その場合、ユーザが指定したデバイス・リストからブート・デバイスをシステムが検索します。ブート可能なシステム・ソフトウェアを最初に見つけたデバイスからシステムが自動的にブートします。代替デバイス名をブート・コマンド行に指定すれば、デフォルトのブート・デバイスを無効にできます。

一般の環境ではデフォルトのブート・フラグで問題ないはずですが。デフォルトのブート・フラグを無効にするには、`-flags` オプションでブート・コマンド行から動的にブート・フラグを指定します。

2.

```
>>>SHOW *PK*
```

```
pka0_disconnect      1
pka0_fast            1
pka0_host_id         7
```

`pk*0_disconnect` パラメータは、ターゲットがコマンドを実行中に、ターゲットの接続を SCSI バスから切断してよいかどうかを指定します。マルチホスト SCSI バスでは、接続が切断されるように、`pk*0_disconnect` パラメータを必ず 1 に設定します。

`pk*0_fast` パラメータは、SCSI コントローラ上の Fast SCSI デバイスが Standard モードと Fast モードのどちらで実行するかを制御します。パラメータを 0 に設定すると、デフォルトの速度が Standard モードに設定されます。`pk*0_fast` パラメータを 1 に設定すると、デフォルトの速度が Fast SCSI モードに設定されます。この例で、SCSI コントローラ `pka0` 上のデバイスは、Fast SCSI モードに設定されます。この場合、このコントローラに接続された Standard と Fast の両方の SCSI デバイスともそのデバイスに応じた速度で実行されず (つまり、Fast か Standard モード)。

`pk*0_host_id` パラメータは、指定したホスト・アダプタのバス・ノード ID を割り当てます。たとえば、`pka0` には SCSI デバイス ID の 7 が割り当てられています。

3.

```
>>>SHOW *POLL*
```

```
scsi_poll            ON
```

コンソール・モードのときに SCSI デバイスのポーリングを有効または無効にします。

ポーリングの ON, OFF は、サイトのニーズと環境に合わせて設定してください。ポーリングを有効にすると、SHOW DEVICE の出力をいつも最新の状態に維持できます。ただし、ポーリングは SCSI バス帯域幅を消費するので (未使用の SCSI ID に比例)、マルチホスト SCSI バスのシステムのどれかが長時間コンソール・モードになるときはポーリングを無効にするほうが良い場合もあります。

ホット・プラグ操作時に、ポーリングは必ず無効にしておきます。SCSI OpenVMS Cluster 環境におけるホット・プラグについては、第 A.7.6 項を参照してください。

4.

```
>>>SHOW *TERM*
```

```
control_scsi_term    external
```

一部のシステム (AlphaStation 400) で、外部コネクタの次の SCSI ターミナータを有効または無効にするときに使用します。ケーブルをバルクヘッドに接続する場合は、`control_scsi_term` パラメータを `external` に設定します。それ以外は、パラメータを `internal` に設定します。

A.6.5 手順 5: OpenVMS オペレーティング・システムのインストール

OpenVMS オペレーティング・システムのインストールについては、OpenVMS Alpha または VAX のアップグレードとインストレーションのマニュアルを参照してください。OpenVMS Cluster システムでシステム・ディスクごとにインストールを 1 回ずつ実行します。ほとんどの構成がシステム・ディスクは 1 枚です。したがって、この手順はどのシステムでも 1 回で済みます。

インストール時に、システムをクラスタ・メンバにするかどうかを確認するプロンプトが表示されたら、Yes と応答してください。次に、『OpenVMS Cluster システム』のガイドラインに従ってインストールを実行してください。

A.6.6 手順 6: その他のシステムの構成

CLUSTER_CONFIG コマンド・プロシージャで追加システムを構成します。このプロシージャは、SCSI バスで構成した二次ホストに対して 1 回実行します。(詳細については、第 A.7.1 項を参照してください。)

A.7 補足情報

以下の項では、SCSI OpenVMS Cluster システムに関する技術的な細部および概念について説明します。

A.7.1 OpenVMS Cluster 構成コマンド・プロシージャの実行

CLUSTER_CONFIG.COM または CLUSTER_CONFIG_LAN.COM コマンド・プロシージャを実行して、OpenVMS Cluster システムでノードをセットアップし、構成します。ブートに DECnet と LANCP コーティリティのどちらを使用するかで使用するコマンド・プロシージャは異なります。CLUSTER_CONFIG.COM では DECnet を使用し、CLUSTER_CONFIG_LAN.COM では LANCP コーティリティを使用します。(両方のプロシージャの使用方法については、『OpenVMS Cluster システム』を参照してください。)

通常は、イニシャル OpenVMS インストレーション・プロシージャで最初のコンピュータを OpenVMS Cluster システムとしてセットアップします (第 A.6.5 項参照)。CLUSTER_CONFIG プロシージャにより、追加ノードを構成します。ただし、前のインストールでクラスタリングを有効にせずに OpenVMS をインストールしていた場

合は、最初に CLUSTER_CONFIG を実行したときに、このプロシージャによってスタンドアロン・システムがクラスタ・システムに変換されます。

SCSI クラスタに新たなノードを構成するには、追加するノードごとに CLUSTER_CONFIG.COM を実行します。追加 SCSI ノードの構成については、表 A-7 を参照してください。

表 A-7 追加ノードのインストール手順

ステップ	手順
1	最初のノードから、CLUSTER_CONFIG.COM 手順を実行し、ADD のデフォルト・オプション[1]を選択します。
2	続行を確認するプロンプトが CLUSTER_CONFIG.COM によって表示されたら、Yes と応答します。
3	既存のシングル・ノード・クラスタに追加するノードの DECnet 名とアドレスを指定します。
4	これを共用 SCSI インターコネクトのノードにすることを確認します。
5	このノードをサテライトにするかどうかを確認するプロンプトが表示されたら No と応えます。
6	他のクラスタ・メンバにサービスを提供する場合、そのディスク・サーバにするノードを構成します。
7	新しいノードのシステム・ルートを所定のデフォルトのデバイスに設定します。
8	新しいノードのシステム・ルートを選択します。最初のノードは SYS0 を使用します。デフォルト (最初の追加ノードは SYS10) を選択するか、自分で番号付けの方式を選択します。SYS1 から SYSn まで選択できます。n は 16 進数の FFFF です。
9	クラスタの新しいノードが最初のノードと同じ ALLOCLASS を使用するよう、デフォルトのディスク割り当てクラスを選択します。
10	クォーラム・ディスクがあるかどうかを確認します。
11	ページ・ファイルとスワップ・ファイルのサイズに関する質問に応答します。
12	CLUSTER_CONFIG.COM が終了したら、新しいシステム・ルートから新しいノードをブートします。たとえば、ディスク DKA200 の SYSFF の場合、以下のコマンドを入力します。 BOOT -FL FF,0 DKA200 BOOT コマンドでは、以下のフラグを使用します。 <ul style="list-style-type: none">• -FL は、ブート・フラグを表します。• FF は、新しいシステム・ルートです。• 0 は、会話型ブートなど特別なブート要件がないことを表します。

CLUSTER_CONFIG.COM プロシージャを実行すると、SCSI クラスタに追加ノードをセットアップできます。例 A-2 に例を示します。

例 A-2 SCSI クラスタにノードを追加

```
$ @SYS$MANAGER:CLUSTER_CONFIG
```

Cluster Configuration Procedure

Use CLUSTER_CONFIG.COM to set up or change an OpenVMS Cluster configuration. To ensure that you have the required privileges, invoke this procedure from the system manager's account.

Enter ? for help at any prompt.

1. ADD a node to a cluster.
2. REMOVE a node from the cluster.
3. CHANGE a cluster member's characteristics.
4. CREATE a duplicate system disk for CLU21.
5. EXIT from this procedure.

Enter choice [1]:

The ADD function adds a new node to a cluster.

If the node being added is a voting member, EXPECTED_VOTES in every cluster member's MODPARAMS.DAT must be adjusted, and the cluster must be rebooted.

WARNING - If this cluster is running with multiple system disks and if common system files will be used, please, do not proceed unless you have defined appropriate logical names for cluster common files in SYLOGICALS.COM. For instructions, refer to OpenVMS Cluster Systems manual.

Do you want to continue [N]? y

If the new node is a satellite, the network databases on CLU21 are updated. The network databases on all other cluster members must be updated.

For instructions, refer to OpenVMS Cluster Systems manual.

What is the node's DECnet node name? SATURN

What is the node's DECnet node address? 7.77

Is SATURN to be a clustered node with a shared SCSI bus (Y/N)? y

Will SATURN be a satellite [Y]? N

Will SATURN be a boot server [Y]?

This procedure will now ask you for the device name of SATURN's system root. The default device name (DISK\$BIG_X5T5:) is the logical volume name of SYS\$SYSDEVICE:.

What is the device name for SATURN's system root [DISK\$BIG_X5T5:]?

What is the name of SATURN's system root [SYS10]? SYS2

Creating directory tree SYS2 ...

System root SYS2 created

(次ページに続く)

例 A-2 (続き) SCSI クラスタにノードを追加

NOTE:

All nodes on the same SCSI bus must be members of the same cluster and must all have the same non-zero disk allocation class or each will have a different name for the same disk and data corruption will result.

Enter a value for SATURN's ALLOCLASS parameter [7]:

Does this cluster contain a quorum disk [N]?

Updating network database...

Size of pagefile for SATURN [10000 blocks]?

.
.
.

A.7.2 マルチホスト SCSI 環境におけるエラー・レポートと OPCOM メッセージ

マルチホスト SCSI バスでホストのブートやシャット・ダウンなどの一般的な操作では、SCSI バスの他のホストにエラーが発生することがあります。また、シングル・ホスト SCSI 構成ではまれにしか発生しないエラーが、マルチホスト SCSI バスで発生する場合があります。

このようなエラーを一時的エラーと言い、OpenVMS がエラーを検出、レポートし、回復処理を行っても実行中のアプリケーションのデータが失われたり影響を受けることはありません。この項では、このようなエラーの発生要件、および、オペレータ・コンソールに表示されるエラー・ログに入力されるメッセージについて説明します。

A.7.2.1 SCSI バスのリセット

電源入力やリブートで SCSI バスに接続されたホストを最初に起動する際は、SCSI バスやその上のデバイスの状態はホストにはわかりません。ANSI SCSI 標準では、BUS RESET という方式でバスやその上のデバイスを既知の状態に強制的に設定する機能を定めています。ホストは、最初のスタートアップとシャットダウンのときに、各 SCSI バス上で通常は RESET シグナルを 1 回以上表明します。RESET を表明しているホストではこれが正常な動作ですが、RESET ではホストの強制終了と、実行中のすべての I/O 操作の再起動が必要になるため、他のホストにとって RESET シグナルはエラー扱いになります。

ホストが正常な処理の途中で、他の方法では修正できない問題を検出したときもホストはバスをリセットします。このようリセットは、一般的ではありませんが、通常はバス上の何かが乱れたときに発生します。たとえば、SCSI デバイスがアクティブなときにホット・プラグしようとしたり(第 A.7.6 項参照)、Ctrl/P でホストのどれかを一時停止しようとする、1 つ以上のホストがバス・リセットを発行しなければならない条件が揃います。

A.7.2.2 SCSI タイムアウト

ホストが SCSI バス上のデバイスとデータを交換するときに、ホストがデバイスや SCSI アダプタの反応を待機しなければならないポイントが数箇所にあります。OpenVMS システムでは、待機中にホストを他の処理に振り向けることができますが、待機時間が長くなりすぎるのを防ぐため、タイマが起動します。SCSI デバイスやアダプタからの応答が来る前にタイマが時間切れになった場合、その状態をタイムアウトと呼びます。

タイムアウトには 3 種類あります。

- 切断タイムアウト (Disconnect timeout) — デバイスがホストからコマンドを受信し、コマンドの処理時にバスとの接続を切断したが、バスを再接続できずトランザクションを終了できなかった。このエラーは、バスのビジーが続いて発生するのが一般的です。詳細については、第 A.7.5 項を参照してください。切断タイムアウト時間はデバイスによって異なりますが、ほとんどのディスクでは約 20 秒に設定されています。
- 選択タイムアウト (Selection timeout) — ホストから SCSI バス上のデバイスにコマンドを送信しようとしたが、デバイスが応答しない。デバイスが存在しないか、バスから取り外されているか、電源が入っていないというのがこのタイムアウトの原因のほとんどです。(このエラーは、マルチ・イニシエータ・システムではあまり発生しません。) 選択タイムアウト時間は約 0.25 秒です。
- 割り込みタイムアウト (Interrupt timeout) — ホストからアダプタに他の何らかの理由による応答が要求されたが、応答がない。このエラーは、通常は SCSI バスのビジー状態を表します。このエラーの原因としては、イニシエータの番号を上位番号 (6 や 7) ではなく下位番号 (0 か 1) にした場合に多く見られます。割り込みタイムアウト時間は、約 4 秒です。

SCSI OpenVMS Cluster システムでは、タイムアウト・エラーは、避けられないわけではありません。しかし、トラフィックがヘビーな SCSI バスやイニシエータが 2 つある SCSI バスでは、さらに発生頻度が高くなります。タイムアウト・エラーが、必ずしもハードウェアやソフトウェアの問題を意味するわけではありません。頻繁に記録されるようなら、SCSI バスの負荷の軽減を検討してください(たとえば、新規バスの追加など)。

A.7.2.3 マウント検査

マウント検査は、ホストがデバイスについて宣言する状態です。バス・リセットやタイムアウトなど数多くの一時的なエラーの兆候が続くと、ホストがこの状態を宣言します。デバイスがマウント検査状態にある場合、正しいデバイスの存在と、デバイスのアクセス可能状態が確認されるまで、ホストは正常な I/O を中断します。再開後、マウント検査処理は正しいデータの読み書きができる状態で、中断していた I/O を再試行します。アプリケーション・プログラムにとって、マウント検査の状態が発生してもマウント検査が終了するまでは認識できません。

ホストが一定の時間内に正しいデバイスをアクセスできない場合、ホストはマウント検査タイムアウトを宣言し、アプリケーション・プログラムにはデバイスが利用できないことが通知されます。ホストがマウント検査タイムアウトを宣言したら、手作業でデバイスをサービス状態に復元する必要があります。マウント検査タイムアウトは、通常はエラーが一時的ではないことを表します。システム管理者は、マウント検査のタイムアウト時間を指定できます。デフォルトでは1時間になります。

A.7.2.4 シャドウ・ボリューム処理

シャドウ・ボリューム処理は、マウント検査に似ていますが、シャドウ・セット・メンバが対象です。シャドウ・セットのどれかのメンバにエラーが発生すると、そのセットはボリューム処理状態になり、OpenVMS がメンバのアクセスを再度確立できるまで I/O はブロックされます。シャドウ・ボリューム処理がタイムアウトになる前にアクセスが確立されると、中断していた I/O が再発行され、シャドウ・セットは正常動作に戻ります。タイムアウトが発生すると、障害の発生したメンバはセットから削除されます。システム管理者は、システム・ディスク・シャドウ・セットのタイムアウト値と、アプリケーション・シャドウ・セットのタイムアウト値を1つずつ設定できます。両方のタイムアウトのデフォルトは20秒です。

注意

SCSI 切断タイムアウトとデフォルトのシャドウ・ボリューム処理タイムアウトは同じです。SCSI バスの負荷が大きく、切断タイムアウトが発生しそうな場合は、シャドウ・ボリューム処理タイムアウトの値を大きくします。(推奨値は60秒です。) これにより、切断タイムアウト・エラーになってもシャドウ・セット・メンバが切断されるのを防ぐことができます。

A.7.2.5 マルチホスト SCSI 環境で表示される OPCOM メッセージ

バス・リセットが発生すると、マウント・ディスクがマウント検査やシャドウ・ボリューム処理を出入りするたびに OPCOM メッセージが表示されます。

ドライブの I/O でタイムアウト・エラーになると、マウント・ディスクがマウント検査やシャドウ・ボリューム処理を出入りするたびに OPCOM メッセージが表示されません。

クォーラム・ディスクや共用 SCSI バスに以上のどちらかのエラーが発生すると、さらに追加の OPCOM メッセージが表示され、クォーラム・ディスクとの接続が失われ、再接続されたことを表します。

A.7.2.6 エラー・ログの基本

OpenVMS システムでは、デバイス・ドライバに異常が発生すると、Error Log コーティリティでその関連情報を保存できます。従来は、ほとんどの例外条件はハードウェア障害、ソフトウェア障害、一時条件(ケーブルのゆるみなど)などのエラーの結果として発生していました。

DCL コマンド SHOW ERROR を入力すると、システムは、システムのブート後に発生したエラーの要約を表示します。以下に例を示します。

```
$ SHOW ERROR

Device                               Error Count
SALT$PKB0:                            6
$1$DKB500:                             10
PEA0:                                  1
SALT$PKA0:                              9
$1$DKA0:                                0
```

この場合、SALT の SCSI ポート B (PKB0) における 6 つのエラーと、ディスク \$1\$DKB500 に対する 10 のエラー、その他が記録されています。

これらのエラーの詳細を確認するには、コマンド ANALYZE/ERROR/SINCE= dd-mm-yy:hh:mm:ss を DCL プロンプトで実行します。このコマンドからの出力では、エラー・ログ・エントリのリストが以下のような情報とともに表示されます。

```
***** ENTRY      2337. *****
ERROR SEQUENCE 6.          LOGGED ON: CPU_TYPE 00000002
DATE/TIME 29-MAY-1995 16:31:19.79      SYS_TYPE 0000000D

<identification information>

      ERROR TYPE          03          COMMAND TRANSMISSION FAILURE
      SCSI ID              01          SCSI ID = 1.
      SCSI LUN             00          SCSI LUN = 0.
      SCSI SUBLUN          00          SCSI SUBLUN = 0.
      PORT STATUS          00000E32    %SYSTEM-E-RETRY, RETRY OPERATION

<additional information>
```

この例では、ERROR TYPE に注目してください。また、場合によっては、PORT STATUS フィールドも重要です。この例で、エラー・タイプは 03、COMMAND TRANSMISSION FAILURE、ポート状態は 00000E32、SYSTEM-E-RETRY となっています。

A.7.2.7 マルチホスト SCSI 環境におけるエラー・ログ・エントリ

この項で示したエラー・ログ・エントリは、マルチホスト SCSI 構成で発生するエラーを想定しており、一般には気にかける必要はありません。ただし、この項で紹介した以外のメッセージのエラー・ログ・エントリはよく確認してください。

- ERROR TYPE 0007, BUS RESET DETECTED

他のシステムが SCSI バスのリセット・シグナルを表明すると発生します。原因は次のとおりです。

- システムの電源入力時の実行診断テストが実行された。
- コンソール INIT コマンドが実行された。
- EISA 構成ユーティリティ (ECU) が実行された。
- コンソール BOOT コマンドが実行された (この場合、数回のリセットが実行されます)。
- システム・シャットダウンが終了した。
- システムがアダプタや SCSI バスに問題を検出した (割り込みタイムアウトなど)。

このエラーが発生すると、マウントされているすべてのディスクについてマウント検査が開始します。

- **ERROR TYPE 05, EXTENDED SENSE DATA RECEIVED**

SCSI バスをリセットしたときに、イニシエータは各デバイスから“センス・データ”を受け取る必要があります。このデータを受け取ったイニシエータは、EXTENDED SENSE DATA RECEIVED エラーをログします。これは正規の動作です。

- **ERROR TYPE 03, COMMAND TRANSMISSION FAILURE
PORT STATUS E32, SYSTEM-E-RETRY**

他のディスクとエラー情報を交換中のディスクに、ホストからコマンドが送信されることがあります。その場合、大半のディスクは SCSI “BUSY”コードを戻します。この SCSI BUSY コードに対し、OpenVMS システムはエラーをログし、操作を再試行します。このエラーが最もよく発生するのは、バスをリセットして間がない状況です。通常、リセット直後には発生しないはずですが、直後に発生する場合は正規の動作なので回避できません。

- **ERROR TYPE 204, TIMEOUT**

割り込みタイムアウトが発生しました (第 A.7.2.2 項参照)。このエラーが発生するとディスクのマウント検査が開始します。

- **ERROR TYPE 104, TIMEOUT**

選択タイムアウトが発生しました (第 A.7.2.2 項参照)。このエラーが発生するとディスクのマウント検査が開始します。

A.7.3 制限事項と既知の問題

複数のホストを同じ SCSI バスに構成した OpenVMS Cluster ソフトウェアには、以下の制限事項があります。

- OpenVMS Alpha バージョン 7.2 より前のバージョンでは、ノードからディスクへのアクセスは、直接 SCSI パスから MSCP サービスによるパスにフェールオーバーしません。

MSCP サービスによるパスから直接 SCSI パスへのフェールオーバーも実装されていません。一般に、この種のフェールオーバーは、重要ではありません。

OpenVMS が直接パスとサービスによるパスの両方を検出すると、OpenVMS は必ず直接パスを選択するためです。ただし、MSCP サービスによるパスが最初のパスとして使用可能になり、直接パスが使用可能になる前に OpenVMS によって選択されるのを避けたい場合もあります。これを避けるには、以下の規則に従います。

- SCSI システム・ディスクまでの直接パスがあるノードは、必ず SCSI ポートから直接ブートし、LAN は経由しないこと。
- ノードが MSCP サーバを実行している場合、第 2 のノードがブート後は、SCSI ディスクをマルチホスト SCSI パスに追加しないこと (物理的な挿入や HSZxx のいずれも不可)。

2 個のノードのブート後にデバイスを追加し、SYSMAN でデバイスを構成すると、直接パスが認識されるまでは、サービスによるパスを介していずれかのシステムに認識されます。イベントのタイミングにもよりますが、以下の手順でこの問題を回避できる場合もあります。

```
$ MCR SYSMAN
SYSMAN> SET ENVIRONMENT/CLUSTER
SYSMAN> IO AUTOCONFIGURE
```

新しいデバイスまでの直接パスを確保するには (HSZxx 仮想デバイスも含む)、デバイスの追加後はノードをリブートしてください。

- OpenVMS Alpha バージョン 7.2 より前のバージョンでは、デバイスまでのパスが 2 本ある場合、\$DEVICE_SCAN システム・サービスと FS\$DEVICE レキシカル関数により、共用バス上のデバイスが 2 回リストされます。非 SCSI システム・ディスクをブートすると、DCL コマンド SHOW DEVICE の出力に共用バス上のデバイスも 2 回リストされます。この二重リストは、表示プログラムのエラーです。これは特に問題ではなく、また、直接 SCSI パスの代わりに MSCP サービスによるパスが使用されることを意味しているわけでもありません。
- システムの電源入力時、ブート時、またはシャット・ダウン時に、システムは SCSI バスをリセットします。このリセットにより、SCSI バス上の他のホストが I/O エラーになります。Files-11 ボリュームの場合、このようなエラーはマウント検査ファシリティによって自動的に回復され、中断していた I/O が完了します。その結果、エラーなしでユーザ・プロセスが続行します。

このレベルのエラー回復は、/FOREIGN 修飾子でマウントしたボリュームには適用できません。そのため、バスのリセット発生時に I/O が中断するとユーザ・プロセスは I/O エラー通知を受け取ります。

マルチホスト SCSI バスには、できるだけフォーリン・デバイスをマウントしないでください。どうしてもフォーリン・デバイスを共用バスにマウントしなければならない場合は、フォーリン・デバイスに I/O を実行中にその共用バス上のシステムが SCSI バス・リセットを表明しないようにしてください。

- マルチホスト SCSI バス上で ARC コンソールが有効になっていると、すべてのローカル・ホスト・アダプタの SCSI ターゲット ID は 7 に設定されます。ホストやデバイスが ID 7 のバスにあると、この設定で SCSI ID 競合が発生します。この種の競合が発生すると、一般にそのバスだけでなく、場合によってはそのバス上のすべてのシステムもハングしてしまいます。

ARC コンソールの使用対象は、KZPSA 構成ユーティリティなど一定のプログラムのアクセスに限定してください。ARC コンソールを実行するときは、まずマルチホスト SCSI バスや SCSI ID 7 でデバイスを接続しているバスからシステムを切り離します。

- システムの電源入力時、ブート時、シャットダウン時に SCSI バス・リセットが発生すると、その SCSI バス上の他のシステムがエラーをログし、OPCOM メッセージが表示されます。これは正規の動作であり、問題ではありません。
- マルチホスト SCSI バス上のシステムを急激にホルトすると (コンソールで Ctrl/P を押すなど)、KZPAA SCSI アダプタは同じバス上の他のホストの操作を干渉するようになります。システムが急激に停止した場合は、できるだけ早く初期化、ブート、または復帰 (CONTINUE) を続行してください。
- マイクロコードを更新中は、ディスク・ドライブに対する I/O をすべて停止してください。これについては、シングル・ホスト環境の場合よりもマルチホスト環境の方で特に注意が必要です。必要な手順については、第 A.7.6.3 項を参照してください。
- EISA 構成ユーティリティ (ECU) を実行すると多数の SCSI バスをリセットできます。このリセットにより、同じ SCSI バス上の他のシステムは、そのバス上の I/O サブシステムが回復するまで一時停止します。ECU が実行中は、共用 SCSI バス上のシステムをシャットダウンすることをお勧めします (必須ではありません)。

OpenVMS Cluster システムでは、SCSI クォーラム・ディスクがシングル・ホスト SCSI バスにあるかマルチホスト SCSI バスにあるかに関わらず、制限事項が 1 つあります。SCSI クォーラム・ディスクは、タグ付きコマンド・キューイング (TCQ) をサポートしなければなりません。この制限事項は、クォーラム I/O が OpenVMS SCSI ドライバで受け取る特別な処理に必ず適用します。

いずれにしてもマルチホスト SCSI バス上のディスクは、すべてタグ付きコマンド・キューイングをサポートする必要があり (第 A.7.7 項参照)、クォーラム・ディスクは一般にはシングル・ホスト・バスでは使用しないのでこの制限事項をそれほど気にする必要はありません。

A.7.4 トラブルシューティング

以下の項では、SCSI インターコネクトを使用する OpenVMS Cluster システムの代表的な問題について説明します。

A.7.4.1 終端問題

すべての SCSI インターコネクトに 2 つのターミネータがあるか確認してください (インターコネクトの両端に 1 つずつ)。外部からは見えませんが、BA350 エンクロージャ、BA356 エンクロージャ、DWZZx、KZxxx の各アダプタには内部ターミネータがあります (第 A.4.4 項参照)。

A.7.4.2 正しくない構成が原因のブート障害やマウント障害

OpenVMS では、この項で説明するエラーが自動的に検出され、バグチェックやディスク・マウントの拒否により、このような構成エラーが原因のデータ・ロスを回避しています。

A.7.4.2.1 ブートストラップ・プロセス間のバグチェック

OpenVMS Alpha バージョン 7.2 より前のバージョンでは、バグチェックを実行する原因になるブート時の構成エラーが 3 種類ありました。バグチェック・コードは、VAXCLUSTER、Error detected by OpenVMS Cluster software です。

OpenVMS がブートすると、すべての SCSI ID に照会コマンドを送信して SCSI バス上のデバイスを認識します。デバイスは照会を受け取ると、ディスク、テープ、またはプロセッサのどれであるかを表すデータを戻します。

プロセッサ・デバイス (ホスト・アダプタ) によっては、オペレーティング・システムの助けがなくてもこの照会に回答しますが、他のデバイスは、オペレーティング・システムが実行していないと回答できません。OpenVMS Cluster システムでサポートしているアダプタには、オペレーティング・システムの助けが必要です。このようなアダプタは、OpenVMS の助けを借りて照会に対する回答により情報を伝え、回答の受信側は、以下の構成エラーを検出します。

- 同じ SCSI バス上に異なるデバイス名がある。

ポート割り当てクラスを使用する場合を除き、SCSI バス上の各アダプタの OpenVMS デバイス名は同一でなければなりません (すべて PKC0 など)。同一にしない場合、OpenVMS Cluster ソフトウェアはホストからストレージまでのアクセスを管理できなくなります (第 A.6.2 項と第 A.6.3 項参照)。

OpenVMS は、照会回答でコントローラ名を送信し、これを自動的にチェックします。ブート・システムはこの回答を受け取り、リモート・コントローラ名とローカル・コントローラ名を比較します。一致しない場合は、OPCOM メッセージが出力され、VAXCLUSTER バグチェックとともにシステムが停止してデータ・ロスを防ぎます。詳細については、Help Message ユーティリティの NOMATCH エラーの説明を参照してください。(Help Message ユーティリティを NOMATCH に使用するには、DCL プロンプトで HELP/MESSAGE NOMATCH と入力します。)

- 割り当てクラスの値が異なるかゼロである。

SCSI バス上の各ホストのディスク割り当てクラスの値は、ゼロ以外の同じ値であるか、対応するポート割り当てクラスの値とします。それ以外の値では、OpenVMS Cluster ソフトウェアはホストからストレージへのアクセスを管理できません (第 A.6.2 項と第 A.6.3 項参照)。

OpenVMS は照会応答の中で、必要な情報を送信してこれを自動的にチェックします。ブート・システムは応答を受け取ると、リモート値をローカル値と比較します。一致しない場合または値がゼロの場合、OPCOM メッセージが出力され、システムは VAXCLUSTER バグチェックを出力して停止し、データ・ロスを防ぎます。Help Message コーティリティの ALLODIFF エラーと ALLOZERO エラーの説明を参照してください。

- サポートされていないプロセッサ

OpenVMS を実行していなかったり、構成のチェックに必要なコントローラ名や割り当てクラスの情報を戻さないプロセッサが、SCSI バスにあります。ブート・システムが照会応答を受け取り、その応答に OpenVMS の特別な構成情報が含まれていない場合、OPCOM メッセージが出力され、VAXCLUSTER バグチェックが実行されます。Help Message コーティリティの CPUNOTSUP エラーの説明を参照してください。

システムの SCSI バスにプロセッサ・デバイスが必要な場合、SYSGEN の特別なパラメータ (この場合は SCISICLUSTER_Pn) については、Help Message コーティリティの CPUNOTSUP メッセージの説明を参照してください。

A.7.4.2.2 デバイス構成上の障害

OpenVMS Alpha バージョン 7.2 では、正しく構成されていないバス上の SCSI デバイス (第 A.7.4.2.1 項参照) は構成されません。そして不正構成を記述するエラー・メッセージが表示されます。

A.7.4.2.3 マウント障害

ディスクのマウントが失敗する原因の構成エラーには 2 種類あります。

まず、共用 SCSI バス上のディスクからシステムがブートするときにシステム・ディスクのマウントが失敗することがあります。これは、同じ SCSI バスにブート済みの別のシステムがあって、問題のシステム・ディスクにそのシステムが別のデバイス名を使用している場合に発生します。(前の項で説明したように、2 つのシステムが使用しているコントローラ名や割り当てクラスの構成が間違っていると、共用バス上のデバイス名に関する競合が発生します。) 前の項で説明したバグチェックを先に実行しない場合は、コンソールに以下のエラー・メッセージが表示されます。

```
%SYSINIT-E- error when mounting system device, retrying..., status = 007280B4
```

この状態をデコードすると次のようになります。

```
VOLALRMNT, another volume of same label already mounted
```


このエラーは、システム・ディスクがすでにマウント済みであり、OpenVMS Cluster システムの他のドライブ名として扱われているようなので再度マウントはできない、ということを示しています。これを解決するには、共用 SCSI バスのノードごとに、コントローラ名と割り当てクラスの色をチェックします。

もう 1 つの構成エラーは、ディスクがタグ付きコマンド・キューイング (TCQ) をサポートしていないと、共用 SCSI バス上の SCSI ディスクは両方のシステムでマウントできません。TCQ が OpenVMS Cluster の状態遷移で必要とされるコマンド整列保証動作を渡すからです。

OpenVMS は、第 A.7.4.2.1 項で説明する機構を利用して認証時に、SCSI バス上に他のプロセッサがあるか判断します。SCSI バスにおける別のホストの存在は、システムをリブートするまで記録として保存されます。

この情報は、非 TCQ デバイスをマウントするときに使用します。デバイスがマルチホスト・バスにない場合、マウントは失敗し、以下のメッセージが戻ります。

```
%MOUNT-F-DRVERR, fatal drive error.
```

同じ SCSI 上の複数のホストでこのドライブをマウントする場合は、TCQ をサポートするドライブに置き換える必要があります。

マルチホスト SCSI バスで最初にブートするプロセッサは、他のホストが OpenVMS をまだ実行していないため、他のホストからの照会応答を受け取りません。したがって、最初にブートするシステムは、同じバスに複数のホストがあることを認識しないので非 TCQ ドライブを共用バスにマウントできます。その SCSI バスの他のホストは、最初のホストを検出しますが、デバイスのマウントはできません。2 つのプロセッサが同時にブートすると、互いの存在を検出し、どちらも共用バス上に非 TCQ ドライブをマウントできなくなります。

A.7.4.3 接地

接地オフセット電圧が高すぎたり、最大 SCSI インターコネクト長を超えると、システム障害が生じたりパフォーマンスが低下したりします。SCSI 接地要件の詳細については、第 A.7.8 項を参照してください。

A.7.4.4 インターコネクトの長さ

シグナルの一貫性を確保するには、SCSI バスの長さを厳しく守る必要があります。バスの長さの推奨値を守らないと、診断が困難な問題が発生します (断続的なエラーなど)。SCSI バスの長さについては、第 A.4.3 項を参照してください。

A.7.5 SCSI アービトレーション上の注意

SCSI バスはどの一瞬をとってみても、制御できるイニシエータ (通常は、ホスト・システム) やターゲット (通常は、周辺機器) は 1 つだけです。複数のターゲットによる SCSI バスのアクセスで混雑しているコンピューティング環境では、これらのターゲットの一部でスループットに関する問題が発生することがあります。この項では、

SCSI バスの制御，その制御がコンピューティング環境に及ぼす影響，および，最良の結果を得るために何ができるかについて説明します。

SCSI バスの制御は常に変化します。イニシエータから SCSI ターゲットにコマンド (READ 等) を発行すると，ターゲットはコマンドを処理している間，通常，SCSI がバスとの接続を切断して，他のターゲットやイニシエータにそのバスを開放します。ターゲットがコマンドに対する応答準備を整えたら，再び SCSI バスの制御を取り戻す必要があります。同じく，イニシエータがターゲットにコマンドを送信するときも，SCSI バスの制御をとる必要があります。

複数のターゲットとイニシエータがバスの制御を同時に要求した場合，バスの所有権は，SCSI 標準に定められたアービトレーションというプロセスで決まります。デフォルトのアービトレーション規則は単純です。つまり，バスの制御は，要求イニシエータや最上位のユニット番号を持つターゲットに与えられます。

以下の項では，アービトレーションの意味と，環境に関係のあるアービトレーション状態に対する応答方法を説明します。

A.7.5.1 マルチディスク環境におけるアービトレーション問題

バスがあまりビジーでなく，バスの競合もあまりない場合は，単純なアービトレーション方式で，システム上のすべてのデバイスに対し，その I/O 要求に十分に答えることができます。しかし，イニシエータからの I/O 要求が増えれば増えるほど，バスの競合がますます通常の状態になってきます。その結果，ID 番号が下位のターゲットは，バス上の他のデバイス (特に ID 番号が上位のターゲット) によって頻繁に I/O 要求を阻止されるため，そのパフォーマンスが低下します。ある程度バスがビジーになると，下位のターゲットは要求を達成することができなくなります。このような状況では，同時に一時停止になるコマンドが増えるため，イニシエータが複数あるシステムで発生しやすくなります。

OpenVMS システムでは，I/O 要求の処理にかかる時間を監視して下位ターゲットが完全にブロックされるのを防いでいます。指定した時間内に要求が処理されない場合，その I/O が終了するまで OpenVMS システムは新しい要求の送信を停止します。このアルゴリズムは，すべてのターゲットがバスを均等にアクセスできることを保証するものではありませんが，下位番号のターゲットが完全にブロックされることはなくなります。

A.7.5.2 アービトレーション問題の対策

I/O に大きな負荷がかかっている時間帯に，サービスを迅速に受けられないディスクが見つかった場合は，サイトの状況に合わせて以下の対策をいくつか実行してみてください。

- DWZZH-05 SCSI ハブをアクセスし，フェア・アービトレーション機能を有効にする。
- 応答時間を短くしたいディスクには上位の ID 番号を与える。
- ディスクを分散させる SCSI バスを増やす。

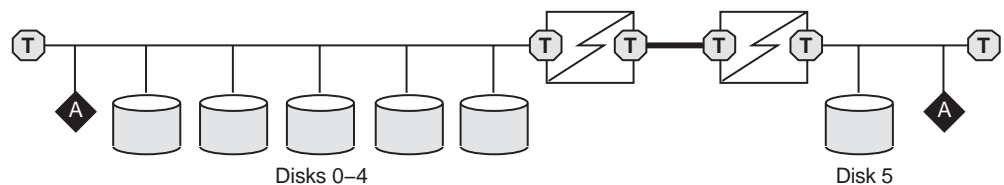
- アクセスの必要なディスクのアクセス元を、非共用 SCSI バス上では 1 つのホストに限定する (たとえばページ・ディスクやスワップ・ディスク)。

下位と上位の ID 番号のディスクに対するサービスをより平均化できるもう 1 つの方法として、ホスト ID を最上位ではなく最下位の番号 (0 か 1) にする方法があります。この方式では、ホストはバスの制御が得られないので、最下位の ID のディスクも含め、すべてのディスクがバスを必要とする限り、新しいコマンドを送信できなくなります。この選択肢では一定の状況下では公正さを達成できるものの、以下のような理由から、多くの場合あまり望ましい構成ではありません。

- 合計スループットが低下することがある。
- コマンドを 2, 3 秒で送信できないとタイムアウト条件になることがある。
- 物理的な構成がむずかしくなる。たとえば、BA350 などの StorageWorks シェルフに ID 7 のディスクを装着するスロットがなく、ID 0 のディスク用のスロットがあると仮定します。ホストを ID 0 に変更すると、BA350 にあるスロット 0 からディスクを取り出さなければなりません、ID 7 には移動できません。ID 0 と 1 の 2 つのホストがある場合、BA350 ではスロット 0 や 1 を使用できません。(ただし、BA353 に ID 7 のディスクを持つことはできます。)

A.7.5.3 アービトレーションとバス・アイソレータ

バス・セグメントを接続する DWZZx などのアクティブ・デバイスでは、あるセグメントのデバイスから別のセグメントのデバイスへのシグナルの受け渡しで、わずかな遅延が生じます。状況によっては、このような遅延がアンフェア・アービトレーションのもう 1 つの原因になることがあります。たとえば、大きなワーク・ロードの下ではディスク・サービス問題 (枯渇) が発生する場合があります。



ZK-7913A-GE

ディスク 5 の ID 番号は最上位ですが、ディスク 5 からのバスへのアクセス順位が最下位になる場合もあります。そのような状態は、下位番号のディスクのどれかがバスの制御をとり、バスの制御が必要な操作を終了した後に発生します。この時点でディスク 5 は、バスが開放されていることを認識できないので、バスの制御のアービトレーションまで待機します。その結果、バスが開放されたことを認識して、バス要求を発行した下位ディスクのどれかがそのバスの制御を得ます。

この種の問題は、以下の対策で削減できます。

- すべてのディスクを同じバス・セグメントに配置する。

- すべてのディスクを同じバス・セグメントに配置できない場合は(たとえば、独立した RZ28 ディスクと HSZxxがある場合)、すべてのディスクのペアの間にアイソレータが1つしかない構成にしてみる。
- ディスク・ペア間にアイソレータが必要な構成の場合(たとえば、距離の要件を満たすため)、各バス・セグメント上のディスク数のバランスを調整する。
- 第 A.7.5.2 項の指示に従って論理バス上のトラフィックの合計数を削減する。

A.7.6 OpenVMS Cluster システムが実行中の SCSI デバイスの取り外しや取り付け

手順が正しければ、バスの実行中の操作を割り込むことなく一定の SCSI デバイスはアクティブな SCSI バスに挿入したり取り外すことができます。この機能をホット・プラグと呼びます。ホット・プラグでは、障害が発生した構成要素を交換中でも、OpenVMS Cluster システムの処理を続行できるような構成が可能です。ホット・プラグ機能がない場合は、SCSI バスに対するデバイスの追加、削除をする前に、SCSI バスを非アクティブにし、SCSI バスの電源を切る必要があります。

SCSI OpenVMS Cluster システムでホット・プラグを利用するには、バス上のすべてのデバイスが一定の電気的特性を備え、SCSI バスに正しく構成できることが前提になります。ホット・プラグを成功させるには、この項で説明する手順に正しく従う必要があります。これらの手順では、アクティブ・バス・シグナルの邪魔にならないよう、ホット・プラグしたデバイスは非アクティブになります。

— ストレージ・コントローラの背後の SCSI バスのホット・プラグ —

この項では、OpenVMS を実行中のホストと同じ SCSI バス上にあるデバイスのホット・プラグ手順を説明します。手順の内容は、ストレージ・コントローラが背後にある HSZxxなどの SCSI バスでは異なります。デバイスのホット・プラグ手順については、ストレージ・コントローラのマニュアルを参照してください。

A.7.6.1 ホット・プラグを説明するための用語

この項で太字になっている用語は、プラグ規則と手順で使用する言葉です。

- SCSI バスセグメントは、2つのターミネータ、その間の導通を果たす電気バス、そして場合によってはスタブを接続して構成します。バス・セグメントは、バス・アイソレータですべて接続でき(たとえば、DWZZx)、これによって論理 SCSI バスや通常の SCSI バスを構成します。
- セグメントの接続方法は2種類あります。それは2つのターミネータ間のバスを分割するバス接続と、スタブのすべて、あるいは一部の接続を切断するスタブ接続です。
- SCSI バス上で、バス・シグナルを1つ以上表明しているデバイスはアクティブです。デバイスがバス・シグナルを表明していないとき、そのデバイスは非アクティブです。

バス・アイソレータに接続されているセグメントは、そのセグメント上のすべてのデバイス (場合によってはバス・アイソレータを除く) が非アクティブのとき非アクティブになります。

- バス・アイソレータ上のポートは、両端が終端されているセグメントに接続され、SCSI-2 要件に準拠した TERMPWR がある場合、正しい終端を持つと言います。

A.7.6.2 ホット・プラグの規則

ホット・プラグの計画と実行にあたっては、以下の 3 つの規則に従ってください。

- ホット・プラグするデバイスと、同じセグメント上の他のすべてのデバイスは、SCSI-3 Parallel Interface (SPI) Standard (作業ドラフト X3T10/855D) の Annex A と A.4 の電氣的要件に準拠しなければなりません。SCSI-2 標準ではホット・プラグの要件が十分に定義されていないのでこのドラフト標準を参照してください。SPI マニュアルでは、ホット・プラグを実行するセグメントのレシーバとターミネータの要件、および、プラグするデバイスのトランシーバ、TERMPWR、終端、電源/接地/シグナルのシーケンスの要件を指定しています。

- ホット・プラグはスタブ接続でのみ実行してください。

ホット・プラグしたデバイスでは、SCSI バスと 1 個でしか接続できず、デバイスで SCSI バスを終端することはできず、またデバイスの接続はスタブの許容最大長を超えることができないことを意味します。図 A-3 を参照してください。有効なホット・プラグである SCSI バス・トポロジの例を図 A-13 に示します。

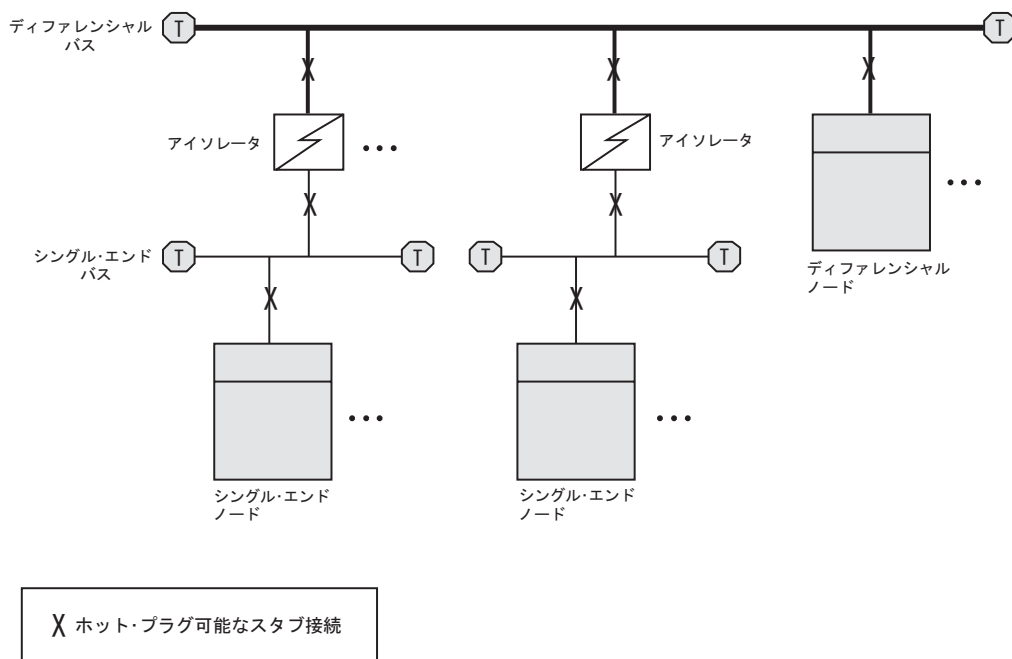
- 静電放電 (ESD) でデバイスが破壊されたり、SCSI バス上のアクティブ・シグナルに障害が発生しないよう配慮してください。SCSI バスの導線が暴露状態にあるときだけでなく、接続や切断プロセスの間は以上の点に注意してください。
- 危険や SCSI バス・シグナル通知の干渉を避けるため、接地オフセット電圧には注意してください。特にシングル・エンドの構成には注意が必要です。接地オフセット電圧の測定と防止手順については、第 A.7.8 項を参照してください。
- ホット・プラグしたデバイスは、接続、切断のプロセス間は非アクティブとします。アクティブな状態でこれらのプロセスを実行すると、SCSI バスがハングすることがあります。OpenVMS は、最終的にはハング・バスを検出してリセットしますが、ハング状態になると一時的に OpenVMS Cluster の操作ができなくなります。

注意

同じ理由から、デバイスは非アクティブな状態で電源を切っておくほうが良いでしょう。

デバイスを確実に非アクティブにする手順については、第 A.7.6.3 項を参照してください。

図 A-13 SCSI バス・トポロジ



ZK8842AGE

- クォーラム・ディスクはホット・プラグできません。クォーラム・ディスクへの I/O を停止する機構がなく、また、交換ディスクに正しいクォーラム・ファイルが格納されないからです。

クォーラム・ディスクを取り外すには、クォーラム・ディスクとしてデバイスを取り外すように OpenVMS Cluster システムを再構成する必要があります。操作手順については、『OpenVMS Cluster システム』を参照してください。

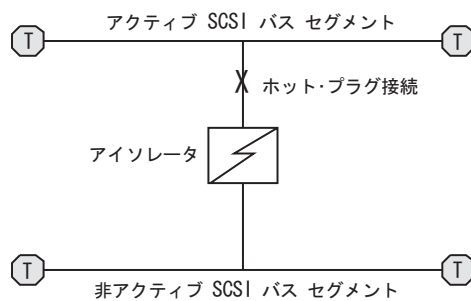
クォーラム・ディスクの可用性を強化するには、HSZxxミラー・セットをクォーラム・ディスクとして使用します。この場合、クォーラム・ディスクの機能を維持しながら、障害が発生したメンバを交換できます。

- ホット・プラグ操作でディスクを取り外し、または交換するときは、前もって論理的にマウントを解除します。これは、そのディスクを非アクティブにし、ファイル・システムの一貫性を保つためです。
- アクティブな SCSI バスに挿入するときは、DWZZxの電源を入れておき、アクティブな SCSI バスに接続されている間は電源を入れたままにしておきます。これは、DWZZxの電源のオン/オフ切り替え時に、接続されているセグメントの操作に悪影響が及ぶのを避けるためです。

- バス・アイソレータに接続されたセグメントは、そのバス・アイソレータ上の他のポートが不正終了したときには必ず非アクティブな状態にしておきます。不正終了したバス・アイソレータ・ポートから他の信号に不正シグナルが転送されるのを防ぐためです。

以上のことから、ホット・プラグ操作では、バス・アイソレータに接続されているセグメントのどれかを(潜在的な)アクティブ・セグメントとして指定し、他は非アクティブにしておきます。図 A-14 を参照してください。セグメントを確実に非アクティブにするための手順については、第 A.7.6.3 項を参照してください。

図 A-14 バス・アイソレータのホット・プラグ



ZK8843AGE

バス・アイソレータには複数のスタブ接続があるので、それぞれにホット・プラグが可能ですが、ホット・プラグ操作では、同時に複数のセグメントをアクティブ・セグメントにすることはできません。

- 電気導線は、必ず相手のコネクタ・ピンと合うか確認してください。コネクタが外されているときだけでなく、切断や接続の最中もよく注意してください。
- デバイスを交換するときは、必ず同じタイプのデバイスと交換してください。OpenVMS Cluster 上のシステムが SCSI ID を DK デバイスまたは MK デバイスとして構成すると、OpenVMS Cluster メンバが実行している限りその SCSI ID には、それぞれ DK デバイスか MK デバイスしか接続できなくなるためです。
同じデバイス・タイプの異なる実装は互換性があります(たとえば、RZ26L は RZ28B と交換できます)。ただし、新しいデバイスのマウントを開始するまで、システムはデバイス・タイプの変更を認識しません。また、ホスト方式のシャドウイングでは、シャドウ・セットのすべてのメンバが同じデバイス・タイプでなければなりません。
- システムがブートしたときに空の SCSI ID は、システムが実行中は空のままにします。この規則は、複数のプロセッサが SCSI バスにあって、MSCP サーバがそのどれかにログオンしているときに適用します(MSCP_LOAD システム・パラメータを 1 に設定しておく、MSCP サーバがロードされます)。

この規則が必要なのは、SCSI バスのノードに、MSCP サービス提供によるパスではなく直接パスでディスクをアクセスさせるためです。新しいデバイスを (SYSMAN IO コマンドで) システム上に構成すると、システムは共用 SCSI バスの二次システムにそのデバイスをサービスします。二次システムは、MSCP サービスによるパスで新しいデバイスを自動的に構成します。この場合、MSCP サービスによるパスから直接 SCSI パスへのフェールオーバーが実装されていないため、二次システムは新しいデバイスに直接 SCSI パスを使用できなくなります。

A.7.6.3 デバイスまたはセグメントを確実に非アクティブにするための手順

デバイスやセグメントを確実に非アクティブにするには、以下のように操作します。

- ディスクを確実に非アクティブにするには、以下のように操作します。
 1. OpenVMS Cluster システムの全メンバ上でディスクのマウントを解除します。
 2. マウントを解除したディスクに発生する可能性のある I/O のマウントが解除されていることを確認します。次に例を示します。
 - ディスクをクォーラム・ディスクとして無効にします。
 - ディスクを (DCL コマンド ALLOCATE で) 占有し、以後のマウントや初期化ができないようにします。
 - 論理 SCSI バス上のすべての一時停止ホストによるコンソール・ポーリングを無効にします (コンソール変数 SCSI_POLL を OFF に設定し、INIT コマンドを入力します)。
 - 電源入力や初期化自己診断テスト、ブート、SCSI バスの構成を行っているホストが論理 SCSI バス上にないか確認します (SYSMAN IO コマンド使用)。
- HSZxx コントローラを確実に非アクティブにするには、以下のように操作します。
 1. OpenVMS Cluster システムの全メンバ上のすべての HSZxx 仮想ディスクについて、マウントを解除します。
 2. 『HS Family of Array Controller User's Guide』の手順に従ってコントローラをシャットダウンします。
 3. HSZxx の電源を切ります (オプション)。
- ホスト・アダプタを確実に非アクティブにするには、以下のように操作します。
 1. システムを一時停止します。
 2. システムの電源を切るか、コンソール変数 SCSI_POLL を OFF に設定し、停止したシステムに INIT コマンドを入力します。これでそのシステムがポーリングしたり、ポーリングに応答するのを防ぐことができます。
- セグメントを確実に非アクティブにするには、セグメント上のすべてのデバイスについての場合と同じ操作をします。

A.7.6.4 StorageWorks SBB ディスクのホット・プラグの手順

SBB (ストレージ構築ブロック) ディスクをアクティブな SCSI バスから取り外すには、以下のように操作します。

1. 接地スタッドに取り付けられている ESD 接地ストラップまたはシステムのどれかのキャビネットの無塗装金属面を使用します。操作手順については、システム・インストール手順を参照してください。
2. 第 A.7.6.3 項の手順に従って、ディスクを非アクティブにします。
3. SBB 横のクリップを押し込み、StorageWorks シェルフからディスクを取り出します。

SBB ディスクをアクティブな SCSI バスに接続するには、以下のように操作します。

1. 接地スタッドに取り付けられている ESD 接地ストラップまたはシステムのどれかのキャビネットの無塗装金属面を使用します。操作手順については、システム・インストール手順を参照してください。
2. そのデバイスに対応する SCSI ID が (ジャンパか、StorageWorks シェルフのドロットで設定) 以下の要件を満たしているか確認します。
 - SCSI ID が論理 SCSI バスに対して一意である。
 - SCSI ID は DK デバイスとして以下のすべてにおいて構成済みである。
 - ID を構成済みの OpenVMS Cluster システムのすべてのメンバ。
 - 同じ SCSI バス上で MSCP サーバを実行中のすべての OpenVMS プロセッサ。
3. SBB を StorageWorks シェルフに挿入します。
4. 必要に応じて SYSMAN IO コマンドでディスクを OpenVMS Cluster メンバに構成します。

A.7.6.5 HSZxxのホット・プラグ手順

HSZxxコントローラをアクティブな SCSI バスから取り外すには、以下のように操作します。

1. 接地スタッドに取り付けられている ESD 接地ストラップまたはシステムのどれかのキャビネットの無塗装金属面を使用します。操作手順については、システム・インストール手順を参照してください。
2. 第 A.7.6.3 項の手順に従って、HSZxxを非アクティブにします。
3. HSZxx電源を切ることができますが、配電系統には接続したままにして接地を維持します。
4. ディファレンシャル・トライコネクタのネジをゆるめて、HSZxxから取り外します。
5. 取り外しの際に、裸のコネクタ・ピンが ESD や他の電気導体に触れないように注意してください。

HSZxxコントローラをアクティブな SCSI バスに接続するには、以下のように操作します。

1. 接地スタッドに取り付けられている ESD 接地ストラップまたはシステムのどれかのキャビネットの無塗装金属面を使用します。操作手順については、システム・インストール手順を参照してください。また、HSZxと、それに取り付けられているすべての構成要素の間の接地オフセット電圧が第 A.7.8 項の指定範囲内にあるか確認してください。
2. 取り外しの際に、裸のコネクタ・ピンが ESD や他の電気導体に触れないように注意してください。
3. HSZxxの電源を入れ、HSZxxに付属しているディスク装置が以下の要件を満たしているか確認します。
 - ディスク装置が論理 SCSI バス上で一意である。
 - ディスク装置は、以下のすべてにおいて必ず DK デバイスとして構成する。
 - ID を構成済みの OpenVMS Cluster システムのすべてのメンバ。
 - 同じ SCSI バス上で MSCP サーバを実行中のすべての OpenVMS プロセッサ。
4. HSZxxアクティブ・セグメントと正しくスタブ接続できるか確認してください。(途中にケーブルを挟まず、トライコネクタを HSZxxコントローラ・モジュールに直結すれば、接続は制約に違反していません。)
5. 位置がずれないように慎重にディファレンシャル・トライコネクタを HSZxxに接続します。ネジを締めます。
6. 必要に応じて、SYSMAN IO コマンドで HSZxx仮想ディスクを OpenVMS Cluster メンバに構成します。

A.7.6.6 ホスト・アダプタのホット・プラグ手順

ホスト・アダプタをアクティブな SCSI バスから取り外すには、以下のように操作します。

1. 接地スタッドに取り付けられている ESD 接地ストラップまたはシステムのどれかのキャビネットの無塗装金属面を使用します。操作手順については、システム・インストール手順を参照してください。
2. これから切り離す接続部がスタブ接続か確認します。スタブ接続でない場合は、ホット・プラグ手順は実行しないでください。
3. 第 A.7.6.3 項の手順に従って、ホスト・アダプタを非アクティブにします。
4. システムの電源は切ってもかまいませんが、接地を保つため配電系統にはプラグを差し込んだままにしておきます。
5. ホスト・アダプタのシングル・エンド・コネクタから“Y”ケーブルを外します。
6. 取り外しの際に、裸のコネクタ・ピンが ESD や他の電気導体に触れないように注意してください。

7. ホストの電源を入れたままホストの内部バスからアダプタをアンプラグしないでください。

これで、アダプタは SCSI バスから切り離されました。アダプタをホストから取り外すには、まずホストの電源を切り、ホストの内部バスからアダプタを取り外します。

ホスト・アダプタをアクティブな SCSI バスにプラグ・インするには、以下のように操作します。

1. 接地スタッドに取り付けられている ESD 接地ストラップまたはシステムのどれかのキャビネットの無塗装金属面を使用します。操作手順については、システム・インストール手順を参照してください。また、ホストと、それに取り付けられているすべての構成要素の間の接地オフセット電圧が第 A.7.8 項の指定範囲内にあるか確認してください。
2. 取り外しの際に、裸のコネクタ・ピンが ESD や他の電気導体に触れないように注意してください。
3. ホスト・アダプタをアクティブ・セグメントと合法的にスタブ接続できるか確認してください(スタブ長は許容範囲内とし、ホスト・アダプタがアクティブ・セグメントの終端にならないこと)。
4. アダプタをホストにプラグ・インします(プラグ・インされていない場合)。
5. システムを配電系統にプラグ・インし、接地が正しく行われているか確認します。必要に応じて電源を入れます。
6. “Y”ケーブルをホスト・アダプタにまっすぐ接続します。

A.7.6.7 DWZZxコントローラのホット・プラグ手順

DWZZxをアクティブな SCSI バスから切り離すには、以下のように操作します。

1. 接地スタッドに取り付けられている ESD 接地ストラップまたはシステムのどれかのキャビネットの無塗装金属面を使用します。操作手順については、システム・インストール手順を参照してください。
2. これから切り離す接続部がスタブ接続か確認します。スタブ接続でない場合、ホット・プラグ手順は実行しないでください。
3. DWZZxxの電源は切らないでください。付属 SCSI バス・セグメントの操作に支障が生ずることがあるからです。
4. 切断後、どの SCSI バス・セグメントをアクティブにしておくかを決めます。第 A.7.6.3 項の手順に従って、他のセグメントを非アクティブにします。

アクティブ・セグメントから DWZZxを取り外すとき、非アクティブ・セグメントからも DWZZxを取り外すまで、あるいは、アクティブ・セグメントから切断された DWZZxポートに適正な終端が復元されるまで、非アクティブ・セグメントは非アクティブのままにしておきます。

5. 次の手順は、DWZZxのタイプとホット・プラグするセグメントによって、以下のように異なります。

DWZZxタイプ	要件	操作
SBB ¹	シングル・エンド・セグメントをアクティブのままにする。	SBB 横のクリップを押し込み、DWZZxを StorageWorks シェルフから取り出します。
SBB ¹	ディファレンシャル・セグメントはアクティブのままとする。	ネジをゆるめて、DWZZxからディファレンシャル・トライコネクタを取り外します。
テーブル・トップ	シングル・エンド・セグメントはアクティブのままにする。	DWZZxのシングル・エンド・コネクタから“Y”ケーブルを外します。
テーブル・トップ	ディファレンシャル・セグメントはアクティブのままとする。	ネジをゆるめて、DWZZxからディファレンシャル・トライコネクタを取り外します。

¹StorageWorks で SBB とは、ストレージ構築ブロックの略称です。

6. 取り外しの際に、裸のコネクタ・ピンが ESD や他の電気導体に触れないように注意してください。

DWZZxをアクティブな SCSI バスにプラグ・インするには、以下のように操作します。

1. 接地スタッドに取り付けられている ESD 接地ストラップまたはシステムのどれかのキャビネットの無塗装金属面を使用します。操作手順については、システム・インストール手順を参照してください。また、DWZZxと、それに取り付けられているすべての構成要素の間の接地オフセット電圧が第 A.7.8 項の指定範囲内にあるか確認してください。
2. 取り外しの際に、裸のコネクタ・ピンが ESD や他の電気導体に触れないように注意してください。
3. DWZZxをアクティブ・セグメントと正しくスタブ接続できるか確認してください (スタブ長は許容範囲内とし、DWZZxがアクティブ・セグメントの終端にならないこと)。
4. DWZZxの電源は入れたままにしておきます。追加する SCSI セグメントを接続して終端を接続します。このセグメント上のすべてのデバイスは非アクティブにしておきます。
5. 次の手順は、DWZZxのタイプと、ホット・プラグするセグメントによって、以下のように異なります。

DWZZxタイプ	要件	操作
SBB ¹	シングル・エンド・セグメントをホット・プラグする。	DWZZxを StorageWorks シェルフに挿入します。

¹StorageWorks で SBB とは、ストレージ構築ブロックの略称です。

DWZZxタイプ	要件	操作
SBB ¹	ディファレンシャル・セグメントをホット・プラグする。	ずれないように慎重にディファレンシャル・トライコネクタを DWZZx に接続します。ネジを締めます。
テーブル・トップ	シングル・エンド・セグメントをホット・プラグする。	ずれないように慎重に“Y”ケーブルを DWZZx に接続します。
テーブル・トップ	ディファレンシャル・セグメントをホット・プラグする。	ずれないように慎重にディファレンシャル・トライコネクタを DWZZx に接続します。ネジを締めます。

¹StorageWorks で SBB とは、ストレージ構築ブロックの略称です。

- 新しい取り付けられたセグメントにストレージ・デバイスがある場合、SYSMAN IO コマンドで必要に応じて OpenVMS Cluster メンバ上に構成します。

A.7.7 マルチホスト SCSI OpenVMS Cluster システムで使用されるデバイスの OpenVMS 要件

表 A-2 にあるのは、現在、マルチホスト SCSI OpenVMS Cluster システムでの使用が認定されているデバイスです。特に使用が認定されていないディスク・デバイスでも、以下の要件を満たしていればマルチホスト OpenVMS Cluster システムで使用できるものがあります。

- 同時マルチ・イニシエータ I/O をサポートする。
- イニシエータ・ベースで以下の状態または要件を適切に管理すること。
 - 同期とネゴシエーションによる状態と通信速度
 - ネゴシエーションによる幅の状態
 - 緊急時の信頼性とユニット・アテンションの要件
- タグ付きコマンド・キューイング。これは、I/O がフラッシュされたかどうかを確認するときに OpenVMS Cluster システムが使用する整列保証動作をするのに必要です。ドライブは、SCSI-2 標準の Section 7.8.2 に準拠したキューイングを実装していなければなりません。以下にその一部を示します。

“...All commands received with a simple queue tag message prior to a command received with an ordered queue tag message, regardless of initiator, shall be executed before that command with the ordered queue tag message.” (強調表記は後から追加したもの。)

[[訳]] “... 整列キュー・タグ・メッセージ付きのコマンドより前に、単純なキュー・タグ・メッセージ付きで受信したすべてのコマンドは、イニシエータとは無関係に、整列キュー・タグ・メッセージ付きのコマンドより前に実行しなければなりません。”

- コマンド切断のサポート。

- SCSI-2 標準の Section 6.1.4.2 の Option b 準拠の再選択タイムアウト手順。また、デバイスは非応答イニシエータの再選択の所要時間を制限する再選択再試行アルゴリズムが実装されたデバイスであること。
- 複数のホストによって同じブロックが不必要に再ベクトル化されるのを防ぐための ARRE (automatic read reallocation enabled) と AWRE (automatic write reallocation enabled) (つまりドライブ方式の不良ブロックの再ベクトル化)。データが壊れるのを防ぐには、ドライブが SCSI-2 標準の Section 9.3.3.6 に準拠していなければなりません。以下にその一部を示します。

“...The automatic reallocation shall then be performed only if the target successfully recovers the data.” (強調表記は後から追加したもの。)

[[訳]] “... 自動再割り当ては、ターゲットがデータの回復に成功した場合にだけ実行されなければなりません。”

- ストレージ・デバイスは TERMPWR を提供しないこと。提供する場合は、1つのセグメント上の TERMPWR のソースが5つ以上にならないことを保証する構成規則の適用が必要です。

最後は、デバイスや同じセグメント上の他のデバイスをホット・プラグする場合、そのデバイスは第 A.7.6.2 項の電氣的要件を満たさなければなりません。

A.7.8 接地要件

この項では、SCSI OpenVMS Cluster システムにおける電装システムの接地要件について説明します。

接地方法が正しくないと、構成要素のエンクロージャ間に接地オフセット電圧という電圧差が生じます。SCSI インターコネクタ上では、たとえ小さな接地オフセット電圧であっても(表 A-8 の手順 3 参照) 構成に混乱を生じ、システム・パフォーマンスの低下やデータの破壊を招きます。

表 A-8 は、正しい接地方法に必要な確認事項をまとめたものです。

表 A-8 正しい接地のための確認手順

手順	説明
1	サイトの配電が地域の電気コードを満たしているか確認します。
2	サイト全体の配電システムが以下の要件を満たしているか認めます。 <ul style="list-style-type: none">• コンセントにはすべて接地接続がある。• コネクタの電源ケーブルには、すべて接地ピンがある。• 電源コンセントの中立接続が実際の接地接続ではない。• 電源コンセントの接地は、すべて同じ配電パネルに接続されている。• 同じサーキット・ブレーカにコンピュータ機器として接続されているデバイスは、すべて UL®または IEC 認定である。
3	これらの要件の検査が困難な場合は、ハンドヘルド・マルチメータで、任意の 2 つのキャビネット間の接地オフセット電圧を測定します。電圧を測定するには、マルチメータのリード線をエンクロージャの無塗装金属面に接続します。次に、電圧が以下の許容接地オフセット限度を超えていないか確認します。 <ul style="list-style-type: none">• シングル・エンド・シグナル通知: 50 ミリボルト (許容最大オフセット)• ディファレンシャル・シグナル通知: 800 ミリボルト (許容最大オフセット) マルチメータ方式では、測定の瞬間だけにデータが得られます。接地オフセットの値は、デバイスが追加されてアクティブになったり、同じ電源に接続されると、時間経過とともに変化することがあります。時間が経っても接地オフセットを許容範囲内に収めるには、有資格の電気技術者による電源の調査をお勧めします。
4	接地状況や、測定したオフセット値が許容範囲を超えているかどうか分からない場合は、有資格の電気技術者によって問題を解決することをお勧めします。オフセットの測定値を削減するには、接地ケーブルをエンクロージャ間に挿入する必要があります。
5	測定したオフセット電圧が許容範囲を超えていたために接地ケーブルも接続したら、再度電圧を測定して許容範囲内に収まっているかどうかを確認してください。許容範囲を超えている場合は、電気技術者の手で接地オフセット電圧の元を探し出し、その電圧を削減するか取り除く必要があります。

MEMORY CHANNEL 技術概要

この付録では、高度なパフォーマンスを実現するクラスタ・インターコネクト・テクノロジーとして MEMORY CHANNEL を説明します。MEMORY CHANNEL は OpenVMS Alpha バージョン 7.1 で導入されたテクノロジーであり、複数の構成をサポートします。

この付録の構成は、以下のとおりです。

項目	内容
製品概要	MEMORY CHANNEL 製品とその長所、ハードウェアの構成要素、構成に関する専門的な概要説明です。
技術概要	MEMORY CHANNEL のはたらきについて、技術的な内容を詳細に取り上げます。

B.1 製品概要

MEMORY CHANNEL は、PCI ベースの Alpha システム向けの高度なパフォーマンスを実現するクラスタ・インターコネクト・テクノロジーです。待ち時間の短縮、広い帯域幅、直接メモリ・アクセスという利点があるため、MEMORY CHANNEL は、OpenVMS Cluster の独特の機能を補い、また、拡張して、1 つの仮想システムとして機能します。

MEMORY CHANNEL は、CI、DSSI、FDDI、Ethernet といった既存のインターコネクトにおけるノード間のクラスタ・トラフィック（ロック管理通信など）の負荷を軽減し、ストレージやネットワークのトラフィック処理の実効性を強化します。MEMORY CHANNEL は、スループットを大幅に増強する一方で、従来の I/O 処理の待ち時間を大幅に短縮します。

MEMORY CHANNEL は、ノード間で大量のデータを移動するアプリケーションに活かすことができます。リアルタイム処理やトランザクション処理など、高速データ通信が必要なアプリケーションのソリューションとして最適です。さらに、MEMORY CHANNEL により、高度なパフォーマンス・データベースや、大量の OpenVMS Lock Manager トラフィックが生成されるアプリケーションでも、スループットを改善できます。

B.1.1 MEMORY CHANNEL の特長

MEMORY CHANNEL テクノロジーには、以下のような特長があります。

- コスト・パフォーマンスが高い

MEMORY CHANNEL は、CI 帯域幅の数倍の 100 MB/s インターコネク트를、最小限の遅延時間とともに提供します。MEMORY CHANNEL アーキテクチャは、業界標準の PCI バスを前提に設計されています。

- 既存のアプリケーションを変更する必要がない

MEMORY CHANNEL は、既存のクラスタ・ソフトウェアとシームレスに運用できるので、既存のソフトウェアを変更する必要がありません。PMDRIVER と MCDRIVER の 2 つの新しい MEMORY CHANNEL ドライバは、既存のポート・ドライバと同じ方式で、OpenVMS Cluster の Systems Communication Services レイヤと統合します。クラスタ・ソフトウェアの上位層にはまったく影響ありません。

- CI、DSSI、SCSI クラスタの LAN の負荷を軽減

ストレージを MEMORY CHANNEL に直結させることはできません。

MEMORY CHANNEL は、CI や DSSI を代替するものではありませんが、これらのインターコネクと併用すれば、そのノード間トラフィックを代替して実行できます。これにより、CI や DSSI をストレージ・トラフィック専用でき、クラスタ全体の通信の最適化が可能です。

SCSI インターコネクと LAN インターコネクを持つクラスタで使用すると、LAN のノード間通信を MEMORY CHANNEL が代替して実行し、LAN で処理できる TCP/IP トラフィックや DECnet トラフィックを増やすことができます。

- 障害隔離動作を提供

システム障害が発生しても、MEMORY CHANNEL ノードは OpenVMS Cluster のノードと同様に対応します。障害が発生したノードは、再びクラスタに復帰するまで、クラスタの他の要素によって代替されます。

B.1.2 MEMORY CHANNEL バージョン 2.0 の特長

OpenVMS バージョン 7.1 で導入した当初、MEMORY CHANNEL は 10 フィートのラジアル・トポロジにより、最大で 4 ノードをサポートしていました。また、1 組の送受信側単位の通信しかできませんでした。MEMORY CHANNEL バージョン 1.5 では、8 ノード、新アダプタ (CCMAA-BA)、全メッセージのタイム・スタンプ、堅牢なパフォーマンスをサポートしています。

MEMORY CHANNEL バージョン 2.0 では、以下の機能を提供します。

- 新しいアダプタ (CCMAB-AA) と新しいハブ (CCMHB-AA と CCMHB-BA) のサポート

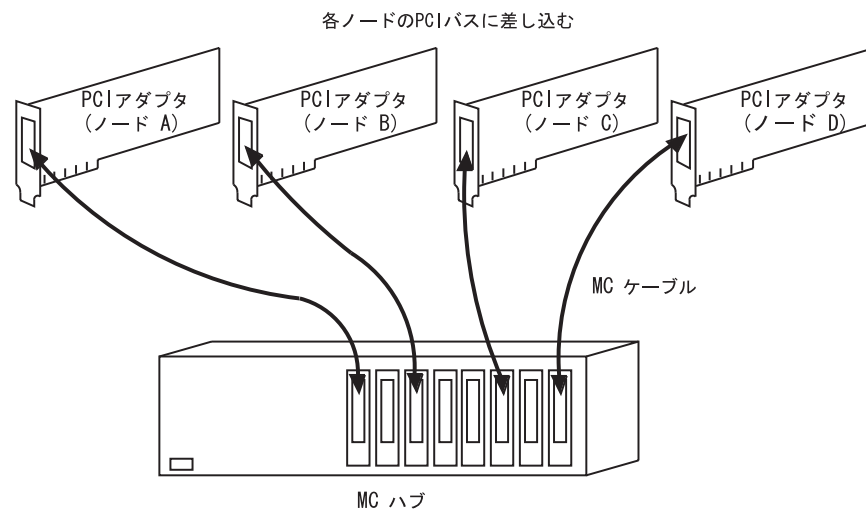
- 4組の送受信ペア間で同時通信をサポート
- 最長3 kmのラジアル・トポロジに対応できる長いケーブル長のサポート

B.1.3 ハードウェアの構成要素

MEMORY CHANNEL クラスタはハブで相互に結合します。ハブは、システムを結合するデスクトップ PC の大きさのユニットです。ハブとシステムの PCI アダプタは、リンク・ケーブルで接続します。図 B-1 は、MEMORY CHANNEL のサポート用にノードに必要な3つのハードウェア構成要素です。

- PCI 対 MEMORY CHANNEL アダプタ
- リンク・ケーブル
- MEMORY CHANNEL ハブ内のポート (2つの PCI アダプタだけをケーブルで接続した2ノード構成を除く)

図 B-1 MEMORY CHANNEL ハードウェアの構成要素

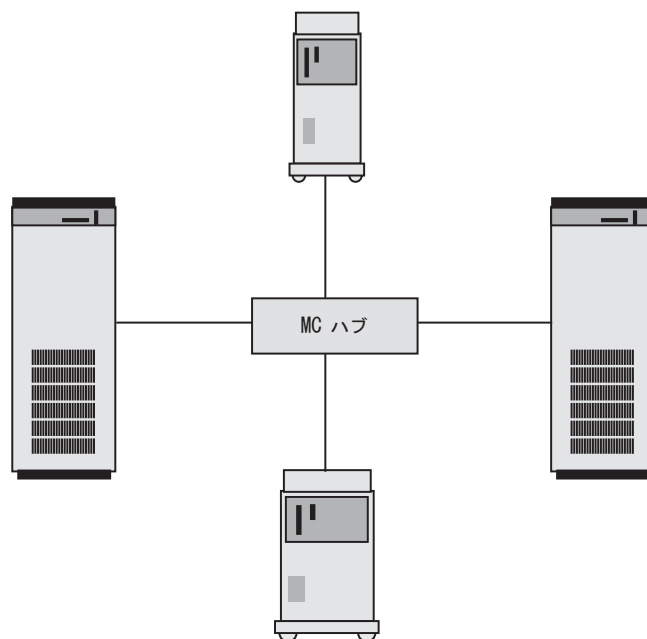


VM-0220A-A1

図 B-1 に示す PCI アダプタのメモリ・マッピング・ロジックでは、MEMORY CHANNEL クラスタ内のシステム同士の通信が可能です。

図 B-2 は、ハブが中央にある4ノード MEMORY CHANNEL クラスタです。

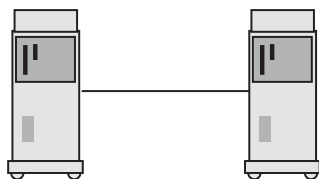
図 B-2 4 ノード MEMORY CHANNEL クラスタ



ZK8582AGE

2 ノードしかないクラスタで MEMORY CHANNEL ハブは不要です。図 B-3 に示すような 2 ノード構成では、同じアダプタとケーブルを使用し、PCI アダプタの 1 つが仮想ハブとして機能します。アダプタとケーブルは、構成を拡張してもそのまま使用できます。

図 B-3 仮想ハブ MEMORY CHANNEL クラスタ



ZK8583AGE

B.1.4 高可用構成のためのバックアップ・インターコネクト

3 ノード以上の構成では、中央ハブが MEMORY CHANNEL に必要です。MEMORY CHANNEL ハブには、電源を供給してアクティブな電子構成要素を組み込みます。電源のシャットダウンや構成要素の障害でハブに障害が発生すると、MEMORY CHANNEL インターコネクトの動作は停止します。この種の障害は、

CI, DSSI, ほとんどの LAN 構成など, この方式以外のクラスタ・インターコネク
トでは発生しません。

MEMORY CHANNEL 構成を使用し, また高可用も求める場合は, 以下の構成で二
次バックアップ・インターコネクトを採用することをお勧めします。

- 一般には, 二次インターコネクトは, LAN (Ethernet または FDDI) でクラ
スタを実装すれば簡単に構成できます。FDDI と 100 Mb/s Ethernet を使用す
れば, MEMORY CHANNEL に障害が発生しても十分なパフォーマンスをインタ
ーコネクトに確保できます。(LAN によるクラスタの実装の詳細については,
『OpenVMS Cluster システム』および『OpenVMS Cluster 構成ガイド』を参照
してください。)
- CI インターコネクトと DSSI インターコネクトは, MEMORY CHANNEL のバ
ックアップとして自動的に機能します。
- MEMORY CHANNEL インターコネクトが 2 本ある構成では, 高いパフォーマン
スを実現でき, また MEMORY CHANNEL インターコネクトの 1 本に障害が発
生しても処理を続行できます。

B.1.5 ソフトウェアの要件

MEMORY CHANNEL を使用する場合, メモリや診断ツールは一定の要件に従って
選択してください。

B.1.5.1 メモリの要件

MEMORY CHANNEL は, 通常の操作でもメモリを使用します。MEMORY
CHANNEL クラスタの各システムには, メモリが少なくとも 128 MB 必要で
す。

B.1.5.2 大容量メモリ・システムにおける NPAGEVIR パラメータの使用

大きな非ページ化プール・メモリを組み込んだシステムでは, MEMORY
CHANNEL が初期化を終了できないことがあります。その場合, コンソールに
は以下のメッセージが繰り返し表示されます。

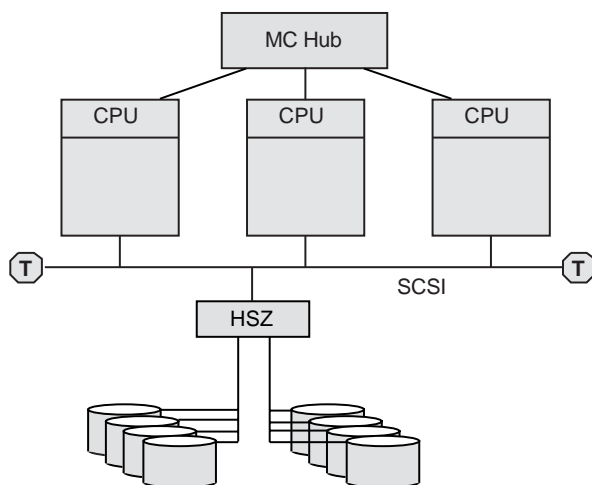
```
Hub timeout - reinitializing adapter
```

この場合は, SYSGEN パラメータ NPAGEVIR の値を調べます。値が 1 GB を超
える場合, 約半分まで値を下げてください。その後システムをリブートすると,
MEMORY CHANNEL による初期化が終了します。

B.1.6 構成

図 B-4 は, ストレージに SCSI インターコネクトを使用する MEMORY CHANNEL
の基本クラスタです。この構成には, MEMORY CHANNEL インターコネクトの高
いパフォーマンス, そして SCSI インターコネクトの経済性という 2 つの利点があり
ます。

図 B-4 MEMORY CHANNEL ベース・クラスタと SCSI ベース・クラスタ

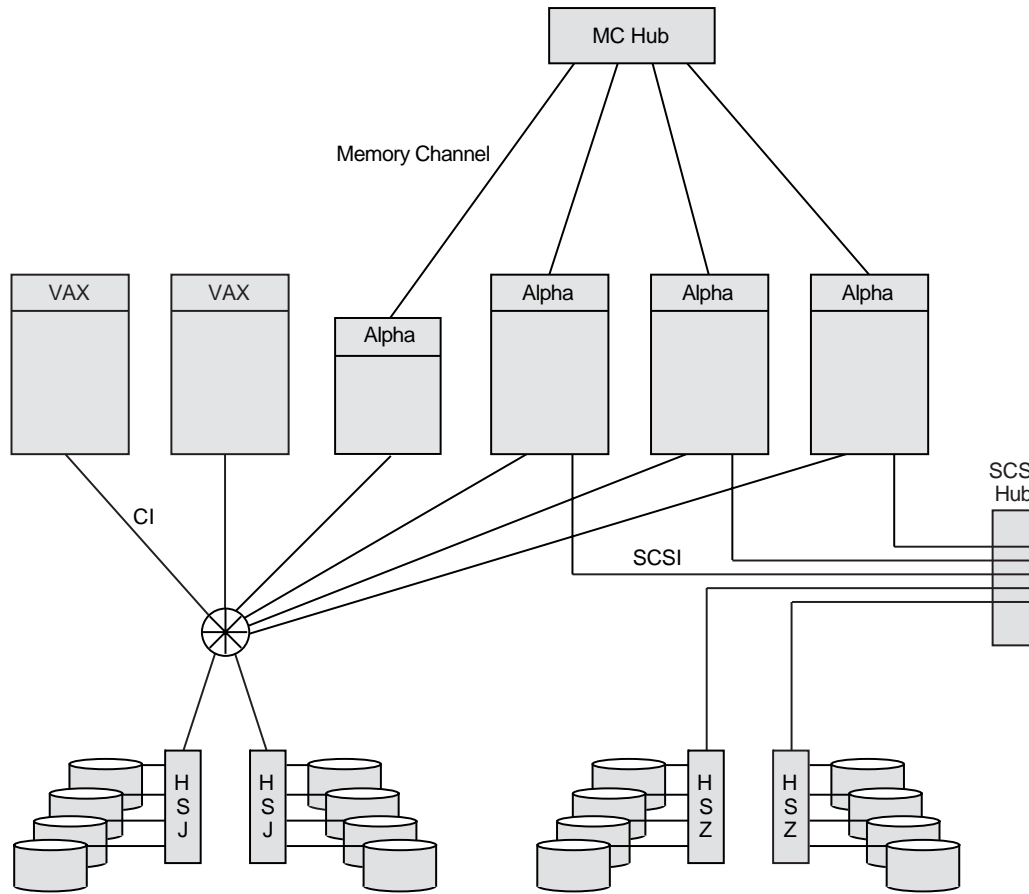


ZK-8777A-GE

図 B-4 に示す構成では、MEMORY CHANNEL インターコネクトがノード間通信を処理し、SCSI バスがストレージ通信を処理します。

MEMORY CHANNEL を現在のシステムに追加して統合することもできます。図 B-5 は、CI ベース・クラスタと SCSI ベース・クラスタの混成アーキテクチャに MEMORY CHANNEL を追加した例です。この例で、BI 方式の VAX システムと XMI 方式の VAX システムは、同じ CI クラスタで PCI 方式の Alpha MEMORY CHANNEL システムに統合されています。

図 B-5 MEMORY CHANNEL の CI 方式と SCSI 方式のクラスタ

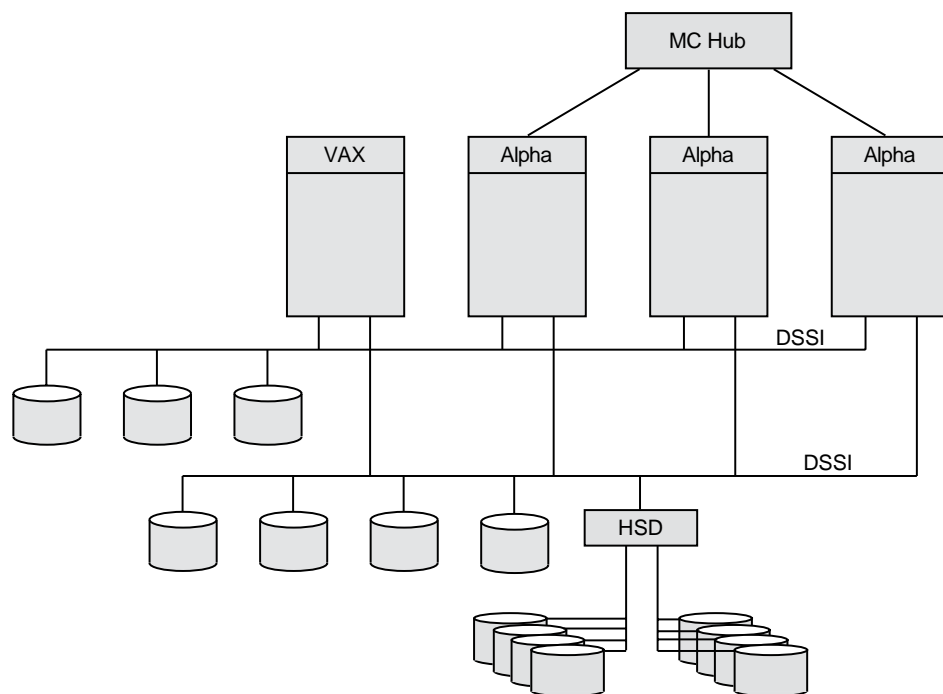


ZK-8756A-GE

MEMORY CHANNEL インターコネクトは、ストレージやブートには使用していないので、ブート・デバイスには他のインターコネクトでアクセスします。たとえば図 B-5 の場合、全ノードが CI でブート・デバイスに直結されているので、ブート・デバイス用には、CI 方式のディスクのどれかを選択するとよいでしょう。

図 B-6 にあるように、MEMORY CHANNEL は既存の DSSI クラスタに統合することもできます。

図 B-6 MEMORY CHANNEL DSSI 方式のクラスタ



ZK-8774A-GE

図 B-6 にある，3つの MEMORY CHANNEL システムと VAX システムからは，HSD コントローラに接続された SCSI ストレージの他，DSSI インターコネクタに直結されたストレージをアクセスできます。この構成で，MEMORY CHANNEL は Alpha ノード間通信を処理し，DSSI はストレージ間通信を処理します。

B.1.6.1 構成サポート

MEMORY CHANNEL は，表 B-1 に示すプラットフォームと構成をサポートしています。

表 B-1 MEMORY CHANNEL 構成のサポート

要件	説明
構成	<p>MEMORY CHANNEL は、以下の構成をサポートしています。</p> <ul style="list-style-type: none">• MEMORY CHANNEL ハブごとに最高で 8 ノード。• 2 ハブ構成の場合、ノード当たり 2 PCI アダプタ。それぞれ異なるハブに接続のこと。• 2 ノード構成の場合、ハブは不要。
ケーブル	<p>MEMORY CHANNEL は、以下のケーブルをサポートしています。</p> <ul style="list-style-type: none">• 最長 10-m (32.8 ft) ラジアル・トポロジの銅ケーブル。• 最長 30-m (98.4 ft) ラジアル・トポロジの HP 製の光ファイバ・ケーブル。最長 3-km (1.8 マイル) ラジアル・トポロジの他のベンダの光ファイバ・ケーブル。
ホスト・システム	<p>MEMORY CHANNEL は、以下のシステムをサポートしています。</p> <ul style="list-style-type: none">• AlphaServer 8400• AlphaServer 8200• AlphaServer 4100• AlphaServer 2100A• AlphaServer 1200• AlphaServer 800

注意

MEMORY CHANNEL バージョン 1.5 ハブと MEMORY CHANNEL バージョン 2.0 ハブのどちらを組み込んだ OpenVMS Cluster システムでもコンピュータを構成できますが、アダプタとケーブルのバージョン番号をハブのバージョン番号に合わせないと MEMORY CHANNEL が正しく機能しません。

つまり、MEMORY CHANNEL バージョン 1.5 アダプタは MEMORY CHANNEL バージョン 1.5 ハブと MEMORY CHANNEL バージョン 1.5 ケーブルに、また、MEMORY CHANNEL バージョン 2.0 アダプタは、MEMORY CHANNEL バージョン 2.0 ハブと MEMORY CHANNEL バージョン 2.0 ケーブルに使用してください。

B.2 技術概要

この節では、MEMORY CHANNEL の働きについて、その技術的な内容を詳しく取り上げます。

B.2.1 従来のネットワークや SMP との比較

MEMORY CHANNEL は、遠距離に分散するシステムを、最高 8 組まで相互接続できる“SMP バスの延長”型と言えます。ただし、複数の CPU が同じ物理メモリに直接アクセスできる SMP 環境と MEMORY CHANNEL は同じものではありません。つまり、ノード同士が MEMORY CHANNEL グローバル・アドレス領域を共用している場合でも、MEMORY CHANNEL では、ノードごとに専用の物理メモリが必要です。

SMP システムと従来のパケット方式のネットワークでは、両者の間にコスト・パフォーマンスの面でかなりの開きがありますが、MEMORY CHANNEL はちょうどその中間的な存在です。表 B-2 は、SMP、MEMORY CHANNEL、通常のネットワークのそれぞれの特性を比較した表です。

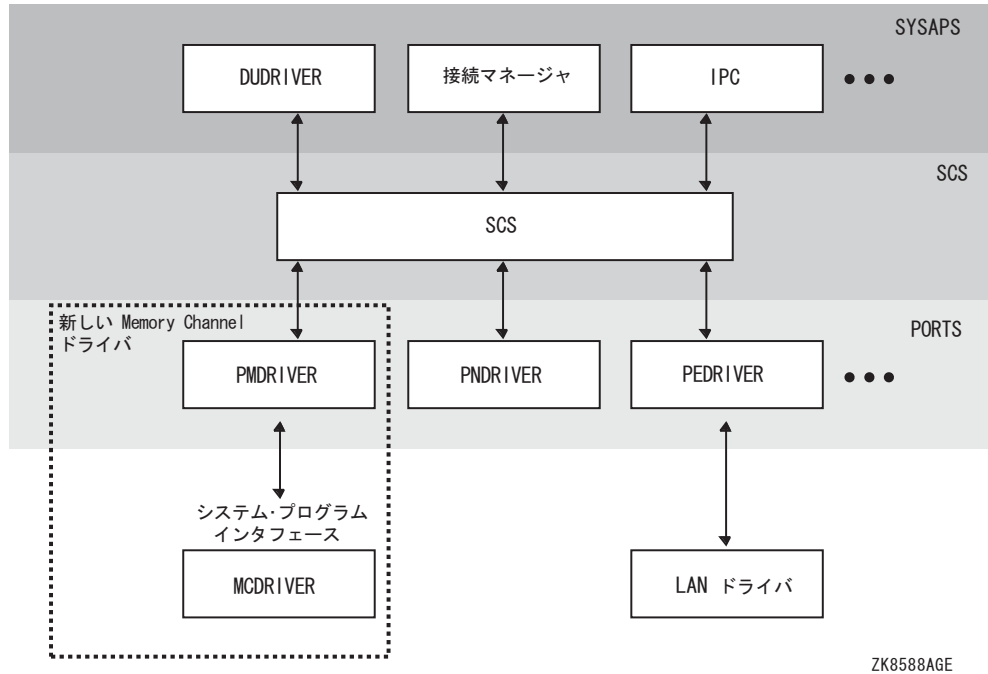
表 B-2 SMP, MEMORY CHANNEL, 通常のネットワークの比較

特性	SMP	MEMORY CHANNEL	通常のネットワーキング
帯域幅 (MB/s)	1000+	100+	10+
待ち時間 (ms/最も単純なメッセージ)	0.5	5 未満	約 300
オーバーヘッド (ms/最も単純なメッセージ)	0.5	5 未満	約 250
ハードウェア通信モデル	共用メモリ	メモリ・マップ	メッセージ伝達
ハードウェア通信プリミティブ	メモリに保存	メモリに保存	ネットワーク・パケット
ブロードキャストのハードウェア・サポート	n/a	可	ときどき
同期のハードウェア・サポート	可	可	不可
ノード・ホット・スワップのハードウェア・サポート	不可	可	可
ソフトウェア通信モデル	共用メモリ	ファースト・メッセージ, 共用メモリ	メッセージ
エラー通信モデル	回復不能	回復可能	回復可能
直接ユーザ・モード通信をサポート	可	可	不可
典型的な物理インターコネク ト・テクノロジー	バックプレーン・エッチ	パラレル銅ケーブル	シリアル光ファイバ
物理インターコネク トのエラー率	非常に少ない 頻度: 1 未満/年	非常に少ない 頻度: 1 未満/年	少ない: 1 日に数回
ハードウェア・インターコ ネク ト方式	特殊目的のコネ クタとロジック	標準 I/O バス・ アダプタ (PCI)	標準 I/O バス・アダプタ (PCI その他)
ノード間距離 (m)	0.3	ハブ構成で 20 (銅) または 60 (光ファイバ), 2 ノード構成で 10 (銅) または 30 (光ファイバ)	50 ~ 1000
ノード数	1	8	数百
プロセッサ数	6 ~ 12	SMP システムで は CPU の最大数 の 8 倍	数千
障害モデル	全体障害	部分障害	部分障害

B.2.2 OpenVMS Cluster アーキテクチャの MEMORY CHANNEL

図 B-7 にあるように、MEMORY CHANNEL の機能は、System Communication Services レイヤの直下に OpenVMS Cluster アーキテクチャで実装されています。このデザインでは、OpenVMS Cluster ソフトウェアの上位レイヤが変更されないの
で、既存のアプリケーションを変更する必要はありません。

図 B-7 OpenVMS Cluster アーキテクチャと MEMORY CHANNEL



MEMORY CHANNEL ソフトウェアは、2つの新しいドライバからなります。

ドライバ	説明
PMDRIVER	クラスタ・ポート・ドライバをエミュレート。
MCDRIVER	MEMORY CHANNEL ハードウェアに MEMORY CHANNEL サービスとインタフェースを提供。

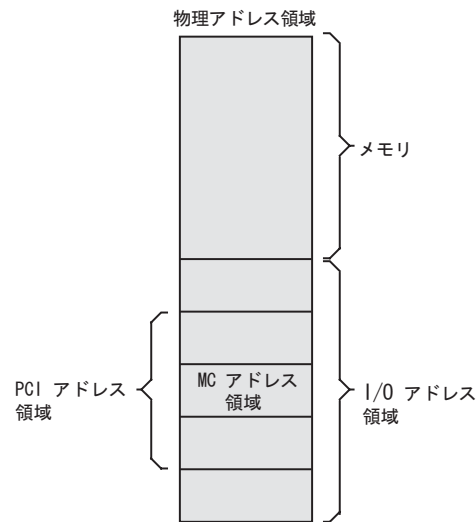
B.2.3 MEMORY CHANNEL アドレス指定

MEMORY CHANNEL 構成では、全ノードがシステムの物理アドレス領域の 1 セクションを共有します。あるシステムがこのアドレス領域にデータを書き込むと、MEMORY CHANNEL ハードウェアもグローバル書き込みを実行し、これによって他のシステムのメモリにもデータが保存されます。つまり、MEMORY CHANNEL アダプタが占有する PCI アドレス領域にノード CPU がデータを書き込むと、データは MEMORY CHANNEL インターコネクトを通じて他のノードにも送信されます。

他のノードの PCI アダプタは、このデータをそれぞれのメモリにマップします。このインフラストラクチャでは、あるシステム上の I/O アドレスに書き込むと、他のシステムの物理アドレスにも書き込むことができます。以下の 2 つの図は、これをさらに詳しく説明したものです。

図 B-8 は、物理メモリにおける MEMORY CHANNEL グローバル・アドレス領域のアドレス指定です。

図 B-8 物理メモリと I/O アドレス領域



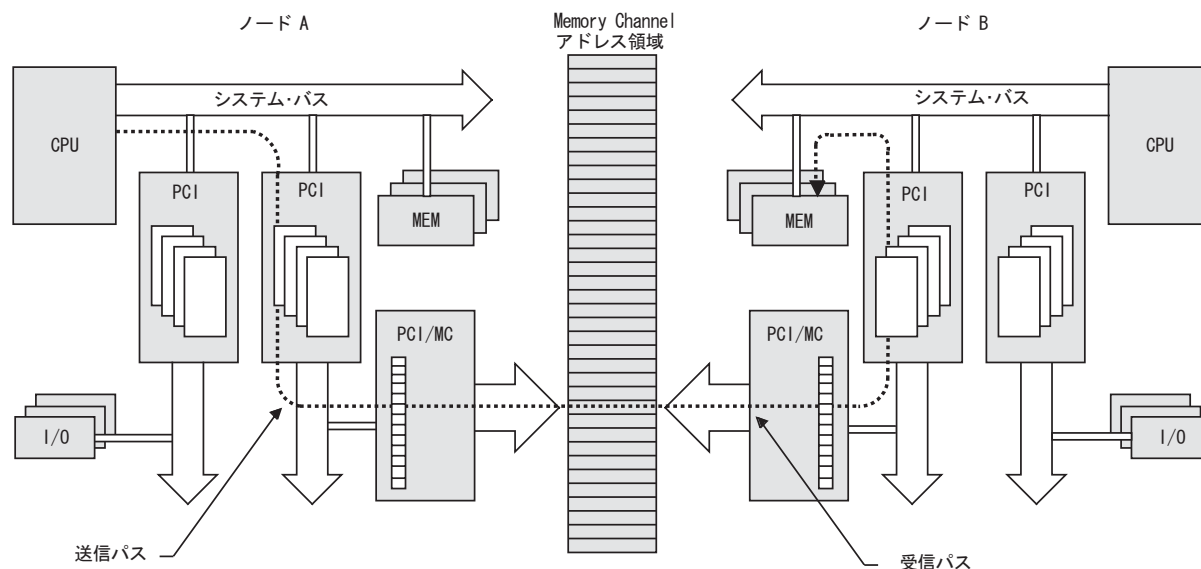
ZK8778AGE

図 B-8 は、システムの典型的なアドレス領域の例であり、物理メモリと I/O アドレス領域に分かれています。PCI I/O アドレス領域で、MEMORY CHANNEL は、128 MB から 512 MB のアドレス領域を消費します。そのため、MEMORY CHANNEL PCI アダプタをこの領域内でアドレス指定でき、この範囲に CPU はデータを書き込むことができます。

MEMORY CHANNEL クラスタの全システムは、このアドレス領域を MEMORY CHANNEL データと通信に割り当てています。このアドレス領域を利用して、CPU は他のノードのメモリにグローバル書き込みを実行します。

グローバル書き込みの例として、図 B-9 では、ノード A とノード B による 2 ノード内部バス・アーキテクチャを示します。

図 B-9 MEMORY CHANNEL バス・アーキテクチャ



ZK8776AGE

図 B-9 に示す図では、以下の順序でノード A がノード B のメモリにグローバル書き込みを実行します。

1. ノード A の CPU は、MEMORY CHANNEL アドレス領域に書き込みを実行します。MEMORY CHANNEL アドレス領域は、PCI アドレス領域の一部です。書き込みは、PCI バスから PCI/MEMORY CHANNEL アダプタに送られ、MEMORY CHANNEL インターコネクトから出力します。
2. ノード B の PCI アダプタはデータを受信します。受信したデータは PCI バスによって取り上げられ、メモリに DMA マップされます。

同様にクラスタ内の全ノードを MEMORY CHANNEL グローバル・アドレス領域にアドレス指定すると、同じアドレス領域と同じデータを仮想的に“共用”できます。このため、MEMORY CHANNEL アドレス領域は図 B-9 で中央アドレス領域として表現されます。

MEMORY CHANNEL グローバル・アドレス領域は、8 KB (8,192 バイト) のページに分割されています。これらを MC ページといいます。この 8 KB ページは、システム間で同じようにマップできます。

MEMORY CHANNEL グローバル・アドレス領域の“共用”機能は、PCI アダプタに、ページ制御テーブルか PCT でセットアップします。PCT には、MC ページごとに設定できる属性があります。これらの属性については、表 B-3 を参照してください。

表 B-3 MEMORY CHANNEL ページ属性

属性	説明
ブロードキャスト	すべてのシステムにデータが送信されます。ノード ID を指定すると、そのシステムにだけデータが送信されません。
ループバック	クラスタ内の他のノードに送信されたデータは、送信ノードの PCI アダプタによってメモリにも書き込まれます。これにより、メッセージの順序が維持されるとともに、すぐれたエラー検出機能を確保できます。
割り込み	現在の MC ページに場所を書き込むと、CPU に割り込みを生成します。これは、他のノードに対する通知に使用できます。
エラー後の送受信機能の停止	現在のページにエラーが発生すると、エラー条件が解消されるまで送受信処理ができなくなることを指定します。
ACK	ページに書き込むと、各受信システムのアダプタが ACK (肯定応答) で応答し、それぞれのホストを割り込むことなくリモート・ノードに書き込み (または他の処理) が発生します。これはエラー・チェックとエラー回復に使用します。

B.2.4 MEMORY CHANNEL の実装

MEMORY CHANNEL ソフトウェアは、OpenVMS Cluster ソフトウェアにバンドルされています。ハードウェアをセットアップした場合は、CLUSTER_CONFIG.COM プロシージャのプロンプトに従って、MEMORY CHANNEL ソフトウェアを構成します。ローカル・コンピュータのノード間通信で MEMORY CHANNEL を有効にするかどうかを確認するプロンプトが表示されたら、“Yes”と応答します。MEMORY CHANNEL が有効かどうかを示すシステム・パラメータの MC_SERVICES_P2 が 1 に設定されます。この設定でドライバ PMDRIVER がロードされ、他の MEMORY CHANNEL システム・パラメータのデフォルト値が有効になります。

MEMORY CHANNEL システム・パラメータの意味については、『OpenVMS Cluster システム』マニュアルを参照してください。

MEMORY CHANNEL ハブ、リンク・ケーブル、PCI アダプタのセットアップの詳細については、『MEMORY CHANNEL User's Guide』(EK-PCIMC-UG.A01)を参照してください。

マルチサイト OpenVMS Cluster

ここでは、マルチサイト OpenVMS Cluster 構成について説明します。ここで説明するマルチサイト OpenVMS Cluster 構成では、約 40 km から 200 km という遠距離を隔てて分散したサイトに、複数のノードが配置されています。この構成方法は、OpenVMS バージョン 6.2 から導入されたものです。一般的な構成方法とともに、複数のサイトを接続するための 3 つのテクノロジーについて説明します。マルチサイト・クラスタの利点は、他の参考資料からの引用とともに参照先も紹介します。

この付録の内容は、『Multiple-Site VMScluster Systems』追補マニュアルの内容を更新したものです。

C.1 マルチサイト OpenVMS Cluster システムとは？

マルチサイト OpenVMS Cluster システムは、メンバ・ノードが物理的に分散して配置された OpenVMS Cluster システムです。使用テクノロジーにより、接続距離が約 800 km を超える場合があります。

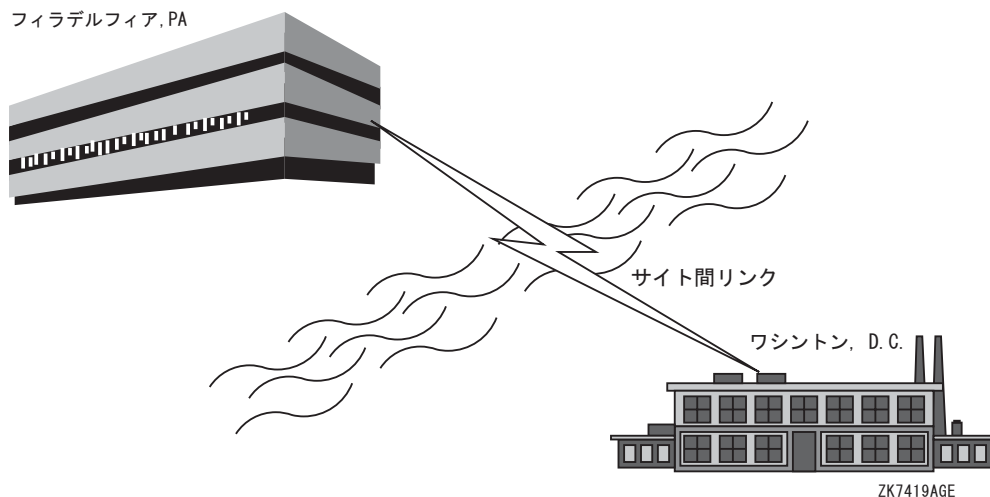
遠距離を隔てて拠点が分散した組織では、マルチサイト OpenVMS Cluster システムを利用すれば、OpenVMS Cluster システムの利点を活かすことができます (たとえば、データをサイト間で共用し、データ・センター・オペレーションは、1 個所に一元化するなど)。

図 C-1 は、製造拠点をワシントン D.C. に持ち、本社がフィラデルフィアにあるマルチサイト OpenVMS Cluster システムの概念を表しています。この構成におけるサイト間の物理的な距離は約 210 km あります。

マルチサイト OpenVMS Cluster

C.1 マルチサイト OpenVMS Cluster システムとは?

図 C-1 フィラデルフィアとワシントン間のサイト間リンク



C.1.1 ATM, DS3, FDDI, [D]WDM サイト間リンク

以下のサイト間リンクのテクノロジーは、OpenVMS VAX システムと OpenVMS Alpha システムに認定されています。

- 非同期転送モード (ATM)
- DS3
- FDDI
- [D]WDM

高度なパフォーマンスを実現するローカル・エリア・ネットワーク (LAN) テクノロジーを ATM, DS3, FDDI, [D]WDM の各インターコネクトと組み合わせれば、ワイド・エリア・ネットワーク (WAN) 通信サービスを OpenVMS Cluster 構成で利用できます。GIGAswitch クロスパー・スイッチと ATM, DS3, または FDDI インターコネクトで構成した OpenVMS Cluster システムでは、ノード同士が何キロメートルも離れていても使用できます。(2 箇所のサイト間の実距離は、サイト間の直線距離ではなく、サイト間のケーブル経路の長さです。) 第 C.4 節では、OpenVMS Cluster システムと WAN 通信サービスをさらに詳しく説明します。

注意

マルチサイト OpenVMS Cluster でディザスタ・トレランスを利用するには、HP 製のシステム管理、および、ソフトウェア・パッケージである Disaster Tolerant Cluster Services for OpenVMS が必要です。

詳細については、弊社のサービス担当者にお問い合わせください。

C.1.2 マルチサイト OpenVMS Cluster システムの長所

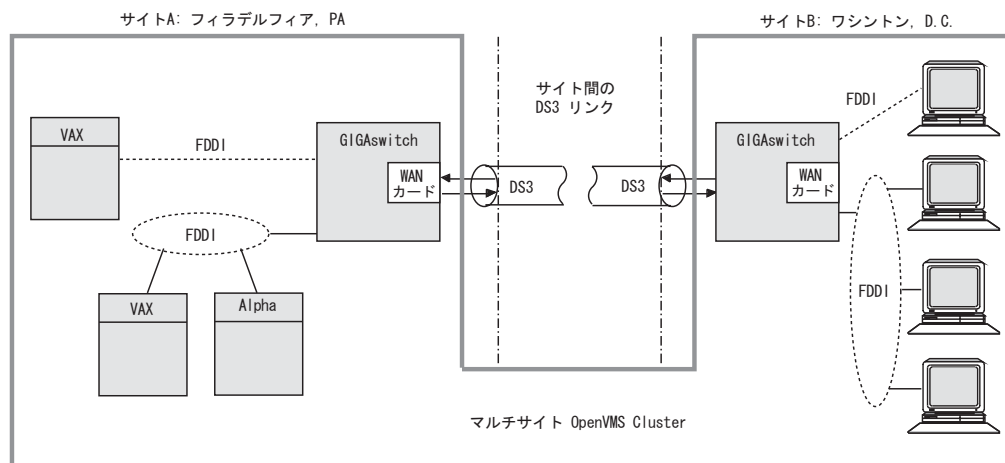
マルチサイト OpenVMS Cluster システムには、次のような長所があります。

長所	説明
リモート・サテライトとノード	図 C-2 に示すように、二次サイトにリモート・サイトを少数配置でき、それらを一元管理できるだけでなく、一次サイトのリソースも利用できます。たとえば、サイト固有デバイスに直接関連付けられた倉庫や小規模工場に、本社のデータ・センターをリンクすることができます。あるいは、一次ビジネス・サイトから都市のビジネス街に数個のエンジニアリング・ワークステーションをインストールすることができます。
データ・センターの管理強化	1 つの管理チームで複数のサイトのデータ・センターにあるノードを管理できます。
リソースの物理的な共用	大容量のコンピュータ、テープ・ライブラリ、ディスク・アーカイブ、フォトタイプセッタなどのデバイスを、複数のサイトで簡単に共用できます。
リモート・アーカイブ	クラスタの任意の場所にあるアーカイバル媒体のバックアップを作成できます。たとえば、一般には、1 個所にディスクやテープを用意してマルチサイト OpenVMS Cluster 内のすべてのサイトのデータをバックアップします。リモート・サイトのデータのバックアップは透過的に処理されます (つまり、リモート・サイトで人が操作する必要はありません)。
強化された可用性	<p>一般に、マルチサイト OpenVMS Cluster は、LAN OpenVMS Cluster の可用性の利点をすべて備えています。また、マルチサイト OpenVMS Cluster 構成では遠距離にある複数のサイトを接続して、さまざまな方法でシステムやシステム要素の可用性を強化します。</p> <ul style="list-style-type: none">論理ボリューム/データの可用性 — ボリューム・シャドウイングや RAID (redundant array of independent disks) を利用すると、両サイトにメンバを配置した論理ボリュームを作成できます。両方のサイトのどちらかが利用できなくなっても、残ったサイトのデータは利用できます。サイト・フェールオーバー — サイトのどれかに障害が発生した場合や、他のサイトとの通信が途絶えたときに自動的に処理を続行するためのサイトを選択するには、VOTES システム・パラメータを調整します。ディザスタ・トレランス — Disaster Tolerant Cluster Services for OpenVMS が提供するソフトウェア、サービス、管理手順と組み合わせると、高度なディザスタ・トレランスが可能です。詳細については、弊社のサービス担当者にお問い合わせください。

マルチサイト OpenVMS Cluster C.1 マルチサイト OpenVMS Cluster システムとは?

図 C-2 は、リモート・サイトからアクセスできるサテライトを備えた OpenVMS Cluster システムです。

図 C-2 リモート・サテライトを備えたマルチサイト OpenVMS Cluster 構成



ZK7235AG E

C.1.3 一般的な構成ガイドライン

LAN の OpenVMS Cluster システムに適用されるものと同じ構成規則が、ATM、DS3、または FDDI サイト間インターコネクトを組み込んだマルチサイト OpenVMS Cluster 構成に適用されます。一般的な LAN 構成規則については、以下のマニュアルを参照してください。

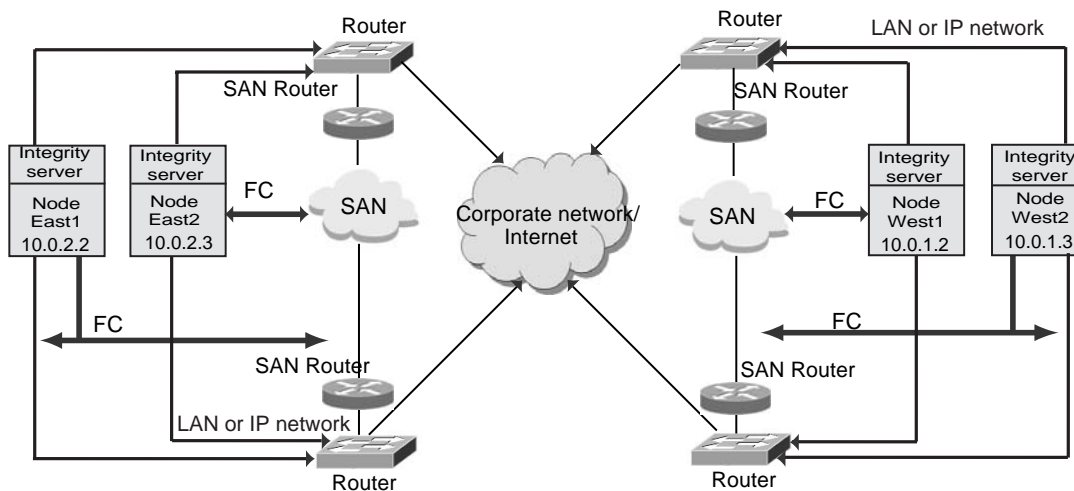
- 『OpenVMS Cluster Software Software Product Description』 (SPD 29.78.xx)
- 本書の第 8 章

マルチサイト OpenVMS Cluster に固有の構成ガイドラインもあります。これらのガイドラインについては、第 C.4.4 項を参照してください。

C.2 Cluster over IP によるマルチサイト OpenVMS Cluster システムの構成

ノードが複数のサイトあるいは複数の LAN に分散して存在している場合、クラスタ通信に IP を使用するのが好まれます。Cluster over IP を使用すると、最大 800 km のサイト間距離でサイトを接続することができます。図 C-3 にストレージの構成に IP インターコネクトを使用したマルチサイト OpenVMS クラスタを示します。

図 C-3 Cluster over IP によるマルチサイト OpenVMS Cluster 構成



ノード East1, East2, West1, および West2 は, 同じ LAN でも異なる LAN でも構いません。クラスタ・トラフィックはルータブルです。東のサイトは Cluster over IP を使用して地理的に離れた西のノードと通信します。ノード East1 およびノード West1 は仮想サーキット (VC) を形成します。VC は, SCS トラフィックのための IP チャネルから構成されます。ただし, 東の 2 つのノードはそれら同士のクラスタ通信に LAN を使用します。ノード East1 は LAN チャネルを使用してノード East2 と仮想サーキットを形成します。

C.3 FDDI によるマルチサイト OpenVMS Cluster システムの構成

VMS バージョン 5.4-3 以降, FDDI は, 2 地点の OpenVMS Cluster サイトを遠距離接続できる最も標準的な方法でした。高速 FDDI 光ファイバ・ケーブルにより, サイト間のケーブル経路の長さが最長で (40 km) 離れたサイト間を接続できます。

サイト間はこれらの FDDI 方式で接続できます。

- パフォーマンスを最大限に活かすには, 各サイトの GIGAswitch/FDDI ブリッジ間で全二重 FDDI リンクを双方向 100 Mb/s で使用して, サイト間帯域幅を最大にします。
- 可用性を最大限に活かすには, 配線コンセントレータまたは GIGAswitch/FDDI ブリッジのデュアル・アタッチメント・ステーション (DAS) ポート間でデュアル FDDI リンクを 100 Mb/s で使用して, 最高のリンク可用性を実現します。
- パフォーマンスと可用性の両方を最大限に活かすには, それぞれ専用のホスト・アダプタと全二重 FDDI サイト間リンクが各サイトで GIGAswitch/FDDI ブリッジに接続された分離 FDDI LAN を 2 つ使用します。

システム管理のその他の説明については、本書と『OpenVMS Cluster システム』の該当箇所を参照してください。これらのマニュアルの最新バージョンの入手方法については、『HP OpenVMS V8.4 リリース・ノート[翻訳版]』を参照してください。

OpenVMS Cluster システム固有の柔軟性と強化された OpenVMS Cluster LAN プロトコルでは、ATM と DS3 のどちらか一方、またはその両方の通信サービスを利用して、複数の OpenVMS Cluster サイトに接続できます。

C.4 WAN サービスによるマルチサイト OpenVMS Cluster システムの構成

ここでは、ATM と DS3 のワイド・エリア・ネットワーク (WAN) サービスの概要と、FDDI インターコネクトを ATM と DS3 のどちらか、または両方の通信サービスにブリッジする方法、そしてこれらのサービスで、マルチサイト OpenVMS Cluster システムを構成する方法を説明します。

ATM サービスと DS3 サービスでは、OpenVMS Cluster システムに構成して WAN 接続ができる長距離のポイント・ツー・ポイント通信が可能です。ATM サービスと DS3 サービスは、たいいていの一般通信事業者や他のソースから利用できます。

注意

ヨーロッパや一部の国々では DS3 を使用できません。また、ATM は新しい、進展中の標準なので、ATM サービスをすべての個所で利用できるとは限りません。

ATM サービスと DS3 サービスは、以下の OpenVMS バージョンで使用できます。

サービス	OpenVMS の認定バージョン
ATM	OpenVMS バージョン 6.2 以降
DS3	OpenVMS バージョン 6.1 以降

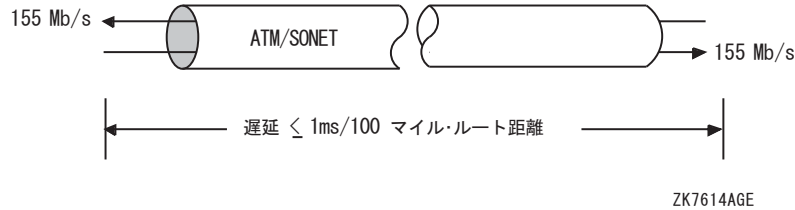
以下の項では、ATM 通信サービスと DS3 通信サービスを説明し、これらのサービスをマルチサイト OpenVMS Cluster システムに構成する方法を説明します。

C.4.1 ATM 通信サービス

SONET 物理レイヤ (ATM/SONET) を使用する ATM 通信サービスは、全二重通信を提供します (図 C-4 に示すように、双方向にビット・レートを同時に使用できます)。ATM/SONET は、複数の標準ビット・レートと互換性があります。FDDI の 100 Mb/s ビット・レートには、155 Mb/s 全二重レートの SONET OC-3 サービスが最適です。ATM/SONET OC-3 は、世界で最も普及している標準サービスです。

ヨーロッパでは、従来の E3 標準に代わるハイ・パフォーマンス標準サービスとして ATM/SONET があります。

図 C-4 ATM/SONET OC-3 サービス

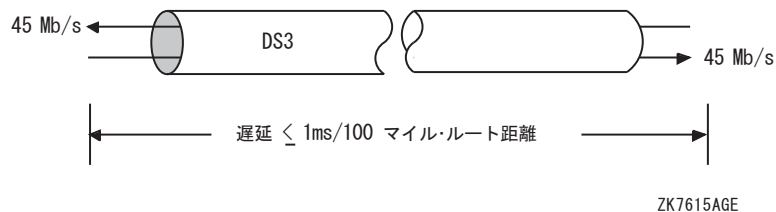


データを伝送するとき、ATM フレーム (パケット) は ATM サービスによって伝送用のセルに分割されます。各セルの大きさは 53 バイトであり、内 5 バイトがヘッダ情報用なので、データに使用できるのは 48 バイトです。伝送先で、セルは ATM フレームに組み立てられます。セルの使用により、ATM サプライヤは、複数のデータ・ストリームを異なるビット・レートで効率的に多重化、多重分離できます。このフレーム対セルの変換は、上位レイヤには透過的に処理されます。

C.4.2 DS3 通信サービス (T3 通信サービス)

図 C-5 に示すように、DS3 通信サービスは全二重通信を提供します。DS3 (T3 ともいう) では、45 Mb/s の T3 標準ビット・レートを利用できます。T3 は、北米や他の多くの国々に普及している標準サービスです。

図 C-5 DS3 Service

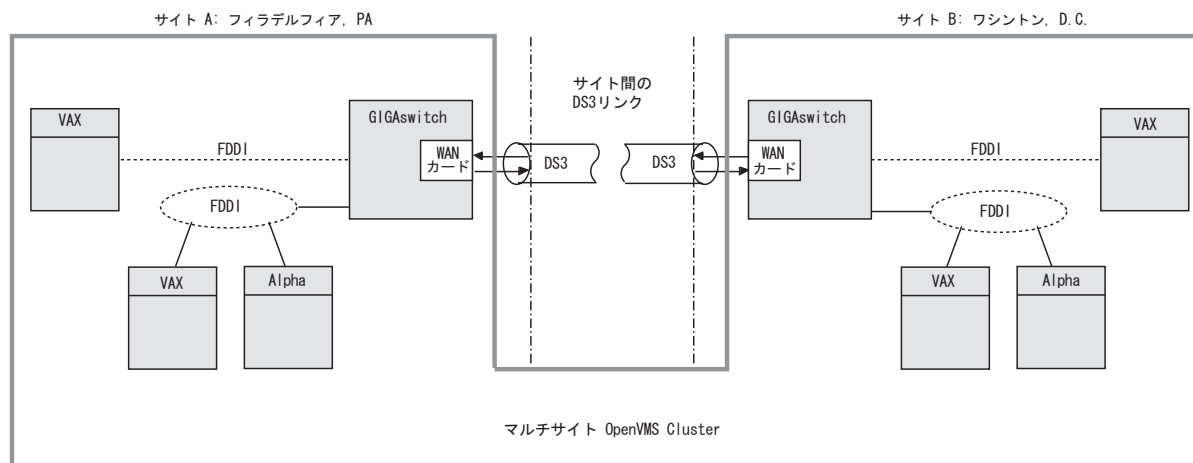


C.4.3 FDDI-to-WAN ブリッジ

図 C-6 に示すように、FDDI-to-WAN (たとえば、FDDI-to-ATM と FDDI-to-DS3 のどちらか、または両方) ブリッジは、距離を隔てて配置されたサイトにノードがある OpenVMS Cluster の構成に使用できます。この図では、各サイトの OpenVMS Cluster ノード同士は、2 つのサイトが FDDI で接続されている場合と同様に通信できます。FDDI-to-WAN ブリッジにより、OpenVMS Cluster ソフトウェアにとって ATM と DS3 の存在は透過になります。

マルチサイト OpenVMS Cluster C.4 WAN サービスによるマルチサイト OpenVMS Cluster システムの構成

図 C-6 DS3 で結合したマルチサイト OpenVMS Cluster 構成



ZK7234AGE

図 C-6 で、FDDI-to-DS3 ブリッジと DS3 は以下のように動作します。

1. ローカル FDDI-to-DS3 ブリッジが、他のサイトのノードにアドレス指定された FDDI パケットを受け取ります。
2. ブリッジは FDDI パケットを DS3 パケットに変換し、そのパケットを DS3 リンクで他のサイトに送信します。
3. 受信側 FDDI-to-DS3 ブリッジは、DS3 パケットを FDDI パケットに変換し、それを、そのサイトの FDDI リングで発信します。

FDDI-to-WAN ブリッジの構築には、GIGAswitch/FDDI システムの使用を推奨します。GIGAswitch/FDDI は、DEFGT WAN T3/SONET オプション・カードと組み合わせて、マルチサイト OpenVMS Cluster システムにおける ATM 通信サービスと DS3 通信サービスの認定テストに使用していました。

C.4.4 OpenVMS Cluster システムにおける ATM と DS3 の構成ガイドライン

マルチサイト OpenVMS Cluster の構成では、サイト間リンクの遅延、帯域幅、可用性、ビット・エラー率の特性がアプリケーションのニーズを満たしているか確認する必要があります。この項では、それらの要件とともに推奨値について説明します。

C.4.4.1 要件

HP の認定構成として、マルチサイト OpenVMS Cluster は以下の規則に従います。

サイト間リンクの最長経路	マルチサイト OpenVMS Cluster のメンバ間で、サイト間リンクのケーブル経路の最長距離は 242 km とします。正確な測定距離については、ATM サブライヤか DS3 サブライヤに確認してください。 この距離の許容上限値は、Disaster Tolerant Cluster Services for OpenVMS を利用することにより、引き上げることができます。Disaster Tolerant Cluster Services for OpenVMS は、OpenVMS ディザスタ・トレラントなクラスタを構成し管理するためのシステム管理機能であり、ソフトウェア・パッケージです。
サイト間リンクの最大利用率	サイト間リンクの任意の 10 秒間の平均利用率は、両方向とも、リンクの帯域幅の 80% を超えないものとします。この利用率を超えると、許容範囲を超えるキュー遅延やパケット・ロスが頻繁に発生する可能性があります。
サイト間リンクの仕様	サイト間リンクは、表 C-3 で指定した OpenVMS Cluster 要件を満たす必要があります。
OpenVMS Cluster LAN 構成規則	LAN 上の OpenVMS Cluster システムの構成規則を構成に適用します。構成規則の内容については、第 C.1.3 項を参照してください。

C.4.4.2 推奨値

DS3 インターコネクトの構成時には、『OpenVMS Cluster Software SPD』(SPD 29.78.nn)、および、本書に掲載してある LAN 接続の OpenVMS Cluster システムの構成ガイドラインに従ってください。各サイトの OpenVMS Cluster メンバには、サテライト、システム、その他 CI や DSSI などのインターコネクトを任意で組み合わせさせて組み込めます。

この項では、マルチサイト OpenVMS Cluster システムの構成に必要なその他の推奨値を紹介します。

DS3 リンク・キャパシティ/プロトコル

WAN T3/SONET オプション・カード付きの GIGAswitch では、全二重の、155 Mb/s ATM/SONET リンクを利用できます。このリンクの全帯域幅は、すべて WAN オプション・カードに充てられます。ただし、GIGAswitch/FDDI の内部デザインは、FDDI に対する全二重拡張機能を基にしています。したがって、GIGAswitch/FDDI のデザインにより、ATM/SONET リンクのキャパシティは両方向とも 100 Mb/s に制限されます。

WAN T3/SONET オプション・カード付き GIGAswitch には、DS3 リンクで使用できるプロトコル・オプションがあります。DS3 リンクは、全帯域幅が WAN オプション・カードに充てられるクリアなチャンネル・モードで使用してください。DS3 リンクのキャパシティは、選択したプロトコル・オプションによって異なります。プロトコル・オプションについては、表 C-1 を参照してください。

表 C-1 DS3 プロトコル・オプション

プロトコル・オプション	リンク・キャパシティ
PLCP ³ 無効の ATM ¹ AAL-5 ² モード	39 Mb/s
PLCP 有効の ATM AAL-5 モード	33 Mb/s
HDLC ⁴ モード (現在は使用可能)	43 Mb/s

¹非同期転送モード
²ATM 適合レイヤ
³物理レイヤ収束プロトコル
⁴高速データリンク制御

リンク・キャパシティを最大限にするには、PLCP を無効にして ATM AAL-5 モードを使用するよう WAN T3/SONET オプション・カードを構成することをお勧めします。

サイト間帯域幅

サイト間帯域幅により、アプリケーション・ロッキング、I/O パフォーマンス (ポリューム・シャドウイングや RAID 設定コピー回数)、ロック・マネージャのパフォーマンスが制約されることがあります。

適切な応答時間を達成するには、平均トラフィックがサイト間リンクの両方向とも、任意の 10 秒間にリンクの帯域幅の 60% を超えないよう設定してください。これを超えると、FDDI-to-WAN ブリッジ内の遅延によりアプリケーション・パフォーマンスが悪化します。

リンクの利用状況を計算するときは、OpenVMS Cluster 通信 (ロッキングや I/O) とネットワーク通信 (TCP/IP, LAT, DECnet など) を必ず考慮してください。

サイト間遅延

サイト間リンクでは、サイト間ケーブル長 100 マイルにつき最高 1 ms の一方向遅延と、両端の FDDI-to-WAN ブリッジにおける遅延を合わせた遅延が生じます。アプリケーション応答時間とスループットに対するサイト間遅延の影響を忘れないでください。

たとえば、サイト間リンクの一方向パス遅延には、以下のような構成要素があります。

- ATM と DS3 のいずれもケーブル経路の 1 方向遅延 1 ms/100 マイル (0.01 ms/マイル)
- FDDI-to-WAN ブリッジ遅延 (約 0.5 ms/ブリッジ, 2 ブリッジ/1 方向トリップ)

ラウンド・トリップ遅延は、以下のように計算します。

WAN ラウンド・トリップ遅延 = 2 x (Nマイル x 0.01 ms/マイル + 2 x 0.5 ms/マイル FDDI-WAN ブリッジ)

MSCP サービスによる I/O 書き込み操作には、少なくともラウンド・トリップ・パケット交換が 2 回必要です。

WAN I/O 書き込み遅延 = 2 x WAN ラウンド・トリップ遅延

したがって、100 マイル WAN リンクによる I/O 書き込みは、短い、ローカル FDDI による同じ I/O 書き込みよりも、少なくとも 8 ms 余分にかかります。

同じく、ロック操作では、パケットのラウンド・トリップ交換が必要です。

WAN ロック操作遅延 = WAN ラウンド・トリップ遅延

I/O 操作において、同期のために N 回ロックをすると、WAN による遅延は以下のようになります。

WAN ロックの I/O 操作遅延 = $(N \times \text{WAN ロック操作遅延}) + \text{WAN I/O 遅延}$

ビット・エラー率

ビット・エラー率 (BER) パラメータは、サイト間リンクにおけるビット・エラーの発生確度を示す重要な尺度です。マルチサイト OpenVMS Cluster を構成するときは、アプリケーション・スループットと応答に対するビット・エラーの影響を考慮する必要があります。サイト間リンク・ビット・エラーが発生すると、パケットが失われたり、アプリケーション I/O 応答時間 (第 C.4.6 項参照) の遅延後に再送されることがあります。ビット・エラーごとのアプリケーション遅延が、200 ミリ秒から 300 ミリ秒になると、パケットが失われる可能性があります。

サイト間リンクの可用性

サイト間リンク・サービスに割り込みがあると、接続が復元されるまで 1 個所以上のサイトのリソースが利用できなくなることがあります (第 C.4.5 項参照)。

システム・ディスク

クォーラム・ボーツに寄与するノードを備えたサイトには、ローカル・システム・ディスクやそれらのノード用のディスクがあります。

システム管理

大規模なマルチサイト OpenVMS Cluster には、各種業務担当者が使用する多数の異種システムからなる環境をサポートできるように、トレーニングを受けたシステム管理スタッフが必要です。

マイクロ波 DS3 リンク

DS3 リンクには部分的に、マイクロ波無線機器を接続できます。第 C.4.6 項の仕様は、DS3 リンクに適用します。マイクロ波無線部分の BER と可用性は、局地的な天候やリンクのマイクロ波部分の長さの影響を受けます。マイクロ波を DS3 リンクの部分として使用する場合は、マイクロ波コンサルタントの協力を得てください。

C.4.5 可用性の考慮

複数のサイトを接続する FDDI-to-WAN ブリッジとリンクが一時的に利用できなくなった場合は、以下のような状況になります。

- サイト間リンク障害は、1 個所または複数の個所で発生することがあり、その場合は、サイト間接続が復元されるまで利用できなくなります。
- サイト間リンクビット・エラー (と ATM セル・ロス) と非可用性は、次の要素に影響があります。
 - システムの応答性
 - システム・スループット (または帯域幅)
 - 仮想回路 (VC) 閉止率
 - OpenVMS Cluster 遷移とサイト・フェールオーバー時間

多くの通信事業者は、パスの多様性、保護の切り替え、その他、サイト間リンクの可用性を大幅に増加するオプションなど、可用性強化オプションを用意しています。

C.4.6 仕様

この項では、WAN 通信サービスにおける通信とパフォーマンスの要件について説明します。

WAN サービス・サプライヤとの交渉上の便宜を考え、この項では、電気通信事業者が広く使用している WAN 仕様の用語と定義を使用します。これらの要件と目標値は、Bellcore Communications Research の仕様と、OpenVMS Cluster のエラーによる影響を弊社で解析した結果の組み合わせから抽出したものです。

表 C-2 で説明する用語は、表 C-3 で使用した Bellcore と OpenVMS Cluster の要件と目標値の説明に使用されます。

表 C-3 で指定した ATM/SONET - OC3 と DS3 サービス・エラー・パフォーマンス (品質) の Bellcore 要件と OpenVMS Cluster 要件は、サービス・サプライヤのサービス品質、可用性、ダウン・タイム、サービス割り込み頻度の目標値がシステムにどう影響するかを評価するときに使用します。

注意

OpenVMS Cluster システムがアプリケーションの応答時間の要件を満たしているかを確認するには、表 C-3 で定めた Bellcore と OpenVMS Cluster の要件および目標値を超える WAN 要件を達成する必要があります。

表 C-2 Bellcore と OpenVMS Cluster の要件と目標値の用語

仕様	要件	目標
Bellcore Communications Research	<p>Bellcore の仕様は、Bellcore Technical Reference TR-TSY-000499 TSGR: Common Requirementsにまとめられた推奨の“一般的なエラー・パフォーマンス要件と目標値”です。これらの仕様は、WAN サプライヤがサービス保証として採用しています。FCC でも、これらを通信事業者間のタリフ・サービスとして採用しています。ただし、サプライヤによっては、顧客の要求に応じてさらに高品質なサービス提供を契約をする場合もあります。</p> <p>他の国々では、Bellcore 仕様とパラメータに匹敵する仕様とパラメータを用意しています。</p>	<p>これらは推奨最低値です。Bellcore では、これらの目標値をTSGR: Common Requirementsマニュアルの“目標値”と呼んでいます。</p>
OpenVMS Cluster	<p>HP が構成を認定するには、表 C-3 の OpenVMS Cluster 要件列に示す値以上の値にパラメータが応じる必要があります。</p> <p>目標値を達成していない値があると、インターコネク ト・エラー回復遅延と、OpenVMS Cluster の状態遷移とサイ ト・フェールオーバーのどちらか、または両方の原因になる VC 閉止により、OpenVMS Cluster で満足なパフォーマンスが得られなくなり ます。</p> <p>これらの値が目標値以上であれば、インターコネク ト・ビット・エ ラー関連の回復遅延があっても OpenVMS Cluster の平均スルー ットが極端に悪化するのを防ぐことができます。また、OpenVMS Cluster の応答時間は、全般的に満足できる値になります。</p> <p>要件が満たされないと、1 時間当たり複数回のアプリケーション一時 停止が発生します。¹</p>	<p>OpenVMS Cluster 操作を最適化するには、すべてのパラメータが OpenVMS Cluster の目標値以上の値を満足するものとします。</p> <p>これらの目標値以上の値が満足されている場合は、インターコ ネク ト・ビット・エラーとビット ト・エラー回復遅延があっても、OpenVMS Cluster の平均スルー ットが極端に悪化するのを防ぐこ とができます。</p> <p>アプリケーションの短時間の一時 停止が毎日2、3 回発生するかもし れませんが、OpenVMS Cluster の応答時間は、全般的に満足でき る値になります。²</p>

¹ビット・エラーによるパケット・ロスのため、アプリケーションの一時停止はほぼ毎時間発生します (OpenVMS Cluster 要件の説明と同様)。

²一時停止の原因は、1 個以上の NISCA 伝送仮想回路上の損失パケットによる仮想回路の再送タイムアウトです。1 回の一時停止は、200 ~ 300 ミリ秒から 2 ~ 3 秒続きます。

マルチサイト OpenVMS Cluster
C.4 WAN サービスによるマルチサイト OpenVMS Cluster システムの構成

表 C-3 OpenVMS Cluster DS3 と SONET OC3 エラー・パフォーマンスの要件

パラメータ	Bellcore の要件	Bellcore の目標値	OpenVMS Cluster の要件 ¹	OpenVMS Cluster の目標値 ¹	単位
エラー秒 (% ES)	<1.0%	<0.4%	<1.0%	<0.028%	% ES/24 hr
	次のように、ES パラメータは、エラー秒でも表すことができます。				
	<864	<345	<864	<24	ES per 24-hr period
バースト・エラー秒 (BES) ²	≤4	–	≤4	Bellcore の目標値	BES/day
ビット・エラー率 (BER) ³	1 × 10 ⁻⁹	2 × 10 ⁻¹⁰	1 × 10 ⁻⁹	6 × 10 ⁻¹²	Errored bits/bit
DS3 チャンネルの使用不可状態	なし	≤97 @ 250 マイル, 24 @ ≤25 マイルまで直線的に低下	なし	Bellcore の目標値	Min/yr
SONET チャンネルの使用不可状態	なし	≤105 @ 250 マイル, 21 @ ≤50 マイル直線的に低下	なし	Bellcore の目標値	Min/yr
チャンネルの使用不可イベント ⁴	なし	なし	なし	1 から 2	Event/year

¹アプリケーション要件は、OpenVMS Cluster 要件列に表示した値よりも厳しくなる場合があります。

²多数の日数の平均値。

³測定期間中に発生したバースト・エラー秒を含みません。

⁴1 年に発生したチャンネル・ダウンタイム期間の平均回数。このパラメータは、チャンネルが使用できなくなる頻度を指定するときに使用します。

テーブル・キー

- 可用性 — 意図した通りの動作を伝送チャンネルが実行した長期のまとまった時間、または時間をパーセントで表した値。可用性を非可用やダウンタイムで表すこともしばしばあります。
- BER (ビット・エラー率)—“BER は、測定時間内に伝送された合計ビット数に対するエラー内のビット数の割合であり、測定時間内のバースト・エラー秒 (下記定義参照) は含みません。バースト・エラー秒内では、ビット・エラーの回数もビット数もカウントしません。”
- BES (バースト・エラー秒)—“バースト・エラー秒は、最低 100 件のエラーを含むエラー秒を指します。”
- チャンネル —SONET リンクや DS3 リンク向けの Bellcore 『TSGR: Common Requirements』マニュアルで使用するリンク用語です。
- ダウン・タイム — 指定した期間 (例、1 年) 内に、伝送チャンネルが使用できない長期間の平均時間 (例、分)。
“... チャンネルの非可用やダウン・タイムは、10 秒間[またはそれ以上の] SES (重大なエラー秒) が連続すると、それをもって開始とし、10 秒間非 SES が発生すると、それをもって終了とします。”
非可用時間は、10-SES シーケンスの最初の SES からカウントします。
“非可用時間の最後は、[非-SES]シーケンスの最初の無障害の 1 秒までカウントします。”
- ES (エラー秒)—“エラー秒は、最低 1 つのエラーがある 1 秒を指します。”
- SES (重大なエラー秒)—“...SES は、BER が10⁻³を超えた秒を指します。”

C.5 マルチサイトにおける OpenVMS Cluster システムの管理

一般に、マルチサイト OpenVMS Cluster の管理には、LAN で接続した OpenVMS Cluster の場合と同じツールおよび技法を使用します。以下の項では、新たな注意事項と、推奨システム管理ツールおよび技法について説明します。

以下の表は、マルチサイト OpenVMS Cluster システムに固有のシステム管理上の注意事項です。

問題点	対処方法
マルチサイト構成では、以下の障害モードの発生する確率が高くなります。 <ul style="list-style-type: none">• サイト間通信リンク障害による OpenVMS Cluster クォーラム・ロス。• 停電やその他ブレイクダウンによるサイト・ロスが発生すると、そのサイトの全システムに影響が出る。	ポーツについては、サイト間通信に障害が発生したり、他のサイトが利用できなくなっても 1 つの優先サイトで定足数を維持し、処理を続行できるだけの十分なポーツを割り当ててください。また、最も重要なサイトを一次サイトとして選択してください。重要度の低いシステムやサテライトが少ないサイトには、十分なポーツを割り当てる必要はありません。
以上のような障害が発生してもローカル・リソースを引き続き利用できるか、ただちに回復することを期待しますが、必ずしもそうなるとはかぎりません。	以下のオプションがユーザの期待に対応できる可能性があります。 <ul style="list-style-type: none">• 管理とユーザの要求については、障害の確率効果を考慮して設定し、クォーラム・ロスやその他の問題でシステムが応答しない場合は、リモート・サイトに対応できるようにユーザをトレーニングしてください。• 管理ポリシーと、これらの障害モードの特定と処理に必要な操作手順を決定してください。この手順には、サイト処理の続行に必要なクォーラムの手動での調節も含まれます。

C.5.1 手法とツール

以下のシステム管理手法とツールでは、リモート・ノードとローカル・ノード両方を管理できます。

- リバース LAT モードによる DECserver でサイト間リンクを使用する場合、リモート・サイト・コンソール・アクセスには 2 つのオプションがあります。
 - 以下のツールは、リモート・コンソールの接続に使用します。
 - + SET HOST/LAT コマンド
 - + POLYCENTER Console Manager
 - + OpenVMS Cluster Console System (VCS)
 - + HP のシステム管理、および、ソフトウェアパッケージ Disaster Tolerant Cluster Services for OpenVMS

- リモート・システム・コンソールのダイアル・アップにはモデムを使用しません。
- リモート・サイト・コンソール・アクセスの代替方法として、各サイトにシステム管理者を配置する方法があります。
- OpenVMS Cluster システム内のすべてのノードでデバイスとプロセッサの制御コマンドを有効にするには、OpenVMS オペレーティング・システムに組み込まれているシステム管理ユーティリティ (SYSMAN) を使用します。

C.5.2 パフォーマンスの監視

マルチサイト OpenVMS Cluster システムのパフォーマンスは、以下のように監視します。

- 仮想回路 (VC) パケット・ロスとラウンド・トリップ時間の値は、SDA (System Dump Analyzer) で監視します。手順については、『OpenVMS Cluster システム』を参照してください。
- サイト間リンクのビット・エラー率 (BER) とパケット・ロスは、ネットワーク管理ツールで監視します。GIGAswitch や WAN T3/SONET オプション・カードの管理情報をアクセスしたり、警告しきい値を設定するには、POLYCENTER NetView や DECMCC などのツールを使用します。必要に応じて、GIGAswitch、WAN T3/SONET カード、POLYCENTER、DECMCC ドキュメントを参照してください。
- Availability Manager、HP ECP (CP/Collect および CP/Analyze)、あるいは Unicenter Performance Management for OpenVMS (以前の Polycenter Performance Solution Data Collector および Performance Analyzer、さらに以前の SPM および VPA) を使用してクラスタの性能を監視することができます。

A

Accounting ユーティリティ	2-5
Advanced Server for OpenVMS	1-6
Alpha システム	
Integrity システムとの共存	3-1
Alpha ワークステーション	9-14, 10-4
ARC コンソール	A-38
ATM (非同期転送モード)	
サイト間リンクの仕様	C-12
通信サービス	C-6
AUTOGEN コマンド・プロシージャ	2-5, 8-6, 9-5
Availability Manager	
OpenVMS Cluster の可用性	8-4

C

CCL	
コマンド・コンソール LUN を参照	
CIFS for OpenVMS	1-6
CLUSTER_CONFIG.COM コマンド・プロシージャ	7-54
CLUSTER_CONFIG_LAN.COM コマンド・プロシージャ	7-54
Cluster over IP	
マルチサイト OpenVMS Cluster	C-4

D

DCL レキシカル	
FSGETDVI	7-28
DECamsds	
操作管理	2-5
DECEvent	
Fibre Channel 構成での使用	7-8
DECnet ソフトウェア	
DECnet-Plus ネットワーク通信	1-6
OpenVMS Cluster エイリアス	8-3
DECram for OpenVMS	1-6, 5-5, 9-23
DS3 サイト間リンクの仕様	C-12
DS3 通信サービス	C-6

E

Ethernet	4-14
Fast	4-1
Gigabit	4-1, 4-18
10 Gigabit	4-1
LAN セグメントのブート	9-17
ガイドライン	9-17
拡張 LAN のガイドライン	9-17
6 サテライト OpenVMS Cluster	9-9, 9-10
12 サテライト OpenVMS Cluster	9-11
サテライトの制約	9-17
スケーリング	9-3
スループット	4-17
51 ノード OpenVMS Cluster	9-13
ルーティングの流れの考慮	4-17
EXPECTED_VOTES システム・パラメータ	9-19, 10-12

F

FSGETDVI レキシカル関数	
キーワード WWID	7-28
Fast Ethernet	4-1
スケーリング	9-3
FCSCP	
Fibre Channel Control Program を参照	
FCTC-II	7-23
FDDI (Fiber Distributed Data Interface)	
Alpha ワークステーション	9-14
スケーラビリティ	9-3
マルチサイト OpenVMS Cluster	C-5
FGx0 アダプタ	7-18
FIBRE_SCAN ユーティリティ	7-63
Fibre Channel Control Program (FCSCP)	7-61
Fibre Channel Tape Controller II	7-23
Fibre Channel WWID	
デバイス ID を参照	
Fibre Channel	
Fibre Channel アドレス指定	7-15
Fibre Channel インターコネクト	4-4, 7-1
サポートされている媒体	7-2
Fibre Channel 構成	7-9
FIBRE_SCAN ユーティリティ	7-63
OpenVMS Cluster システムに対するノードの追加	7-54
SHOW DEVICE コマンド	7-47

Fibre Channel 構成 (続き)	
wwidmgr コマンド	7-41
オンライン再構成	7-60
共用 FC システム・ディスク	7-54
クラスタの有効化	7-54
情報の表示	7-41
性能データ	7-63
ダンプ	7-44
デュアル・ポート・ストレージ・コントローラ	6-25
パス識別子	6-33, 7-19
ブート	7-44
マルチパス	6-25
Fibre Channel 構成のマルチパス・テープのサポート	
Fibre Channel ディスク	
ダンプ	7-44
ブート	7-44
Fibre Channel デバイス	
Fibre Channel ディスク, Fibre Channel テープ・デバイス, および MDR も参照	
FGx0 アダプタ	7-18
HSG コントローラ	7-52
PGx0 SCSI 擬似デバイス	7-18
スイッチ	6-1, 7-1
ストレージ・アダプタ名	7-18
ネーミング規則	7-5
Fibre Channel テープ・デバイス	
移動	7-37
交換	7-38
構成	7-28
提供	7-38
名前の変更	7-37
物理位置の検索	7-39
命名	7-23
Fibre Channel のテープ・ブリッジ	
スワップ	7-38
G	
Gigabit Ethernet	
スケーリング	9-3
10Gigabit Ethernet	4-1
10 Gigabit Ethernet	
構成上のガイドライン	4-18
スケーリング	9-3
H	
HP Insight Management Agents for OpenVMS	
	2-5
HP Integrity システム	
	3-1
HP OpenVMS Integrity オペレーティング・システム	
	3-1
HSG60 コントローラ	7-2
HSG80 コントローラ	6-3, 6-25, 7-2
HSG コントローラ	
OpenVMS の構成	7-52

HSG コントローラ (続き)	
デバイスが構成されない	7-60
HSV110 コントローラ	6-3
HSV110 ストレージ・コントローラ	7-52
HSZ70 コントローラ	6-3
HSZ80 コントローラ	6-3
HSZ 割り当てクラス	
割り当てクラスを参照	

I	
I/O スループット	
MONITOR IO コマンド	9-25
MONITOR MSCP コマンド	9-25
キャッシュ	9-24
スケーリング	9-2, 9-3, 9-21 ~ 9-26
ディスク・テクノロジー	9-23
パケット・サイズ	9-24
ホット・ファイル	9-25
ボリューム・シャドウイング	9-25
読み込み/書き込みの比率	9-23
I/O パス	
可用性	8-3
フェールオーバー	8-3
I/O 要件	2-3
Integrity システム	
Alpha システムとの共存	1-1
IO FIND_WWID SYSMAN コマンド	7-27
IO FIND_WWID コマンド	7-30
IO REPLACE_WWID SYSMAN コマンド	7-27, 7-38
IP (インターネット・プロトコル)	4-19
ISL (インタースイッチ・リンク)	7-2

L	
LAN (Local Area Network)	
アダプタ	
LAN パスに接続された	9-17
複数	8-10
高可用性	8-6
構成	8-10
セグメント	
OpenVMS Cluster 構成における 2 つの	8-10
OpenVMS Cluster 構成における 3 つの	8-12
複数	8-10
複数バス	4-15
マルチサイト OpenVMS Cluster	8-21, C-3
利用率の削減	10-4
LAT ソフトウェア	
OpenVMS Cluster の可用性	8-3
LOCKDIRWT システム・パラメータ	9-19

M

MCDRIVER	B-12
MDR (Modular Data Router)	7-2f, 7-22, 7-41
MEMORY CHANNEL	4-6
NPAGEVIR パラメータ	B-5
OpenVMS Cluster アーキテクチャの	B-12
SMP との比較	B-10
アダプタ	4-7
アドレス指定	B-12
技術概要	B-10
グローバル書き込み	B-13
高可用性 OpenVMS Cluster システム	8-15
高可用性の構成	B-4
構成要件	B-8
構成例	B-5
サポートしているシステム	B-9
実装	B-15
従来のネットワークとの比較	B-10
スループット	4-7
製品概要	B-1
定義	4-5
特長	B-2
ドライバ	B-12
ハードウェア	B-3
ページ制御テーブル	B-14
メモリの要件	B-5
Modular Data Router	
MDR を参照	
MONITOR IO コマンド	
ホット・ディスクの分離	9-25
MONITOR MSCP コマンド	
ホット・ファイル用	9-25
Monitor ユーティリティ	2-6
MOP サーバ	
可用性のための選択	8-10
MPDEV_AFB_INTVL システム・パラメータ	6-32
MPDEV_D1 システム・パラメータ	6-32
MPDEV_ENABLE システム・パラメータ	6-32
MPDEV_LCRETRIES システム・パラメータ	6-32
MPDEV_POLLER システム・パラメータ	6-32
MPDEV_REMOTE システム・パラメータ	6-32
MSCP_LOAD システム・パラメータ	9-19
MSCP サーバ	
I/O のオーバーヘッド	9-22
OpenVMS Cluster での役割	1-4
サテライトの追加	9-4
サービス対象のパケット	9-22
動的負荷バランス調整	9-17
負荷分散	4-3
読み込み/書き込みの比率	9-24
利点	9-22

N

NETNODE_REMOTE.DAT ファイル	10-7
NETNODE_UPDATE.COM ファイル	10-7
NETPROXY.DAT ファイル	10-6, 10-7
NFS (ネットワーク・ファイル・システム)	
サーバ・ソフトウェア	1-6
NISCS_MAX_PKTSZ システム・パラメータ	9-19
NPAGEDYN システム・パラメータ	9-19
NSR (Network Storage Router)	7-22, 7-29, 7-41
IO FIND_WWID コマンド	7-30
IO REPLACE_WWID コマンド	7-38
SCSI-LUN	7-39

O

OpenVMS Cluster システム	
CLUSTER_CONFIG.COM	A-29
DECnet-Plus 要件	1-8
Fibre Channel 構成のマルチパス・テープのサポート	7-40
SCSI 構成上の要件	A-4
SCSI ストレージ・インターコネクト	A-1
SCSI のインストール	A-21
SCSI ハードウェア構成	A-11
SCSI パフォーマンス	A-8
TCP/IP 要件	1-8
移行サポートと保証サポート	10-14
エイリアス	8-3
環境上のリスク	8-5
共通環境	10-2, 10-7
共有 SCSI ストレージ	4-8
共用 SCSI ストレージの概念	A-7
高可用性	8-15
構成	
一般規則	1-7
構成要素	
オペレーティング・システム	1-4
ソフトウェア	1-3
ハードウェア	1-2
混成アーキテクチャ・サポート	10-14
混成バージョン・サポート	10-14
サイズと複雑さ	8-6
修正キット	10-14
将来の機能拡張計画	2-2
シンプルと複雑	10-1
セキュリティ	10-10
ソフトウェアの変更	8-6
単一ホストに接続された SCSI ストレージ	A-3
ツールとユーティリティ	2-5
ディザスタ・トレラントな	8-22
ハードウェアの変更	8-6
ビジネス要件の決定	2-1

OpenVMS Cluster システム (続き)	
複数のホストに接続された SCSI ストレージ	
ジ	A-3
分割	10-6
ホスト, ノード, コンピュータ	A-1
マルチ環境	10-2
長所	10-9
OpenVMS Integrity Cluster システム	4-8
OpenVMS オペレーティング・システム	
構成要素	1-4
OpenVMS ユーティリティ	
Monitor ユーティリティ	2-6
Show Cluster ユーティリティ	2-6

P

PAGEDYN システム・パラメータ	9-19
PGx0 SCSI 擬似デバイス	7-18
PMDRIVER	B-12, B-15

Q

QDSKINTERVAL システム・パラメータ	10-13
QMANSMASTER.DAT ファイル	10-6, 10-7, 10-11

R

RAID (Redundant Arrays of Independent Disks)	
SCSI ストレージ	A-3
ディスク・ストライピング (レベル 0)	1-6, 5-6, 9-23
ホット・ファイル	9-25
パリティ付きディスク・ストライピング (レベル 5)	1-6
ボリューム・シャドウイング (レベル 1)	1-6
RECINXINTERVL システム・パラメータ	9-19, 10-13
Redundant arrays of independent disks	
RAID を参照	
RIGHTSLIST.DAT ファイル	10-7
RMS	
OpenVMS Cluster での役割	1-4

S

SAS インターコネク	
アダプタ	4-14
サポートデバイス	5-8
ストレージ	5-8
定義	4-13
SCC モード	7-22
SCS (システム通信サービス)	
OpenVMS Cluster での役割	1-4
SCSI_PATH_VERIFY コマンド	
FC ディスク・デバイス	7-60
SCSI Command Controller (SCC) モード	7-22

SCSI インターコネク	A-1
ANSI 標準	A-1
TERMPWR ライン	A-10
アダプタ	4-12
アービトレーション規則	A-41, A-42
インストール	A-21
概念	A-7
ケーブル配線と終端	A-10
コンソール・パラメータの表示と設定	A-26
最大距離	4-11, A-9
サポートされているデバイス	5-9
サポートできるデバイス数	A-7
スケーラビリティ	9-5
ストレージ	5-9
制御	A-41
接地要件	A-22
単一ホストに接続された	A-3
定義	4-7
データ転送速度	A-9
デバイスのホット・プラグ	A-44
電源入力と検査	A-25
トラブルシューティング	A-39
バスのリセット	A-32, A-38
パフォーマンス	A-8
複数のホストに接続された	A-3
利点	4-10

SCSI 構成

BA350/BA353 StorageWorks エンクロージャによる構築	A-12
CLUSTER_CONFIG.COM による構成	A-29
DWZZx コンバータによる SCSI OpenVMS Cluster システムの構築	A-14
OpenVMS Cluster システムにおけるシングル・ホスト SCSI アクセス	A-3
OpenVMS Cluster システムにおけるマルチ・ホスト SCSI アクセス	A-3
SCSI OpenVMS Cluster システム	A-1
SCSI OpenVMS Cluster システムのインストール	A-21
SCSI OpenVMS Cluster システムのトラブルシューティング	A-39
SCSI OpenVMS Cluster システム向けの CLUSTER_CONFIG の使用	A-29
SCSI インターコネク	A-4
SCSI ノード ID の構成	A-22
SCSI ハードウェア	A-11
固有の SCSI デバイス ID	A-8
シングル・エンドとディファレンシャルの SCSI シグナル通知	A-9
デバイス ID の構成	A-23
3 ノード Fast Wide OpenVMS Cluster	9-6
SCSI ハブを備えた	9-7
4 ノード Ultra SCSI OpenVMS Cluster	9-7
ハードウェア	A-11
BA350/BA353 StorageWorks エンクロージャ	A-12
DWZZx コンバータ	A-14
HSZ40 コントローラ	A-14

SCSI 構成		
ハードウェア (続き)		
アドオン SCSI アダプタ	A-12	
マルチバス	6-3	
マルチパス	6-15	
透過的フェールオーバー	6-15	
要件	6-8	
マルチポート・ストレージ・コントローラ	6-17	
要件	A-4	
SCSI コンソール		
パラメータの設定	A-26	
SCSI ディスク		
アクセス	A-2	
構成上の要件	A-4	
固有のデバイス ID	A-8	
操作モード	A-8	
SCSI の概念		A-7
SCSI バス		
SCSI インターコネクトを参照		
SET DEVICE/[NO]POLL コマンド	6-38	
SET DEVICE/PATH コマンド	6-28	
SET UNIT PREFERRED_PATH コマンド	6-39	
SHADOW_MAX_COPY システム・パラメータ	9-19	
Show Cluster コーティリティ	2-6	
SHOW DEVICE/FULL DCL コマンド		
WWID の表示	7-28	
SHOW DEVICE/FULL コマンド	6-34	
SHOW DEVICE/MULTIPATH_SET コマンド	6-34	
Small Computer System Interface		
SCSI を参照		
SMP	9-2, 9-4	
SONET OC-3 サイト間リンクの仕様	C-12	
StorageWorks		
BA350/BA353 エンクロージャ	A-12	
定義	5-1	
StorageWorks Data Replication Manager	7-7	
SYLOGICAL.COM 手順		
システム・ディスク上の	10-6	
SYSSYSTEM:SYSSDEVICES.DAT ファイル		
Fibre Channel デバイス名	7-26	
SYSMAN コマンド		
IO FIND_WWID	7-27	
IO REPLACE_WWID	7-27, 7-38	
System Dump Analyzer (SDA)		
コマンド		
FC 性能	7-63	
SYSUAF.DAT ファイル	10-6, 10-7, 10-8	
複数	10-10	

T

T3 通信サービス	C-7
TCP/IP Services	
ネットワーク通信	1-6
TCP/IP Services for OpenVMS ソフトウェア	
ア	1-6
TIMVCFAIL システム・パラメータ	10-13
TMSCP サーバ	
OpenVMS Cluster での役割	1-4

U

UDID	7-24
Ultra SCSI 構成	
SCSI 構成を参照	

V

VIOC	
仮想 I/O キャッシュを参照	
VMSMAIL_PROFILE.DATA ファイル	10-7
VMS クラスタ	
OpenVMS Cluster を参照	
Volume Shadowing for OpenVMS	
I/O スループットに対する影響	9-25
MSCP のオーバーヘッド	9-25
OpenVMS Cluster の可用性	8-4
RAID レベル 1	1-6
SCSI OpenVMS Cluster 構成でサポーター	
ト	A-5
スケラビリティ	9-3
ストライピングとの組み合わせ	5-6
説明	5-7
読み込み/書き込みの比率	9-26
読み込みパフォーマンス	10-3
VOTES システム・パラメータ	9-19

W

WAN (Wide Area Network)	
マルチサイト OpenVMS Cluster システムにおける利用	C-2
WAN T3/SONET オプション・カード	C-8, C-9
管理について	C-16
プロトコル・オプション	C-10
WAN サイト間リンクの仕様	C-12
WAN 要件	
マルチサイト・クラスタ	
Bellcore Communications Research	C-12
OpenVMS Cluster	C-12
WWID	
デバイス ID を参照	
Fibre Channel	
wwidmgr -quickset コマンド	7-44

wwidmgr -show wwid コマンド	7-42
『Wwidmgr Users' Manual』	7-25

X

XFC

拡張ファイル・キャッシュを参照

ア

アーキテクチャ

Alpha と Integrity	3-1
Alpha と VAX	10-14
Integrity と Alpha	1-1

アクセス・パス

システム・ディスク	5-7
複数の	5-10
ローカル・アダプタ	5-10

アダプタ

MEMORY CHANNEL	4-7
SAS	4-14
SCSI	4-12
アドオン SCSI	A-12
マルチ LAN	8-9
ローカル	5-10

アービトレーション規則

SCSI インターコネクットの制御	A-41
SCSI バスの制御	
影響の変更	A-42

アプリケーション

可用性を強化するコーディング	8-6
要件	2-3

アレイ

ディスクとテープのスケーリング用	9-3
------------------	-----

イ

移行サポート	10-14
--------	-------

インストール手順

SCSI OpenVMS Cluster システム	A-21
SCSI ノード ID の構成	A-22

インターコネクト

ATM	C-6
Cluster over IP	4-1
DS3	C-6
Ethernet	4-14
FDDI	C-5
Fibre Channel	4-4
IP	4-19, C-4
MEMORY CHANNEL	4-5
SAS	4-13
SCSI	4-7, A-1
SCSI ストレージのアクセス	A-2
SCSI のトラブルシューティング	A-39
混成	4-4
スケーリング	9-3
ストレージ・デバイスへの接続	5-2
タイプ	1-3, 4-1

インターコネクト (続き)

定義	4-1
特徴	4-2
フェールオーバー	8-3
複数	4-3

インタースイッチ・リンク

ISL を参照

インターネット・ネットワーキング・プロトコ

ル	1-6
---	-----

エ

エイリアス

OpenVMS Cluster システムを参照

エラー処理要件	C-12
---------	------

エラー・ログ・エントリと OPCOM メッセー

ジ	A-38
---	------

オ

オブティマイザ

ストレージの可用性	5-7
ディスク・パフォーマンス	5-5

カ

拡張 LAN

構成のガイドライン	9-17
-----------	------

拡張ファイル・キャッシュ	5-6
--------------	-----

仮想 I/O キャッシュ	5-6
--------------	-----

可用性

LAN OpenVMS Cluster システムにお

ける	8-6
----	-----

LAN 構成要素の障害後	8-9
--------------	-----

MEMORY CHANNEL OpenVMS Cluster システ	
------------------------------------	--

ムにおける	8-15
-------	------

MOP サーバの選択を通じて	8-10
----------------	------

SCSI ストレージ領域	8-15
--------------	------

強化手法	8-4
------	-----

サテライト OpenVMS Cluster システムにお

ける	8-17
----	------

システム・ディスク	5-7
-----------	-----

ストレージ・オブティマイザ	5-7
---------------	-----

ディスク	5-7
------	-----

データ	5-7
-----	-----

データとアプリケーション	1-2
--------------	-----

バックアップ・コピー	5-8
------------	-----

フェールオーバー機能	3-2
------------	-----

ブート・サーバの増加	10-4
------------	------

保守の強化方法	8-5
---------	-----

ボリューム・シャドウ	5-7
------------	-----

マルチサイト OpenVMS Cluster システムにお

ける	8-20
----	------

要件	2-2
----	-----

レベル	2-2, 8-1
-----	----------

環境

共通	10-2, 10-8
----	------------

環境 (続き)	
マルチ	10-2, 10-9
環境ファイル	
管理	10-2

キ

記憶容量	
増加量を計算する	5-4
定義	5-3
キャッシュ	5-6, 9-24
キャパシティの計画	
可用性について	8-5
スケラビリティに関する注意事項	9-4
キューイング遅延	
拡張 LAN における	9-18
キュー・マネージャ	
複数	10-10
共通環境 OpenVMS Cluster システム	10-8
長所	10-8
別々のディスク	10-8
共有 SCSI ストレージ	
OpenVMS Integrity サーバの 2 ノード・クラス	
タ	4-8
共用ストレージ領域	8-4

ク

クォーラム	8-5
定義	10-11
クォーラム・ディスク	
ウォッチャ	10-11
手法	10-11
クラスタ	
OpenVMS Cluster を参照	
クラスタ全体のプロセス・サービス	
OpenVMS Cluster での役割	1-4

コ

構成	
Fibre Channel 構成または SCSI 構成を参照	
構成要件	
Fibre Channel テープ・デバイス	7-28
コマンド・コンソール LUN (CCL)	
デバイス識別子の設定	7-53
混成アーキテクチャ OpenVMS Cluster システ	
ム	1-1, 3-1, 10-14
混成バージョン OpenVMS Cluster システ	
ム	10-14
コントローラ	
HSZ40 コントローラ	A-14
スケーリング	9-3
ディスク・キャッシュ	5-6
コントローラ方式のキャッシュ	9-24
コンバータ	

コンバータ (続き)	
SCSI OpenVMS Cluster システムにおける	
DWZZx の使用	A-14

サ

再送タイムアウト率	9-18
サイト間リンクの仕様	C-9, C-12
サイト固有のユーティリティ	
ストレージの必要要件	5-3
サテライト	
サーバとの近さ	9-5
定義	8-17, 9-9
サテライト OpenVMS Cluster システム	
高可用性	8-17
サード・パーティ製品	
ストレージの必要要件	5-3
サーバ	
キャパシティ	9-4
サテライトとの近さ	9-5

シ

磁気ディスク	9-23
資源	
共用	1-2
システム	
スケーリング	9-2
システム管理	
AUTOGEN コマンド・プロシージャ	2-5
可用性の保守手法	8-6
事前対策	8-6
手法とツール	C-15
製品	2-5
操作環境	10-7
ソフトウェア	1-7
ツール	2-5, 8-6
マルチサイト OpenVMS Cluster システ	
ム	C-15
容易さ	1-2
システム仕様	3-2
システム通信サービス	
SCS を参照	
システム・ディスク	
OpenVMS Cluster 可用性のためのシャドウイン	
グ	8-4
OpenVMS Integrity システムでの構成	7-48
可用性	5-7
手法	10-2
冗長化	5-7
シングル	10-2, 10-7
シングル対マルチ	10-7
マルチ	10-7
長所	10-3
システム・パラメータ	
EXPECTED_VOTES	9-19, 10-12
LOCKDIRWT	9-19
MPDEV_AFB_INTVL	6-32

システム・パラメータ (続き)	
MPDEV_D1	6-32
MPDEV_ENABLE	6-32
MPDEV_LCRETRIES	6-32
MPDEV_POLLER	6-32
MPDEV_REMOTE	6-32
MSCP_BUFFER	6-46
MSCP_LOAD	9-19
NPAGEDYN	9-19
PAGEDYN	9-19
RECNXINTERVL	9-19
SHADOW_MAX_COPY	9-19
VOTES	9-19
更新	9-5
サテライトを備えた OpenVMS Cluster	
の	9-19
ポーリング	10-13
ポーリング・タイマ	8-6
マルチパス	6-31
システム・ファイル	
管理	10-2
システム・ルート	10-6
シャドウイング	
Volume Shadowing for OpenVMS を参照	
状態遷移	
ガイドライン	10-13
冗長化	
システム・ディスク	5-7
ディスク	5-7
データ	5-7
バックアップ・コピー	5-8
冗長性	
Ethernet セグメント	9-3
Ethernet ブリッジ	9-3
I/O パス	8-3
LAN アダプタ	9-17
インターコネクト	8-3
構成要素	8-2, 8-4
スケーリング	9-3
ブート・サーバ	8-3
ジョブ・コントローラ	
分散	1-4
シングル・エンド・シグナル通知	
SCSI の	A-9
信頼性	
ハードウェア	8-5

ス

スケーラビリティ	
I/O	9-21
IP	9-20
OpenVMS Cluster の設計	2-2
SCSI OpenVMS Cluster システムにおけ	
る	9-5
拡張次元	9-1
サテライトを備えた OpenVMS Cluster システム	
における	9-9

スケーラビリティ (続き)	
実現手法	9-4
処理とストレージ	1-2
定義	9-1
プロセッサのキャパシティ	9-4
要件	2-2
スター・カブラ	
スケーリング	9-3
ストライピング	
ディスク・ストライピングを参照	
ストレージ	
アクセス	9-4
スケーリング	9-3
パフォーマンス・オブティマイザ	5-5
ホスト方式	5-10
ローカル・アダプタ	5-10
ストレージ域拡張ソフトウェア	1-6
ストレージ・コントローラ	
Fibre Channel	
デュアル・ポート	6-25
マルチポート	6-17
ストレージ・デバイス	
ストレージの必要要件を参照	
インターコネクトへの接続	5-2
可用性の必要要件の決定	5-7
設置必要要件	5-3
定義	1-3
必要容量の見積もり	5-3
ストレージの必要要件	
OpenVMS オペレーティング・システム	
の	5-3
アーカイブ	5-4
サイト固有のユーティリティ	5-3
サード・パーティ製品	5-3
増加量	5-4
データベース	5-4
ページ・ファイル, スワップ・ファイル, ダン	
プ・ファイル	5-3
ユーザ作成プログラム	5-3
ユーザ・データ	5-4
ユーザ・レベル	5-4
容量	5-3
レイヤ製品	5-3
ストレージ領域	
共用, 直接アクセス	8-4
スループット	
Ethernet	4-17
MEMORY CHANNEL	4-7
SCSI	4-11

セ

成長の要件	2-2
性能	
データ	
Fibre Channel 構成	7-63

セキュリティ要件	2-3
接地	
SCSI 要件	A-22
トラブルシューティング	A-41
設置面積の必要要件	
ストレージ・デバイス	5-3
接続マネージャ	
OpenVMS Cluster での役割	1-4

ソ

相互依存	
緩和	9-4
ソフトウェア	
可用性の強化手法	8-4, 8-5

タ

帯域幅	
拡張 LAN における	9-18
タイマ	
ポーリング	8-6
ターミナル・サーバ	
可用性	8-3
ターミネータ	
SCSI 用	A-10
単一点障害の要因	8-4
緩和	9-4
サーバとディスク	8-6

ツ

通信サービス	
OpenVMS Cluster システムにおける	
ATM	C-6
OpenVMS Cluster システムにおける	
DS3	C-6

テ

ディザスタ・トレラントな OpenVMS Cluster 構成	8-22
ディスク	
SCSI	A-1
SCSI のアクセス	A-2
概念	A-7
構成上の要件	A-4
可用性	5-7
キャッシュ	5-6, 9-24
データ	
シャドウイング	8-4
パフォーマンス・オブティマイザ	5-5
半導体	5-5
必要量の見積もり	5-3
ページとスワップ	10-6
ディスク・サーバ	
フェールオーバ	8-3

ディスク・ストライピング	
RAID を参照	
ディスク・テクノロジー	9-23
RAID も参照	
ディファレンシャル・シグナル通知	
SCSI の	A-9
データ	
OpenVMS Cluster 環境にアクセス	1-2
可用性	5-7
冗長化	5-7
データベース	
ストレージの必要要件	5-4
データ保管ソリューション	7-7
デバイス ID	
SCSI ディスクも参照	
Fibre Channel	7-15
WWID	7-15, 7-24
ストレージ・アダプタ	7-18
ストレージ・アダプタ名	7-15
ストレージ・デバイス	7-19
テープ・デバイス	7-24
HSZ 割り当てクラス	6-22
SCSI の構成	A-23
ノード割り当てクラス	6-21
ポート割り当てクラス	6-21
マルチバス	
パラレル SCSI	6-20
テープ・デバイス	
Fibre Channel	
命名	7-23
テープ・ライブラリ・ロボット	7-23

ト

ドライバ	
DKDRIVER	1-5
DUDRIVER	1-5, B-12
Ethernet E*driver	1-5
FDDI F*driver	1-5
FG*driver	1-5
MCDRIVER	1-5, B-12
PADRIVER	1-5
PBDRIVER	1-5
PEDRIVER	1-5, B-12
PG*DRIVER	1-5
PIDRIVER	1-5
PK*DRIVER	1-5
PMDRIVER	1-5, B-12, B-15
PNDRIVER	1-5, B-12
TUDRIVER	1-5
ポート	1-5
トラブルシューティング	
SCSI 構成	A-39

ナ

内部バス 5-10

ネ

ネットワークング・ソフトウェア 1-6

ノ

ノード
OpenVMS Cluster システムにおける A-1
ノード・カウント
制限 9-4
ノード割り当てクラス
割り当てクラスを参照

ハ

媒体チェンジャ・デバイス 7-23
バックアップ
可用性のための 5-8
重要性 8-6
冗長化 5-8
パケット・サイズ
拡張 LAN における 9-19
バス
内部 5-10
バス
現在の 6-36, 6-39
ポーリング 6-38
複数の
選択 6-39
優先順位 6-39
プライマリ 6-36
ポーリング 6-37
バス識別子
デバイス ID も参照
Fibre Channel 6-33, 7-19
Fibre Channel マルチバス 6-33
パラレル SCSI 6-33
バス選択
テープ・ドライブのマウント 6-41
バス名
バス識別子を参照
バッチ・キュー
可用性の仕組み 8-3
クラスタ全体での共用可能 1-4
ハードウェア
BA350/BA353 StorageWorks エンクロージャ
 A-12
DWZZx コンバータ A-14
HSZ40 コントローラ A-14
SCSI OpenVMS Cluster システムでの A-11
アドオン SCSI アダプタ A-12
可用性の強化手法 8-4

ハードウェア (続き)

構成要素 1-2
スケーラビリティの手法 9-4
ハードウェア要件
Fibre Channel テープ・デバイス 7-22
パフォーマンス
SCSI ストレージ A-8
アプリケーション 10-3
システム 10-3
データ転送速度 A-9
パフォーマンスの要件 C-12
パラメータ
システム・パラメータを参照
半導体ディスク 5-5, 9-23

ヒ

ビジネス要件
可用性 2-2
決定 2-1
将来の成長率 2-2
スケーラビリティ 2-2
セキュリティ 2-3
物理的な制約 2-3
予算 2-1

フ

ファイル
環境 10-7
システムと環境 10-2
ストレージの必要要件 5-3
ページとスワップ 10-2
ホット 9-25
フェールオーバ
HSx モード
透過的 6-3, 6-11
マルチバス 6-3, 6-12
MSCP サービス対象から直接 SCSI への 6-5
OpenVMS Cluster のメカニズム 8-2
計画 8-5, 8-6
直接 SCSI から MSCP サービス対象へ
の 6-2, 6-5, 6-6
直接 SCSI から直接 SCSI への 6-2, 6-3, 6-6
定義 4-3
透過的 6-3, 6-15
複数のアクセス・バス 5-7
マルチバス 6-3
マルチバス
OpenVMS による実行方法 6-42
フォーリン・デバイス
バスのリセット A-37
負荷分散
定義 4-3
複数サイトの管理 C-15
ブート
LAN 用シーケンス 10-6

ブート (続き)	
マルチ・システム・ディスクの処理時間の短縮	10-3
ブート・サーバ	
サテライト用	10-6
フェールオーバ	8-3
ブート・デバイス・パス	
FC 上での構成	7-50
ブリッジ	
LAN セグメント間の	9-17
制限	9-18
プリント・キュー	
可用性の仕組み	8-3
クラスタ全体での共用可能	1-4
プロセッサ	
可用性	3-2
要件	2-3
プロセッサ設定	
OpenVMS Cluster システム	3-2
プロセスの割り当て量	
増加	10-10
分散ジョブ・コントローラ	
ジョブ・コントローラを参照	

へ

ページ・ディスクとスワップ・ディスク	10-6
ベースバンド・ネットワーク	
Ethernet のインターコネクトを参照	

ホ

保証サポート	10-14
ホスト	
OpenVMS Cluster システムにおける	A-1
ホスト ID	
SCSI の構成	A-22
ホスト方式の RAID	
SCSI OpenVMS Cluster 構成でサポート	A-5
ホスト方式のキャッシュ	9-24
ホスト方式のシャドウイング	
SCSI OpenVMS Cluster 構成でサポート	A-5
ホスト方式のストレージ	5-10
ホット・ファイル	9-21
の処理の解析	9-25
ホット・プラグ (SCSI デバイス)	A-44
ボトルネック	
緩和	9-4
削減	5-6
ポート割り当てクラス	
割り当てクラスを参照	
ボリューム・シャドウイング	
Fibre Channel デバイス	7-7
マルチパス・デバイス	6-10, 7-7

ポーリング・タイマ	
タイマを参照	
ポーリング・パラメータ	10-13

マ

マルチサイト OpenVMS Cluster システム	8-20
マルチパス SCSI サポート	
SCSI 構成を参照	
マルチパス・セット	6-2

メ

メモリ	
サテライトへの追加	9-17
追加	2-4, 9-2
メモリ要件	2-3

ユ

ユーザ作成プログラム	
ストレージの必要要件	5-3
ユーザ・データ	
ストレージの必要要件	5-4

ヨ

読み込み/書き込みの比率	
I/O スループット	9-23

リ

リソース	
可用性について	8-5

ル

ルート	
システム	10-6

レ

レイヤ製品	
ストレージの必要要件	5-3

ロ

ロック・マネージャ	
分散	1-4

ワ

ワイド・エリア・ネットワーク	
WAN を参照	
ワークステーション	
Alpha	9-14, 10-4

割り当てクラス	
HSZ	6-9, 6-22
SCSI 構成の設定	A-6, A-22, A-24
ノード	6-21
ポート	6-21
HSZ 割り当てクラスの	6-9
割り込みコアレス機能	7-61

OpenVMS Cluster 構成ガイド

2010年10月 発行

日本ヒューレット・パカード株式会社

〒102-0076 東京都千代田区五番町7番地

電話 (03)3512-5700 (大代表)

AA-RNJ9C-TE.2

