

Grid MP 4.0

最先端のセキュアグリッドソリューション

内 容:

- Grid MP 4.0 のご紹介
- MPIプログラム実行について

2003年8月19日

United Devices, Inc.

住商エレクトロニクス株式会社

1

United Devices 社 企業概要

- 創立 : 1999年、テキサス州 オースティン市
- セキュアなグリッドソリューションのマーケットリーダ
 - エンタープライズやグローバルグリッド
 - グリッドの提供、アプリケーションフレームワーク、セキュリティ
- 市場: ライフサイエンス、地球科学、政府官公庁、その他
 - 顧客: Novartis, GlaxoSmithKline, Sanofi, ADA...
 - パートナー: Gateway, Accelrys, Landmark, SSE
- 企業背景
 - マネジメントチーム - IBM, Dell, Intel, Microsoft
 - 技術チーム - [SETI@Home](#) や Distributed.net に携わる
 - 従業員、50名弱



2

Grid MP 4.0 の概要

Grid MP4.0 はグリッド・コンピューティングに求められる以下のような機能を実装したパッケージ・ソリューションです。

認証、セキュリティ、 計算機の登録・管理、
 ジョブの適切な計算環境への配布、実行、 最適なスケジューリング、
 グリッド環境のモニタリング、 基盤ソフトにWebサービスを適用

ネットワークに接続されているさまざまな既存の計算機を集約し、遊休のリソースを利用してグリッドのジョブを実行します。

計算機の所有者(利用者)にグリッドの実行負荷を感じさせることのない機能を実現しています。

Grid MP4.0 はWebサービスを基盤とし、ネットワークのセキュアなポート (80番、433番)をコミュニケーションに使用しますので組織のネットワークのポリシーを変更することなくファイアウォールやローカルIP環境下での運用が可能です。

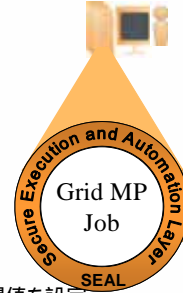
バーチャルな計算環境

- **ヘテロリソースへのバーチャルなインターフェイスを提供**
 - ネットワークに接続されている既存の計算リソース(クラスタ、ワークステーション、サーバ、デスクトップなど)を集約して管理
 - 単一のアクセスポイントから計算リソースのバーチャルな利用が可能
 - 既存リソースの遊休CPU時間を利用してグリッドのジョブを実行
 - 並列 MPI ジョブを実行可能
- **アプリケーション バイナリ エミュレーション (サポート予定)**
 - LinonWin: Linux バイナリを Windows環境で透過的に実行

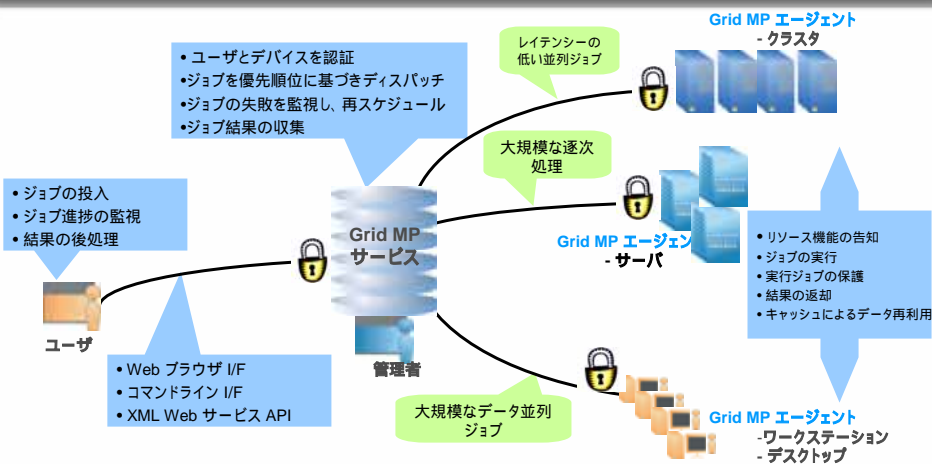


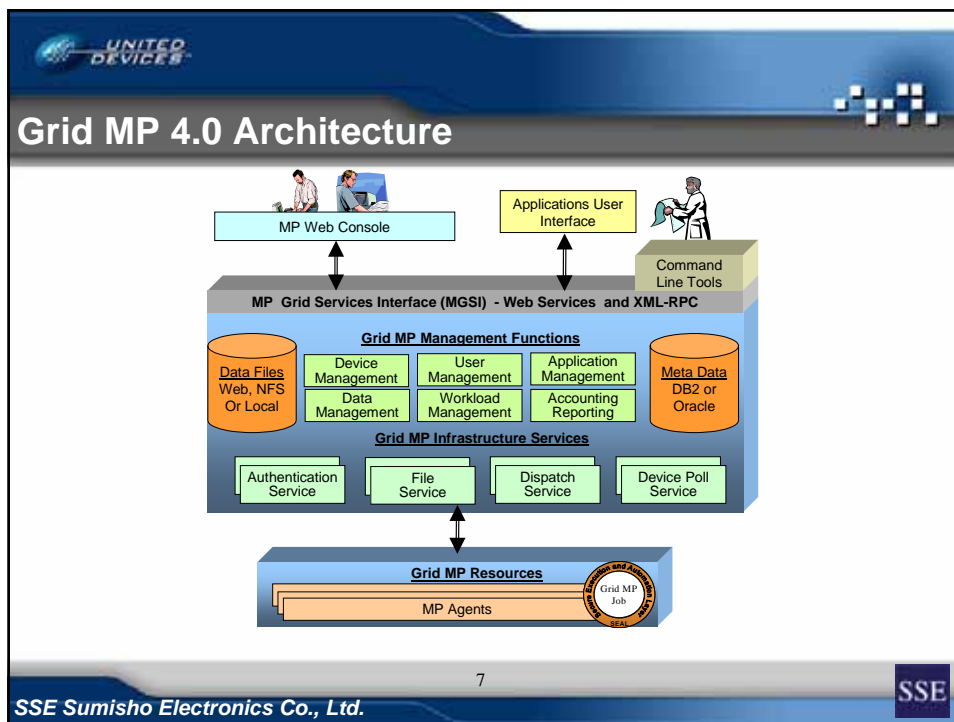
SEAL テクノロジ – “Secure Execution and Automation Layer”

- セキュアなランタイム「Sandbox」環境**
 - 計算リソースに投入されたジョブの実行を妨害行為から保護
- ジョブデータの自動暗号化と圧縮**
 - 計算リソースに保存されるジョブデータを改ざんから保護
 - Triple-DESアルゴリズムによる暗号化
 - トークンベースのセッション・キーの更新
- リソースでの負荷制御と自動監視**
 - リソース上で実行するグリッド・ジョブのCPU、メモリ、ディスクへの閾値を設定
 - Grid MP エージェントが閾値に基づきジョブによるリソースの使用状況を監視
- デジタル署名による証明**
 - アプリケーションはデジタル署名化され、実行前に検証される



Grid MP[®] 動作モデル

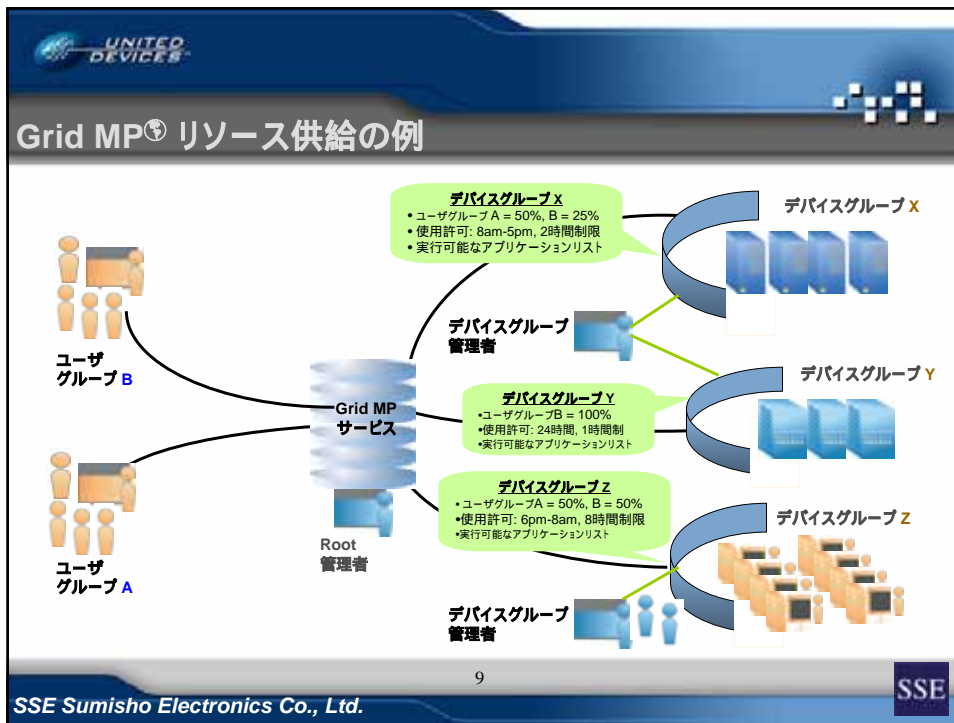




分散リソースとワークロードの管理

- リソースとユーザ**
 - リソースやユーザはそれぞれの属性に応じてグループを設定
 - グループ化されたリソースやユーザ毎にグリッドでの利用の閾値、優先順位、権限などITポリシー設定可能
- リソース管理**
 - ジョブの失敗やアクシデント時に再スケジューリングを制御するタイムアウト・パラメータ
 - 日毎・週毎のタイマーベースによるリソースのグリッドへの提供が可能
 - アプリケーションリストに基づいたリソース上のグリッド・ジョブ実行制御
- ユーザの許可**
 - ユーザや管理者にロールに基づくアクセス管理と権限を付与

The slide is presented by **SSE Sumisho Electronics Co., Ltd.**



UNITED DEVICES

グリッド アプリケーション フレームワーク

- 集約化されたアプリケーションとデータ管理**
 - アプリケーションとデータセットはGrid MPサービスに登録、暗号化し利用管理
 - 登録されたアプリケーションとデータセットはグリッドのポリシーに基づき運用
- 高い柔軟性を持つグリッド インターフェイス**
 - ジョブの投入と管理は Web ブラウザやコマンドラインより実行
 - カスタム GUIやワークフローを開発するための Web サービス I/F
- 並列アプリケーション サポート**
 - MPI によるアプリケーションは MPIRUN により修正することなく実行可能
 - パイプラインなどの複雑なワークフローを持つシングル・マルチステップ並列ジョブ
- グリッド ソフトウェア開発キット (SDK)**
 - ドキュメント、ツール、ライブラリの豊富なセットにより、アプリケーション開発を迅速化

10

SSE


SSE Sumisho Electronics Co., Ltd.

ライフサイエンス アプリケーション

- ゲノミクス
 - ホモロジー、シーケンス比較、系統学
 - 事例: HMMER、BLAST、RepeatMASKER など
- プロテオミクス
 - たんぱく質量み込み、たんぱく質ホモロジー
- ハイスループット スクリーニング
 - ドッキング、バーチャル スクリーニング
 - 事例: GOLD、Ligandfit
- 臨床試験
 - 有望ポートフォリオの確率論的シミュレーション
 - 事例: Pharsight
- データマイニング
 - OmniViz などのデータ可視化アプリケーション

11

エンタープライズ事例: Novartis

- **事業目標**
 - 有望案件の特定を、10倍以下のコストで10倍迅速化する
- **UD グリッドソリューションによる対応**

 - **拡張性:** 3,000 台の Windows デスクトップにて早期にの運用開始が可能に
 - **統合性:** シームレスなインテグレーション、バンド幅の消費は最小(2%以内)
 - **セキュリティ、不侵入性:** デスクトップ PC のユーザには、パフォーマンス上の影響なし。すべての End-to-End のセキュリティテストをクリア
 - **計算容量:** 導入後7日間にて、トータル 3.18 年に相当する計算時間を集約
 - **競合比較:** 他社グリッド製品との比較の結果、Grid MP が選択される
- **結果**
 - 大規模なスクリーンジョブやその他のジョブを日常的に実行
 - スクリーニングの処理時間を「月」から「日」に削減
 - すべてのデスクトップに展開を計画: 他のアプリケーションもグリッド化予定

12

UNITED DEVICES

Example: バイオテロリズム対策

- Virtual High-Throughput Screening
 - 炭疽菌の致死因子プロジェクト例
 - UDのGMPで初回のパスが5日で完了
 - Dr. Anna Johnson-Winegar (化学・生物兵器防衛担当国防長官補) と R. Stephen Day (核拡散防止及び化学・生物兵器防衛プログラム・ディレクター) がこの結果を受け取った
- DoDの要請で天然痘にフォーカスした類似のスクリーニングを実施
- 天然痘ウイルスの専門家たちが参画
 - USAMRIID, DoD 及び学術機関
 - 最初の72時間で100,000件の結果が得られた

“これらの結果は画期的なものでした。もし、従来の方法で行っていたとしたら、4週間ではなく数年を必要としていたでしょう。”

Dr. Graham Richards, Chairman of the Chemistry Department at Oxford University and the Director of the Centre for Computational Drug Design. (炭疽菌のスクリーニングについてのコメント)

13

SSE Sumisho Electronics Co., Ltd.

UNITED DEVICES

その他のアプリケーション

- 政府・官公庁
 - 暗号解読
 - 核、武器シミュレーション
 - ロケット噴出炎分析
 - 地理空間リモートセンシング
 - 情報収集
 - 画像分析
- エンターテイメント・メディア
 - 図書エンコーディング
 - グラフィックス レンダリング
 - 画像圧縮
- コンピュータ製造 (EDA)
 - 検証シミュレーション
 - シンセシス
- 工業製造業
 - 有限要素法
 - 計算流体力学
 - 衝突シミュレーション テスト

14

SSE Sumisho Electronics Co., Ltd.



Grid MP 4.0

Grid MP WEBサービス環境下の MPIプログラム実行例のご紹介

15



特徴

既存の計算資源と新規の計算資源の統合利用やヘテロな
仮想クラスタシステムの実現

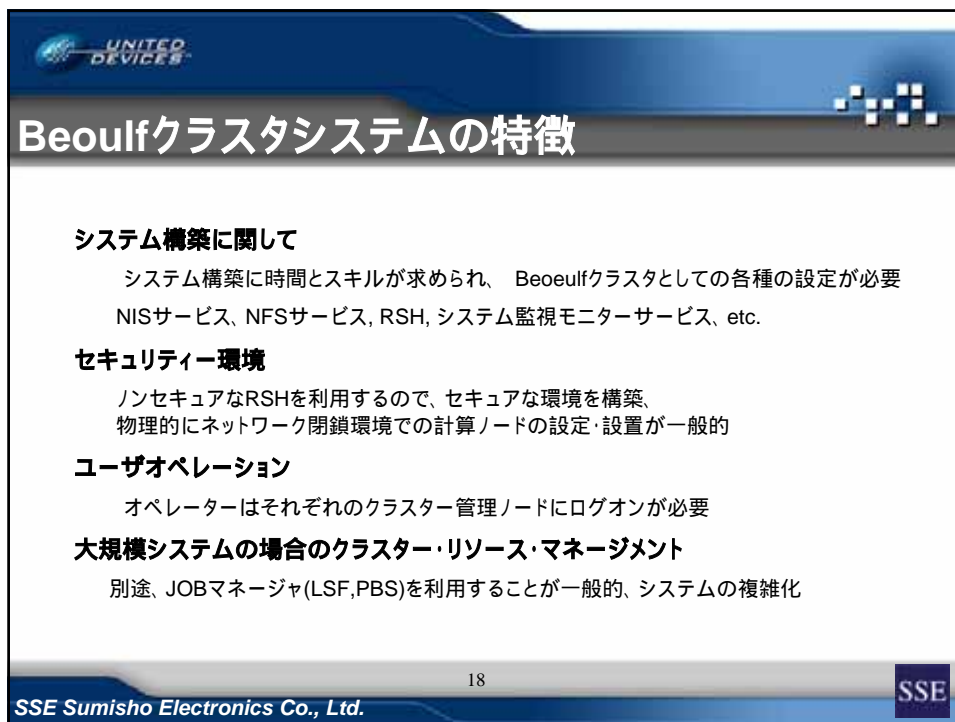
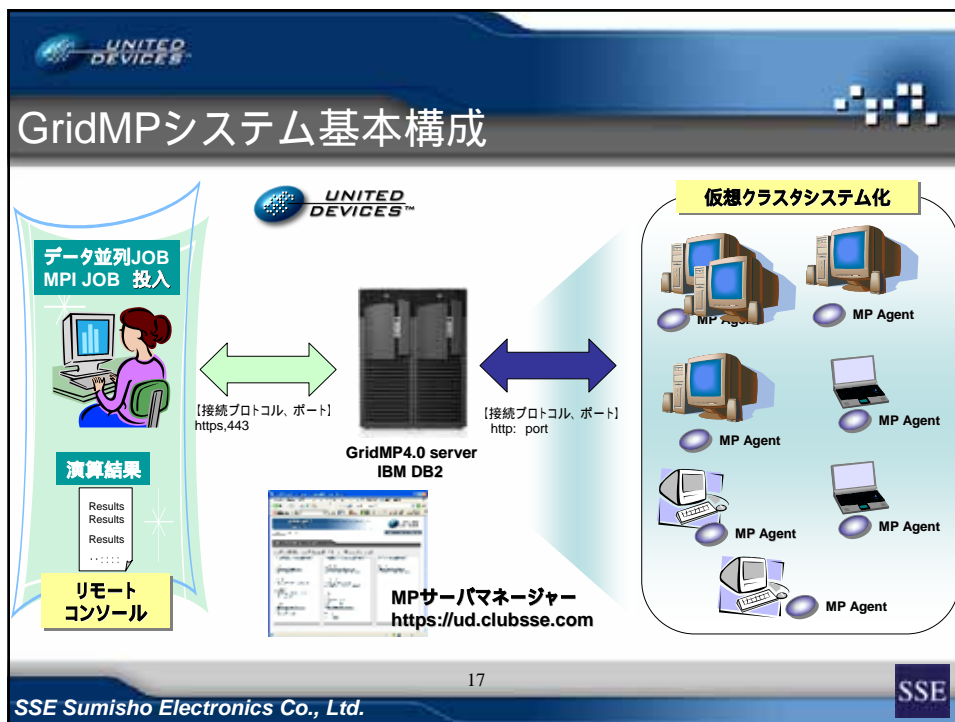
シンプルなユーザオペレーション

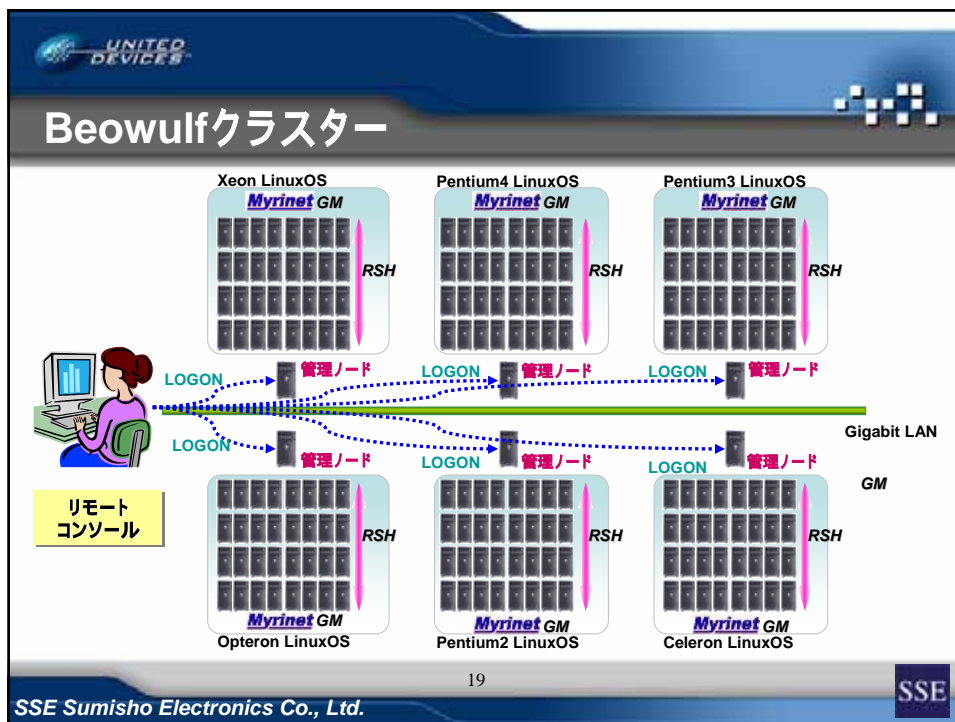
MPIプログラム実行環境の容易なインストレーション
Beowulfクラスタの構築方法と比べて、GridMPIはLinuxOSに
MP Agentをインストールだけ

管理者やユーザの負荷軽減によるTCOの削減効果

グリッドのフレームワークにおけるセキュアな
クラスタ・コンピューティング実行環境の実現

16





GridMPによるクラスタ・システムの特長

システム構築に関して

- 既存のBeowulfクラスタシステムに追加導入が可能
- 新規システムではOSをインストールし、各ノードにMP Agentをインストールするだけ

セキュリティー環境

- RSHを利用しないセキュアな通信を利用しているため、セキュリティー確保の為に物理的にネットワーク閉鎖環境の計算ノードの設定・設置を行なう必要がありません
- HTTPS 433, 80 ポートのみ利用
- 物理的にネットワーク閉鎖環境の計算ノードの設置環境でも使用可能

ハイパフォーマンスコンピューティング

- MPIのみならずデータ並列としての利用も可能
- MyrinetTCP/IP機能と併用したハイパフォーマンスネットワークの実現も

SSE Sumisho Electronics Co., Ltd.

UNITED DEVICES

GridMPによるクラスタ・システムの特長(2)

ユーザオペレーション環境

オペレータはGridMPサーバにシングルサインオン、そして全仮想クラスタ資源を即座に利用可能

インハウスMPIプログラム実行の際には、ソースプログラムのリコンパイル(gcc3.0)が必要ただし、リコンパイルの方法は通常のMPICHのコンパイル方法と同様です

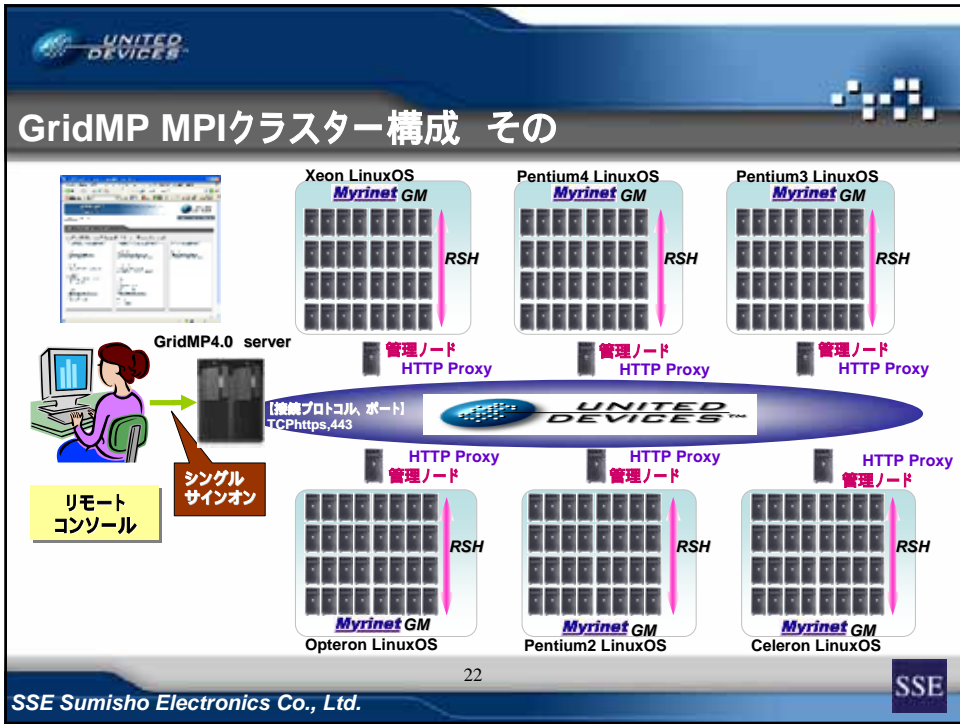
全システムリソースマネージメント

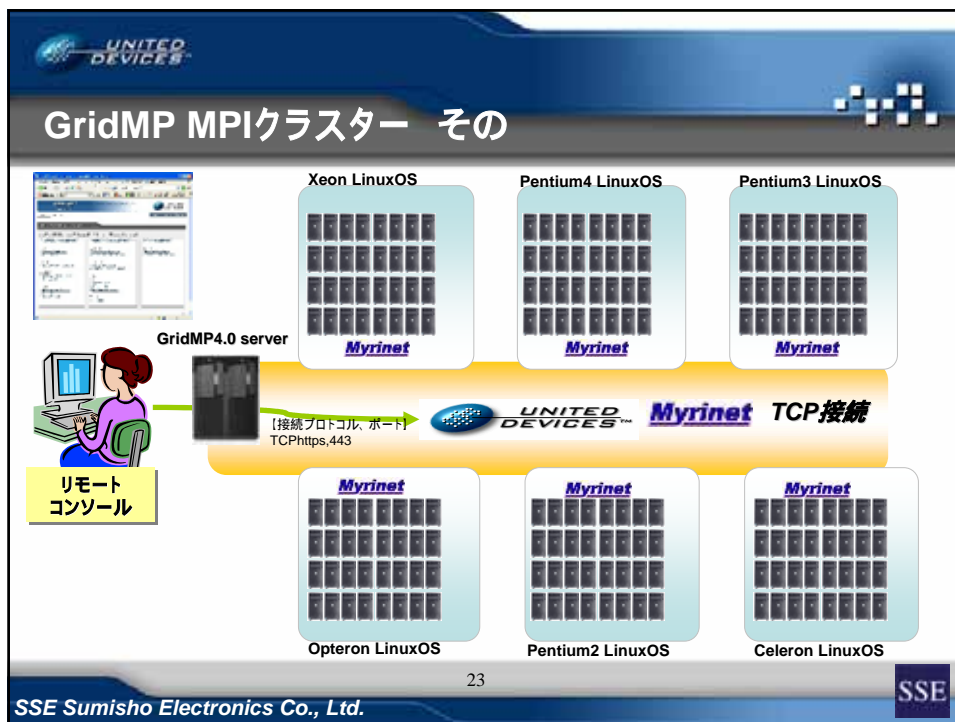
GridMPコンソール(WEB)によるシステム管理・監視ができます

GridMPに接続されているコンピュータのステータス(稼働率など)も即座に確認できます

21

SSE Sumisho Electronics Co., Ltd. SSE





GridMP MPIプログラム実行手順

テストプログラム

姫野ベンチ

ソースコード: C + MPI, static allocate version

URL: <http://w3cic.riken.go.jp/HPC/HimenoBMT/program2.htm>

MPI通信ライブラリ

Untied Devices提供 MPICH-1.2.5

計算環境

仮想クラスタ環境

The slide includes the "UNITED DEVICES" logo at the top left, the number "24" at the bottom center, and the "SSE Sumisho Electronics Co., Ltd." logo at the bottom right.

GridMP MPIプログラム実行手順

コンパイル方法

```
# mpicc himenobmtxps.c -o M112c
```

テストプログラム実行

```
# ./ud_mpirun.pl -np 2 M112c
```

テストプログラム実行結果取得

```
# ./mpresult.pl 63 ← JOB ID
```

25

Grid MP MPI動作イメージ

26

MPI JOB投入

MPI JOB Dispatch



コンソール

```
bash-2.05a$ ./ud_mpirun.pl -np 2 M112c
Checking for MPI Application.
MPI Application already exists.
Creating new MPI Program.
Creating new ProgramVersion.
Creating new ProgramModule.
Targeted 0 DeviceGroups.
Creating new ProgramModuleVersion.
Uploaded /tmp/A09Fh9pEvD to https://172.28.52.232:28443/mgsi/filesvr.fcgi
Uploaded /tmp/NPQ7WnBnma to https://172.28.52.232:28443/mgsi/filesvr.fcgi
Uploaded /tmp/PnsC3lprM5 to https://172.28.52.232:28443/mgsi/filesvr.fcgi
Created Job 63
```

GridMP MPI実行コマンド

```
# udmpirun -np <ノード数> (-dgfile <デバイスグループ名>) <プログラム名>
```

27

演算結果 取得

MPI JOB Retrieve



コンソール

```
bash-2.05a$ ./mpresult.pl 63
processing job 63...
result saved in result-003203-job-000063.tar
result saved in result-003204-job-000063.tar
no errors for Job 63, JobStep 9E38E3FD-08FB-400A-BC01-
2F89CE98C86E
```

MPI-JOB演算結果取得コマンド

```
# mpresult.pl (JOB ID)
```

28

演算結果 表示

```

Sequential version array size
mimax = 129 mjmax = 129 mkmax = 257
Parallel version array size
mimax = 129 mjmax = 129 mkmax = 131
imax = 128 jmax = 128 kmax = 129
I-decomp = 1 J-decomp = 1 K-decomp = 2
Start rehearsal measurement process.
Measure the performance in 3 times.

MFLOPS: 52.182520 time(s): 7.882245 1.667103e-03

Now, start the actual measurement process.
The loop will be executed in 22 times
This will take about one minute.

cpu : 57.403322 sec.
Loop executed for 22 times
Gosa : 1.548856e-03
MFLOPS measured : 52.545966
Score based on Pentium III 600MHz : 0.640804
    
```

29



www.ud.com

SSE

United Devices 社 国内正規代理店
 〒101-8453 東京都千代田区神田錦町3丁目11番地
 住商エレクトロニクス カンパニー
 HPC ソリューション部
 PHONE: 03-5217-5142
<http://www.sse.co.jp/UD/>

30