

データマネジメントを意識した ストレージソリューションWG 検討結果報告書

2008年5月

サイエンティフィック・システム研究会
データマネジメントを意識したストレージソリューションWG

目次

1. はじめに	1
1.1. 背景とその活動内容	1
1.2. 活動メンバー	2
1.3. WG 活動経過	2
1.4. 本報告書の構成	4
2. ストレージソリューションの現状	5
2.1. ストレージソリューションの発展	5
2.2. 非構造型データの認識とデータマネジメント	7
2.3. ストレージソリューションにおけるデータマネジメント	7
2.3.1. データプロテクション	8
2.3.2. アーカイブ・HSM	8
2.3.3. ストレージ管理	8
2.3.4. データアクセス管理	9
2.3.5. デバイス管理	10
2.3.6. データレプリケーション	10
2.3.7. ファイルシステム	10
2.3.8. その他(圧縮、暗号化など)	10
3. データマネジメントの現状とストレージソリューションへの期待	13
3.1. データマネジメントの現状	13
3.1.1. JAXA における数値シミュレーションデータの管理の特徴	13
3.1.2. 天文台における天文観測データの管理の特徴	17
3.1.3. 中京大におけるコンピュータ演習用ストレージシステムの特徴	20
3.1.4. JAIST におけるストレージデータ管理の特徴	24
3.1.5. 理研におけるデータマネジメントの特徴	27
3.2. 調査結果の整理と分析	33
3.2.1. 各機関保有データの特徴	33
3.2.2. 各機関の現システムにおける対応状況	34
3.2.3. 課題の抽出	36
3.3. ストレージソリューションへの期待	39
4. ストレージソリューションの今後	41
4.1. 富士通におけるストレージソリューションとデータマネジメント	41
4.1.1. データプロテクション	41
4.1.2. アーカイブ・HSM	42
4.1.3. ストレージ管理	43
4.1.4. データアクセス管理	44
4.1.5. デバイス管理	44
4.1.6. データレプリケーション	44
4.1.7. ファイルシステム	45
4.1.8. データ暗号化	46
4.2. ストレージソリューションの方向性	46
5. おわりに	48
添付資料	
参考資料	

1. はじめに

1.1. 背景とその活動内容

情報処理性能の飛躍的な向上による科学技術計算の大規模化や、情報処理システムの入力元となる観測機器および測定機器の技術革新による高精度化によって、日々非常に大容量のデータが生み出されている。さらにブロードバンドの普及により Web システムなどにおけるデータのマルチメディア化にも拍車がかかり、これまでとは比較にならないデータ量の増加を招いている。

このような環境の中、大学・研究所においては、大規模なストレージシステムが必要とされており、しかも、日常の教育・研究活動におけるデータの相互活用の重要性が高まっている。この為、これらのストレージシステムは、個々のサーバシステムに閉じて配置された分散形態ではなく、データの相互活用を意識した統合ストレージシステムとして構築されるケースが増えてきている。こうしたストレージシステムの要求要件の変化により、高速かつ高可用にストレージにアクセスできるという基本性能に加えて、如何にそのシステムマネジメントを実現するかについての期待も高まっている。

このような背景から、2001 年度より 2 年間の活動を行なった『ネットワーク時代の統合ストレージマネジメント WG』では、SAN/NAS の技術動向とその融合した形態を調査検討し、その結果として、会員の機関をベースに研究所モデル 2 件/大学モデル 2 件を作成・評価した。この活動では、ハードウェア技術の進歩に焦点を当て、SAN/NAS を利用したシステムの評価を実施したが、その一方、高可用性、運用管理、バックアップ、データ共用、ストレージの仮想化など、運用性向上を目指す「ストレージマネジメント」に関して、さらに深く掘り下げた議論が必要であるという声が強まった。これを受けて、2003 年度より新たに『ストレージを中心としたシステムマネジメント WG』が発足し、2 年間の活動を行った。ここでは、会員機関において、要求される管理・セキュリティ・データ保全・仮想化などのキーワードを基に、システム全体の問題点の洗い出しと分類整理を行い、ハードウェア/ソフトウェアを含めたストレージシステム全体として何が問題か、適切にマネジメントする上で何を考慮しなければならないかについて議論した。また、その結果を、ストレージシステム管理者が留意すべきポリシー項目として整理し、各機関がポリシー項目達成度を評価できるようにするとともに、他機関システムとの比較・参照を可能とすることに取り組み、「ストレージシステムポリシー評価ワークシート」を作成した。

さらに、過去 4 年間の活動を受け、2006 年 2 月より『データマネジメントを意識したストレージソリューション WG』として本 WG の活動を開始した。

本 WG では、大学・研究所における代表的な業務データを洗い出し、分類を行った上で、それらをどのようなアプリケーションやミドルウェアで管理しているか、または管理すべきかを検討するとともに、大学・研究所が目指すデータマネジメントを意識したストレージソリューションの明確化とその対応の検討を行い、備えるべき機能・構成を明確化しその対応を検討した。また、前 WG「ストレージを中心としたシステムマネジメント WG」での成果物であるストレージシステムマネジメントポリシー評価ワークシートについて、ユーザの利便性をさらに追求した「ストレージシステム導入時のチェックシート」として、管理者がより実践的に利用できるよう改良・作成を行い、「ストレージシステム設計・導入にあたってのガイドライン」として冊子にまとめた。本活動は、サブ WG 活動と位置付け、通常の会合とは別に検討活動を行い、2007 年 9 月 11 日に SS 研 HP 上に公開した。なおガイドラインは、本報告書に別冊として添付している。

1.2. 活動メンバー

【会員】	担当幹事	鈴木 富男(理化学研究所)
	まとめ役	磯 直行(中京大学)
		藤田 直行(宇宙航空研究開発機構)
		水本 好彦(国立天文台 天文データセンター)
		松澤 照男(北陸先端科学技術大学院大学)
		長谷川 忍(北陸先端科学技術大学院大学)
		黒川 原佳(理化学研究所)
【富士通】	まとめ役	森屋 光弘(科学ソリューション事業本部科学ソリューション統括部)
		小野 英司(文教ソリューション事業本部)
		吉田 真和(文教ソリューション事業本部文教ソリューション統括部)
		松井 泰敏(文教ソリューション事業本部文教ソリューション統括部)
		服部 和徳(ストレージシステム事業本部ストレージインテグレーション統括部)
		社本 昇(次世代テクニカルコンピューティング開発本部ソフトウェア開発統括部)
		松本 一志(サーバシステム事業本部 Linux 技術開発統括部)
		小林 英一(ソフトウェア事業本部プラットフォームソフトウェア事業部)
		大内 敦夫(ソフトウェア事業本部事業計画統括部)
		竹内 一(富士通北陸システムズ 公共・文教システム部)
オブザーバ		長倉 浩士(ソフトウェア事業本部データマネジメント・ミドルウェア事業部)
		有川 保仁(ストレージシステム事業本部ストレージシステム事業部)
		金子 悟(ストレージシステム事業本部ストレージソリューション事業部)
		坂口 吉生(科学ソリューション事業本部科学ソリューション統括部)
		林 良隆(富士通北陸システムズ)
		澤田石 稔(富士通北陸システムズ 公共・文教システム部)
		長田 昭夫(富士通北陸システムズ 公共・文教システム部)
		平沢 健一(マーケティング本部カスタマーリレーション部)

1.3. WG 活動経過

本 WG では、まず、大学・研究機関のデータ特性の分析、および R&D 関係の構造化データ/非構造型データの観点から整理・検討を行い、業務指向ストレージソリューションの明確化を行った。次に、大学・研究所向けストレージソリューションにとって将来あるべき姿に向けた改善項目の検討および備えるべき機能・構成を明確化しその対応を検討し、その結果を本報告書とした。各会合では、富士通より情報提供を行い、最新の技術情報の調査と議論を行った。

また、前 WG「ストレージを中心としたシステムマネージメント WG」での成果物であるストレージシステムマネージメントポリシー評価ワークシートについて、ユーザの利便性をさらに追求した「ストレージシステム導入時のチェックシート」として、管理者がより実践的に利用できるよう改良し、『ストレージシステム設計・導入にあたってのガイドライン』としてまとめた。ガイドライン作成については、サブ WG 活動とし、通常の会合とは別に検討活動を行い、2007 年 9 月 11 日に、SS 研 HP に公開した。

(<http://www.sskn.gr.jp/wg/str/guideline2007/contents.html>)

各開催 WG での概要は、表 1-1、および表 1-2 に示す。

表 1-1 WG 会合の概要

	日時	場所	活動内容
第 1 回	2006 年 4 月 6 日(木) 14:00~17:40	富士通本社	・今後の WG 活動の進め方/内容検討 ・<技術情報> 新ストレージ・プラットフォーム ETERNUS8000、ETERNUS4000
第 2 回	2006 年 7 月 6 日(木) 14:30~19:00	富士通本社	・構造化/非構造型データの整理 ・ポリシー評価シートのエンハンスの検討 ・<技術情報> ストレージ新市場への展開
第 3 回	2006 年 10 月 12 日(木) 14:30~19:00	富士通本社	・会員各機関のデータ特性の整理と今後 ・ポリシー評価シートのエンハンスの検討 ・<技術情報> Shunsaku
第 4 回	2007 年 1 月 26 日(金) 14:00~17:40	富士通 北陸営業本部	・データ特性の整理のまとめ(1) ・サブ WG 活動状況報告 ・次年度の WG 活動の進め方の検討 ・<技術情報> ETERNUS 性能監視
第 5 回	2007 年 4 月 6 日(金) 14:30~18:00	富士通本社	・データ特性の整理のまとめ(2) ・サブ WG 活動状況報告 ・2007 年度活動の検討 ・<技術情報> 科学技術向け次期ファイルシステム 紹介
第 6 回	2007 年 6 月 6 日(水) 14:30~17:40	富士通 東海支社	・データ特性の整理のまとめ(3) ・サブ WG 活動状況報告 ・<技術情報> ETERNUS NR1000V NAS ゲートウェイ紹介 ETERNUS 2000 ディスクアレイ紹介
第 7 回	2007 年 10 月 24 日(水) 14:30~17:40	富士通本社	・成果報告書 作成の検討 ・<技術情報> フィールドイノベーション 富士通のストレージシステム ETERNUS と関連ソリューション 紹介
第 8 回	2008 年 1 月 11 日(金) 14:00~18:00	富士通 北陸営業本部	・成果報告書のレビュー①
第 9 回	2008 年 3 月 24 日(月) 14:30~17:40	富士通本社	・成果報告書のレビュー②

表 1-2 サブ WG 会合の概要

	日時	場所	活動内容
第 1 回	2006 年 8 月 22 日(火) 10:00~17:30	富士通本社	・ポリシー評価シートのエンハンス検討
第 2 回	2006 年 11 月 9 日(木) 13:30~17:30	富士通本社	・チェックシートのレビュー ・今後のスケジュール ・公開方法の検討
第 3 回	2007 年 8 月 27 日(月) 14:00~19:00	富士通本社	・ガイドラインのレビュー
	2007 年 9 月 11 日(火) 13:15~19:30	富士通本社	システム技術/研究教育環境分科会 共催 2007 年度第 1 回会合「教育現場への ICT 活用」 ストレージシステム設計・導入にあたってのガイドライン 発行 成果報告 & Web 公開

1.4. 本報告書の構成

本報告書の構成について、概要を記述する。

2章では、データマネジメントを意識した視点から現状のストレージシステムの課題・要望が分析できるように、ストレージソリューションを構成する代表的な8つのカテゴリについて解説する。

3章では、本WG参加メンバーの所属機関におけるデータマネジメントの現状を調査し、その結果を基に今後求められるストレージソリューションへの期待をまとめている。また、調査に使用した「データの整理分類のためのワークシート」は添付資料としている。

4章では、3章において導出されたストレージソリューションへの期待を受け、各カテゴリにおける富士通の状況について紹介し、ストレージソリューションの技術分野とその今後の方向性について展望する。

また、WG 会合時に富士通から情報提供された、ストレージ関連の技術資料を参考資料として、掲載している。

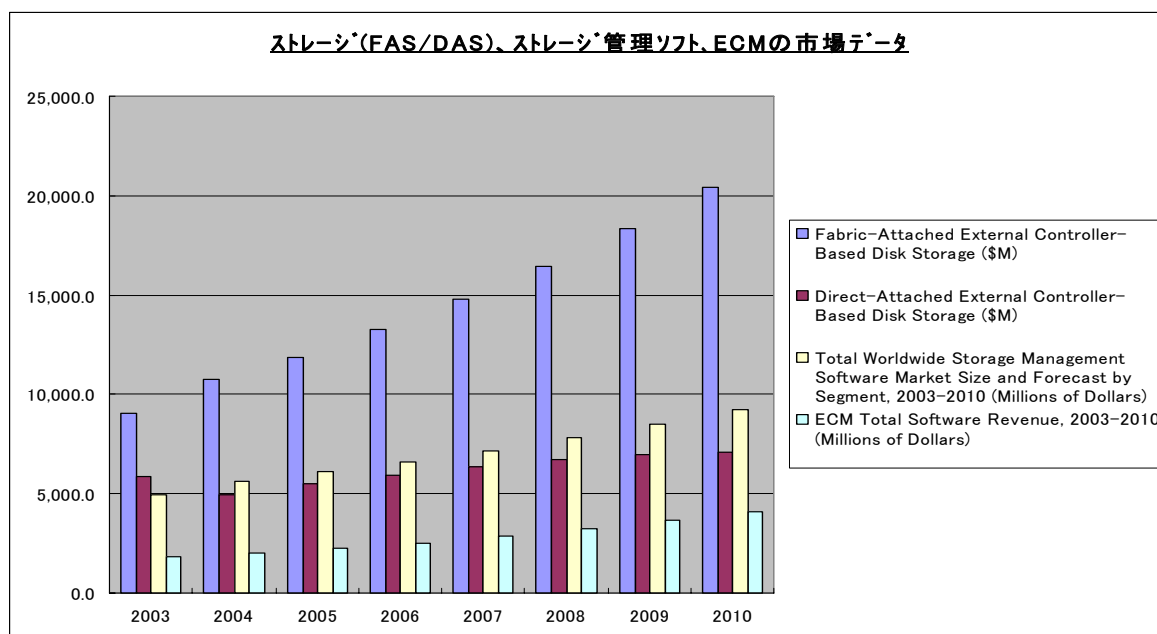
2. ストレージソリューションの現状

本章では、データマネジメントを意識した視点から現状のストレージシステムの課題・要望が分析できるように、ストレージソリューションを構成する代表的な 8 つのカテゴリについて解説する。

2.1. ストレージソリューションの発展

2006 年度版 Gartner Dataquest レポート「ストレージ市場動向」を参考に、今後の需要予測を含めたトレンドを図 2-1 に示した。そこから読み取れる市場動向を以下に要約する。

- ・ストレージ装置の出荷(需要)が DAS¹から SAN²、NAS³、CAS⁴を含めたネットワーク型ストレージシステム(FAS⁵)にシフトしている。
- ・サーバに直接接続する DAS の CAGR⁶(2005-2010)は 5.1%にとどまるが、一方、FAS として認識される SAN、NAS などのネットワークストレージの市場は CAGR(2005-2010)=11.5%で伸長すると予測されている。
- ・ネットワークストレージ市場の拡大はストレージマネジメントソフトウェアの需要との相関が見受けられる。
- ・ストレージマネジメントソフトウェアの内容としては、HSM⁷、アーカイブ分野の伸長が 30%と大きく、また、ストレージ管理と関連する情報管理ミドルウェアである ECM も相関的に拡大している。



※ECM(Enterprise Content Management):データ管理ミドルウェアの一分野。企業規模での情報管理を支援する仕組みとして、ファイル管理、レコード管理、プロセス管理などが統合されている。

図 2-1:ストレージ市場のトレンド(Gartner 社 2006 年度版レポートを基に富士通で作図)

上記では、ストレージ業界全体のトレンドについて概観したが、以下に特徴的な技術動向の

¹ DAS: Direct-Attached External Controller-Based Disk Storage は Gartner の分類名。サーバ直結のストレージの意。

² SAN: Storage Area Network

³ NAS: Network Attached Storage

⁴ CAS: Content Aware Storage の略。画像・映像・音声・ドキュメント・ログ・メールなどの変更や更新をしないフィクストコンテンツを最適に長期間保管するためのストレージ装置。

⁵ FAS: Fabric-Attached External Controller Based Disk Storage は Gartner の分類名。

⁶ CAGR: Compound Average Growth Rate

⁷ HSM: Hierarchical Storage Management

例として EMC 社の M&A⁸による技術統合について考察する。

2003 年の EMC 社による Documentum 社の買収は、同社のソフトウェア化戦略を象徴している。同社から公表された技術系会社の M&A は、図 2-2 に示すように 20 社に及んでいる。ストレージの品揃え(SAN、NAS、CAS)の完成⁹後は、バックアップソフトウェア、コンテンツ管理ソフトウェア、仮想化ソフトウェアまでを統合し、EMC 社はデータマネジメントに必要な技術幅広く取り揃えることに成功している。

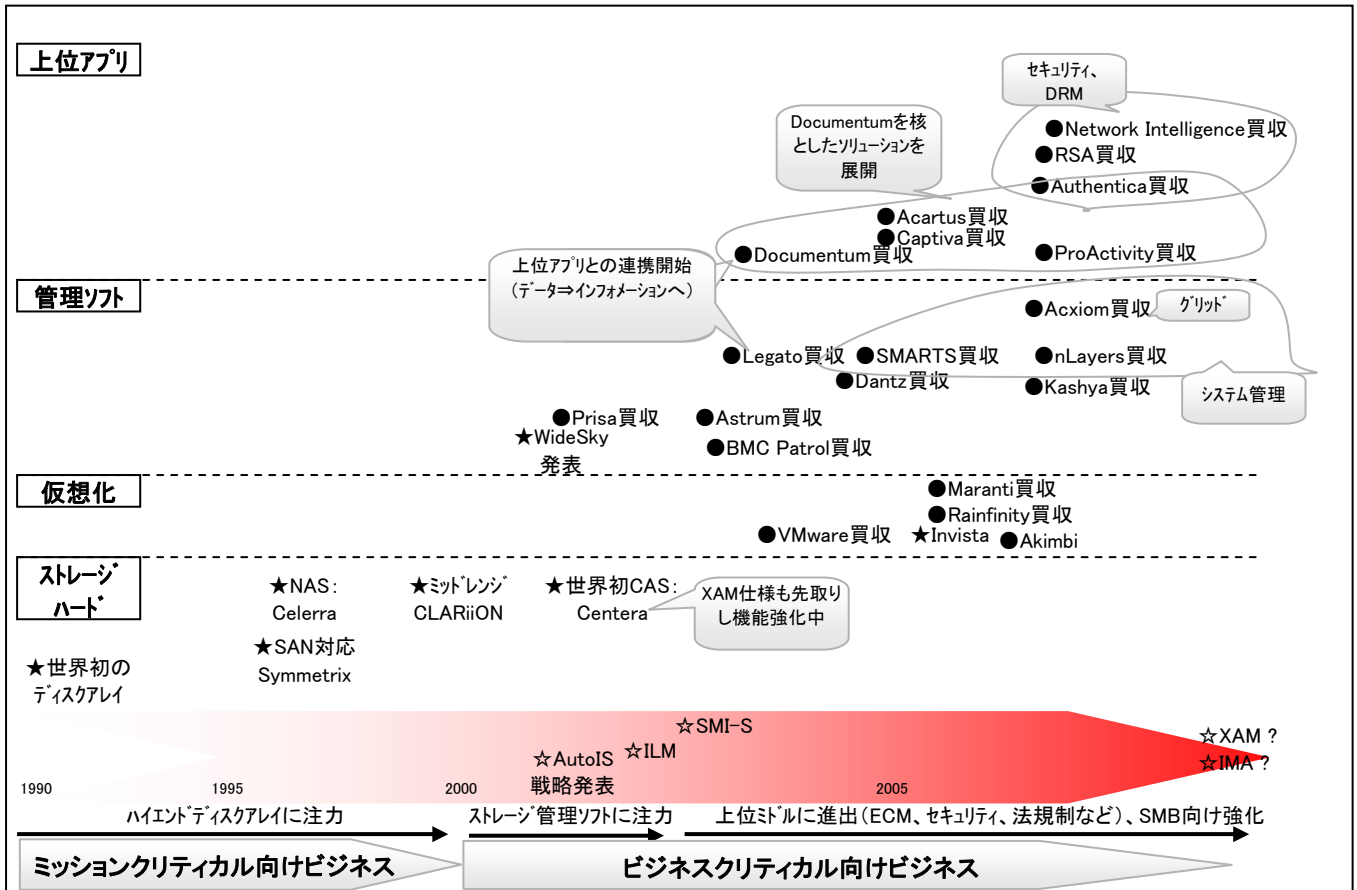


図 2-2: EMC 社におけるストレージ関連技術の獲得状況

また、同社以外にも 2006 年の IBM 社による FileNet 社の買収や Oracle 社による Stellent 社の買収も EMC 社と共通点がある以下のような戦略に基づく、企業行動であると考えられる。

- ・顧客のストレージソリューションへの期待が装置機能や性能からストレージに格納するデータをマネジメントする方法、ツール、容易性、管理対象規模に関する期待に変化しており、これを顧客要件として取り組むことを可能にするソフトウェア技術力の確保が重要であるとの判断がある。
- ・データマネジメントに関するソフトウェア技術は、ストレージ管理ソフトウェア、バックアップ管理ソフトウェア、データ管理ミドルウェア、およびデータセキュリティ機能として認識できる。

このように、一部のストレージベンダは、ハードウェアシステムのみならず、データマネジメントから経営に直結する情報マネジメントまで総合的なソリューションを自社製ソフトウェアと M&A を通じて獲得した製品・技術によって提供してきている。さらにこれらのベンダは、個別の顧客要件に対しても、製品をコアにしたソリューションやシステム構築サービスを一貫して提供できる能力を持っている。

⁸ M&A: Mergers and Acquisitions 企業の合併・買収

⁹ ストレージインタフェースの拡大も企業買収によって実施。

2.2. 非構造型データの認識とデータマネジメント

データマネジメントについてのユーザ要件を具体的に理解するために、多くのストレージベンダが言及している「非構造型データ」について WG を通じて議論を深めた。この非構造型データの定義および議論の中から抽出したデータマネジメントの特徴について、以下に示す。

- ・構造化データの対義語として非構造型データという類型的な呼称を使用した。非構造化データという呼称もあるが、「非構造化(テキスト)データ処理」の意味で使われることもあり、ストレージソリューションにおけるデータマネジメントの共通要件を抽出する目的を明確にしつつ混同を防止するため「非構造型データ」と呼称した。
- ・構造化データは RDB などに格納された「データベース」を指す。データベースは情報を分解したデータをレコードとして管理し、各レコードにはデータ処理を目的とするフィールドの設定やインデックスのような内部構造を持たせている。データベースのデータ処理は、レコードの更新、検索、ソートなどの形式的な変形処理を指す。一方、データベースのテーブル管理やテーブルを格納するストレージ空間や装置の管理機能はデータ保全を目的とした機能として区別できる。
- ・共通構造を持たないという意味で構造化されていないデータ群を「非構造型データ」と呼称したが、実際はサーバにおけるオペレーティングシステムが提供するファイルシステムに格納されるファイルと同一視できる。
- ・オープンシステムによる分散処理のデータ管理はファイルを基本単位にしており、ファイルサーバの普及とあいまって、ファイル群を効率よく管理するデータマネジメント手法が必要となっている。
- ・PCなどの普及によって、ファイル形式で保存されるデータは著しく増加している。また企業や組織の業務にとって重要なデータがファイルとして保存される傾向にある。
- ・ファイルとして保存されるデータは特定のアプリケーションソフトウェアにおいて処理するものであり、複数のファイルを群として扱うこともある。たとえば、メールサーバは非常に大量のメールデータをファイルとして処理しており、データベース処理に匹敵する変形処理と保全処理機能を持っている。
- ・論理的にはアプリケーションソフトウェアとファイルが揃えばユーザは処理が実行できるので、サーバ、ストレージなどのプラットフォームに依存しない仮想的処理環境が必要となるだろう。
- ・大学・研究所が持っている学術データや高精細画像など長期保存を目的としたデータアーカイブでは、データ保存だけでは不十分で、将来においてデータを活用するための処理系の保存も重要である。
- ・アプリケーションからみて一つの業務単位を構成するファイル群をまとめて管理できるコンテンツ管理というミドルウェアの考え方は学術データの管理に応用できる可能性がある。

2.3. ストレージソリューションにおけるデータマネジメント

2.1.節に紹介したように、ストレージベンダはディスク装置の提供から、コンピュータのオペレーティングシステムが担ってきたデータ管理の機能、ネットワーク管理の機能、およびこれらの技術をストレージソリューションとして揃えている。

このストレージソリューションの変化を捉えることは、既に 1 章で述べた以下の本 WG の活動趣旨と合致している。

「大学・研究所における代表的な業務データを洗い出し、分類を行った上で、それらをどのようなアプリケーションやミドルウェアで管理しているか、または管理すべきかを検討するとともに、大学・研究所が目指すデータマネジメントを意識したストレージソリューションの明確化とその対応を検討する。」¹⁰

¹⁰ 「データマネジメントを意識したストレージソリューション WG,2007 年度活動計画」より <http://www.sskn.gr.jp/wg/str/plan2007.html>

大学・研究所における代表的な業務データの整理・分析には、IDC¹¹が定義している以下の 8 つのストレージソリューションのカテゴリを採用した。

- (1) データプロテクション
- (2) アーカイブ・HSM(階層ストレージ管理)
- (3) ストレージ管理
- (4) データアクセス管理(ストレージインフラストラクチャ)
- (5) デバイス管理
- (6) データレプリケーション
- (7) ファイルシステム
- (8) その他(圧縮、暗号化など)

2.3.1. データプロテクション

ストレージソリューションにおけるデータマネジメントの重要目的の一つは格納されたデータがそのままの状態でも参照できることを保障することであり、これをデータプロテクションと呼ぶ。

データプロテクションは、ストレージシステムの大容量化、および格納されるデータ特性や種類の多様化により、ディスク全体のバックアップから始まり、個別データのアーカイブによる選択的保護、さらに更新されたデータのみを自動検出して保護するまでに発展している。

データプロテクションの機能は、RAID¹²による実装からデータバックアップソフトウェア、リモートコピー、アーカイブなど、ストレージソリューションを構成する様々な機能レイヤで実現している。なかでも、データバックアップソフトウェアは早期から整備され、対象とするデータは、データベースやファイルシステムだけでなく、メールボックスやリポジトリなどへ拡張されている。

これらの技術は、保護すべきデータの選択、データ保護レベルの客観化、保護処理プロセスの自動化とカスタマイズなど、IT システムにおけるデータ処理量の拡大と運用管理の複雑さに対応する IT インフラ管理技術の方向性と一致している。

2.3.2. アーカイブ・HSM

アーカイブと HSM は異なる技術であるが、データマネジメントの観点からは同一視できる技術である。

バックアップがデータの復元手続きを前提とするデータ保護技術であることに比べ、アーカイブは処理手続き後のデータ保護技術である。アーカイブ処理手続きは、保護機能が特別に整備されている領域にデータを格納することでデータ保護される特徴を持つ。

HSM はファイルシステム空間の拡張を実現する技術である。この技術は、拡張された空間に対して透過的ファイルアクセスを実現させる特徴を持ち、アーカイブと区別することができる。

最近では、ポリシーベースのアーカイブという半自動化したアーカイブ技術が注目され、透過的ファイルアクセスを実現させる HSM とアーカイブのユーザメリットが近接してきている。

2.3.3. ストレージ管理

大規模化、複雑化するストレージシステムをサーバに接続し、ファイルサーバやリポジトリサーバとしてアクセス可能にするための論理的な記憶空間(ストレージ空間やファイルシステム空間)、アクセスパス、可用性レベルなどのサービスを提供する仕組みが、ストレージ管理で

¹¹ International Data Corporation 社

¹² RAID: Redundant Arrays of Inexpensive (もしくは Independent) Disks の略。ハードディスクなどの記憶装置を複数台用いて、アクセスを分散させることにより、高速、大容量で信頼性の高いディスク装置を実現するための技術。

ある(図 2-3)。

ストレージシステムの管理ソフトウェアは、各ストレージベンダから提供されているが、SNIA¹³が提唱する SMI-S¹⁴によってストレージベンダ各社のストレージを同一インタフェースで制御できるような標準化技術も成長してきている。

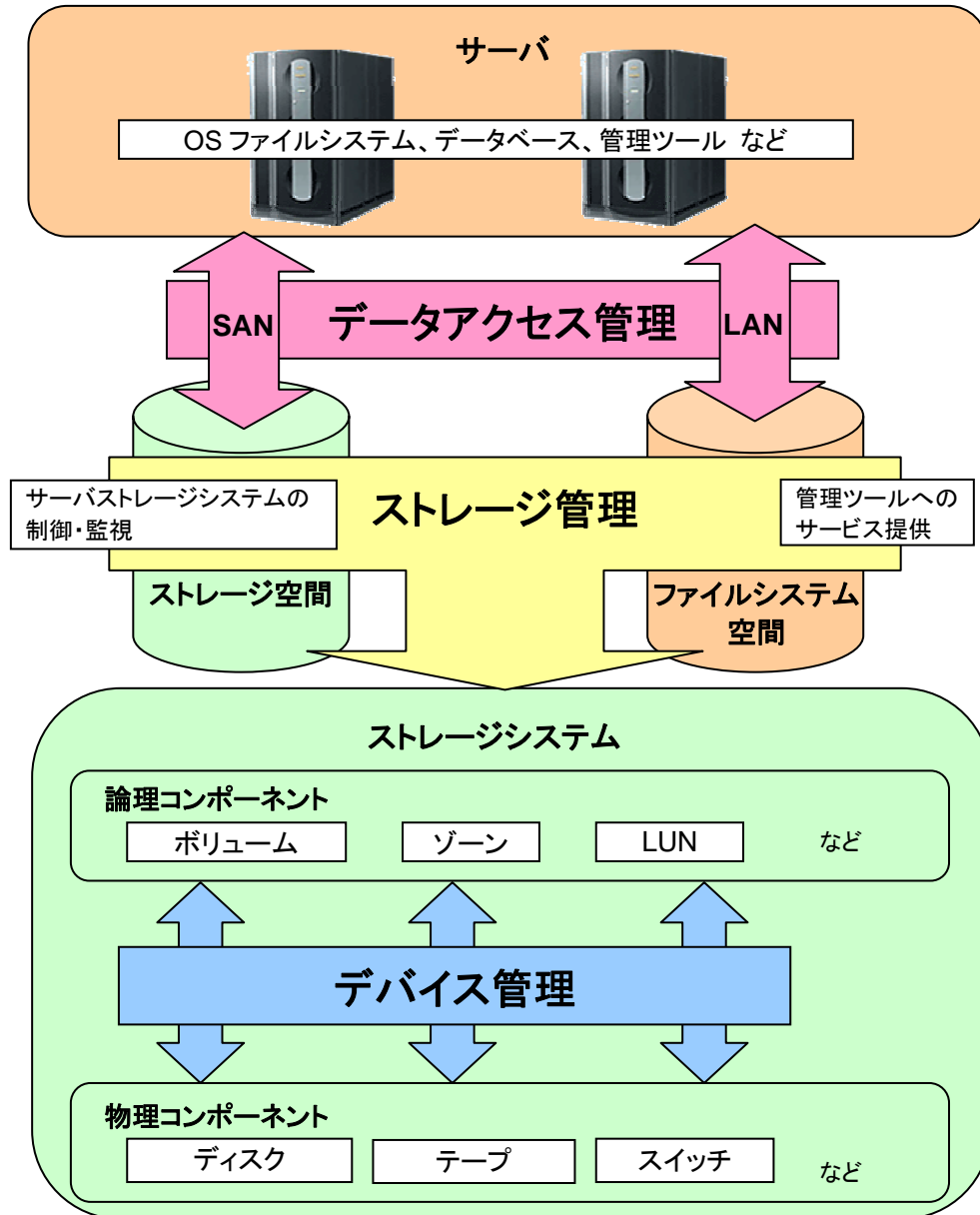


図 2-3:ストレージ管理、データアクセス管理、デバイス管理

2.3.4. データアクセス管理

図 2-3 に示すように、代表的なデータアクセス管理に対応するソフトウェアは、SAN におけるマルチパスドライバと LAN における NFS¹⁵/CIFS¹⁶制御プログラムである。いずれも可用性維

¹³ SNIA: Storage Networking Industry Association

¹⁴ SMI-S: Storage Management Initiative Specification

¹⁵ NFS: Network File System の略。TCP/IP 環境で一般的に使用されるネットワーク経由のファイルサービス。

¹⁶ CIFS: Common Internet File System の略。SMB をベースとした、インターネット上のリモートファイルシステムにアクセスするためのプロトコル。

持および性能確保のために複数のネットワークパスを制御し、ストレージやファイルサーバの記憶空間とアプリケーションを処理するサーバとの接続を制御するものである。QoS¹⁷によるデータ転送帯域保障やサービスレベルの保障がデータアクセス管理に求められる重要な要件となる。

2.3.5. デバイス管理

保存データが大量になるとこれらを格納するストレージシステムを大規模化させるか、広域に分散化させるなどして大容量化する必要がある。どのようなアーキテクチャでストレージシステムを大容量化したとしても、ストレージシステムは個々のデバイスであるディスクドライブ、コントローラ、ネットワークで構成されることになる。

複雑化するストレージ関連装置を構成しているモジュールの動作状況を監視し、システムとしての可用性を維持するためには、構成定義の変更や装置の追加、交換、さらにはデータ移行などを容易にするための仕組みが必要であり、これらの機能をデバイス管理と呼ぶ(図 2-3)。

2.3.6. データレプリケーション

データを複製(レプリケーション)する目的は二つある。

第一の目的はデータアクセスを容易にし、処理性能を向上させるために、データを必要とするシステムに近接させることである。例えば、データウェアハウスにおけるデータ複製がこれにあたる。

第二の目的はデータの保護レベルを向上させることである。2001年9月11日に起こった同時多発テロの実例が有名である。世界貿易センタービル崩壊によって失われたデータは、あらかじめ定期的に複製されていたことにより無事復元された。このようにITの災害対策ソリューションにはデータレプリケーションを含む場合が多い。

なお、ストレージソリューションの観点から言えば、サーバや業務の状況によらずデータを複製することで、第一のメリットが活かせる場合が多いために、データレプリケーションが活用されてきた経緯がある。

2.3.7. ファイルシステム

ファイルシステムは、ストレージ記憶空間をアプリケーションソフトウェアから POSIX などの標準インタフェースでアクセスさせるためのインタフェースソフトウェアである。通常はサーバ OS の重要なコンポーネントとして提供されるが、ファイルシステムを内蔵しているファイルサーバも導入が容易で需要が高い。

ファイルシステムの技術によって、ストレージシステムにおける可用性、拡張性、高速性を統合している。これまでのファイルシステム技術は、複数のボリュームを仮想化技術で束ねるボリューム管理ソフトウェアなどと連携してより高い可用性、拡張性、高速性を追及してきた。さらに、NFS や CIFS に代表されるネットワーク共有技術により、データアクセス管理における利便性を向上してきた。

2.3.8. その他(圧縮、暗号化など)

データマネジメントを情報マネジメントとして拡大して捉えると、上記 7 カテゴリーのほかにもデータマネジメントレベルで実装できる技術がいくつかある。

たとえばデータ圧縮である。ファイルの同一性検査の仕組みを発展させ、ファイル内の同一

¹⁷ QoS: Quality of Service

ブロックを検出し、これらを共用ブロックとして管理すれば、ストレージに保存されているデータ量を圧縮することが可能となる。この仕組みは、近年 De-dup 技術¹⁸として注目されている。

また、情報のアクセス権限管理と暗号化機能を組み合わせてデータアクセス管理に適用すれば、ストレージのセキュリティ管理機能を強化することができる。

次に、ストレージソリューションにおけるデータ暗号化は、データの記憶媒体をストレージシステムから持ち出す可能性に対して、不用意なアクセスを防止する情報漏洩対策機能として実装されている。データの暗号化は、ネットワーク管理やファイル管理において実装されることが一般的であり、一般ユーザのプライバシーとシステム管理者などによる内部監査や情報統制までに対応する必要がある技術分野である。

¹⁸ De-dup 技術: データ圧縮技術の一つ。近年、増大する非構造型データの中にはファイルの重複、データの部分的な重複が多く含まれていることに注目して、重複部分を排除して保存する技術を導入した製品が市場投入されてきた。組み込みソフトウェアやアプライアンスとして提供されることが多い。

3. データマネジメントの現状とストレージソリューションへの期待

本章では、大学・研究所におけるデータマネジメントの現状を調査し、その結果を基に今後求められるストレージソリューションについてまとめる。なお、調査対象は本 WG 参加メンバーが所属機関において、担当または管理しているシステムおよびデータとした。

調査対象システムおよびデータを以下に記述する。

- ・ 宇宙航空研究開発機構(以下、JAXA と呼ぶ):
NSIII¹⁹およびメールサーバ
- ・ 国立天文台(以下、天文台と呼ぶ):
ハワイ観測所および三鷹本部にあるすばる望遠鏡のためのデータアーカイブと解析処理のためのシステム
- ・ 中京大学(以下、中京大と呼ぶ):
学内の理系学部(情報理工学部・生命システム工学部)の教育研究用コンピュータ演習室システムとメールサーバおよび公開 Web サーバ
- ・ 北陸先端科学技術大学院大学(以下、JAIST と呼ぶ):
情報科学センターおよび遠隔教育研究センターの管理するシステム
- ・ 理化学研究所(以下、理研と呼ぶ):
情報基盤センターが管理するシステムのデータ領域における研究系データ

3.1. データマネジメントの現状

データマネジメントの現状を認識するため、まず本 WG 参加メンバー所属機関が保有しているデータの特性を添付資料の「データの整理・分類のためのワークシート」に記述した。ワークシートの項目は、保有データのデータ種類とその説明、データ管理の特性として構造／非構造、発生頻度、保存期間、管理者、管理ソフトウェア、年間容量等とした。また、データ発生から消去までの管理方法と補足説明のために備考を加えた。

さらに、データマネジメントの状況を知るため、会員機関ごとに調査した主要データの特性、使用ストレージシステムの要件／特徴等をまとめ、課題を検討した。

なお、課題は 2.3 節で述べた以下の 8 カテゴリに基づいて検討した。

- (1) データプロテクション
- (2) アーカイブ・HSM
- (3) ストレージ管理
- (4) データアクセス管理
- (5) デバイス管理
- (6) データレプリケーション
- (7) ファイルシステム
- (8) その他(圧縮、暗号化など)

以下に、会員機関ごとの調査結果を記す。

3.1.1. JAXA における数値シミュレーションデータの管理の特徴

3.1.1.1. はじめに

JAXA では、ワークシートに例示したように、数値シミュレーションデータ・実験データ・事務データ(Web コンテンツ、電子メール)と、幅広いデータが蓄積されている。ここでは、数値シミュレーションデータに絞ってシステム概要と特徴を述べるとともに現状の課題について考察する。

¹⁹ NSIII: Numerical Simulator III

3.1.1.2. 数値シミュレーションにおけるストレージシステム

計算サーバの高速化に伴い、大規模数値シミュレーションが行われるようになってきた。複雑形状の航空機全機まわりのナビエ・ストークス解析計算、乱流の直接シミュレーション、燃焼化学反応流の計算等がその例である。計算サーバはこれらの数値シミュレーションを現実的に処理できるまでにその CPU 性能が向上してきた。一方で、計算過程における I/O 処理が数値シミュレーション全体のボトルネックとなっている。現有の JAXA スーパーコンピュータシステムでは、この I/O 処理のボトルネックを解決できるシステム構築を行っている。図 3-1 に JAXA スーパーコンピュータシステムの概要を I/O 処理を司るストレージシステムを中心に図示する。

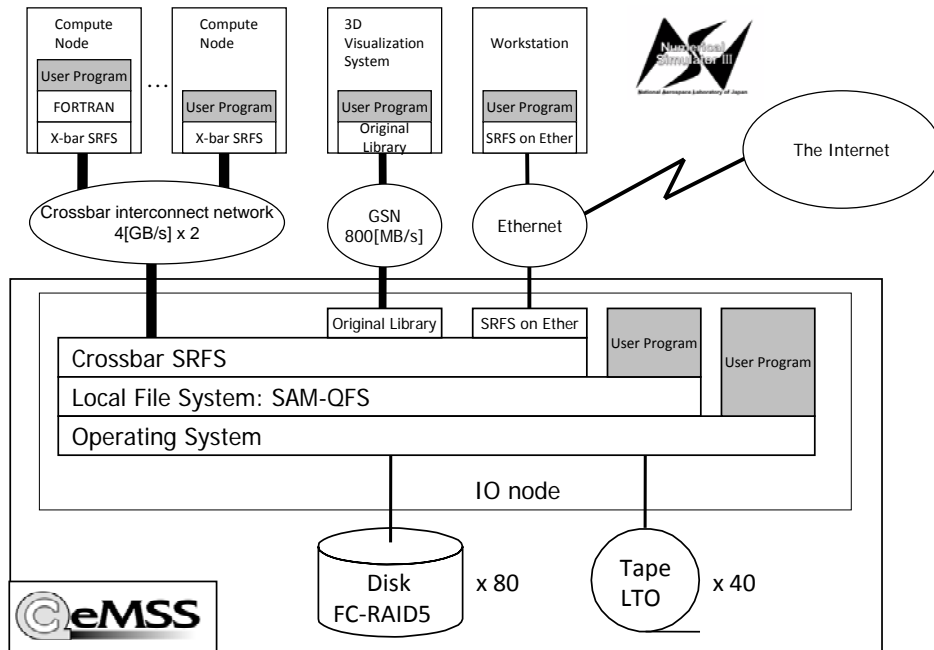


図 3-1: JAXA スーパーコンピュータシステムにおけるストレージ

JAXA スーパーコンピュータシステムは、NSⅢと呼ばれているが、ストレージシステム部を特に、CeMSS²⁰と呼んで区別している。CeMSS はデータセントリックという考え方にに基づき、NSⅢのストレージシステムとして構成され、クロスバネットワークで直結されている計算サーバ(CeNSS²¹)から、インターネット上に存在するクライアントまで広範なクライアントに対してストレージサービスを提供している。計算サーバの FORTRAN プログラムから実効 8Gbps、三次元可視化サーバ(CeViS²²)から 4Gbps、構内ネットから NFS より数倍速い速度で、さらには、実験・開発的にインターネット上に存在するクライアントからアクセス可能なファイルシステムの構築を行っている。図 3-1 に示す”User Program”とある位置から測定したストレージシステムの性能評価の詳細については参考文献²³をご覧ください。

ファイルの管理は、階層ストレージ管理(HSM)システム SAM-FS により行っている。高速にアクセスできるストレージとして RAID5 ディスク装置が 50TB あり、その後ろに 600TB のテープ空間が一元的に管理されている。ディスク装置は大規模 I/O にチューニングされており、ストライプ技術により 8Gbps の実効性能を実現している。また、50TB のディスクから溢れ出してしまったファイルは HSM システムにより 600TB のテープ空間へ自動的に移動され、必要な時に自動的に取り出される。また、テープ装置の速度(非圧縮時ピーク性能 120Mbps)がデ

²⁰ CeMSS: Central Mass Storage System

²¹ CeNSS : Central Numerical Simulator System

²² CeViS : Central Visualization System

²³ 【参考文献】”Storage Devices, Local File System and Crossbar Network File System Characteristics, and 1 Terabyte File IO Benchmark on ”Numerical Simulator III”,” Proc. IEEE/NASA Mass Storage Systems and Technologies, pp.72-76

ディスク装置の速度(実効性能約 8Gbps)より遅いため、ディスク～テープ間のデータ移動においては、テープドライブの速度がボトルネックにならないように、テープ媒体 4 個への同時書き込み処理(ストライプ)を行うことにより最大 480Mbps のテープへの I/O を行える機能を備えている。

3.1.1.3. 大規模数値シミュレーションにおけるストレージシステムの特徴

大規模数値シミュレーションのストレージシステムに必要なものは、第一に大容量と高速性、第二に広範な利用環境の提供である。表 3-1 に CeMSS の特徴とその実現手段をまとめた。大容量と高速性は既にその実現方法を述べたので、ここでは、広範な利用環境の実現について述べておく。

表 3-1:大規模数値シミュレーション用ストレージシステムの特徴

特徴	実現手段
大容量	HSM によるテープ媒体を用いた 600TB のファイルスペース
高速性	RAID5 のストライプによる実効 8Gbps の I/O 性能
広範な利用環境	<ul style="list-style-type: none"> ・スーパーコンピュータ用インターコネクティブインターフェイス ・GSN インタフェース ・インターネットにインタフェースできるミドルウェア SRFS on Ether インタフェース

NSⅢのファイルシステムは、当初スーパーコンピュータ用インターコネクティブであるクロスバネットワーク上でのみ実現されていた。ITBL²⁴ 等、インターネットを用いた仮想研究開発組織の活動には、インターネット上で利用可能な遠隔ファイルシステムの実現が不可欠であった。また、CeViS とのデータ交換においても高速なファイル共有の仕組みが必要であった。これら二つの要求を満たすために、Ethernet(図 3-1 の"SRFS on Ether")と GSN²⁵ (同"Original Library")のインタフェースを NSⅢのファイルシステムに追加した。

また、管理面・セキュリティ面の必要性から、ユーザ ID・グループ ID(以下、まとめて ID と呼ぶ)のマッピング機能を実装した(図 3-2)。これは、サーバにアクセスしてくるクライアントが使用している ID を、サーバ管理者の責任の下で、サーバで管理している ID にマッピングしてしまう機能であり、ファイルシステム上での簡便なアクセスコントロールであると考えられる。サーバ管理者は、サーバ上での ID のみを意識して、ファイルシステムのアクセス権限を管理することができる。NSⅢのような中央集中型のシステムの場合、GSI²⁶のような複雑な仕組みに替わって、この ID マッピングのようなシンプルな方法でユーザの一元管理を実現することも考えられる。

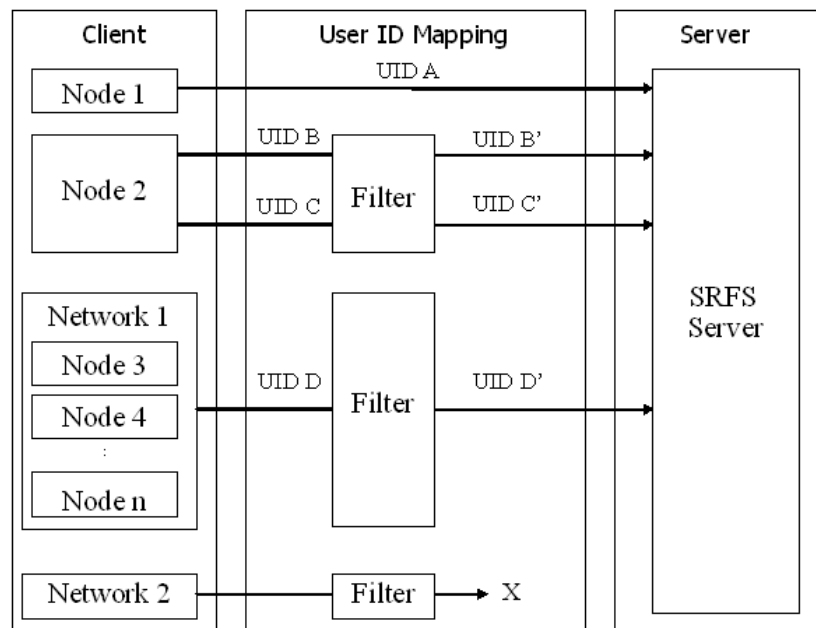


図 3-2 CeMSS におけるファイルシステムの UID、GID マッピング機能概要

²⁴ ITBL: IT-Based Laboratory

²⁵ GSN : Gigabyte System Network

²⁶ GSI : Globus Security Infrastructure

3.1.1.4. 現状の課題

ここでは 2 章で定義されたデータマネジメントにおける 8 カテゴリに則って、NSⅢの現状の課題を述べる。

以前の数値シミュレーションにおいては、“データは計算し直せばよい”という考え方があり、保護すべきはソースコードとジョブ投入コマンド群である、との認識が少なからずあった。しかし、最近の数値シミュレーションの大規模化や、非定常計算への関心の高まりにより、多くの CPU タイムを使用して得たシミュレーション結果ファイルや、急速に増加している非定常解析用データファイル群が現れ、“データは保管するもの”という考え方が広がってきた。NSⅢでは大規模なストレージシステムを整備し、この要求に応えてきた。

(1) データプロテクション

以前のシステムに比べれば、HSM を大規模に導入しデータプロテクションの機能は向上したと考えられる。バックアップソフトウェアではなく HSM によりデータプロテクション機能を実現しているため、ユーザは今まで通りのファイル管理を行うだけで、データプロテクションのための操作を新たに覚える必要はない。システムが自動的にバックアップとリストアを行っている。具体的には、ディスクに書き込まれたファイルを一定のポリシーに従ってテープ媒体 2 本にコピーする。一定のポリシーに従ってディスクから削除されたファイルを使用する場合は、システムが自動的にテープ媒体にコピーしたファイルをディスクにコピーしユーザが利用可能な状態にする。一つのテープ媒体が破損していた場合は、もう一方のテープ媒体を使用するようになっている。従来のバックアップ/リカバリの考え方や操作とは異なる方法ではあるが、このような方法でデータプロテクションを実現している。

この運用で大きな課題は発生していないが、あえて課題を挙げると次のようなものがある。

- (a) 全てのファイルを分け隔てなくバックアップしてしまうため、バックアップ不要なファイルまでバックアップしてしまっている可能性がある。
- (b) ユーザはこれから使おうとしているファイルがテープ媒体上に存在することを意識していない場合があり、このとき、ファイル I/O 開始までの待ち時間が通常の場合と比べて非常に長時間となってしまう。現在はバッチジョブのスケジューリングの仕組みを工夫して、この問題が発生しないようにしている。

(2) アーカイブ・HSM

NSⅢの場合、データプロテクション、アーカイブ、HSM の区別が明確ではないが、上記(1)で示したシステムの形態で運用している。

HSM の便利さに気が始めたユーザからは、スーパーコンピュータシステム内だけでなく、構内ネットワークやさらにはインターネットからもその機能を使えないかという声が始まり始めており、遠隔ファイルシステム等と連携したデータプロテクションシステムの実現が望まれる。また、ストレージシステムが広く利用されるようになると、その構築コストにも目が行く。現在は比較的高価な RAID 装置を用いているが、装置自体の低価格化が望まれる。また、HSM ソフトウェアのライセンス料も課題の一つである。HSM 管理される容量に対して従量課金するライセンス形態の場合、大規模なシステムの構築に支障が生じる可能性があるため、定量課金ライセンス形態の方がシステム構築し易い場合が多い。

(3) ストレージ管理

ストレージ管理での課題の一つは、ストレージハードウェア故障時の交換対応である。単体ディスクが故障した場合は RAID 構成となっていれば運用を停止することなく復旧可能であるが、RAID コントローラや複数故障が発生した場合、運用を停止する必要がある。大容量が進めばこの課題は今まで以上に顕在化してくると思われる。

また、ストレージ換装時のデータ移行が次の課題として挙げられる。ディスク装置だけでペタバイトのシステムを持てるようになった現在では、今までの手法だとデータ移行に長期間を要してしまう。オーガニックストレージ等大容量のストレージを長期間にわたり連続的に運用管理できる仕組みの構築が望まれる。

(4) データアクセス管理

データアクセス管理で一番の課題は、データ共有である。スーパーコンピュータのみならず、他のコンピュータからも、スーパーコンピュータにあるファイルにアクセスしたいという要求である。上記(2)でも述べたが、このアクセス範囲は、構内ネットワークに留まらず、外部の組織との連携を考え、インターネット越しのアクセス要求も出てきている。また、JAXA の次期スーパーコンピュータシステムでは、JAXA 内の複数の事業所からのスーパーコンピュータ利用が想定されており、この実現に向けた具体的なシステム構築が望まれている。

次の課題としては、データアクセス権の細かい制御である。現状でも拡張アクセスコントロール機能を利用すれば実現できるものであるが、広く利用されていない現状を考えると、その利用には課題がいくつかありそうである。具体的な要求としては、メインフレームで実現しているような、個人単位のアクセス許可や、個人が複数のグループに属した場合の分かりやすいアクセス制限のかけ方のようなものである。

(5) デバイス管理

スーパーコンピュータシステムの場合、一度システムを構築してしまったら、ストレージシステムの構成変更はあまり発生しない。一方で、大規模なシステム構築をすることから、日々の動作状況監視が重要となる。今後のシステム設計のための参考資料としても、この日々の動作監視が重要であると考えられる。

(6) データレプリケーション

NSⅢでは、データレプリケーション機能は実現していない。スーパーコンピュータの役割として、スーパーコンピュータでなければできない数値シミュレーションの実行を主眼としてきているためであると考えられる。今後も、この方針は変わらないと考えられるが、ユーザニーズの多様化や、遠隔地事業所からの利用等の観点で、データレプリケーション機能が必要になってくる可能性がある。

(7) ファイルシステム

ファイルシステムに要求されるものは、高速性と共有性である。高 FLOPS 化する計算エンジンの入出力に耐えられるファイルシステム、遠隔地事業所からのスーパーコンピュータ利用や共同開発体でファイル管理を実現する遠隔共有性の実現が望まれる。

(8) その他(圧縮、暗号化など)

NSⅢでは、データ圧縮や暗号化は行っていない。今後、多岐にわたるユーザが単一システムであるスーパーコンピュータを利用する場合、暗号化は利用者から要求される可能性がある。具体的には、システム内でのデータ漏洩防止やデータをシステム外に持ち出す場合のセキュリティの観点からファイルの暗号化が必要になる可能性がある。また、データ量が増大する中、ストレージの媒体総量を圧縮するためにデータ圧縮技術を利用することが期待される。

3.1.2. 天文台における天文観測データの管理の特徴

3.1.2.1. はじめに

天文台では米国ハワイ島のすばる望遠鏡を用いて、2000 年 12 月の共同利用観測開始以降、最遠方銀河の発見や銀河系形成過程の解明などに大きな成果をあげている。天文観測は原理的に再観測が不可能であるため、天文観測データの長期保存は学問上極めて重要である。コンピュータシステムには原本の保障やユーザへのデータ配布に関していくつかの工夫がなされている。ここでは、すばる望遠鏡で取得した観測データの生成から保存・提供までのコンピュータシステムの特徴を紹介するとともに現状の課題について考察する。

3.1.2.2. 天文観測データの管理とその流れ

システムは図 3-3 に示すように、ハワイ島マウナケア山頂、ハワイ観測所山麓施設、および三鷹本部の 3 拠点から構成されている。すばる望遠鏡は標高 4,200m のマウナケア山頂に設置

されている。すばる望遠鏡には現在 7 つの観測装置がある。観測装置によって取得された天体観測データは望遠鏡の指向方向や時刻などの情報を加えて、FITS²⁷規約に則ったフォーマットのファイルに即座に変換され、その管理情報が ORACLE データベースに登録される。同時に観測データの保全のためテープライブラリにバックアップを行う(データ①)。即ち、山頂システムのテープ媒体をオリジナルデータ(原本管理)としている。

さらに、これらの観測データは専用回線によって直ちに山麓施設に送られる。山麓システムにあるすばる観測データアーカイブシステム(STARS²⁸)は、受け取った観測データファイルのヘッダの内容やデータサイズなどの確認後、FITS ヘッダから抽出した情報と管理情報をデータベースに登録する。また、後述するデータの整合性を保障するために MD5²⁹(ここで算出した MD5 値を MD5s と呼称、以下同じ)を算出・登録する(データ②)。ハワイに居る研究者は直接 STARS を通して観測データを検索し、データ要求を行う。STARS はヒットしたデータをクライアント側に転送する。この時、観測データの原本保証性を高めるため依頼のあったデータを読み込む際、再度 MD5(MD5u)を算出し、MD5s=MD5u の場合のみユーザ側への転送を行うようにしている。なお、STARS を構成しているアーカイブシステムは現在、SFG³⁰連携の磁気テープライブラリ Sony Peta Site/Serve を採用している。

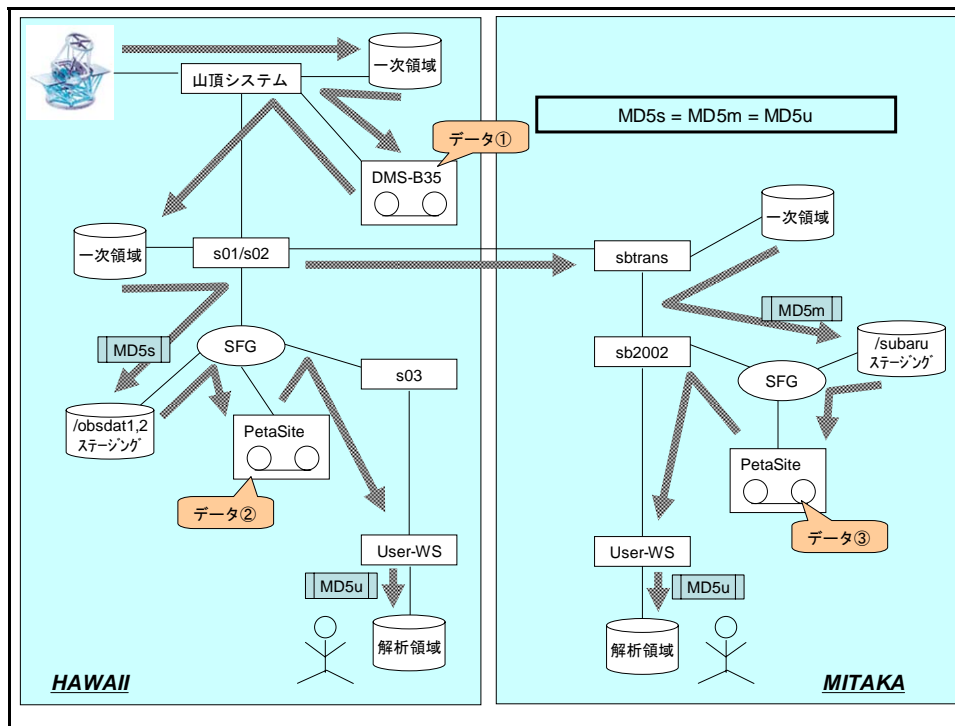


図 3-3:天文台 天文観測データシステム

一方、日本に居る研究者は三鷹の MASTARS³¹を介して容易にすばる望遠鏡で取得された観測データを検索できる。MASTARS も STARS と同様に Peta Site/Serve から構成されている。STARS 側に観測データの追加があると、その観測データは専用回線を通してハワイから三鷹へも転送される。三鷹に送られた観測データは STARS と同様な仕組みで MASTARS に登録されると共に、ここでも MD5(MD5m)が算出され DB にその値が登録される(データ③)。

このように、すばる望遠鏡のデータアーカイブシステムは三つの独立したシステムがデータ

²⁷ FITS: The Flexible Image Transport System

²⁸ STARS: Subaru Telescope ARchive System

²⁹ MD5: Message Digest Algorithm5 :与えられた入力に対して 128 ビットのハッシュ値を出力するメッセージダイジェストアルゴリズム。R.Rivest, "The MD5 Message-Digest Algorithm", RFC1321, April 1992.

³⁰ SFG: SafeFILE/Global

³¹ MASTARS: Mitaka Advanced STARS

のバケツリレー方式で連携している。通常は山頂で観測データが取得されると数分後には三鷹のMASTARSからそのデータを取得することができるリアルタイム系を構成している。一方、通信回線の故障時にはそれぞれが独立したオフライン系として機能する設計となっている。

3.1.2.3. 天文観測データの保全と整合性の保障

すばる望遠鏡の観測は前世紀末から始まったが、如何に高性能の望遠鏡でも地上望遠鏡の宿命として観測時の気象条件の影響を強く受けるため同じ観測は二度と出来ない。また、超新星爆発のような時間的に急激に変化する現象もある。つまり同じ観測データは二度と撮れないことから、観測データを逸することは計算機のダウン以上の衝撃となる。このため、国立天文台では他の多くのシステムで用いているように、基本は原本のバックアップ(上記データ①)によってデータ保障性を確認しつつ、さらに災害・天災時や経年劣化やコンピュータシステムの不具合および、人間の誤操作を考慮してハワイと日本での観測データの二重持ちを行っている(上記データ②とデータ③)。つまり言葉を変えれば、ディザスタリカバリを実現していることになる。これは一見過剰投資なようであるが、観測データの保全・活用が最大の目的であることから、限られた予算・資源の中でCPU性能を犠牲にしてもデータストレージを最重要資源とする自然な選択である。

ところが、「二重持ち」にしているとハワイ側で持っている観測データと日本側で持っている観測データの同一性保証が問題となる。最新のSTARSおよびMASTARSでは上述したMD5をそれぞれ独立に算出してDBに登録する機能が追加されている。併せて、両者の値を定期的にチェックする仕組み、即ちMD5sとMD5mの整合性を監視するように改良されている。加えて、研究者に観測データを配布する際にも同様に読み出したタイミングでMD5を算出してデータの保障性を高めている。

3.1.2.4. 現状の課題

ここでは2章で定義されたデータマネジメントにおける8カテゴリに則って、現状の課題を述べる。特にデータの移行を容易にするための仕組み(以下(5)項)、アーカイブデータおよびバックアップデータの識別・同一性保障(以下(6)項および(8)項)が重要な課題であると認識している。

(1) データプロテクション

天文観測は既に述べているように原理的に再観測が不可能であることに加えて、観測装置の進歩によりデータが質・量共に巨大化している。このため、データの保護は最重要課題である。データ保護の基本はバックアップであるが、巨大化した量のバックアップ採取時間の短縮と確実なリカバリ手段の提供を期待している。また、データ量が巨大化するとこれまでのバックアップとは異なるデータプロテクション方式の開発が望まれる。

(2) アーカイブ・HSM

現在のSTARS/MASTARSは主にコストを抑えるために透過的アクセス方法を提供しているHSMシステムを採用している。最下層のテープ媒体をバックアップと低価格大容量記憶として利用することにより大容量のデータアーカイブを比較的 low cost で実現している。しかし、レスポンスが一定でないことや磁気テープとディスク間のデータコピー時の信頼性への疑問などのデメリットも多い。本来は高速なランダムアクセスが可能なディスクベースのデータアーカイブが理想である。安価で信頼性の高いストレージ装置の実現を期待する。

(3) ストレージ管理

天文観測データは一つの観測目的に対して複数の観測データが対応する(つまり複数のファイルが出力される)ことから複数ファイルをワンセットのコンテンツとして管理することが求められる。ファイルの集合を論理的に管理し可用性レベルを向上させるサービスを提供する仕組みが必要である。

(4) データアクセス管理

負荷分散とディザスタリカバリのために地理的に離れた複数のデータアーカイブを連携して運用している。それらのデータアーカイブを参照するため、利用者にファイルシステムを意識させず、接続方式が違って互いに干渉することなく一定のアクセススピードを保証するデータアクセス管理が必要である。

(5) デバイス管理

天文観測データは原則永久保存であり、データ量は常に増加するため、ストレージの容量の増強や別のストレージ装置へのデータの移行が必要である。また、現状では計算機システムのリプレースに伴い、全観測データの移行が数年に一度発生し、データ移行に数ヶ月の時間を要する。データの引越しを容易にするストレージ装置やバックアップ装置(その媒体)、およびその仕組みとサービスの提供が必要である。

(6) データレプリケーション

(a) 天文分野ではアーカイブデータの二次利用が盛んである。大量の天文観測データの複製が世界中で作られている。デジタルデータは複製時に劣化しないという常識はどこまで通用するのだろうか。文書データとは違い、天文観測データはホワイトノイズに近いため複製ミスを人間が発見することは困難である。生物が DNA の複製ミスを修復する機能を進化させたように、デジタルデータの複製方式も進化させる必要があるだろう。

(b) これらの複製が地理的に分散すると、何処に何があるか把握できない状態に陥る。さらにそれぞれの複製に独自の付加情報が加わるとデータ自身が独自に進化し始めてしまう。従来のバージョン管理方式では困難な、オリジナルは何処の何か等の識別がつくようなミドルウェアが必要になってくるのではないか。さらに、データ自身に所有者を管理させるとか、データ自身に自己消滅機能を持たすとか、従来とは逆の思想の新たな方式が必要であろう。

(7) ファイルシステム

現在では、OS やベンダによって様々なファイルシステムが提供されているが、用途に応じて使い分けするのではなくユーザインタフェースが統一化されており、高性能・高可用性のものが一般的に必要であると考えている。

(8) その他(圧縮、暗号化など)

データレプリケーション技術の範疇かも知れないが、地理的に分散された(アーカイブされた)データとかバックアップされているデータ(媒体とかに)の同一性保障をどのようにするのが重要である。STARS/MASTARS では本文で述べているがアプリケーションによってデータの生成や移動、提供のタイミングでMD5値を算出し、データの同一性の保証を行っている。これでは維持管理するコストも嵩み、ユーザの負担が大きい。データの同一性を保証する仕組みがどこかのレイヤーが必要である。

3.1.3. 中京大におけるコンピュータ演習用ストレージシステムの特徴

3.1.3.1. はじめに

高等教育機関としての大学では、初学年で全員が共通の基礎教育を受け、その後専門科目の教育を受ける。そのため、コンピュータ演習室を使った初学年の必修講義・演習科目では、多くの学生が一斉にストレージシステムを利用する。近年では一般社会がコンピュータネットワークの利用を前提とする場面が多くなり、学生もコミュニケーションツールとして学内のコンピュータを使用する機会が増えている。そのため、初学年のリテラシ教育だけでなく、最終学年の卒業論文作成、そして就職活動までさまざまな場面でコンピュータ演習室が利用され、ストレージシステムにも多種のデータが蓄積・保存されるようになってきた。

ここでは、私立大学における教育利用ストレージシステムの例として、中京大理系学部で稼動しているコンピュータ演習室用ストレージシステムを紹介するとともに、現状の課題について考察する。

3.1.3.2. コンピュータ演習室の利用

学生にとってコンピュータ演習室は、講義・演習の場であるとともに、講義終了後はレポート作成や、コミュニケーションをとる場にもなる。最近では、紙媒体ではなくファイルとしてレポートを提出することが多くなっている。それは教員にとっても学生にとってもネットワークにつながっている限りどこからでも課題提示、レポート提出が可能になったからである。また、学生課、教務課等の事務部門から学生への呼び出しや休講連絡も、個人情報保護と伝達の確実性を目指して昔ながらの掲示板からポータルサイト等の電子媒体へ移行しつつある。学生はコンピュータ演習室やコンピュータが置いてある自習室を普段から利用し、情報収集・保存に努めることが多くなっている。そして、学生が扱う各種データはすべてコンピュータ演習室用ストレージシステムへ保存することになる。

3.1.3.3. システムに求められる要件

以下では、大学におけるコンピュータ演習室用ストレージシステムに必要な要件について述べる。

1. 高可用性

私立大学では技術職員が少なく、講義や演習に関するシステムを教員が管理することが多く、機器が不調になると担当教員は講義どころでなくなる。従って、高可用性なストレージシステムが求められる。

2. 高速性

コンピュータ演習室で行われる演習系科目では、全員がログイン完了するまで講義を始められない。そのため、講義開始時の一斉ログイン時の待ち時間の短縮が求められる。その逆に、遅くとも休憩時間内には学生全員が教室移動できるように、講義時間内にファイルをストレージシステムにすべて保存し終えなければならない。

3. 大容量ファイルへの対応

最近では映像データのようなマルチメディアデータを利用する講義が増え、それらを高速に扱えることが求められる。講義中は作業用として一人当たり数 G バイトのファイルを扱うことを許可し、その領域への読込、保存が高速であることが必要である。

4. 学年進行に伴うユーザの多様化

教育用システムは学年進行とともに用途が変化する。低学年のリテラシ教育は Windows 主体であるが、高学年の講義や研究室(ゼミ)活動では Macintosh や UNIX(Linux、FreeBSD、Solaris 等)が用いられ、これらのプロトコルにも対応する必要がある。

5. データの完全保証

教育用システムでは、レポート提出等の直接成績に関わる部分について特に慎重な取扱いを必要とする。例えば、ストレージシステムを使ってレポート提出が行われるが、その提出の有無によって学生の成績が左右されることがある。よって、「講義中に教員が課題を提示」でき、「学生がいつでもレポートを提出」でき、そして「教員が提出されたレポートを確実に受理」できることを保証しなければならない。機器故障による提出レポート紛失等の不測の事態は絶対に避けなければならない。そのため、高速かつ確実なバックアップシステムも要求される。また、提出されたレポートファイルは、他の学生からは閲覧できない強固なセキュリティやアクセス制御も必要である。

3.1.3.4. システム概要

上記の要件を満足させるために、以下のシステムを導入し、運用している。

ファイルサーバ : NR1000 F500

ディスク容量(物理) : 3.6TB (=144GB/10,000rpm×25 本)

論理容量(RAID 構成後) : 約 1.7TB

ホームディレクトリとして学生一人当たり 130GB の利用制限を実施し、1 学年約 600 名の学生に対応。その他の領域は課題提示・レポート提出用として利用。

メモリ容量 : 3GB

ネットワーク : 10/100/1000BASE-T×4、10/100BASE-TX×1
 ファイル共有 : NFS、CIFS
 バックアップ装置 : LT130(ドライブ数 : 1、テープ巻数 : LTO3×30 巻)
 バックアップソフトウェア : NetBackup

DNS、メール、Web サーバ等の各種サーバ群と、学生用 Linux サーバ、演習室端末がこのファイルサーバを共有している。アクセスする端末の OS は、Windows、Linux、Solaris、Macintosh(CIFS 使用)がある。

ネットワークトラフィックを分散させる目的で、1000BASE-T×4 のネットワークインタフェースを、演習室外接続用一つと、三つの端末サブネット(赤、緑、青)に接続している。このようにすることで、3 サブネットが敷設された各コンピュータ演習室ではネットワーク障害発生時でも他の回線に接続された端末を使用することであわてずなんとか講義を続けることができる。端末本体の前面カラーとネットワークの色を合わせているので管理者側としても障害時の切り分けが容易である。

バックアップ装置は LTO ライブラリ装置をファイルサーバに直接接続し、NetBackup を用いて、テープ媒体へのバックアップを行っている。

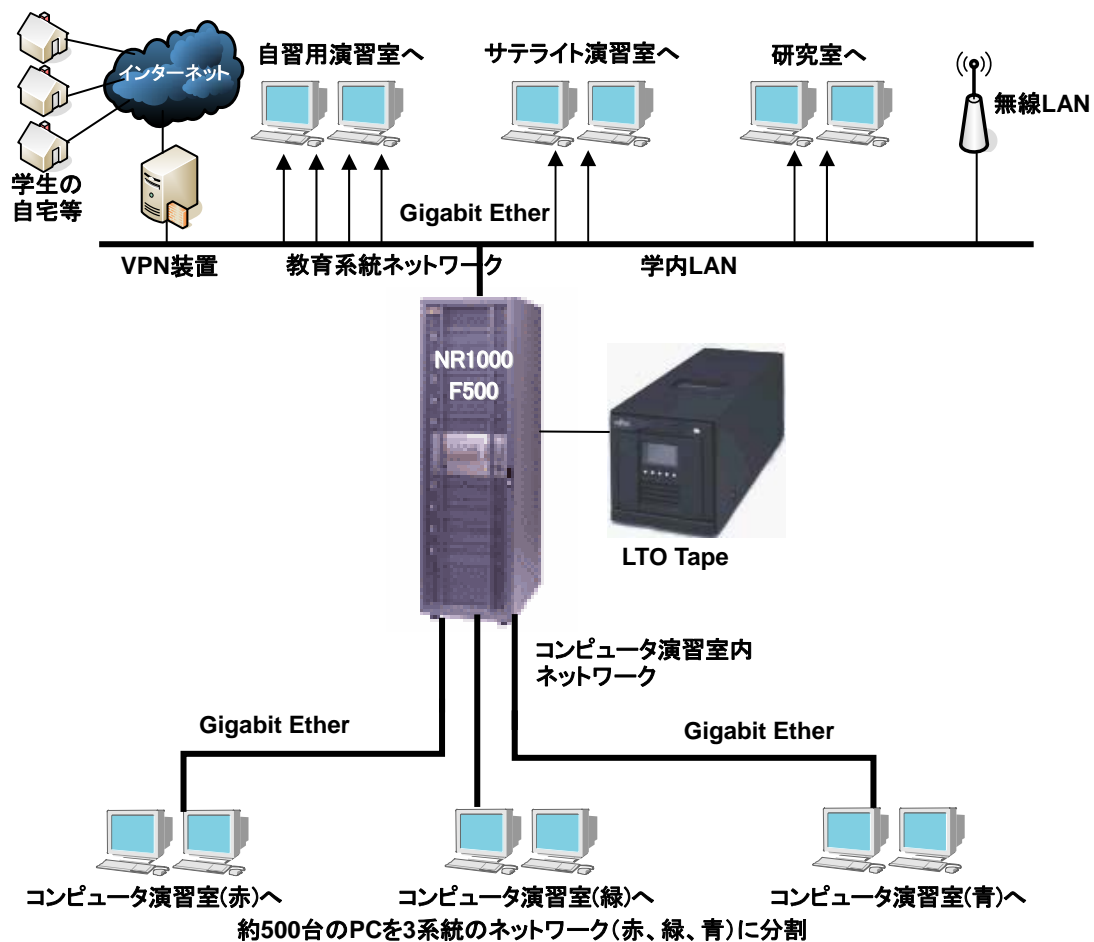


図 3-4: 中京大学情報理工学部のストレージシステム構成

3.1.3.5. 現状の課題

ここでは、2章で定義されたデータマネジメントにおける8カテゴリに則って、中京大理系学部コンピュータ演習室で運用されているストレージシステムの現状の課題について述べる。なお、大学のコンピュータ演習室の主なユーザは学生であるが、通常の講義だけでなく課題提出や卒業研究等でも利用されるため、ユーザとして教員も考慮しなければならない。

大学のコンピュータ演習室用ストレージシステムでは、特に(1)と(4)が重要な課題と考えられる。

(1) データプロテクション

大学では、ユーザの多くが学生であり、ほぼ4年で入れ替わっていく。在学期間中の4年間は課題提出や卒業論文作成等で学生はファイルサーバを頻繁に利用するため、基本的に毎日バックアップを行っている。学生が卒業した後は、保存用バックアップを取った後、ストレージシステム上から削除される。近年、卒業後に必要なファイルを閲覧したいという要望が寄せられることが多くなり、その都度手動でテンポラリ領域へリカバリしファイルを探索しているが、大きなシステムを必要とせずテープから必要なファイルを容易に探索できると良い。また、1年間のスパンで考えてみると、講義期間中(春学期、秋学期)は多くの利用があるが、休暇期間中(特に夏季、春季休暇)は、研究用途を除いてストレージシステムはほとんど利用されない。このことから、バックアップの頻度や範囲を運用に合わせて調整することが可能である。これまでは人的な操作により調整しているが、システムの稼動状況を監視しながらバックアップ頻度・範囲の調整を自動化できると良い。

(2) アーカイブ・HSM

前述の(1)データプロテクションにおける課題を実現するために、アーカイブ、HSMが可能なシステムを導入することも検討されているが、現状のストレージシステムでは、導入および運用コストの都合で導入されていないのが現状である。

(3) ストレージ管理

講義期間中は多くの講義で利用され、特に、期末に集中する課題提出については学生の成績に影響するため、ユーザが期限内に確実に提示・提出(指定領域に保存・移動)できることを保証するストレージシステムが必要である。そのため、ストレージシステムには高可用性が求められる。一方、休暇期間中は自習用設備を除いてコンピュータ演習室を閉鎖し比較的自由にメンテナンス作業することができる。RAIDの構成変更等、停止時間を長く必要とする作業についてはこの期間に実施することができ、コンピュータ演習室におけるストレージ管理運用についてはあまり意識していない。

(4) データアクセス管理

中京大のコンピュータ演習室では、主に各種サーバ向けの NFS と、端末 PC 向けの CIFS によりストレージシステムを開放している。コンピュータ演習室内では学生が一斉に同一課題に取り組むが、端末 PC から CIFS でストレージシステムへアクセスする場合と、Linux 等のアプリケーションサーバを介して NFS でファイルサーバへアクセスする場合があります。各ユーザ領域のアクセス制御は重要な課題である。それはストレージシステムで提供される NFS と CIFS でアクセス制御に関して微妙に違いがあり、一般ユーザがそれを認識できないからである。例えば NR1000F では、CIFS に対して NFS のアクセス制御が暗黙的に上書きするため、Windows 等から CIFS でアクセスするユーザから、アクセスできるはずのファイルにアクセスできないという事態がよく発生する。

(5) デバイス管理

大学におけるコンピュータ演習室では、前述の(3)ストレージ管理で述べたように、システムの大規模化、複雑化に対応するためのリプレース、構成変更等の大きな作業は休暇期間中に実施している。他の企業や研究所等と異なり、ユーザへの影響をあまり考えずに作業することができる。ただし、最近では卒業見込み学生の就職活動早期化等の影響で、春期休暇中の利用希望が増えており、年間通してのシステム提供を考えると、今後短時間のリプレース作業およびデータ移行ができるシステムの導入を検討しなければならない状況になっている。

(6) データレプリケーション

中京大のコンピュータ演習室では、データ複製等のデータレプリケーションは行っていない。キャンパスが複数あり、学生がキャンパス間を行き来するような大学のコンピュータ演習室ではデータレプリケーションは必要な機能と思われる。例えば、常に各キャンパスのストレージシステムが

互いに複製を持つことで端末 PC との距離が近くなり、アクセス速度が向上するとともに、ディザスタリカバリ等が自然に実現できる可能性があるからである。

(7) ファイルシステム

ストレージシステムをユーザ側から見たときにユーザ側で認識できるものはファイルシステムであり、その動作・運用には高可用性、高速性が要求される。また、コンピュータ演習室のような講義で使用する場合にはファイル共有による課題提示・提出がストレージシステム上で行われるため、ファイル・ディレクトリへのアクセス制御がユーザにより容易に設定できることが必要である。

(8) その他(圧縮、暗号化など)

中京大のコンピュータ演習室では、ストレージシステムに圧縮や暗号化は行っていない。ストレージシステムに保存されているデータについて、以前はほとんどが共通課題に対するレポートであったが、最近は電子メール等の私的なファイルが保存されることが増え、それらの多くは個人情報であるため暗号化等による保護も検討しなければならない。しかし大学の場合、ユーザの多くは学生であり 4 年以上の長期にわたって保存し続けることは稀であることと、教育と言う立場から教員が学生を指導するという関係、そしてファイルへのアクセス速度のことを考えると導入に踏み切れないのが現状である。

3.1.4. JAIST におけるストレージデータ管理の特徴

3.1.4.1. はじめに

JAIST のストレージシステムは、特徴的な機能を持つ複数のファイルサーバシステムから構成されており、毎年いずれかのファイルサーバシステムの更新を実施している。ここでは、大学における研究・教育・事務処理等の活動を対象とするストレージシステムの一例として、教職員・学生・研究員のホーム領域、研究プロジェクトのテンポラリ領域等のデータを対象としたストレージサービスの概要と JAIST におけるストレージシステム全体に関する現状および将来的な課題について考察する。

3.1.4.2. 取り扱うデータの概要

JAIST では研究・教育・事務処理等の活動において生成・利用されるデータを、情報科学センターが提供するストレージシステムで一元的に管理することにより、ユーザのストレージに対する管理コストを軽減するとともに、広帯域な学内ネットワーク環境と組み合わせることで、ユーザが学内のどこにいても自身の環境で作業を行えるロケーションフリーなサービスを実現している。

こうした教職員・学生・研究員をはじめとする大学構成員の日々の活動に対するストレージシステムのサービス適用範囲は多岐に渡る。例えば、教職員や学生の Unix ワークステーション・Windows ターミナルサーバから共有されるホーム領域では、設定ファイルやドキュメント、メールデータ、マルチメディアファイルなど日々生成・利用されるデータのリード/ライトが頻繁に発生している。さらに、これらの重要なデータを保護するための定期的なバックアップも必須の機能となる。一方、表 3-2 に示すように、取り扱うデータの種類によっては、上記サービスとは性格の異なるストレージサービスが要求されることも少なくない。例えば、ある研究プロジェクトのテンポラリ領域として一時的に非常に大きなストレージ領域を提供するケースや、大規模計算機のワーク領域としてバックアップは行わないものの高速なライト性能が要求されるケース、ワークステーションやターミナルサーバのアプリケーション領域としてほぼリードのみの利用に対する高速性が要求されるケースなどが存在している。

このような状況の下で、センターがサービスするストレージ領域は年々、大容量・高性能・高信頼性が要求されるようになってきている。しかしながら、すべての要件を同時に満たすストレージシステムを実現するコストは膨大なものであり、現実的ではない。このため、ユーザの利

用目的に対して適切な容量・性能・信頼性を提供可能なサービスをいかに構成するかが重要な課題と言える。

表 3-2: JAIST で取り扱うデータの種類の種類と要求されるストレージサービスの特徴

データの種類の種類	ストレージサービスへの要求				
	容量	高速性	可用性	信頼性	バックアップ
ホーム領域	◎	◎	◎	◎	◎
テンポラリ領域	◎	△/○	○	△/○	△/○
ワーク領域	◎	◎	○	○	△
アプリケーション領域	△	◎	◎	◎	○
サーバ領域	◎	○	○/◎	○	◎
バックアップ領域	◎	△	△	△	-
ログデータ領域	○	△	○	○	○

△ : 要求が小さい、○: 要求が大きい、◎: 特に要求が大きい

3.1.4.3. ストレージサービスの概要

従来は、特徴の異なる 3 種類のファイルサーバシステムを組み合わせることでサービスすることによって、ホーム領域で要求される高速性・高可用性・高信頼性だけでなく、ワーク領域やテンポラリ領域で要求される大容量のサービス領域、多様なクライアント環境にサービスを行うために要求される柔軟性・安全性等といった要求要件を実現してきた。しかしながら、大学におけるインフラとしてのストレージシステムは 24 時間/365 日にわたる安定したサービスを要求されており、分散したファイルサーバシステムによる運用管理の煩雑さは日を迫るごとに大きなものとなっていた。

これらの問題を解消するために、本学では複数の用途での利用を想定した統合ストレージシステムを導入している。図 3-5 に示す統合ストレージシステムは、教員・研究生・学生のホーム領域、研究プロジェクトのテンポラリ領域、バックアップシステムのバックアップ領域等をサービスすることを目的とし、以下の要素で構成される。

- (a) ホーム領域向け:
高可用性(クラスタシステム)、高速性(高性能ディスク、複数コントローラ、大容量キャッシュメモリ)、マルチプラットフォーム(NFS および CIFS プロトコル)を実現する高速ファイルサーバ
- (b) テンポラリ領域向け:
容量重視(高速ファイルサーバよりやや低めのスペック)、バックアップ取得あり・なし(重要度に応じて領域割り当て)を実現する大容量ファイルサーバ
- (c) バックアップ領域向け:
サービスに影響しないオンラインバックアップ(SnapMirror 機能)、容量重視(大容量ファイルサーバと同一)を実現するディスクバックアップ

本システムは、ホーム領域のように小さな設定ファイル等のリード/ライトが頻繁に発生するケースだけでなく、テンポラリ領域のように一定の信頼性のもとで大きな容量が必要なケースにも一つのシステムで対応することが可能であり、スケールメリットによるコストダウンを図りつつ、集中的に運用管理可能なシステムを実現している。

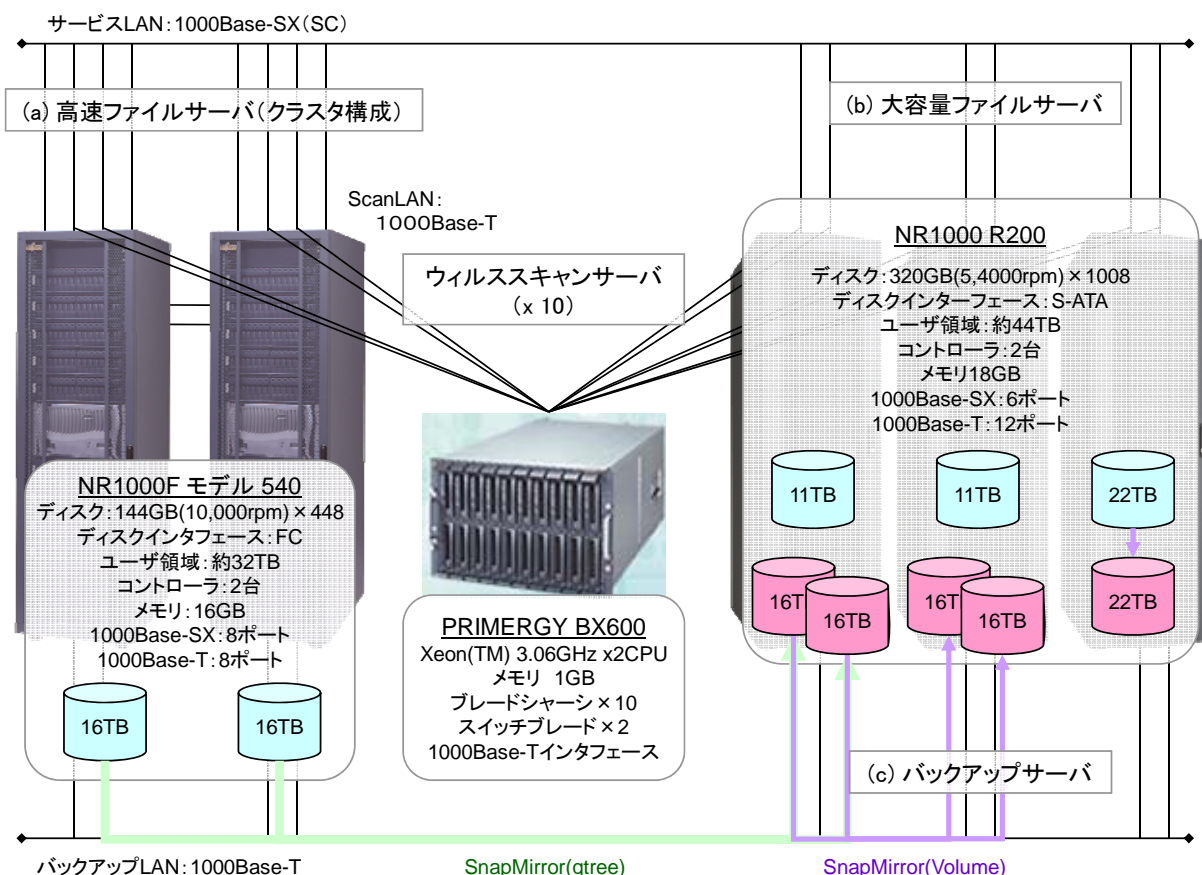


図 3-5: JAIST における統合ストレージシステムの概要

3.1.4.4. 現状および将来的な課題

2章で定義されたデータマネジメントにおける8カテゴリに従って、JAISTストレージシステムに対する現状の課題および将来的な課題を優先度の高いものから順に述べる。

●現状の課題

① (2)アーカイブ・HSM

学内情報環境の大部分のストレージサービスを中央に集約するモデルを採用する JAIST においては、非常に大きなディスク領域をコスト効果的な形で提供することが必要不可欠である。こうしたサービスを実現するためには、取り扱うデータの目的に応じて適切な容量・性能・信頼性を確保可能なディスク空間を提供する HSM 的ストレージサービスを実現できることが望ましい。図 3-5 に示す構成は、目的別のディスク領域を提供する状態に過ぎないが、ポリシーに基づくデータ移動の半自動化やそれに伴う透過的なアクセスが可能な構成が実現できれば、ユーザはディスク領域の性能や信頼性、ステージングを意識せずにストレージサービスを利用することが可能となる。

② (4)データアクセス管理

ストレージサービスを中央に集約することにより、学内ネットワーク経由でどのクライアントからもデータにアクセスすることが可能な環境においては、性能および可用性を確保するためのマルチパス制御が必要不可欠である。また一方で、成績データに代表される個人情報や特許関連の機密情報の保護についても、アクセス管理技術やストレージ管理技術、暗号化技術を組み合わせることで実現することが要求される。さらに、データを利用するクライアント端末が多様化している現状では、異なる種類のクライアントであっても同様のオペレーションでデータにアクセス可能であることが望ましい。

③ (5) デバイス管理

本項では JAIST における統合ストレージサービスの例を挙げたが、実際にはこの他にも導入時期や利用目的が異なるマルチベンダによる複数のストレージシステムが稼働している。管理者側としての立場では、これら全体の運用管理をいかに効率的に行うかについても大きな課題となっている。また、上記統合ストレージシステムについては更新時期がせまっており、サービスへの影響を最小限にしながらデータの移行を行うための方法を今後検討していく必要がある。これらのことから、ストレージシステムのハードウェア的な管理のフレームワークを提供するデバイス管理技術への期待は非常に大きい。なお、データの移行にあたっては、膨大なデータ量をいかに移行するかという問題もさることながら、ファイルサーバ室の容積や電源・冷房設備などといった物理的な制約も大きな問題となることに留意が必要である。

④ (1) データプロテクション

近年のディスク容量の急速な拡大により、バックアップ領域が従来のテープからディスクに変更されつつある。これにより、バックアップの速度や利便性は向上したが、一方で電源・冷房等の設備面の要件は大きなものとなっている。また、バックアップシステムに要求される信頼性はサービス領域と比較すると低いものとなるが、バックアップデータが必要な時に確実にリストアが可能でなければならない。このため、省電力・省スペースかつ一定の信頼性を持つバックアップシステムの実現に向けた研究開発が期待されることである。また、ユーザのホーム領域をサービスするシステムの運用局面においては、フルリストアよりもユーザがロストしたファイルをサルベージする機会が多くみられることから、こうした管理作業の負荷を軽減する機能も有用であろう。

● 将来的な課題

① (6) データレプリケーション

近年のクライアント端末のストレージ容量の増加に伴い、個々人が持つユーザデータは日々増大している。一方で、ユビキタス環境の進展により(オフラインを含む)あらゆる場所から多様なクライアントで情報が共有される状況も間近に迫ってきている。これらのことから、今後のストレージシステムには、ファイルサーバで閉じた機能だけでなく、OS やミドルウェアとの連携を通じてデータのレプリケーションやメタデータ、インデックス、バージョン管理等による付加価値を考慮に入れることが重要になるのではないだろうか。また、これらの付加価値を提供するための強力な機能をじっくり開発するアプローチだけでなく、基礎的な API を組み合わせ、システム管理者のレベルで、状況にあわせてデータに対する付加価値をつけるための枠組みを容易に設定できることが望ましい。

② (7) ファイルシステム/(8) その他

多様なクライアント端末からデータを取り扱う状況においては、オープン(かつセキュア)なファイルシステムが必要不可欠であり、ストレージベンダ、ネットワークベンダ、OS ベンダ等の垣根を超えた取り組みをさらに推し進めてもらいたい。また一般に、暗号化されたデータは、機器の更新におけるデータの移行やセキュリティのチェックなどが困難になる。こうした問題についても社会的コンセンサスを得ながら、管理と保護のバランスをとった解決策が求められていると言えよう。

3.1.5. 理研におけるデータマネジメントの特徴

3.1.5.1. はじめに

本項では理研の研究系データに対する管理や処理等について説明する。現在、研究所内で情報基盤センターが関連する大規模な研究系データには3種類の系統が存在する。一つはスーパーコンピュータシステムで生成されるシミュレーションデータ、一つはライフサイエンス系データベースのミラーデータ、最後に高エネルギー物理実験データとなる。それらについて現状と課題等を紹介する。

3.1.5.2. シミュレーションデータ

現在、理研で運用されているスーパーコンピュータシステム(RSCC³²)のデータ系ネットワークを図 3-6 に示す。RSCC は PC クラスタを計算資源の主力とし、ベクトル計算機や MD³³専用計算機を接続し、それらをシームレスに利用可能としたシステムである。RSCC のデータマネジメントとして特徴的なものは、データ処理系ストレージへのアクセス方法として 4 種類用いている。一つ目は FC による SAN 接続による GFS によるアクセス、二つ目は NFS によるアクセス、三つ目はバッチジョブ処理用の PC クラスタの各計算ノードがユーザホーム領域をマウントせず、ジョブ実行前に必要なデータをローカル HDD にステージングし、ジョブ終了後ステージアウトするアクセス、最後に HPSS³⁴に対するアクセスである。

本システムはいわゆる集中ストレージ領域(Home 領域など)を全計算ノードからアクセスさせる I/O 帯域を持たせるのではなく、計算ノードに接続される HDD を有効活用することで、安価に I/O 性能を確保することにある。ステージイン・アウトに要する時間もステージングを行う I/O ノードの数(現システムは 8 ノード)で調整可能である。現在の運用において、NFS 領域へのアクセス性能が低いという本質的な問題はあるが、ステージング・システムはステージング時間という本質的な問題はさておき、運用上の問題は起きておらず、並列ジョブの I/O 性能の振れも少なく、比較的安定で良好な運用性を保っている。

しかし、全ノードからホーム領域が見えないことで、従来の PC クラスタからの移行においてバリアがあり、ユーザ、特に演算性能をサービスしている計算システムの利用に慣れていないユーザは使いにくさを訴える場合が多い。また、システムを構築する側から見ると I/O 性能を如何に節約し、演算性能に振れるかも重要な構築指針である。今後計算システムのオンラインストレージを構築するに際して、I/O 性能と信頼性と利便性とコストのバランスシートはもう少しストレージ側を考慮する必要があると考えている。

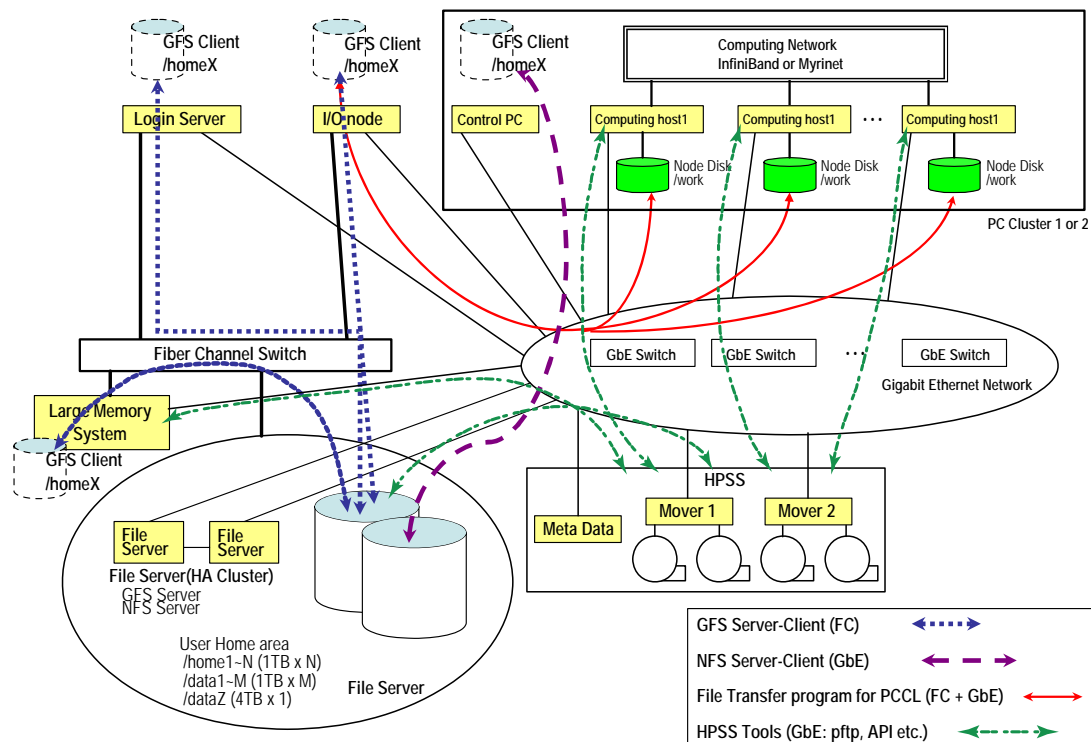


図 3-6: RSCC データ系ネットワーク図

³² RSCC: Riken Super Combined Cluster

³³ MD: Molecular Dynamics

³⁴ HPSS: High Performance Storage System

3.1.5.3. ライフサイエンス系データベースのミラーデータ

理研の主要な研究領域としてライフサイエンスがある。ライフサイエンスでは、生物の遺伝子またはタンパク質の配列データの相同性検索やアノテーション情報等が重要となっており、実験によって得られたデータや注釈づけされたデータが、データベースとして世界中で管理されているため、研究者各々は必要なデータベースを各サイトからダウンロードして利用することになる。そのため以前は、理研ドメインから同一のデータベースサーバへ多数のアクセスが集中し、データベース管理者にも、理研以外のデータベースの利用者にも非常に不評であった。

そのため、理研の情報基盤センターでは、理研から各データベースへのアクセスを一元化および理研内の研究者に対する利便性の向上を目的として、データベースミラーサーバを用意している。ミラーデータシステムの運用ポリシーとしては、更新があったデータベースに対して1日以内に最新のデータベースをミラーし、理研内ユーザに提供することを努力目標としている。またスーパーコンピュータ(RSCC)を用いて解析を行うものについては、同スーパーコンピュータのストレージシステムへのミラーも行っている。

現行のシステム構成としては、1+2 台のサーバと NAS(リダンダント構成)であり、3 台のうち1 台はデータベースのダウンロード専用サーバとして、世界中に点在する FTP サーバ等にアクセスし、更新の確認や最新データのダウンロードを行っている。その他2 台はユーザサービス(リダンダント構成)を行っている。なお、昨年度まで利用していた旧サーバの各ボリュームをソフトウェアで仮想テープ装置と見立て、1 世代分の全てのデータベースをバックアップしている。現在運用しているシステムは後3 年利用する予定であるが、現状大きな問題は起きていない。システムパフォーマンスもある程度満足できる状況にある。

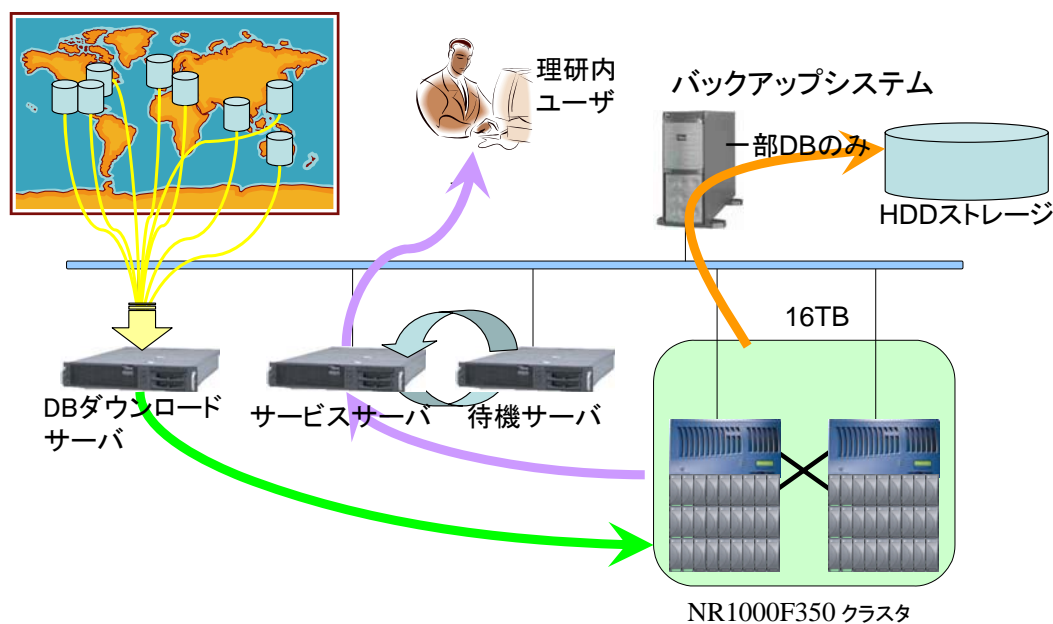


図 3-7: ライフサイエンス系データベースのミラーデータ

3.1.5.4. 高エネルギー物理実験データ

高エネルギー物理学実験から出力されたデータは粒子の加速エネルギーの増加や検出器の高性能化により爆発的に増加している。理化学研究所は米国のブルックヘブン国立研究所(BNL)に研究センターを持ち、同研究センターが主導して BNL にある RHIC³⁵加速器を用いた PHENIX 実験を行っている。PHENIX 実験は原子核・原子核衝突により、宇宙初期と同様の高温・高密度状態(QGP)を生成し、その性質を研究・解明することを目的としている。実験で生成されたデータは完全に分散解析可能なデータで、Embarassingly Parallel で処理される。

³⁵ RHIC: Relativistic Heavy Ion Collider

実験データの出力状況はこの2、3年は1RUN(1回の衝突実験ではなく、スケジュール上の1回)あたり200~300TB程度の生データが出力される。現時点でのデータ量は1.2PB程度である。2年前まではデータを入れたテープを空輸していたが、最近はインターネット経由でのデータ転送を行っている。およそ5TB/day程度の転送を行っているが、今年度から10TB/day程度の転送性能が得られるような準備を行っている。2005、2006年度での総転送量は約550TBとなった。生数値データをRSCCで一次解析を行い、再構成されたデータとして35TB程度のデータをBNLに転送している。

データの格納にはHPSSを用いている。HPSSはHSMの一種であり、HDDストレージとテープで構成されている。HPSSは実験データと上記RSCCのデータストレージとして用いている。テープドライブ14台(9940B:200GB,30MB/sとT10000:500GB,120MB/s)とキャッシュ用HDDストレージで構成されている。テープはライブラリ装置(StorageTek PowderHorn9311)をスーパーコンピュータ用と実験データ用に用いており、現状9000本程度のテープが納められている。

実験データ処理グループは、現在、一次解析後の再構成されたデータを用いた計算処理中に用いるデータもHPSSを用いて、データ処理中にテープからローカルHDDへ移動させて計算処理を行っている。しかし、次期システムではさらに計算ノードのローカルHDDを有効に使い、予めデータを分散して配置し、データがあるところでデータ処理を行う方式である、Data Intensive Computingに移行したいという希望を持っており、そのような場合、どのようなシステム構成がありえるのかを検討中である。

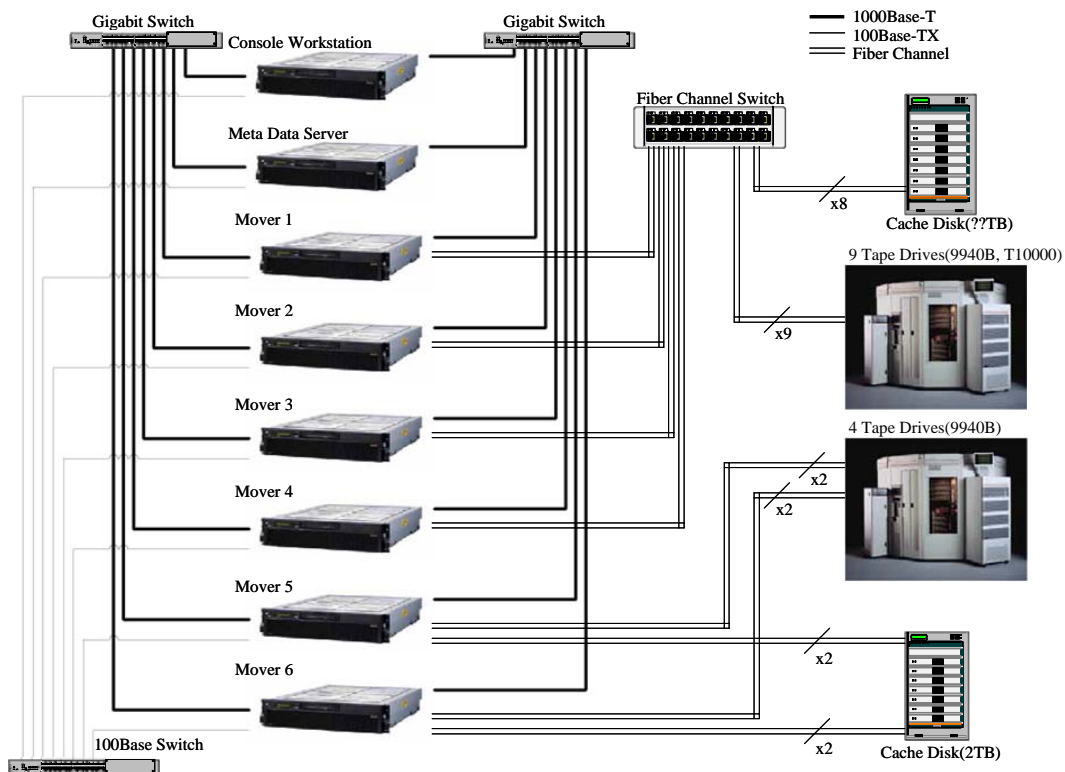


図 3-8:HPSS システム

3.1.5.5. 現状の課題

2章で定義されたデータマネジメントにおける機能範囲である8カテゴリに基づいて現状の課題を述べる。

(1) データプロテクションの観点

上述の理研におけるデータマネジメントにおいて、データプロテクションを行う必要があるのは、スーパーコンピュータシステムにおけるユーザデータの一部のみである。ライフサイエンス系データベースや高エネルギー実験データは、元のデータがあって、その複製を保持しているため、自動的にバックアップやディザスタリカバリがされている。ただ、ライフ系データは元データを保管している相手先には伝わっていないため、本質的にプロテクションとは言えないが、理研でプロテクションを意識する必要はない。高エネ実験データは BNL－理研間双方向でリカバリされているという前提の元でデータ格納が運用されている。

スーパーコンピュータで出力されるデータのプロテクションの課題としては、出力データは元になるデータやソースコードがあれば、基本的に再取得可能であり、その出力データをどこまでプロテクションするかと言う点である。元データやソースコードは容量としてそれほど大きなものではないため、一般的な方法でバックアップ可能である。また、出力データをバックアップするかは基本的には運用ポリシーとコストの問題であり、バックアップデータ量とデータを出力可能な時間および計算機コストおよびバックアップ機器費用等から、何かしらか一般的な指針や指標の策定が望まれる。

(2) アーカイブと HSM の観点

スーパーコンピュータ用アーカイブと高エネルギー物理学実験のデータ格納用に HPSS を用いている。理研で用いている HPSS では最下層の媒体としてコストの問題でテープを用いている。テープを用いる HSM は全般として、多くの課題がある。読み込みのレイテンシが非常に大きなブレがあること。粒度の細かい大量のデータを扱うのが苦手であること。データ移行に時間がかかる、特に細かいデータ移行は膨大な時間がかかること。また、テープドライブとその媒体の進歩で読み書きの互換性が無くなる場合があること、等が挙げられる。多くの場合、運用でカバーしているが、データ格納を必要とし、利用者がそれを意識しているプロジェクトにおいては、それほどのネガティブ要因ではないが、特にスーパーコンピュータシステム内で利用する HSM システムとして考えた場合、ユーザに不便を感じさせ、クレーム対象となることが多く、また、管理者として頭の痛い問題である。これらの問題を一つでも解決するソフトウェアが有ればよい。

テープを用いるアーカイブの本質的な問題は、PB 級のデータを保持していく場合の将来的なデータ移行・保持の問題が非常に大きい。それには三つの問題があり、一つ目はデータ移行に膨大な時間を要すること。二つ目はドライブと媒体の互換性の問題、最後にデータを保持する期間を明確にする必要があり、そのコストをどうするのかという問題がある。最後の問題はポリシーの問題であるため、ここでは触れない。一つ目の問題はテープからのデータ移行の本質的な問題であり、技術でどうにかなる問題かどうか不明である。ただ、無法地帯で利用された HSM のテープからのデータ移行は致命的な問題を発生することが明かであり、運用は十分に考えておく必要がある。そのため、ある程度粒度の大きなデータをユーザ透過に使う方法が欲しいところである。二つ目の問題が大きな課題であるが、テープに記録されたデータの寿命等を考えた場合、テープ to テープのデータ移行はやむを得ないのかもしれない。ただ、テープドライブ・媒体の互換性が十分考慮されることを期待する。

(3) ストレージ管理の観点

現在のシステムにおけるストレージ管理の課題は構成が正常に動作することを確認可能とすること、スーパーコンピュータシステムのファイルシステムの拡張性が欠けることである。容量・速度・可用性などを確保するためにストレージシステムは年々複雑な構成となってきている。スーパーコンピュータシステムやデータ保管やサービスを行っているサーバのストレージシステム構成を考えても、RAID やコントローラやパス、サーバ等の二重化は必ず行われている。これらの構成はハードウェア的に組むことは簡単であるが、サーバソフトウェアとの連携等ハードウェア構成だけでは閉じない事が多い。問題はストレージ構成が予定通りに機能しているかをいかに検証できるかが重要で、ストレージシステムを構築し、構築したストレージ構成が必ず意図したサービスが実現できるか、機能するかどうかを検証できるソフトウェアが必要だと考えている。

また、年々増加するデータを管理する場合、後述のファイルシステムやテープの管理本数のフレキシビリティが非常に重要となる。ファイルシステムの場合、運用停止をすることなく、HDD の拡張やファイルシステムの拡張が出来ることが重要となる。テープの場合、テープの管理(マイグレーション等)が柔軟かつ効率よく、テープの投入が容易であることが必要であり、HSM ソフトウェアの柔軟さやテープライブラリ装置の使いやすさがポイントである。

(4) データアクセス管理の観点

データアクセスを行うストレージシステムへのリーチャビリティの管理という観点でスーパーコンピュータシステムでもバイオ系データや実験データでもほぼ同じ事が言える。RAID 装置とコントローラおよびサーバから構成されるシステムでは、その間の接続状態をサービスレベルとともにモニターする必要があり、トータルシステムとしてユーザサービスが問題なく動作しているのか、それとも設計と異なる動作をしているのかを判定できる管理ソフトウェアおよびその状況を管理者に通知する仕組みが必要であり、ユーザにはサービスレベルを通知する方法が必要となる。データアクセス管理とは少し異なるが、ユーザにインターネット経由でデータをサービスするような場合、あるいは、インターネット経由でデータを更新する必要があるような場合、インターネットの回線品質(遅延、帯域等)も含めてモニタできるようなシステムが望ましい。

(5) デバイス管理の観点

ストレージシステムはスーパーコンピュータシステムにせよ、バイオ系や高エネ実験データシステムにせよ、システムを中心に位置するものであり、構成するデバイスの管理・設定変更・システム増強を正常に動作・変更可能とするストレージシステムを運用することは、システム全体の安定動作に不可欠となる。その中で RAID やコントローラ、テープドライブ、ライブラリ装置の個々管理は当然として、トータルシステムとしてのデバイス管理という観点を持った管理ソフトウェアが必要である。

また、巨大データの移行性をこの先どのように確保するかがきわめて重要であるが、移行時間を考慮すると事実上不可能な場合が多い。ただ、そのような場合、データ移行が出来ないという物理的な問題あるいはテープ媒体やドライブの互換性の問題やファイルシステムの柔軟性の欠落から、全くデータ移行が出来なくなる可能性が非常に高い。この状況の一番大きな問題は性能が良く、低価格な別製品への移行性も奪い、競争原理が働く可能性が高く、由々しき問題である。このため、巨大データの移行性あるいはバックアップを考えていくことは、今後のデータの移行性やサービスの性能向上の上でも重要であり、検討が必要な問題である。

(6) データレプリケーションの観点

スーパーコンピュータではローカル HDD にデータをステージング(レプリケーション)することで、ストレージシステムの集中を防ぎ、ローカル HDD のアクセス性能を引き出し、並列計算時のストレージ性能を十分に引き出している。ただし、これはユーザビリティの観点では、ユーザがデータ移動のイメージを持つ必要があり、UI を工夫し使いやすい利用環境を提供しているが、バリアを持ったシステムである。理想的にはステージングなどせずに十分な性能を持つストレージシステムを設計すべきであるが、これは価格性能のバランスを考慮していくしかない。並列計算時の十分な性能を提供するには、ローカル HDD を利用するのがリーズナブルであるが、ステージングのためにいかに簡単な UI を提供できるかが重要である。

また、バイオ系データや高エネ実験データに関しては、上記でディザスタリカバリの観点で記述したが、それ以外には仮にインターネットの帯域が十分にあるような場合は特に今のように二重化のような構成は必要ない。数十から数百 TB クラスのデータをインターネット経由で自由に移動出来る帯域を持つことは無いと考えられ、これからも性能の問題としてもレプリケーションは必要である。インターネットを通したレプリケーション時のデータの一貫性を如何に確保するかである。

(7) ファイルシステムの観点

ファイルシステムにおいて POSIX 標準 I/F を有することは言うに及ばず、一定の可用性を満たした上で、拡張性と高い性能を確保することは導入時の設計に非常に重要な問題である。拡張性

と高性能はファイルシステムだけではなく、いくらハードウェアに拡張性があり、高性能であってもファイルシステムによって、その拡張性や高性能が犠牲にされることがしばしばある。

現在スーパーコンピュータシステムのストレージは EXT3 をベースとしており、最大のファイルシステム作成サイズが 1TB と非常に小さく、拡張性に欠ける。ストレージとして大きなサイズを用意しても、活用できていない。単一ファイルシステムのサイズは大きければ大きいほどよい。その中でユーザがサイズを設定できるようなファイルシステムが望ましい。また、I/O 帯域も 1 台のサーバで FC の本数や外部 I/O 能力でまかなえる帯域で縛られるのではなく、ストレージを直結したサーバ台数の増減によって帯域の調整や可用性を高められる構成が取れる事が望ましい。

バイオ系データベースのサーバは NR システムの OS である Data ONTAP GX 10.0 を用いており、拡張性はあると思われるが、現時点では拡張したことがないため、どんな問題が発生するかどうかは不明であり、問題がないことを期待する。

(8) その他(圧縮、暗号化など)の観点

理研のシステムの多くは暗号化も圧縮も必要ない。それよりも Grid 的なファイルシステムを管理できるストレージマネジメントを期待したい。具体的には、スーパーコンピュータシステムに PC クラスタを用いる場合、多くでローカル HDD を持つ。ローカル HDD の総容量は 10TB を越えるが、ステージングのみでは有効に活用できていない。細かいステージングを行うことで、多くの部分は活用できるが、ローカル HDD でクラスタファイルシステムが構築できれば、ステージングの細かさをかなり緩和することが可能で、帯域もスケラブルに増強可能となる。また、データ Grid で用いられるデータインテンシブ・コンピューティングなどでも活用でき、ジョブ・スケジューラとの連携で有効性は非常に高いと考えられる。

また、ファイルシステムとして、PC の増加とともに容量をスケラブルに増加できる拡張性の高いファイルシステムを構築できれば、年々データが増加するような実験データなどにも柔軟に対応可能であり、そのようなストレージシステムを容易に導入できることを期待する。

3.2. 調査結果の整理と分析

本節では 3.1. 節に述べた各機関のデータ特性、システムの特徴などの調査結果を分析し、さらに各機関に共通する課題を抽出し、カテゴリごとの要望にまとめた。

3.2.1. 各機関保有データの特性

調査対象のシステム内の主なデータを以下に記す(詳細は添付資料を参照)。

- ・ JAXA : スーパーコンピュータ上の数値シミュレーション入出力データ、風洞試験・飛行試験・観測等の実験データ、公開 Web 用コンテンツデータ、電子メール
- ・ 天文台 : ハワイ観測所における観測データ・望遠鏡エンジニアリングデータ・環境・気象データをハワイおよび三鷹本部に保管、野辺山 45m・10m 望遠鏡の観測データ
- ・ 中京大 : 理系学部の教育研究用コンピュータ演習室に関するデータであり、学生演習用個人データ、課題・レポート提出領域、個人用ホーム、およびメールスプール
- ・ JAIST : 個人用ホーム、スーパーコンピュータ上の数値シミュレーション入出力データ、遠隔教育用コンテンツデータ、外部公開用 Web・FTB 領域、メール、各種ログデータ
- ・ 理研 : スーパーコンピュータ上の数値シミュレーション入出力データ、高エネ物理学実験データ、および公開遺伝子・タンパク質データ

保有データの調査の結果から以下のデータ特性がわかった。

- ・ 研究・教育用のシステムが対象のため(事務関係システムは調査対象外)、データは殆どが非構造型である。
- ・ 非構造型データの内、観測／実験データのように公開／共有されているデータは、メタデータの

み市販や独自の DBMS で構造化し管理している。また、永久保存のものが多い。

- ・大学のデータは目的別に領域を分けて管理している。
- ・学生のディスク上保有データは在籍期間に限られる。
- ・データの種別により要求される要件が異なる。

数値シミュレーション入出力データ: 高速性

観測データ: 大容量、保水性

ホーム領域データ: 可用性、高速性

ワーク領域データ: 大容量、高速性

今回の調査対象機関も同様であるが、一般的に大学・研究機関の共同利用システム内に保有されているデータのもっとも特徴的なことは非構造型データが多いということである。この非構造型データを如何に管理するかが、ストレージシステム管理を効率化するためのポイントと言える。

特に天文学観測データ、実験データおよびライフサイエンスデータは膨大な数のファイル群として非構造型データのまま保存されており、データアクセスの容易性を向上させるため管理データのみを管理ツールや DB 内に持ちメタデータデータベースを構築している。これらのデータは多くの利用者に活用されており、たとえば天文観測データは複数のファイルデータを処理することが多く、またそのデータから派生するものや、コピーされるものがある。さらに、データが不要になったときには、個々のファイルだけでなく複製を含めて一括削除したいとの要望があるため、データ管理者はこれらのデータの状況を把握し管理する必要に迫られている。

また、スーパーコンピュータなどのシミュレーション結果は、関連のある複数の結果が集まって初めて意味をなすこともあり、それらをまとめて構造的に管理したいとの要望がある。

データマネジメントを意識してストレージシステムを構築する観点から、このようなファイル群をコンテンツとしてまとめて管理するデータマネジメントの概念(コンテンツ管理)が浮き彫りになった。

3.2.2. 各機関の現システムにおける対応状況

各機関は、システムの目的や状況変化に応じてストレージシステムをより使いやすくするため、限られたコストで最大の効果を狙い、様々な対応を取っている。各機関の現システムの状況から抽出した特徴を表 3-3 に示す。研究機関は大量の観測データや大規模シミュレーション入出力データがあるため、データ容量が大規模になっており、ストレージ接続形態は DAS や SAN が多くなっている。また、複数ファイルサーバを保有し、大人数でストレージシステムを共用することが多い教育システムでは NAS を選択されている。このように、各機関とも要件特性にあった、ストレージ接続形態やストレージの導入をはかっている。

表 3-3: 現システムの特徴

	項目	JAXA	天文台	中京大	JAIST	理研
1	データ容量	大	大	小	中	大
2	ストレージ接続形態	DAS	SAN	NAS	NAS	SAN+NAS
3	観測データ提供	有	有	無	無	有
4	HSM	有	有	無	無	有

まずは各機関が現状のストレージシステム・データマネジメント導入時の課題に対してどのような対応を行ったかを代表的なデータ種別ごと(数値シミュレーション入出力データ、観測データ、ホーム領域データ)に以下に記述する。

3.2.2.1. 数値シミュレーション入出力データ

JAXA、JAIST、理研が対象機関であるが、各機関ともストレージへのアクセスの高速性を重視している。また、併せて大容量も必須条件となっており、これらの対応としてデータアクセス管理とファイルシステムに工夫をこらしている。その対応の代表例として JAXA の例を以下に記す。

① (4)データアクセス管理、(2)アーカイブ・HSM

JAXA においては高速性を確保するため、ストレージを集約しシステムの中核に据えるというデータセントリックの考え方にに基づき構成し、クロスバネットワークで直結されている計算サーバから、インターネット上に存在するクライアントまで広範なクライアントに対してストレージサービスを提供している。また、HSM を導入し、高速にアクセスできるストレージとして RAID5 ディスク装置があり、その後ろに 600TB のテープ空間を一元的に管理している。ディスク装置は大規模 I/O にチューニングされており、ストライプ技術により 8Gbps の実効性能を実現している。また、ディスクから溢れ出してしまったファイルは HSM システムにより 600TB のテープ空間へ自動的に移動され、必要な時に自動的に取り出される。また、テープ装置の速度がディスク装置の速度より遅いため、ディスク～テープ間のデータ移動においては、テープドライブの速度がボトルネックにならないように、テープ媒体 4 個への同時書き込み処理を行うことにより高速性の向上をはかっている。

② (7)ファイルシステム

計算サーバの FORTRAN プログラムから実効 8Gbps、三次元可視化サーバから 4Gbps、構内ネットワークから NFS より数倍速い速度でアクセスでき、さらには、実験・開発的にインターネット上に存在するクライアントからアクセス可能なファイルシステムの構築を行っている。

3.2.2.2. 観測データ

観測データは、JAXA の風洞実験データ等が 100TB、天文台の天文観測データが 100TB、理研の高エネルギー実験データが 1PB、と大容量であることが特徴である。また、観測データは再度同じデータを取得できないため、原本保障が重要視されている。HSM の代表例として理研を、データプロテクションの例として天文台を以下に記す。

① (2)アーカイブ・HSM

理研では高エネルギー物理学を BNL にある研究センターで行っている。実験データの出力データは 1RUN あたり 200～300TB 程度発生する。現時点でのデータ量は 1.2PB 程度である。2 年前まではデータを入れたテープを空輸していたが、最近はインターネット経由でのデータ転送を 5TB/day 程度行っている。この大容量データの格納には HSM の一種である HPSS を用いている。HPSS はテープドライブ 14 台(9940B:200GB, 30MB/s と T10000:500GB, 120MB/s)とキャッシュ用 HDD ストレージで構成されている。

② (1)データプロテクション

天文観測データは、同じデータを二度と撮れないことから、データプロテクションと原本保障を第一に考えている。基本は原本のバックアップによってデータ保障性を確保しつつ、さらに災害・天災時や経年劣化やコンピュータシステムの不具合、および人間の誤操作を考慮してハワイと日本に観測データの二重持ちしてディザスタリカバリを実現している。しかし、「二重持ち」にしているとハワイ側で持っている観測データと日本側で持っている観測データの同一性保証が問題となる。両方の観測データをメッセージダイジェストアルゴリズムによってハッシュ値を計算し、両者の値を定期的にチェックする仕組みで、整合性を監視するように改良している。加えて、研究者に観測データを配布する際にも同様に読み出したタイミングでハッシュ値を算出してデータの保障性を高めている。

3.2.2.3. ホーム領域データ

大学の教職員や学生が使用するインフラとしてのストレージシステムは、24 時間/365 日にわたる安定したサービスを要求されており、Unix ワークステーション・Windows ターミナルサーバから共

有されるホーム領域は講義や演習の妨げにならないよう、可用性・高速性が要求されている。中京大および JAIST を例に記す。

① (3)ストレージ管理

中京大では講義期間中は多くの講義で利用され、特に期末に集中する課題提出については学生の成績に影響するため、ユーザが期限内に確実に提示・提出(指定領域に保存・移動)できることを保証するストレージシステムが必要である。そのため、ストレージシステムには高可用性が求められており、ストレージシステムと演習室を3系統で接続することにより冗長化している。JAIST では分散したファイルサーバシステムによる運用管理の煩雑さを解消するため、ストレージシステムを中央に集約し、学内ネットワーク経由でどのクライアントからもデータにアクセスすることが可能な環境にした。可用性を向上させるためクラスタ構成のファイルサーバを導入し、さらに LAN への接続数を増加させた。

3.2.3. 課題の抽出

各機関は、3.2.2.項のように状況に合わせ様々な対応をとっているが、まだまだ現状および将来への課題は多く残っている。そこで前 3.1.節にある課題のカテゴリについて、各機関にとってどのカテゴリを課題として重要視しているか調べた。その結果を表 3-4 に示す。

なお、カテゴリごとに最重要課題(◎ : 5 点)、重要課題(○ : 3 点)、将来対応すべき課題(△ : 1 点)、対応不要(－ : 0 点)を選択し、重要度の順位付けをした。

表 3-4: 課題の重要度一覧

		JAXA	天文台	中京大	JAIST	理研	得点
1	(4)データアクセス管理	◎	◎	◎	◎	◎	25
2	(2)アーカイブ・HSM	◎	○	△	◎	◎	19
3	(1)データプロテクション	△	◎	◎	○	○	17
4	(7)ファイルシステム	◎	△	○	△	◎	15
5	(5)デバイス管理	○	△	△	◎	◎	15
6	(3)ストレージ管理	○	△	○	○	◎	15
7	(6)データレプリケーション	△	○	－	△	△	6
8	(8)その他(圧縮、暗号化など)	△	－	－	－	△	2

課題の第一位は「(4)データアクセス管理」であり、全機関とも重要課題との認識であった。次いで「(2)アーカイブ・HSM」、「(1)データプロテクション」、「(7)ファイルシステム」、「(5)デバイス管理」、「(3)ストレージ管理」と続き、「(6)データレプリケーション」、「(8)圧縮、暗号化」はコストを意識し優先度付けした結果、将来対応すべき課題となっている。

内容としては、エンドユーザへのサービスレベル向上、管理者の負荷軽減、状況認識の即時性向上が主なものである。以下に各カテゴリの具体的内容を重要度の得点上位カテゴリから記述する。

① (4)データアクセス管理

各機関とも利用者への利便性の向上のための課題をあげており、一番の課題は利用者からの一定のアクセススピードの確保、二番目はデータアクセス時のオペレーションである。

研究所は、インターネットでの大量データ配信や他研究機関とのデータ交換、大学はキャンパスの分散などで、データ共有のエリアがより広域化している。また、接続方法も多様化している。このような状況において場所や接続方法が異なっても一定のアクセススピードを保証しサービスレベルを確保したいというのが一番目の課題である。そのために、接続状況やサービス状況を常にモニタし、設計と異なる動作をしたとき管理者へ自動通報し、利用者へはサービスレベルを通知するような管理ソフトウェアが望まれている。

また、二番目の課題であるデータアクセス時のオペレーションに関しては、現在は使用クライアント端末、接続方法、使用ファイルシステムによりオペレーションやアクセス制御が異なっていることが課題として挙げられている。これらの場合でも同様なオペレーションでのデータアクセスが望まれている。

一方、個人情報や機密情報の保護についてより強化したいとの要望も挙げられている。そのためには、データアクセス権の細かい制御がより簡単に設定できるようにしたい、というものである。その他、多様なクライアント端末からデータを取り扱う場合や、用途に応じて異なるファイルシステムを使用しなければならない場合も多い。OS やベンダの垣根を越え、ファイルシステムのユーザインタフェースの統一化を期待する声も多い。

② (2)アーカイブ・HSM

大規模データを所有する 3 研究所は、いずれもディスクに比べ比較的安価な二次媒体にテープを使用した HSM を導入して、大容量ストレージの提供を行っている。しかし、HSM に対する改善要求は多くある。

一番の要求は、速度と信頼性の向上である。HSM は読み込み時のレイテンシのバラツキが大きく利用者へのサービス向上のためには一定化を望む声が多い。また、信頼性に関しては、データの保証をするために、テープとディスク間のコピー時およびテープ自体の信頼性の向上が望まれている。

次にテープドライブや媒体自体の互換性を望む声がある。これはシステム更新時などのデータ移行時間が多大にかかっており、それを解消するのが主な目的である。

その他としては、遠隔ファイルシステムとの連携によるデータプロダクションの実現、HSM のライセンスの定量課金化、等が望まれている。

以上のように HSM 自体への要望が多いが、HSM とコストが同程度のディスクの提供が各機関の本音と考える。

③ (1)データプロテクション

各機関ともバックアップ・リカバリ時間短縮と手間の簡素化を望んでいるが状況・課題は様々である。各機関の状況および課題を以下に記す。

JAXA は大規模 HSM により、自動バックアップ／リカバリを既の実現しており、現状の課題はバックアップ不要ファイルの自動認識を挙げている。

天文台はハワイ山頂システムでバックアップし原本管理をしている。また、山麓、三鷹へ送られることによりディザスタリカバリを実現している。天文観測データは再観測が不可能であり、観測装置の進歩によりデータの質量ともに巨大化している。そのため、バックアップ時間の短縮と確実なリカバリ手段の提供が課題となっている。

中京大は、卒業生から学生時代のファイルの閲覧希望が増えており、その都度手動でテンポラリ領域へリカバリし、ファイルを探索している。この作業を簡略化するため、大きなシステムを必要とせずテープから直接必要なファイルを簡単に探索する仕組みを望んでいる。また、講義期間中と休暇期間中で使用頻度の格差が大きく、システムの稼働状況を監視しながらバックアップ頻度・範囲の調整を自動化したい、との要望もある。

JAIST ではバックアップ領域をテープからディスクへ変更しつつある。速度や利便性は向上してきたが、電源・冷房装置等の設備面の負荷が拡大しているため、省電力・省スペースかつ一定の信頼性を持ったバックアップシステムの実現を望んでいる。

理研ではデータプロテクションを行う必要があるのは、スーパーコンピュータシステムにおけるユーザデータの一部のみである。その課題としては、出力データは元になるデータやソースコードがあれば、基本的に再取得可能であり、その出力データをどこまでプロテクションするかと言う点である。バックアップデータ量とデータを出力可能な時間および計算機コストおよびバックアップ機器費用等から、指針を出そうとしている。

④ (7)ファイルシステム

ファイルシステムについては、高速性、拡張性の向上が望まれている。CPU(特にスーパーコン

コンピュータ等のハードウェアがいくら高速化され拡張性を持っていても、ファイルシステムがボトルネックとなりその性能や拡張性を犠牲にしてしまうことがしばしばある。ハードウェアの機能・性能を損なわないファイルシステムが望まれている。

その他、ファイルディレクトリへのアクセス制御設定の容易化、遠隔地や複数機関でのファイル管理を実現する遠隔共有性向上の期待がある。

⑤ (5)デバイス管理

導入時期や利用目的が異なるマルチベンダによる複数のストレージシステムが稼働している。管理者側としての立場では、これら全体の運用管理を如何に効率的に行うかについても大きな課題となっている。更新時にはサービスへの影響を最小限にしながらデータの移行を行うための方法を今後検討していく必要がある。これらのことから、ストレージシステムのハードウェア的な管理のフレームワークを提供するデバイス管理技術への期待は非常に大きい。また、マルチベンダかつ各種ストレージ(RAID やコントローラ、テープドライブ、ライブラリ装置)の個々管理は当然として、トータルシステムとしてのデバイス管理という観点を持った管理ソフトウェアが必要である。

また、短期間でのデータ移行、物理的制約(容積・電源・設備)の軽減も挙げられている。

次の課題は、構成が正常に動作することを確認可能とすることである。RAID、コントローラ、パスやサーバ等の二重化は必ず行われている。これらストレージ構成が予定通りに機能しているかを如何に検証できるかが重要で、構築したストレージ構成が必ず意図したサービスを実現しているかどうかを検証できるソフトウェアを望む声もある。

⑥ (3)ストレージ管理

ストレージ管理での課題の一番目は高可用性である。単体ディスクが故障した場合は、RAID 構成となっていれば運用を停止することなく復旧可能であるが、RAID コントローラや複数故障が発生した場合は、運用を停止する必要がある。大容量化が進めばこの課題は今まで以上に顕在化してくると思われる。

⑦ (6)データレプリケーション

デジタルデータの複製時に劣化しない保障はない。文書データとは違い、天文観測データはホワイトノイズに近い複製ミスを見逃すことが人間が発見することは困難である。デジタルデータの複製方式も進化させる必要がある。

⑧ (8)その他(圧縮、暗号化など)

今回調査した会員機関では、データ圧縮や暗号化は行っていないが、今後多岐にわたるユーザが単一システムを利用する場合、暗号化は利用者から要求される可能性が高い。具体的には、システム内でのデータ漏洩防止やデータをシステム外に持ち出す場合のセキュリティの観点から、ファイルの暗号化が必要になる可能性がある。また、データ量が増大する中、ストレージの媒体総容量を圧縮するためにデータ圧縮技術を利用することが期待されるが、ファイルへのアクセス速度の劣化を不安に思っている。

課題の抽出時、8 カテゴリに分類しがたい課題も挙げられた。ログ／解析データ統合管理とデータ Grid への対応である。以下にこの2点について述べる。

ログ／解析データ統合管理は、マルチベンダかつ多種のストレージを保有している中、ログが各ストレージから出力されるため大量過ぎて活用できていない。それらのログを統合的に管理し、システムの状態を的確に把握したいという要望がある。二点目のデータ Grid への対応については、高エネルギー物理学実験の分野ではペタバイト規模のデータ解析が必要であり、効率的なデータ連携が必要となっている。しかし、Grid コンピューティングからの類推としてデータ保存ノードとデータ処理ノード間での連携技術が課題となり、ネットワーク性能やデータ規模に対応するデータ管理技術がペタバイト規模のデータ処理のために必要となる。

3.3. ストレージソリューションへの期待

本節では 3.2.節で述べた会員機関の保有データ特性、現用システムの状況および課題から、今後のストレージソリューションへの期待について考察する。

8つのデータマネジメントカテゴリごとの課題調査においては、表 3-4 に示されているように、「(4)データアクセス管理」が最重要課題との認識は全ての参加機関で一致している。また、より特徴的な傾向は、「(1)データプロテクション」と「(2)アーカイブ・HSM」にあると言える。

「(1)データプロテクション」を実現する手段として代表的なものにバックアップ・リストアがあり、これらの仕組みは階層型ストレージでも使われていることから「(2)アーカイブ・HSM」は「(1)データプロテクション」の要素技術はほぼ同じものと考えられる。このことから、「(1)データプロテクション」と「(2)アーカイブ・HSM」の二つのカテゴリを併せて考えると「(4)データアクセス管理」と同等の重要度と参加機関の全てが認識しているといえる。この3カテゴリをはじめとした8カテゴリの課題へ対応するソリューションが期待されており、本質的なソリューションは複数のカテゴリやその他のシステム要素と複雑に絡んでいるといえる。例えば、「(4)データアクセス管理」で一番目に挙げられている「利用者からの一定のアクセススピードの確保」にしてもファイルシステム、デバイス自体や接続方式・形態等が関わっている。

このように、個々の課題解決方法（ソリューション）とともに、より多くの機能と連携した幅広い概念のソリューションも期待されており、提供側であるベンダもこれに対応し、2章に書かれているようにストレージベンダもソフトウェア・コンテンツベンダとの吸収・合併等を行い、体制を整えてきている。

会員機関の課題指摘からは、ストレージシステム／データ管理者が念頭に置いているマクロな要件は以下の3点と読み取れる。

- ・ 利用者の利便性向上
データアクセス管理、ファイルシステム、ストレージ管理に関係し、アクセススピード確保、オペレーションの簡素化、ハードウェア・ソフトウェア機能の最大限の活用、遠隔地との連携強化が挙げられている。また、データ Grid の進展も望まれている。
- ・ ストレージシステム／データ管理者の手間削減
アーカイブ・HSM、データプロテクション、デバイス管理に多く、最も重要視しているのはデータ移行であり、システム移行時などの大量データ移行には大きな問題意識が感じられる。また、運用時の課題としてはマルチベンダおよび各種ストレージの統合的な状況把握、バックアップ／リカバリ時間短縮、信頼性向上が挙げられている。さらに、3.2.1.項の各機関保有データの特性で挙げた、非構造型データの効率的な管理およびコンテンツ管理機能の向上に大きな期待を寄せられている。
- ・ 限られたコストで最大のパフォーマンス
各カテゴリが関係するが、ストレージそのものの価格低下、および省電力・省スペースを望む声が多い。

以上のようにストレージソリューションに対する要望・期待は数多くある。個々のハードウェアの性能、容量の伸長と制御ソフトウェアの機能の向上により、これらの要望に応えていくことになる想定している。しかし、非構造型データの効率的な管理やコンテンツの関係付けのように、ユーザが保有する大量データの管理課題を総合的に解決する手段は単なる製品レベルの向上では実現できない。これらは、ハードウェア、ミドルウェア、OS など総合的なシステムレベルの概念や技術が必要となってくる。使い勝手がよく、柔軟な管理ができ、かつコストパフォーマンスのよいストレージシステムの実現に向け、インタフェースの標準化を含めた対応が望まれている。

4. ストレージソリューションの今後

最近の 10 年において、IT システムを構成する業務用サーバ、ストレージ、及びネットワークは、その毎秒処理速度、転送速度、データ容量がギガバイトオーダからテラバイトオーダに伸長し、ペタバイトクラスの容量を必要とするデータ管理システムが話題になる状況が出現している。本章では、3 章において導出されたストレージソリューションへの期待を受け、各カテゴリにおける富士通の状況について紹介し、複数のカテゴリに跨がる大量データの管理課題を総合的に解決するための今後の方向性について考察する。

4.1. 富士通におけるストレージソリューション

富士通は総合 IT システムベンダとして、社会システムに要求されるような高品質、高性能な製品を開発・製造し、サービス・アウトソーシングを含めた幅広い IT 需要に対応している。

ストレージソリューションにおけるカテゴリと富士通カテゴリ、対応製品を対比して表 4-1 に示す。

本節では、本 WG の活動を通じて情報提供した、各カテゴリにおいて富士通が提供するストレージソリューションについて概説する。

表 4-1: ストレージソリューションと富士通製品

ストレージソリューション	富士通カテゴリ	富士通製品 <small>網掛け部分はハードウェア製品</small>	他社製品
データプロテクション	バックアップ・リストア	ETERNUS SF AdvancedCopy Manager ETERNUS SF Recovery Manager for Oracle ETERNUS SF TSM ETERNUS Disk Array	IBM/TSM Symantec/NetBackup Legart/Networker Bakbone/Netvault
アーカイブ・HSM (E-Mail アーカイブを含む)	アーカイブ	Mail Archiver ETERNUS AS500	Symantec/Enterprise Vault MailArchive, FileSystemArchive Share Point Server AXS-One Centera File Archiver 各社 ECM のアーカイブ機能
	HSM	ETERNUS Disk Array ETERNUS LT200 Tape Library	IBM/TSM-SM HPSS Sun/SAM-QFS EMC/DiskExtender
ストレージ管理	ストレージリソース管理	ETERNUS SF Storage Cruiser ETERNUS SF QM ETERNUS SF Disk Space Monitor ETERNUSmgr ETERNUS マルチパスドライバ ETERNUS VS900	Symantec/CommandCentralStorage EMC/Control Center IBM/TotalStorageControlCenter
データアクセス管理			
デバイス管理			
データレプリケーション	レプリケーション	ETERNUS SF AdvancedCopy Manager ETERNUS SF Replicator ETERNUS SF XL-DATA/MV ETERNUS Disk Array(AdvancedCopy)	SnapMirror (NR 向け) Softek TDMF DoubleTake/NSIsoftware Symantec/NetBackup
ファイルシステム	ファイルシステム	PRIMECLUSTER GFS/ PRIMECLUSTER GDS Parallelnavi SRFS ETERNUS NR1000F	IBM/GPFS Sun/ZFS Symantec/ Storage Foundation
その他	暗号化、 データ圧縮など		

4.1.1. データプロテクション

富士通が提供する ETERNUS ディスクアレイはデータの保護機能を備えた大容量キャッシュにより、高速応答性と高度なデータ保護を両立させており、社会基盤を支えるミッションク

リティカルシステムでの使用に耐える高性能、高品質を備えている。

富士通は、ETERNUS ディスクアレイに搭載したコントローラの高機能化を進めており、柔軟で高速なデータバックアップ処理を支援できるアドバンスド・コピー機能を提供している。ETERNUS SF AdvancedCopy Manager は、Symfoware、Oracle、Microsoft SQL server などの各種データベースシステム(DBMS)と連携しデータ保全プロセスを実現する富士通が提供しているバックアップソフトウェアである (図 4-1 参照)。

各種 DBMS は、個々のトランザクションを保持し運用の継続性を維持するとともに、万が一のシステムトラブルに備えたデータバックアップと迅速なリカバリのためのソフトウェア製品やソリューションを用意しており、最も実績がある分野である。

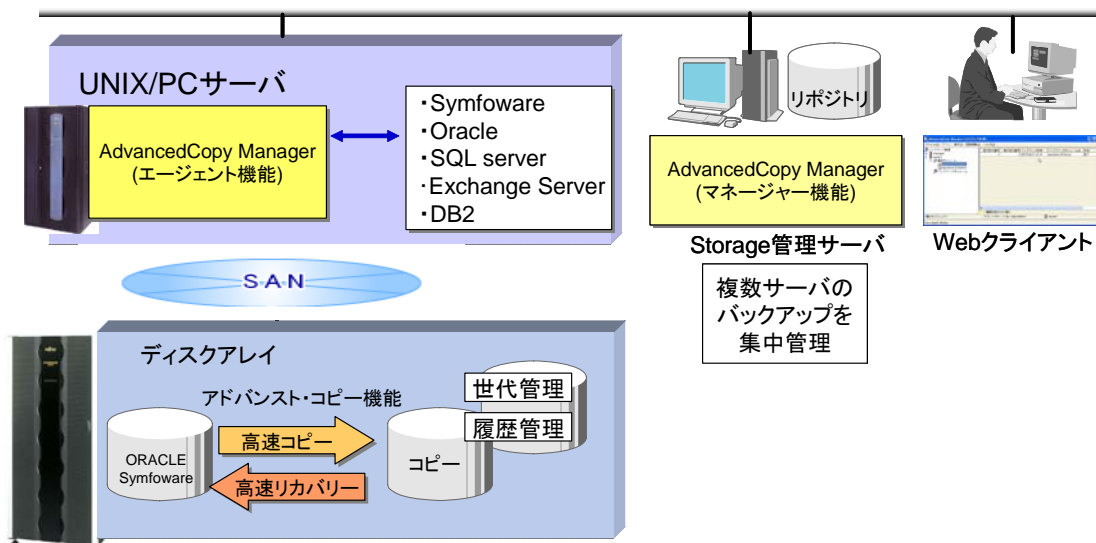


図 4-1: アドバンスド・コピーによる高速バックアップ

今日、ファイルシステムに格納される非構造型データにおいて著しい増大傾向が見られ、ディスク・ボリュームごとのデータ保全から個々のファイルを対象としたデータ保全プロセスとして、OS 連携を含めたプロセスの自動化が必要となっている。たとえば Microsoft の VSS³⁶ のようなデータ保全に関する統合化された OS インタフェースのサポートはストレージソリューションとして必要である。

4.1.2. アーカイブ・HSM

富士通のストレージソリューションとしては、ETERNUS AS500 アーカイブストレージとこれに対応したアーカイブデータ管理を必要とするメールアーカイブをはじめとした業務ソフトウェア群がある (図 4-2 参照)。

ETERNUS AS500 アーカイブストレージはコンテンツ管理型ストレージ(CAS)であり、コンテンツの二重化によるデータ保護、WORM³⁷ 構造およびハッシュ値によるデータ変更の検出、製品世代を超えたデータ移行機能備えている。特に、ポリシーベースのファイルシステムアーカイブソフトウェア(Symantec Enterprise Vault Filesystem Archiving など)との連携は非構造型データの長期保存に有効なストレージソリューションである。

一方、HSM については、日本国内において R&D ユーザで数 100TB 規模でのサポート実績があり、SAM-QFS、Symantec Enterprise Vault Filesystem Archiving などのデータアーカイブを含む HSM ソフトウェアを用いて対応している。

³⁶ VSS: Volume Shadow copy Service の略。Windows Server のストレージ管理機能として Snapshot データの採取と連携ミドルウェアの前後処理を規定したサービス機能。

³⁷ WORM: Write Once Read Many の略。一度書き込まれたデータの消去・変更ができない追記型の記憶メディアを指す。

であるが、仮想化されたサーバや業務データの特性に則したマネジメントの自動化については、今後の発展が期待される分野である。

4.1.4. データアクセス管理

SAN における FC や i-SCSI による接続制御はマルチパスドライバソフトウェアを提供している。Windows などにおいては、ドライバソフトウェアは MPIO として標準化されてきている。富士通では、ETERNUS ディスクアレイ向けに、これらのドライバソフトウェアを提供している。

NAS においては、NFS、CIFS プロトコルによりデファクトスタンダード化されているが、広域ネットワークでのファイル共有では、ネットワーク負荷の増大とレスポンスの低下が懸念されるため、WAFS³⁸などの広域ネットワーク共有における負荷低減を狙ったプロキシ制御を実現する技術開発が行われている。ETERNUS NR1000 ネットワークディスクアレイはこれらのアクセスプロトコルをサポートしている。

また、IP ネットワークを経由したデータアクセスにはエージェントソフトウェア経由でのデータアクセスがある。非構造型データを対象としたマネジメントの一つとしてコンテンツ管理型のデータアクセスを制御するエージェントは Web ブラウザによる対話型インタフェースのものと API ライブラリによる組み込み型がある。ETERNUS AS500 アーカイブストレージは API ライブラリをサポートしている。

4.1.5. デバイス管理

富士通はストレージデバイス管理機能として、以下の三つの仕組みを提供している。

ETERNUS SF Storage Cruiser(ストレージ管理ソフトウェア)は運用管理者向けの機能であり、ディスクアレイの制御インタフェースにアクセスすることで、アクセスパス、チャネル、スイッチ、コントローラの状態を監視することが可能である。

ETERNUS ディスクアレイの設定ソフトウェアである ETERNUSmgr は ETERNUS ディスクアレイの内部制御インタフェースにアクセスする GUI ソフトウェアであり、ETERNUS ディスクアレイ内部の詳細な設定を参照、変更することができる。本機能はシステム設計の反映やシステム構成を変更する場合に使用することが目的である。

富士通の提供する REMECS は公衆回線を経由した IT 機器の運用監視サービスであり、24 時間監視体制のインフラとなっている。

4.1.6. データレプリケーション

データレプリケーションの方法として、富士通では 3 種類の方法を用意している。

第一の方法は、ETERNUS ディスクアレイのアドバンスド・コピーにより、データを複製するものであり、ETERNUS SF AdvancedCopy Manager のレプリケーション指定により採取することができる(図 4-5 参照)。この方式では、ファイバチャネルなどの高速ネットワークを使用したデータ転送による高性能や OPC³⁹のような仮想的データ制御により、高速なデータレプリケーションを実現している。

³⁸ WAFS: wide area file services

³⁹ OPC: One Point Copy の略。必要になった時点で業務ボリュームの全データを秒単位の短時間で複製ボリュームにコピーする機能。

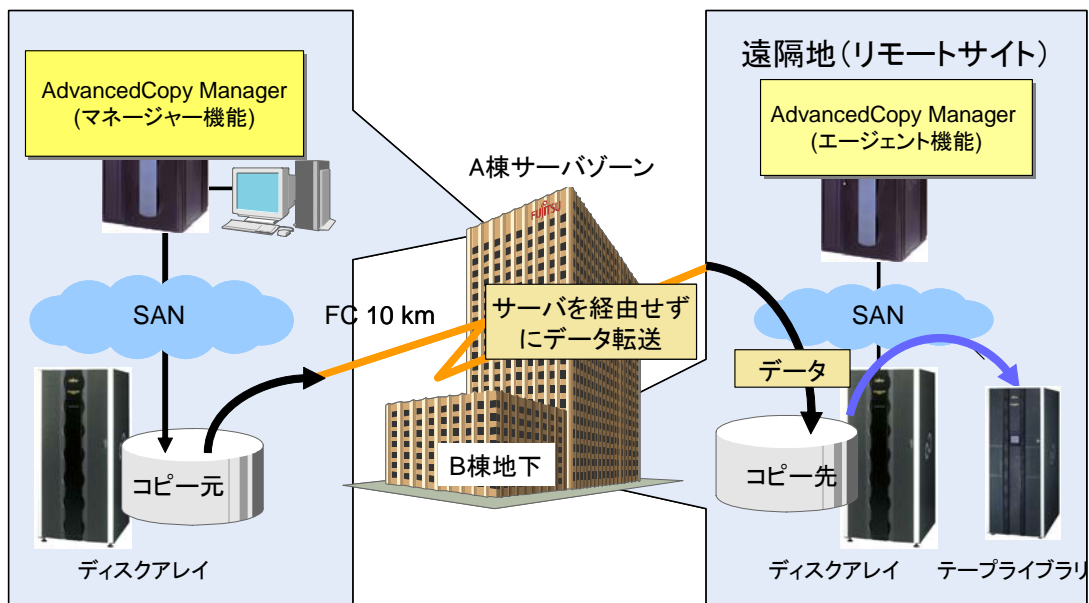


図 4-5: リモートコピーによる遠隔地筐体間コピー

第二の方法は、ETERNUS SF Replicator によるサーバのドライバレベルのソフトウェアコピーであり、コピー元の業務運用中にスナップショットデータを採取することができる。データ転送経路が LAN であり、データ転送手順がサーバで実行されるなど、現状ではデータ転送性能は第一の方法に比べて低い。

第三の方法は、PRIMECLUSTER GDS や VERITAS Volume Replicator のようなボリュームマネージャソフトウェアを使用することである。

ETERNUS ディスクアレイのアドバンスド・コピー以外の方法は、LAN(Ethernet)を経由したデータ転送であり、ストレージを接続するサーバにてデータ管理と転送制御をするため、異機種ストレージ間でのデータレプリケーションが可能である。

4.1.7. ファイルシステム

富士通はファイルシステムとして、PRIMECLUSTER GFS を提供している。

PRIMECLUSTER GFS は、一貫性／整合性を保持した同時共用ファイルシステムである。特徴は、ファイルシステムを共用している各サーバにおいてファイルデータの SAN 直接アクセスによる高性能と高可用性を実現し、複数のサーバ間でファイルシステムのキャッシュ整合性を取っていることで、アプリケーションの分散化が可能な点である(図 4-6 参照)。

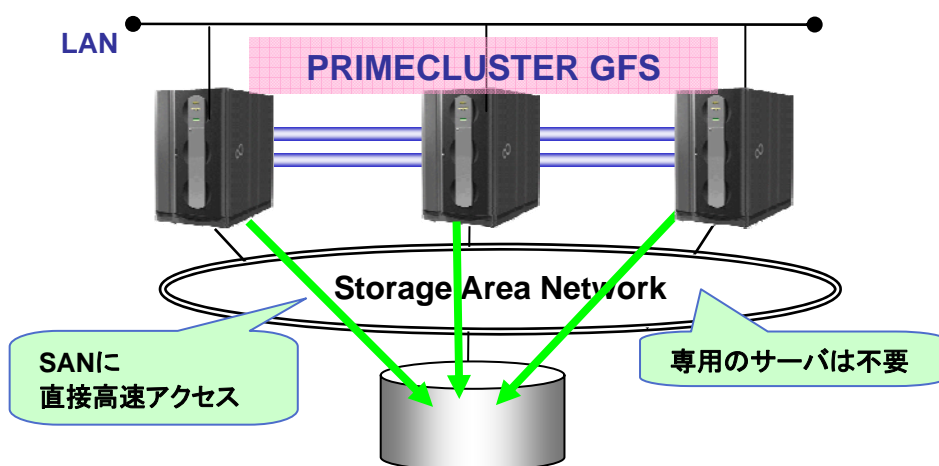


図 4-6: 富士通のファイルシステム PRIMECLUSTER GFS

また、ファイルシステムを内蔵した NAS 製品である ETERNUS NR1000F は、ストレージ管理、バックアップ支援、仮想ストレージ管理機能を持ち、統合化した運用機能を備えている(図 4-7 参照)。

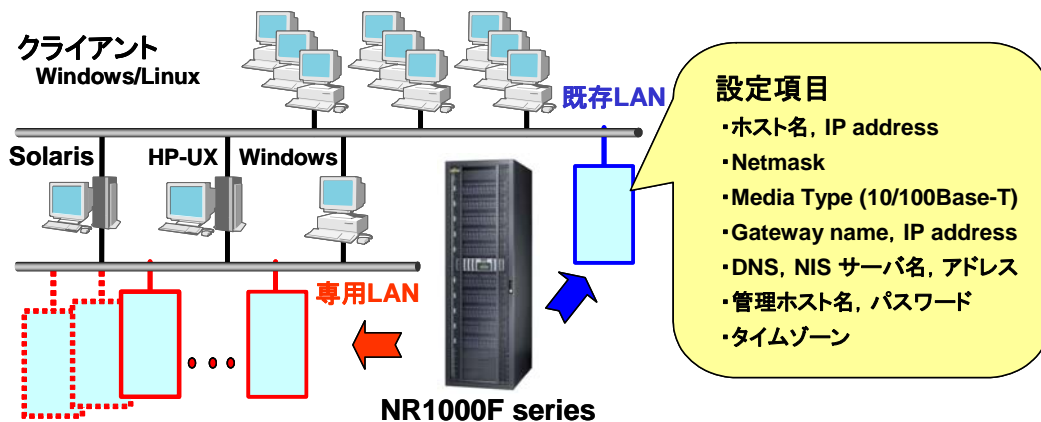
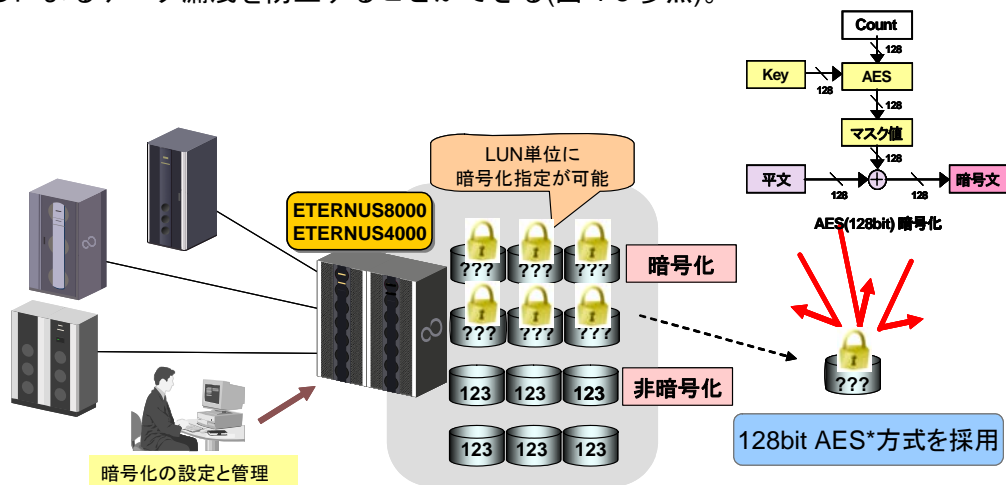


図 4-7: ファイルシステムを内蔵した NAS 製品 ETERNUS NR1000F

4.1.8. データ暗号化

富士通のストレージソリューションでは、ETERNUS ディスクアレイにおけるデータ暗号化機能により、ディスク上に保存されるデータは暗号化することができ、不用意なディスクの持ち出しによるデータ漏洩を防止することができる(図 4-8 参照)。



*AES: Advanced Encryption Standard 米国連邦情報処理標準の暗号、総務省および経済産業省が、「電子政府推奨暗号リスト」に掲載

図 4-8: 情報漏洩を防止するデータ暗号化

また ETERNUS LT シリーズ・テープライブラリ装置ではオプション製品により、LTO テープへのデータの保存時に暗号化する機能を備えている。

4.2. ストレージソリューションの方向性

本節では、3.3 節で導出された大量データの管理課題を総合的に解決するストレージソリューションの技術分野とその方向性について展望する。

大学・研究所が多く保有する非構造型データの管理を効率化する技術は ILM⁴⁰とストレージの仮想化である。

⁴⁰ ILM: Information Lifecycle Management (情報ライフサイクル管理)

ILM は情報やデータの価値を技術的に測定し、その価値に見合うストレージサービスに対応させる、ある種のバランスマネジメントであり、データアクセススピード、データ保護・信頼性機能などのストレージサービスの選択と情報管理の自動化機能の導入によって構築される。

しかし、このような情報管理システムの実現については、関連するシステムサービス・設備投資の方針に基づく、管理機能の継続的な整備が必要である。さらに、データ管理、ファイル管理を超えるコンテンツ管理のレベルで情報を管理する必要性も、以下の会員調査結果からも明らかになっている。

- ・ 意味のある情報を構成する複数のファイルを一括して処理するコンテンツという概念で管理したい。
- ・ コンテンツは情報として共有されるものであり、ネットワークアクセスが前提であり、データの格納場所、格納装置については意識することなく透過的にアクセスしたい。すなわち、データ配置の仮想化が重要である。
- ・ データマネジメントの自動化ツールとして、情報検索を支援するアノテーションの技術はこれから取り組むべき技術領域である。

一方、ストレージの仮想化は、ボリュームレベルの仮想化とファイルレベルでの仮想化があり、「4.1.3.ストレージ管理」で説明したように仮想ストレージ管理機能として提供されはじめている。

ボリュームレベルの仮想化機能は、ストレージ装置内の空きスペースなどの割り当ての自由度が増すため、ストレージ装置の有効利用に効果がある。また、ストレージ装置のリプレースに伴うデータ移行においても、容易に作業できるようになる。

今後、仮想ストレージ管理の Thin Provisioning 制御技術⁴¹や仮想ディスクプール制御が提供され、データの保存が必要になったときに動的にストレージ領域が割り当てられ、効率的な運用が可能になる。

ファイルレベルの仮想化機能は、Global Name Space と呼ぶ仮想的に統一化されたファイル名前空間管理機能により、複数の NAS が提供するファイルシステム空間を連結した単一のアクセスパスを提供する。Global Name Space により、ファイル管理がシンプルになることから大量のファイルを処理するアプリケーションの拡張性が高まる方向に技術開発が展開されていく。

ストレージの仮想化技術は ILM が必要とするストレージ装置の使い分けやファイルデータの再配置を容易化し、巨大ファイルシステム空間や大規模データアーカイブシステムの構築に必要な技術である。

このようにストレージの仮想化はリソースの有効活用やデータの移行ばかりでなく、ILM と組み合わせることにより、参加機関の誰もが悩んでいる人的負荷の軽減をもたらし、さらには TCO⁴²の削減に貢献できる。

⁴¹ Thin Provisioning 制御技術: ストレージ空間やファイルシステム空間がそのままストレージ装置上に割り当てられているのではなく、必要な領域を集めてストレージ空間を構成する機能

⁴² TCO: Total Cost of Ownership (総所要コスト)

5. おわりに

「データマネジメントを意識したストレージソリューション WG」の活動報告書として、ここに「データマネジメントを意識したストレージソリューション WG 検討結果報告書」をお届けする。

ストレージシステムは、どの情報システムにも搭載されている重要なシステムである。その機能は単純であり表には出てこないが、システムの性能を決めてしまうほど大きな能力と影響力を持っている。最近ではちょっとした SOHO の現場でも大容量のストレージシステムが活躍しているが、本報告書で扱うストレージシステムは特に研究所や大学といった大規模・高速・高可用と言われているシステムを対象とした。しかし、ブロードバンドの普及によりインターネットが家庭に入り込んでいる状況を考えると、一般ユーザもすぐに「ストレージシステム管理・運用の悩み」に直面するに違いない。それは、家庭用ストレージシステムの低価格化と、そこで扱うデータのほとんどが本報告書でも扱ったメールや Web のような非構造型データであるためである。もしかしたら「あの時使ったファイルはどこか？」とか「〇〇と△△は関連しているファイルなのか？」といった大量のファイル要求の嵐にすでに SOHO や一般家庭が巻き込まれているのかもしれない。

本 WG では、普段からストレージシステムの管理・運用に悩んでいる 3 研究機関と 2 大学を題材に、ストレージシステムの設計段階からデータ廃棄までの一連のマネジメント業務と機能の検討を行った。この中には近未来の一般家庭のストレージシステムに当然のように搭載されるであろう機能もあるだろうし、その逆に、その時代に応じた最先端の研究用途にのみ適用可能な機能もあるだろう。しかし、この段階でどのようなストレージシステムにどのようなマネジメントシステムが適合するかを検討するのは大変重要なことである。それは、ストレージシステムが日々進化、高性能化しており、その方向性をいまこそ決定していく必要があるためである。もう一つ高い視点から、すなわち「データマネジメントを意識したストレージソリューション」をいま検討すべきと考えたからである。

また、本 WG はサブ WG 活動として「ストレージシステム設計・導入にあたってのガイドライン」をとりまとめ公開した。本報告書の別冊として添付するこのガイドラインは現在のストレージシステムの管理・運用に携わる技術者のために編纂したものである。特にこれから最新のストレージシステムに挑もうとしている若い技術者に対してより実践的に利用してもらえよう、できるだけ新しい技術と用語を盛り込んだ。時間が経てばその内容は古くなってしまふかもしれないが、各項目のコンセプトはぜひ読み取っていただきたい。サブ WG に参加していた時はあまり気にしていなかったが、今思うとこれもいまますべきことのひとつだったと思っている。

さて、この WG は多くの方に支えられて活動することができた。前 WG の会員側まとめ役だった松澤 照男先生(北陸先端科学技術大学院大学)には忙しい中、毎回 WG に参加いただき貴重な意見を伺うことができた。研究所の大規模システムについて深い洞察をしていただいた水本 好彦氏(国立天文台)、黒川 原佳氏(理化学研究所)と担当幹事の鈴木 富男氏(理化学研究所)にも感謝する。また、特にサブ WG の活動を支えていただいた藤田 直行氏(宇宙航空研究開発機構)、長谷川 忍先生(北陸先端科学技術大学院大学)の活躍は素晴らしいものがあつた。さらに、最新の製品技術の紹介には森屋 光弘氏(富士通株科学ソリューション事業本部)並びに服部 和徳氏(富士通株ストレージシステム事業本部)をはじめとする各事業部の方々の協力が不可欠であつた。そして、WG 活動全体の調整をしていただいたサイエンティフィック・システム研究会事務局の平沢、築瀬、原の 3 氏にも大変お世話になった。感謝する。

読者には、本報告書を通してストレージシステムとそのマネジメントの重要性をあらためて認識していただくことに期待したい。

磯 直行(中京大学)

添付資料

- ・ データの整理分類のためのワークシート

データの整理・分類のためのワークシート

○宇宙航空研究開発機構(JAXA)

No.	機関名	保有データ		データ管理の特性										データ発生から消去までの管理	備考	
		データ種類	説明	構造/非構造	発生頻度	保存期間	原本の保障性	管理者	共有	DBMS	ミドルウェア	年間容量	保存媒体			現在容量
1	JAXA	公開用Web	公開用Webコンテンツ(静的)	非構造	1回/週程度の更新	千差万別	元データがFirewall内サーバに有り	コンテンツ作成者	単一サーバを参照	なし	なし	不明(数GB程度?)	ディスク	不明	コンテンツ作成者がそれぞれの分担を管理	過去のコンテンツの保管はこのシステム上ではなく、コンテンツ作成者が各自のローカルな環境にバックアップする。
2	JAXA	数値シミュレーションデータ	スパコンの入出力データ	非構造	千差万別	千差万別	仕組みは無し	ジョブの所有者	スパコンシステム内では共有可能	なし	なし	不明	HSM(ディスク+テープ×2)で保存。上記に加えて、ローカル環境にコピーする場合もある。	300TB	ジョブの所有者が管理	HSMのポリシー管理はスパコン管理者が実施。
3	JAXA	実験データ	風洞試験データ、飛行試験データ、観測データ等	非構造	1回/週程度か?(厳密には不明)	不明(今までの実験データは削除せずテープにアーカイブして保管している)	仕組みは無し	実験実施者	実験計測システムとアーカイブシステムはファイルシステムのクロスマウントにより共有を実現。	なし	無し(但し、市販アプリケーションのインターフェースを通じて管理)	30TB程度	HSM(ディスク+テープ×2)で保存。実験測定直後は、実験計測システムの高速度RADIディスクに保管している。	100TB程度	実験実施者	ネットワークファイルシステムがローカルの高速ディスク並みに高速であれば、直接書き込みも考えられるが、現時点では、実験計測システムのデータ出力速度に耐えられるネットワークファイルシステムが見当たらない。
4	JAXA	電子メール	業務遂行上のメール	非構造	12000通/日程度	千差万別	仕組みは無し	メールの受信者/メールサーバ管理者	無しIMAPによるメール処理は共有とどう?	なし	なし	不明	ハードディスク。ローカル環境にバッチ処理でコピーして保管。	300GB以上	各人で管理	

○国立天文台

No.	機関名	保有データ		データ管理の特性										データ発生から消去までの管理	備考	
		データ種類	説明	構造/非構造	発生頻度	保存期間	原本の保障性	管理者	共有	DBMS	ミドルウェア	年間容量	保存媒体			現在容量
1	国立天文台 ハワイ観測所	すばる望遠鏡観測データ	観測データおよび望遠鏡エンジニアリングデータ(STARS)、環境・気象データ	コンテンツ管理	常時	永久	バックアップ有、一部のエンジニアリングデータはバックアップなし	計算機・データ管理部門	Webベースでデータ検索/要求	ORACLE	UIは独自開発、ミドルウェアとしてSafeFile/Global(PetaServe)を使用	10TB	Disk+Tape(PetaSite)	60TB	消去せず、Read-Only	バックアップは山頂でテープ保存および三鷹MASTARSが該当する。
2	国立天文台 (三鷹)	すばる望遠鏡観測データ	観測データおよび望遠鏡エンジニアリングデータの一部(MASTARS)	コンテンツ管理	常時	永久	バックアップ有	計算機・データ管理部門	Webベースでデータ検索/要求	ORACLE	UIは独自開発、ミドルウェアとしてSafeFile/Global(PetaServe)を使用	5TB	Disk+Tape(PetaSite)	30TB	消去せず、Read-Only	バックアップは山頂でテープ保存およびハワイ観測所STARSが該当する。
3	国立天文台 野辺山 宇宙電波観測所	野辺山45m観測データ	観測データ	コンテンツ管理	常時	永久	バックアップ無	計算機管理部門	複数サーバから共有	ORACLE	独自にAPIを提供	2.0TB	Disk+Tape(STKライブラリ)	3.7TB	消去せず、Read-Only	
4	国立天文台 野辺山 宇宙電波観測所	野辺山10m観測データ	観測データ	コンテンツ管理	常時	永久	バックアップ無	計算機管理部門	複数サーバから共有	ORACLE	独自にAPIを提供	1.0TB	Disk+Tape(STKライブラリ)	4.7TB	消去せず、Read-Only	

○中京大学

No.	機関名	保有データ		データ管理の特性										データ発生から消去までの管理	備考	
		データ種類	説明	構造/非構造	発生頻度	保存期間	原本の保障性	管理者	共有	DBMS	ミドルウェア	年間容量	保存媒体			現在容量
1	中京大学	個人用データ	個人用ホーム領域(学部学生)	非構造	適宜, 特に講義時	4年	あり	NW管理者	複数サーバ間で共有	ディスククォータによる使用上限設定	なし	合計 1400GB	Disk+Tape (LTOライブラリ)	合計 510GB	4年で消去	
2	中京大学	個人用データ	個人用ホーム領域(教員・教員研究用・大学院生用)	非構造	適宜	退職・修了まで	あり	NW管理者	複数サーバ間で共有	なし	なし	160GB	Disk+Tape (LTOライブラリ)	合計 83GB	退職後特に依頼が無ければ1年で消去	
3	中京大学	講義課題用データ	講義用課題提示・レポート提出用領域	非構造	適宜, 特に講義・レポート作成時	1年	あり	NW管理者	複数サーバ間で共有	ディスククォータによる使用上限設定	なし		Disk+Tape (LTOライブラリ)		1年で消去	個人用ホーム領域(教員, 教員研究用, 大学院生用)と共用 高可用性が必要
4	中京大学	個人用データ	個人用Profile領域(学生・教員共用)	非構造	適宜, 特に講義時	退職・卒業まで	なし	NW管理者	クライアントへ共有	なし	なし	100GB	Disk	60GB	学生: 4年で消去 教員: 退職後消去	
5	中京大学	テンポラリー領域	テンポラリー領域	非構造	適宜, 特に講義時	次回PowerONまで	なし	ユーザ	なし	なし	なし	7.8GBx390 +30GBx95	Disk	(7.8GBx390 +30GBx95)	各ユーザが管理	各PCに用意されているテンポラリーディスク(Dドライブ). 講義中に発生する中間データ一時保存用
6	中京大学	メールスプール	メールスプール(学生・教員共用)	非構造	適宜 2000通~5000通/日程度	退職・卒業まで	なし	NW管理者	複数サーバ間で共有	なし	なし	55GB	Disk	8GB	学生: 4年で消去 教員: 退職後特に依頼が無ければ1年で消去	

○北陸先端科学技術大学院大学(JAIST)

No.	機関名	保有データ		データ管理の特性										データ発生から消去までの管理	備考	
		データ種類	説明	構造/非構造	発生頻度	保存期間	原本の保障性	管理者	共有	DBMS	ミドルウェア	年間容量	保存媒体			現在容量
1	JAIST	ホーム領域	教員・学生用ホーム 事務職員用ホーム	非構造	随時	所有者アカウントが存在する期間	バックアップ有	情報科学センター	WS, TS, 並列計算機等から共有 事務用TSから共有	特になし	特になし	6.5TB	Disk	30.5TB	所有者アカウントが存在する間は保存, アカウントがなくなると圧縮して保管	[fs3,fs8: 統合ストレージシステム] バックアップは約3ヶ月分
2	JAIST	テンポラリー領域	申請により作成される研究用データ領域/申請なしで利用できるデータ領域	非構造	随時	申請期間/不定	バックアップ有/無	情報科学センター	WS, 並列計算機等から共有	特になし	特になし	7TB	Disk	11TB	申請期間は保存	[fs1,fs3: 統合ストレージシステム] 必要に応じ人手で容量追加
3	JAIST	アプリケーション領域	WS用アプリケーション領域	非構造	ほぼ更新無し	長期間	バックアップ有	情報科学センター	WSから共有	特になし	特になし	0.1TB	Disk+Tape	0.5TB	ほとんど更新無し	[fs9] バックアップは約3ヶ月分
4	JAIST	ワーク領域	計算データ	非構造	随時	不定	バックアップ有	情報科学センター	なし	特になし	特になし	3TB	Disk+Tape	9TB	あふれそうになったらお知らせ	[個別Disk]
5	JAIST	サーバ領域	外部公開用Web, FTPサーバ	非構造	随時	所有者アカウントが存在する期間	バックアップ有	情報科学センター	なし	特になし	特になし	7TB	Disk+Tape	8TB	あふれそうになったらお知らせ	[個別Disk]
6	JAIST	サーバ領域	メールスプール領域	非構造	随時	不定	バックアップ有	情報科学センター	なし	特になし	特になし	0.1TB	Disk+Tape	0.5TB		[個別Disk]
7	JAIST	ログデータ領域	各種サーバから収集したログ	非構造	随時	不定	バックアップ有	情報科学センター	なし	特になし	特になし	0.5TB	Disk+Tape	0.5TB		[個別Disk]
8	JAIST	サーバ領域	講義アーカイブデータ	非構造	約10件/日	長期間	バックアップ有	遠隔教育研究センター	複数配信サーバから共有	特になし	特になし	1TB	Disk	5TB	1年後に別ディスクに移動	[個別Disk] 移動は人手

○理化学研究所

No.	機関名	保有データ		データ管理の特性										データ発生から消去までの管理	備考	
		データ種類	説明	構造/非構造	発生頻度	保存期間	原本の保障性	管理者	共有	DBMS	ミドルウェア	年間容量	保存媒体			現在容量
1	理研	ライフサイエンス	公開されている遺伝子・タンパク質データ	コンテンツ管理	適宜	1世代位	一部バックアップあり	サーバ管理者	なし	独自	なし	数TB	Disk	8TB	新規DBが公開されるまで保持 消去しない	バックアップは基本的にしない
2	理研	実験データ	高エネ物理実験データ	コンテンツ管理	適宜	永久	本データは元データのレプリカ	サーバ管理者	複製	なし	なし	約300TB	Disk+Tape	1PB		元データはUS BNLに存在する
3	理研	シミュレーションデータ	スパコンユーザの利用領域	非構造	適宜	不明	ホーム領域のみ	ユーザ	なし	なし	なし	2TB	Disk+Tape	約37TB	ユーザ次第	テープへのバックアップはホーム領域のみ。容量の内訳はディスクが約7TB、テープが約30TB

参考資料：富士通からの情報提供

- a. 新ストレージ・プラットフォーム ETERNUS8000,ETERNUS4000 ご紹介
- b. Interstage Shunsaku Data Manager ご紹介
- c. ETERNUS 性能監視 ご紹介(ESC,SQC)
- d. 科学計算向け次期ファイルシステム ご紹介(SRFS)
- e. ETERNUS2000 ディスクアレイ ご紹介
- f. ETERNUS NR1000V NAS ゲートウェイ ご紹介
- g. フィールドイノベーション
富士通のストレージシステム ETERNUS と関連ソリューション ご紹介

富士通ストレージシステム
ETERNUS


新ストレージ・プラットフォーム ETERNUS8000, ETERNUS4000 の御紹介

2006年4月6日
富士通株式会社
ストレージシステム事業本部



Fujitsu Proprietary and confidential 2006. All Rights Reserved. Copyright FUJITSU LIMITED 2006.

■ ETERNUSストレージ製品一覧



Fujitsu Proprietary and confidential 2006. All Rights Reserved. Copyright FUJITSU LIMITED 2006.

■ ETERNUS8000, ETERNUS4000開発の狙い

スケーラビリティ、ビジネスの継続性

- 世界最大容量
- 世界最高性能
- 卓越した信頼性の提供 (冗長化、活性保守、予兆検出)

データ保全性、セキュリティ強化

- バックアップ機能の強化
- 遠隔災害対策 (ディザスタリカバリ対応強化)
- 情報漏洩を防止する「データ暗号化」

柔軟な運用管理、TCO削減

- LUN動的拡張、仮想ストレージ対応
- ストレージ統合、接続性強化

Fujitsu Proprietary and confidential 2006. All Rights Reserved. Copyright FUJITSU LIMITED 2006.

■ 製品ラインナップ

Model	ETERNUS8000	Model 700	Model 900	Model 1100	Model 500	Model 80	Model 100	Model 300	ETERNUS4000
タイプ	エンタープライズ・ストレージ	ミッドレンジ・ストレージ	ミッドレンジ・ストレージ	エンタープライズ・ストレージ	ミッドレンジ・ストレージ	ミッドレンジ・ストレージ	ミッドレンジ・ストレージ	ミッドレンジ・ストレージ	ミッドレンジ・ストレージ
ドライブ数	2,760	80	480	1,020	240	15	30	120	240
記憶容量	828TB(FC) 1,365TB(NL)	4.5TB(FC) 7.5TB(NL)	144TB(FC) 236TB(NL)	306TB(FC) 502TB(NL)	72TB(FC) 120TB(NL)	4.5TB(FC) 7.5TB(NL)	9TB(FC) 15TB(NL)	36TB(FC) 60TB(NL)	72TB(FC) 120TB(NL)
コントローラ	4/6/8/12/16CPU	2(2CPU)	2(4/8CPU)	4(8CPU)	2(2CPU)	2(2CPU)	2(2CPU)	2(2CPU)	2(2CPU)
キャッシュ	64-256GB	4-16GB	8-64GB	32-128GB	8-32GB	1GB	2GB	4/8GB	16GB
バス数	32-128	4-16	4-64	16-64	16	FC4G: 4	FC4G: 4	8	4/8
接続性	FC4G: 4-64, FCLink: 4-64	FC4G: 4-8, FCLink: 4-8	FC4G: 4-32, FCLink: 4-32	FC4G: 16-64, FCLink: 4-32	FC4G: 4-16, FCLink: 4-8	FC4G: 4	FC4G: 4	FC4G: 2-8, FCLink: 2/4	FC4G: 4-16, FCLink: 4-8

FC: Fibre Channel disk, NL: Near Line disk

Fujitsu Proprietary and confidential 2006. All Rights Reserved. Copyright FUJITSU LIMITED 2006.

富士通ストレージシステム
ETERNUS

スケーラビリティ、ビジネスの継続性

データ保全性、セキュリティ強化

柔軟な運用管理、TCO削減

Fujitsu Proprietary and confidential 2006. All Rights Reserved. Copyright FUJITSU LIMITED 2006.

■ 世界最大容量、最高速のストレージシステム

1PBを超える世界最大容量の実現

エンタープライズ・ストレージ

ETERNUS8000 モデル700, 900, 1100, 2100

- 最大1,365TB (2,760ドライブ) 従来: 1,020ドライブ
- 最大256GB/512GBキャッシュ 従来: 128GBキャッシュ
- 最大128FCバス (4Gbps) 従来: 64FC
- 最大64FCLinkバス 従来: 32FCLinkバス

ミッドレンジ・ストレージ

ETERNUS4000 モデル80, 100, 300, 500

- 最大120TB (240ドライブ) 従来: 240ドライブ
- 最大32GBキャッシュ 従来: 16GBキャッシュ
- 最大16FCバス (4Gbps) 従来: 8FCバス

Fujitsu Proprietary and confidential 2006. All Rights Reserved. Copyright FUJITSU LIMITED 2006.

■世界No.1性能の確保: 処理IO数

ETERNUS8000
ETERNUS4000

単位時間処理IO数(IOPS)世界最高

SPC-1 (Storage Performance Council 1) 目標値

ETERNUS8000 model 2100: 300,000 iops
ETERNUS8000 model 1100: 150,000 iops
ETERNUS6000 model 1100: 108,745 iops
ETERNUS4000 model 500: 60,000 iops
ETERNUS3000 model 700: 41,202 iops

エンタープライズ・ストレージ
ミッドレンジ・ストレージ

http://www.storageperformance.org/results

高性能化により、より大規模なストレージ統合や通常業務とバックアップ業務の多重処理に対応

■性能・拡張性を担う先端技術

ETERNUS8000
ETERNUS4000

高性能CPU構成による卓越したIOPS性能

- ETERNUS8000 : CM当り3.6GHzの2CPU構成、装置当り最大16CPU
- ETERNUS4000 : CM当り2.0GHzのSingle or Dualコア、装置当り2 or 4CPUコア

スケラビリティを実現するモジュラーアーキテクチャ

- CMとCA/DAを一体化したモジュール構造によりデータ転送処理を高速化
- ETERNUS8000 : 最大8CMをFRT(PCI Expressスイッチ)経由で接続
CMとディスク間をBRT(FC Fabricスイッチ)経由で接続
- ETERNUS4000 : CM間を2GB/s × 2の高速専用バス(PCI Express)で相互接続

CA: Channel Adapter, CM: Control Module, DA: Device Adapter, FRT: Front-end Router, BRT: Back-end Router, PBC: Port Bypass Circuit

■最新ディスクドライブの採用

ETERNUS8000
ETERNUS4000

高性能オンラインFCディスク
(15,000 rpm)
FC 4 or 2 Gbps

- ディスク容量: 146 GB, 73 GB, 36 GB
- 円板/ヘッド数: 4枚/8個, 2枚/4個, 1枚/2個

大容量オンラインFCディスク
(10,000 rpm)
FC 2 Gbps

- ディスク容量: 300 GB, 146 GB, 73 GB
- 円板/ヘッド数: 4枚/8個, 2枚/4個, 1枚/2個

ニアラインFCディスク
(7,200 rpm)
FC 2 Gbps

- ディスク容量: 500 GB
- 円板/ヘッド数: 4枚/8個

低価格でバックアップとアーカイビングに最適

■冗長構成と活性保守による卓越した信頼性

ETERNUS8000
ETERNUS4000

- 完璧な冗長化構成
 - コントローラ部のロジック、電源、ファン等、全てのコンポーネントを多重化
 - ドライブ部のコンポーネントを二重化
 - コントローラ部とドライブ部の接続を多重化
- 活性交換・活性増設
 - ハードウェアの主要コンポーネント全てが対象
 - ファームウェアの交換もシステム移動中に可能

■RAID6(デュアルパリティ)のサポート

ETERNUS8000
ETERNUS4000
OPEN稼働

RAID6,5構成を各DEに分散して、信頼性をさらに強化

- RAID6構成をサポート
 - 同一RAID内のHDD2本故障を救済
 - 推奨構成は6D+2P
 - 14D+2Pもサポート
- RAID5構成の拡張
 - RAID5を独立したDEに分散
 - 3D+1Pと7D+1Pの選択が可能(従来は3D+1Pのみ)
 - ETERNUS4000では4D+1Pを推奨

RAID5	RAID6
リード性能比	1 : 1
ライト性能比	1 : 0.67
総合性能比	1 : 0.92
(R/W比 3:1)	

富士通ストレージシステム ETERNUS

スケラビリティ、ビジネスの継続性

データ保全性、セキュリティ強化

柔軟な運用管理、TCO削減

■ニアラインFCディスクを活用したD2Dバックアップ

ETERNUS8000
ETERNUS4000
OPEN専用

同一筐体内で、D2Dバックアップの複数世代管理

運用中ポリウム
オンラインFCディスク (146GB/300GB)
最新世代バックアップポリウム
ニアラインFCディスク (500GB)

他社ニアラインディスクよりも優れた信頼性、性能、可用性

	当社	他社
インタフェース	FC (2ポート)	SATA (1ポート)
FCディスクとのDE内接続	可	不可 (別DE要)
アクセス性能比 (シタLW/R)	1.0	0.3~0.5

D2D : Disk to Disk
OPC : One Point Copy

■用途・業務に応じたOPCの多機能化

ETERNUS8000
ETERNUS4000
OPEN専用

QuickOPC

バックアップデータを全面保持 ⇒ 故障耐性が高い

SnapOPC (Copy on Write)

コピー先容量を軽減 ⇒ 複数世代のバックアップに最適 (ハード故障復旧は不可)

■電力消費量の削減

ETERNUS8000
ETERNUS4000
OPEN専用

MAID技術の応用による省エネ対応

MAID : Massive Arrays of Inactive Disks ⇒ 未アクセスディスクの回転を停止

- 時刻設定により特定ディスクの回転停止期間をスケジュール
- 指定ディスクへのアクセスが10分(可変)以上無い ⇒ 停止状態へ
- 指定ディスクへのアクセスが発生 ⇒ 回転を起動し、約1分でレディ状態へ

活用例

データポリウム
300GB オンラインディスク 60本
RAID0+1, 7.5TB

バックアップポリウム
500GB ニアラインディスク 60本
RAID5, 22TB

三世帯バックアップ

全体(60本)を5時間のみオン

全体の電力消費量を20%削減!!

■ディザスタリカバリ(DR)機能の強化

ETERNUS8000
ETERNUS4000
OPEN専用

- ミッドレンジで非同期モードをサポート
- エンタープライズ・ストレージとミッドレンジ間の接続が可能

小規模システムでの遠距離転送を実現 (従来100km ⇒ ∞)

エンタープライズ・ストレージとミッドレンジ間の機能ギャップを解消

メインサイト (ETERNUS 8000) ↔ LAN/WAN (東京・大阪間、等) ↔ リモートサイト (ETERNUS 4000)

■ディザスタリカバリ(DR)の容易な導入

ETERNUS8000
ETERNUS4000

ネットワーク直結のディザスタリカバリ機能を提供

- ETERNUS間をiSCSIインタフェースで接続 ⇒ 高価なFCスイッチやSANゲートウェイ装置が不要
- ミッドレンジ&エンタープライズ・ストレージ、オープン&メインフレームディスクともにサポート
- IPsecを搭載し、ネットワーク上の暗号化も可能

従来 FCから直結

今後 ネットワーク直結

■情報漏洩を防止するデータ暗号化

ETERNUS8000
ETERNUS4000

ディスク媒体の不正なデータアクセスを防止

- 市場動向
 - 金融機関での動向 : 金融庁一全国銀行協会一都銀・地銀等
 - 情報保護法令の影響 : データ漏洩を防止するための保護措置が必要であり、現状はHDDのデータ消去に対応
- ストレージシステム内部でディスクの暗号化を実現 (安全性の高い128bit AES*方式を採用)

暗号化の設定と管理

暗号化 (AES(128bit)暗号化)

非暗号化

ディスク持ち出し時のデータ漏洩に対応

富士通ストレージシステム
ETERNUS

スケーラビリティ、ビジネスの継続性

データ保全性、セキュリティ強化

柔軟な運用管理、TCO削減

FUJITSU Proprietary and confidential 2006 18 All Rights Reserved, Copyright FUJITSU LIMITED 2006

ETERNUS8000
ETERNUS4000
OPEN専用

■ LUNの動的移行・拡張

LUNの移行や容量拡張に柔軟に対応

LUNの移行

LUNの容量拡張

FUJITSU Proprietary and confidential 2006 19 All Rights Reserved, Copyright FUJITSU LIMITED 2006

ETERNUS8000
ETERNUS4000
OPEN専用

■ 仮想化ストレージ機能

VS900による仮想ストレージ機能の提供

■ 必要時にストレージプールから仮想ディスクの容量を追加

サーバ

ETERNUSディスクアレイ

FUJITSU Proprietary and confidential 2006 20 All Rights Reserved, Copyright FUJITSU LIMITED 2006

ETERNUS8000

■ GS/OPENストレージ統合

GS/OPENの統合により、ハード・運用管理コストを削減

- 異種サーバによる装置共用が可能 ⇒ ハード+運用コスト(消費電力、スペース)削減
- SAN環境の共用が可能 ⇒ 柔軟なSAN構成の構築、装置移行、装置の仮想化が容易に

異種サーバのストレージ装置統合

運用ボリューム
ETERNUS8000

FUJITSU Proprietary and confidential 2006 21 All Rights Reserved, Copyright FUJITSU LIMITED 2006

ETERNUS8000

■ ストレージ統合による一元管理

GS/OPENの統合により、バックアップ、監視、管理を一元化

- バックアップ運用の統合が可能
- 監視・管理の統合が可能 ⇒ 単一Viewによるストレージの監視・管理

運用ボリューム

バックアップボリューム
ETERNUS8000

VSS: Volume Shadow Copy Service

FUJITSU Proprietary and confidential 2006 22 All Rights Reserved, Copyright FUJITSU LIMITED 2006

ETERNUS8000
ETERNUS4000
OPEN専用

■ マルチプラットフォーム対応

IBM、HPを中心に他社サーバにも対応

サポートOS

- グローバルサーバ
 - OS/IV/MSP, OS/IV/F4 MSP
 - OS/IV/XSP
 - AVM/EX, AVM/EXS
- UNIXサーバ
 - Solaris 8/9/10
 - HP-UX 11i1/11i2
 - AIX 5.1/5.2/5.3
- IAサーバ
 - Windows Server™ 2003
 - Red Hat Enterprise Linux AS/ES v4.0
 - SUSE LINUX Enterprise Server 9
 - VMware ESX Server 3.0

FUJITSU Proprietary and confidential 2006 23 All Rights Reserved, Copyright FUJITSU LIMITED 2006

■運用管理ソフトウェア

ETERNUS8000

ETERNUS4000

OPEN専用

豊富なストレージ管理ソフトウェアをラインナップ

ストレージソフトウェア基盤 ETERNUS SFシリーズ		
用途	製品名	特徴
確実な管理/監視	ETERNUS SF Storage Cruiser	ストレージシステムにおける構成管理/関係管理/障害監視/性能監視
	Systemwalker Resource Coordinator	ストレージシステムのプロビジョニング
業務影響の少ないバックアップ及びデータ移行	ETERNUS SF Advanced Copy Manager	ETERNUSが提供するOPC-EC機能を制御し、高速にバックアップを作成
	ETERNUS SF Recovery Manager for Oracle	Oracleデータベースの整合性を保証した高速バックアップと高速リストア・リカバリーを実行
	ETERNUS SF Replicator	業務を継続しつつ、装置ヘブリカを作成(データ移行)
ストレージ容量管理	ETERNUS SF TSM	各種システム構成に対応したバックアップ
高速なデータ転送	ETERNUS SF Space Advisor	SANストレージ容量のリアルタイム表示としきい値監視
	ETERNUS SF XLデータムーブ	異なるプラットフォーム間でのデータ交換

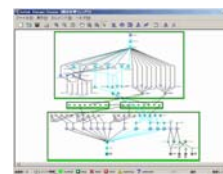
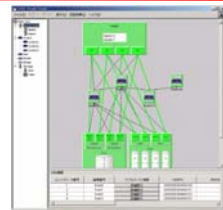
Fujitsu Proprietary and confidential 2006

24

All Rights Reserved, Copyright FUJITSU LIMITED 2006

■ETERNUS SF Storage Cruiser

- SAN環境のストレージリソース管理
 - アクセスパスの設定
 - 整合性の確認
 - 機器の状態監視と表示
- 関係管理機能
 - ストレージ上とサーバ上のリソースの対応関係表示
- 第三者研究調査機関による導入効果評価
 - 5年間で3000万円の運用管理コスト削減



◆新規機能

- ◆ETERNUS8000、ETERNUS4000サポート
- ◆仮想スイッチVS900管理機能
- ◆SQCと連携した性能分析機能

*Systemwalker Service Quality Coordinator

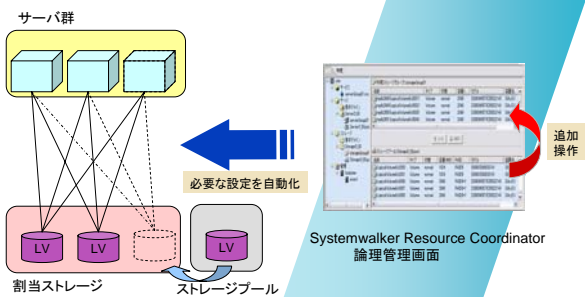
Fujitsu Proprietary and confidential 2006

25

All Rights Reserved, Copyright FUJITSU LIMITED 2006

■Systemwalker Resource Coordinator

- ETERNUSのボリュームをストレージプールとして管理
- 画面上的操作でサーバに簡単にボリューム割当可能



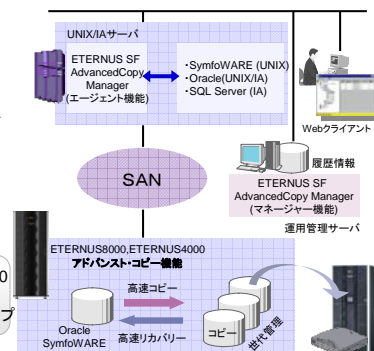
Fujitsu Proprietary and confidential 2006

26

All Rights Reserved, Copyright FUJITSU LIMITED 2006

■ETERNUS SF AdvancedCopy Manager

- 高速バックアップ管理を行い、システムの24時間365日運用を実現
- 富士通のETERNUSディスクアレイと連携し、アドバンスド・コピー機能による高速バックアップ/リストア、レプリケーション運用を実現



◆新規機能

- ◆ETERNUS8000、ETERNUS4000 SnapOPC、QuickOPCサポート
- ◆MS Exchange オンラインバックアップ

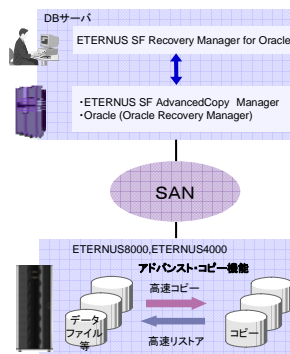
Fujitsu Proprietary and confidential 2006

27

All Rights Reserved, Copyright FUJITSU LIMITED 2006

■ETERNUS SF Recovery Manager for Oracle

- 富士通ETERNUSディスクアレイ装置のアドバンスド・コピー機能およびOracle Recovery Managerと連携することで整合性を保証したOracleデータベースの高速オンラインバックアップと高速リストア・リカバリーを実現
- 高信頼なバックアップ資産を取得可能
- 簡単・安全・確実なリカバリーを提供



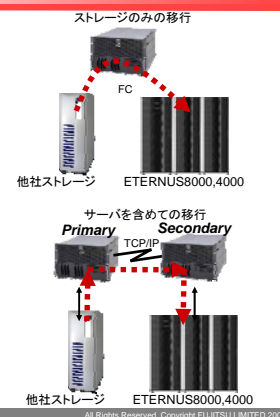
Fujitsu Proprietary and confidential 2006

28

All Rights Reserved, Copyright FUJITSU LIMITED 2006

■ETERNUS SF Replicator

- 既存資産からのデータ移行、データ複製を行うソフトウェア
- データ移行中は業務運用可能
- ストレージベンダーに依存せず、多くのプラットフォームをサポート
 - UNIX (Solaris™ OS, HP-UX, AIX)
 - Windows
- お客様の環境、目的に最適な多様な複製アプローチを提供
 - 1対1、1対多数、多数対1の実現
 - 非同期転送
 - 転送速度制御、スケジューリング



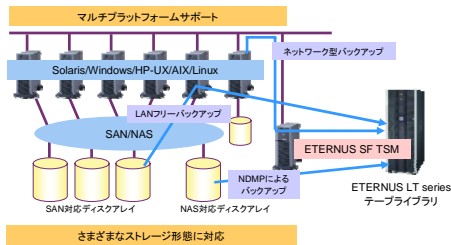
Fujitsu Proprietary and confidential 2006

29

All Rights Reserved, Copyright FUJITSU LIMITED 2006

■ ETERNUS SF TSM

- 複雑なITシステム環境の統合バックアップ管理を実現
- SAN、LAN、WANに対応して自動的なバックアップ管理を実行
- 幅広いプラットフォーム、ネットワーク、ストレージデバイスをサポートすることで、複雑なIT環境において高い柔軟性を発揮し、重要なデータを確実に保護



Fujitsu Productivity and Confidential 30 All Rights Reserved, Copyright FUJITSU LIMITED 2006

■ ETERNUS SF Space Advisor

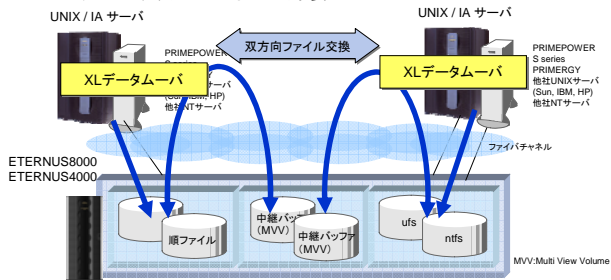
- 容量のしきい値監視
 - 監視対象となるディスクやデータベースの使用量が「しきい値」を超えた場合には、管理者へ通知
 - この通知をもとに、システム全体の重大なサービス低下を事前に予防
- リアルタイムな容量の表示
 - 収集/蓄積した容量情報の「ドリルダウン」表示、および「即時レポート」表示可能
 - 可視化されたリアルタイムな容量情報による問題発生時の状況把握をサポート
 - 空き領域の不足が発生しているサーバやファイルシステムの特長
- 定期レポートの出力
 - 収集/蓄積した容量情報を、「日報」、「週報」、「月報」などのフォーマットで定期的な出力
 - システム管理者の定常的な監視作業における報告レポートの作成を支援
 - 作成したレポートのデータをダウンロードし、分析、評価に利用可能



NASの容量監視やPDRと連携したSLA/MON/メーションにはSoftex Storage Managerを利用してください。 Fujitsu Productivity and Confidential 31 All Rights Reserved, Copyright FUJITSU LIMITED 2006

■ ETERNUS SF XLデータムーバ

- 異種サーバ間を高速にデータ転送
- データ転送処理時間を短縮(1GBのデータを2分以下で転送可能)
- サーバのCPU負荷は、FTPと比較して1/3以下
- ネットワークリソースはまったく不要



Fujitsu Productivity and Confidential 32 All Rights Reserved, Copyright FUJITSU LIMITED 2006



技術力・適切・安心
富士通のミドルウェア

THE POSSIBILITIES ARE INFINITE FUJITSU

データマネジメントを意識した ストレージソリューションWG 資料

Interstage Shunsaku Data Manager ご紹介

～あるがままに格納し、思いのままに検索する～

2006年10月12日
富士通株式会社

技術力・適切・安心

Interstage Shunsaku Data Manager

Shunsaku とは

All Rights Reserved. Copyright © FUJITSU LIMITED 2005-6

技術力・適切・安心

Shunsakuとは？

- Shunsakuはデータベースです。
- スキーマ定義やインデックス設定の必要がありません。
- 「あるがままのデータ活用」を実現します。

項目の型、サイズ、名前、順番、テーブルのフォーマットがバラバラでも自在に登録できる

検索できる項目を限定せず、どんな条件でも安定して利用できるシステム構成をとることができる

受注・売上
検査進捗
出荷情報

登録

Shunsaku

XML形式レコード

検索

業務アプリ
Web検索アプリ

AsIs (アズイズ) にデータを管理・活用

※ AsIsとはas it isの略で、Shunsakuがコンセプトとする「あるがままのデータ活用」を意味します。

All Rights Reserved. Copyright © FUJITSU LIMITED 2005-6

技術力・適切・安心

Shunsakuの特長① ～柔軟に格納する～

- あるがままの形式でデータを格納可能
 - ✓ レコード毎にスキーマが異なっていても問題なし。
 - ✓ 統合スキーマの定義不要。

多様
可変

北海道支社
東京本社
沖縄支店

項目の個数、順番、長さに制約なし

多種多様なデータを混在させられる

複数の不揃いのデータであっても、「あるがままの形」で混在して格納することができます

All Rights Reserved. Copyright © FUJITSU LIMITED 2005-6

技術力・適切・安心

事例：アフターセールス情報の統合DB

適用のポイント

- 形式の異なるデータを統合
- 業務ごとに柔軟な活用
- 更新を含めた多重実行性能を確保

要件

- 社内に散在する各種データの活用

課題

- システムにデータ形式が異なるため、統合は困難
- 業務ごとに閉じた検索
- 更新を含めた更新が困難

導入効果

- 欲しい情報を欲しいスピードで検索
- サービスの均質化やスキルアップ

データをXML化して一元集約

Interstage Shunsaku Data Manager (統合DB)

用品ニュース
技術レポート
不具合・整備情報
サービスパーツニュース
故障診断データ

ETLツール

検索

検索コンテンツ指定
ロサービスマニュアル
用品ニュース
サービスパーツニュース
故障診断データ

検索キーワード指定

車種名 フジツ
ジャンル BAISS-100*
キーワード キャンペーン 限定

All Rights Reserved. Copyright © FUJITSU LIMITED 2005-6

技術力・適切・安心

Shunsakuの特長② ～自在に検索する～

- 思うがままに検索
 - どんな項目でも同じ速度で検索。
 - どんな方式 (完全/部分一致等) でも同じ速度で検索。
 - インデックスレスのためチューニング不要。

～単純明快な検索方式～

- 全レコード、全項目を一方向・逐次ベタ読み検索
- インデックスレス
- 性能見積りが明快で分かりやすい

All Rights Reserved. Copyright © FUJITSU LIMITED 2005-6

事例：従来のRDBとShunsaku 構築・運用の違い

弊社現場SEやSIパートナー様の声から抜粋しています。

	従来型DB技術の課題 (スキーマ/正規化前提)	Shunsakuのデータ管理方式 (スキーマレス/インデクスレス)
変化する要求や 変化するデータ への対応	<ul style="list-style-type: none"> 正規化再設計、インデクス再設計が必要 場合によってはDB再構成からやり直す。 設計に膨大な時間を要する上、性能チューニングもやり直し、しかも今ある運用を止めなくてはならない。 	<ul style="list-style-type: none"> 正規化設計、インデクス作成が不要。 全てのデータ項目があるが任意に登録可能なので、DB拡張(項目追加)に容易に対応できる。 しかも、運用を止めずにDB拡張ができる。
安定性能の確保	<ul style="list-style-type: none"> 性能は実測するまで安心できない。 条件の複雑化やアクセス増大で性能が極端に劣化する事がある。 DBを作ってみても、試行錯誤を繰り返して最適化しなければならない。 	<ul style="list-style-type: none"> 性能チューニング不要で、事前に性能の予測がつく。 複雑条件、アクセス増大に対しても、安定したレスポンスが期待できる。
運用(メンテナンス)の 容易性	<ul style="list-style-type: none"> 定期的なDB、インデクスのメンテナンスが面倒。 データ量の増大に対応するには、物理配置再設計、性能設計のやり直しや、データの移行が必要。 	<ul style="list-style-type: none"> メンテナンスが楽である。 データ量の増大に対しては、ハードウェアの追加のみで対応が可能となった。 しかも、運用を止めない活性増設である。

Shunsakuの特長⑥ ~データの保全性を確保~

一連の更新処理をCommit/Rollbackにより確定・取消し制御

Shunsakuの特長⑦ ~データベースの信頼性~

任意時点・最新状態への復旧
操作ミスなどによるデータ消失時でも最後のバックアップ以降の任意の時刻を指定してのリカバリが可能

Interstage Shunsaku Data Manager 事例

事例：デジタルアーカイブソリューション

適用のポイント

- 学芸員・研究員グループ毎に異なる収蔵資料情報の管理項目を複数定義。(例: 文書と考古それぞれ独自の設定が可能)
- 検索条件を細かく設定し、どんな検索条件に対してもレスポンスを悪化させることなく検索。

課題

- せっかくの収蔵資料をもっと活用したい
- 論文やデジタル写真など多様な情報を一元管理したい

導入効果

- 収蔵資料管理項目の追加、変更時にデータベースの作り直し不要
- 複雑な項目での検索や収蔵点数の増加時もチューニング不要

検索条件入力画面 | 検索結果一覧画面

複雑な条件を入力し、高度な検索 | 検索結果は、テキストファイルに出力(エクスポート) | 検索結果は、一覧形式またはサムネール一覧で表示

やりたくてもできなかった統合データベースを短期構築

既存資産をそのまま活用し、目的・用途の異なる複数データの統合を短期間で構築

事例：受注データ統合

異なる形式の受注データをXML形式で統一
頻繁に変わるデータ形式に容易に対応

販売管理
情報分析

Shunsaku

生ログ形式で受け付け格納
XMLデータ
標準フォーマットで取だし

受注データは各ショップ毎にバラバラのCSV形式
統一フォーマットに変換！

しかも頻繁に変更

18

事例：商品情報の統合

旧システムへの不満
「各情報単体の検索しか出来ない」
「検索速度が遅い」
「型名一括・複数条件による検索が出来ない」等

商品検索システム
利用者の不満解消!! 大好評!!

検索レスポンスの大幅な改善を実現!! どんな検索もほぼ1秒

全製品の型名価格と供給量、スペック情報を1システムで検索可能!!

データ統合

Interstage Shunsaku Data Manager

【システム構築者のメリット】
既存DBの管理していた項目名やテーブル形式がバラバラでも、そのままShunsakuへの登録が可能
複雑な条件での検索や、アクセスが集中しても、レスポンスをほぼ一定に保つことが可能

19

事例：病院内の現場の見える化

今までやりたくてもできなかった複数システムの横断的なデータ活用を可能にした「あるがままにデータを格納」スクラップ&ビルドの容易さ

● 分析したいDBの知識はあまりない (難しい設計しないと思えないDBは困る)
● 本格的なDWH+BIシステムへの思い切った投資をする前に、効果を見たい
● 欲しい情報を持ってきて自分で実行確認できる

● 業者への不要な支払いの削減
データのレイアウトを公開することにより各々で必要なデータの分析・取得ができ、メーカー(業者)へ依頼していた費用の削減が可能

● 横断的な情報活用による新たな気づきと早期対応
今までシステム毎に管理の範囲しか参照・分析ができなかったが、広範囲・他角度からのデータに隠れている新たなデータの関係の発見を早く行うことができる

● 新たな視点での分析による経営指標の見直し
今まで既成システムパッケージではできなかった経営の指標となる新たなデータ間の関係の発見を早く行うことができる

● ノウハウの価値を高め、リスクを未然に回避
リスク情報やナレッジの蓄積・公開を行うことにより、スキルが不足している職員への情報提供ができる

院内情報統合サーバ
Interstage Shunsaku Data Manager

見える化サーバ

20

ご参考：その他の適用例

業種	適用場面		
製造	構成表(レシピ)管理	FAQシステム	コールセンター
	電子精密部品調達	住宅プランニング支援	顧客管理
	副資材購買	営業支援	電子電話帳
	共通部品管理	保守点検管理	商品検索
流通	品質情報管理	医薬品情報/安全性情報管理	メール監査
	トレーサビリティ	ソースコード検索	権限資産管理
文教	商品特長検索	商品検索	購買分析
	伝票検索	原材料・仕込み管理	顧客管理
官庁・自治体・公共	大学情報DB	歴史資料データベース	収蔵管理システム
	求職情報検索	旅券発給システム	企業情報検索
金融・保険証券	特許・実用新案検索	電子調達	会員情報検索
	ゲノム検索	電子戸籍	大学情報検索
情報・サービス	マニュアル検索	顧客情報管理	ビジネスマッチング
	デジタルアーカイブ	ロコミサイト(レストランガイド)	記事検索
電力・ガス	クリッピングサービス	商品検索	商品検索
	リース履歴管理	営業支援	設備管理
			監査証跡

21

第三者機関による評価

(国内) ソフトウェア・プロダクト・オブ・ザ・イヤー

独立行政法人情報処理推進機構(略称:IPA、理事長:藤原 武平太氏)、および財団法人ソフトウェア情報センター(略称:SOFTIC、理事長:安西 邦夫氏)が2004年8月28日に発表した「ソフトウェア・プロダクト・オブ・ザ・イヤー2004」システム分野において、2004年9月10日に表彰されました。

独立行政法人情報処理推進機構(IPA) : <http://www.ipa.go.jp/>
財団法人ソフトウェア情報センター(SOFTIC) : <http://www.softic.or.jp/>

SOFTWARE PRODUCTS OF THE YEAR 2004
ソフトウェア・プロダクト・オブ・ザ・イヤー

(海外:中国) 優秀ミドルウェアソリューション設計賞

中国でもっとも知名度が高いITサイトであるComputer Worldが、2004年10月28日に開催した優秀ミドルウェア製品評価において、100以上の製品の中からTOP100に選ばれ、<優秀ミドルウェアソリューション設計賞>を受賞しました。


22

FUJITSU

THE POSSIBILITIES ARE INFINITE

ストレージ基盤ソフトウェア
ETERNUS SF THE POSSIBILITIES ARE INFINITE FUJITSU

ETERNUS 性能監視のご紹介



2007年1月26日
富士通株式会社

1. はじめに

- **ご紹介ソフトウェア**
 - **ESC ETERNUS SF Storage Cruiser**
ストレージシステム全体の構成、性能、障害を管理するソフトウェア製品
ストレージシステムを構成する機器から詳細な性能情報を採取できます。
 - **SQC Systemwalker Service Quality Coordinator**
システム全体の性能情報を様々な角度から監視/分析するソフトウェア製品
様々な性能情報を多面的に分析することができます。

ESCとSQCを連携させることで、複雑なSAN(Storage Area Network)の
性能管理を容易に行えます。

FUJITSU CONFIDENTIAL. Internal Use Only

2. ストレージ性能管理のレベル

性能管理の運用には以下のレベルがあります

- **性能トラブル発生時に対処する**
- **性能トラブルが表面化する前に対処する**
実現方法: ストレージ性能の**定常監視(閾値)**を行う。
- **性能トラブルにつながるような予兆をより事前にとらえ
計画的に対処する**
実現方法: **レポート**によるストレージ性能分析を行う

**レポートによる分析を繰り返し、
定常監視で検出される問題を発生させないことが、
ストレージの安定運用につながります。**

FUJITSU CONFIDENTIAL. Internal Use Only

3. 性能監視の目的と項目

	目的	監視する性能項目
定常監視 (定常的)	<ul style="list-style-type: none"> ・ 想定外の業務アクセス増加の検出 	<ul style="list-style-type: none"> ・ レスポンスタイム ・ Diskビジー率
レポート (中期的・ 長期的)	<ul style="list-style-type: none"> ・ 業務I/Oバランスの変化に伴う、ストレージ負荷の変化の検出 ・ 業務の拡大によるストレージ負荷上昇の兆候の検出 	<ul style="list-style-type: none"> ・ IOPS ・ Disk 使用率 ・ スループット ・ CM 使用率

FUJITSU CONFIDENTIAL. Internal Use Only

4. ESCで採取するストレージ性能情報

＜採取可能な性能情報＞

- ◆ RAIDグループ/論理ディスク単位

項目	意味
IOPS (I/O Per Sec)	1秒間のRead及びWrite回数
スループット	1秒間のデータ転送量 (MB)
レスポンスタイム	平均I/O処理時間 (msec)
キャッシュヒット率	CMのキャッシュにヒットした割合 (%)
- ◆ 物理ディスク単位

ディスクビジー率	ディスクの使用率 (%)
----------	--------------
- ◆ モジュール単位

CMビジー率	CMのCPU使用率 (%)
CMコピー残量	CMのコピー残量 (GB)
CAビジー率	CAのCPU使用率 (%)
CA IOPS	CAの1秒間のRead及びWrite回数
CA スループット	CAの1秒間のデータ転送量 (MB)
DAビジー率	DAのCPU使用率 (%)
DA IOPS	DAの1秒間のRead及びWrite回数
DA スループット	DAの1秒間のデータ転送量 (MB)

FUJITSU CONFIDENTIAL. Internal Use Only

4. ESCで採取するストレージ性能情報(補足)

- **用語説明**
 - CM (Controller Module)
 - CA (Channel Adapter)
 - DA (Device Adapter)
 - Disk使用率 (=Diskビジー率)
 - CM使用率 (=CMビジー率)
 - スループット
 - IOPS
 - キャッシュヒット率
 - レスポンスタイム

ETERNUS 6000に搭載されるコントローラサーバとのインターフェースを制御するモジュール。ディスクとのインターフェースを制御するモジュール。(Disk1本あたり) 単位時間あたりのI/O処理時間 (%)。(CM1個あたり) 単位時間あたりのI/O転送量 (MB/s)。各LogicalVolume のread/writeデータ転送量 (MB/s)。各LogicalVolume の1秒間のI/O発行回数。各LogicalVolume のread/writeキャッシュヒット率 (%)。各LogicalVolume のread/writeの応答時間 (ms)。

FUJITSU CONFIDENTIAL. Internal Use Only

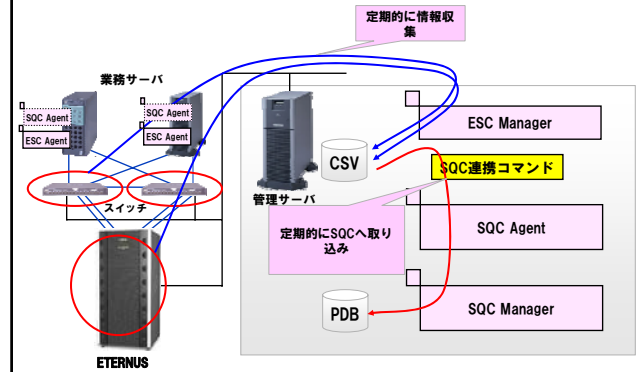
5. SQCのレポートの種類

- ◆ レポート単位
日次、週次、月次
- ◆ レポート形式
時系列、回帰、相関、等高線、過去現在、推移比較、複合表示

6

FUJITSU CONFIDENTIAL. Internal Use Only

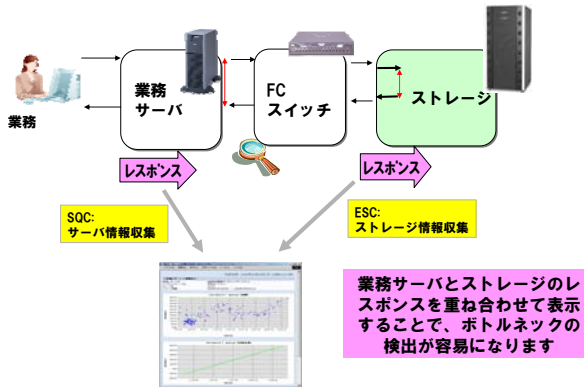
6. ESC/SQCの性能情報連携イメージ①



7

FUJITSU CONFIDENTIAL. Internal Use Only

6. ESC/SQCの性能情報連携イメージ②



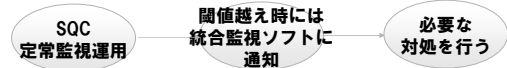
8

FUJITSU CONFIDENTIAL. Internal Use Only

7. 定常監視運用

- 運用方法:
- SQCの閾値監視機能を使用し、
 - ・レスポンスタイム
 - ・Disk使用率
 を定期的に監視します。あらかじめ、Systemwalker Centric Manager等の統合監視ソフトに閾値越え通知のアラームが飛ぶように設定しておきます。なお、SQCでは、以下のような通知手段があります。
 - ・トラップ通知
 - ・イベントログ出力
 - ・コマンド実行

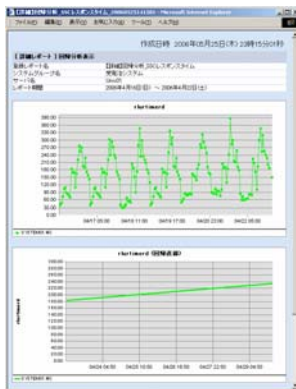
運用フロー:



9

FUJITSU CONFIDENTIAL. Internal Use Only

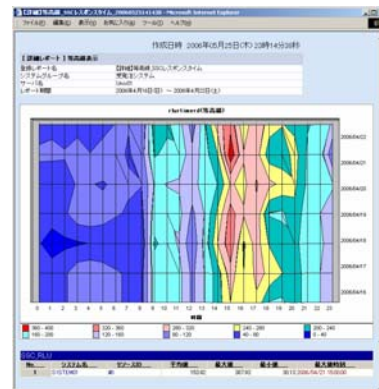
8. 中長期レポート分析①



10

FUJITSU CONFIDENTIAL. Internal Use Only

8. 中長期レポート分析②



11

FUJITSU CONFIDENTIAL. Internal Use Only

8. 中長期レポート分析③



過去比較レポート
(ある一日の状況を過去と比較)

グラフ1
棒グラフ：ある1日の24時間の推移
折れ線：過去一週間の推移(1日ごと)

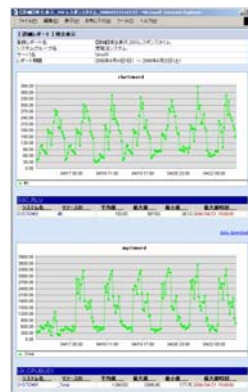
グラフ2
棒グラフ：ある1日の24時間の推移
折れ線：過去1カ月の時間帯ごとの
最大・最小・平均

過去の動作実績を軸に、
「ある1日」の状況を評価するのに便利。

12

FUJITSU CONFIDENTIAL. Internal Use Only

9. 性能トラブル分析



複合レポート
(任意の二つの性能値の関係を分析)

グラフ1
(例) ストレージのレスポンスタイム

グラフ2
(例) DBサーバのディスクビジー率

レスポンス悪化の原因調査など、関連するリソースとの依存関係を分析するのに便利。下図のような、任意の二つの性能値の相関分析も可能。



13

FUJITSU CONFIDENTIAL. Internal Use Only

FUJITSU

THE POSSIBILITIES ARE INFINITE

www.fujitsu.com THE POSSIBILITIES ARE INFINITE FUJITSU

データマネジメントを意識したストレージソリューション
WG資料

科学計算向け次期ファイルシステム
ご紹介

2007年4月
サーバシステム事業本部
Linuxソフトウェア開発統括部

高速ファイル共有システムSRFS FUJITSU

- InfiniBandで高速にI/OノードとファイルI/O
- 複数I/Oノードによる静的負荷分散

計算サーバ群

高速インタコネクタ (4GB/s~28GB/s)

SRFS

IOノード

IOノード

IOノード

GFSによる並列ファイルシステム

SANを使った高性能ブロックI/O

共用ファイル(GFS) ストライピング・仮想ボリューム(GDS)

次期以降の展開 FUJITSU

基本方針: 単一アクセスの高速化
データの集中管理

は従来通りで、多様化するデータアクセス要件に対応

- 継続運用対応
 - IOサーバダウン時の切換えによる運用継続(SAN共用FSとの連携により実現)
- 計算サーバ台数増加の対応
 - 通信トラフィック軽減・フロー制御
- 大容量データ(PetaByte)への対応
 - 階層管理(HSM)システムとの連携

次期HPCシステム構成 FUJITSU

計算サーバ群

高速インタコネクタ (4GB/s~28GB/s)

SRFS

SRFS

SRFS

SRFS

計算ノード

運用管理ノード
運用IOノード

待機管理ノード
運用IOノード

待機管理ノード
運用IOノード

待機管理ノード
運用IOノード

IOサーバダウン時にIOサーバ切り替えにより運用を継続

内蔵ストレージ
ディスクキャッシュ

計算ノード数万オーダーの計算ノードサポート時のイメージ (将来検討)

共用ファイル・仮想ボリューム(GDS)

HSM

データ元管理層

DLMを横断的な階層管理とするならば、SRFSは横断的な階層管理を適宜して行きます。

高性能ファイルシステム <SRFS V1.0> FUJITSU
SAN共用ファイルシステム <GFS>

性能(ご参考)

SRFS+GFS性能(InfiniBand®)

SRFS+GFS性能(GigabitEthernet)

- SRFSクライアントノード - PRIMERGY RX200S3
- SRFSサーバノード - PRIMEQUEST580
- インターコネクタ - InfiniBand 1本
- DETERNUS8000M900 4GbFC 8本

- SRFSクライアントノード - PRIMERGY RX200S2
- SRFSサーバノード - PRIMEQUEST480
- インターコネクタ - Gigabit Ethernet 6本
- DETERNUS4000 2GbFC 4本

SRFS : Shared Rapid File System

□ PRIMERGY RX200S2 8台からのTHROUGH PUT性能
□ SRFS独自通信ドライバによる6本のGigaEthernetをストライピング

InfiniBandはInfiniBand® Trade Associationの商標です。

高性能ファイルシステム <SRFS V2.0> FUJITSU
SAN共用ファイルシステム <GFS>

性能(ご参考)

SRFS+GFS性能(InfiniBand®)

- SRFSクライアントノード - PRIMERGY RX200S3
- SRFSサーバノード - PRIMEQUEST580
- インターコネクタ - InfiniBand 1本
- DETERNUS8000M900 4GbFC 8本

- SRFSクライアントノード - PRIMEQUEST580
- SRFSサーバノード - PRIMEQUEST580
- インターコネクタ - InfiniBand 8本
- DETERNUS4000 4GbFC 8本

SRFS : Shared Rapid File System

□ SRFS独自通信ドライバによる8本のInfiniBandをストライピング
チューニング中のデータのため参考

InfiniBandはInfiniBand® Trade Associationの商標です。

THE POSSIBILITIES ARE INFINITE FUJITSU

ETERNUS2000

ディスクアレイ ご紹介



2007年 5月
富士通株式会社

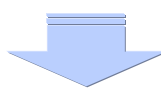
Copyright 2007 FUJITSU LIMITED

エントリーディスクアレイ市場の伸長

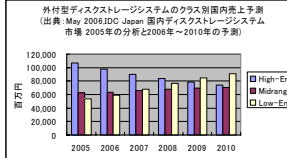
FUJITSU

■ エントリーディスクアレイ市場動向

- 2005年~2010年の年間平均成長率11%
- ハイエンド/ミッドレンジに比べ高成長



高まるエントリーディスクアレイのニーズ



外付型ディスクストレージシステムのクラス別国内売上予測
(出典: May 2006 IDC Japan 国内ディスクストレージシステム
市場 2005年の分析と2006年~2010年の予測)

ストレージシステムの低価格化

ストレージに対する高信頼・拡張性の要求

- サーバ内蔵ディスクから外付けストレージへ

中小・中堅企業でのサーバ・ストレージ集約

- 運用の簡素化

- 低価格・コンパクト・低消費電力
- 高い拡張性と多様な接続性
- 柔軟な運用管理と容易な導入
- ミッドレンジクラスの高信頼性


Copyright 2007 FUJITSU LIMITED

ETERNUS2000 ラインナップ

FUJITSU

■ 中小企業から大企業の部門用途、研究開発など、お客様の規模や利用シーンに合わせて、選択可能な幅広いラインナップ


最大12ドライブ



モデル50

ドライブ数: 12
記憶容量: 3.6TB(SAS)
9.0TB(ニアラインSATA)
コントローラー: 2
キャッシュ: 1GB


最大24ドライブ
アドバンスド・コピー
サポート



モデル100

ドライブ数: 24
記憶容量: 7.2TB(SAS)
18.0TB(ニアラインSATA)
コントローラー: 2
キャッシュ: 2GB

最大72ドライブ
アドバンスド・コピー
サポート



モデル200


ドライブ数: 72
記憶容量: 21.6TB(SAS)
54.0TB(ニアラインSATA)
コントローラー: 2
キャッシュ: 4GB

2 Copyright 2007 FUJITSU LIMITED

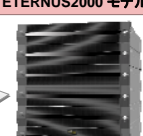
従来モデルとの対応

FUJITSU

ETERNUS4000 モデル100




ETERNUS2000 モデル200




最大容量の増加

ETERNUS4000 モデル80




ETERNUS2000 モデル100




アドバンスド・コピー
サポート

ETERNUS SX300



ETERNUS2000 モデル50



Windows/Linuxに
加えてUNIXサポート

3 Copyright 2007 FUJITSU LIMITED

ETERNUS2000の特長

ETERNUS2000の特長

- コンパクト・低消費電力
- 高い拡張性と多様な接続性
- 柔軟な運用管理・容易なセットアップ
- ミッドレンジクラスの高信頼性

4 Copyright 2007 FUJITSU LIMITED

ETERNUS2000 ディスクアレイの特長

FUJITSU

コンパクト・低消費電力・静音性

高い拡張性と多様な接続性

柔軟な運用管理・容易なセットアップ

ミッドレンジクラスの高信頼性

5 Copyright 2007 FUJITSU LIMITED

コンパクト・省エネ設計



省スペース・軽量・低消費電力

ETERNUS4000 モデル80



482x592x176mm

ETERNUS2000



482x660x88mm

	ETERNUS4000	ETERNUS2000	備考
サイズ	4U	2U	1/2
消費電力 ⁽¹⁾	580W	340W	40%減
質量	50kg	35kg	30%減

¹: 2コントローラー + 300GB ドライブ12本構成、稼働時

6

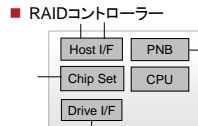
Copyright 2007 FUJITSU LIMITED

部品点数、基板サイズの大幅な縮小化

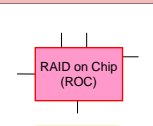


部品点数60%削減、基板サイズ60%の削減

ETERNUS4000 モデル80



ETERNUS2000



コントローラーボード



基板外形サイズ 420×300mm
部品点数 3,992点(442種)



基板外形サイズ 415x107mm
部品点数 1,490点(201種)

7

Copyright 2007 FUJITSU LIMITED

エコモード



MAID技術の応用による省エネ対応

MAID: Massive Arrays of Inactive Disks ⇒ アクセスしていないディスクは回転を停止

- 時刻設定により特定ディスクの回転停止期間をスケジュール
- ディスクへのアクセスが10分(可変)以上無い ⇒ 停止状態へ
- ディスクへのアクセスが発生 ⇒ 約1分でレディ状態へ

活用例



エコモード使用により電力消費量を15%削減 !!

8

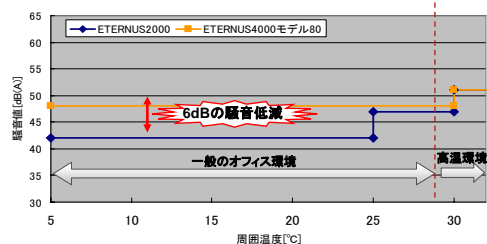
Copyright 2007 FUJITSU LIMITED

優れた静音性



冷却効率の向上、多段階ファン制御方式の採用

外気温に応じてファン回転速度を多段階で制御することにより、業界最高水準である42dBという図書館並みの静音性を実現



9

Copyright 2007 FUJITSU LIMITED

ETERNUS2000 ディスクアレイの特長



コンパクト・低消費電力・静音性

高い拡張性と多様な接続性

柔軟な運用管理・容易なセットアップ

ミッドレンジクラスの高信頼性

10

Copyright 2007 FUJITSU LIMITED

最先端のディスクドライブを採用



高性能(15,000rpm)ディスクドライブとバックアップ用途に最適な大容量ディスクを採用

SASディスク
(15,000 rpm)
SAS 3Gbps



ニアライン
SATAディスク
(7,200 rpm)
SATA 3Gbps



11

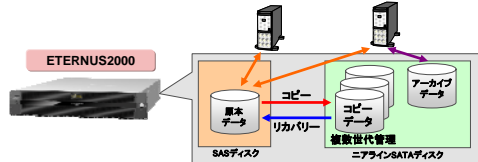
Copyright 2007 FUJITSU LIMITED

用途に応じた搭載ディスクドライブ選択 FUJITSU

同一筐体(同一DE*)内に異なるドライブを混在可能



世代管理バックアップが可能



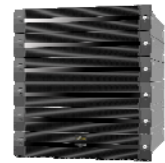
12

Copyright 2007 FUJITSU LIMITED

最大54TBまでの拡張が可能 FUJITSU

業務拡張に柔軟に対応

ETERNUS2000 モデル200



基本構成

~12ドライブ
最大3.6TB (SAS)
最大9TB (SATA)

最大構成

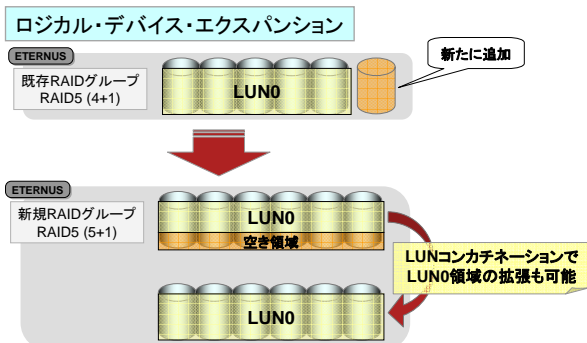
~72ドライブ
最大21.6TB (SAS)
最大54.0TB (SATA)

13

Copyright 2007 FUJITSU LIMITED

LUNの動的拡張 FUJITSU

運用停止せずに拡張が可能



14

Copyright 2007 FUJITSU LIMITED

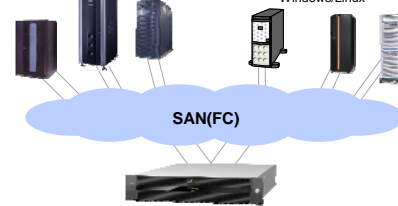
サポートサーバ・サポートOS FUJITSU

エントリークラスでも全サーバに対応

富士通、IBM、HP等のUNIX/PCサーバに対応

自社・他社UNIXサーバ

自社・他社PCサーバ



サポートOS

UNIXサーバ

■ Solaris™

■ HP-UX

■ AIX

PCサーバ

■ Windows Server® 2000/2003

■ Red Hat Enterprise Linux AS/ES

■ SUSE Linux Enterprise Server

■ VMware ESX Server

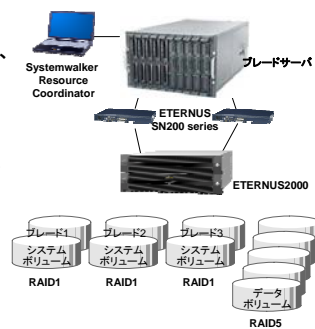
15

Copyright 2007 FUJITSU LIMITED

SAN Bootシステムによる安定運用 FUJITSU

モデル50, 100, 200すべてのモデルでSAN Bootに対応

- サーバの起動ディスクを含むすべてのデータを一元管理し、信頼性の向上と管理作業の効率化を実現
- 万が一のディスク障害発生時の迅速な対応が可能
- アドバンスド・コピー機能を利用したバックアップの高速化



注: ETERNUS2000のすべてのモデルでサポート

16

Copyright 2007 FUJITSU LIMITED

ETERNUS2000 ディスクアレイの特長 FUJITSU

コンパクト・低消費電力・静音性

高い拡張性と多様な接続性

柔軟な運用管理・容易なセットアップ

ミッドレンジクラスの高信頼性

17

Copyright 2007 FUJITSU LIMITED

豊富な高速コピー機能



■ アドバンスド・コピー機能のサポート

- ディスク to ディスクの高速なデータバックアップを実現する「アドバンスド・コピー機能」をサポート機種を拡大(モデル100, 200)
- 用途に応じて選択できる三つのコピー機能をサポート



OPC	原本データを全て複製する機能。複数世代のフルバックアップが可能のため、世代管理でのバックアップ運用に最適。
QuickOPC	一度原本データを作成した後に、更新部分のみのバックアップを行う機能。バックアップ時間の短縮が求められる、データベースシステムなどのシステムに最適。
SnapOPC	原本データの、更新部分の更新前の状態のみを複製する機能。ディスク容量を抑えることができるとともに、複数世代のバックアップが可能で、開発データなどのバックアップに最適。

18

Copyright 2007 FUJITSU LIMITED

RAIDマイグレーション

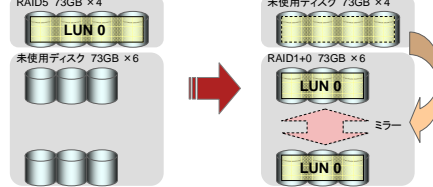


運用停止せずにRAIDグループ間のデータ移行が可能

■ 大容量ディスクへの移行



■ 信頼性の高いRAIDレベルへの移行



19

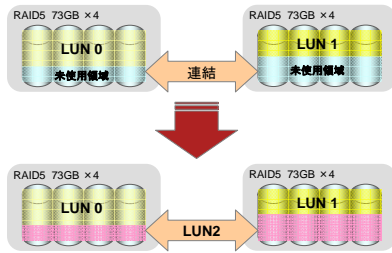
Copyright 2007 FUJITSU LIMITED

LUNコンカチネーション



分散した未使用領域のLUNを集め、ディスク容量を効率的に使用

■ 未使用領域の連結



20

Copyright 2007 FUJITSU LIMITED

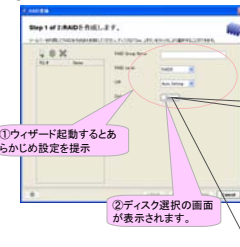
日本語GUIメニューで簡単に設定、運用、管理が可能



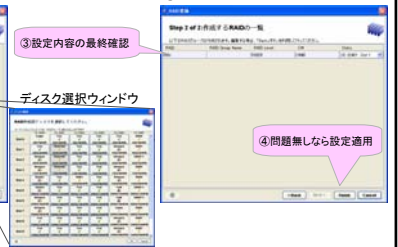
- 初期導入時: スタートアップウィザードが全て必要な設定をアシスト
- 設定変更、増設時: 各種ウィザードで設定可能 (初期設定、RAID設定、Volume設定、Host Affinity設定)
- Webブラウザ(Java有効)のみの簡易環境で動作

例) RAID設定手順

STEP1



STEP2



21

Copyright 2007 FUJITSU LIMITED

ETERNUS2000 ディスクアレイの特長



コンパクト・低消費電力・静音性

高い拡張性と多様な接続性

柔軟な運用管理・容易なセットアップ

ミッドレンジクラスの高信頼性

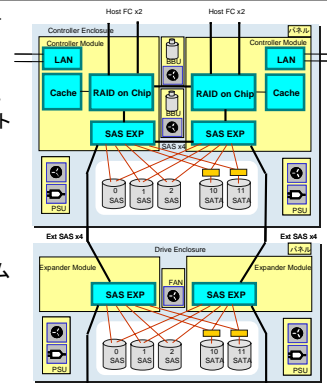
22

Copyright 2007 FUJITSU LIMITED

主要コンポーネントの二重化



- コントローラー、電源、ファンなどの主要コンポーネントを二重化
- 万が一の障害発生時でもシステム停止することなくコンポーネントの交換が可能
- ディスクドライブはシステム稼働中の活性増設が可能
- システム稼働中でも最新ファームウェアへの交換可能



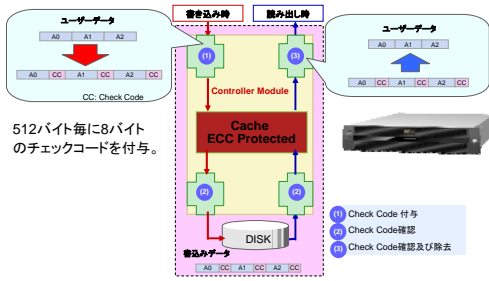
23

Copyright 2007 FUJITSU LIMITED

ブロックガード機能によるデータ保護

FUJITSU

格納する全てのデータの正常性を保証



24

Copyright 2007 FUJITSU LIMITED

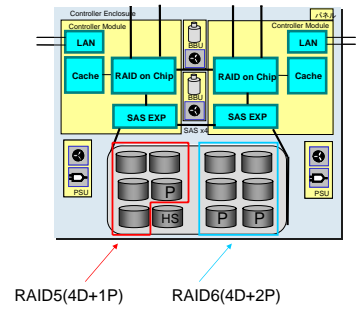
RAID6(ダブルパリティ)のサポート

FUJITSU

RAID6のサポートにより、信頼性をさらに強化

RAID6構成をサポート

- 同一RAID内のHDD2本故障を救済



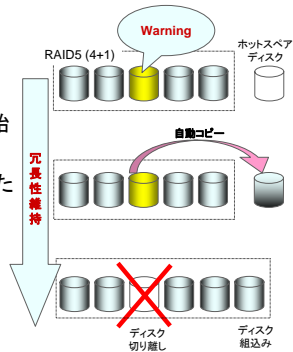
25

Copyright 2007 FUJITSU LIMITED

リダント・コピー

FUJITSU

- ディスクドライブの故障の兆候を検出すると自動的にコピーを開始
- コピーの完了時点で兆候のあったディスクを切り離し



26

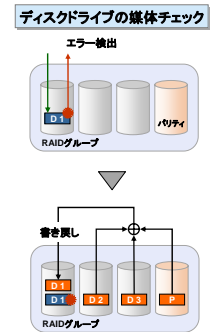
Copyright 2007 FUJITSU LIMITED

ディスクドライブパトロール

FUJITSU

パトロール機能の動作

- バックグラウンドでディスクドライブからデータを読み出し
- エラー検出時には、RAIDグループのディスクからデータ再構築を実行
- エラーが発生したディスクドライブの別ブロックへ書き戻し
- アクセス頻度の低い領域のドライブ故障も検出可能



27

Copyright 2007 FUJITSU LIMITED

エントリー向けストレージソフトウェア

FUJITSU

運用管理ソフトウェア一覧

FUJITSU

豊富な運用管理ソフトウェアをラインナップ

用途	製品名	特長
確実な管理/監視	ETERNUS SF Storage Cruiser	ストレージシステムにおける構成管理/関係管理/障害監視/性能監視
	Systemwalker Resource Coordinator	ストレージシステムのプロビジョニング/ストレージ、サーバ等のITリソースを一元管理して、自律的に全体最適化を実現
業務影響の少ないバックアップ及びデータ移行	ETERNUS SF AdvancedCopy Manager	ETERNUSが提供するOPC-EC機能を制御し、高速にバックアップを作成
	ETERNUS SF Recovery Manager for Oracle	Oracleデータベースの整合性を保証した高速バックアップと高速リストア・リカバリを実行
	ETERNUS SF Replicator	業務を継続しつつ、装置レプリカを作成（データ移行）
ストレージ容量管理	ETERNUS SF TSM	各種システム構成に対応したバックアップ
	ETERNUS SF Disk Space Monitor	SANストレージ容量のリアルタイム表示ときき値監視
高速なデータ転送	ETERNUS SF XL-DATA/MV	異なるプラットフォーム間でのデータ交換

28

Copyright 2007 FUJITSU LIMITED

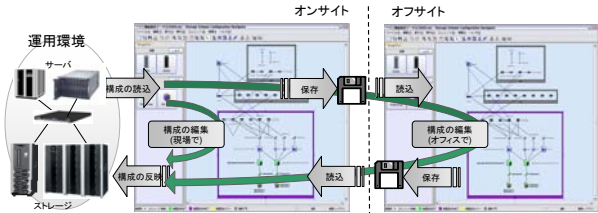
29

Copyright 2007 FUJITSU LIMITED

ETERNUS SF Storage Cruiserの新機能 FUJITSU

異なるストレージの異なるボリューム作成を共通化。
また、グラフィカルな操作画面により、ボリューム作成作業の負担を軽減します。

- サーバからストレージまでの関係をグラフィカルに把握しながら構成の設計ができます。
- わずかなパラメーター入力により、最適な構成を自動的に作成できます。
- 設計したストレージ装置の構成に問題がないかを事前に自動的に検証できます。
- 運用環境の構成をファイルに保存してオフサイトで修正可能。
また、修正した構成は、オンサイトで一括反映ができます。

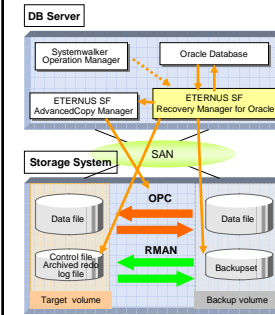


30

Copyright 2007 FUJITSU LIMITED

Recovery Manager for Oracleのプラットフォーム拡充 FUJITSU

Oracleデータベースの簡単・安全・確実なバックアップ・リカバリーソフトウェア



- 信頼性の高いバックアップ資産を取得
- 簡単で安全・確実なリカバリー操作
- ETERNUS連携で高速バックアップ・リストア (Recovery Manager for Oracle EEのみ)

PRIMERGY向け商品を拡充

- Windows版
Recovery Manager for Oracle SE 提供済
Recovery Manager for Oracle EE 計画中
(2007年秋提供予定)
- Linux版
Recovery Manager for Oracle SE 2007/06提供
Recovery Manager for Oracle EE 2007/06提供

31

Copyright 2007 FUJITSU LIMITED

FUJITSU

参考資料

32

Copyright 2007 FUJITSU LIMITED

ETERNUS2000 ディスクアレイ 装置仕様 FUJITSU

	モデル50	モデル100	モデル200
コントローラー数		2	
CPU周波数		800MHz	1.2GHz
キャッシュ容量	1GB	2GB	4GB
データバックアップ時間	48h (1BBU) / 96h (2BBU *1)		
ホストインターフェース (*2)	FC		
ドライブインターフェース	4Gbps × 20r4		
サポートRAID	SAS/SATA 3Gbps		
搭載ディスク数	4~12	6~24	6~72
搭載ディスク	SAS	300/146/73GB 15,000回転	
	SATA	750/500GB 7,200回転	
ホスト接続数	8	64	128
LUN数	128	512	1,024
ラックマウントサイズ	2U	2U~4U	2U~12U
電源条件	単相, AC100~240V		

*1: 2BBUはオプション

33

Copyright 2007 FUJITSU LIMITED

ETERNUS2000 ディスクアレイ 機能仕様 FUJITSU

	ETERNUS2000	ETERNUS SX300	ETERNUS4900 モデル60, 100	ETERNUS4900 モデル300
OPC/Quick OPC(*)		×	モデル100のみ	
Snap OPC (*)	モデル100, 200のみ	×	×	○
EC (*)		×	×	
エコモード	○	×	×	○
RAIDマイグレーション	○	×	○	○
ロジカル・デバイス・エクステンション	○	×	○	○
LUNコンカネーション	○	×	×	○
ブロックガード	○	×	○	○
RAID6	○	×	×	○
リダンダントコピー	○	×	○	○
ディスクドライブ/ポートロール	○	○	○	○
SAN Boot	○	×	○	○
HDD番号化	×	×	×	○
ISCSI	×	×	×	○
リモートコピー	×	×	×	○

*: 各種コピー機能はオプション

34

Copyright 2007 FUJITSU LIMITED

ETERNUS2000とサーバ内蔵RAIDの比較 FUJITSU

ETERNUS2000



- コントローラー 二重化
- ブロックガード機能あり
- 最大72ドライブ
- 最大4GBキャッシュ
- 複数サーバ共用可
- RAID6サポート

サーバ内蔵RAID



- コントローラー 単一構成
- ブロックガード機能なし
- 最大9ドライブ
- 最大256MBキャッシュ
- サーバ内蔵のみ

35

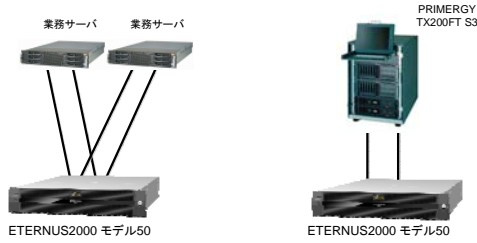
Copyright 2007 FUJITSU LIMITED

システム構成例 (1)

FUJITSU

小規模なシステムに適したストレージ

- オフィス環境に適した省電力・コンパクト
- サーバの内蔵ディスクドライブに比較して高性能・高信頼に優れています



36

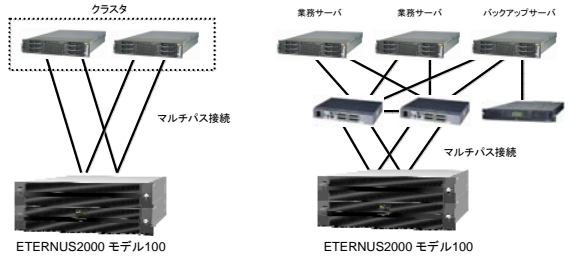
Copyright 2007 FUJITSU LIMITED

システム構成例 (2)

FUJITSU

中規模オフィスや小規模データセンターに適したストレージです

- 企業の電子メールのアーカイブシステムに最適です
- 運用が簡単、拡張性に優れています



37

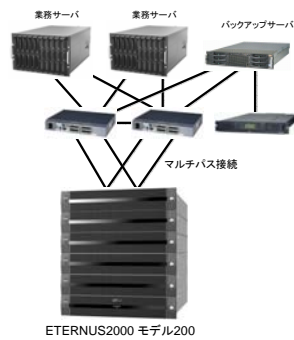
Copyright 2007 FUJITSU LIMITED

システム構成例 (3)

FUJITSU

ブレードサーバによるサーバ集約に適したストレージ

- 内部統制を行うために必要なITインフラを構築可能です
- SAN Bootに対応しており、耐障害性が高いシステムが構築可能です



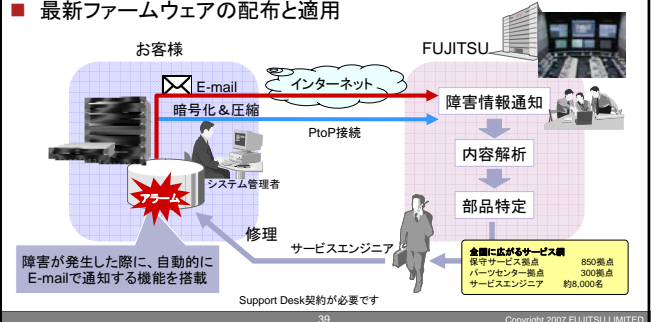
38

Copyright 2007 FUJITSU LIMITED

リモートメンテナンス

FUJITSU

- 障害発生時の迅速な保守対応
- 障害の予兆監視
- 専用線を利用したPtoP接続も提供
- 最新ファームウェアの配布と適用



39

Copyright 2007 FUJITSU LIMITED

ETERNUS

インターネット情報サービス
<http://storage-system.fujitsu.com/jp/>

●記載されている会社名、商品名は各社の登録商品または商標です。

FUJITSU

THE POSSIBILITIES ARE INFINITE

THE POSSIBILITIES ARE INFINITE FUJITSU

ETERNUS NR1000 V

NASゲートウェイご紹介

2007年6月6日
富士通株式会社
ストレージソリューション事業部



Copyright©2007 FUJITSU LIMITED

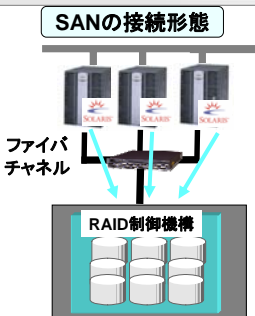
FUJITSU

SANとNAS

Copyright©2007 FUJITSU LIMITED

接続形態

SANの接続形態

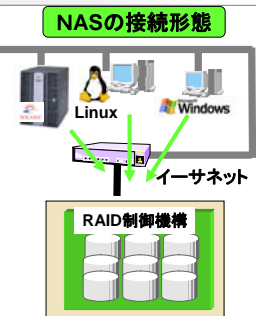


ファイバチャンネル

RAID制御機構

同一ベンター下で
ストレージの共有

NASの接続形態



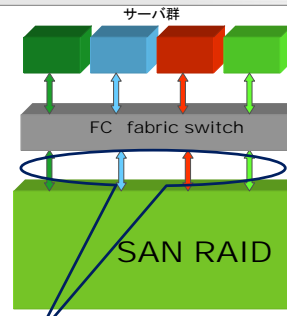
Linux Windows
イーサネット

RAID制御機構

異機種間で
ファイルの共有

Copyright©2007 FUJITSU LIMITED

インターフェースの特徴

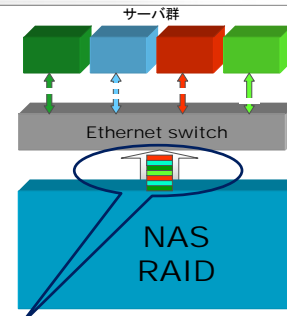


サーバ群

FC fabric switch

SAN RAID

アクセス帯域保証



サーバ群

Ethernet switch

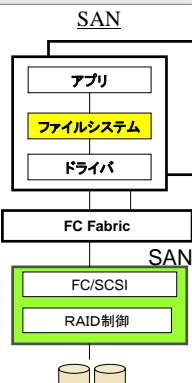
NAS RAID

帯域使用効率最大化

Copyright©2007 FUJITSU LIMITED

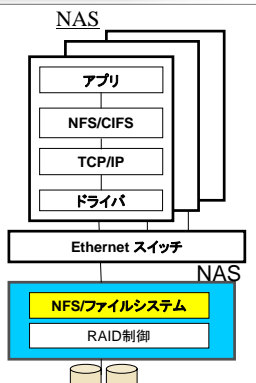
アクセス方式

SAN



SAN

NAS



NAS

Copyright©2007 FUJITSU LIMITED

業務区分

	SAN	NAS
目的/狙い	ストレージ統合、装置統合によるTCO削減、大規模システムへの対応	
適性業務	同一ベンター下のストレージ ミッションクリティカル	異種サーバ混在LAN環境 ファイル共有、頻繁な増設
ファイル共有	同一サーバ/同一OS 環境下でのみ共有	複数のサーバ(UNIX/ Windows)でファイルを共有

機種	ETERNUS8000 ETERNUS4000	NR1000F series
----	----------------------------	----------------

SANの一部を
NASとして利用する方法はないものか?...

Copyright©2007 FUJITSU LIMITED

NASゲートウェイによるSANディスク利用用途の拡大



	SAN	NASゲートウェイ
目的/狙い	ストレージ統合、装置統合によるTCO削減、大規模システムへの対応	
適性業務	同一ベンダー下のストレージ ミッションクリティカル	異種サーバ混在LAN環境 ファイル共有、頻繁な増設
ファイル共有	同一サーバ/同一OS 環境下でのみ共有	複数のサーバ(UNIX/ Windows)でファイルを共有

装置共有	ETERNUS8000 ETERNUS4000	NR1000 V3020
------	----------------------------	--------------



NASゲートウェイ

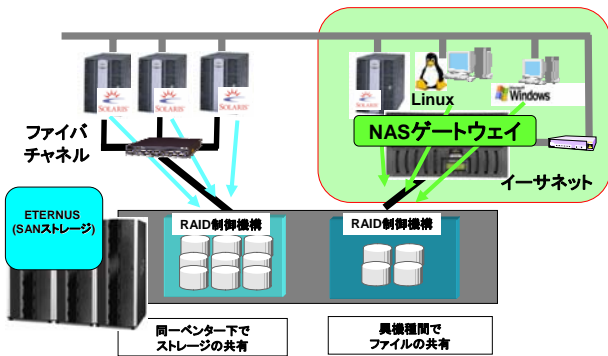


NASゲートウェイ

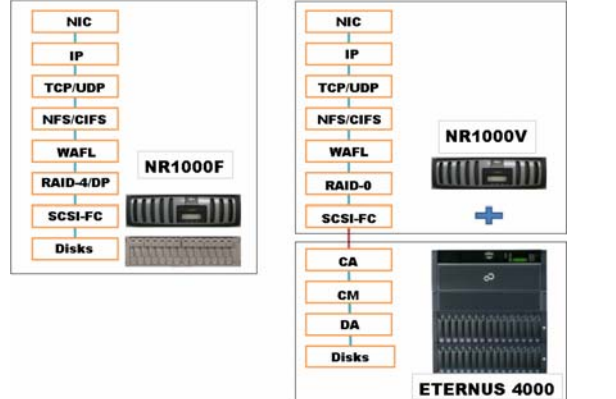
NASゲートウェイ



SANストレージの一部領域をNAS領域として利用できる



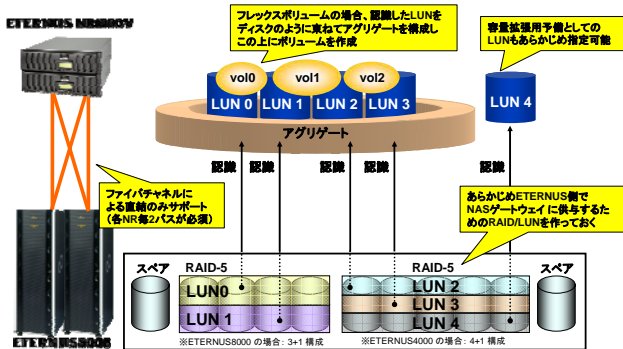
NR1000V 制御構造



NASゲートウェイ "NR1000V" Vol構造



SAN Storage 側に用意された LUN を束ねて使用



NASゲートウェイの特性(NR1000Fとの差)



- 性能
 - NASデータ転送速度: NR1000Fと同等
(注: ETERNUSがネックにならない構成)
 - ランダムアクセス: 測定予定。過去の他社例ではNR1000Fと同等にすることが可能。
(注: ETERNUSがネックにならない構成にする必要あり)
- 機能
 - NR1000Fの機能は、ディスク周りの機能を除きすべて可能
 - できないものは、SnapLock Compliance、RAID-DPなど
- 仕様
 - 最大容量: 最大LUNサイズ500GBの制約があるため、最大容量はNR1000Fより小さくなる。

NR1000V+ETERNUS8000/4000の利点 FUJITSU

- NASとしては、世界No.1シェア: NetApp社による NR1000の機能をフルに活用
- RAIDとしては、世界No.1クラス性能・容量・富士通 ETERNUS 8000/4000を活用
- 最強の組み合わせで、SANの利用にもNASの利用にも最適なストレージを提供

12

Copyright©2007 FUJITSU LIMITED

製品仕様

13

Copyright©2007 FUJITSU LIMITED

NR1000 V3020装置仕様 FUJITSU

		容量
対応ネットワークプロトコル		NFS、CIFS、HTTP、各プロトコル
検索容量*1	ETERNUS8000接続時 最大容量*2	87.3TB
	ETERNUS4000接続時 最大容量*3	59.3TB
ユーザー容量*4	ETERNUS8000接続時 最大容量*2	37.6TB
	ETERNUS4000接続時 最大容量*3	33.3TB
コントローラ構成	最大ディスク数(10/100/1000BASE-T)	8
	最大ディスク数(SAS)	2
	標準接続用SCSIポート(LVD)	2
	最大ディスク数	6
	組み合わせた最大枚数	6
インターフェースカード	デュアルギガビットイーサネット(前面または背)	6
最大搭載数(PCI-X)	10ギガビットイーサネット(前)	4
	10ギガビットイーサネット(後)	4
	LVD SCSI	6
	Fibre接続用ファイバチャネル(2ポート)	6

		仕様
外形寸法 (W×D×H) (mm)		449×618×130
質量(kg)		34
電源条件	電圧	AC100V/200V(±10%)
	周波数	単相
	消費電力	50W/60W(+2~4%)
	最大電力 (100V時) (W)	250
	最大電力 (100V時) (kVA)	900
	動作環境条件 (動作時)	温度 10~35℃
		湿度 10~80%RH(結露がないこと)

NR1000 V3020は、NASストレージ装置のリーディングベンチマークであるNetwork Appliance社 iFPOEM製品です。
 *1: 検索容量はNR1000 V3020で使用可能なLLNの合計容量で、本容量は18Byte/1,024Byteで計算しております。
 *2: 本容量は、300GBファイバチャネルディスクドライブ RAID5(3+1)の RAIDグループにFC/LINE作成した場合の容量です。
 *3: 本容量は、300GBファイバチャネルディスクドライブ RAID5(3+1)の RAIDグループにFC/LINE作成した場合の容量です。
 *4: 本容量は、フルクォータリウムでアプリケーション用Snapshot容量を9%、Snapshot容量を20%とした場合のユーザー容量で、180Byte/1,024Byteで計算しております。
 *5: オプションカード最大搭載数は搭載構成によって異なります。
 *6: コントローラ部分の構成です。NR1000 V3020はコントローラとなります。
 *7: NR1000 V3020の17インチの最大搭載容量、および17インチの最大搭載容量は16TBとなります。

14

Copyright©2007 FUJITSU LIMITED

ETERNUS8000, ETERNUS4000接続仕様 FUJITSU

モデル	ETERNUS4000	ETERNUS8000		
	モデル500	モデル900	モデル1100	モデル2100
サポート接続形態	FC-SWを介さない直接接続のみ			
サポートディスクドライブ	ファイバチャネルディスクドライブ(注1) (146GB/10,000rpm, 300GB/10,000rpm)			
サポートRAIDレベル(注2)	RAID5 (4+1)	RAID5 (3+1)		
サポートNR1000 V3020ク ラスタセット数	最大 1セット			

注1) SATAドライブ検出中
 注2) RAID6検出中

15

Copyright©2007 FUJITSU LIMITED

接続可能ETERNUS8000 装置仕様(モデル900以上) FUJITSU

		モデル 900	モデル 1100	モデル 2100
コントローラ数		2 or 4	4	4, 6 or 8
コントローラ単体のCPU数		3.6GHz × 2	3.6GHz × 2	3.6GHz × 2
キャパシティ		8 ~ 64GB	32 ~ 128GB	64 ~ 256GB
ホストIF	FC	4Gbps × 4 ~ 64	4Gbps × 16 ~ 64	4Gbps × 32 ~ 128
	ISCSI	1Gbps × 4 ~ 32	1Gbps × 4 ~ 32	1Gbps × 4 ~ 64
	OCLINK	4 ~ 32	4 ~ 32	4 ~ 64
	FCLINK	4 ~ 32	4 ~ 32	4 ~ 64
CA数		4 ~ 16	4 ~ 16	4 ~ 32
BRT数		2 or 4	4	4 or 8
ドライブIF		FC 2 or 4Gbps × 16 or 32	FC 2 or 4Gbps × 32	FC 2 or 4Gbps × 32 ~ 64
サポートRAID(注)		1, 0+1, 5, 6	1, 0+1, 5, 6	1, 0+1, 5, 6
DE数		4 ~ 32	24 ~ 68	~ 184
搭載ディスク数		~ 480	~ 1,020	~ 2,760
搭載ディスク		300/146GB 10,000rpm	300/146GB 10,000rpm	300/146GB 10,000rpm
システムディスク数		4 or 8 (専用)	16 (専用)	32 (専用)
ロッカ数		1 ~ 3	3 ~ 6	~ 16
外形寸法	W	740~1,980mm	740~3,840mm	~5,080mm + 4,960mm
	D	955mm	955mm	955mm
	H	1,800mm	1,800mm	1,800mm
電源条件		単相AC200~240V	単相AC200~240V	単相AC200~240V

(注)ETERNUS8000がサポートするRAID+1は、RAID+0+RAID10と書かれる方式と同様、データをミラーリングして、複数本のディスクドライブをトラッキングする方式。

16

Copyright©2007 FUJITSU LIMITED

接続可能ETERNUS4000 装置仕様(モデル500) FUJITSU

		モデル 500
コントローラ数		2
コントローラ単体のCPU数		2.0GHz × 2
キャパシティ		8 ~ 32GB
ホストIF	FC	4Gbps × 4 ~ 16
	ISCSI	1Gbps × 4 or 8
CA数		2 or 4
ドライブIF		FC 2 or 4Gbps × 4 or 8
サポートRAID(注)		0, 1, 0+1, 5, 6
DE数		2 ~ 28
搭載ディスク数		8 ~ 420
搭載ディスク	OnLine	300/146GB 10,000rpm
システムディスク数		4 (共用)
ロッカ数		1 ~ 3
ロッカ寸法	W	580~1,770mm
	D	990mm
	H	1,800mm
電源条件		単相 AC100~120V 単相 AC200~240V

(注)ETERNUS4000がサポートするRAID+1は、RAID+0+RAID10と書かれる方式と同様、データをミラーリングして、複数本のディスクドライブをトラッキングする方式。

17

Copyright©2007 FUJITSU LIMITED

他社比較



	NR1000 V3020	EMC NS704G	NEC iStorage NV7300G	SANRISE Adaptable Modular Storage 500 (参考)
クラス構成	Active/Active	Active x 3+Standby x 1	Active/Active	Active/Active
最大論理容量	67.3TB	48TB	84TB	64.7TB
ネットワークインターフェース	1000BASE-T x 24 (増設カード搭載時)	1000BASE-T x 8 (2port/Data Mover)	1000BASE-T x 8 (4port/node)	1000BASE-T x 8 (4port/NAS Blade)
TCO削減をサポートする機能	Snapshot機能 volume単位最大255個 (差分管理で参照のみ可、高速)	△ SnapSure (Copy on Writeで低速)	○ 最大255個 (Copy on Writeで低速)	○ 最大124個 (Sync Image: 差分管理で参照のみ可、Copy on Writeで低速)
	Snapshotからの一括復元 SnapRestore (ボリューム又はファイル単位、高速)	○ SnapSure(Copy on Writeで低速)	×	×
	更新可能複製ボリュームの作成 FlexClone (差分管理、高速)	×	×	△ ShadowImage (ボリューム複製、低速)
ビジネス継続性をサポートする機能	ミラードボリューム SyncMirror	×	×	×
	災害復旧対策 SnapMirror (LAN/WAN接続、運用が簡易)	△ Celerra Replicator (LAN/WAN接続、設定が複雑)	×	△ TrueCopy remote replication (FC接続、設定が複雑)
	D2Dバックアップ SnapVault(LAN/WAN接続)	×	△ (NetVaultをインストールして仮想テープソフトウェアへバックアップ)	×
WORM機能	○ SnapLock (OS/SANESサポート、ファイル単位で簡易)	×	○ 仮想WORMファイルシステム (NFSのみ)	○ Data Retention Utility (LUN単位)

18

Copyright©2007 FUJITSU LIMITED



THE POSSIBILITIES ARE INFINITE

THE POSSIBILITIES ARE INFINITE FUJITSU

フィールドイノベーション 富士通のストレージシステム ETERNUSと関連ソリューション (ご紹介)

2007年10月
富士通株式会社
ストレージシステム事業本部
インテグレーション統括部

Copyright 2007 FUJITSU LIMITED

THE POSSIBILITIES ARE INFINITE FUJITSU

Agenda

- 経営とIT、IT基盤の要件、ストレージの役割
- 企業活動、内部統制、セキュリティ
- 大量データの長期保存
- コンテンツ管理ストレージ ETERNUS AS500
- アーカイブ需要
- ファイルサーバにおけるアーカイブソリューション

Copyright 2007 FUJITSU LIMITED

THE POSSIBILITIES ARE INFINITE FUJITSU

経営とITの一体化

現場の革新を支え、
現場と現場、現場とマネジメントの
リアルタイムな連携を実現

Business Agility Business Continuity Business Efficiency

Copyright 2007 FUJITSU LIMITED

THE POSSIBILITIES ARE INFINITE FUJITSU

お客様のIT基盤への要件

Business Agility (俊敏性)	<ul style="list-style-type: none"> ● End To Endのリアルタイムなサービス連携 ● 横断的な情報のスピーディーな戦略活用 (攻めの経営)
Business Continuity (継続性)	<ul style="list-style-type: none"> ● ITシステムの安全・安定稼働/災害対策 ● 中長期のインクリメンタルな拡張/成長性 ● 内部統制/セキュリティへの対応
Business Efficiency (効率性)	<ul style="list-style-type: none"> ● ライフサイクルを通じたIT投資の最適化 ● 環境に配慮したシステム構築と運用

変化に対応できるITプラットフォームを提供

Copyright 2007 FUJITSU LIMITED

THE POSSIBILITIES ARE INFINITE FUJITSU

ITプラットフォームの活用

- 組織を存続・発展させ、ステークホルダーに向けた社会的責任を果たすために、Agility、Continuity、Efficiencyが必要
- これらの要件を実現するマネジメントシステムには、企業の神経系を司るITの戦略的活用が不可欠
- マネジメントに必要な情報を「収集」して、「活用」するためには、情報を蓄える「ストレージシステム」を起点にした解決が有効

ITプラットフォーム
ストレージシステム

Copyright 2007 FUJITSU LIMITED

THE POSSIBILITIES ARE INFINITE FUJITSU

内部統制/セキュリティへの対応

Copyright 2007 FUJITSU LIMITED

企業活動の要件

FUJITSU

- 企業が正しく活動していることを開示
 - ステークホルダー＋法令、基準、行動規範
 - 個人情報保護法、e-文書法、新会社法、金融商品取引法、...
 - 各種認証規格など
- 経営者への支援として、企業価値の向上、リスクの回避・低減
- GRC (Governance, Risk, Compliance)の一元管理
 - ガバナンス: 企業に適した戦略的方向性の提示
 - リスク管理: リスクが存在する領域とその潜在的影響の特定
 - コンプライアンス: リスクを軽減するために実施する活動

6

Copyright 2007 FUJITSU LIMITED

大量データの長期保存

FUJITSU

- メール、ファイル、ログデータの確実な保存
 - 規制強化に向けた長期保存、改ざん防止
 - 既存システムへの影響が極力ないこと
 - システム構成、運用、ユーザーの利用環境への影響排除
 - データ増にも、迅速かつ容易な対応
 - 長く安心な利用
- 必要時のアクセス保証
 - 適切な記憶媒体・場所の選択とデータ移動
- 保存コストへの配慮
 - コスト増の抑止（機器、運用軽減）

7

Copyright 2007 FUJITSU LIMITED

コンテンツ管理ストレージ ETERNUS AS500 アーカイブストレージ

THE POSSIBILITIES ARE INFINITE FUJITSU

Copyright 2007 FUJITSU LIMITED

CAS(Content Aware Storage)とは？

FUJITSU

コンテンツアドレスによるシンプルな管理

コンテンツの物理的な格納場所をアプリケーションから管理する必要が無く、固有のコンテンツアドレス (ID=識別子) によりコンテンツへのアクセスが可能。

仮想化されたストレージ領域

異なる特性を持つストレージを意識することなくコンテンツアドレスによるデータ管理が可能であり、コンテンツの完全性を保証しつつ容易なストレージ容量の拡張性を実現。

価値に応じた最適なストレージへのデータ配置

参照頻度の高いコンテンツを高速アクセス可能なストレージに配置し、参照頻度の低くなった古いデータはコストパフォーマンスに優れたストレージにのみ配置する事で、保管コスト削減が可能。

長期保管に最適なコンテンツ管理ストレージ

9

Copyright 2007 FUJITSU LIMITED

ETERNUS AS500 アーカイブストレージの特徴

FUJITSU

- ハイブリッド・ハードウェア
 - ディスクアレイ、テープライブラリ、コンテンツ管理コントローラー
- 内蔵型コンテンツ管理ソフトウェア
- 専用APIによる、セキュアアクセス
- 自律的階層管理
- コンプライアンスへの対応
 - 保管期間設定・変更、完全削除、等

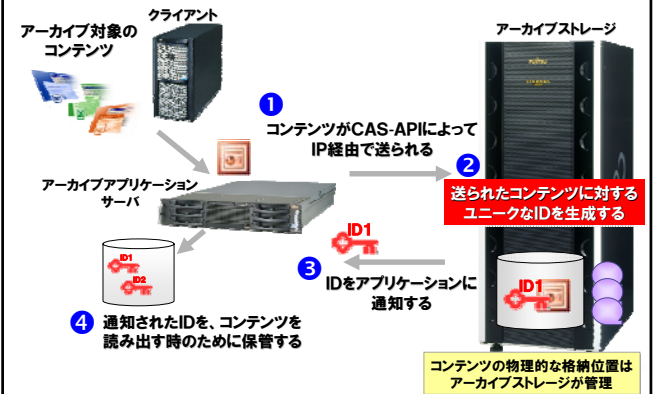


10

Copyright 2007 FUJITSU LIMITED

アーカイブストレージによるコンテンツ管理

FUJITSU



11

Copyright 2007 FUJITSU LIMITED

THE POSSIBILITIES ARE INFINITE FUJITSU

コンテンツ管理、データアーカイブ 応用例 メールアーカイブ から ファイルシステムアーカイブ へ

Copyright 2007 FUJITSU LIMITED

FUJITSU

メールアーカイブソリューション

- 既存メールシステムに追加するだけでアーカイブが可能
 - 検索・表示の操作がWebから容易に実行可能
 - 本文および添付ファイル名検索が可能
 - ETERNUS アーカイブストレージのWORM機能との連携により、メールの改ざん防止を実現
 - 新陳代謝機能との連携により長期保存を実現

Copyright 2007 FUJITSU LIMITED

FUJITSU

Exchange 管理者が抱える課題

Microsoft Exchange Serverの長期運用に伴うメールボックスの肥大化

- メールボックスの容量制限に伴う運用コスト増大
 - 容量制限から必要なメールを退避・維持するための運用が別途必要
- バックアップ、リストアの長時間化
 - データ量の増大から終業時間から始業時間までにバックアップが終わらない

電子情報に関する法令整備に伴うコンプライアンスへの対応

- メール履歴、メール本文、添付資料の長期保存が必要
- メール履歴、メール本文、添付資料のテーマに沿った検索が必要
- 監査時に監査用のシステムを新たに作成する等のコストが発生

Enterprise Vault
 メールボックス拡張機能による解決

Enterprise Vault
 ジャーナルアーカイブ機能による解決

Copyright 2007 FUJITSU LIMITED

FUJITSU

システム管理者にとってのメールボックスのアーカイブ

- 古いメールデータをExchange からEnterprise Vault にアーカイブして、Exchange のデータベースを軽量化
- ポリシーに従い自動的にメールを選択してアーカイブすることで、日々の運用に管理者の関与不要
- アーカイブ後にはユーザーのメールボックスにショートカットを配置することにより、Enterprise Vaultへのシームレスなアクセスが可能
- 保存期間を過ぎ、かつ一定期間アクセスが無かったアイテムを自動的に削除

Copyright 2007 FUJITSU LIMITED

FUJITSU

ジャーナルアーカイブの全体像

- Exchange 組織内のメール/外部メールを書き換え不可能な媒体へ保存する事により、メールの改ざんを防止
- アーカイブ後は複写用のメールボックスストアからメールを削除
- アーカイブしたメールは横断検索可能

Copyright 2007 FUJITSU LIMITED

FUJITSU

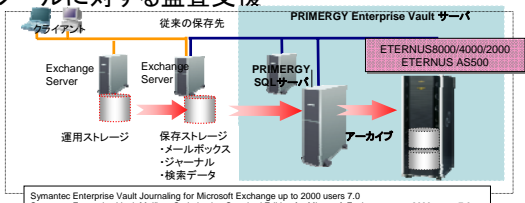
Enterprise Vaultによるメールアーカイブ事例

- Exchange ServerのTCO削減
 - メールボックスの容量を減らしてTCO削減
- ストレージ装置の効率的な利用
 - 圧縮保存するため、ストレージ装置を効率的に利用
- 保存メールは高速検索
- 保存データの削除
 - 保存期間の管理はEnterprise Vaultで実施するので、メンテナンスも最小

Copyright 2007 FUJITSU LIMITED

Enterprise Vaultによるメールアーカイブ FUJITSU

- Exchange ServerのTCO削減
 - 個人用メールボックスデータのアーカイブにより、Exchangeシステムの効率的運用が可能。
 - アーカイブデータを圧縮保存するため、ストレージ容量効率が向上
- 保存データの削除
 - 保存期間の管理はEnterprise Vaultで実施するので、Exchangeシステムと独立したメンテナンスが可能。
 - E-Discovery対応
- 保存メールに対する監査支援



Symantec Enterprise Vault Journaling for Microsoft Exchange up to 2000 users 7.0
Symantec Enterprise Vault Mailbox Optimization Standard Edition for Microsoft Exchange up to 2000 users 7.0

ファイルサーバ・システム管理者が抱える課題 FUJITSU

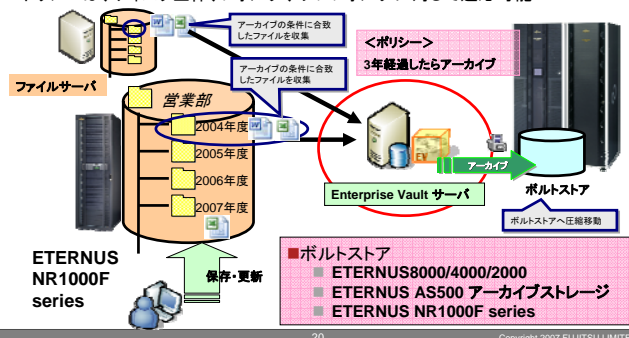


- 古くなったデータの存在で高価なストレージスペースのムダ使いが発生
 - ファイルサーバの容量制限に伴う運用コスト増大
 - 容量制限から必要なデータを退避・維持するための運用が別途必要
 - バックアップ、リストアの長時間化
 - データ量の増大から終業時間から始業時間までにバックアップが終わらない
- 電子情報に関する法令整備に伴うコンプライアンスへの対応
 - 内部統制に向けた対応が必要
 - コンプライアンス要件やリカバリ用に、複数のバージョンのファイルを保持したい
 - システムのアクセス記録を全て保存

**Enterprise Vaultファイルシステムアーカイブ機能
Systemwalker Centric Managerログ統合管理**
で解決

ファイルシステムアーカイブのシステム概要 FUJITSU

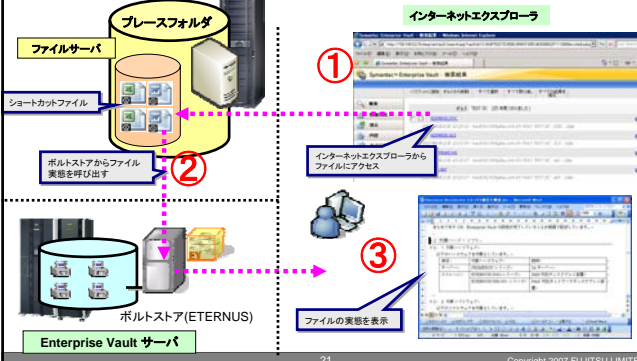
- 古いデータをファイルサーバからEnterprise Vaultにアーカイブ(圧縮移動)して、ファイルサーバを軽量化
- ポリシーは、ディスク全体、フォルダ、サブフォルダに対して適応可能



- ポルトストア
 - ETERNUS8000/4000/2000
 - ETERNUS AS500 アーカイブストレージ
 - ETERNUS NR1000F series

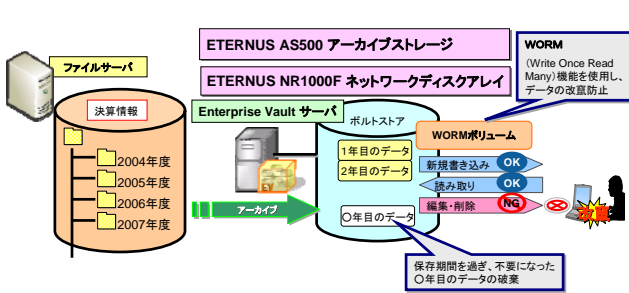
アーカイブデータへ、透過的にアクセス FUJITSU

- ショートカットファイルにアクセスすることにより、透過的にアーカイブ済みのファイルにアクセス可能



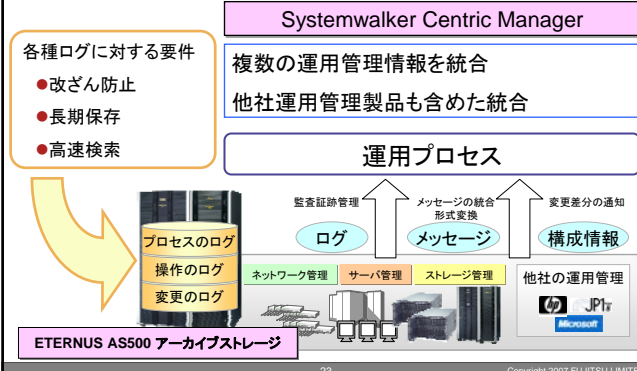
コンプライアンスへの対応 FUJITSU

- WORM機能を使用したポルトストアにアーカイブすることでデータ改竄を防止→有事の際の証拠能力向上
- 保存期間を過ぎ不要となったデータは自動的に破棄



システム運用ログの統合管理 FUJITSU

- インフラの見える化に必要なシステムを統合監視、複数の運用統合






THE POSSIBILITIES ARE INFINITE FUJITSU

ETERNUS ストレージ製品群

Copyright 2007 FUJITSU LIMITED

ETERNUS ストレージ・プロダクト FUJITSU

Tape Libraries	RAID Systems	Software & Switches
 <p>ETERNUS LT series</p>	<ul style="list-style-type: none"> ● SAN Products ETERNUS8000 ETERNUS4000 ● NAS Products ETERNUS NR1000F series 	<ul style="list-style-type: none"> ● Software ETERNUS SF ● Virtualization Switch ETERNUS VS900 ● Fibre Channel Switch ETERNUS SN200 series
<p>コンプライアンス/ILM</p> <div style="display: flex; justify-content: space-around;"> <div style="text-align: center;">  <p>アーカイブ ストレージ</p> <p>ETERNUS AS500 Archive Storage</p> </div> <div style="text-align: center;">  <p>仮想テープ</p> <p>ETERNUS VT600 - CentricStor -</p> </div> </div>		

Copyright 2007 FUJITSU LIMITED

ETERNUS ディスクアレイ ラインナップ FUJITSU

Enterprise storage systems

ETERNUS8000

モデル700 No. of drives: 60 17.4TB(FC), 43.5TB(NL) Cache Capacity: 8-16GB	モデル900 No. of drives: 480 141.8TB(FC), 354.2TB(NL) Cache Capacity: 8-64GB	モデル1100 No. of drives: 1,020 301.7TB(FC), 763.5TB(NL) Cache Capacity: 32-128GB	モデル2100 No. of drives: 2,760 518.5TB(FC), 2,047.1TB(NL) Cache Capacity: 64-256GB
--	---	--	--

Mid-range storage systems

ETERNUS4000

モデル300 No. of drives: 120 36.0TB(FC), 88.2TB(NL) Cache Capacity: 8GB	モデル500 No. of drives: 420 126TB(FC), 313.2TB(NL) Cache Capacity: 32GB
--	---

Entry storage systems

ETERNUS2000

モデル50 No. of drives: 12 3.6TB(SAS), 9.0TB(SATA) Cache Capacity: 1GB	モデル100 No. of drives: 24 7.2TB(SAS), 18.0TB(SATA) Cache Capacity: 2GB	モデル200 No. of drives: 72 21.6TB(SAS), 54.0TB(SATA) Cache Capacity: 4GB
---	---	--

Copyright 2007 FUJITSU LIMITED

ETERNUS テープライブラリ ラインナップ FUJITSU

ULTRIM™	Uncompressed capacity/cartridge (GB)	Data transfer speed (MB/s)
Ultrium2	200	30~35
Ultrium3	400	80
Ultrium4	800	120

Enterprise class

ETERNUS LT270

Tape drive	: LTO Ultrium3, 4
No. of drives	: 20
No. of cartridges	: 170
Storage capacity	: 549.6TB*
Host interface	: FC

Mid-range class

ETERNUS LT250

Tape drive	: LTO Ultrium3
No. of drives	: 2
No. of cartridges	: 40
Storage capacity	: 16TB**
Host interface	: SCSI (Ultra160)

Entry class

ETERNUS LT220

Tape drive	: LTO Ultrium2
No. of drives	: 1
No. of cartridges	: 8
Storage capacity	: 1.6TB**
Host interface	: SCSI (Ultra160)

ETERNUS LT210

Tape drive	: LTO Ultrium2
No. of drives	: 1
No. of cartridges	: 8
Storage capacity	: 1.6TB**
Host interface	: SCSI (Ultra160)

Copyright 2007 FUJITSU LIMITED

ETERNUS アーカイブストレージラインナップ FUJITSU

階層タイプ

ベーシックタイプ コンテンツ二重タイプ ディスクタイプ

ディスクタイプ: Disk to Disk, Disk to Tape

ベーシックタイプ	コンテンツ二重タイプ	ディスクタイプ
Total Capacity : 18.3TB No. of disk drives : 36	Total Capacity : 30.4TB* Disk Cache : 30.4TB No. of disk drives : 72 No. of cartridges : 152	Total Capacity : 274.7TB** Disk Cache : 38.1TB No. of disk drives : 72 No. of cartridges : 1,314

ETERNUS AS500 アーカイブストレージ

Copyright 2007 FUJITSU LIMITED

まとめ FUJITSU

規制強化 個人情報保護法

e文書法

富士通のストレージソリューションは、
お客様の悩みを解消します。

データセキュリティ コンプライアンス コスト増

情報漏洩 改ざん防止

Copyright 2007 FUJITSU LIMITED

【発行】 サイエнтиフィック・システム研究会

【編集】 データマネジメントを意識したストレージソリューション WG



<連絡先> サイエнтиフィック・システム研究会 事務局
〒105-7123 東京都港区東新橋 1-5-2
TEL:03-6252-2582 FAX:03-6252-2934
e-mail:ssken@ssken.gr.jp

ホームページ : <http://www.ssken.gr.jp/>

- ※ 記載の会社名、製品名、名称などの固有名詞は各社の商標または登録商標です。
- ※ 著作権は各原稿の著者または所属機関に帰属します。無断転載を禁じます。
- ※ SS 研会員限定情報を含むため、SS 研会員にのみ配布しております。