



pNFSの紹介とPanasasの取り組み

スケールラブルシステムズ株式会社



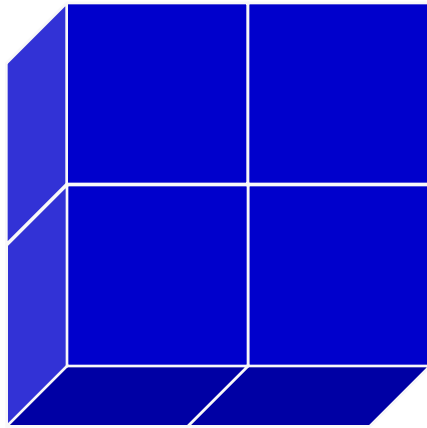
この資料について

- この資料は、米国 Panasas社の製品紹介の目的のために、米国 Panasas 社から提供された資料をスケーラブルシステムズ株式会社が日本語化したものです。
- 掲載してありますPanasas社の製品に関しては、Panasas社のホームページ <http://www.panasas.com> をご覧ください。



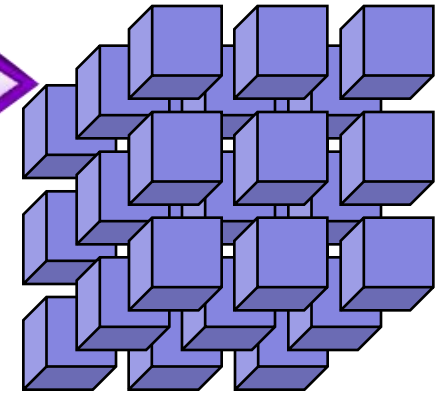
コンピュータシステムの変遷

メインフレーム
スーパーコンピュータ



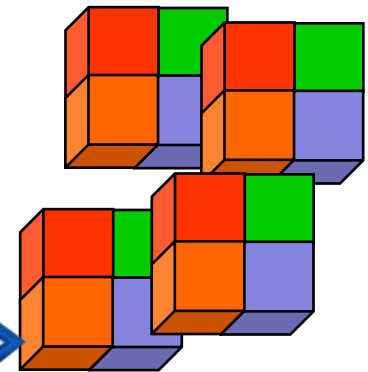
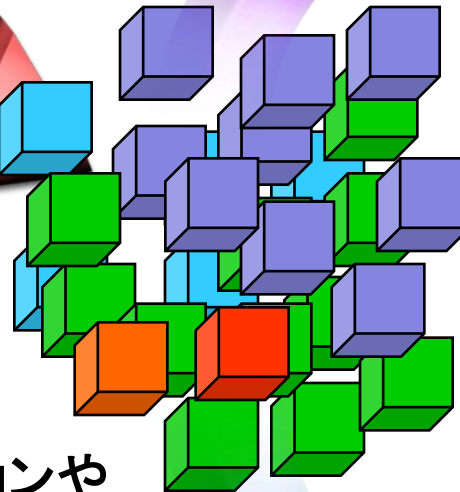
(リソースの集中と管理)

■ クラスタによる仮想コンピュータ



分散したリソースの管理

■ ワークステーションや
サーバによる分散処理

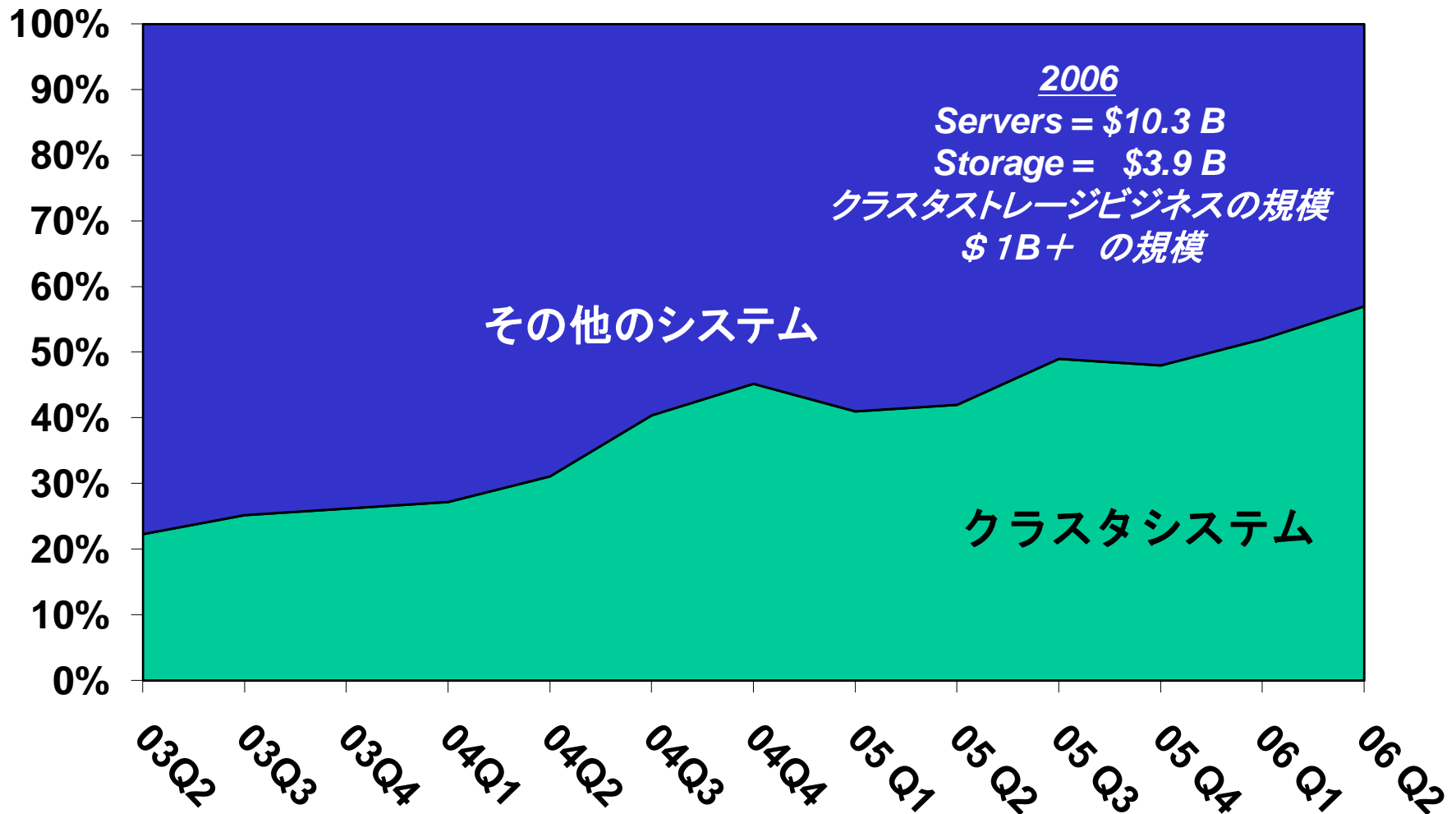


■ 仮想化によるサーバ・コンソリデーション



HPCクラスタがもたらすビジネス

既にクラスタはHPCシステムの主流となっています





パラレルコンピューティングには、パラレルストレージは必須

- x86/Linuxシステムの導入
 - コストを削減し、標準構成部品でのシステム構築が可能
 - 小規模から大規模なシステムまで構築が可能（ブロック構成）
 - システム規模の拡大
- 計算環境での並列処理の一般化
 - 計算機システムでのクラスタの一般化
 - プロセッサのマルチコア化の急速な浸透
 - アプリケーションや利用環境の整備による並列処理の容易な利用が可能



➡ ストレージシステムもこのような並列処理の進歩に対応し、その経済性の向上を図る必要がある



TCO : Total Cost of Ownership

開発の遅れによる機会損失
最新技術の導入機会損失

機会損失コスト

運用管理
トラブル対応
トレーニング
設備費用

調達コスト 20%

オペレーションコスト 80%

Source: Gartner Group 2005



顧客が直面する問題

- セットアップは非常に苦痛
 - クラスタの設定に時間がかかり、実際にシステムの利用が可能になるまでの時間が無駄になる
- システムのアップデートが煩雑
 - 既存のITインフラとHPCシステムの互換性の問題
 - Windows環境とLinux環境（セキュリティ、ユーザ管理）
- ジョブ管理
- アプリケーションの運用
 - 並列アプリケーションの効率的な運用
 - プリ・ポストとの連携
 - ストレージシステム



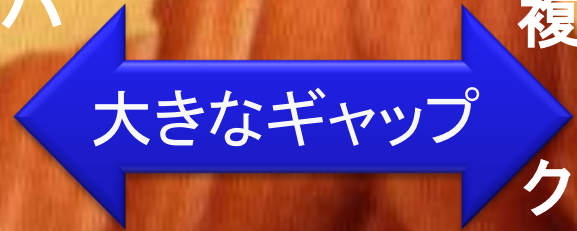
ハイパフォーマンスシステム

SMP (Shared Memory Systems)

ワークステーションやサーバ
PA-RISC, POWER5,
Itaniumなどのプロセッサ
によるSMPサーバ

クラスタシステム

システムの構築には、
高いITスキルが要求される
運用管理コストが高い
複雑なオペレーション環境
複数のOS
クラスタファイルシステム
ソフトウェア、インストール
やアップグレードなど



ワークステーション
サーバ

クラスタ



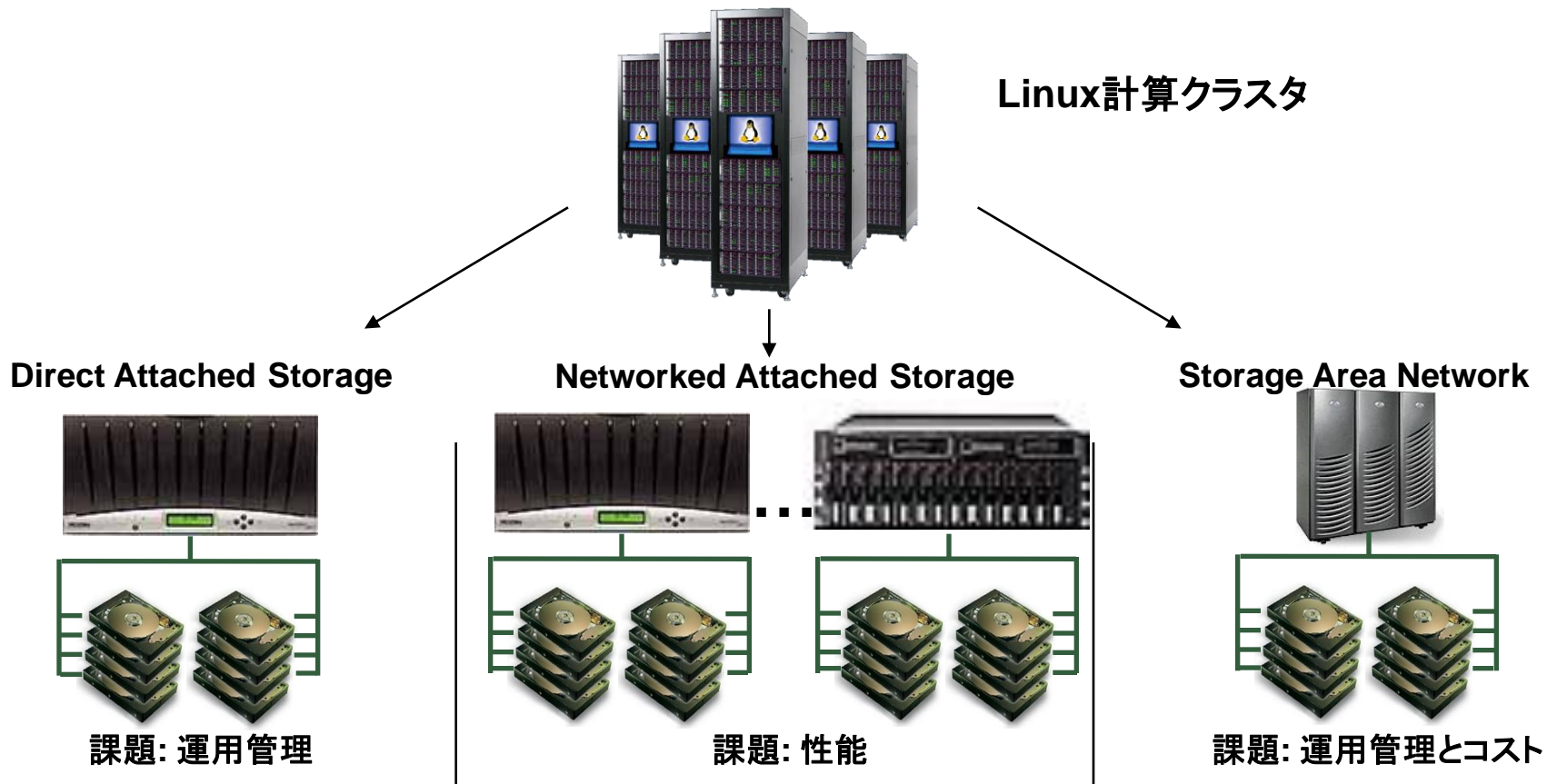


顧客の要求

- エンドユーザ：
 - シンプルにセットアップが可能
 - アプリケーションが豊富で、簡単に導入出来る
 - 簡単なジョブの投入とモニターリング
 - 高い実行性能とスケーラビリティ
- 運用管理者：
 - IT環境をシンプルに
 - シンプルにクラスタを導入し、シンプルな運用で、高い生産性を実現できる環境の構築
- アプリケーション開発者：
 - 生産性の高い開発環境
 - シンプルなシステムデザイン
 - 標準のライブラリやAPI環境



ストレージアーキテクチャ



共通の問題

性能、運用管理の容易さ、スケーラビリティ、コスト



パラレルI/Oの標準化問題

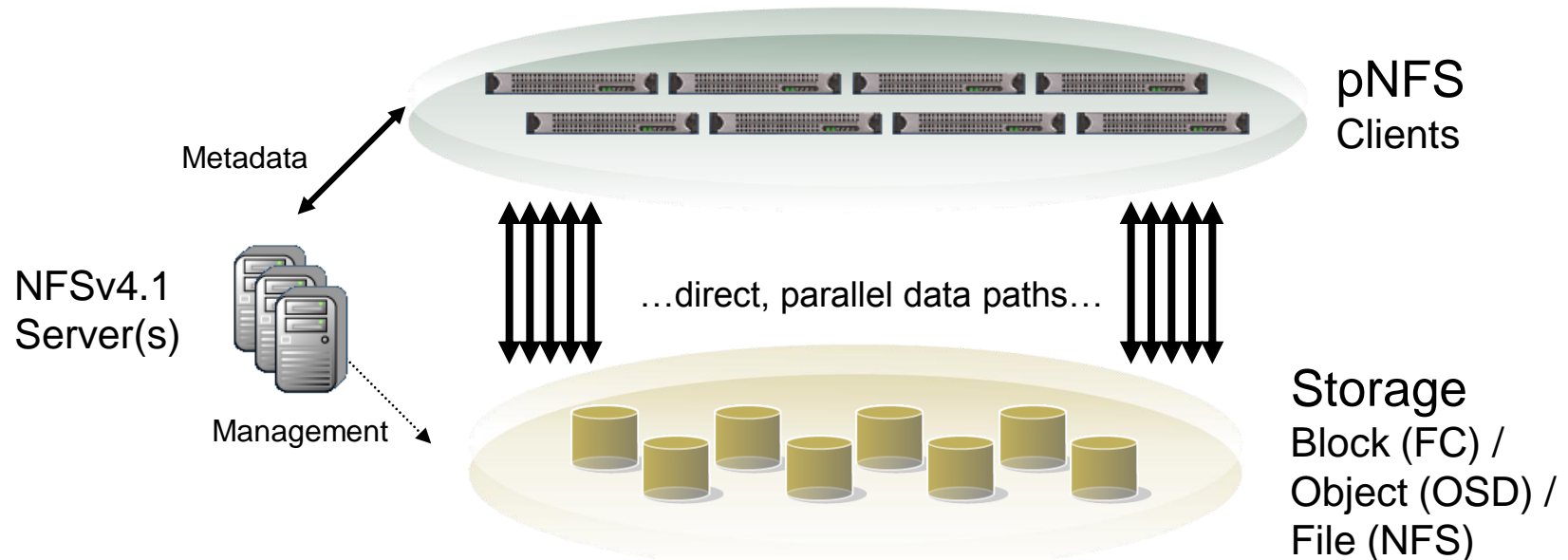
- パラレルストレージを提供するベンダー間で互換性の欠如
 - Panasas PanFS
 - IBM GPFS
 - EMC MPFSi (High Road)
 - IBRIX Fusion
- オープンソースの活動も独自に進展
 - Red Hat GFS
 - PVFS
 - Lustre
 - オープンソースの製品間でも、相互に互換性がない

パラレルI/Oの標準化は、コスト削減やユーザの選択肢の
広がりなどの点で多くの利点がある



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





pNFSへの各社の取り組み



- Panasas (Objects, based on Panasas Storage Cluster OSDs)
- Network Appliance (Files over NFSv4)
- IBM (Files, based on GPFS)
- EMC (Blocks, based on HighRoad MPFSi)
- Sun (Files over NFSv4)
- U of Michigan/CITI (Files over PVFS2, Files over NFSv4)



pNFSによるユーザの利点



- 高いアプリケーション性能
 - 標準化された高性能パラレルストレージの利用が可能



- 選択肢の広がりとリスクの低減
 - パラレルI/Oが可能なストレージ製品を複数のベンダーから提供
 - より多くのクライアントとクライアントOSからのパラレルアクセスが可能（Linux、Windows、AIX、Solarisなど）



- 運用管理が容易
 - 共通のグローバルネームスペース
 - 負荷分散が容易



Panasas: “The pNFS Company”

- pNFSは、PanasasのCTOの Garth Gibsonによって提案
- pNFS はPanasas DirectFLOW クライアントアーキテクチャを活用
- DirectFLOW クライアントは実質的には pNFS v0.9
 - pNFS 標準に準拠
- pNFSの利点の多くは、現在のPanasasプラットフォームが実証
 - 高いバンド幅
 - スケーラビリティ
 - 容易な管理・運用
 - アプリケーションでの有用性の証明



DirectFLOW クライアント ソフトウェアのオープンソース化

- PanasasはDirectFLOW クライアントソフトウェアのコアをオープンソース化
 - どのようにパラレルI/Oでのボトルネックを解消するかのリファレンス
 - キーとなるPanasasのコンポーネント：
 - Storage Access Mgr, OSD client, Object iSCSI Panasas libraries (common, rpc, sec) www.pnfs.com a community resource site.
- Panasas pNFS 開発センター
 - Panasasの開発経験を最大限に活用
 - pNFS Object layout driver, iSCSIなど
 - なぜ、Panasasはオープンソース化？
 - pNFSの普及の加速（Linux カーネルなどへの早急な対応）
 - pNFSベースの製品をより早く市場投入

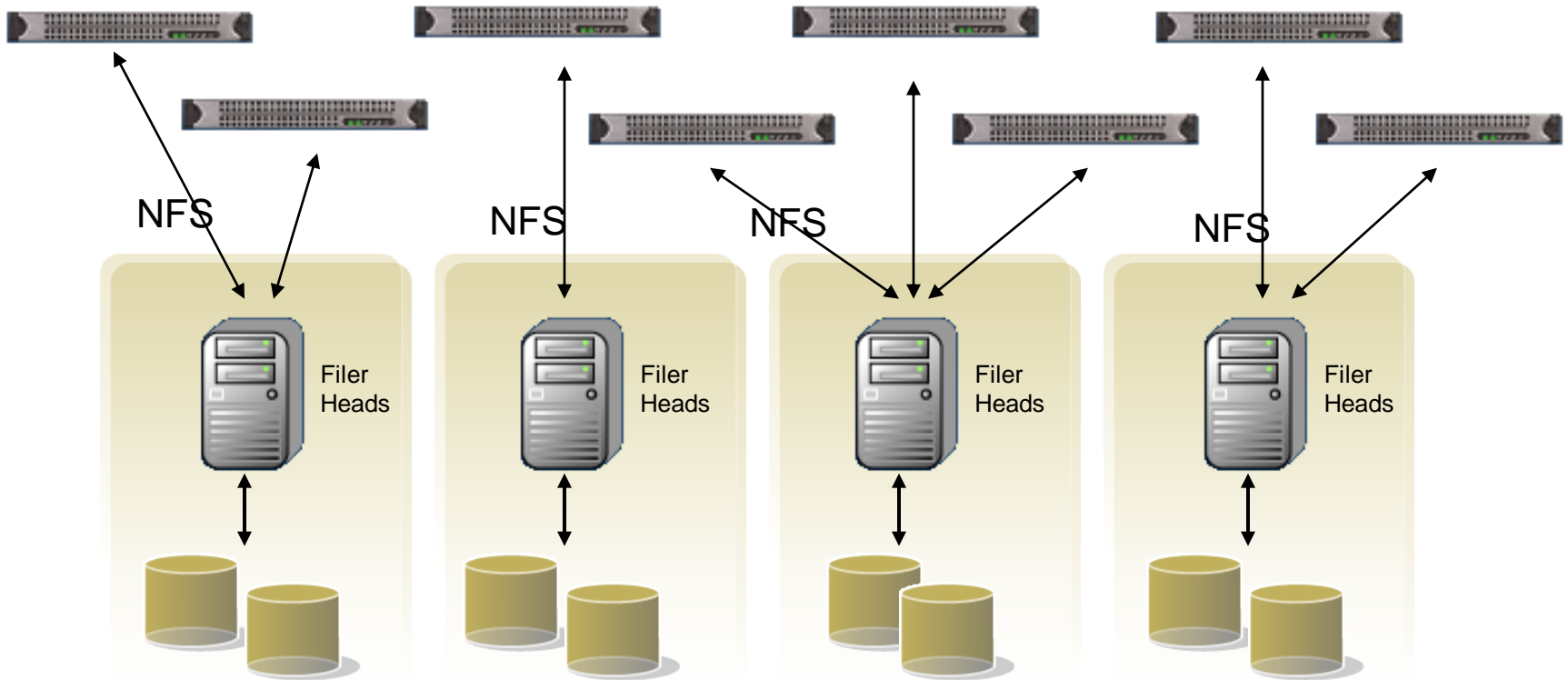


パラレルストレージでの実績

- システムアーキテクチャは完全なパラレル対応
 - pNFSへのアップグレードは容易
 - pNFSによるより広範囲なマーケットへの対応
- 商用製品の開発経験
 - 4年以上の商品出荷実績
 - 100以上のユーザサイトで稼動
- オブジェクトベース pNFS実装
 - より優れた性能：ストリームとランダムI/O
 - 容易な運用管理と導入（短時間でのインストール、自動プロビジョニング、負荷分散、RAIDマネージメント）
 - 可用性：FailOver、予防診断、パラレル再構築
 - ペタスケールの拡張性



1980年代後半：NAS



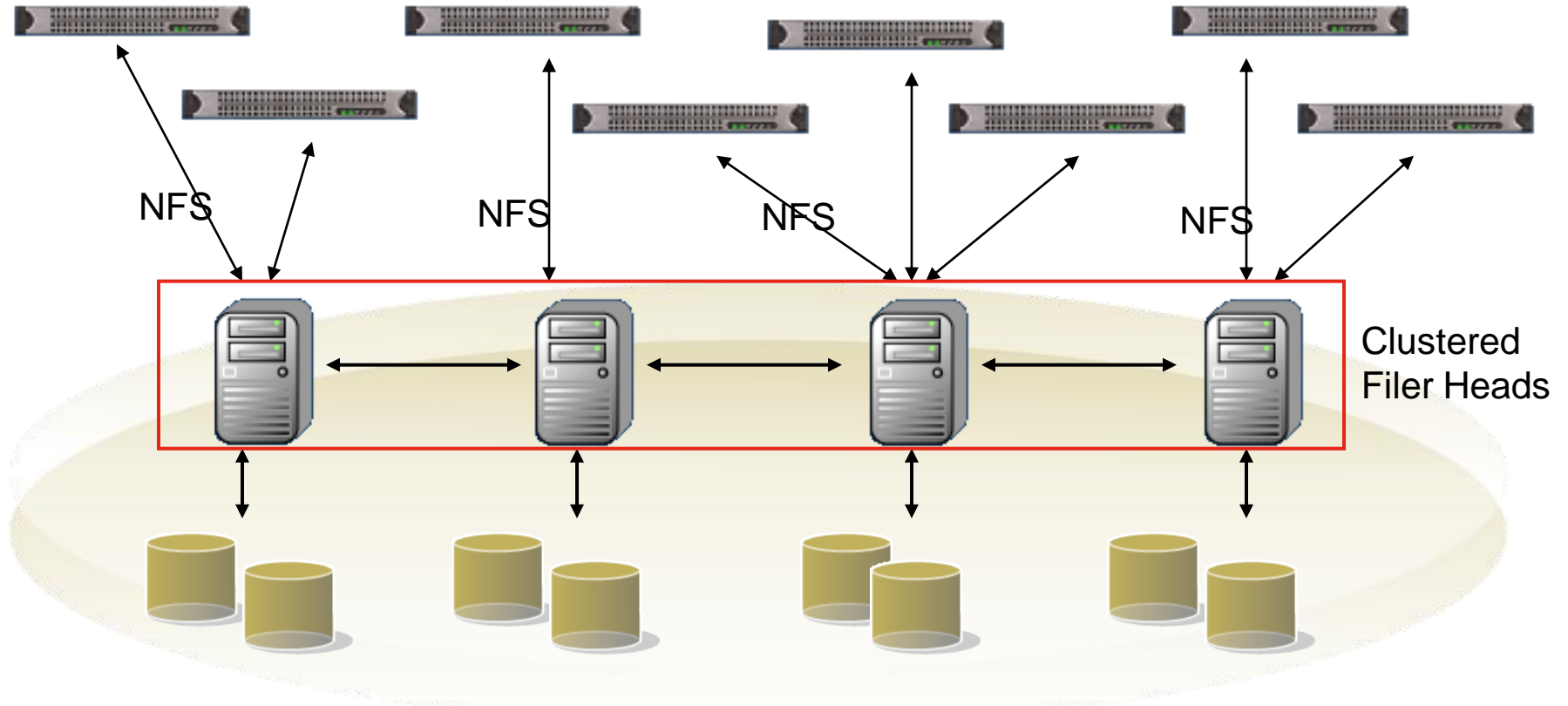
“ストレージは独自に点在”

Filerヘッドが、I/O 性能のボトルネックとなる

複数のストレージの運用管理は容易ではない



2000年以降：クラスタNAS

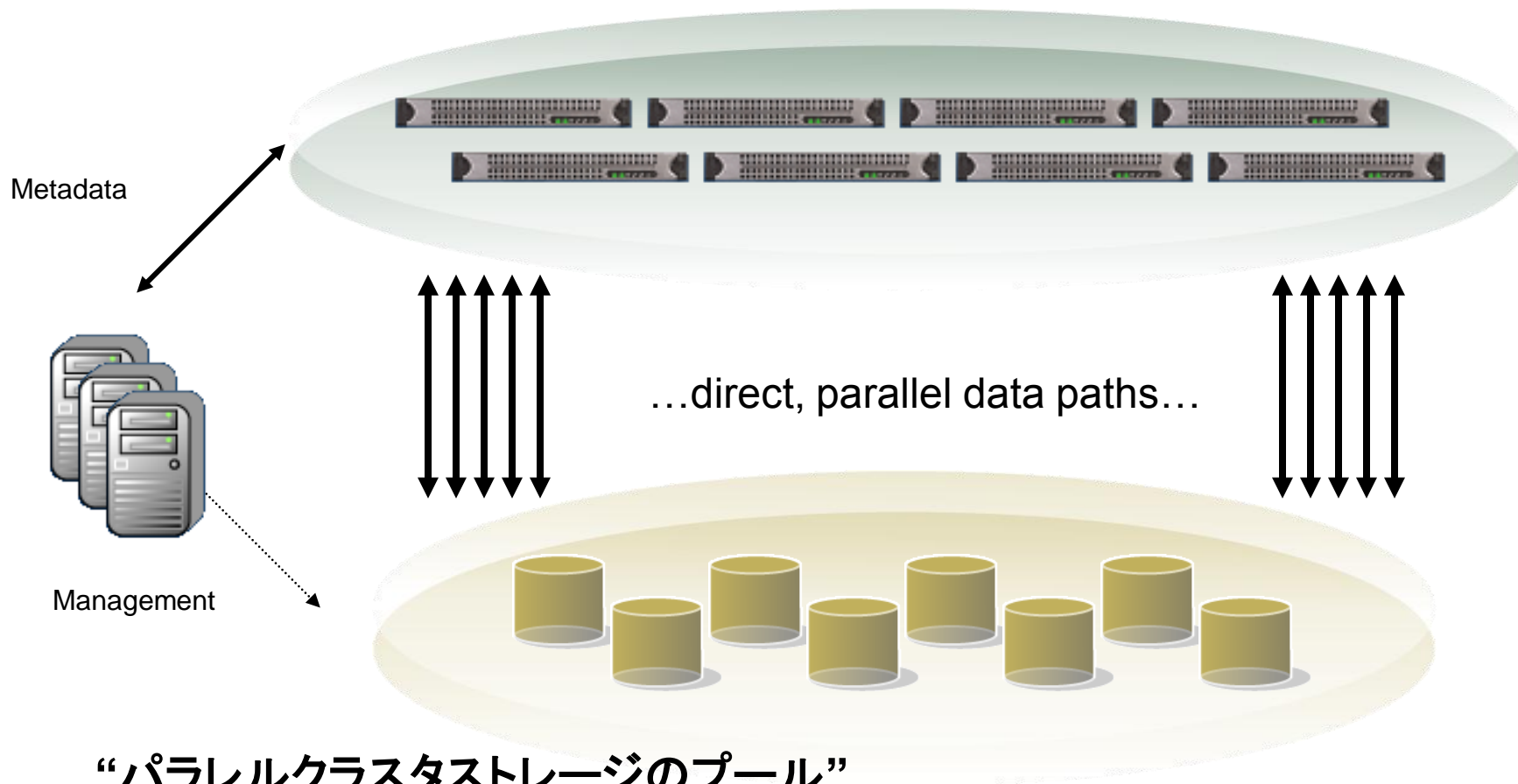


“点在するストレージ間での架け橋の構築”

“In-band” Filer ヘッドの同期処理がI/O性能のボトルネックとなる
負荷分散の管理が性能とシステムの生産性を左右



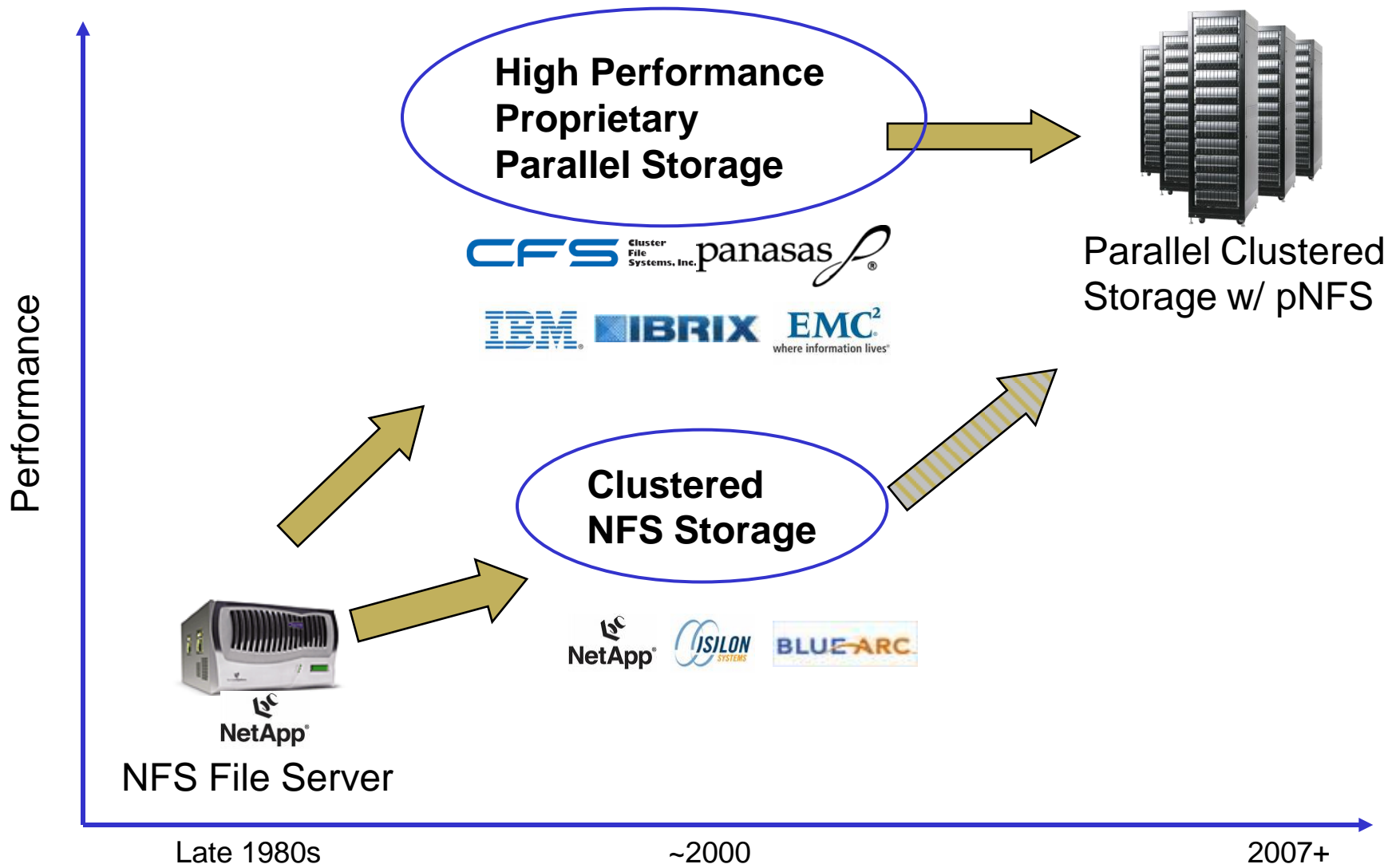
パラレルクラスタストレージ



Filerをデータパスから排除することで、I/O性能のボトルネックと運用管理の問題を解決

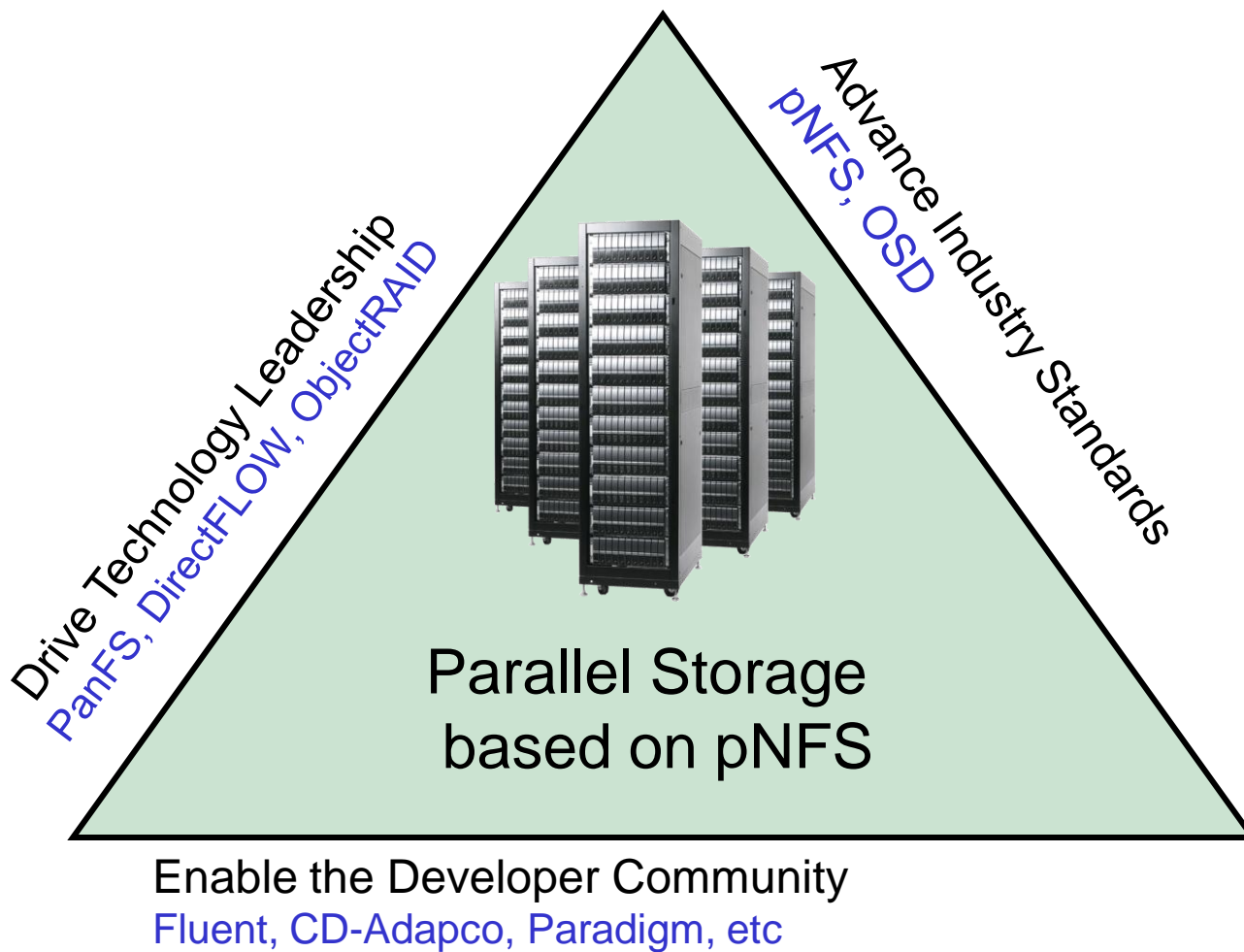


HPC ストレージマーケット





Panasas パラレルストレージ





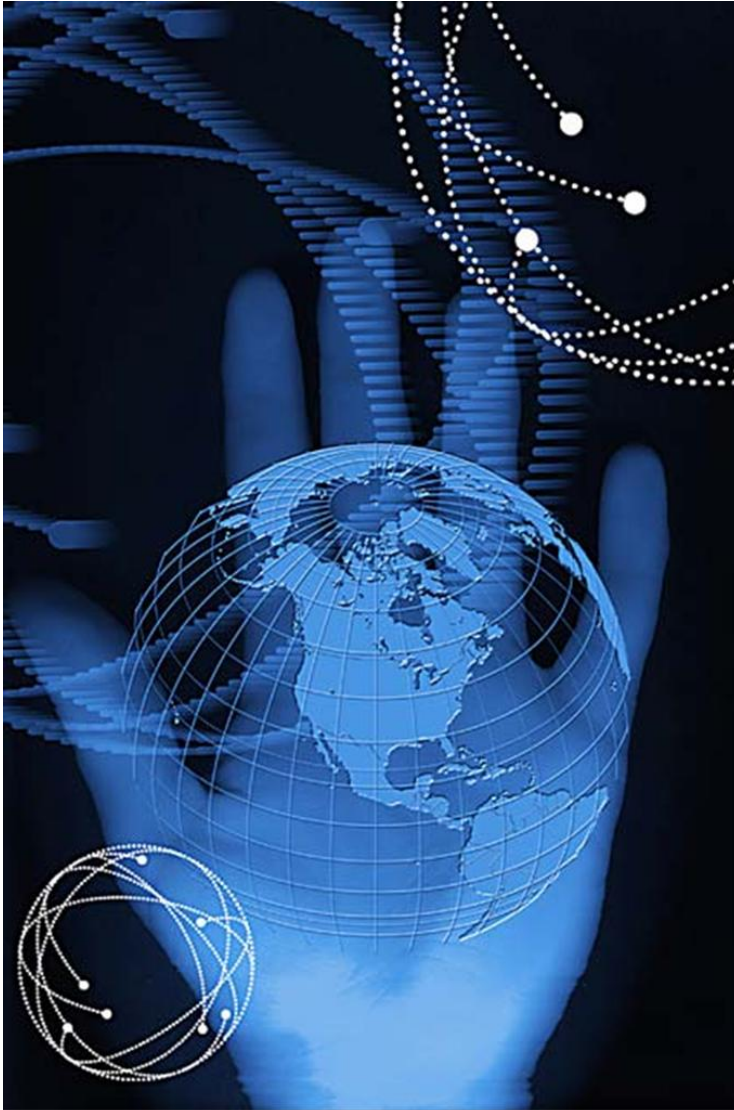
この資料について

この資料の無断での引用、転載を禁じます。
社名、製品名などは、一般に各社の商標または登録商標です。なお、本文中では、特に®、TMマークは明記していません。

In general, the name of the company and the product name, etc. are the trademarks or, registered trademarks of each company.

Copyright Scalable Systems Co., Ltd. , 2007. Unauthorized use is strictly forbidden.

10/22/2007





さらに詳しい情報や最新情報は.....

ホームページにて公開しています。ホームページには、お問い合わせ窓口も開設しておりますので、ご利用ください。

コンサルテーション

<http://www.sstc.co.jp>

製品技術

<http://www.hp2c.biz>

10/22/2007

