

グリッドに対する日立製作所の取組み

ー サーバ、ストレージ、ミドルウェア ー

May 31, 2007

HITACHI
Inspire the Next

株式会社日立製作所
ソフトウェア事業部 先端情報システム研究開発本部
担当本部長 吉野 松樹

本日のご説明内容

1. 日立のグリッドへの取組み
2. グリッドを実現する日立の技術のご紹介
 - 2.1 日立サーバ仮想化機構 *Virtage*
 - 2.2. Services Oriented Storage Solutions
 - 2.3 運用管理ソフトウェア JP1 の自律制御機能
3. グリッド技術の適用シーン

1. 日立のグリッドへの取組み

グリッド

狭義のグリッド

コンピューティンググリッド
(高速計算サービス)

データグリッド
(大規模データ処理サービス)

ビジネスグリッド
(高信頼Webサービス)

PCグリッド
(遊休PCの活用)

ネットワーク上の計算機を
仮想的に統合した並列分散処理

仮想計算機センタサービス

大規模データの解析

グリッドデータファーム

企業間連携、電子政府など

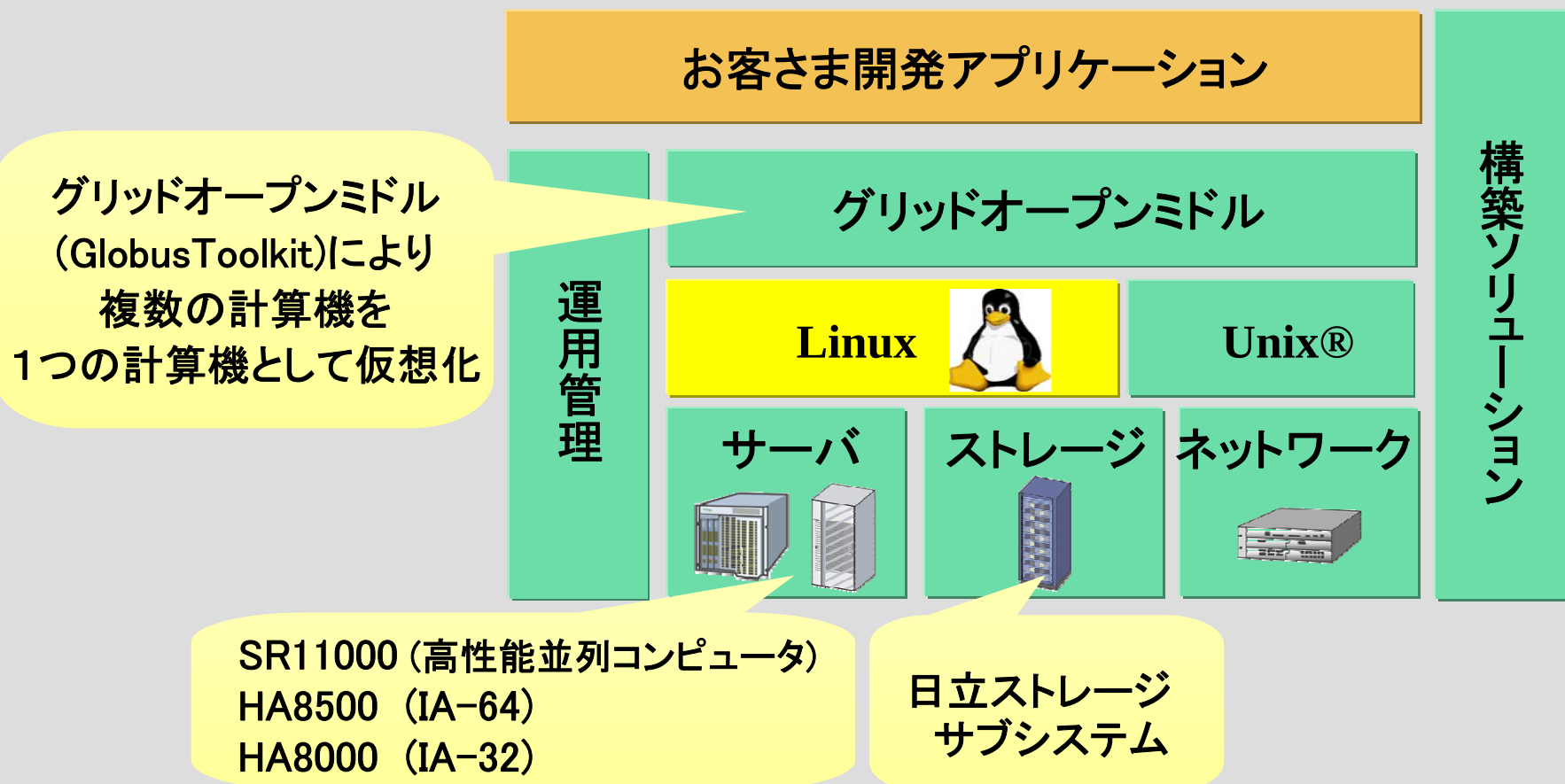
企業内PCによるイントラグリッド

SETI@HOMEなど

1.2

日立のサイエンス分野のグリッドへの取り組み

仮想化された環境で、資源の能力を最大限に活かすために、
お客様の開発アプリケーションの並列化作業を代行

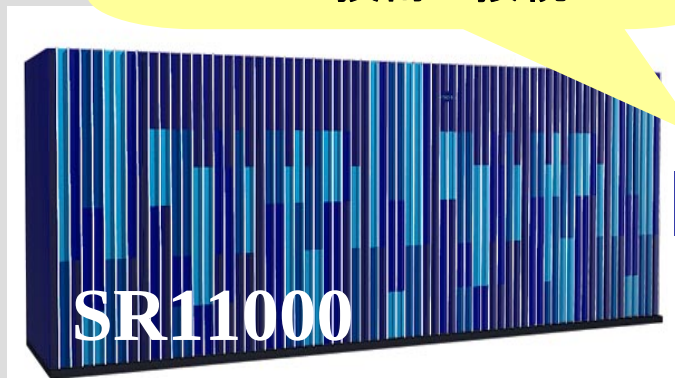


1.3

サイエンスグリッド事例

世界最大級のグリッドコンピューティングシステム (岡崎国立共同研究機構)

スーパーコンピュータと
Linux PCクラスタを
Grid技術で接続



50

5.4 TFLOPS

Gridミドルウェアには、
UNICORE,
Globus Toolkitを使用



HA8000

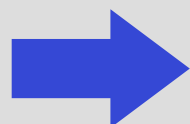
409

5.0 TFLOPS



ノード数

システム規模

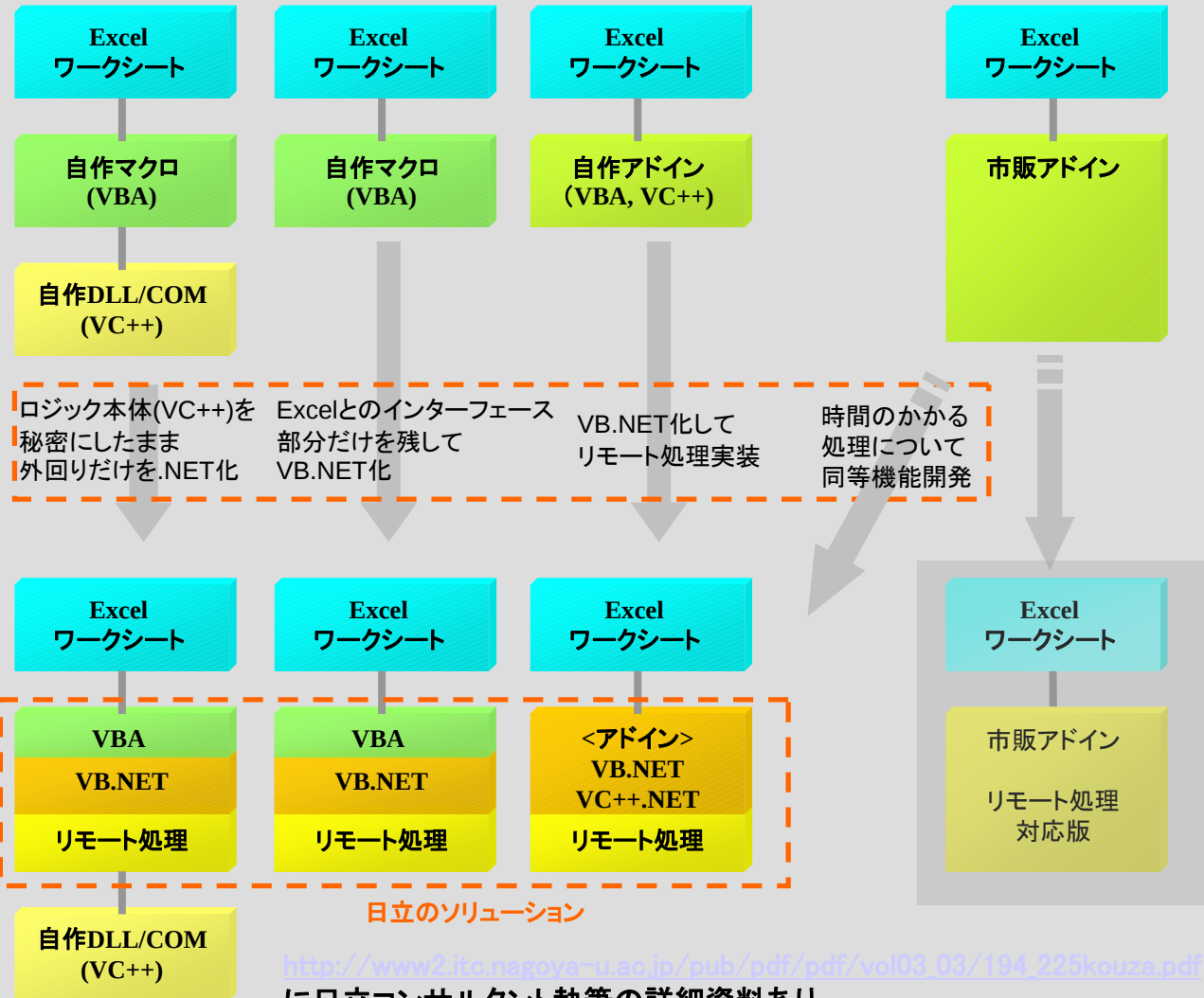
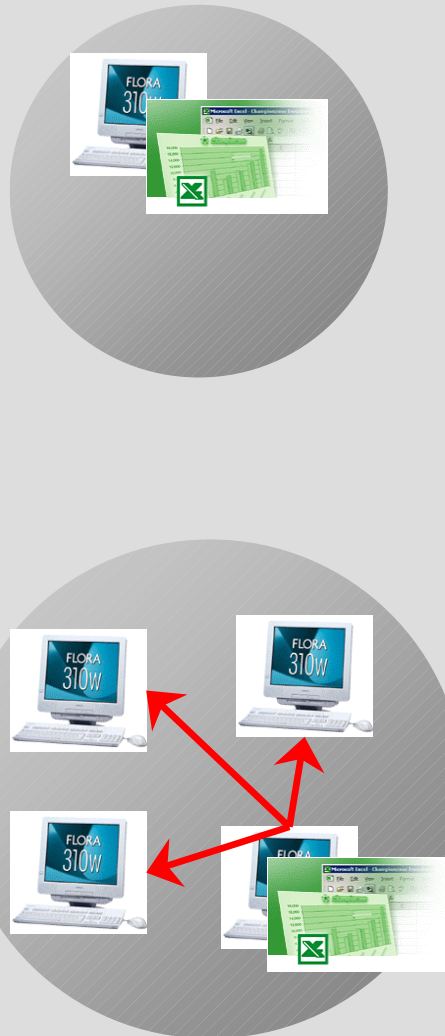


双方の計算機の特性を活かし
複数の大規模科学計算プログラムを連結して実行可能に

1.4

PCグリッド事例 — Microsoft® Excel® の Windows® 並列処理化ソリューション

Excel上のアプリケーションのソースコードがあれば.NET Remotingを利用した並列化が可能。
アドイン形式の市販アプリケーションについても、時間のかかる特定部分の並列化などの対応が可能。



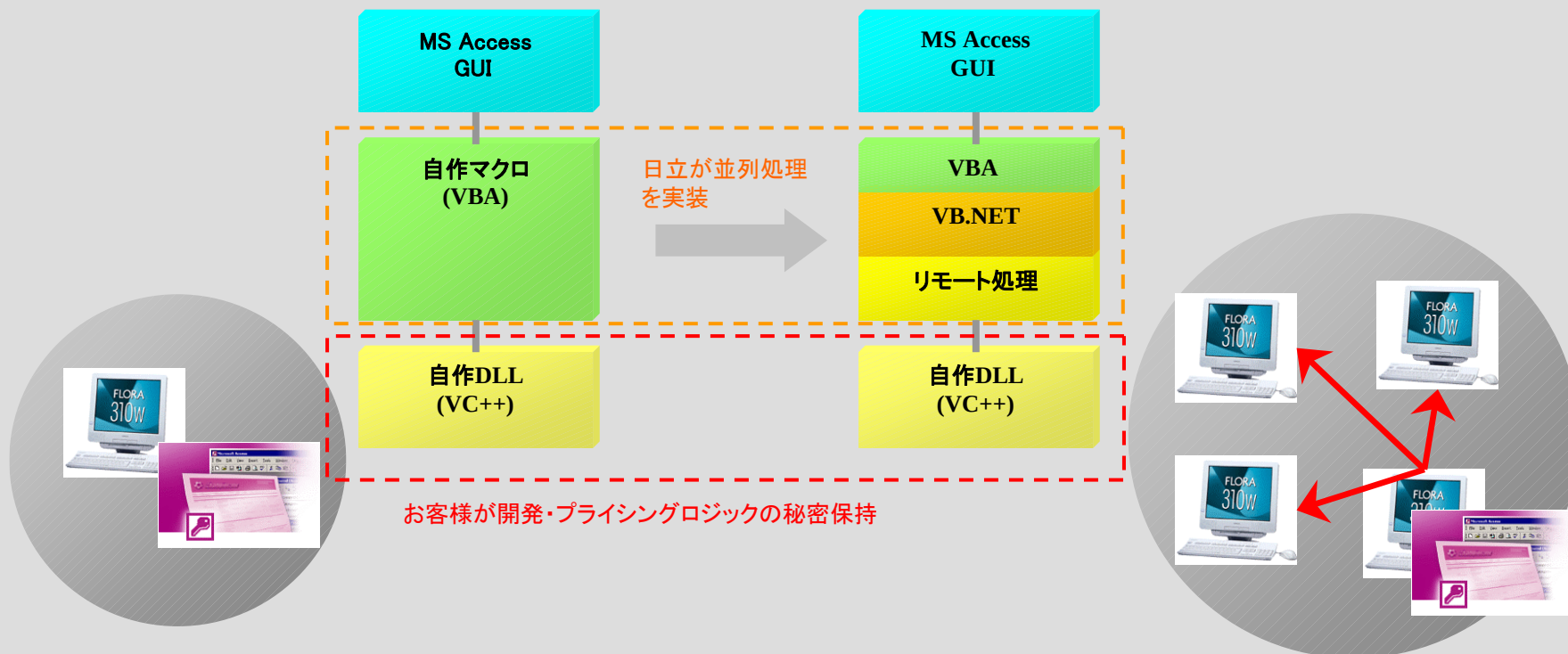
http://www2.itc.nagoya-u.ac.jp/pub/pdf/pdf/vol03_03/194_225kouza.pdf
に日立コンサルタント執筆の詳細資料あり

1.5

PCグリッド事例: 金融商品プライシングシステム

GUIをMicrosoft Access/VBA、プライシングロジックをVisual C++で開発したシステムを

- ・VBAを一部VB.NET化してリモート処理を実装
- ・Visual C++のDLLはまったく変更しない形で並列処理を実現
プライシングロジックはブラックボックスのままで、並列処理を実現可能
- ・同様の事例として、天候デリバティブのプライシングを行う「HiDeri」がある



夜間の遊休PCを有効活用し、かつ省電力を実現する研究用の高速計算用グリッドコンピューティングシステムを実現

■システム構成

◆サーバ...BladeSymphony BS1000 サーバモジュール4台

OS: **Red Hat Enterprise Linux ES** ←Grid MPを動作させる為の前提OS

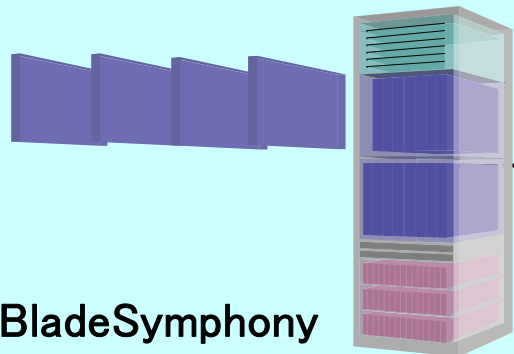
◆グリッドコンピューティングに用いる機器...デスクトップPC 500 台

OS: **CentOS (Linux)** / Windowsをインストール

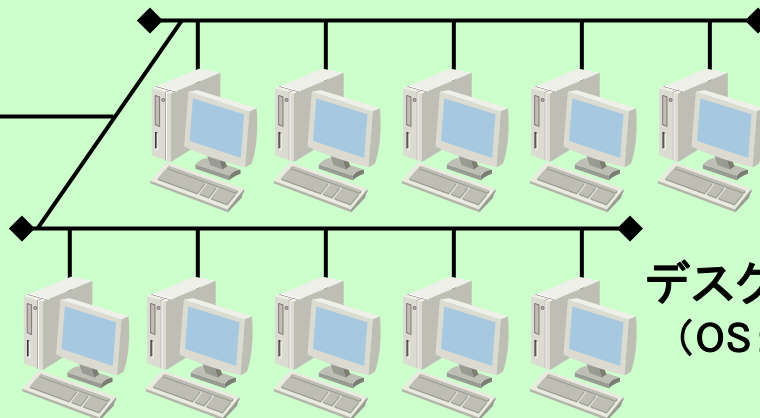
CPU: Pentium 4 3.6GHz、メモリ: 1GB、HDD: 160GB

◆ミドルウェア...Grid MP (サーバ/エージェント) ←お客様からの指定

連携ツール ←日立が開発。複数台の教室PCをリモートで電源ON/OFFする配信サーバと連携



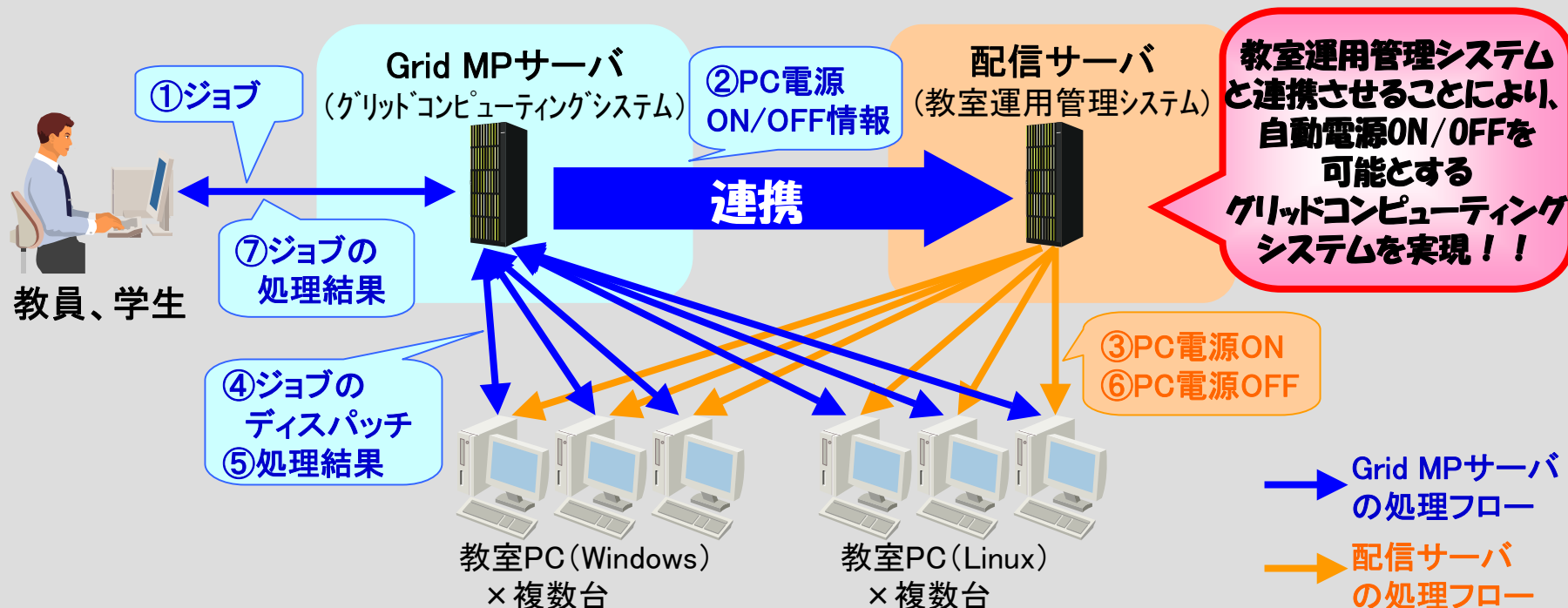
BladeSymphony
サーバモジュール × 4台
(OS: **Red Hat Enterprise Linux ES**)



デスクトップPC × 500台
(OS: **CentOS (Linux)**
/ Windows)

1.7

大学の夜間遊休PC活用事例－運用イメージ



- ①日中に、Grid MPサーバへジョブを投入する。
- ②Grid MPサーバの連携ツールにて、ジョブの実行に必要なPCの台数、起動OS (Windows/Linux)、起動時間、等を算出し、PC電源ON/OFF情報として配信サーバへ転送する。
- ③Grid MPサーバより取得したPC電源ON/OFF情報に従い、配信サーバより教室PCを夜間にリモート起動する。
- ④Grid MPサーバより、起動したPCにジョブをディスパッチする。
- ⑤各PCにてディスパッチされたジョブを実行し、その処理結果をGrid MPサーバへ返す。
- ⑥PC電源ON/OFF情報に従い、配信サーバより起動中のPCのリモートシャットダウンを行う。
- ⑦Grid MPサーバにて、各PCから返された処理結果を収集・統合し、翌朝にジョブの処理結果を得る。

目 的

ネットワークに接続された多数のIT 資源(サーバ、ストレージ等)を組織の境界を越えて稼働させ、IT資源の自在な活用等を可能にするグリッドコンピューティング技術を、ビジネスの分野で実用化

ポイント

- (1) ヘテロジニアス(=異機種混在)な環境への対応
- (2) 柔軟性が高く、高信頼で安全性の高いシステムを低コストで実現

実施内容

- (1) 基盤ソフトウェアの開発
- (2) ユーザ企業の実運用環境での実証実験の実施
- (3) 成果の一部をオープンソフトウェアとして公開し、普及を推進
- (4) 国際的な標準化団体の活動にも積極的に参加し、仕様の提案等を通じて標準化活動に貢献

プロジェクト参加メンバー

富士通(株)、日本電気(株)、(株)日立製作所、(独)産業技術総合研究所

プロジェクト実施期間

2003年度から3年間。2006年3末で終了。

ビジネスグリッド推進コンソーシアム

富士通、日立、日本電気、IPAを発起人とし、プロジェクトで開発した技術をベースに、

- ◆相互接続性確保のための標準化活動の推進

- ◆幅広いユーザニーズの吸い上げ

- ◆実践的な利活用シーンの深耕

を目的として2006年7月発足

1.10 ビジネスグリッド技術をHarmonious Computingへ

日立はビジネスグリッドの技術を
Harmonious Computing関連製品に適用しています

BUSINESS GRID COMPUTING PROJECT


Harmonious Computing

JP1^{Version 8}

HiRDB^{Version 8}

Cosminexus
コスミネクス
Version 7

openTP1^{Version 6}
Distributed Transaction Processing Facility

1.11

サービスプラットフォームコンセプト Harmonious Computing

ビジネス環境の変化に自ら即応するサービスプラットフォーム

ユビキタス
情報社会



業種サービス・ソリューション

プラットフォーム
ソリューション

サービス
プラットフォーム

Harmonious
Computing

アプリケーションフレームワーク

オープンミドルウェア

サーバ・ストレージ・ネットワーク

ユビキタスアクセス
・フレームワーク

ユビキタスアクセス



1.12

サービスプラットフォームコンセプト

Harmonious Computingの狙い

- ・TCO削減によるROIの向上
- ・サービスプラットフォームの存在を意識せずにコア・ビジネスに集中

運用が容易
であること

サーバ、ストレージなどが
散在し、運用管理
コストと手間がかかる



複雑なシステムを
一元的に運用管理

急な性能要求に
対応できること

リソース増強にコストと
時間がかかり、業務拡大
に柔軟に対応できない



必要なときに必要な
リソースを追加

高可用システム
であること

障害の原因究明に
時間がかかり、短時間で
業務を再開できない

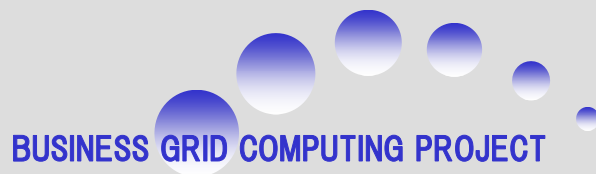


システム全体としての
高可用性を確保

1.13

ビジネスグリッドとHarmonious Computing

Harmonious Computingは、ビジネスグリッドプロジェクトの目指す技術的ソリューションを具現化します



リソースの共用／全体最適化を可能とする
統合的なITリソース管理を用いた
ITリソース有効活用

負荷変動や障害に対応する
自律的なリソース管理による
安定したシステム基盤の提供

システムの維持コストを下げ、新たな
価値創造のための投資継続を可能とする
ITリソース基盤の提供

Harmonious Computing

必要なときに必要なリソースを追加でき、
ビジネスの変化や成長に柔軟に対応

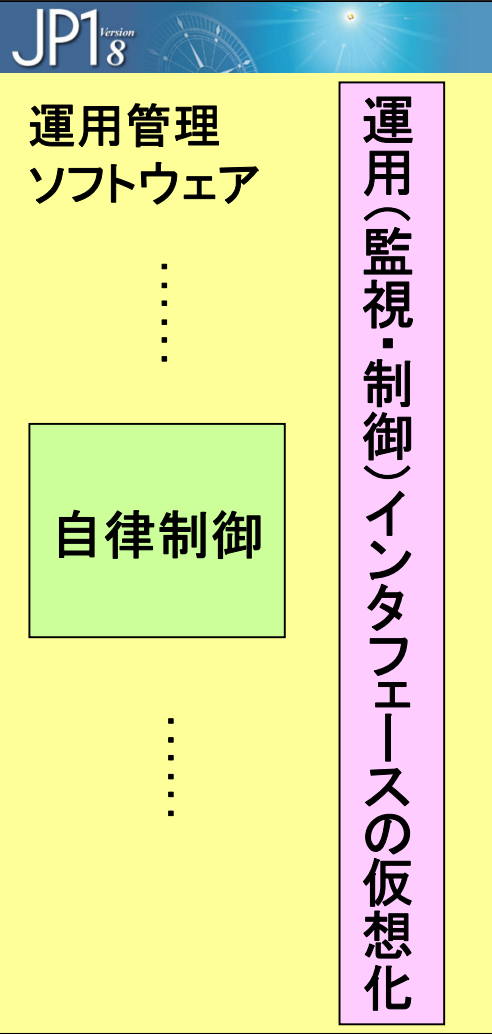
システム全体としての
高信頼性・高可用性を確保し、
止まってはならないビジネスを強固に支える

情報システムに必要な機能を集約し、
複雑なシステムを一元的に運用管理

2.グリッドを実現する日立の技術のご紹介

仮想化：運用環境と実行環境

グリッド＝仮想化によりアプリケーション/サービスと物理リソースの組み合わせを自由にし、仮想化された運用インタフェースで自律的に最適制御すること



2.2

ITプラットフォームの進化

Harmonious Computing

ビジネス環境の変化に自ら即応し、
その存在を意識させない、
柔軟かつ高信頼なIT基盤を提供

仮想化

統合化

分散化

分散化のデメリット
・異機種混在
・TCOの増大

技術の進歩と環境変化

- ・リソース利用効率の追求
- ・省電力化の要求
- ・運用性、信頼性の高度化

e-文書法

省電力

マルチコア

個人情報保護法

Web 2.0

J-SOX法

セキュリティ

10Gbps Ethernet

2.3

仮想化への取組み

1980年代

メインフレーム仮想化

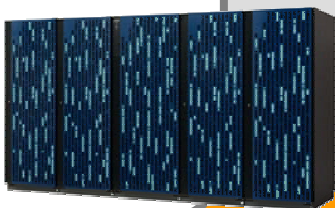
- 命令アシスト機構
- I/Oアシスト機構

2004/9

ストレージ仮想化

SANRISE USP

- 機種隠蔽
- 運用効率向上
- 資産の継承



2007/2

仮想化技術の継承

サーバ仮想化

統合サービスプラットフォーム
BladeSymphony

- 効率性
- 信頼性
- 運用性



運用の統合

- システム運用管理ソフト

ITプラットフォームの仮想化へ

- 即応性・柔軟性・信頼性の向上

2.4

ITプラットフォームの仮想化

ビジネス変化に即応しサービス提供するプラットフォーム

お客様価値

即応性

柔軟性

信頼性

サービスレベルの向上

JP1

業務システム

運用の統合
(運用の容易化)

サーバ管理

ストレージ管理

ネットワーク管理

リソースの仮想化
(利用効率向上)

サーバ
仮想化

ストレージ
仮想化

ネットワーク
仮想化

統合システム運用管理

BladeSymphony

日立ストレージ
サブシステム

GRシリーズ

仮想化の進化

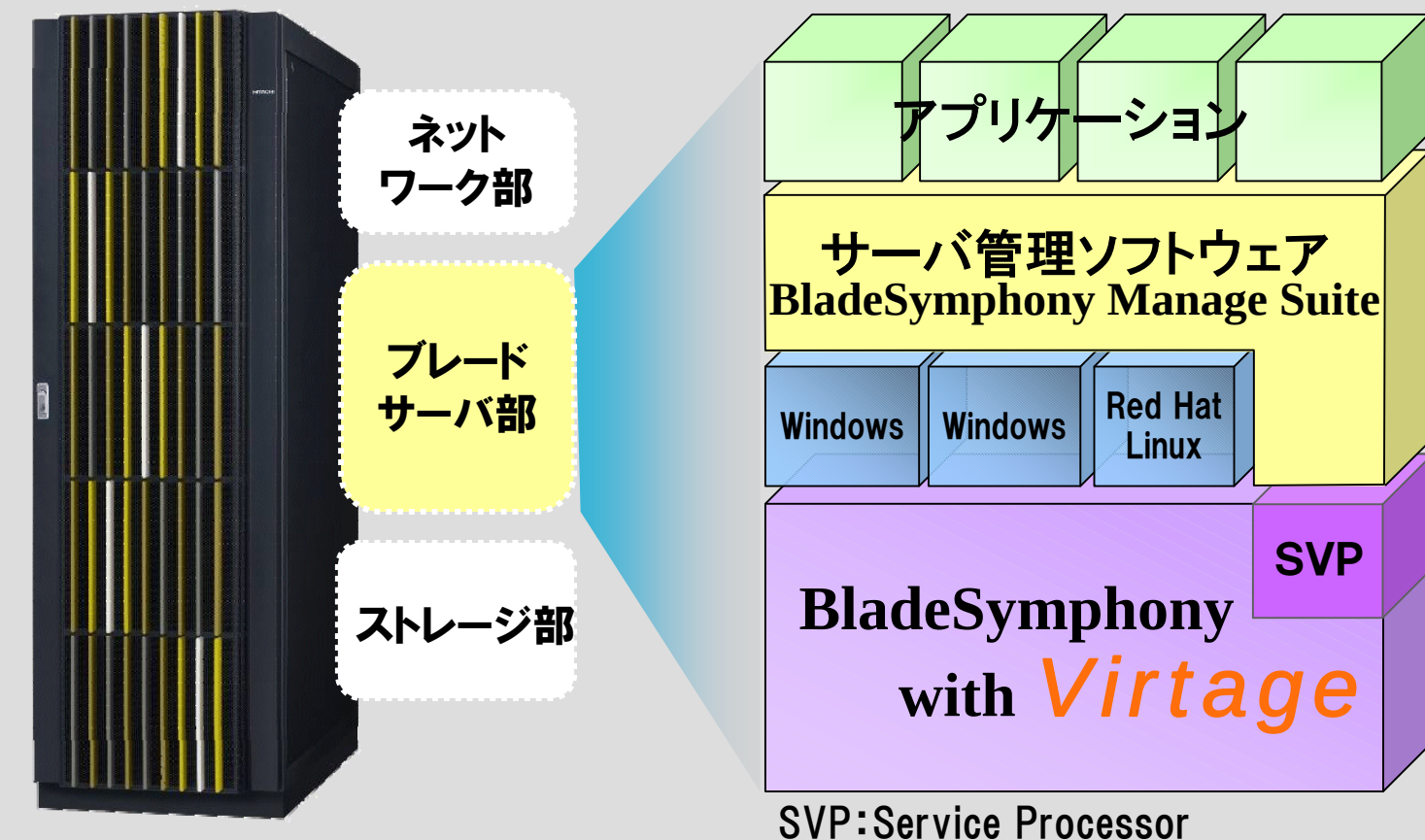
2.1 日立サーバ仮想化機構 *Virtage*

2.1.1

BladeSymphonyと日立サーバ仮想化機構 *Virtage*

ハードウェアアシストとSVP連携による
優れた効率性・運用性・信頼性提供

日立のサーバ仮想化アーキテクチャ



2.1.2

日立サーバ仮想化機構 *Virtage* の特長

特 長

効率性

- ・仮想化アシスト機構でI/O仮想化オーバヘッドを低減

運用性

- ・優れた物理サーバとの互換性
- ・SVP連携による一元管理

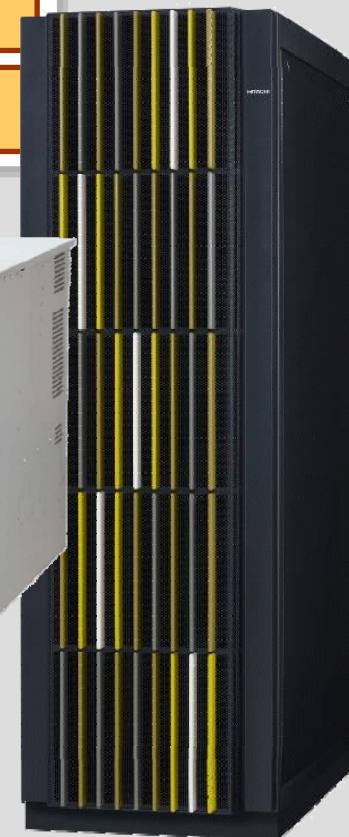
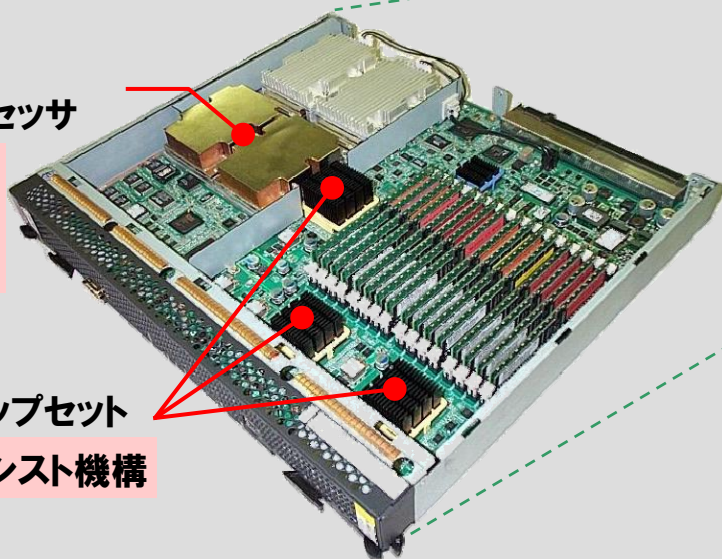
信頼性

- ・仮想化アシスト機構で論理サーバの独立性確保

デュアルコア
Itanium2 プロセッサ

Intel
Virtualization
Technology

自製チップセット
H/Wアシスト機構



BladeSymphony

2.1.3

Virtage が実現する価値

独自技術によるサーバ仮想化機構

Virtage (バタージュ)*

効率性

サーバコスト35%低減!!*²

高いI/O性能が多様な業務を統合可能に

➡ 効率の良いサーバ集約と省電力化を実現

運用性

TCO30%低減!!*³

OS(Linux/Windows)やドライバは改変不要

➡ システム構築時の検証コスト・運用コストを低減

信頼性

サーバと一体化したサポート

➡ 安定稼動によりビジネス機会損失を低減

*「Virtage」(バタージュ): Virtual Stageの造語。

仮想化(Virtualization)を新しいステージ(Stage)に導く製品であることを連想させる弊社独自仮想化機構のブランド名

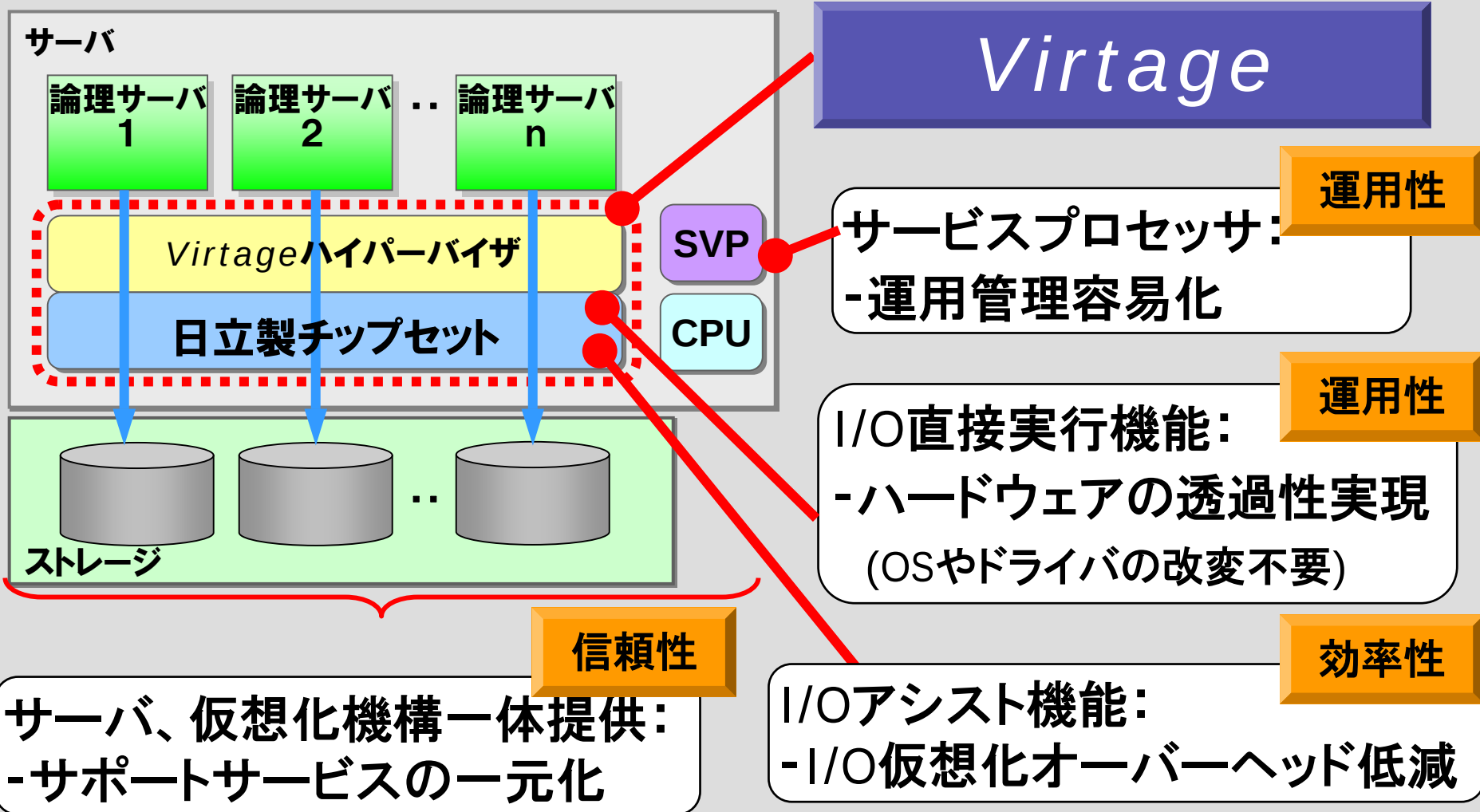
*²待機系サーバ集約モデルに基づく弊社評価

*³サーバ統合における弊社評価

Copyright © Hitachi, Ltd. 2007. All rights reserved.

2.1.4

ハードウェアによる高性能な仮想化環境



2.1.5

I/O仮想化のオーバーヘッド低減

「仮想化アシスト機構」による高効率利用

従来の課題

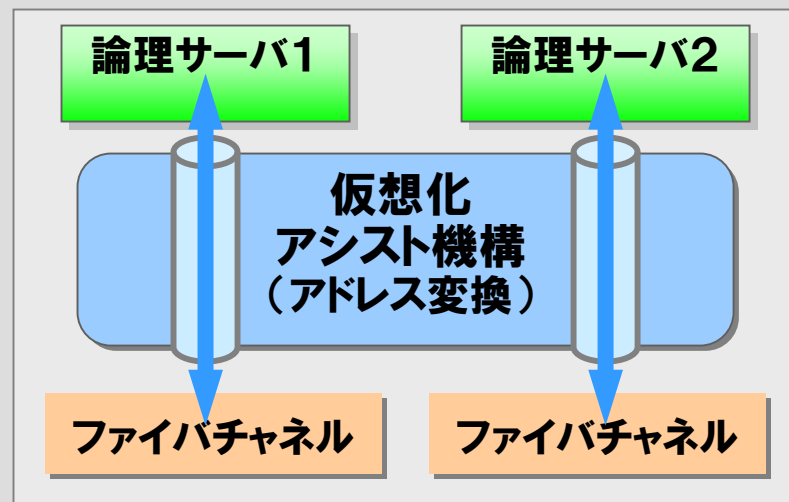
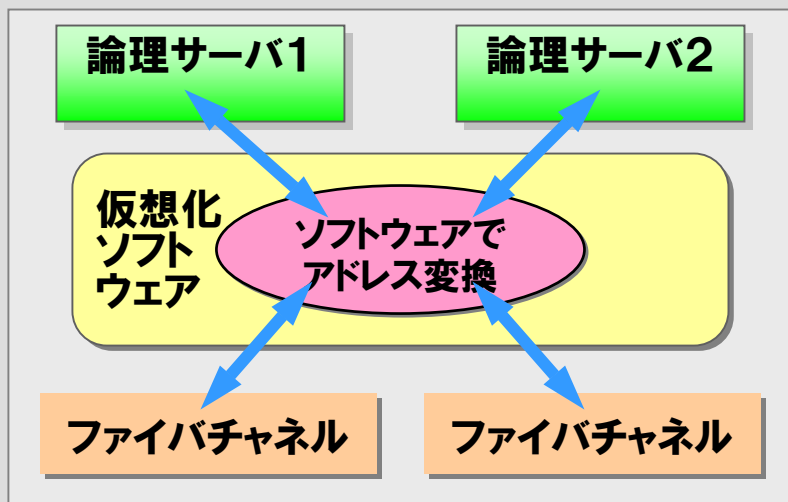
- I/Oアクセス時、**仮想化ソフトウェア**によりアドレス変換実施

➡ ソフトウェア処理のため
オーバーヘッド大

Virtageによる解決

- I/Oアクセス時、**仮想化アシスト機構**によりアドレス変換実施
(U.S. PATENT 7,080,291)

➡ I/Oトランザクション処理量が
従来比**最大40%向上**



2.1.6

優れた物理サーバとの互換性

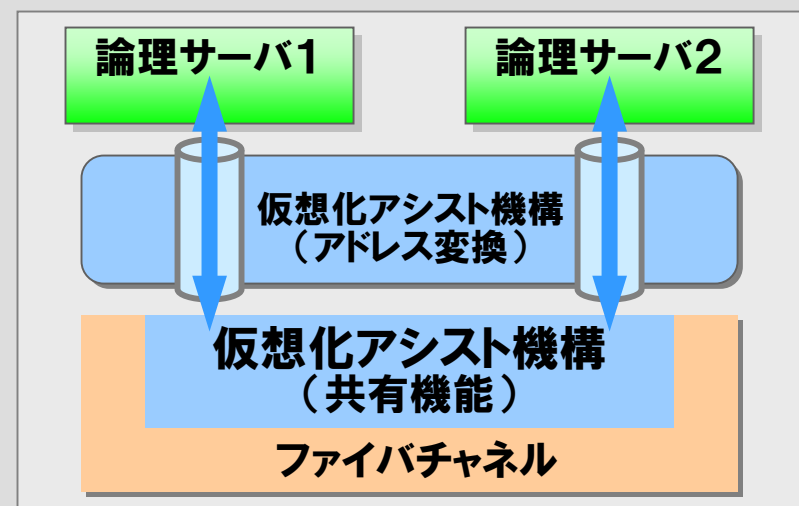
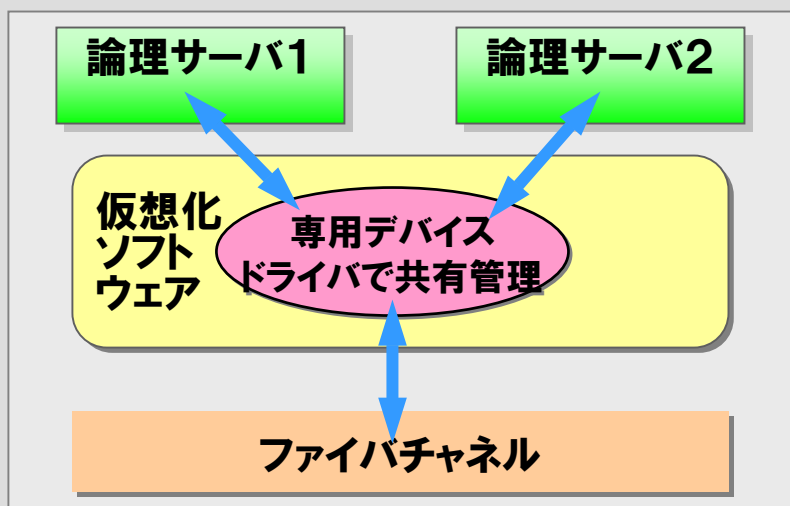
■ 物理サーバ環境(OS、デバイスドライバ)をそのまま利用可能

従来の課題

- ファイバチャネルの共有管理は**専用デバイスドライバ**で実施
- ➡ 使用I/Oデバイスに**制約あり**
- 物理サーバと**異なる**インストール作業

Virtageによる解決

- ファイバチャネルの共有管理は**仮想化アシスト機構**で実施
- ➡ 使用I/Oデバイスに**制約なし**
- 物理サーバと**同一の**インストール作業



2.1.7

論理サーバの独立性確保

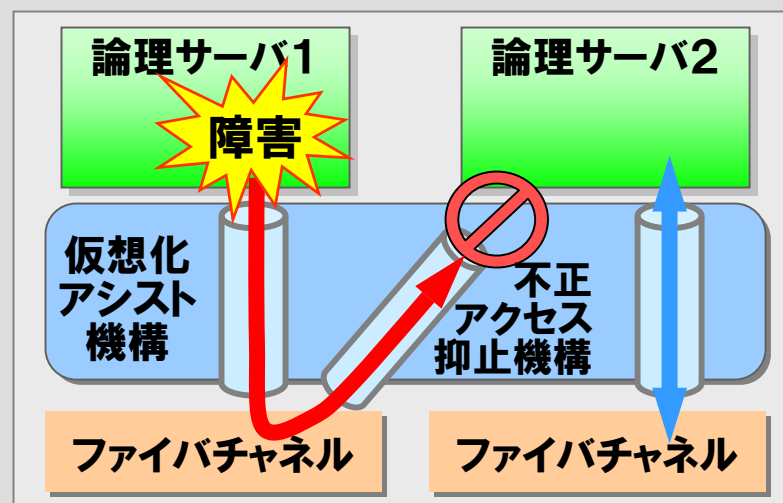
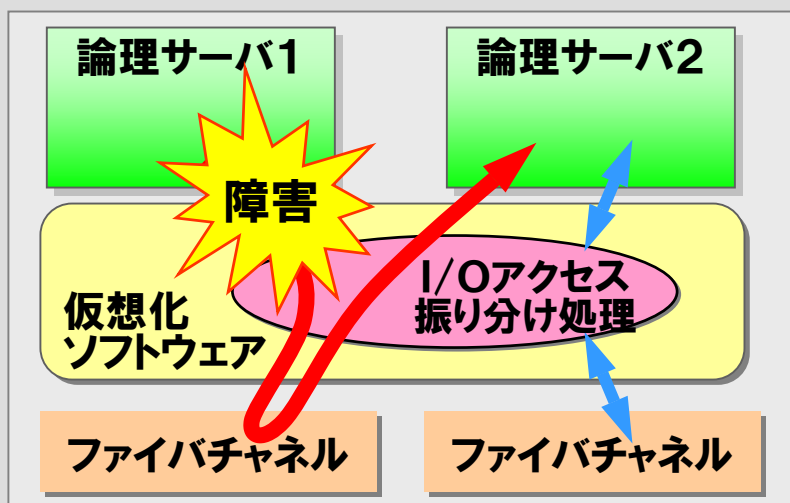
■ 論理サーバにまたがる障害波及の防止

従来の課題

- 仮想化ソフトウェアによるI/Oアクセス振り分け処理
- ➡ ウィルス、障害による不正I/Oアクセスの波及可能性あり

Virtageによる解決

- ハードウェアによる不正アクセス抑止機構
- ➡ 不正I/Oアクセスの波及をハードウェアにより確実に抑止



➡ : 正常I/Oアクセス ➡ : 不正I/Oアクセス

2.1.8

SVP連携による容易な運用管理

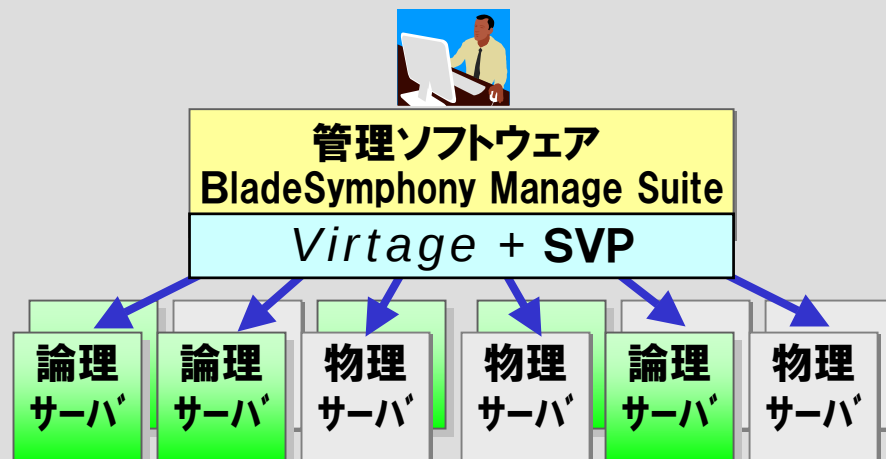
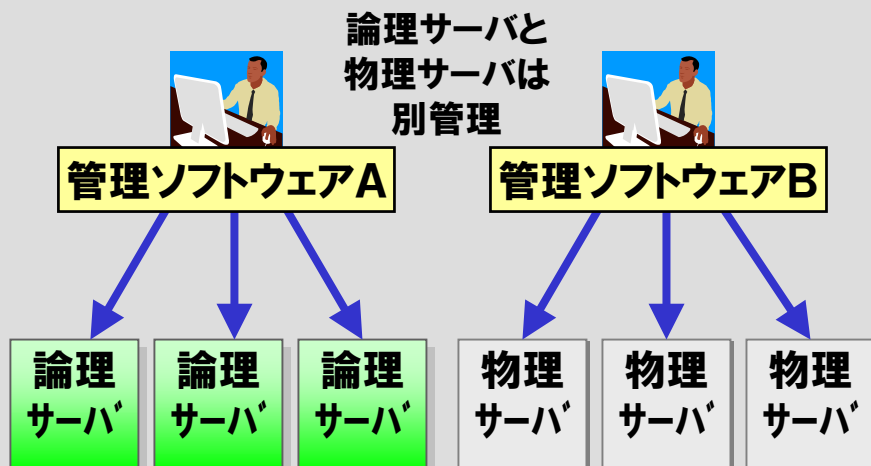
■ 物理サーバ/論理サーバの一元運用管理

従来の課題

- 物理/論理サーバの管理インタフェースが異なる
- ➡ 物理/論理サーバを意識した別々の管理ソフトによる管理

Virtageによる解決

- 物理/論理サーバの管理インタフェースをSVPで統一
 - ➡ 物理/論理サーバを意識しない一つの管理ソフトによる管理
- N+1コールドスタンバイもサポート



SVP:Service Processor

2.1.9

Virtage を支援する管理ソフトウェア

日立サーバ仮想化機構を支援する運用管理ソフトウェア

プラットフォームソリューション for BladeSymphony
ベストプラクティススイート for BladeSymphony

ミドルウェア

HiRDB

Cosminexus

主要ISV製品

仮想化環境上でも
ベア環境と
同等な機能を提供

統合システム運用管理

JP1

システム管理ソフトウェア

BladeSymphony Manage Suite

上位レイヤに
ハードウェアの
仮想化環境を
意識させない

ブレードサーバ



内蔵ディスク
アレイ装置



内蔵FCス
イッチ

内蔵LAN
スイッチ



F5、
CISCO
等IHV
製品

**JP1 / ServerConductor
で実現**



BladeSymphony

2.1.10 システム管理ソフトウェアによる容易なシステム運用・管理

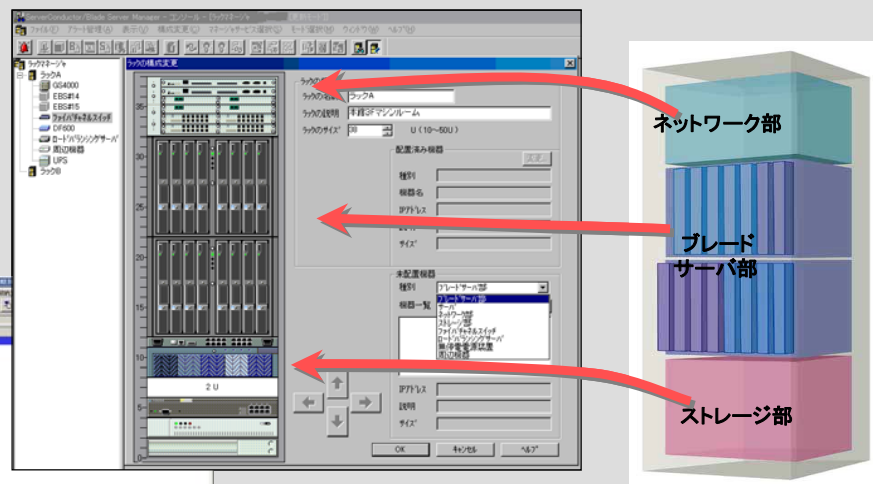
システム管理ソフトウェア「BladeSymphony Manage Suite」

システムの複雑さを意識することなく、システム全体の一元管理を実現。
1つのコンソールから、各ハードウェアの初期設定や構成変更が行え、
しかも、障害に関する情報も一元管理。万一の障害時にも迅速に対応できます。

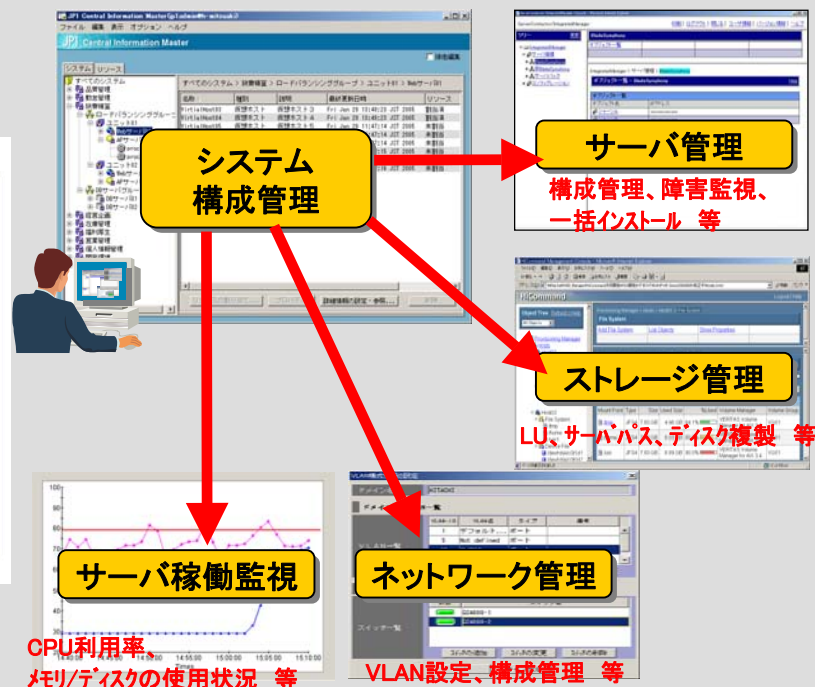
- たとえば、◆OSやアプリケーションの一括インストールがGUI操作で可能
- ◆VLAN構成情報の設定や変更がGUI操作で一度にまとめて可能
- ◆ストレージ部の設定・管理を論理的な階層構成のもとGUI操作で可能

ラック管理

BladeSymphonyのラック全体を操作性に優れたGUI画面で簡単に管理可能



ServerConductor/Blade Server Manager Plus



2.2. Services Oriented Storage Solutions

2.2.1

「Services Oriented Storage Solutions」

お客様のビジネス、ITシステム価値を高めるために、ハード/ソフト/サービスを組合せ、お客様の業務に対して最適な統合ストレージソリューションを提供します。
(2007年5月15日発表)



2.2.2

日立ストレージ仮想化機能

ディスクアレイ装置上での仮想化技術(世界初)により、
機種異なる複数のストレージ統合を実現



*Quality of Service

2.2.3

USP Vが提供する新たな顧客価値

ストレージ仮想化をさらに進化させて、
多様化する業務環境に最適なストレージ環境を提供するとともに、
データ量増大やサービス拡大にも柔軟に対応できるストレージ基盤を提供し
お客様のストレージTCOを削減します。

ボリューム容量の仮想化

ストレージ使用効率を向上、
管理を容易化

スケーラビリティ向上

仮想化容量 最大 247PB
ボリューム数 最大 64k



プロセッサのロードバランス

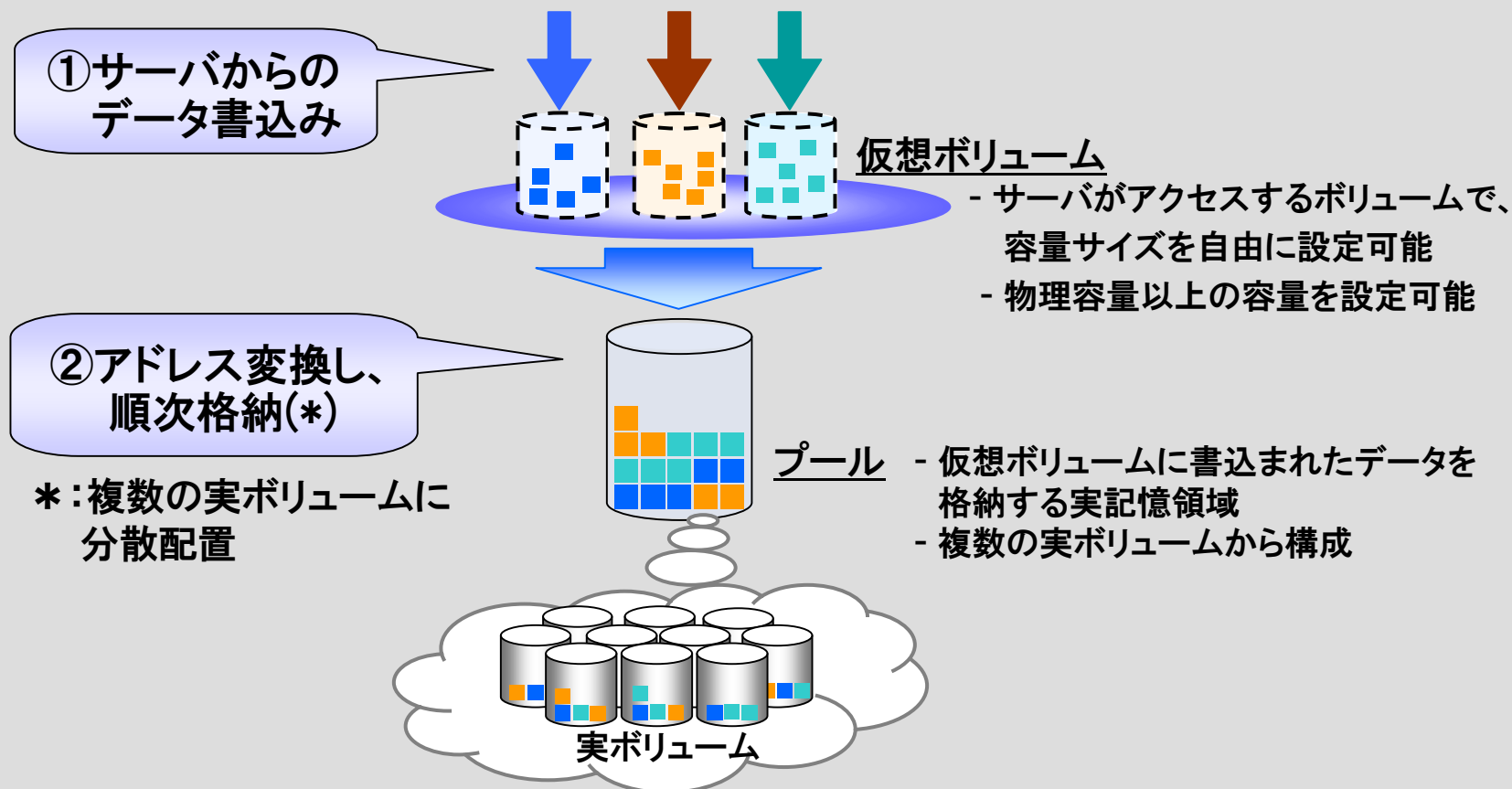
パフォーマンスを最大化

Hitachi Universal Storage Platform V

2.2.4

ボリューム容量仮想化機能

- 自由なサイズのボリュームを仮想ボリュームとして定義
- 仮想ボリュームに書込まれたデータはプールに順次格納(分散配置)

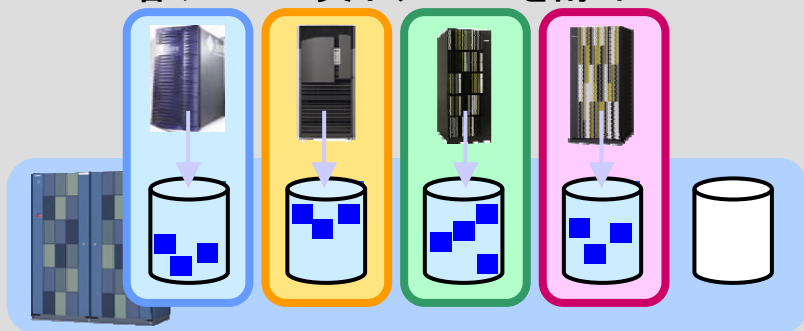


書込みデータ分のみ物理容量を消費する技術により、ストレージの使用効率向上します。

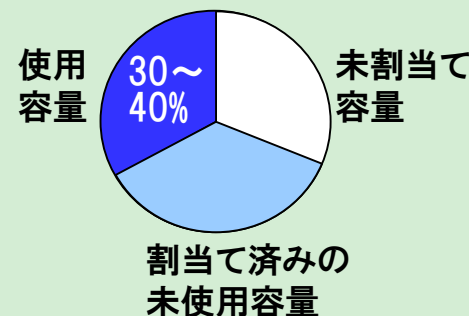
2.2.5 ボリューム容量仮想化による使用効率向上

従来

各サーバに実ボリュームを割当て



割当て済み未使用容量は
別用途での利用不可

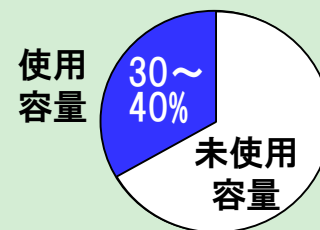


ボリューム容量仮想化 適用時(社内適用事例)

各サーバに仮想ボリュームを割当て

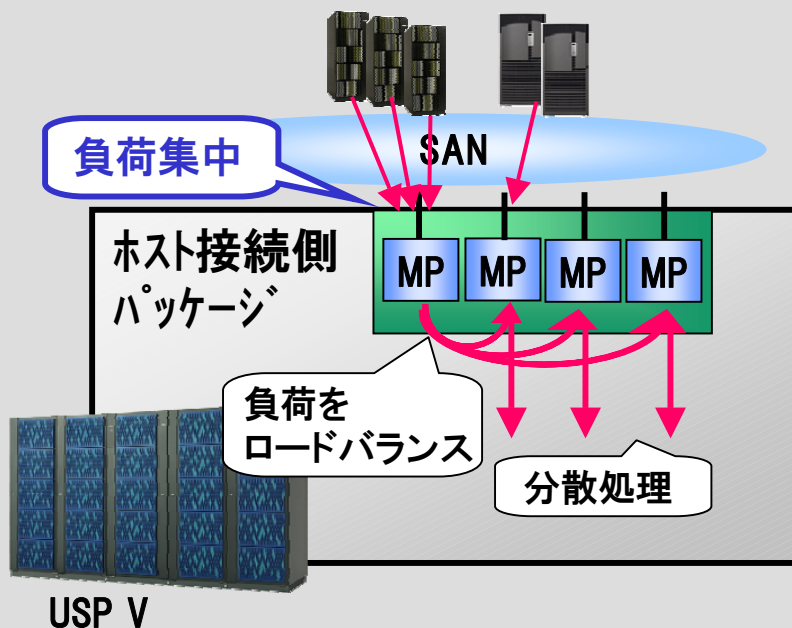


別用途でも利用可能な
未使用容量を最大化



2.2.6 プロセッサ有効活用による高性能の実現

- サーバアクセスに対し、プロセッサ間で負荷のロードバランスを実施
- 外部ストレージへのアクセスに対しても、プロセッサ間でロードバランスを実施し、スループットを最大化



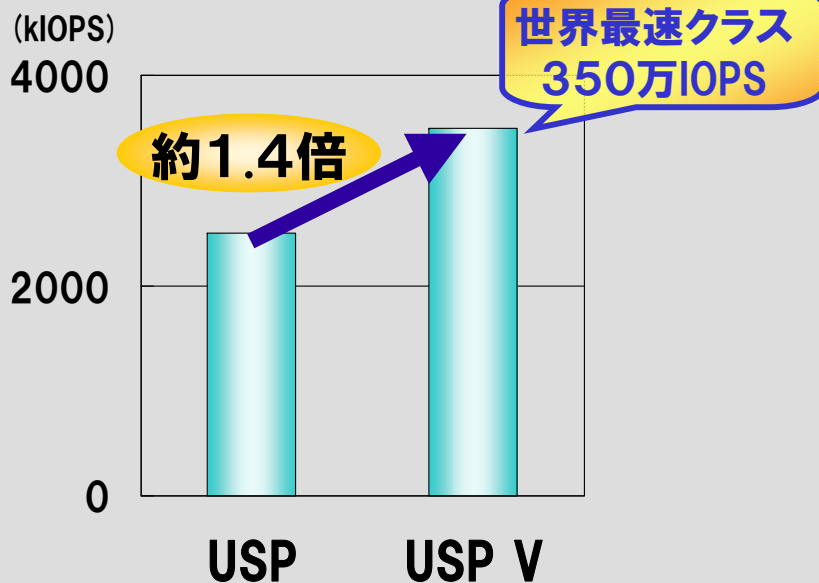
- 1パスにアクセスが集中する業務の処理時間を短縮化
- パス毎のアクセス負荷が異なってもストレージ側の処理をロードバランス
- パス毎のアクセス負荷を意識した性能事前設計を不要化

特定パスへのアクセス集中時も、負荷をロードバランスする新技術によりパフォーマンスを最大化し、お客さまのストレージ投資対効果を向上

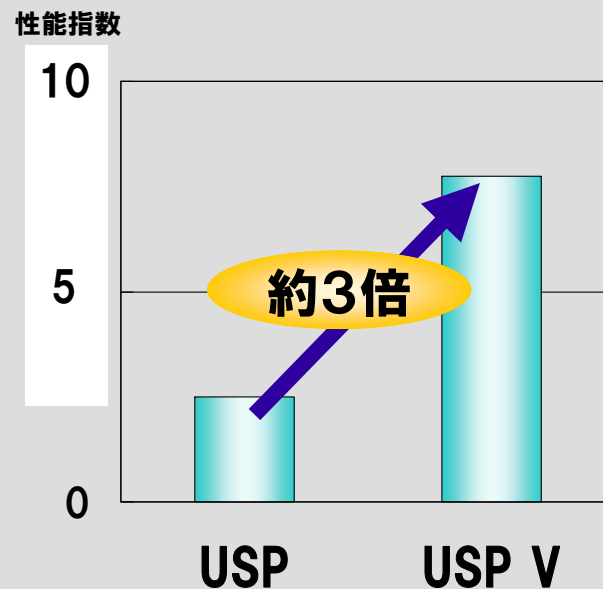
2.2.7 プロセッサのロードバランス による性能向上

最大システム性能向上に加え、1パス性能も向上

DBトランザクション処理性能 (Read)
(最大FCパス構成、キャッシュヒット環境)



DBトランザクション処理性能 (Read)
(1FCパス構成、キャッシュミス環境)



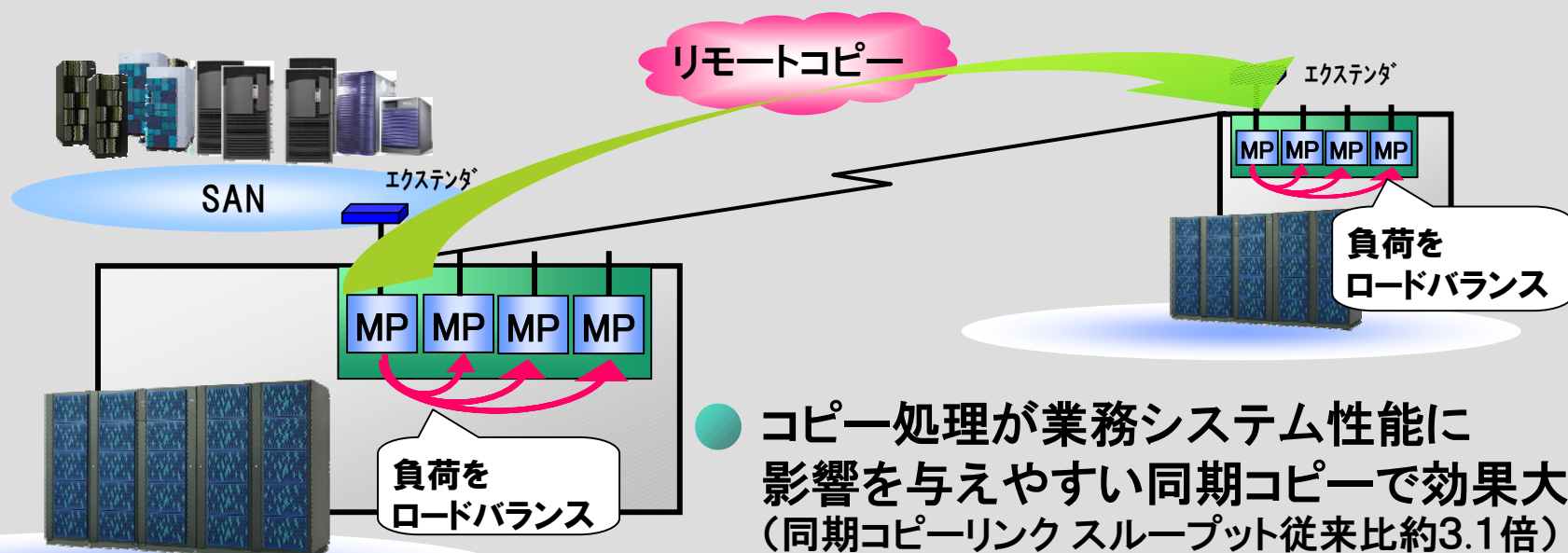
パス毎のアクセス負荷を意識した性能事前設計を不要化

2.2.8

リモートコピー構成のパフォーマンス最大化

リモートコピーリンクの処理スループットを向上

- コピー処理に対し、プロセッサ間で負荷のロードバランスを実施



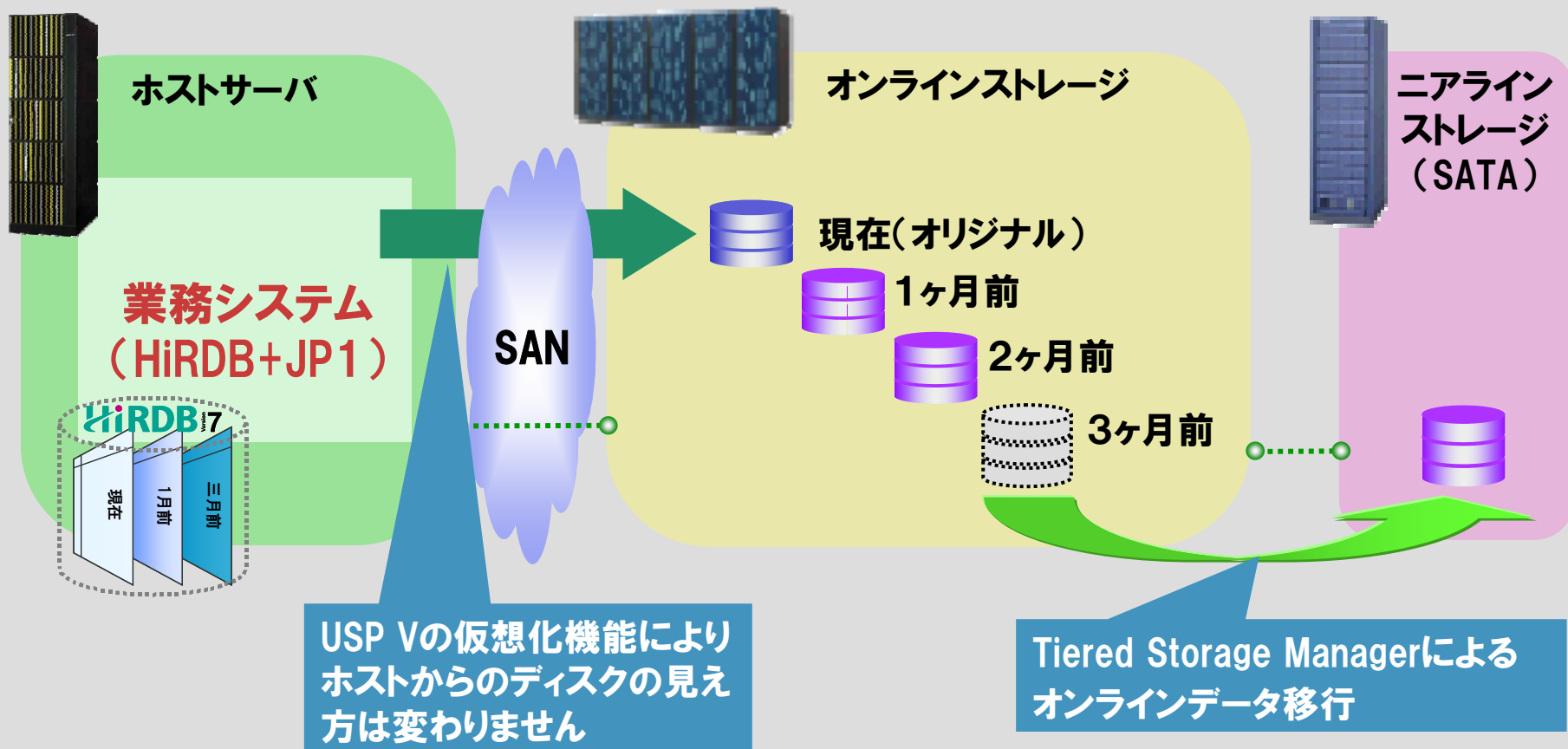
- コピー処理が業務システム性能に影響を与えやすい同期コピーで効果大（同期コピーリンク スループット従来比約3.1倍）

コピーリンク数が限られるリモートコピーシステムでも
スループットを最大化でき、ストレージ投資対効果を向上

2.2.9

仮想化を活用したストレージ階層管理

- ◆ 業務システムで月次管理され、利用頻度が低くなったデータを JP1 / HiCommand Tiered Storage Manager により **業務を止めずに低価格ストレージへデータを再配置**
- ◆ データの価値に応じたストレージデバイスの効率的な使い分けとコスト最適化が可能



2.3 運用管理ソフトウェア JP1の 自律制御機能

ビジネスレベル運用実現に向けて

- ビジネスレベルのPDCAサイクルを実現する4つの製品カテゴリ



2.3.2

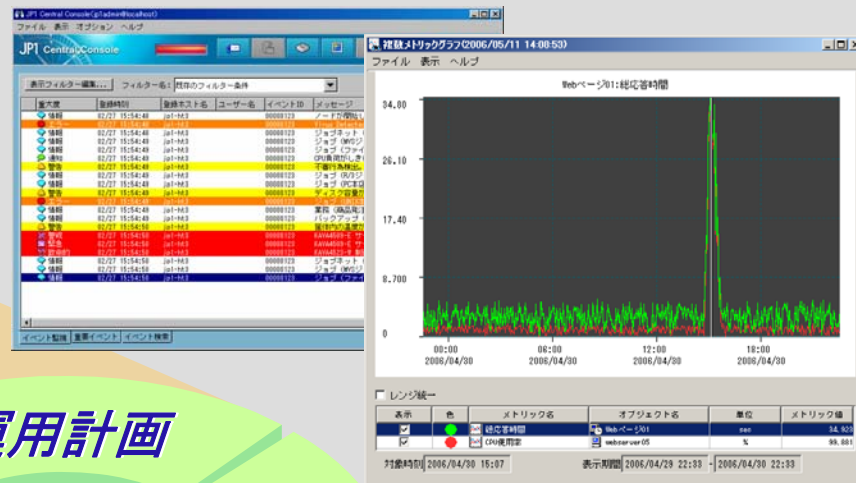
SOAシステムの運用PDCAサイクル最適化～モニタリング～

SOAシステムにおける運用PDCAサイクルを最適化するために
監視から判断、対処に至る一連の流れをシステム化

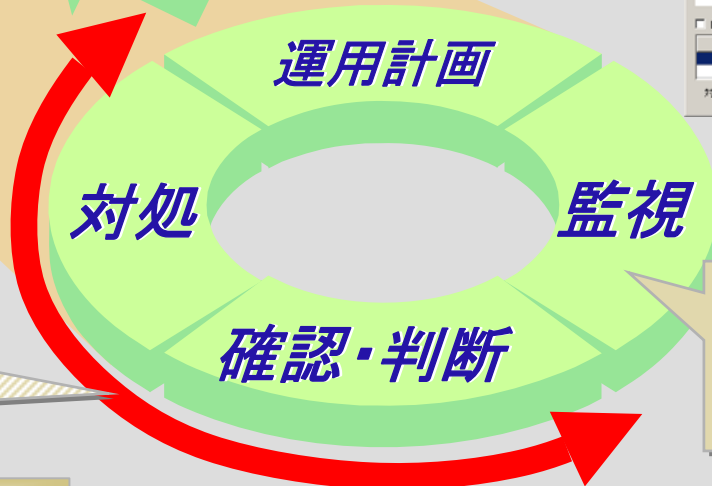
SOAシステムの
運用PDCAサイクル



最適化するため



従来
人の経験則や運用手順書
に沿った手動対応



SOAシステムの監視
・Webサービスレスポンス
・J2EEサーバー

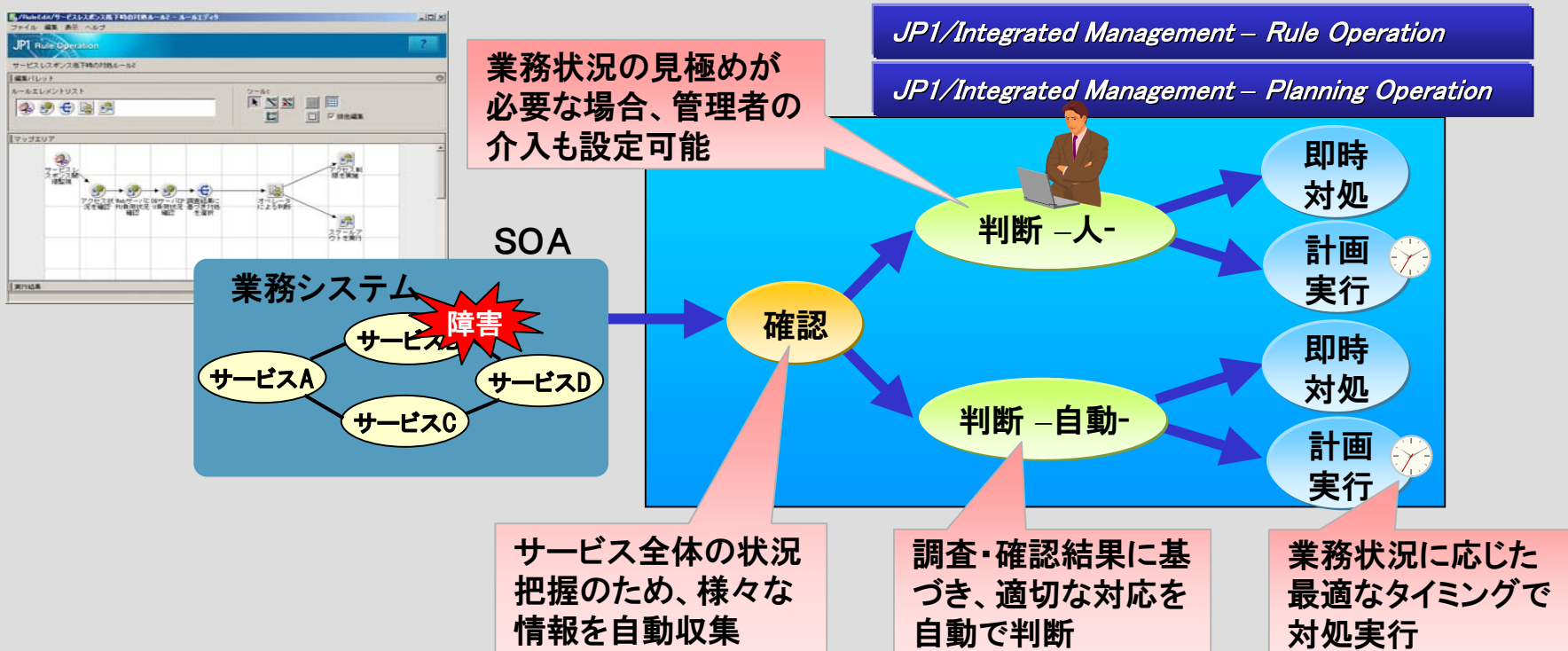
JP1

監視から判断、対処までを自動化し柔軟な運用を実現

2.3.3

自律制御によるPDCAサイクルの実現

確認・判断・対応の運用ノウハウを「ルール」としてシステム化
SOAシステムの問題発生から対応までの作業を効率化！



JP1
による自動化



従来

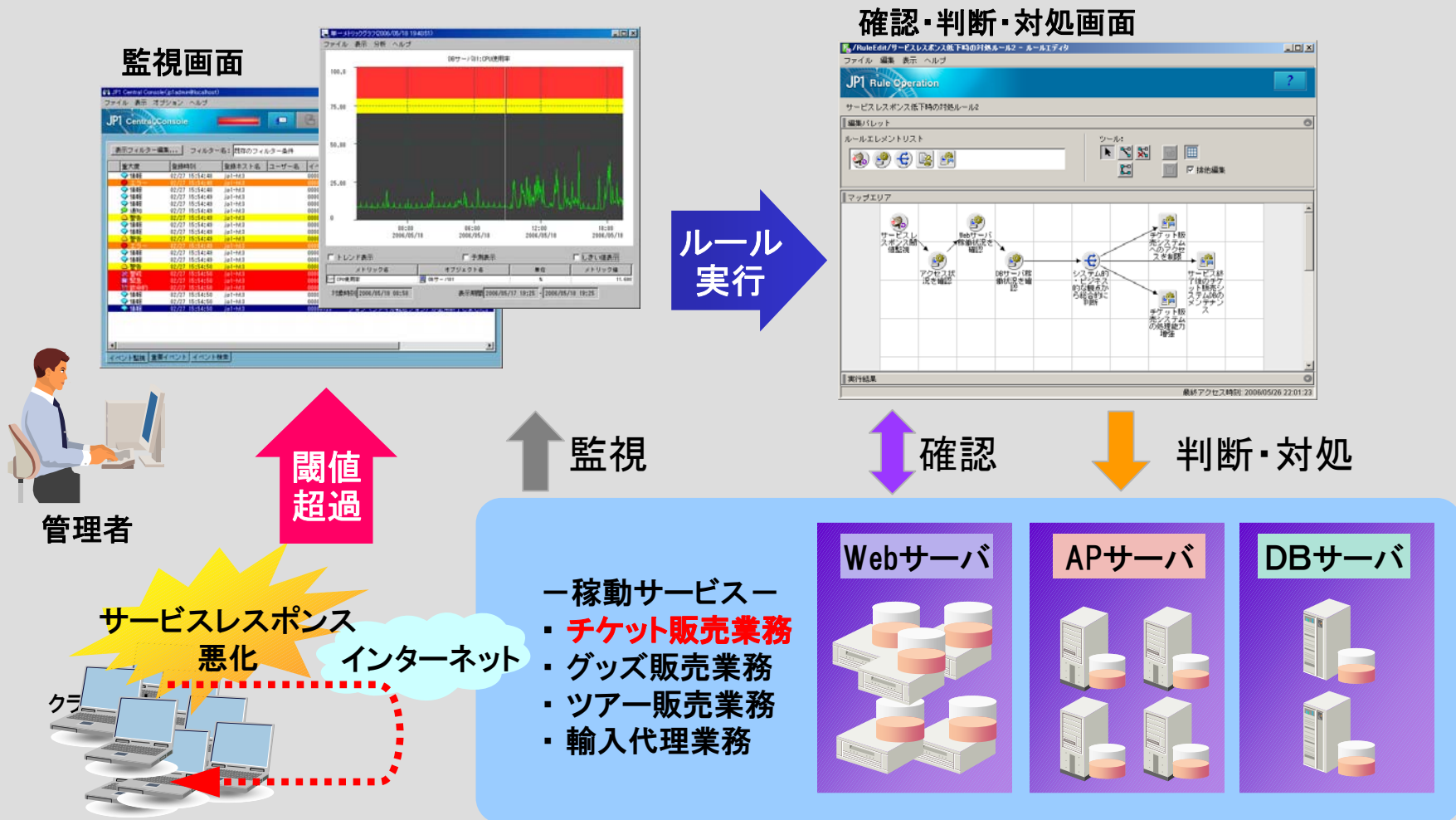


対応者のスキルに依存

2.3.4

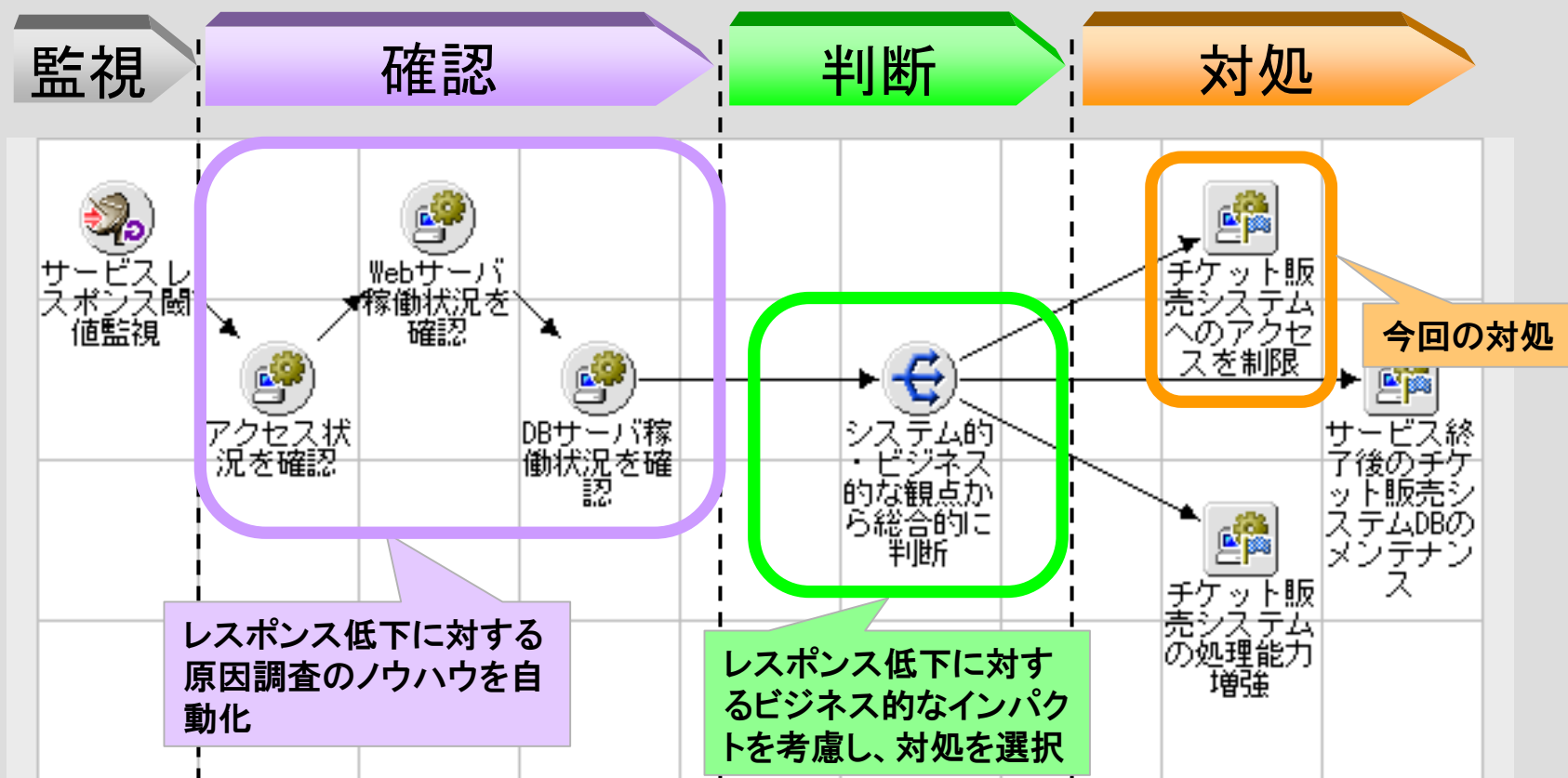
自律制御の適用例

- SOAで構築したチケット販売業務サービスにおいて、枚数限定の希少チケットの販売サービスで、Webサービスのレスポンス低下を検知した場合



2.3.5 ルールの定義例

● JP1 のモニタリングによる自動化

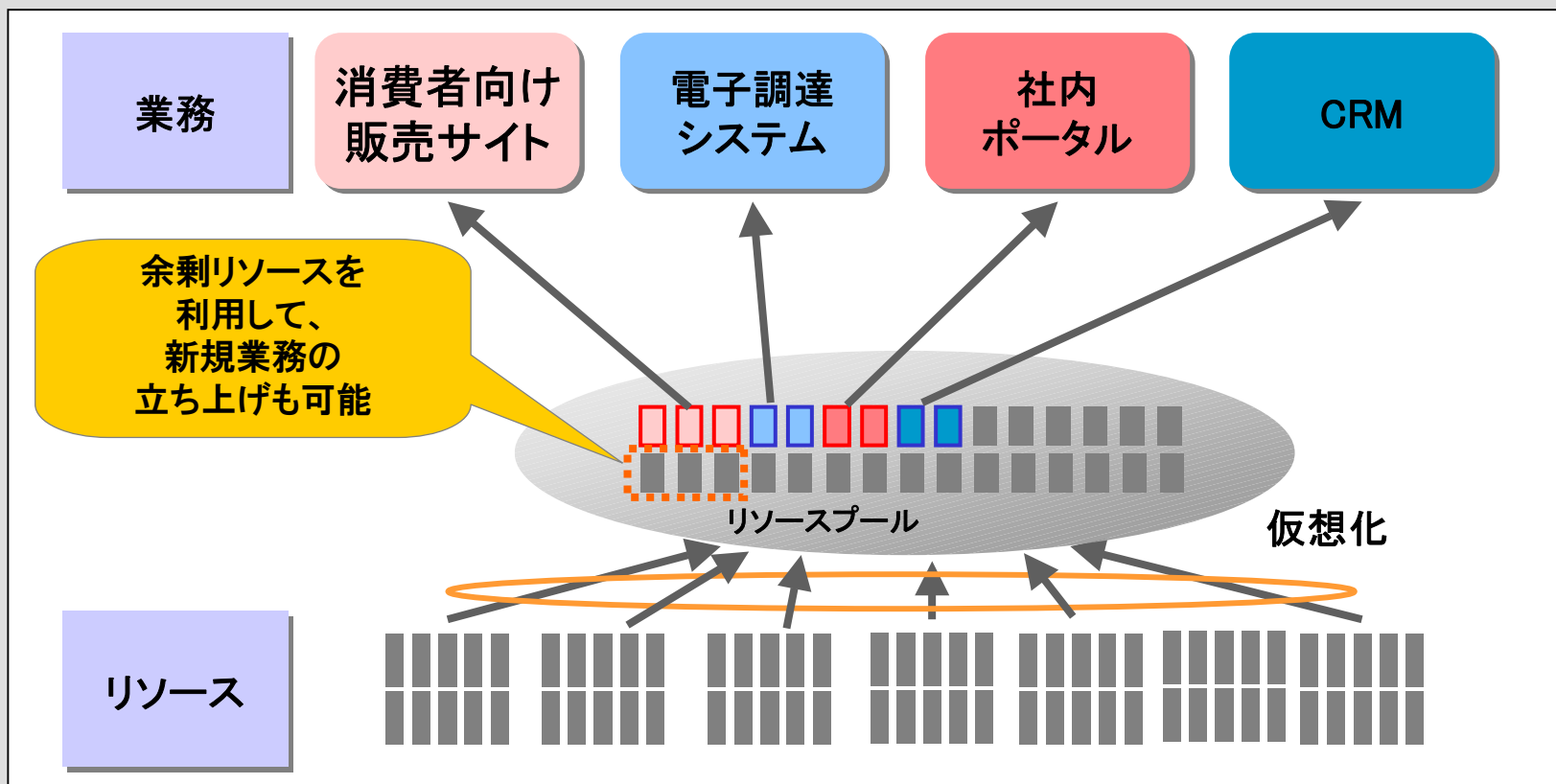


ビジネスへの影響も調査・判断した対応で
SOAシステムのサービス品質を維持

3. グリッド技術の適用シーン

仮想化技術を駆使した大規模データセンタ

- ・ 負荷のピークが異なる複数の業務を組み合わせて、リソース有効活用
- ・ 業務量の予測が難しい新規事業も余剰リソースを利用して、立上げのハードルを下げる
- ・ 負荷変動への対応、故障リソースへの対応等の自動化により運用コストを低減

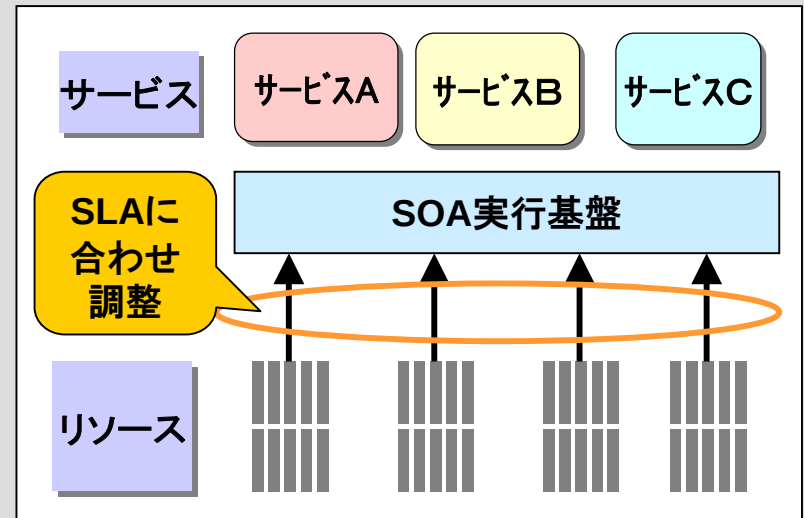


3.2

SOA/SaaS

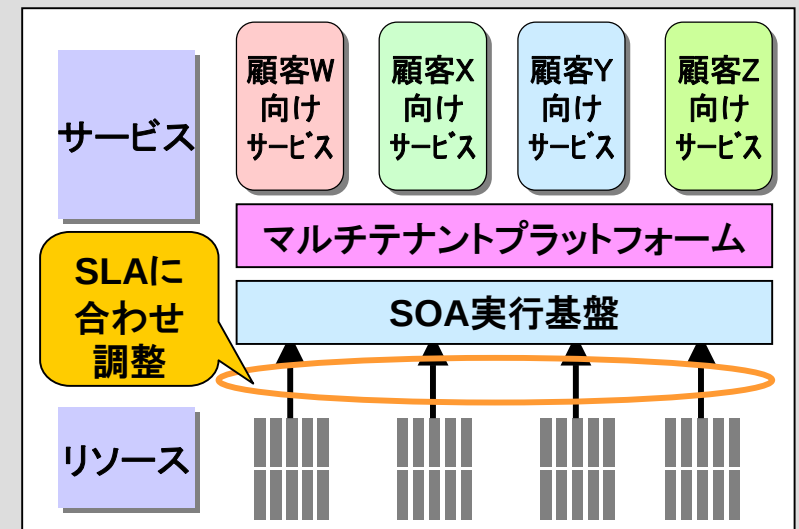
●SOA適用大規模情報システム

- ・各サービスのSLAを守るための自律運用を実現
- ・サービスのスケーラビリティ、可用性等の非機能要件を確保する
- ・運用コストを削減する



●SaaSプロバイダ

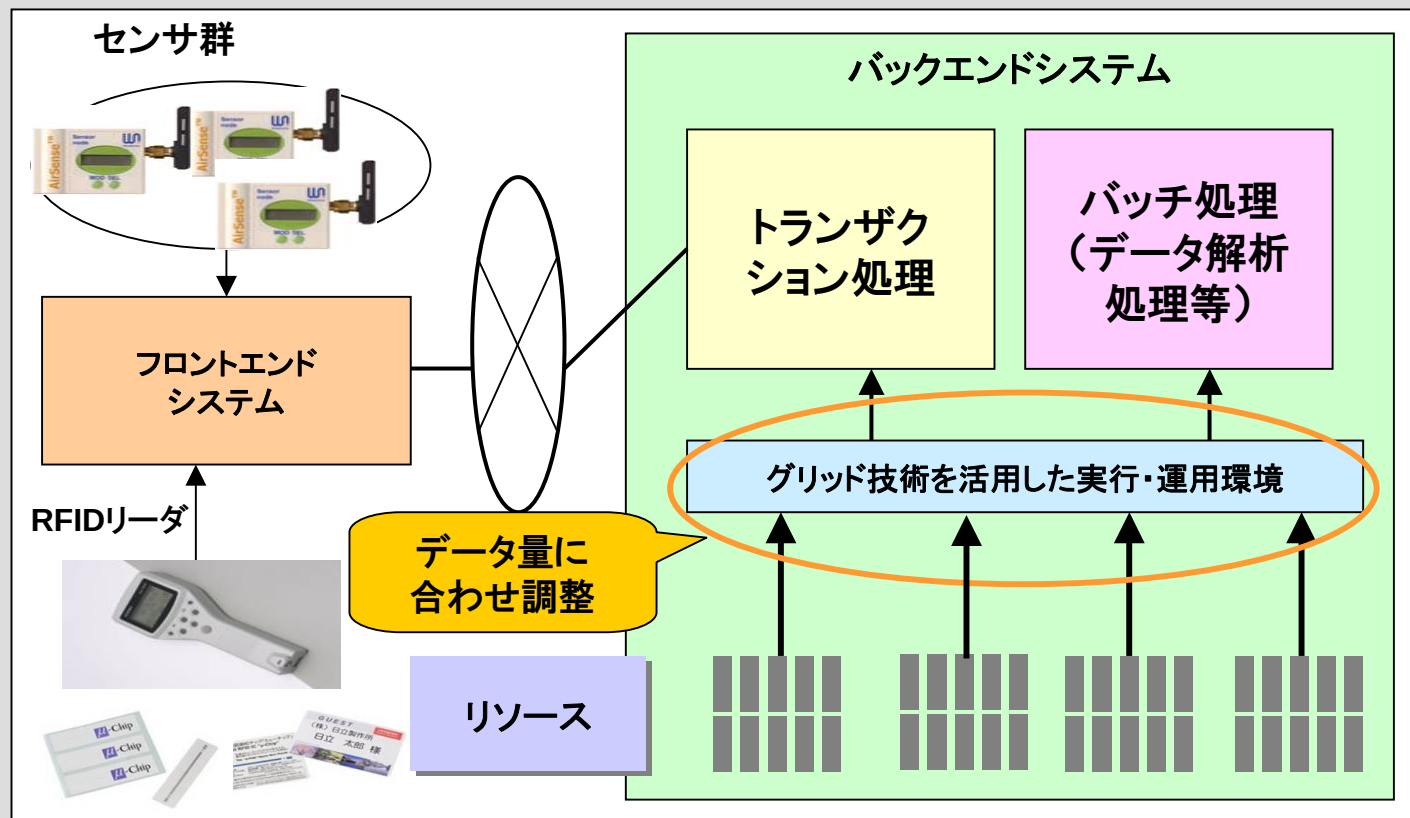
- ・顧客とのSLAを守るための自律運用を実現する
- ・顧客数の増大、顧客の業務量の変化に柔軟に対応する
- ・運用コスト削減により収益を確保する



3.3

ユビキタス情報システムのバックエンドシステム

- ・センサ、RFID等から発生する膨大なデータを安定して処理
- ・予測困難なデータ量の変化がある場合にもリソースの自動割当てにより対応
- ・トランザクション処理とデータ解析処理等のバッチ処理とのリソース割当てを自動調整
- ・自律運用により運用コストを削減



他社商品名、商標等の引用に関する表示

- ・インテル、Intel、Pentium、Itaniumは、アメリカ合衆国およびその他の国におけるIntel Corporationまたはその子会社の商標または登録商標です。
- ・Linuxは、Linus Torvaldsの米国およびその他の国における登録商標あるいは商標です。
- ・Unixは、X/Open Company Ltd.が独占的にライセンスする米国ならびに他の国における登録商標です。
- ・Microsoft、Windows、Windows Server、Microsoft Excel、Microsoft Access、Visual C++、Windowsロゴは、米国Microsoft Corporationの米国およびその他の国における商標または登録商標です。
- ・Red HatならびにShadow Manロゴは、米国およびその他の国でRed Hat, Inc.の登録商標もしくは商標です。
- ・その他記載の会社名、製品名は、それぞれの会社の商標もしくは登録商標です。

END

グリッドに対する日立製作所の取組み
ー サーバ、ストレージ、ミドルウェア ー

2007/05/31

株式会社日立製作所 ソフトウェア事業部
先端情報システム研究開発本部 担当本部長
吉野 松樹