スケーラブルなCAEシュミレーションのためのストレージテクノロジーについて

スケーラブルシステムズ株式会社

目次

- ストレージに関する課題
- CAEにおけるI/O処理での課題
- ボトルネックの解消
  - スケーラブルなI/O処理の実現
  - ワークフローの改善
- Panasas採用事例
- ストレージクラスタ製品事例
- まとめ

ストレージに関する課題

クライアント（エンドユーザ）

クラスタ
- 計算クラスタはI/O処理の終了まで計算を中断
- I/O処理は、クラスタの利用率の低下を引き起こす
- ノード数を増やした場合のスケーラビリティの維持の問題

クライアント
- ジョブの実行終了を待つ
- ユーザ数が増えた場合のスケーラビリティの問題
- ユーザ間でのコラボレーションやデータの共有の問題

BOTTLENECK

クラスタ

BOTTLENECK

従来のネットワークストレージ

BOTTLENECK

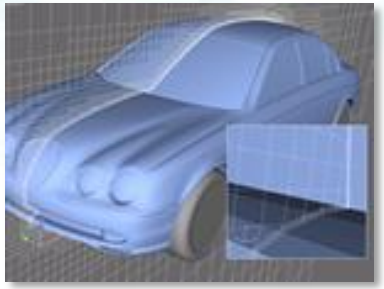
バックアップ/リストア

バックアップ/リストア
- バックアップ処理のためのストレージシステムの負担
- バックアップ実施のタイミング
- 高速でのバックアップの問題

CAEにおけるI/O処理での課題

CAEシュミレーションでのリアリズムの追及への高い要望

- より大きなモデル＋より多くの機能の追加要求=高精細なCAEシュミレーション



高精細なCAEシュミレーションの実現のための分散並列処理

- 高い並列度でのシュミレーション処理の一般化
- X86プロセッサのマルチコアによるコストの削減

高い並列度での処理では、I/Oストリームとファイルサイズの増大が問題

- 計算は高い並列度で分散し、I/Oが逐次処理の場合のボトルネックの問題

CAEにおけるI/O処理での課題

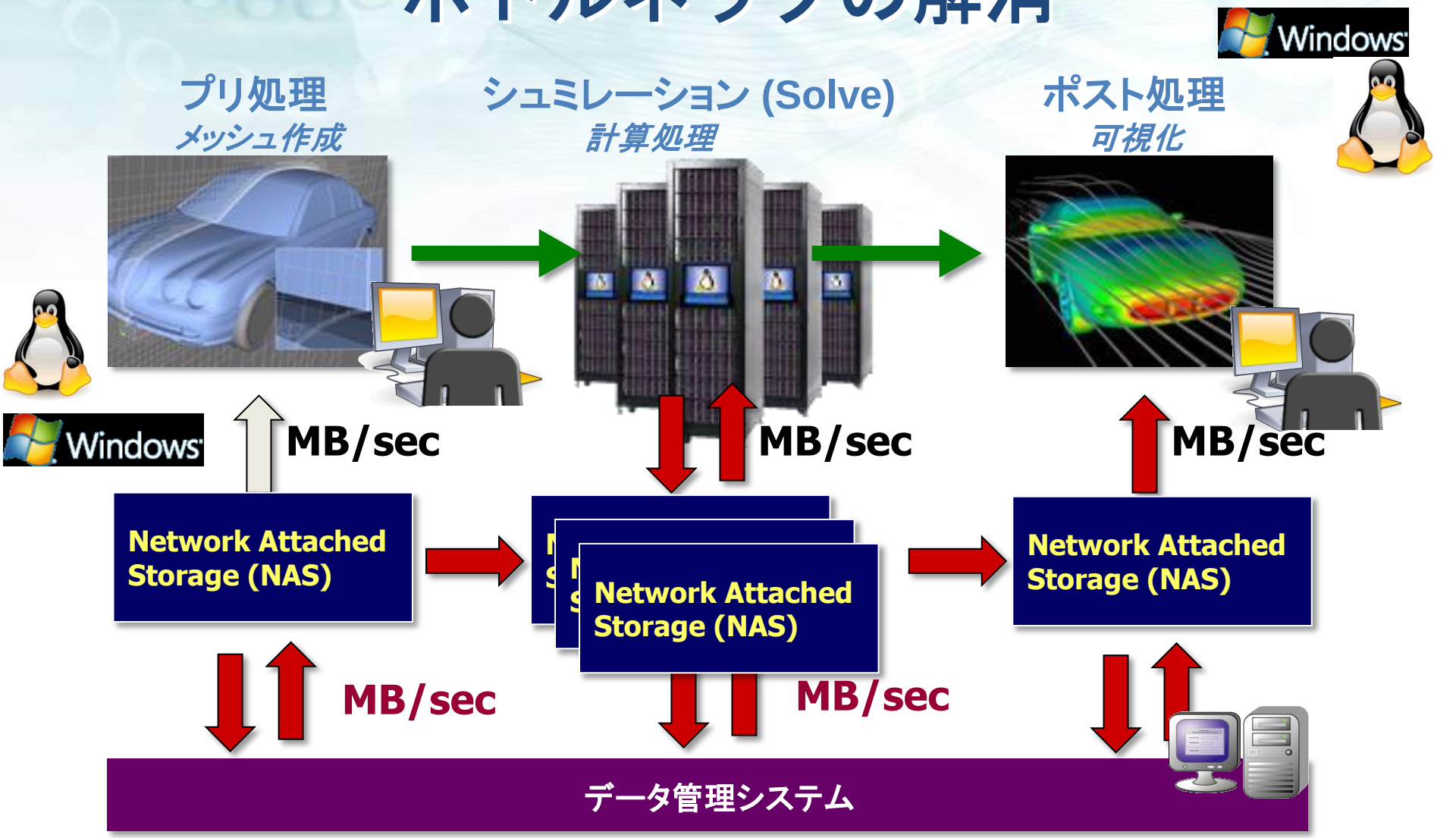
ワークフロー(ユーザのコラボレーションタスク)

- CAEソルバーでのメッシュ分割などのプリ処理の時間
- ネットワークでのファイル転送の負荷
- CAEシュミレーションの解析モデルや計算結果の管理

ワークロード(クラスタでの並列処理)

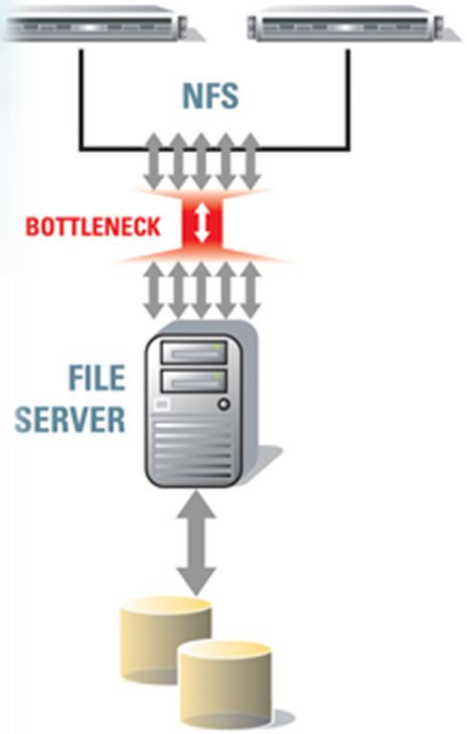
- I/Oリソースに対して様々なワークロード処理
- 非定常計算などでのより頻度の高いデータ書き出し要求
- より自由度の大きな計算処理における‘Out-of-Core’ ソルバーの利用
- 最適化計算などにおけるモデリングの自動化とパラメータ解析
- マルチスケール、多変量、統合モデリングシュミレーションへの対応

CAEにおけるI/O処理での課題
ボトルネックの解消



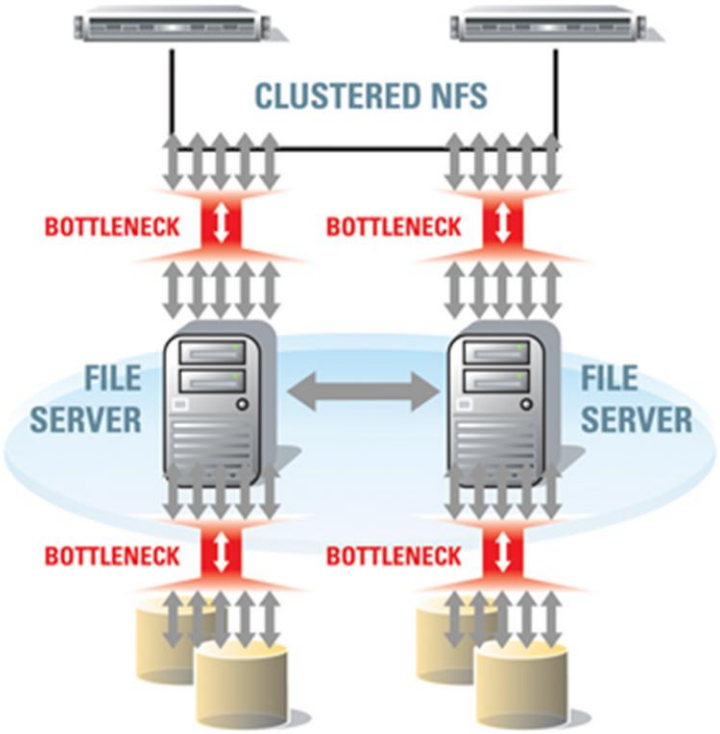
ストレージアーキテクチャ

パラレルストレージの価値の検証



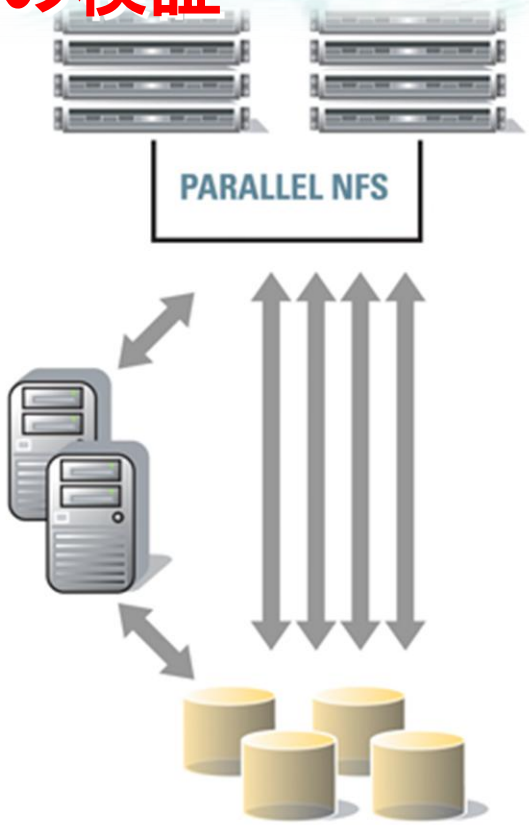
NAS

Network Attached Storage
シリアルI/Oがボトルネック



CLUSTERED STORAGE

複数のNASを統合的に運用管理
個々のNASサーバでのシリアルI/O
がボトルネック

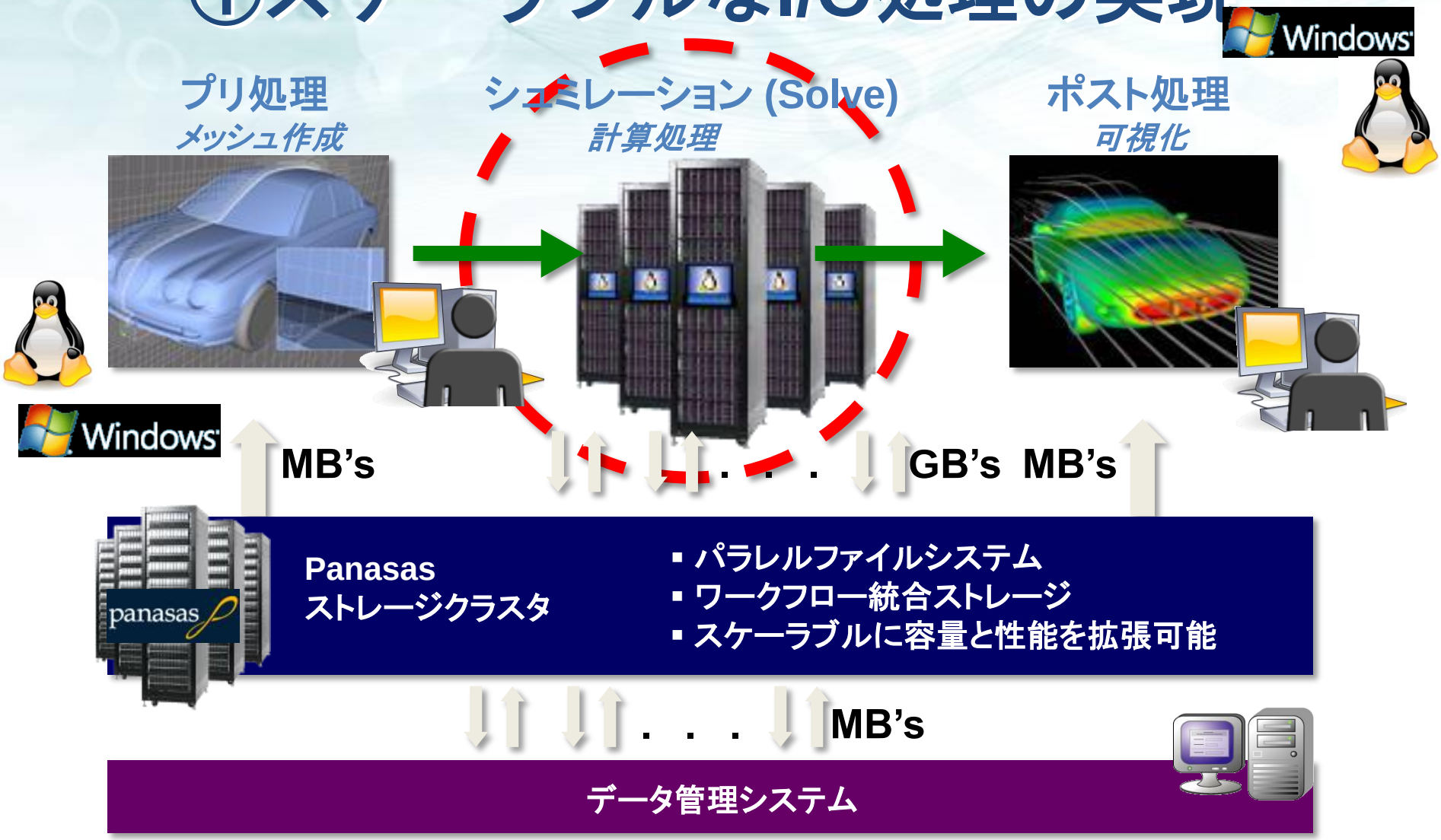


PARALLEL STORAGE

ファイルサーバを経由しないデータ
転送パス
シリアルI/Oのボトルネックの解消と
容易なシステム全体の運用管理

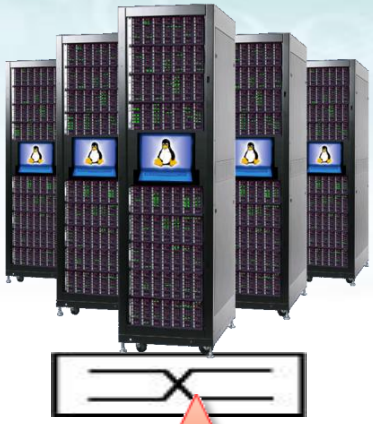
スケーラブルシステムズ株式会社

ボトルネックの解消
①スケーラブルなI/O処理の実現

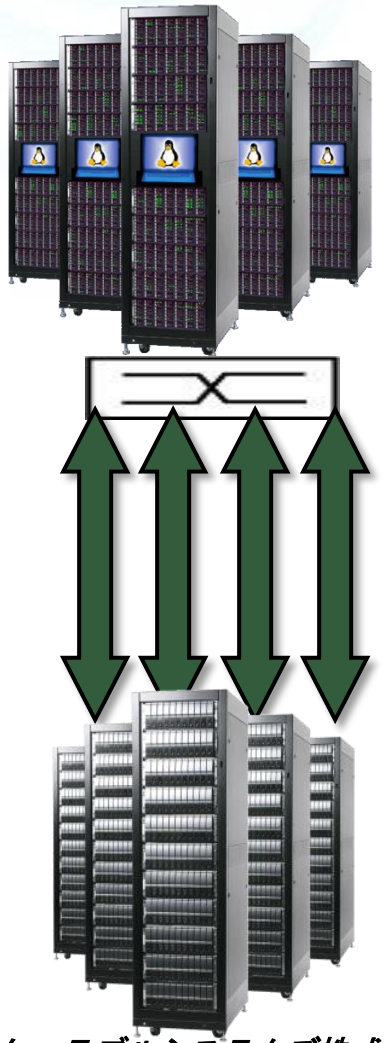


クラスタ利用時のボトルネック

クラスタ⇒パラレルコンピューティング⇒ パラレルIOが必要



一般的なNAS	*パラレルストレージ*
ストレージに対する単一のデータパス	ストレージに対するパラレルなデータパス
スケーリング	スケーラビリティ
限定されたIOバンド幅	高いIOバンド幅
システム拡張の限界	グローバルネームスペース
柔軟性の欠如	容易な運用管理
高価なシステム構成	低価格



CAEにおけるI/Oボトルネック

CAEでのシングルジョブのI/O処理の比重

1999: Desktops

5% I/O

2004: SMP Servers

注意: 以下の推定ではジョブのI/O時間は一定で、並列処理によって、非I/O処理部分の時間は短縮されると仮定しています.....

30% I/O

2009: HPC Clusters

63% I/O

注意: I/O処理部分に関して、性能向上や並列化などの改善がないという極端な仮定での推定であり、実際のCAEでのシングルジョブのI/O処理を完全にシュミレーションした結果ではありません。

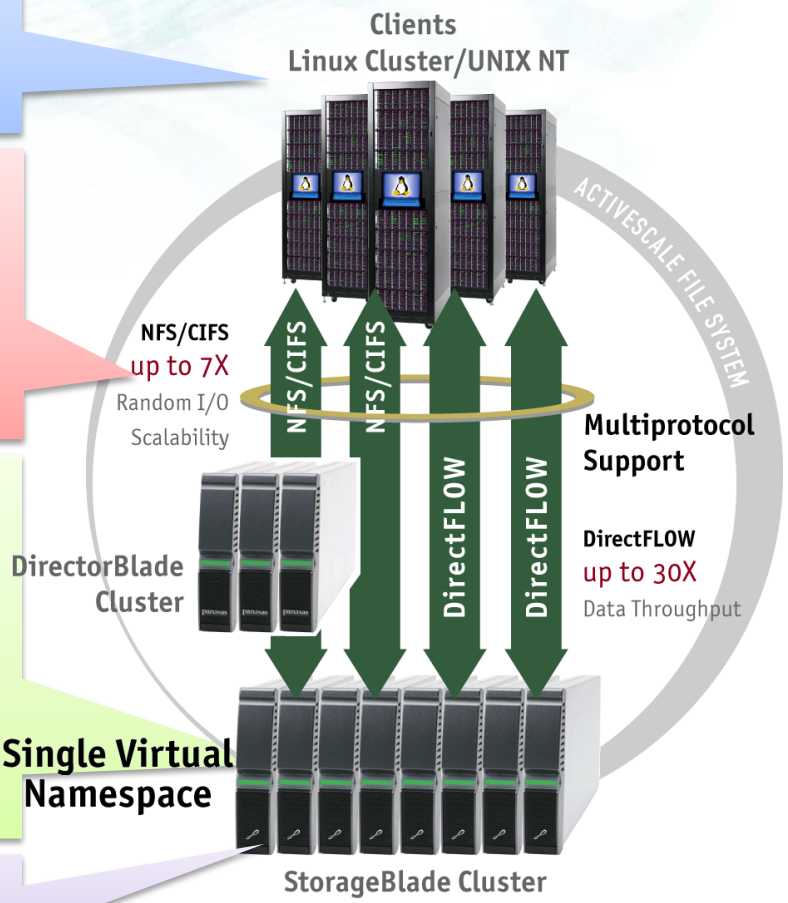
Panasasストレージクラスタ

DirectFLOW クライアントS/W
- *クライアントからの同時アクセスを並列に処理可能*
- RedHat,SUSEなどの主要なLinuxディストリビューションで利用可能
- pNFSにも対応可能

スケーラブルな NFS/CIFS/NDMPサーバ
- 負荷を自動的にストレージクラスタ全体に分散
- クライアント数の増加に合わせてスケーラブルな性能増強が可能
- 全てのDirectorBladeが全てのファイルにアクセス可能

シングルネームスペース
- 同一データへのいずれのプロトコルでのアクセスも可能
- シングルファイルシステム
- DirectFLOW/NFS/CIFS/NDMP間での完全なコヒレンシの実現
- 非Linuxのデバイスをシステムに統合
- *グローバルネームスペースによるシステムの容易な拡張と運用の容易さ*

オブジェクトベース
- 優れたスケーラビリティ、信頼性、運用管理
- ***Panasas Tiered Parityによるデータ保護の強化***



スケーラブルシステムズ株式会社

LS-DYNA 性能評価事例

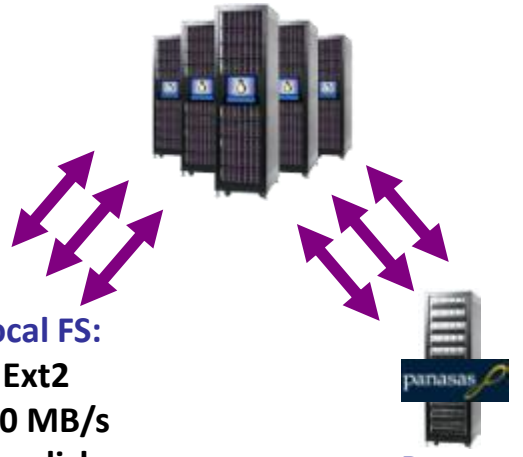
University of Cambridge

UNIVERSITY OF CAMBRIDGE

HPC Service, Darwin Supercomputer

- **Univ of Cambridge DARWIN Cluster**
  - http://www.hpc.cam.ac.uk
- **Vendor:**
  - Dell ; 585 nodes; 2340 cores; 8 GB per node; 4.6 TB total memory
- **CPU:**
  - Intel Xeon (Woodcrest) DC, 3.0 GHz / 4MB L2 cache
- **Interconnect:**
  - InfiniPath QLE7140 SDR HCAs; Silverstorm 9080 and 9240 switches,
- **File Systems:**
  - Panasas PanFS -- 4 shelves AS3000 XC, 20 TB capacity;
  - NFS – Chelsio T310 10Gb ethernet NIC, PERC 5/E RAID Dell MD 1000 SAS 10TB capacity
- **Operating System:**
  - Scientific Linux CERN SLC release 4.6

DARWIN 585 nodes; 2340 cores



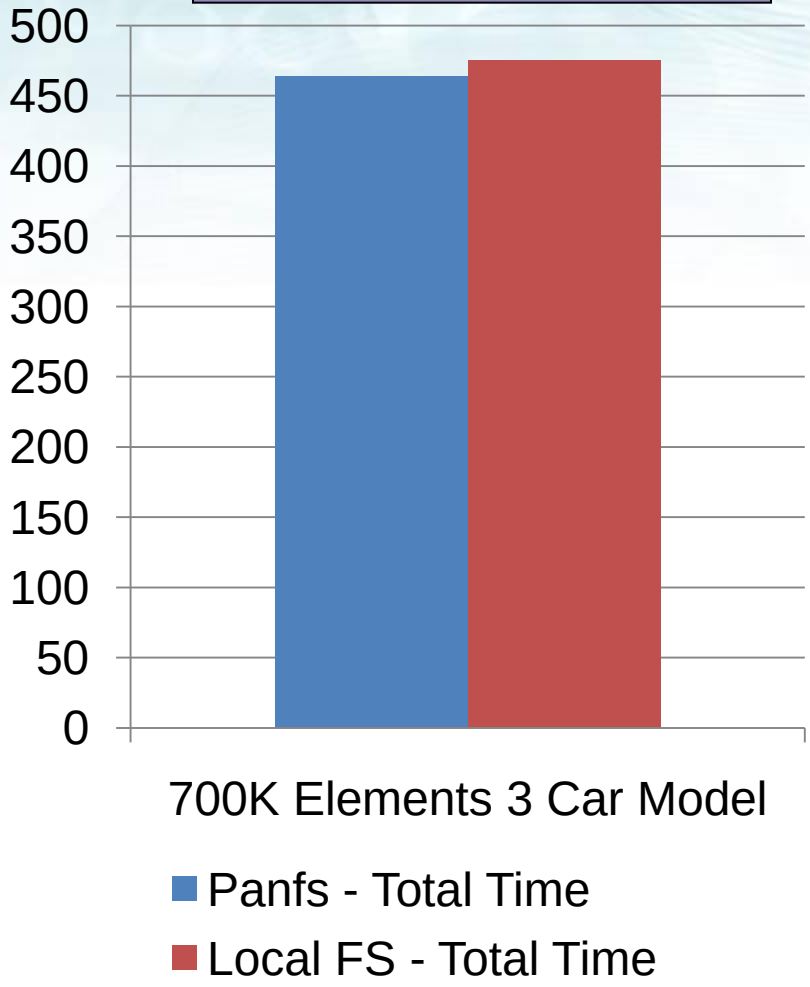
Local FS:
Ext2
~80 MB/s per disk

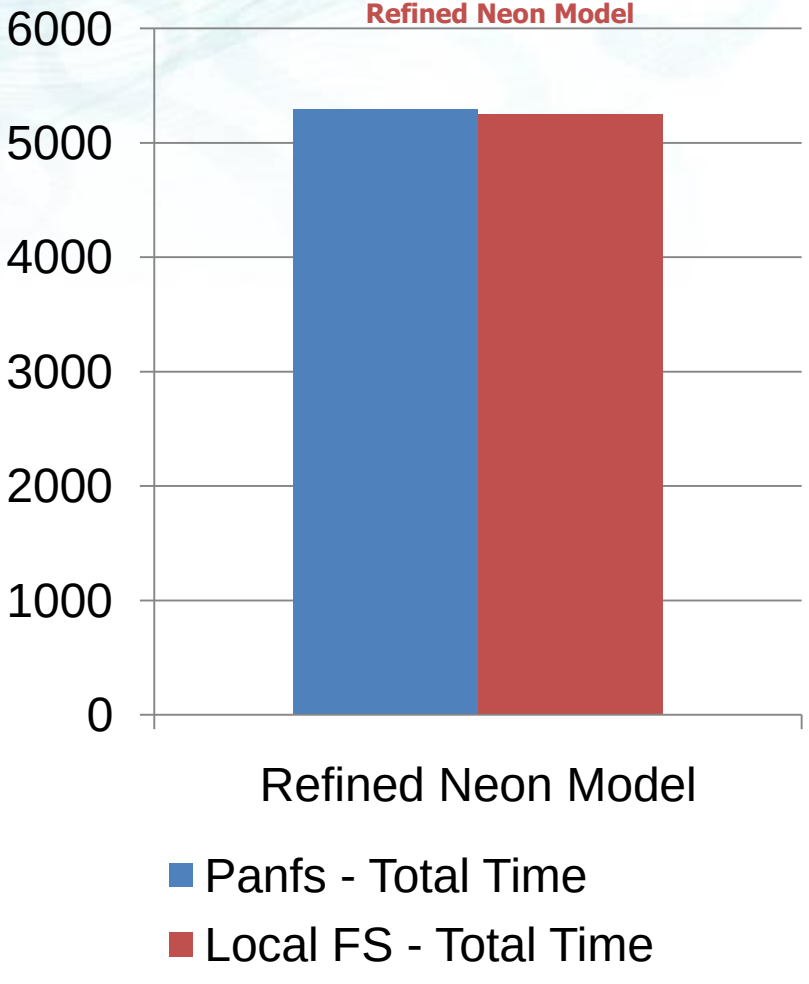
Panasas:
4 Shelves,~ 1 GB/s

Cluster Scalability of Implicit and Implicit-Explicit LS-DYNA Simulations Using a Parallel File System
Stan Posey, Bill Loewe, Ph.D.Panasas Inc, USA Paul Calleja, Ph.D.University of Cambridge, Cambridge, UK
7th European LS-DYNA Conference

LS-DYNA 性能評価事例

LS-DYNA 971 Explicit

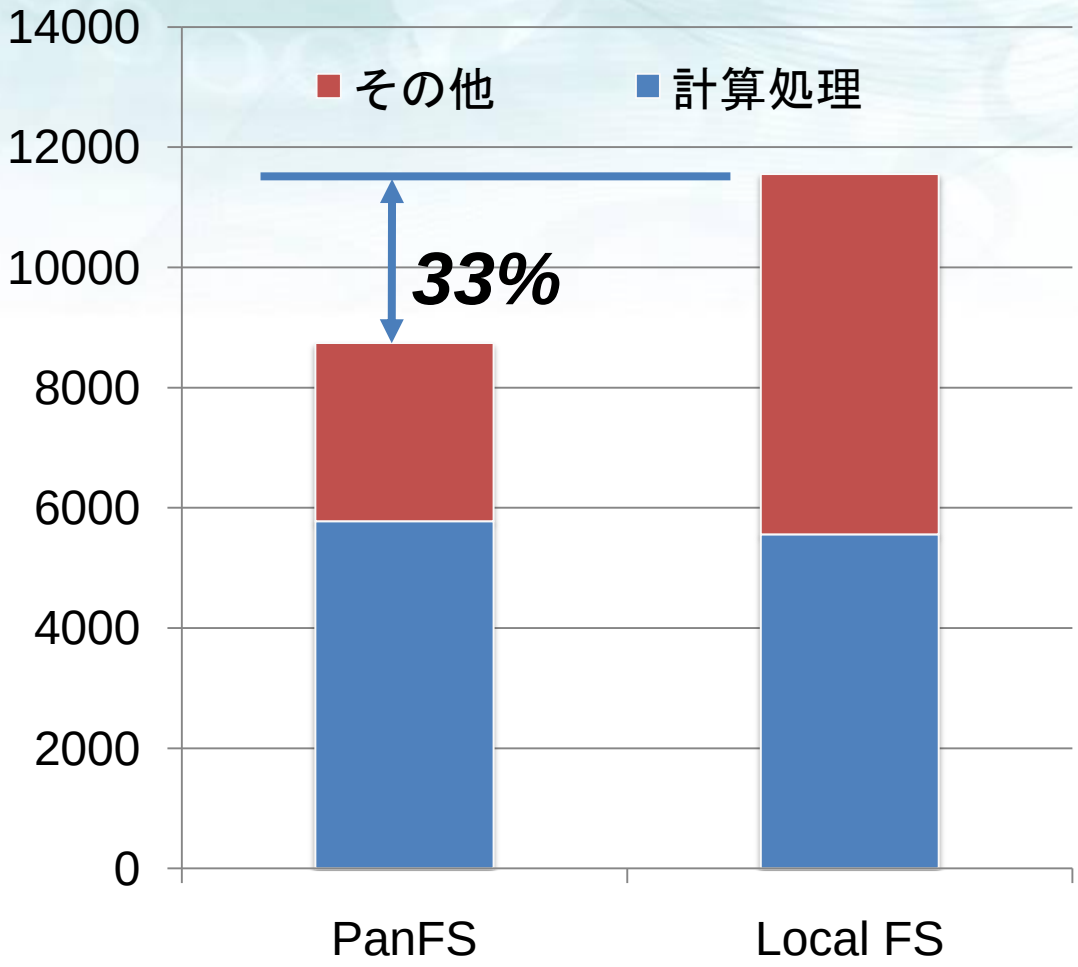


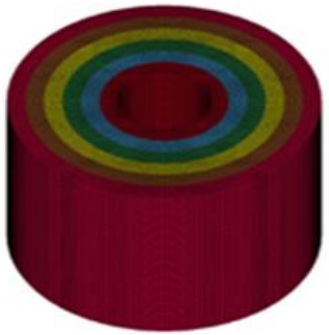


シングルジョブ16コア並列処理

LS-DYNA 性能評価事例

LS-DYNA 971 Implicit





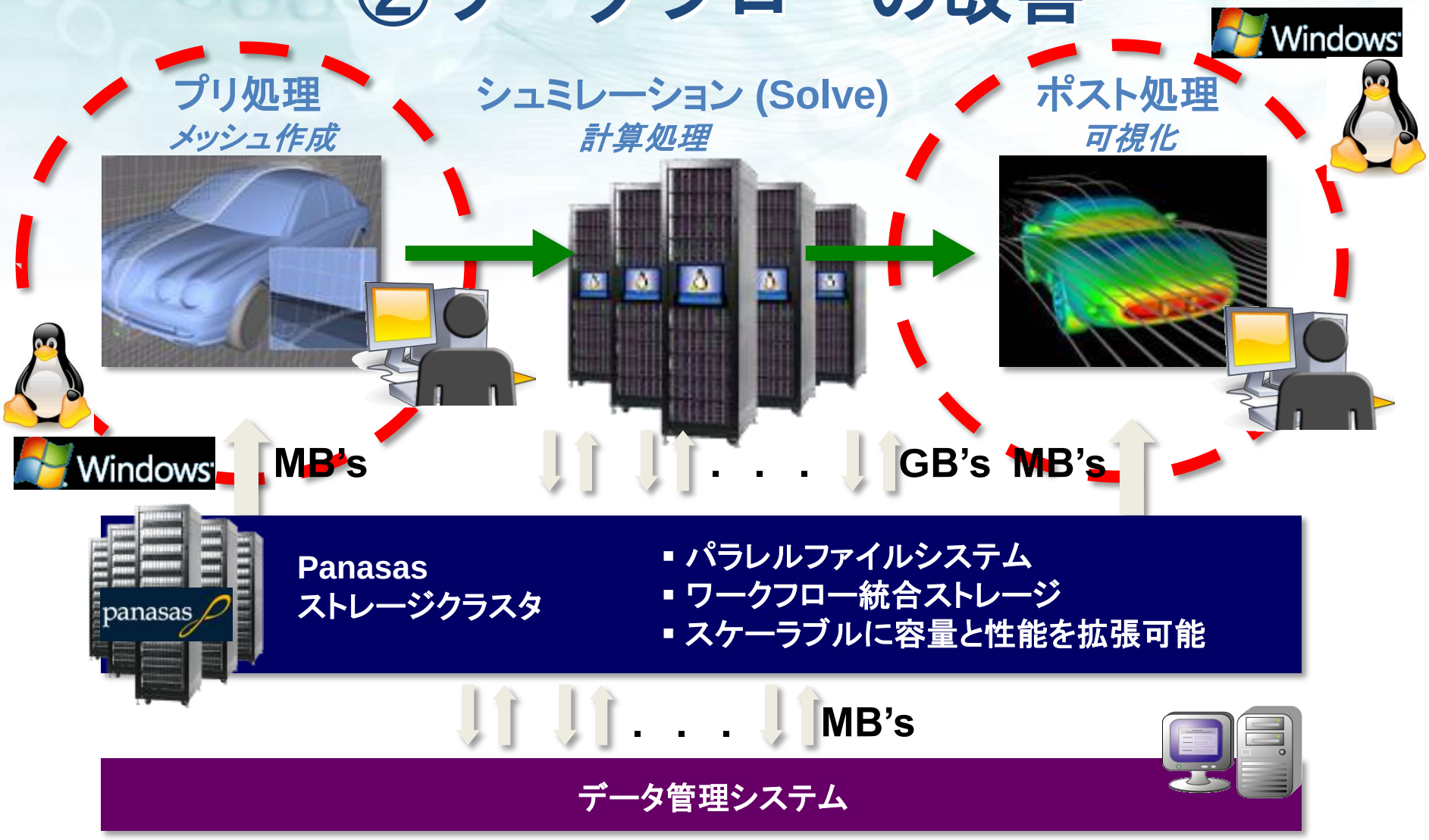
ベンチマークモデル – CYL1E6
- *LS-DYNA v971 Implicit*
- *6 nested cylinders (相互接触)*
- *921,600 ソリッド要素*
- *1,014,751 ノード*
- *3,034,944 (マトリックスサイズ)*
- *非線形時間ステップ（1）,Factor(2), Solvers(2), Force Comp(4)*

シングルジョブ16コア並列処理

Panasas性能評価事例

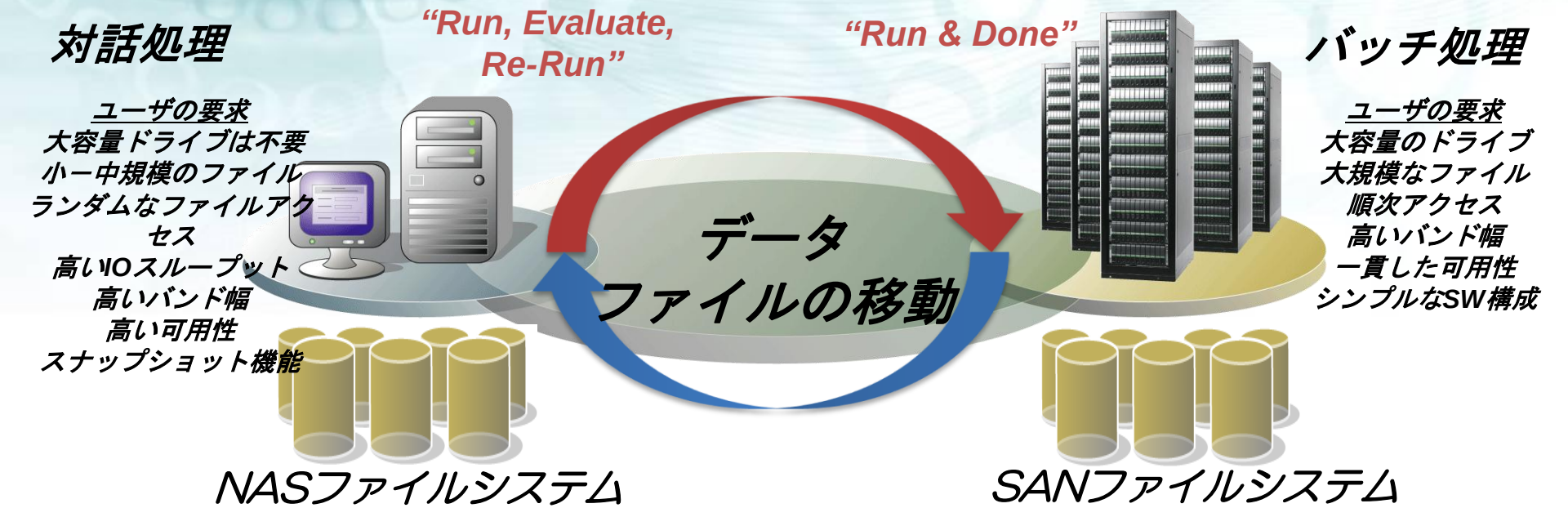
ISV Software	Model Size	#Cores	Advantage
FLUENT 12 ANSYS	111M Cells	128	> 2x vs. NAS
STAR-CD 4.06 CD-adapco	17M Cells	256	1.9x vs. NAS
CDP 2.4 Stanford	30M Cells	512	1.8x vs. NAS
Abaqus/Std 6.8-3 SIMULIA	5M DOFs	multi-job	1.4x vs. DAS
ANSYS 11 ANSYS	SP1 Suite	16	“best NAS ”
LS-DYNA 971 LSTC	3M DOFs	16	equal to DAS

ボトルネックの解消
②ワークフローの改善



スケーラブルシステムズ株式会社

ストレージに対する要求



対話処理	バッチ処理
地下資源探査結果の評価	地質探査解析（大規模データ処理）
EDA デザイン	チップシュミレーション&Tapeout
モデル化、解析結果評価	空力解析、衝突解析
アニメーション処理	レンダリング
トレーディング/ポートフォリオ	リスクマネージメント

ワークフロー統合ストレージ

対話処理

ユーザの要求
大容量ドライブは不要
小ー中規模のファイル
ランダムなファイルアクセス
高いIOスループット
高いバンド幅
高い可用性
スナップショット機能

"Run, Evaluate, Re-Run"

"Run & Done"

バッチ処理

ユーザの要求
大容量のドライブ
大規模なファイル
順次アクセス
高いバンド幅
一貫した可用性
シンプルなSW構成

共有データへの高速で、容易なアクセスが可能
結果が得られるまでの時間を短縮・データの多重保持が不要

対話処理	バッチ処理
地下資源探査結果の評価	地質探査解析（大規模データ処理）
EDA デザイン	チップシュミレーション&Tapeout
モデル化、解析結果評価	空力解析、衝突解析
アニメーション処理	レンダリング
トレーディング/ポートフォリオ	リスクマネージメント

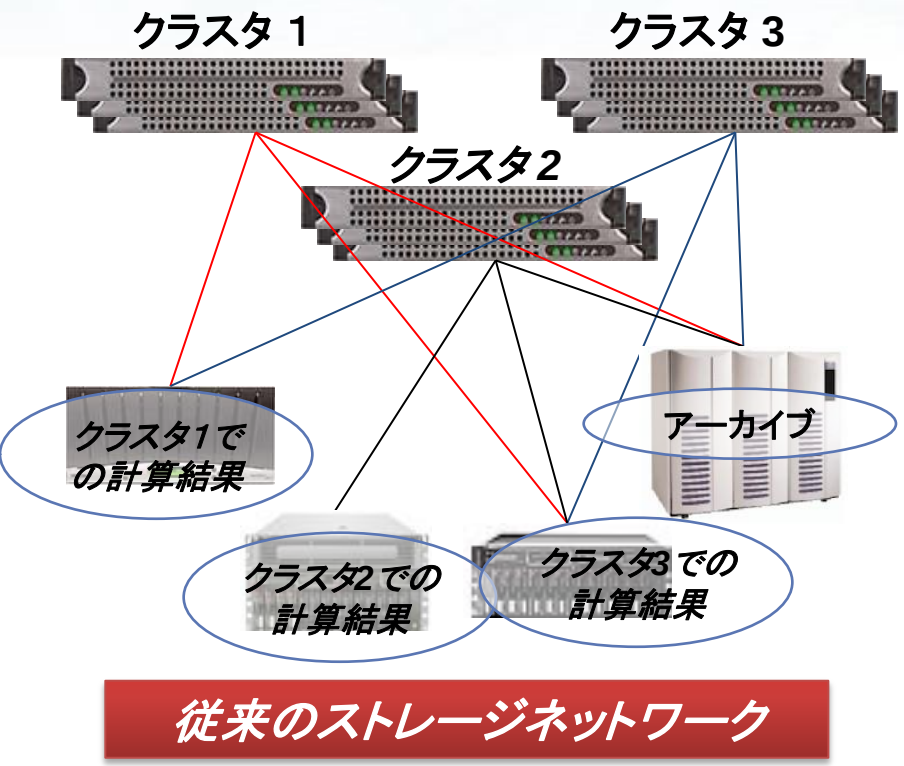
スケーラブルシステムズ株式会社

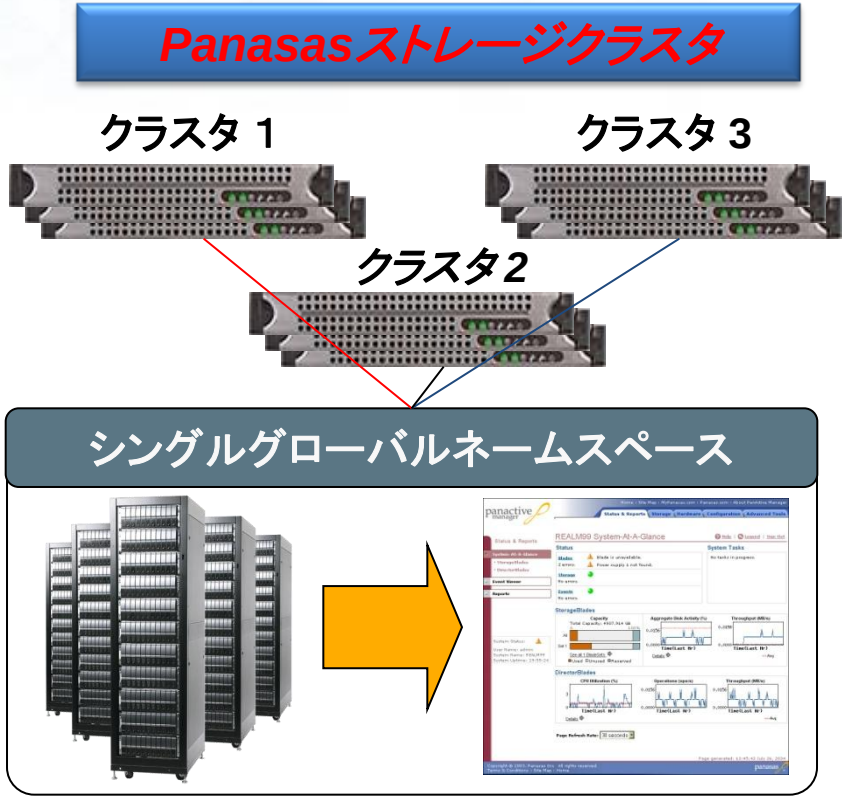
シングルグローバルネームスペース

物理的な境界も論理的な境界も存在しない

クラスタ間でのクロスマウントやデータの移動の排除

自動的プロビジョニング：追加したブレードは自動認識され、ストレージプールに追加される





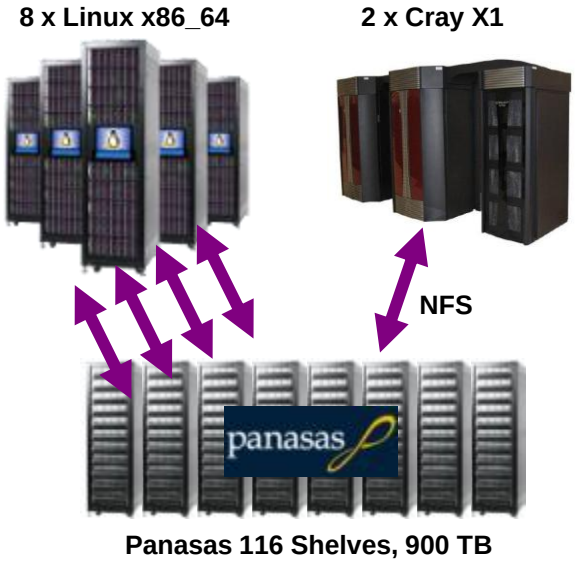
スケーラブルシステムズ株式会社

Pansas採用事例

Boeing Company

CAG & IDS, Locations in USA

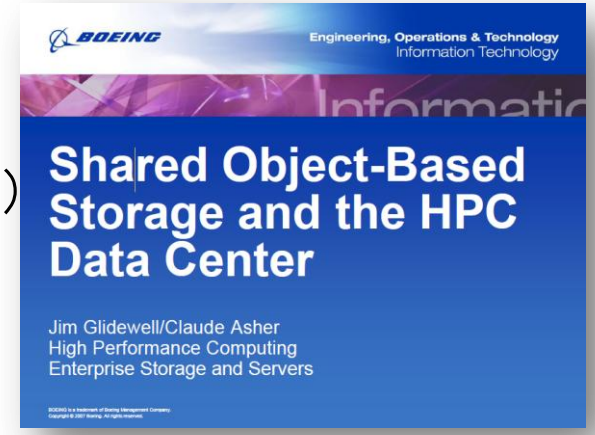
- 利用用途
  - HPCクラスタシステムでの利用
  - 非常に多くのユーザの様々なCAEシュミレーション
    - CFD (Overflow; CFD++; FLUENT)
    - CSM (MSC.Nastran; Abaqus; LS-DYNA)
    - CEM (CARLOS)
- HPCシステム
  - Linuxクラスタ (~3600プロセッサコア)＋Cray X1 (512プロセッサコア)
  - Panasas PanFS, 112 storage systems, > 900 TBs
- 利用効果
  - 高いスケーラビリティと複数ジョブ、複数ユーザの様々なワークロードに対する効率的な処理



Pansas採用事例

Panasasの採用理由

- パラレルファイルシステムが要求要件
  - I/O負荷の大きなジョブと複数のジョブのI/O処理を同時に効率良く処理可能
  - システム全体で高いI/Oバンド幅の要求
- “Production-Ready” ソリューション
  - 導入が容易で直ぐに既存のコンピュータ環境に組み込み利用可能
  - 増設が容易でシステムがスケーラブル
  - システムの負荷分散を動的に実行可能
  - 高い可用性
- TCO削減
  - 導入コスト（コモディティコンポーネント）
  - GbE,10GbE,InfiniBandなどの選択肢
  - 管理運用が容易



Panasas ActiveStor SERIES 製品ライン

Performance Module

ホットスワップ可能
ブレードアーキテクチャ

DirectorBlade
メタデータ処理

StorageBlade
オブジェクトデータ処理

セカンドネットワーク
スイッチ（オプション）

GbE、10GbE
ネットワークスイッチ

バッテリモジュール
電源バックアップ

冗長化電源
ホットスワップ可能

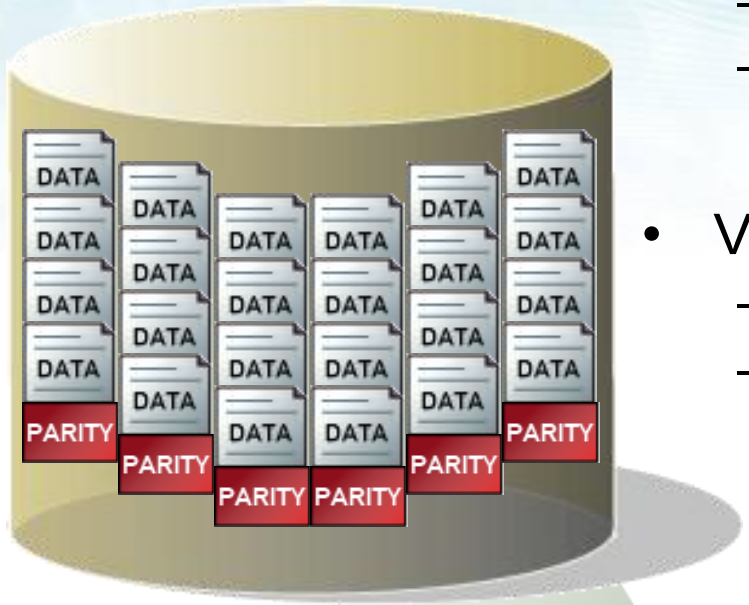
Panasas ActiveStor

スケーラブルストレージアーキテクチャ

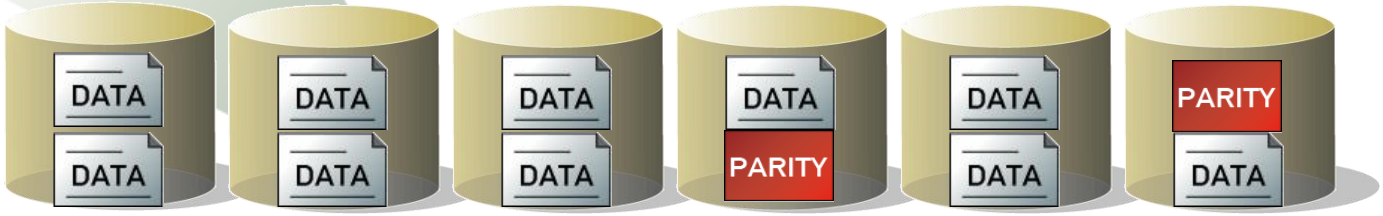
スケーラブルシステムズ株式会社

Panasas Tiered Parity

- Horizontal Parity
  - 従来からのRAIDに相当
  - PanasasのObjectRAIDは、最先端のRAID技術の選択機能と性能と信頼性の向上を図る再構築技術を提供
- Vertical Parity
  - 個々のドライブ内での”RAID“構成
  - ディスクメディアの高密度化が進んで、メディアエラーの発生頻度の確率が大きくなっても、その問題に対する有効な対策



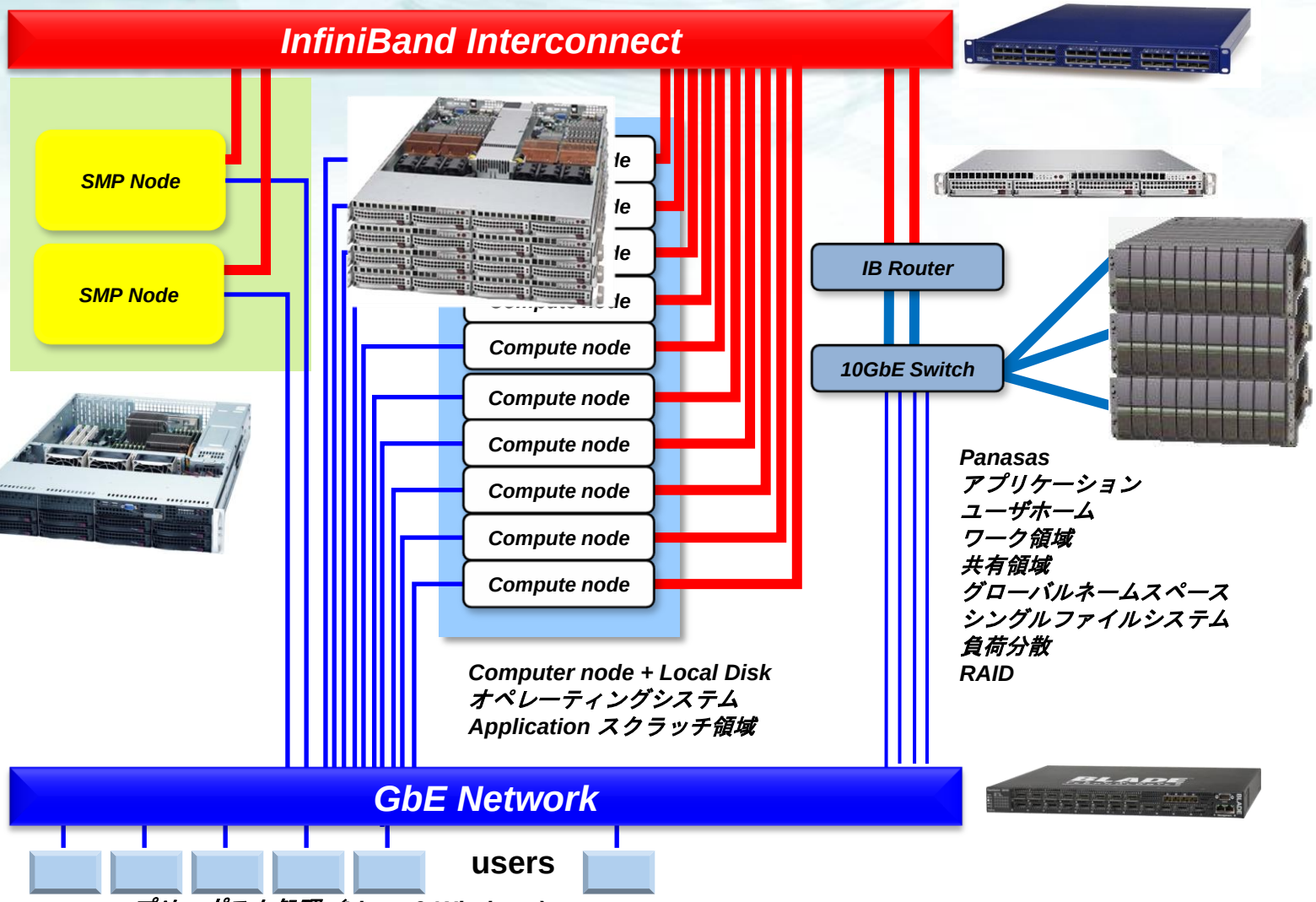
Vertical Parity



Horizontal Parity

All-In-One HP^{2}C システム構成

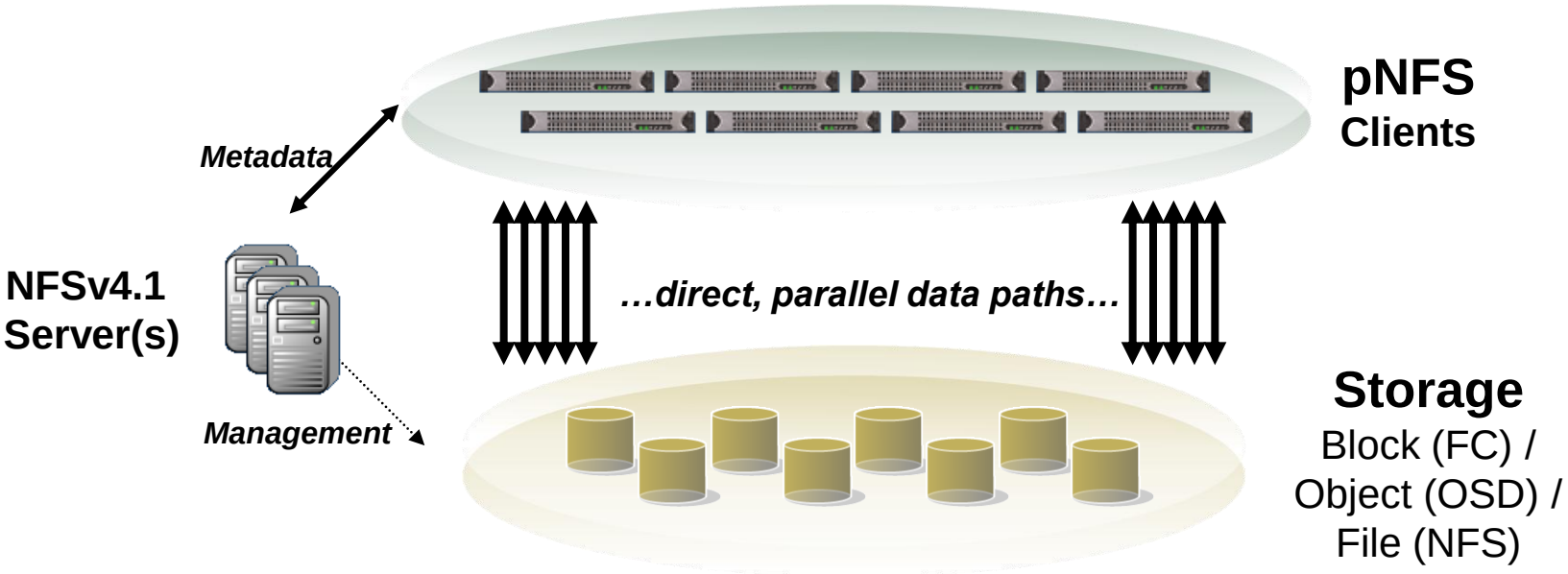
InfiniBand & Panasas IB Router



スケーラブルシステムズ株式会社

pNFS: 標準パラレルNAS

- pNFS は、Network File System v4 プロトコル規格の拡張
  - パラレルかつダイレクトでのデータアクセスが可能
  - ストレージデバイスは、複数のストレージプロトコルをサポート
  - NFSサーバはデータパスに直接介在しない



Panasas ActiveStorの特徴

機能とその利点	Panasas ActiveStor	NAS	SAN
ターゲットとするアプリケーション	バッチ処理 +対話処理	対話処理	バッチ処理
高いバンド幅	◎		◎
クライアント数のスケーラビリティ	◎		◎
ストレージ容量のスケーラビリティ	◎		◎
NFSとCIFSのサポート	◎	◎	
統合システム	◎	◎	
可用性	◎	◎	
高いランダムIO性能	◎	◎	

まとめ：CAEワークロードでの利点

ワークフロー統合ストレージ:CAEワークフローでのコラボレーション

- CAEシュミレーションとプリ・ポスト処理のデータ共有の効率化
- 複数プロトコルサポートによるプラットフォーム非依存での共有データへのアクセス

パラレルI/O:高いI/O性能とボトルネックの解消

- CAEシュミレーションでの生産性の向上
- 高いシングルジョブ性能(スケーラビリティ）
- 複数ジョブでのスループット

NFSとCIFSサポート:システムインテグレーション

- 異機種混在のCAE環境における複数プロトコルのサポートによる容易なシステム導入と運用

シングルネームスペース:ITマネージメントのオーバヘッドの低減

- ストレージマネージメントとデータ管理をシンプルに実行可能な運用管理機能と増設時の容易なオペレーション

お問い合わせ

0120-090715

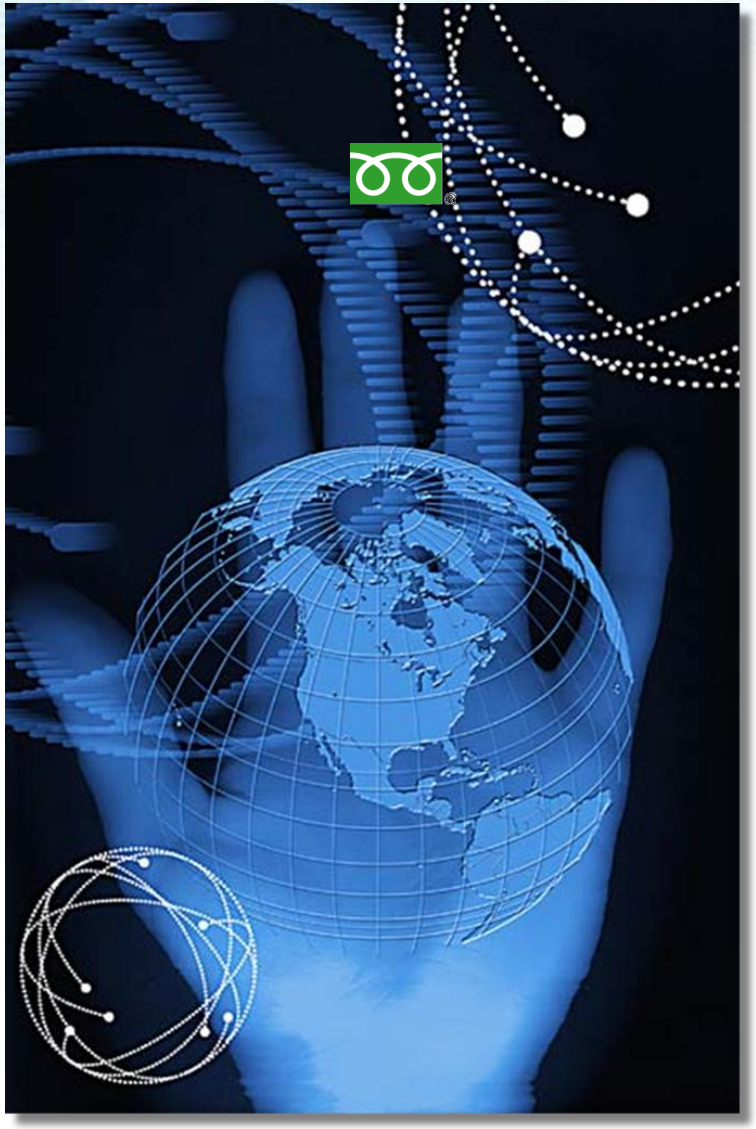
携帯電話・PHSからは（有料）

03-5875-4718

9:00-18:00 (土日・祝日を除く）

WEBでのお問い合わせ

www.sstc.co.jp/contact



10/28/2009

スケーラブルシステムズ株式会社