

「FEFS」の機能のご紹介

■ FEFsの特長的な機能

- ラウンドロビン分散機能(高バンド幅のI/O)
- ユーザー間フェアシェア機能/ノード間優先制御機能
- 冗長機能
- ETERNUS連携(高速バックアップ)

■ その他の機能

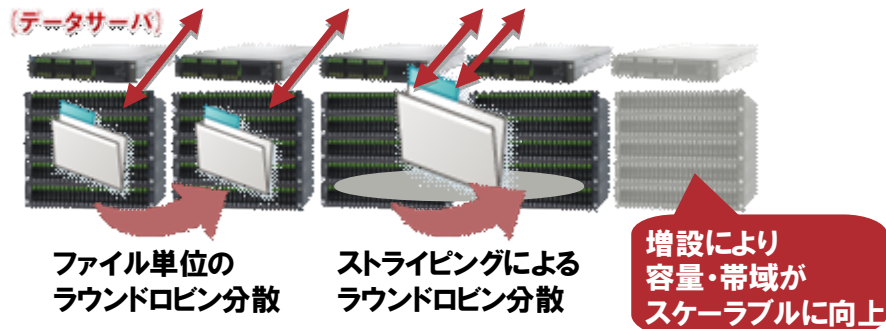
- データ移行
- 他社接続

FEFSの機能紹介

■ FEFSの特長的な機能のご紹介

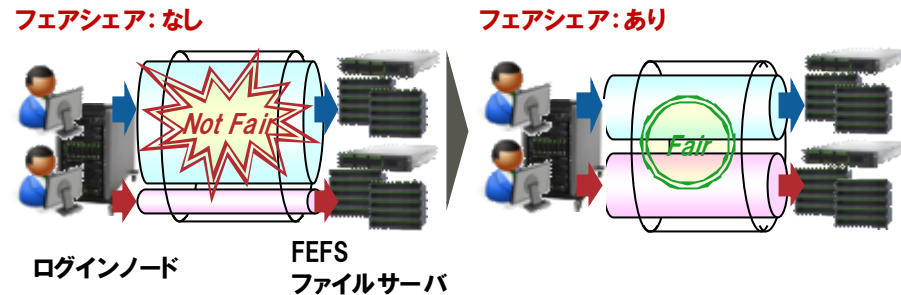
ラウンドロビン分散機能(高バンド幅のI/O)

システムトータルで実効数百GB/s超のスループット実現
 ファイルをラウンドロビンで分散格納しサーバ全体を並列稼働



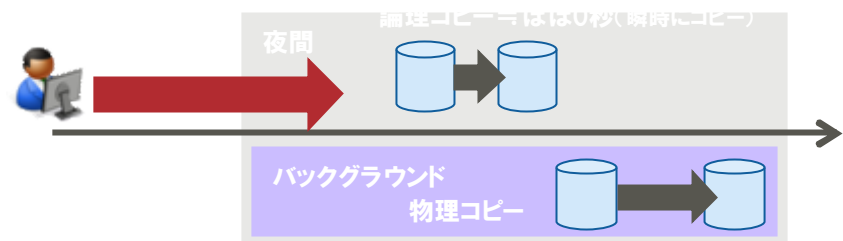
ユーザー間フェアシェア機能/ノード間優先制御機能

特定ユーザーにI/O帯域(サーバ処理能力)を占有させない
 クライアント側:各ユーザーのI/O要求をサーバに均等に発行
 サーバ側:各ユーザーのI/O要求を均等に処理



ビジネス分野で実績あるテクノロジーを採用
 単点故障のないハード構成、サーバの動的交替で実行中ジョブ継続

冗長機能

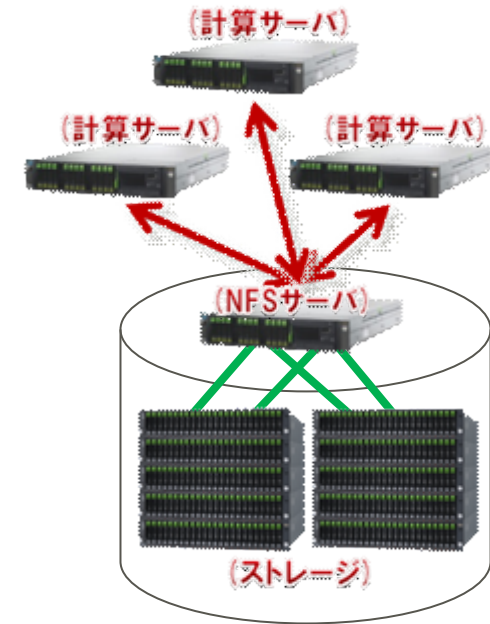


ETERNUS連携(高速バックアップ)

ラウンドロビン分散機能(高バンド幅のI/O)

■ 現状の問題点

- PCクラスタシステムでは、プログラムやデータを共有するためにNFSを使用することは、極当たり前の選択。
- PCサーバ(計算ノード)のコストパフォーマンスが向上し、数十台規模のPCクラスタ構成になる。
- 全体の処理時間の中でNFSサーバが処理する時間が占める割合が高くなり、I/Oボトルネックになる。

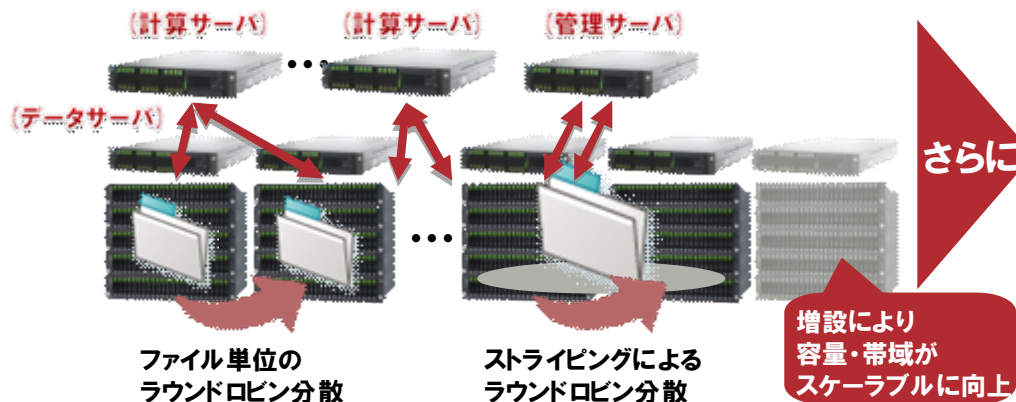


■ 解決するには

- NFSサーバを使用せず解消するために、PCサーバとSAN対応ディスクアレイを組み合わせ、スケーラブルな分散アクセスのファイルシステムが必要。

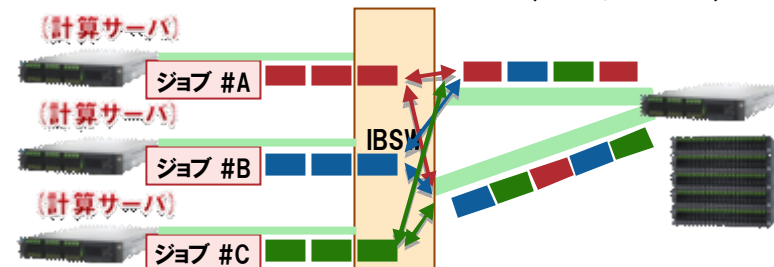
■ FEFSラウンドロビン分散機能

- ファイルをラウンドロビンで分散格納しサーバ全体を並列稼働
 - ・ 小ファイル(～TB) → ファイル単位で分散格納
 - ・ 大ファイル(TB～PB) → ファイルをノード間ストライピングして格納



■ InfiniBandマルチレール機能 (InfiniBandマルチパス構成時の機能)

複数InfiniBandを使用してI/Oデータの実効バンド幅を向上
(マルチレール)



ユーザー間フェアシェア機能 / ノード間優先制御機能

■ 現状の問題点

- さまざまなプロジェクト、グループ、部門の設計・解析者がPCクラスタシステムを共同利用する環境になる。
- 高いI/O帯域の業務に占有され、均等なアクセスレスポンスが得られなくなる。

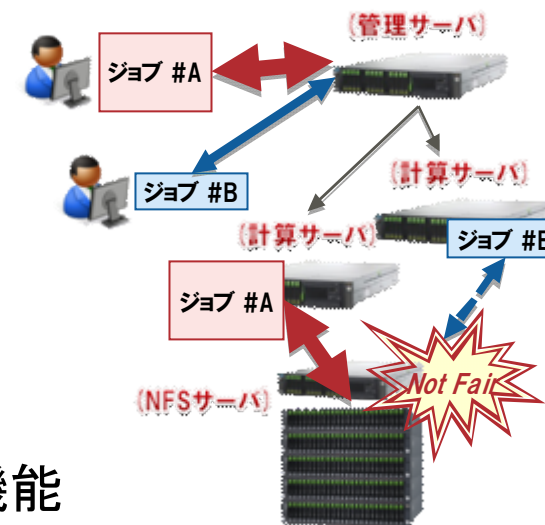
■ 解決するには

- ユーザー間/ノード間の均等なアクセスレスポンスを得るために、I/O帯域の制御ができるファイルシステムが必要。

■ FEFSユーザー間フェアシェア機能 / ノード間優先制御機能

※下記例はユーザー間フェアシェア機能

- ユーザーA: 19ノード・ジョブ
- ユーザーB: 1ノード・ジョブ → 10,000ファイルの生成 / 削除にかかる時間を測定



ユーザー B 10000 Files	フェアシェア無 シングルユーザー	フェアシェア無 マルチユーザー	フェアシェア有 マルチユーザー
Create Files	4 sec	10 sec	4 sec
Remove Files	4 sec	14 sec	4 sec



冗長機能

■ 現状の問題点

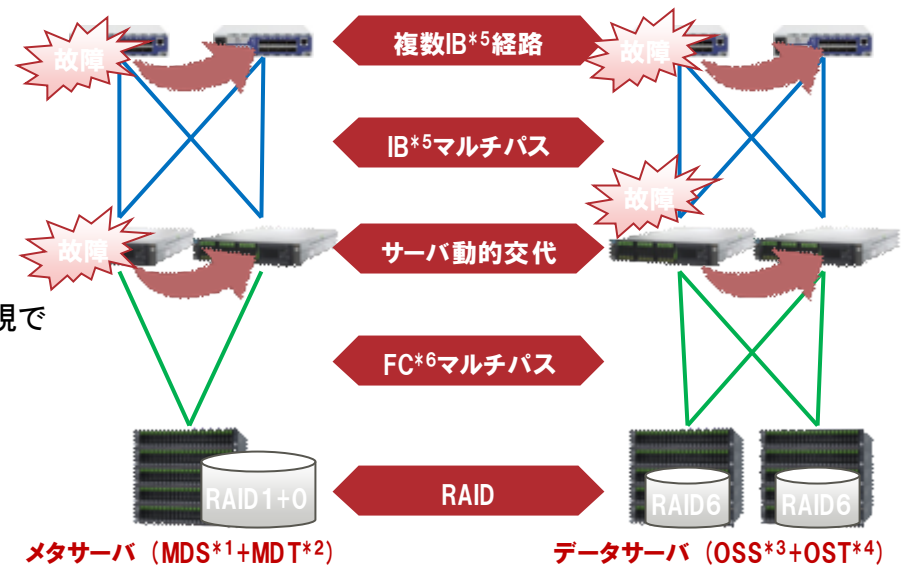
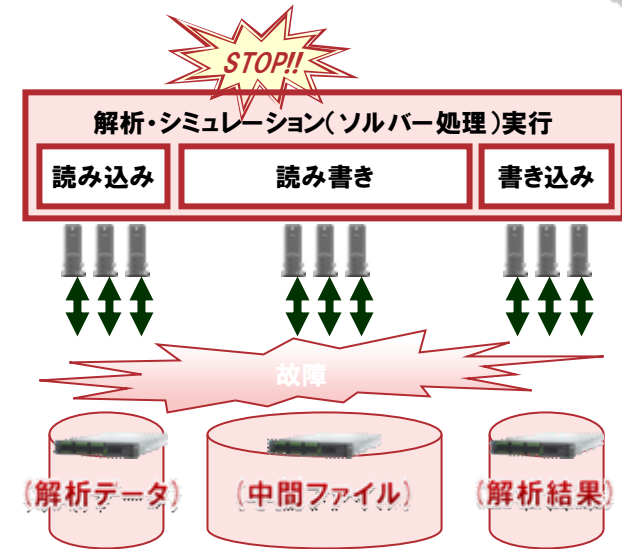
- さまざまなプロジェクト、グループ、部門の設計・解析者がPCクラスタシステムを共同利用する環境になる。
- ファイルシステムのどこかが故障ただけで解析・シミュレーション業務がストップしてしまう。

■ 解決するには

- 各点で冗長構成・自動切り替えができるファイルシステムが必要。

■ FEFS冗長機能

- 単点故障のないハード構成(冗長構成)
 - ・ InfiniBand/ファイバチャネルパス、サーバの二重化による耐故障性強化
 - ・ RAIDによるデータ保護(メタサーバ:RAID1+0、データサーバ:RAID6)
- サーバの動的交替(MDS/OSS)
 - ・ 交替時間はシステムの構成にもよるが、おおよそ 数分で切り替え。走行中のジョブは継続
 - ・ 他のファイルシステム場合、サーバ/クライアント間の監視でタイムアウト時、I/Oをキャンセル(=ジョブキャンセル)



ETERNUS連携(高速バックアップ)

■ 現状の問題点

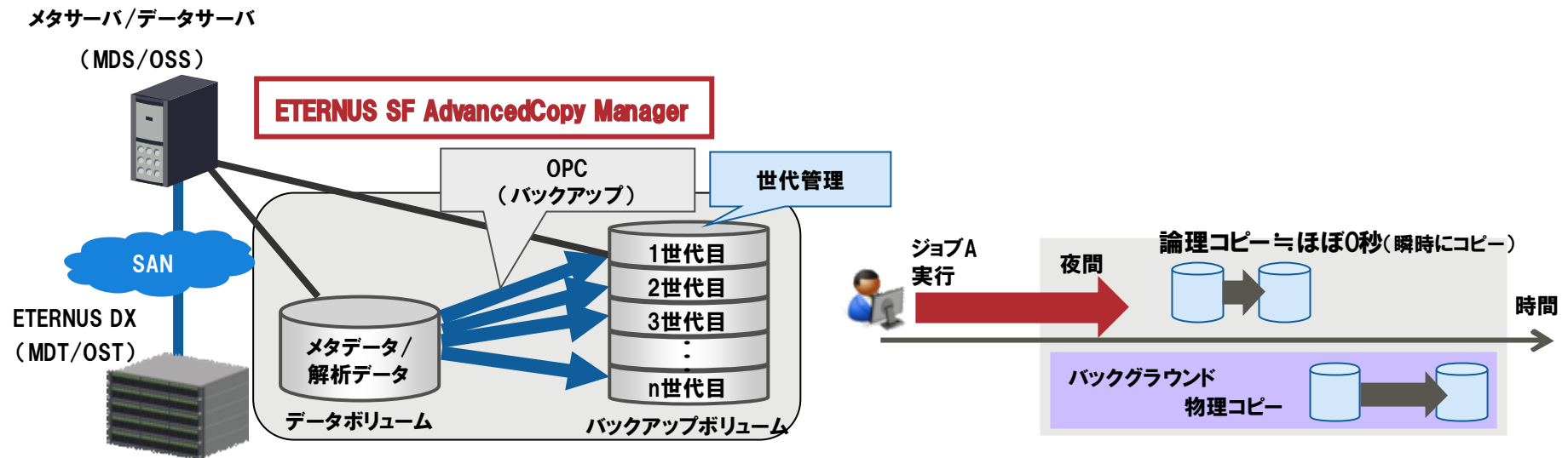
- PCサーバ(計算ノード)のコストパフォーマンスが向上し、数十台規模のPCクラスタ構成になると同時に、大きな規模のモデル解析、小さな規模で多くのモデル解析を実行できる環境になる。
- ファイルシステムに蓄積される解析・シミュレーションのデータが大容量となり、データが持つ重要性からデータ保護のためのバックアップに時間を要し、解析・シミュレーションが実行できない時間が増える。

■ 解決するには

- 解析・シミュレーションの実行にほとんど影響を与えない、高速バックアップができるファイルシステムが必要。

■ FEFSのETERNUS連携(高速バックアップ)機能

- OPCと組み合わせることにより、ファイルサーバなどに求められる多世代バックアップ要件にも対応。

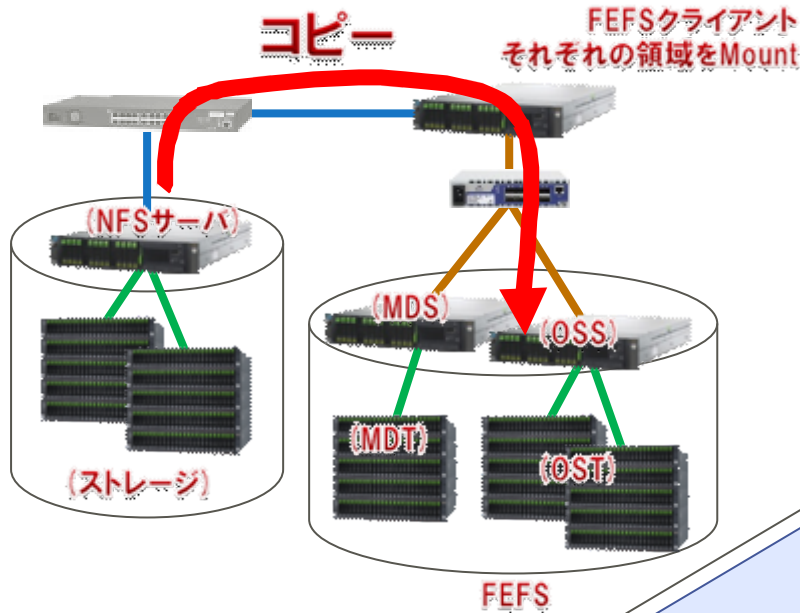


データ移行

■ 良くご利用されているファイルシステムからのデータ移行例をご紹介します。

NFS → FEFS データ移行

NFSのファイルをFEFSにネットワーク経由でコピーします。



転送帯域“1Gbps”でのコピーのため、データ移行に時間を要します。

転送帯域“40Gbps”でのコピーのため、速くデータ移行ができます。

データ移行に際し、

データ移行の調査ポイント

データの持ち方

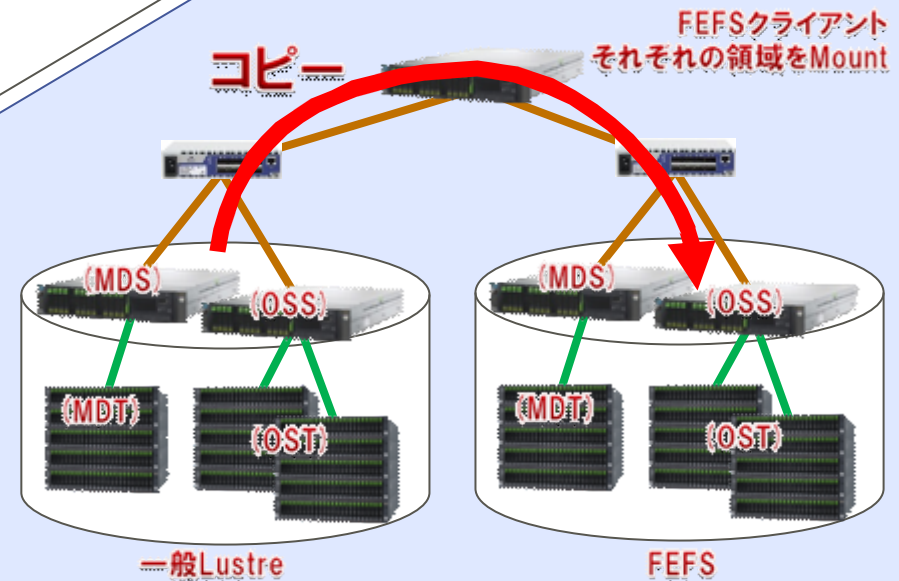
データの形式

データの容量

データ加工
移行ツール
移行方式

データ移行
スケジュール

ご紹介のデータ移行の2つ例にあてはまらない場合はデータ移行方法について別途、ご相談ください。

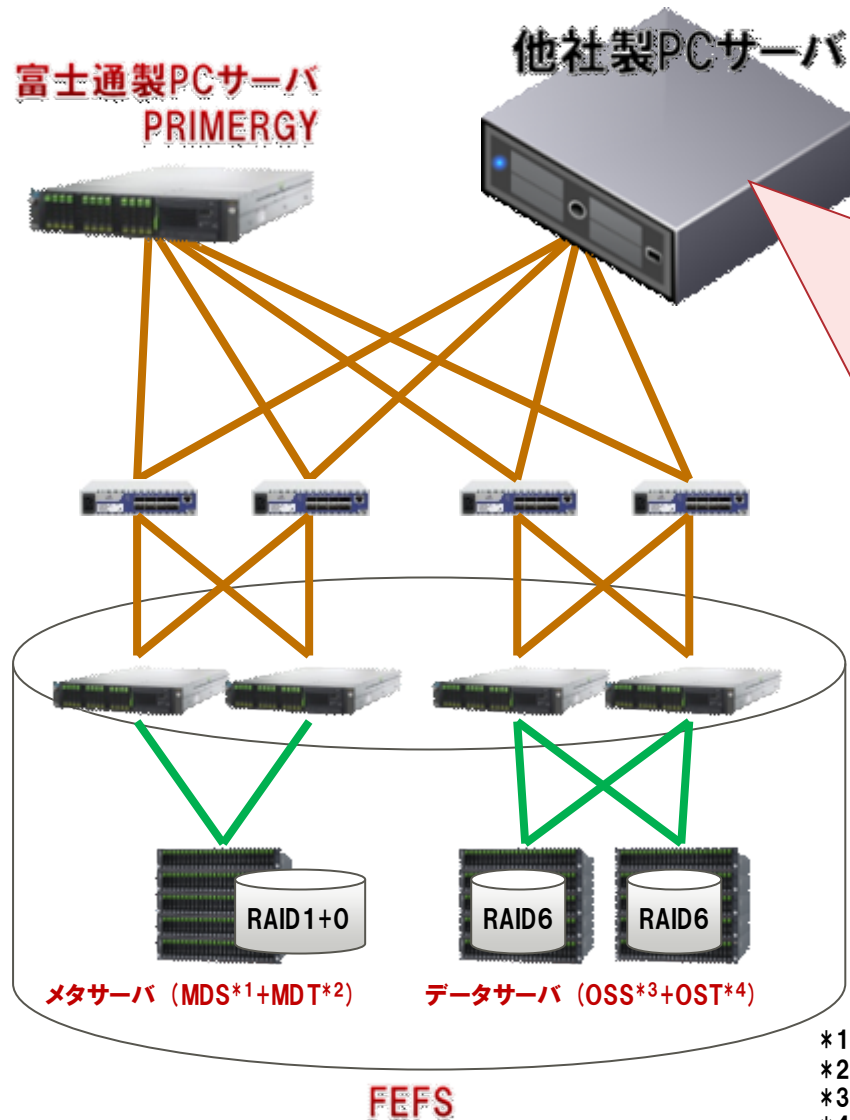


一般Lustreと互換があり、FEFSクライアントにログインしてコピーします。

一般Lustre → FEFS データ移行

他社接続

- 他社PCサーバをクライアントとして、FEFSをご利用できます。



他社製PCサーバをクライアントとしてFEFSをご利用するにあたり、以下を留意する必要があります。



InfiniBand HCAカード

➡ Mellanox社の InfiniBand (QDR) HCAカードである必要があります。



OS

➡ Red Hat Enterprise Linux 5.6で富士通がサポートしているカーネルバージョンである必要があります。

その他詳細についてはご相談ください。

- *1 MDS: Meta Data Server (メタデータを管理するサーバ)
- *2 MDS: Meta Data Target (MDSに接続するストレージ)
- *3 OSS: Object Storage Server (ファイルデータを制御するサーバ)
- *4 OST: Object Storage Target (OSSに接続するストレージ)

《ご参考》FEFSへ移行後の導入・運用効果（例）

従来 スループット性能の制限により、システムごとに独立したファイルシステムを構築

設計部構造システムグループ



PCクラスタ

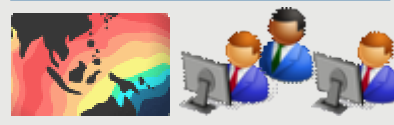


数十MB/s~500MB/s

従来
File
System



設計部流体システムグループ



PCクラスタ



数十MB/s~500MB/s

従来
File
System



設計部材料システムグループ



PCクラスタ



数十MB/s~500MB/s

従来
File
System

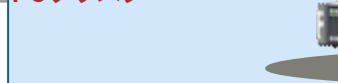


FEFS スループット性能向上により、ファイルシステム統合による一括管理が可能に

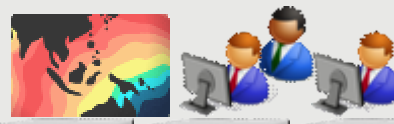
設計部構造システムグループ



PCクラスタ



設計部流体システムグループ



設計部材料システムグループ



1GB/s~1TB/s

FEFS

メタデータ

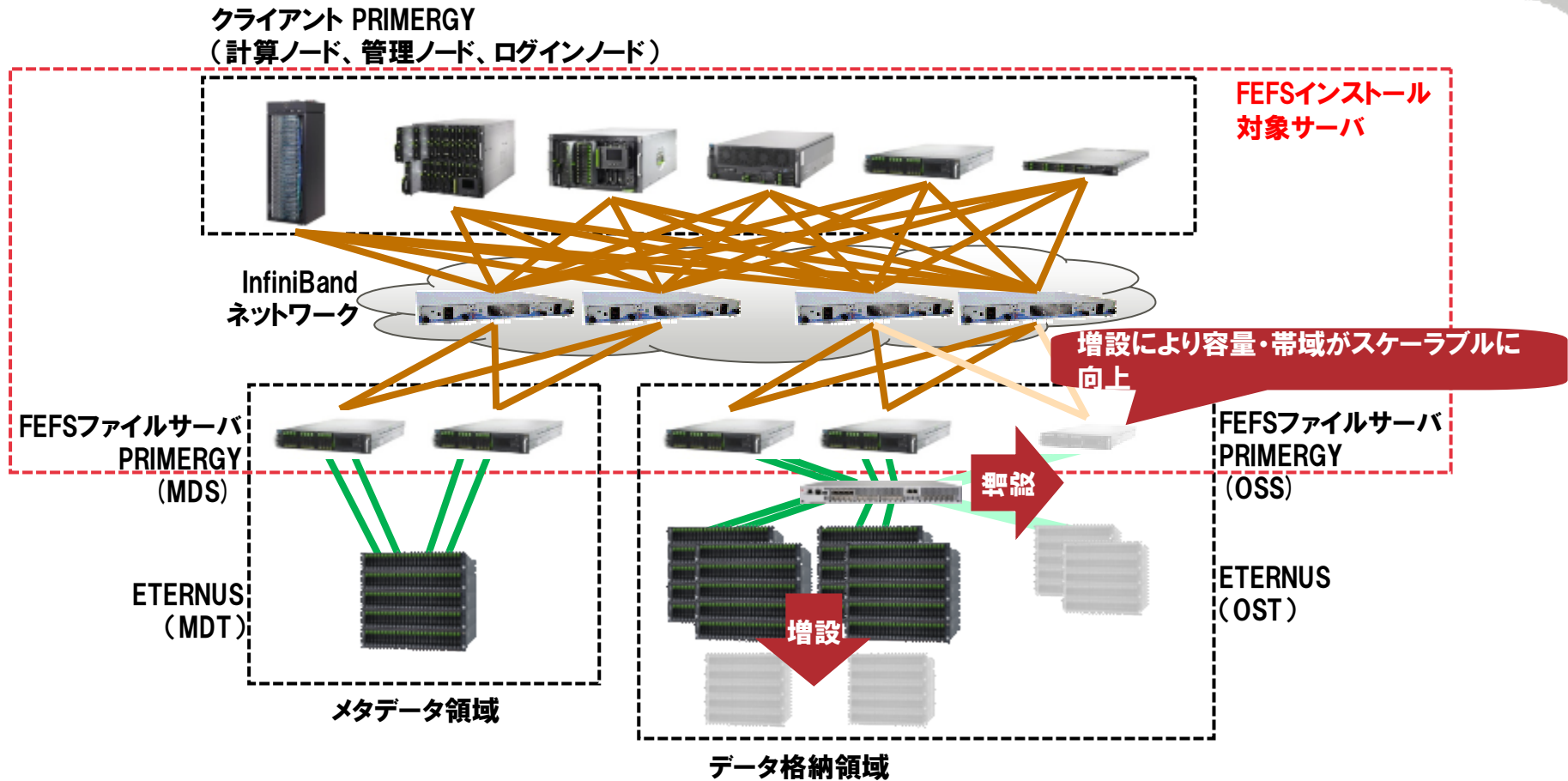
I/Oデータ



- 複数システムからのI/O処理にも耐えうるスループット性能
- フェアシェア機能により、特定ユーザ/ノードの占有を回避

「FEFS」の製品情報

製品・価格体系



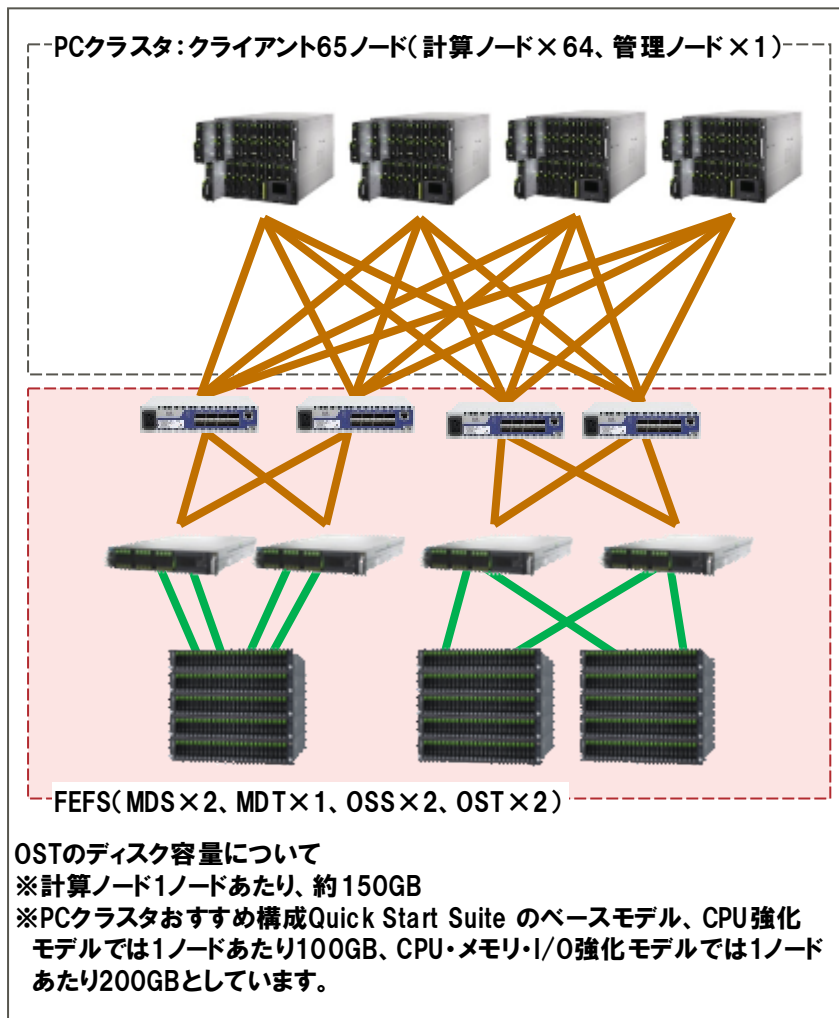
商品名	適用OS	価格
FEFS V10 基本 (4ノードライセンス付)	Linux(EM64T) ※Red Hat Enterprise Linux 5.6(for Intel64)	¥ 3,000,000
FEFS V10 追加 1ノードライセンス	Linux(EM64T) ※Red Hat Enterprise Linux 5.6(for Intel64)	¥ 100,000

仕様一覧

Item	FEFS	一般Lustre
最大ファイルシステムサイズ	8,000,000TB(8EB)	64,000TB(64PB)
最大ファイルサイズ	8,000,000TB(8EB)	320TB
最大ファイル数	9,220,000兆ファイル	4,000百万ファイル
最大ボリューム数	20,000ボリューム	8,150ボリューム
最大ボリュームサイズ	1,000TB(1PB)	16TB
最大クライアント数	1,000,000クライアント	131,072クライアント
最大ストライプ数	20,000ストライプ	160ストライプ
最大ファイル数(クォータ)	9,220,000兆files	4,000百万files
最大ファイルサイズ(クォータ)	8,000,000TB(8EB)	320TB
最大ACL数	8,191	32
ディレクトリ単位のクォータ指定	あり。	--
QOS機能	あり。 ユーザー間, FEFSCクライアントIPアドレスによるQOSが可能	--
ファイルシステムのブロックサイズ	4KB ~ 512KB	4KB
カーネルパッチが必要か？	不要。 但しサポートするカーネルを使用する場合	必要

システム参考価格(システム構成例・構成一覧)

■ システム構成例



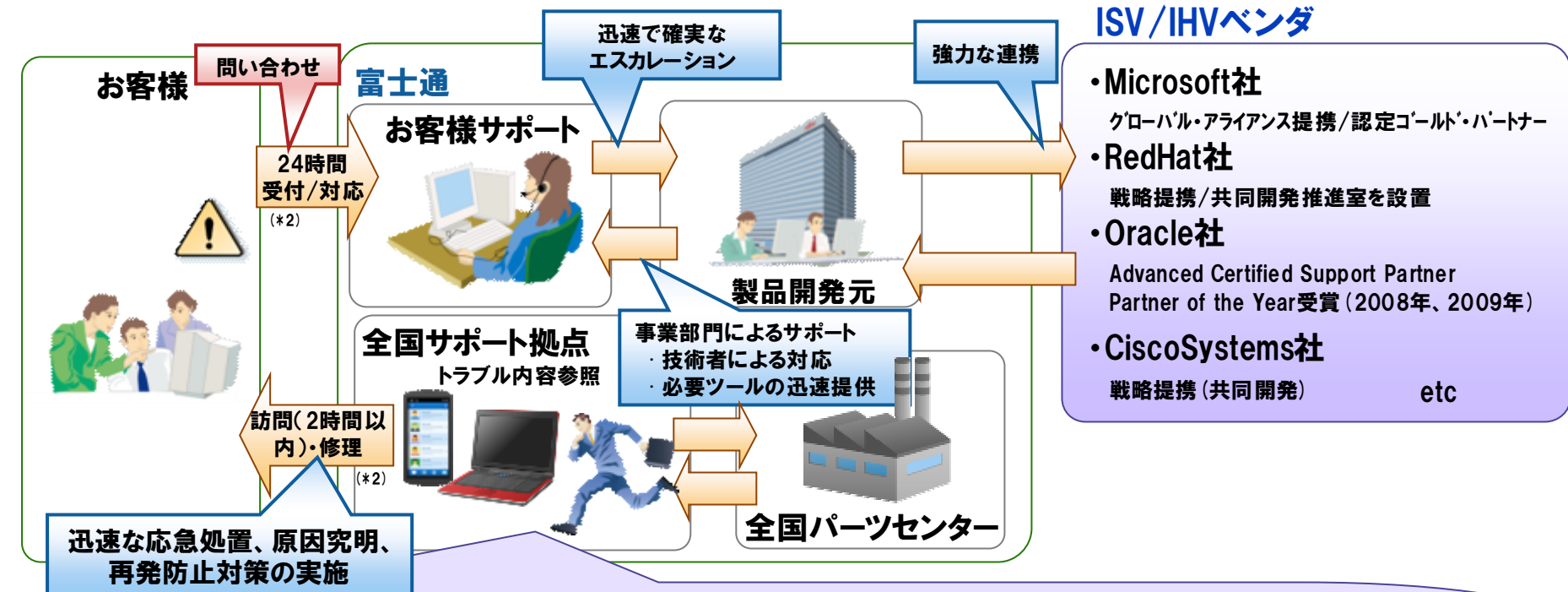
※InfiniBandは、インターコネクトとFEFSパスを兼用します。
 ※ラック、無停電電源装置は含まれておりません。
 ※現調費、搬入費、保守費なども含んでおりません。

■ 構成一覧

商品名	個数	価格
PRIMERGY BX900 S1 ブレードシャーシ(IB搭載)	4	¥ 11,719,200
管理ノード PRIMERGY BX922 S2 サーバブレード(IB搭載)	1	¥ 1,291,000
計算ノード PRIMERGY BX922 S2 サーバブレード(IB搭載)	64	¥ 82,624,000
IB電気ケーブル(16)、ツイストペアケーブル(12)	28	¥ 673,600
OS/ソフトウェア	一式	¥ 3,320,000
PCクラスタ小計価格		¥ 99,627,800
商品名	個数	価格
Mellanox社 8ポート InfiniBandスイッチ IS5022	4	¥ 1,387,320
SR-X316T1 ブレインストールモデルV01	1	¥ 88,000
FEFS V10 基本(4ノードライセンス付)	1	¥ 3,000,000
FEFS V10 追加 1ノードライセンス	65	¥ 6,500,000
※本構成例でのFEFSの年間サポート費用は、¥ 1,520,000です。合計価格には含まれておりません。		
PRIMERGY RX200 S6 (MDS)	2	¥ 2,015,000
ETERNUS DX80 S2 (MDT)	1	¥ 3,804,000
PRIMERGY RX300 S6 (OSS)	2	¥ 2,119,200
ETERNUS DX80 S2 (OST) 容量:9.6TB	2	¥ 7,328,000
FEFS小計価格		¥ 26,241,520
合計価格		¥ 125,869,320

SupportDeskサービス

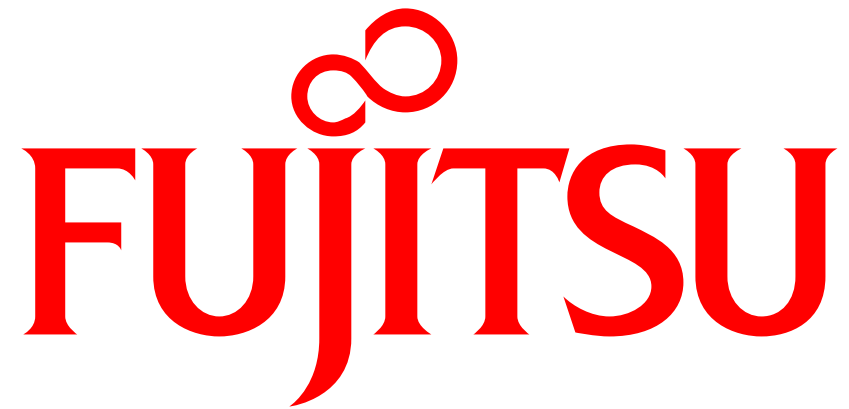
- ハードウェア障害からソフトウェアトラブル、Q&Aまでを富士通サポートセンター（OSC^(*1)）の専門技術者がワンストップで対応します。



国内最大級のサポート体制

・日本全国約850ヵ所のサービス拠点、24時間365日対応のOSC^(*1)（川崎/大阪）、24時間対応の保守パーツ・ロジスティック網など、約8000名のサービスエンジニアが日々お客様の安定稼働を支援（2008年3月現在）

(*1) OSC: One-stop Solution Center (SupportDesk Standardご契約のお客様専用の総合サポートセンター)
 (*2) サービス時間帯やサービス内容はご契約により異なります。



shaping tomorrow with you