

オープン・クラウド・プラットフォーム 構築の秘訣

OpenSource People Seminar

ビジネスの現場であなたを成功に導くオープンソース

日本IBM株式会社

システムズ&テクノロジー・エバンジェリスト

新井真一郎



@araishin

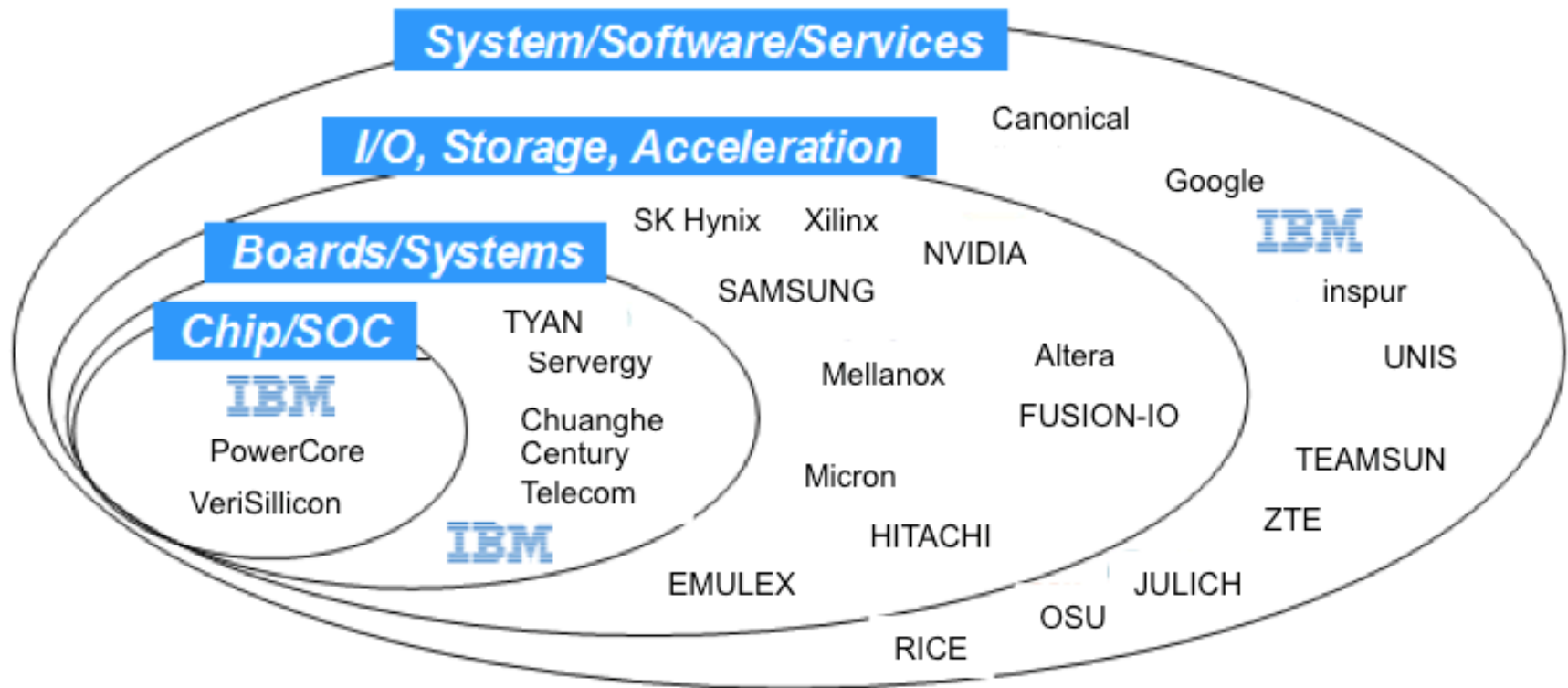
クラウド・プラットフォームのオープンな開発コミュニティ

OpenPOWER Foundation

OpenPOWER Foundation

サーバー史上初のオープンコンピューティングプロジェクト

- POWERアーキテクチャーをチップレベルからシステムレベル、さらにはその上のソフトウェア・スタックまでオープン化
- コミュニティによる開発が次世代データセンターのイノベーションを加速



各レイヤーに属するメンバーは、POWER8テクノロジーと
自社テクノロジーのコラボレーションによる新たな価値の創出を実現

ビッグデータの課題解決に向けて キャッシュ / バンド幅を最大5倍に強化した POWER8

	Sandy Bridge EP	Ivy Bridge EP	Ivy Bridge EX	POWER 7+ Systems	POWER8
Clock rates	1.8–3.6GHz	1.7-3.7GHz	1.9-3.4 GHz	3.1–4.4 GHz	3.0-4.1 GHz
SMT options	1,2*	1, 2*	1, 2*	1, 2, 4	1, 2, 4, 8
Max Threads / socket	16	24	30	32	96
Max L1 Cache	32KB	32KB**	32KB**	32KB	64KB
Max L2 Cache	256 KB	256 KB	256 KB	256 KB	512 KB
Max L3 Cache / socket	20 MB	30 MB	37.5 MB	80 MB (eDRAM)	96 MB (eDRAM)
Max L4 Cache	0	0	0	0	128 MB (eDRAM)
Memory Bandwidth	31.4-51.2 GB/s	42.6-59.7 GB/s	68-85*** GB/s	100 – 180 GB/sec	230 - 410 GB/sec

* Intel calls this Hyper-Threading Technology (No HT and with HT)

** 32KB running in "Non-RAS mode" Only 16KB in RAS mode

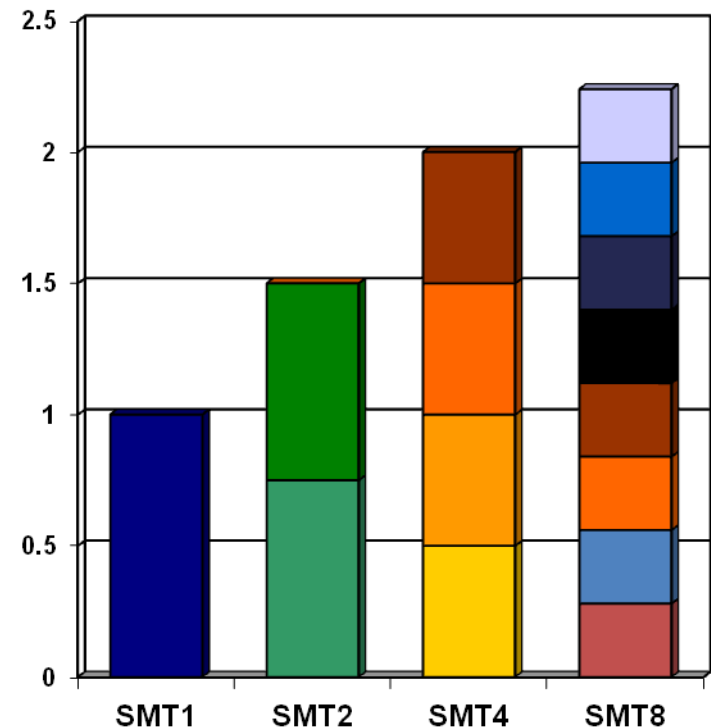
*** 85GB running in "Non-RAS mode" = dual-device error NOT supported

3 ~ 5 倍以上の帯域を確保

ビッグデータ/クラウド時代のITインフラを支えるテクノロジー

(1)スレッド高速処理：マルチスレッド

- 同時に複数の要求に応えることを可能にするマルチ・スレッド
 - SMT 8 に拡張
 - 実行ユニットを効率的に使用
 - 動的にモードを変更可能
- 他社の 4倍のスレッドを処理可能

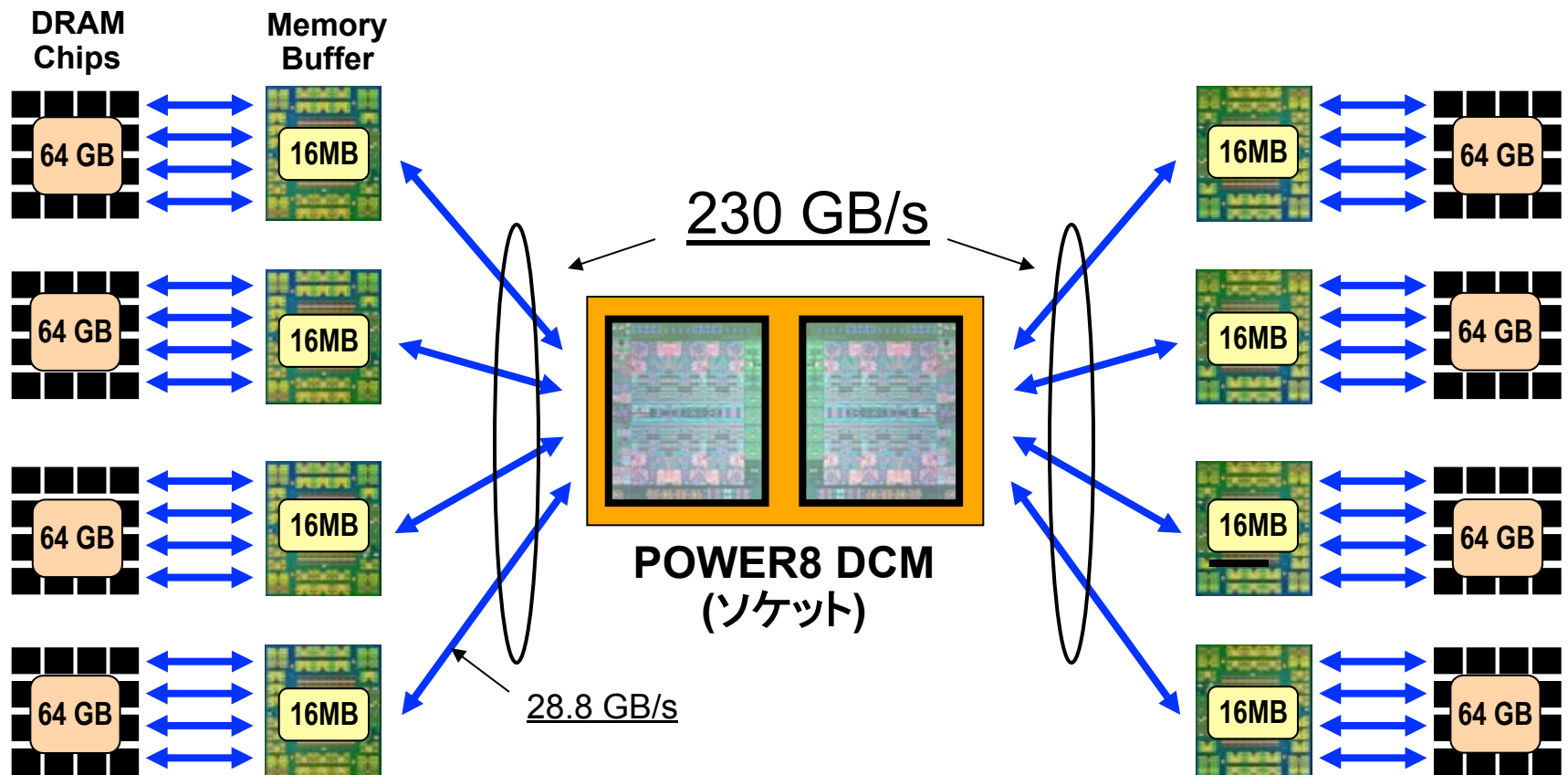


SMT 利用時のイメージ

ビッグデータ/クラウド時代のITインフラを支えるテクノロジー

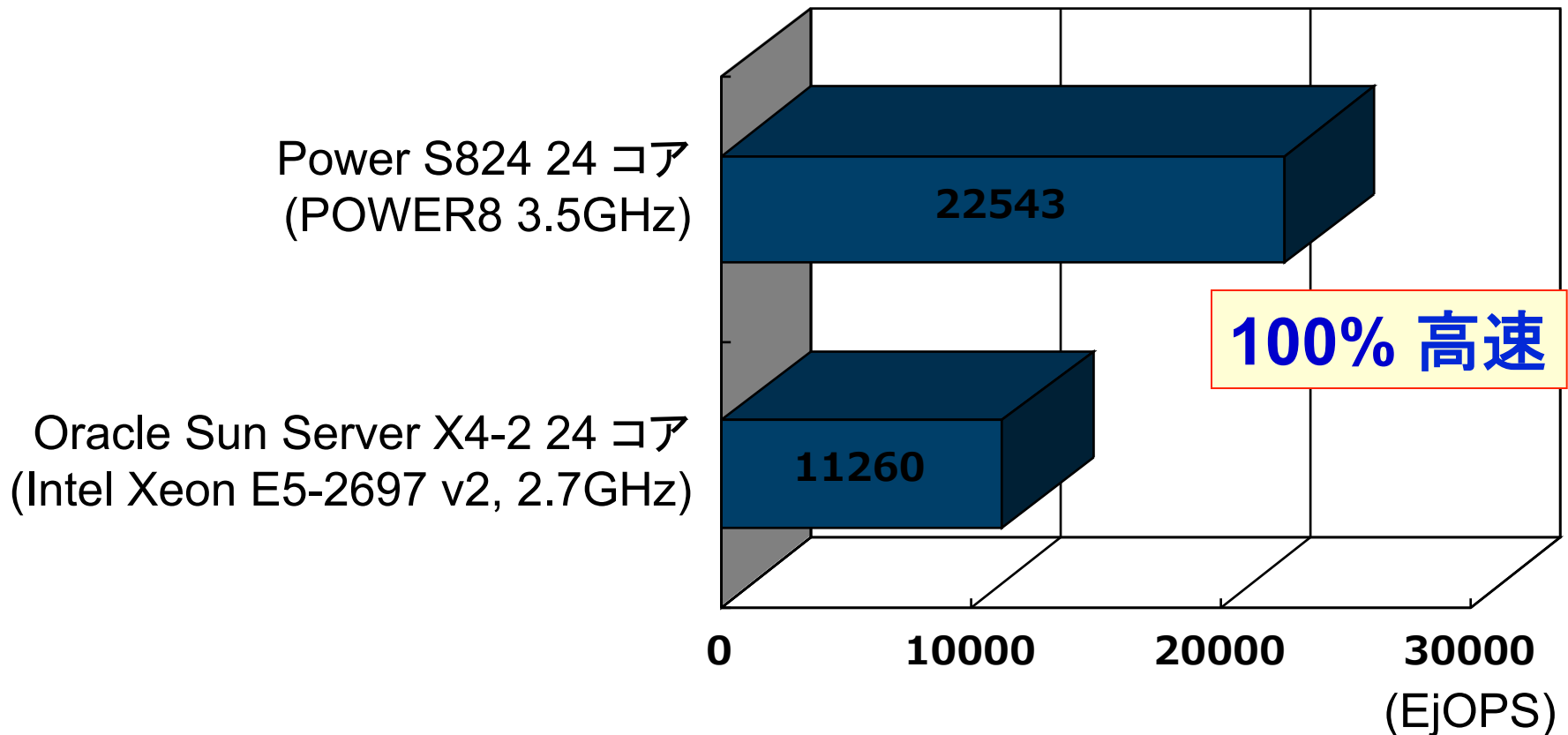
(2) メモリー高速処理：広帯域メモリーバンド幅

- POWER8 メモリー構成 (最大構成時^(*))
 - 230GB/s のメモリーバンド幅を実現



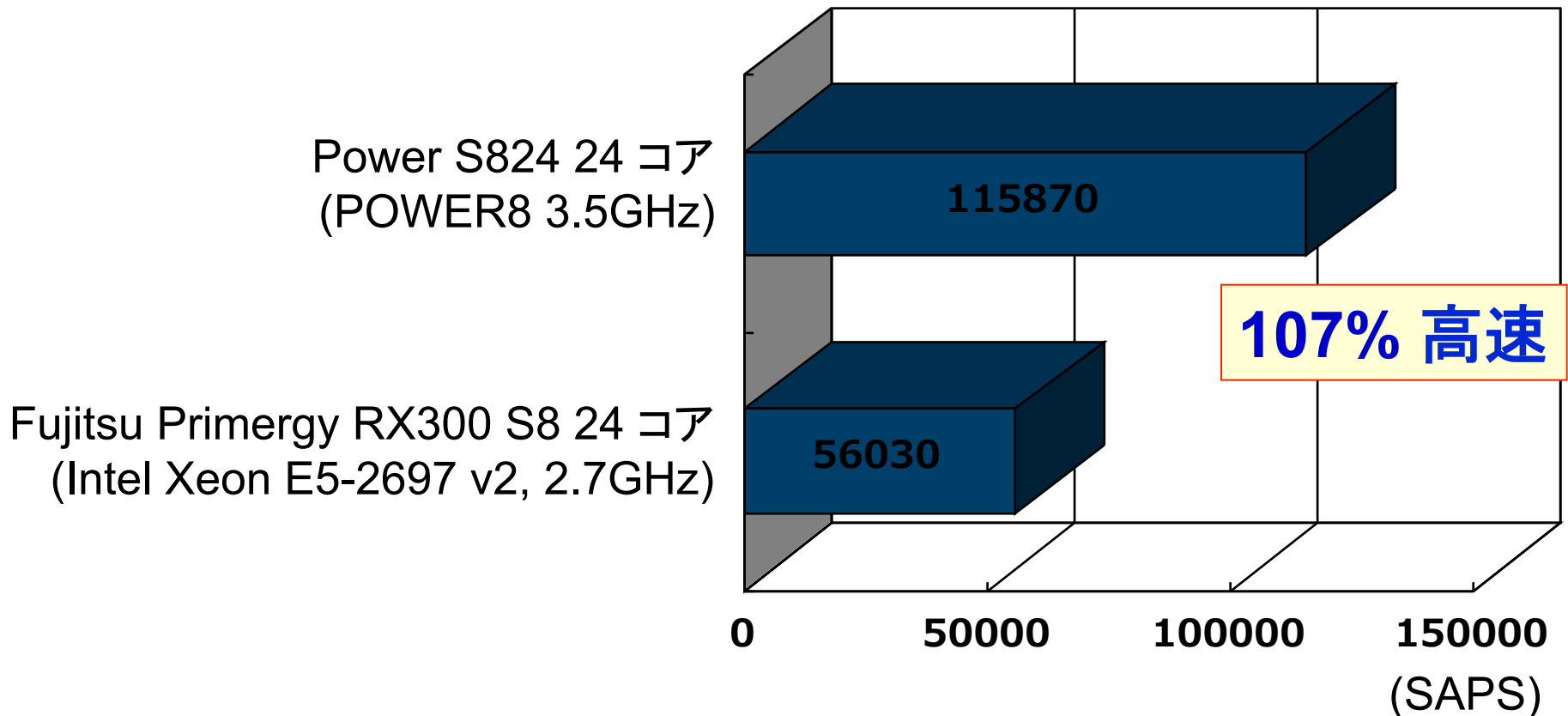
パフォーマンス比較:SPECjEnt2010 (JAVAトランザクション)

POWER8 搭載マシンと他社最速機との比較(24 コア構成)



パフォーマンス比較: SAP SD (商用トランザクション)

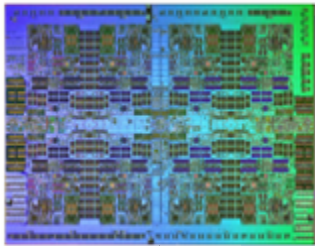
POWER8 搭載マシンと他社最速機との比較(24 コア構成)



ビッグデータ/クラウド時代のITインフラを支えるテクノロジー

(3) I/O高速処理：PCIeデバイス高速接続

POWER7



GX
Bus

I/O
Bridge

PCIe G2

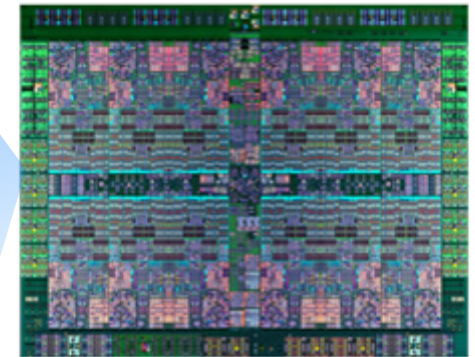
PCIe Gen 3 のネイティブサポート

- プロセッサとの直接結合
- 低いレイテンシー
- Gen3 x16 バンド幅

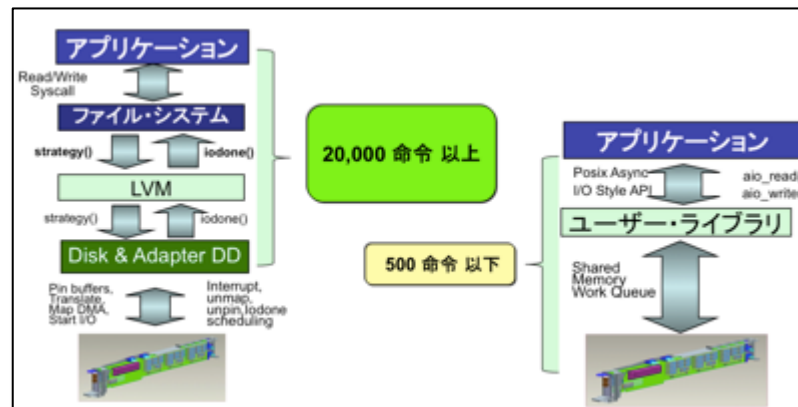
CAPI Protocol トランスポート・レイヤー

- 柔軟かつ迅速にイノベーションを活用できる業界初のテクノロジーCAPI(Coherent Accelerator Processor Interface)
- PCIe 経由のプロセッサと機器との密接接続
- PCIe でのプロトコルのカプセル化

POWER8



PCIe G3



OpenPOWERコミュニティの成果物

Google

新サーバー用マザーボードに
POWER8を採用

NVIDIA

NVIDIA GPUアクセラレーター
-初のJava向け GPUアクセラレーター
-Hadoopアナリティクス・パフォーマンスが飛躍的に向上

Mellanox

POWERにおけるRDMAの活用
-スループット10倍
-Key-Value Storeアプリケーションの遅延を改善

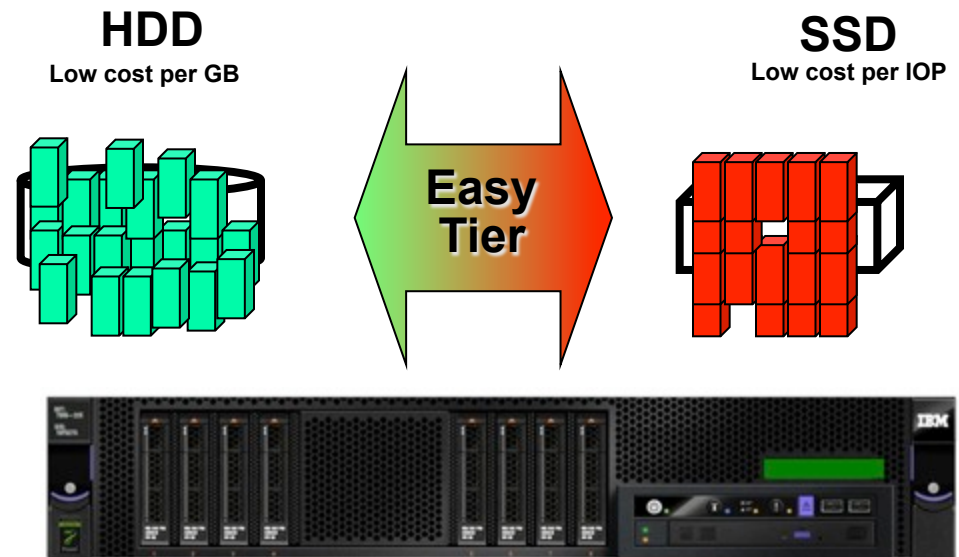
XILINX
ALTERA

CAPI接続FPGAアクセラレーター
-消費電力あたりの性能を35倍
-200倍の高速化を実現する金融分野向けモデル

ビッグデータ/クラウド時代のITインフラを支えるテクノロジー

(4) ディスク性能最適化：ディスク自動階層化

- サーバー本体内存蔵ディスク内で、自動的にホット・データをSSDに配置、コールド・データをHDDに再配置する“Easy Tier”機能をサポート
 - アプリケーションの変更は不要
 - RAIDアレイとして構成
 - 前提OS:AIX V7.1 TL3 SP3, or later; AIX V6.1 TL9 SP3, or later; RHEL 6.5, or later; SLES 11 SP3, or later; or VIOS 2.2.3.3, or later



ビッグデータ活用を徹底的に追求して、 設計、開発されたプラットフォーム

Power Systems 

ビッグデータを超高速処理するための設計

高い経済性と信頼性を備えた**クラウド**基盤

イノベーションを加速する**オープン**プラットフォーム

POWER8 プロセッサ搭載の 「IBM® Power System」新製品を発表

(工場出荷日: 2014年6月10日*1)

スケールアウト型 クラス



Power S812L

- 1ソケット
- 最大12コア
- メモリ: 512GB
- 2Uラックマウント



Power S822L

- 2ソケット
- 最大24コア
- メモリ: 1TB
- 2Uラックマウント



Power S822

- 2ソケット
- 最大20コア
- メモリ: 1TB
- 2Uラックマウント



Power S814

- 1ソケット
- 最大8コア
- メモリ: 512GB
- 4Uラックマウント



Power S824

- 2ソケット
- 最大24コア
- メモリ: 1TB
- 4Uラックマウント

1 & 2 Sockets

POWER8



Power **KVM** Power **SC** Power **VP**
Power **HA** Power **VC** Power **VM**

*1 Power S812Lのみ工場出荷は第三四半期を予定

簡単かつ効果の大きい、x86/Linux 向け Java アプリケーションの POWER8 対応事例



JBアドバンステクノロジー 株式会社

WebReport

- ✓ ビッグデータを利活用する
情報分析・予測分析のBIツール

ポーティング理由

- ✓ データ分析の**高速処理**
- ✓ さらなる**信頼性・堅牢性**の確立



- ✓性能処理は **最大12倍**
- ✓非常に**スムーズなポーテ
ィング**

株式会社アイズファクトリー

bodais

- ✓ ビッグデータ自動解析プラットフォーム

ポーティング理由

- ✓ **ビッグ・データの処理**に適し
た**クラウド**基盤
- ✓ より強固な**セキュリティ**



- ✓ポーティングは**5時間**で完了
- ✓優れた**IT利用効率**
 - ✓大量データ処理でも増えない**メモリー使
用量**
 - ✓**CPU使用効率**の向上

OSSでオープンなクラウド プラットフォームを構築する秘訣

これからの企業で使うクラウドは 適材適所で複数サービスを使い分ける



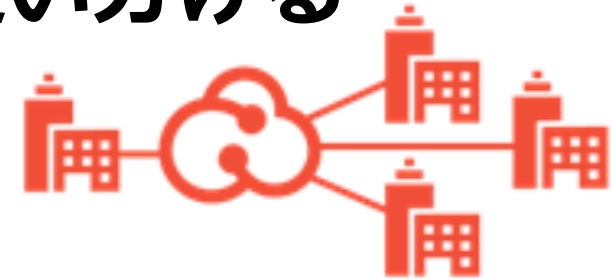
Private cloud

企業の占有形態によるクラウド・モデル。
自社ですべてを保有することも、第三者に管
理・運用を委託することも可能



Hybrid IT

従来型のITシステムとクラウド・モデルを独立させ
たまま、データとアプリケーションの共通利用を行
えるようにしたITシステム



Public cloud

クラウド・サービスを提供するプロバイ
ダーの従量制サービスを利用する共有型
クラウド・モデル



Traditional IT

従来型のハードウェア、ソフトウェア、ネットワーク
設備などを統合して構築したITシステム。

クラウドは適材適所に使い分ける

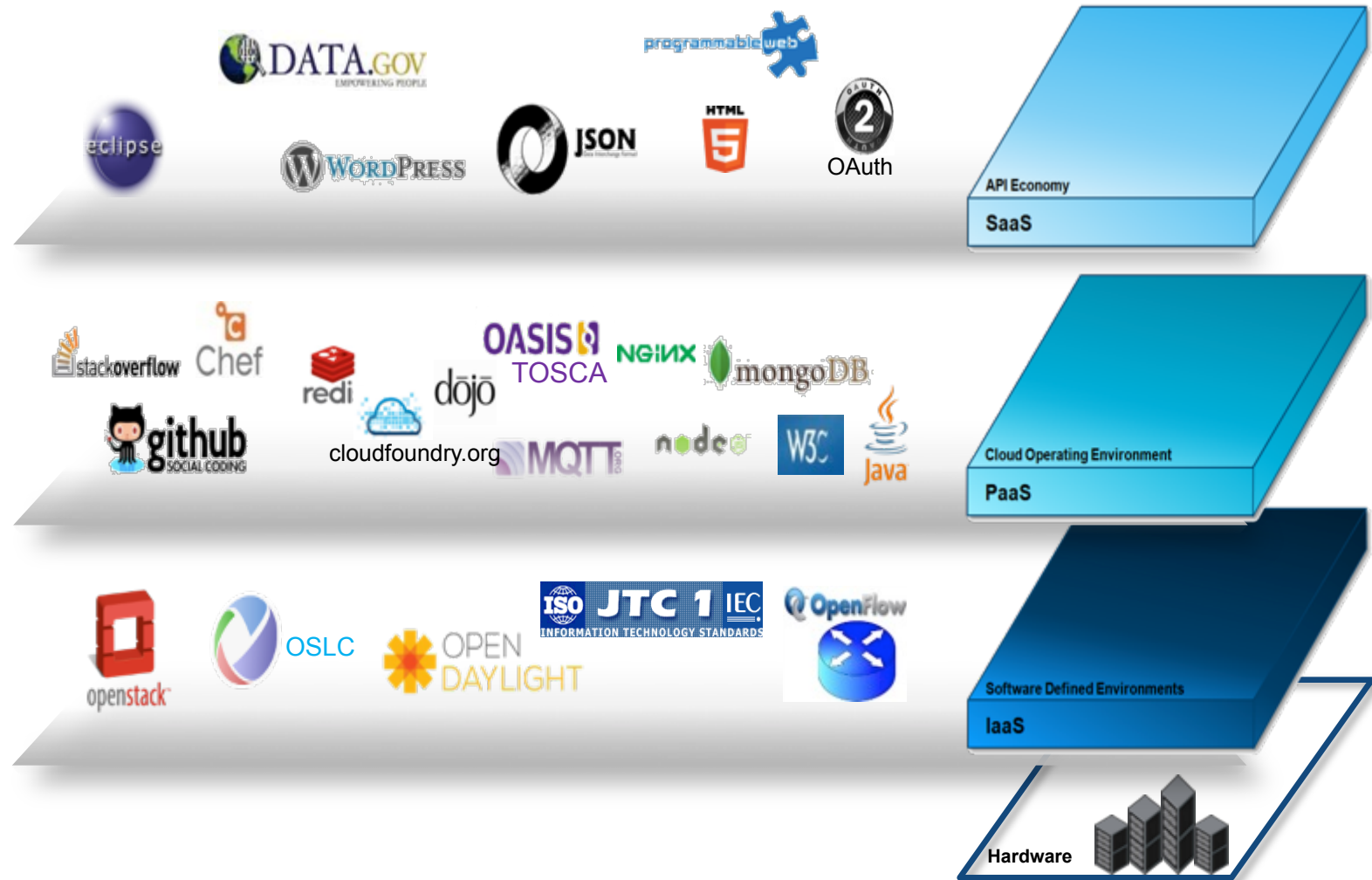
- ✓ オンプレミス・システムとパブリッククラウドは二者択一ではありません。
- ✓ 業務毎に適材適所に使い分けることが重要です。

選択基準の例

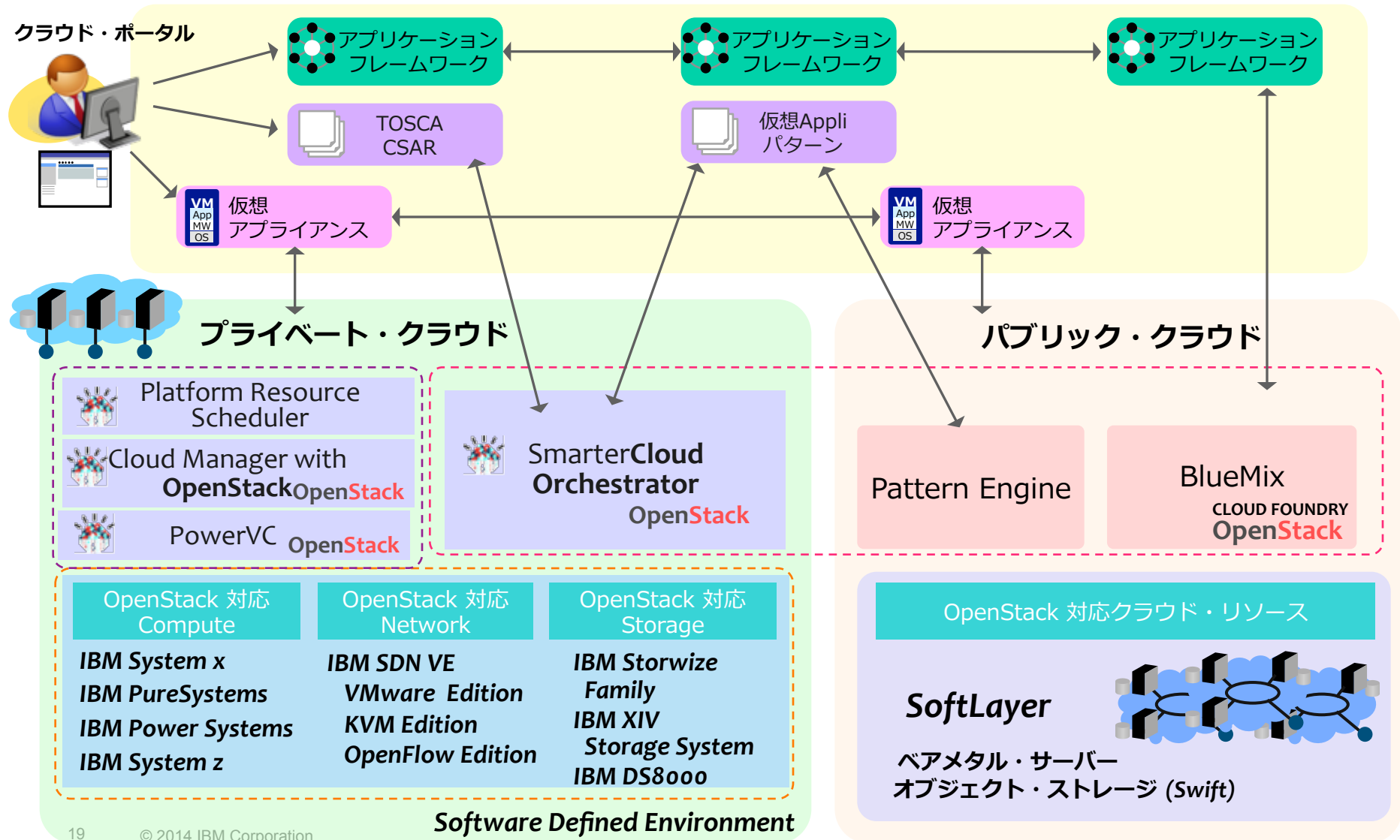
項目	内容
パフォーマンス	<ul style="list-style-type: none">・ 個々の仮想マシンに割当てるリソース量を柔軟にコントロールしたい・ パフォーマンスを要する場合、仮想マシンを追加するか、リソースコントロールを行うか
ストレージ	<ul style="list-style-type: none">・ 必要なストレージ性能に対して最適化されたストレージサービス、リソースを割当てたい・ IO要件が厳しい場合のシステム構成をどのように選択するか
SLA	<ul style="list-style-type: none">・ サービスレベルは、どのレイヤー（OS, ミドルウェア, アプリケーション, 業務システム）まで必要とされるか・ 厳しいサービスレベルが必要とされているか
セキュリティ	<ul style="list-style-type: none">・ 自社ルールを継続適用するのか、サービス提供者の規定したルールに準拠するのか
ソフトウェア・ライセンス	<ul style="list-style-type: none">・ 運用方法に適したライセンス体系になっているか・ 特にピーク時対応の確認が必要

オープンクラウドアーキテクチャ

さまざまなオープンな技術を活用してクラウドを実現します。

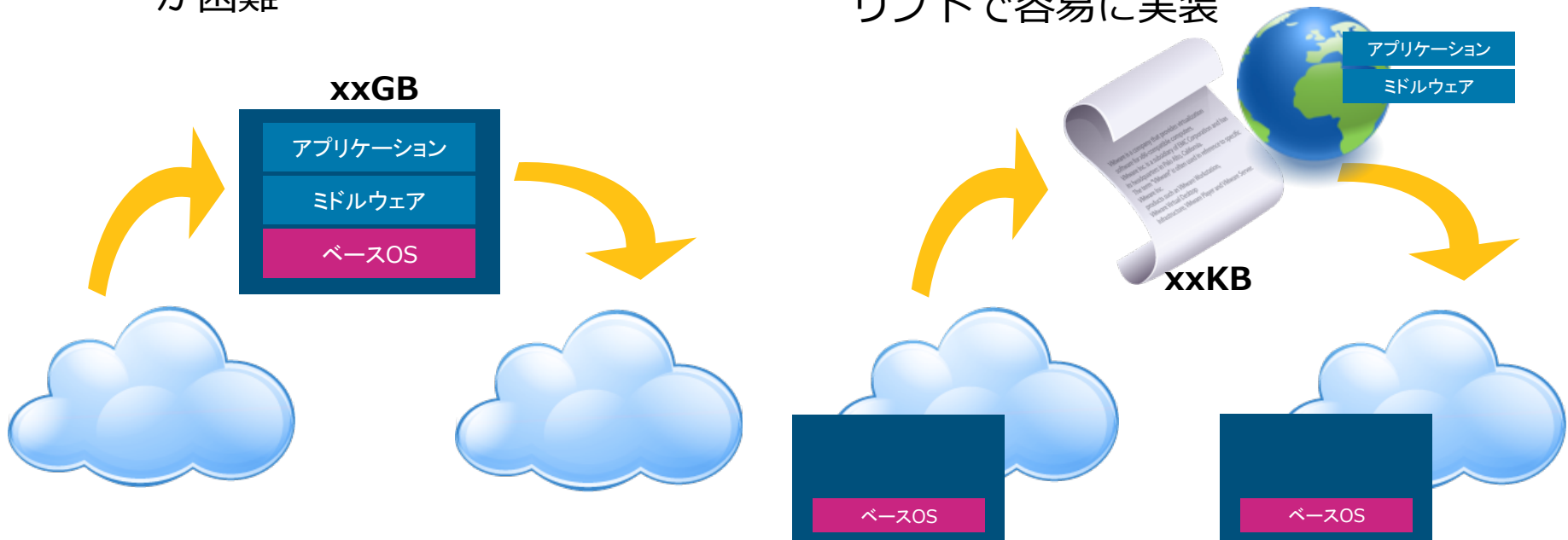


OpenStackで複数のクラウドを透過的に



クラウド環境に向けたデプロイメント方式

- 仮想アプライアンス方式
 - マシン・イメージのファイル・サイズが巨大になるためクラウド間の可搬性に欠ける
 - マシン・イメージ単位でパッケージ化するため各種コンポーネントのアップデート保守が困難
- スクリプト導入方式
 - アプリケーション、ミドルウェアの導入・構成はスクリプトで自動化
 - クラウド間を移動するのは、システム構築のスクリプトのみ。移動先のクラウドですぐに目的のシステムを再現。
 - コンポーネントのアップデートもスクリプトで容易に実装



OSSを活用したオープンなクラウド環境が ビジネスのイノベーションを加速



市場における競争力の強化



ポリシー・ベースでの自動運用

- 新サービスを生むサービス同士のAPI連携
- サービス特性を捉えたポリシーベース自動運用
- サービス特性に応じた最適なクラウド環境の自動選択



個客マーケティング



複数サービスの連携



業務システムの標準化

- 業務システム単位のパターン化による標準化
- パターンによる業務システムのデプロイ
- パターンのカタログ管理



柔軟な災害対策

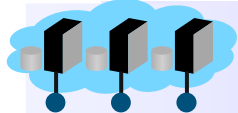


運用管理コスト削減



ハイブリッド・クラウド化

- 複数のクラウドの統合管理
- ID統合、ポータル統合、データ統合、監視統合
- クラウドをまたぐクロス・オーケストレーション



プライベート・クラウド

- 仮想マシンのデプロイメント自動化
- SDNを活用したネットワークの仮想化
- 仮想マシン払い出しスケジュール機能の実装
- クラウド・ユーザー向け/管理者向けポータルの提供



パブリック・クラウド



仮想化



- ハードウェア・リソースの仮想化

Power Systems



ワークショップ、セッション、および資料は、IBMまたはセッション発表者によって準備され、それぞれ独自の見解を反映したものです。それらは情報提供の目的のみで提供されており、いかなる参加者に対しても法律的またはその他の指導や助言を意図したものではなく、またそのような結果を生むものでもありません。本講演資料に含まれている情報については、完全性と正確性を期するよう努力しましたが、「現状のまま」提供され、明示または暗示にかかわらずいかなる保証も伴わないものとします。本講演資料またはその他の資料の使用によって、あるいはその他の関連によって、いかなる損害が生じた場合も、IBMは責任を負わないものとします。本講演資料に含まれている内容は、IBMまたはそのサプライヤーやライセンス交付者からいかなる保証または表明を引きだすことを意図したものでも、IBMソフトウェアの使用を規定する適用ライセンス契約の条項を変更することを意図したものでもなく、またそのような結果を生むものでもありません。

本講演資料でIBM製品、プログラム、またはサービスに言及していても、IBMが営業活動を行っているすべての国でそれらが使用可能であることを暗示するものではありません。本講演資料で言及している製品リリース日付や製品機能は、市場機会またはその他の要因に基づいてIBM独自の決定権をもっていつでも変更できるものとし、いかなる方法においても将来の製品または機能が使用可能になると確約することを意図したものではありません。本講演資料に含まれている内容は、参加者が開始する活動によって特定の販売、売上高の向上、またはその他の結果が生じると述べる、または暗示することを意図したものでも、またそのような結果を生むものでもありません。パフォーマンスは、管理された環境において標準的なIBMベンチマークを使用した測定と予測に基づいています。ユーザーが経験する実際のスループットやパフォーマンスは、ユーザーのジョブ・ストリームにおけるマルチプログラミングの量、入出力構成、ストレージ構成、および処理されるワークロードなどの考慮事項を含む、数多くの要因に応じて変化します。したがって、個々のユーザーがここで述べられているものと同様の結果を得られると確約するものではありません。

記述されているすべてのお客様事例は、それらのお客様がどのようにIBM製品を使用したか、またそれらのお客様が達成した結果の実例として示されたものです。実際の環境コストおよびパフォーマンス特性は、お客様ごとに異なる場合があります。

IBM、IBM ロゴ、ibm.com、Power Systems、PureSystems、およびSystem zは、世界の多くの国で登録されたInternational Business Machines Corporationの商標です。

他の製品名およびサービス名等は、それぞれIBMまたは各社の商標である場合があります。
現時点での IBM の商標リストについては、www.ibm.com/legal/copytrade.shtmlをご覧ください。

Linuxは、Linus Torvaldsの米国およびその他の国における登録商標です。