

教育・研究分野のストレージソリューションに 期待するもの

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成果報告

2008年8月26日
サイエンティフィック・システム研究会
システム技術分科会2008年度第1回会合

中京大学情報理工学部
磯 直行

 SIST Chukyo Univ.

ストレージシステム

- ストレージシステムは
縁の下の力持ち
- ストレージシステムは表には出てこないが
システムの全体性能を左右することもある



その設計・導入時点の検討は大変重要

 SIST Chukyo Univ.

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

- 2001年度～
「ネットワーク時代の統合ストレージマネジメントWG」
- 2003年度～
「ストレージを中心としたシステムマネジメントWG」
- 2005年度～
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」



研究・教育機関における
大規模ストレージシステムの管理、運用について議論

ストレージシステムに求められること

- **高速・高可用性**

いまとなつては
当然の技術

観測データやマルチメディアデータ等の大規模データを高速かつ
高可用性に利用できる環境を提供

SAN/NAS

- **データマネジメント**

なかなか
難しい

一貫したデータ管理方法の検討
設計・運用時に考慮すべきポリシーの設定

仮想化、データセキュリティ

- **ソリューション**

観測データや電子メール等、増大する「非構造型データ」への対応
バックアップ

ミドルウェア

最近問題に
なつてきた

構造型データと非構造型データ

- **構造型データ**

データ処理は比較的容易

RDBのレコードのように、あらかじめ意味をもたせてデータを作成しておくことで、インデックス作成等の処理が可能

- **非構造型データ**

ファイル名だけではどうやって取り扱えば良いかわからない！

観測データやWebコンテンツ、電子メールのように、種類や大きさ等のデータ間の関係が不明確な大量なデータ

いずれもストレージシステムにファイルとして保存できるが、非構造型データは適切なものを適切な時間内で探して利用するのは難しい

従来のストレージ技術だけでは不十分

 SIST Chukyo Univ.

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

- **活動主旨**

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

- **活動内容**

会員機関(研究所・大学)におけるデータ特性の分析および研究開発系の構造型データと非構造型データの整理と検討

- **成果**

各業務におけるストレージシステムに求められるソリューションを明確化

 SIST Chukyo Univ.

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 活動メンバー

• 会員側WGメンバー

鈴木富男 (担当幹事・理化学研究所)
磯 直行 (まとめ役・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)
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)

• 富士通側WGメンバー

森屋光弘 (まとめ役)
小野英司
吉田真和
松井泰敏
服部和徳
社本 昇
松本一志
小林英一
大内敦夫
竹内 一

(敬称略)

 SIST Chukyo Univ.

ストレージソリューションを構成する 8つのカテゴリ

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)

データマネジメントを
意識した視点から
ストレージシステムを
分析するためのカテゴリ

 SIST Chukyo Univ.

1.データプロテクション

- 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証すること

一般家庭にも
普及しつつある

- バックアップ
- アーカイブ
- データの選択的保護
- RAID

2.アーカイブ・HSM

- **アーカイブ**
特別に整備された領域にデータを格納する
データ保護技術
- **HSM**
ファイルシステム空間の拡張を実現する技術

研究所ではよく利用されている
多くの大学は利用していないが
重要であることは認識されている

3.ストレージ管理

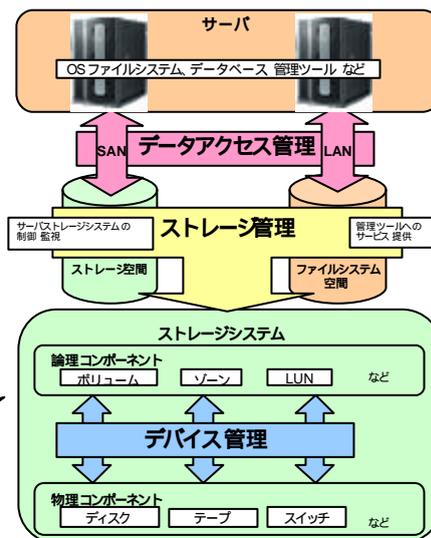
- 大規模化 ,複雑化するストレージシステムをサーバに接続しファイルサーバ等としてアクセス可能にするための論理的な記憶空間 (ストレージ空間やファイルシステム空間)を提供
- アクセスパス ,可用性レベルなどのサービスを提供

階層化されているということ?
よくわからない? それでは...

3.ストレージ管理

- 論理的な記憶空間を提供
 - ・ストレージ空間 (SANの場合)
 - ・ファイルシステム空間 (LANの場合)
- アクセスパスを提供
- 可用性レベルを提供

各種サーバと
ストレージシステムの
関係を図にしてみました !

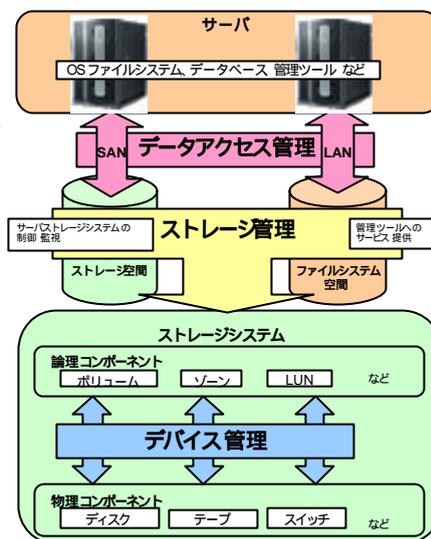


4. データアクセス管理

- 可用性と性能維持のため複数のネットワークパスを制御
- QoSによるデータ転送帯域保証

地味だけれどもとても重要

調査した研究所・大学のすべてで最重要課題

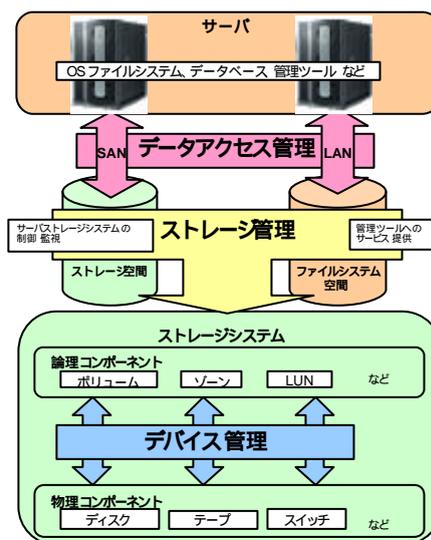


SIST Chukyo Univ.

5. デバイス管理

- システムの可用性を維持
- モジュールの動作状況を監視
- 構成定義の変更
- 装置の追加 交換
- データ移行の容易化

リプレース時や学生が一斉に入れ替わる大学では特に重要



SIST Chukyo Univ.

6 .データレプリケーション

- 「データの複製」のこと
- データアクセスを容易にすることで
処理性能の向上
データの保護レベル向上

データレプリケーションの実現で
「9.11同時多発テロ」でも
ほとんど被害なし

7.ファイルシステム

- 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アクセスさせるためのインタフェースソフトウェア
- 高速性・拡張性の向上が望まれる

CPUの高速化により
ファイルシステムが意外と
ボトルネックになっていることが多い

8.その他 (圧縮 暗号化など)

- データ圧縮
 - ・ファイルブロックの同一性を利用した圧縮
 - ・非構造型データに対してファイルの重複，データの部分重複が多く含まれることを利用したDe - dup技術

- 暗号化

すでにネットワーク管理や
ファイル管理に利用

重複部分を排除した
効率の良い圧縮が可能

8つのカテゴリの重要度

- 参加機関の8つのカテゴリの重要度

データアクセス
管理が最重要

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

8つのカテゴリの評価

- 全機関とも「4.データアクセス管理」が最重要
- 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

8つのカテゴリ以外に考慮すべきもの

- ログ / 解析データの統合管理
- データGridへの対応 など

ストレージソリューションに 期待するもの

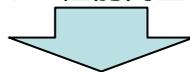
- **利用者の利便性の向上**
 - アクセススピード確保
 - オペレーションの簡素化
 - ハードウェア・ソフトウェア機能の最大活用
 - 遠隔地との連携強化
- **ストレージシステム / データ管理者の手間削減**
 - データ移行時のマルチベンダ・各種ストレージの統合的な状況把握
 - バックアップ / リカバリ時間短縮
 - 信頼性向上
 - 非構造型データの効率的な管理
 - コンテンツ管理機能の向上
- **限られたコストで最大パフォーマンス**
 - ストレージシステムの価格低下
 - 省電力・省スペース

ストレージソリューションに 期待するもの

特に,

- ・非構造型データの効率的な管理
- ・コンテンツの管理機能向上 が望まれているが...

単なる製品レベルの性能向上では対応できない



ハードウェア・ミドルウェア・OSなどの総合的な技術が必要

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に優れたストレージシステムの実現が望まれる

非構造型データを 効率良く管理するには？

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

- 仮想化技術

プラットフォームに依存しない処理環境を提供
ファイルデータの再配置が可能
巨大ファイルシステム空間の実現
大規模データアーカイブシステムの構築

ストレージシステムの
仮想化

- コンテンツ管理ミドルウェア

データ管理, ファイル管理を超えた
ユーザにとって意味のあるファイル群を1つにまとめて管理する
ソリューション

意味のある情報含む
複数のファイルを一括処理

SubWGの紹介

ワークシートガイドラインの作成

- 2003年度～「ストレージを中心としたシステムマネージメントWG」
「ストレージシステムポリシー評価ワークシート」



- 2005年度～「データマネージメントを意識したストレージソリューションWG」
「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」

ストレージシステム導入時に的を絞る
技術解説を加えて管理者の利便性を追及

SubWGとしてWGとは別に活動

SubWGの紹介

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

- 目標
研究・教育機関での大規模ストレージシステム導入
検討および設計段階において、その仕様を策定する
際に検討すべき事項をひとつずつチェックできる

- ガイドラインの構成

チェックシート

技術解説

各項目について互いにリンクを設定

技術解説では、具体的な製品までたどることが可能

SubWGの紹介

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

- 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」の成果のひとつ
- 項目の多さから、単に高速・大容量の観点だけで設計・運用はできないことを再認識
- ストレージシステムは表面には出てこないがシステムの全体性能を左右することもあり、その設計・導入時点の検討は重要