

Platform Computing

ITリソースを最適化する
Platform Computingのグリッドソリューションのご紹介

July, 15, 2004
Platform Computing KK
Tsuyoshi Kitamura

Accelerating Intelligence

Platform Computing

Platform Computing Inc.

- 1992年に設立
- 本社はカナダ(トロント)
- DRM(分散コンピューティング)で12年以上の実績
- 従業員
約350名(ワールドワイド)
- 北米・ヨーロッパ・アジアに20箇所の拠点
- プラットフォーム コンピューティング株式会社
 - 1999年12月設立
 - Platform Computing Inc. 100%出資
 - 従業員: 11名
 - オフィス: 東京(新宿)

2 © Platform Computing KK 2004

Accelerating Intelligence

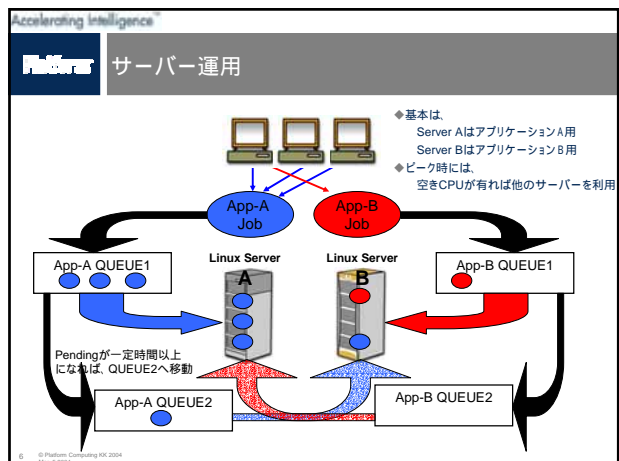
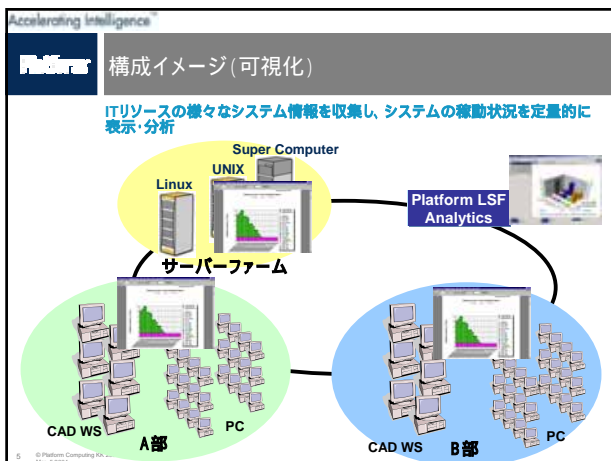
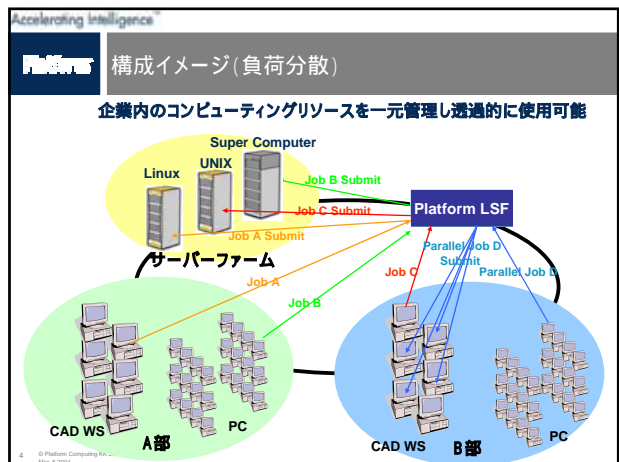
Platform Computing

Platform Computingについて

Platformの顧客(全世界で約1,500サイト)

Computer Manufacturing	Financial Services	Industrial Manufacturing	Life Sciences	Government & Research
<ul style="list-style-type: none"> AMD ATI Technologies Cisco HP IBM NVIDIA SUN Microsystems Texas Instruments 	<ul style="list-style-type: none"> Deutsche Bank Fidelity Investments JPMorgan Chase Morgan Stanley Royal Bank of Canada Société Générale TD Bank 	<ul style="list-style-type: none"> BMW Bombardier Daimler Chrysler General Electric General Motors Lockheed Martin Pratt & Whitney Volkswagen 	<ul style="list-style-type: none"> AstraZeneca Bristol Myers-Squibb Celera Entelos Incyte Genomics Monsanto Sanger Institute 	<ul style="list-style-type: none"> ASCI Grid CERN DoD DoE ENEA Fleet Numeric TACC

3 © Platform Computing KK 2004



Accelerating Intelligence™

Platform Work Station Grid

ピーク需要に合わせた計算リソースが必要??

- 計算需要にはムラがある
- ピークに合わせて計算サーバーを持つのではなく、必要最小限のサーバーリソース

7 © Platform Computing INC 2004

Platform Products及び適用例のご紹介

Platform

Accelerating Intelligence™

Platform Platform Product Family

9 © Platform Computing INC 2004

Accelerating Intelligence™

Platform LSF Product Family Architecture

10 © Platform Computing INC 2004

Platform LSF

Platform

Accelerating Intelligence™

Platform Platform LSF

Enterprise Grid上のコンピューターリソースの負荷情報を収集・監視し、負荷状態及びジョブの特性に応じて最適なホストへ割り当て

- ◆ ネットワーク上のマシンを仮想クラスタ化
- ◆ 1台を「LSFマスター」にアサイン
- ◆ マスターが各マシンの負荷情報を収集
- ◆ ジョブ投入

12

Accelerating Intelligence

Platform LSF

- Scalability
 - 8,000+ CPUs/Cluster
 - Dynamic Node Membership
 - Multi Platform Support
 - UNIX, Linux, Windows, MacOS, SX, Cray, etc...
 - 100+ Clusters/MultiCluster
- Performance
 - 500,000+ Jobs/Cluster
 - 4kbyte/Job
 - Job Array, Chunk Job

13 © Platform Computing UK 2004

Accelerating Intelligence

Platform LSF

- 高スケジューリング機能
 - フェアシェアスケジューリング
 - GroupA (7) → User4 (2), User2 (4), User6 (4)
 - GroupB (4) → User1 (1), User2 (1), User3 (1), User7 (7)
 - GroupC (3)
 - リソース予約
 - 割り込みスケジューリング

	Checkpoint	Restart
低優先度		
高優先度		
 - etc.

14 © Platform Computing UK 2004

Accelerating Intelligence

Platform LSF

- 耐障害性
 - マスターのフェイルオーバー
 - ジョブの自動復元 (lsb.eventsファイル)
 - 二重イベントログ

15 © Platform Computing UK 2004

Accelerating Intelligence

Platform LSF

Extensive Application Support

Electronics		Industrial Manufacturing	Life Sciences
Agilent Technologies	SYNOPSYS	Cadabra	FLUENT
ANSYS	simplex	Mentor Graphics	TECHPOINT
cadence	Axis SYSTEMS	Verisity	ANSYS
Si SOLI	neoliner	Aptix	Schlumberger
SEQUENCE	Model Technology	Exa	accelrys
Supernaticity			Exa

16 © Platform Computing UK 2004

Accelerating Intelligence

Platform LSF

製造業での適用例 (Work Station Grid)

WSグリッド実現のための要件
優れたスケジューリング機能が不可欠

実行ノードの自動選択

- 使用可能なEWSを認識
 - 夜間・休日
 - 離席中
- 並列数のみ指定
- 同一グループ内のノードを選択
- etc.

17 © Platform Computing UK 2004

Accelerating Intelligence

Platform LSF

Goal-oriented SLA driven policies

What is it?

ユーザのSLA (Service Level Agreement) に基づく Goal Driven型ポリシー:

- Deadline based goal: ジョブのグループの終了時間を定義
- Velocity based goal: 同時実行が保証されるジョブの数を定義
- Throughput based goal: 1時間あたりの終了ジョブ数を定義

プロジェクトが時間内に完了することを支援
プロジェクトおよび管理のコストを軽減
プロジェクトの進捗を可視化
管理者のフォーカスを、資源の配分方法ではなく、どのワークをいつまでに完了させるかに集中させる

18 © Platform Computing UK 2004

Accelerating Intelligence®

Platform Goal-Oriented SLA-Driven Scheduling (Cont.)

使用例:

- 課題: 午後15:00までにすべての解析ジョブを終了させなければならない
- Solution: lsb.serviceClassesファイルdeadline service classを定義

```

Begin ServiceClass
  NAME=xyz
  PRIORITY=100
  GOALS = [deadline timeWindow (13:00 - 15:00)]
  DESCRIPTION = A simple deadline demo
End ServiceClass

```

- ジョブの投入およびモニタリング
 - \$bsub -sla xyz -W 30 -J A[1-50] mySimulation
 - \$date.bsla

```

Wed Aug 20 14:00:16 EDT 2003
SERVICE_CLASS_NAME: xyz
GOAL: DEADLINE ACTIVE_WINDOW: (13:00 - 15:00)
STATUS: Active:OnTime
DEAD_LINE: (Wed Aug 20 15:00)
ESTIMATED_FINISH_TIME: (Wed Aug 20 14:30)
Optimum Number of Running Jobs: 5
NJOBS PEND RUN SSUSP USUSP FINISH
  50   25   5           20

```

19 © Platform Computing INC 2004
PLC-200404

Accelerating Intelligence®

Platform Job Level Exception Management

What is it?

例外に基づくエラー検知により、是正処置を自動的に実行
ジョブの確実性と予見性を向上させる
ジョブおよびシステムエラーの検知
管理負荷と費用を軽減

New
Platform 15.0

20 © Platform Computing INC 2004
PLC-200404

Accelerating Intelligence®

Platform Job-Level Exception Management (Cont.)

使用例 1:

要件: 特定のホストで過去5分間に30以上のジョブが異常終了が発生した場合、ホストをクローズし、管理者にホスト名を通知

Solution:

- host exceptionsの定義 (lsb.hostsのEXIT_RATE)

```

Begin Host
  HOST_NAME MXJ   EXIT_RATE # Keywords
  Default       !    30
End Host

```

- lsb.paramsにJOB_EXIT_RATE_DURATION = 5を定義(デフォルト値=10分)

21 © Platform Computing INC 2004
PLC-200404

Accelerating Intelligence®

Platform Job-Level Exception Management (Cont.)

使用例 2:

要件: ジョブの実行時間が3時間以上になった場合、管理者にジョブIDを通知

Solution:

- lsb.queuesにjob exceptionsの定義

```

Begin Queue
  ...
  JOB_OVERRUN = 3*60 # run time in minutes
End Queue

```

22 © Platform Computing INC 2004
PLC-200404

Platform LSF HPC

Platform

23 © Platform Computing INC 2004
PLC-200404

Accelerating Intelligence®

Platform Platform LSF HPC

Platform LSFをベースにHPC(High Performance Computing)向けに最適化

- 並列ジョブ・ランチャーのサポート
LAMMPI, MPICH-GM, MPICH-P4, POE, SCALI, CHAMPION PRO, etc
- ハイパフォーマンス・インターコネクトのサポート
Myrinet with GM
IBM SP Switch
Quadrics QsNet
SGI Numa Flex SGI NumaLink


24 © Platform Computing INC 2004
PLC-200404

Accelerating Intelligence™

Platform Extensive Hardware Support

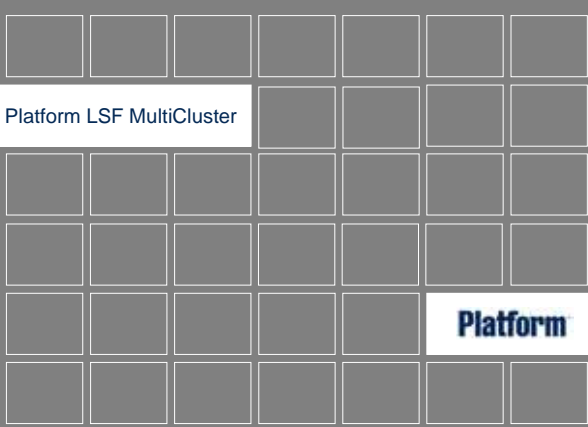
HP HP AlphaServer SC HP XC HP Superdome HP-UX 11i	SGI SGI IRIX SGI TRIX SGI Altix, SGI Propack
IBM IBM RS/6000 AIX IBM SP2/SP3	Linux IA-64 systems with RedHat Intel, AMD 32-bit systems with LINUX kernel
Sun SUN Solaris	

High Performance Interconnects
Myrinet with GM
Quadrics QsNet
SGI Numa Flex SGI NumaLink
IBM SP Switch



25 © Platform Computing UK 2004
UK-2004

Platform LSF MultiCluster



Platform

Accelerating Intelligence™

Platform Platform LSF MultiCluster

複数の Platform LSF cluster を横断したバッチアプリケーションワークロードをポリシーベースによりインテリジェントに処理

- 地理的に遠距離にあるクラスタを結合させる
- ローカル管理と資源の所有権を維持しつつ、世界中からのアクセスと協調を可能とする
- エンタープライズグリッドを構築するためのPlatformの商用ソリューション – スタンドベースおよびOGSI準拠

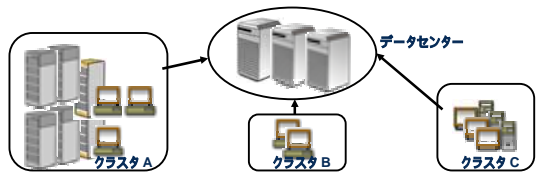
27 © Platform Computing UK 2004
UK-2004

Accelerating Intelligence™

Platform ジョブ フォワードリング モデル

ローカルクラスタよりジョブを自動的に、高性能サーバーを持つ大型のデータセンタークラスタに転送する

- ユーザから見た透過的な共有を向上
- サーバの統合を促進
- コンピューティングリソースの生産性や能力を増加



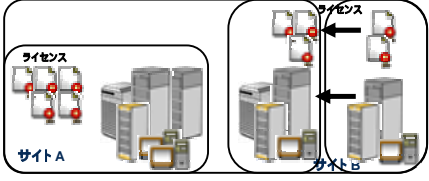
28 © Platform Computing UK 2004
UK-2004

Accelerating Intelligence™

Platform リソース リーシング モデル

クラスタを横断した共有により、ローカルクラスタは、リモートクラスタより資源を「借りる」ことができる

- 貸与された資源をローカル資源として管理
- ユーザから見た透過的な共有
- グリッドの拡張性を向上

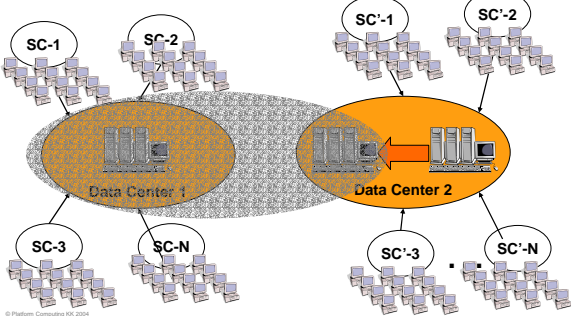


29 © Platform Computing UK 2004
UK-2004

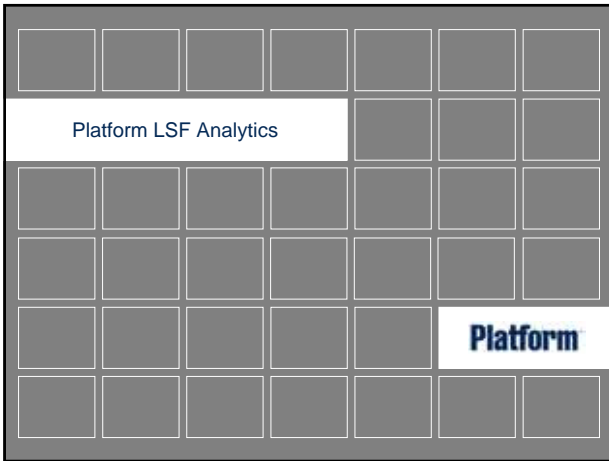
Accelerating Intelligence™

Platform Platform LSF MultiCluster適用例

大手半導体製造メーカー



30 © Platform Computing UK 2004
UK-2004



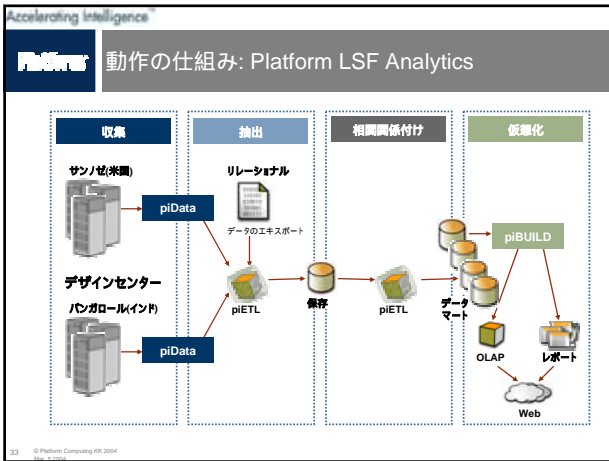
Accelerating Intelligence[®]

Platform Platform LSF Analytics

IT資産のIntelligenceツール

- 意思決定支援ソフトウェア
- Platform LSF と Macrovision FlexLMより、データを収集し定量化
- 設計プロジェクトを管理・計画する
 - 履歴レポート
 - 分析
 - 予測
- プロジェクト管理者による使用実績に基づきより良き判断を支援:
 - プロジェクトの計画立案
 - 資源の供給
 - 購買
 - 統合化
 - 資源の枯渇の予測と管理

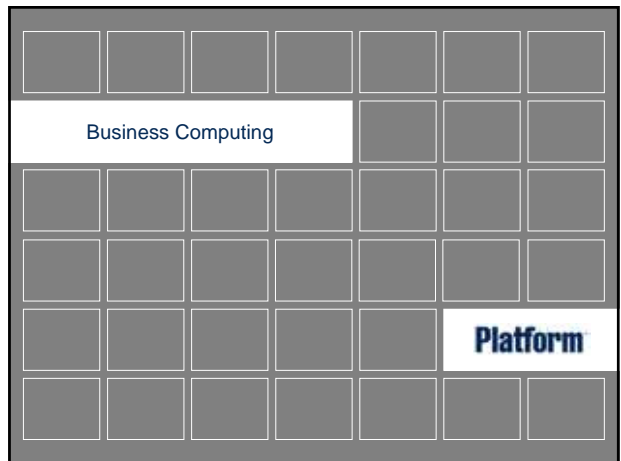
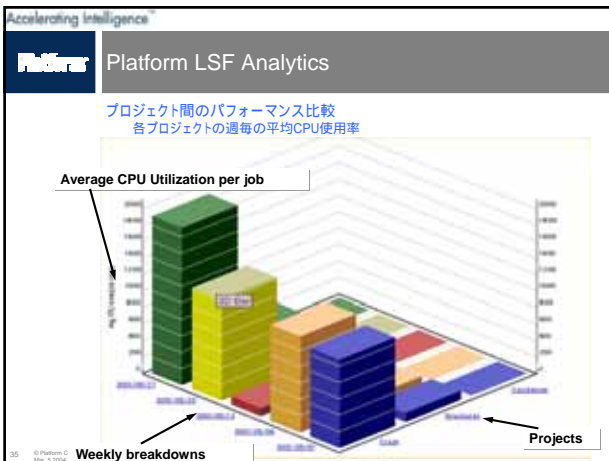
32 © Platform Computing Inc 2004
PLC-000004



Accelerating Intelligence[®]

Platform Platform LSF Analytics出力サンプル

34 © Platform Computing Inc 2004
PLC-000004



Accelerating Intelligence

Platform 金融業適用例1

JP Morgan Chase

JPMorganChase

概要

JPMC海外業務のサポート:

- 確定利付きデリバティブ
- 金融市場
- 外国為替
- 途上国市場
- 普通株

課題

HPC需要の高まり:

- デリバティブ価格付けモデルの複雑さ
- デリバティブポートフォリオの増加
- 新しい複雑な金融商品の導入
- 金融市場という新しい市場への参入

ソリューション

Platform Symphony

- JPMCの「コンピュータバックボーン」をサポート
- HPCリソースの幅広い共有をサポートする共有サービス

結果

より高いサービスレベルの達成

- 運用コストの削減
- アプリケーションおよびハードウェアの利用率の向上
- より生産的なソフトウェア開発

「当行内のさまざまな部門で働いている技術者からアイデアや設計に関するインプットを提供してもらった。このプロジェクトで最も誇らしく思うことのひとつは、これを実現するためにJPMorgan Chase銀行内に確固たる協力関係が築かれたことだ。」

Steve Neiman,
JPMC Investment Bank
ハイパフォーマンス コンピューティング担当副社長

37 © Platform Computing KK 2004

Accelerating Intelligence

Platform Platform Symphony

優先割り込みスケジューリング

イベント

J1/優先度(2) J3/優先度(3) J3/優先度(1) J3/完了 J1/完了 時間

J2/優先度(1) 発行

J1/完了

J2/完了

J1 (20 CPU)

J1 (10 CPU)

J3 (15 CPU)

J2 (10 CPU)

J2 (5 CPU)

J1 (20 CPU)

J2 (30 CPU)

J1 (10 CPU)

J2 (10 CPU)

38 © Platform Computing KK 2004

Accelerating Intelligence

Platform 金融業適用例2

ドイチェ モルガン グレンフィル銀行(DMG)

ドイツ イギリス アメリカ

日本の夜間を利用し、バッチ処理の実行

もし、全サーバがダウンした場合は、他の国のサーバで処理

Application例 Vendor Summit (Risk Management) Misys

計算サーバ

アプリケーション

39 © Platform Computing KK 2004

Accelerating Intelligence

Platform 金融業適用例2

ドイチェ モルガン グレンフィル銀行(DMG)

Enterprise Gridの構築

- 500~1,000のシナリオでのポートフォリオのリスク分析
- ドイツ、イギリス、アメリカ、日本、地帯を有効利用
- 低コスト
- メインフレームの追加購入無し/ソフトウェアライセンスのみ
- 耐障害性の向上/システムの安定性
- 既存サーバを有効利用し、1台のメインフレームから数十台のサーバへシステムを移行(同アーキテクチャであるため、稼働の自動化も実行可能)
- ジョブ実行の信頼性
- ジョブの再実行等の自動化/ジョブスケジューリングの自動化

40 © Platform Computing KK 2004

Accelerating Intelligence

Platform Summary

- Workload Management Systemはグリッドコンピューティングに必要な不可欠な機能
- Reliabilityの高いグリッドコンピューティングを構築するためには、優れたWorkload Management Systemが必要
- Platform Computingは、Workload Management技術で12年、1,500サイト以上の実績

41 © Platform Computing KK 2004

Accelerating Intelligence

Platform

ご清聴有難うございました。

Accelerating Intelligence™

プラットフォームコンピューティング

日本語ホームページ

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

42 © Platform Computing KK 2004