

分散計算機環境における ジョブ実行支援システムに関する研究

電気電子工学専攻 川田研究室

MT043225 藤生英昭

発表内容: 修士論文目次

第1章	序論
第2章	ジョブ実行支援システム
第3章	動的システム管理とシステム構築支援
第4章	適用例
第5章	まとめ

研究の背景

- 分散環境での計算機利用
 - 資源の状態把握や有効活用が困難
- グリッドシステム
 - 一般ユーザにとって、インストール・セキュリティ設定が非常に難しい
 - 特別なミドルウェアが必要な場合が多い
- ロバスト性
 - 分散システムにおいて、システム管理とデータストレージなどはロバスト性が要求される

研究の目的

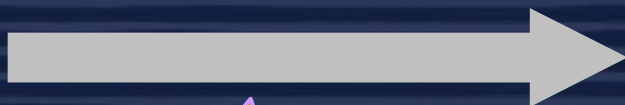
- ・ ジョブ実行支援システム
 - ローカルな分散計算機環境で構築
 - 動的かつロバストなシステム管理
 - MPIジョブをサポート
 - ロバストなデータストレージシステム
 - ・ 結果データの2重化
 - システム構築支援
 - ・ ユーザによる容易な構築

分散環境の問題

- ・ 利用可能な計算機の把握
 - ジョブ投入可能サーバ・リソース要件
- ・ ジョブ設定
 - 利用する全サーバで実行前処理などが必要



利用者



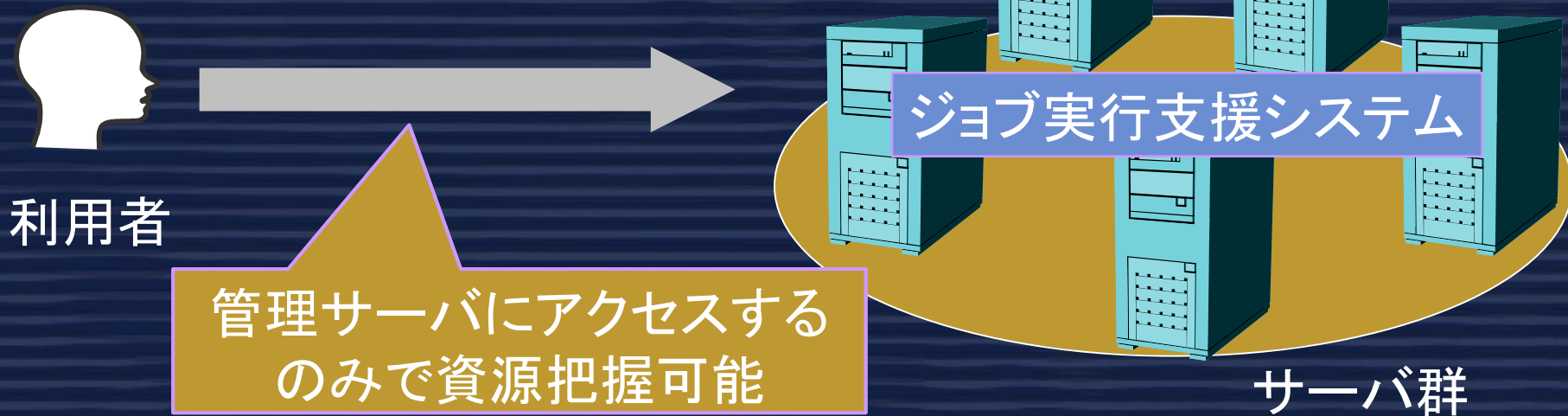
利用可能でリソース条件が
合うサーバを探す必要



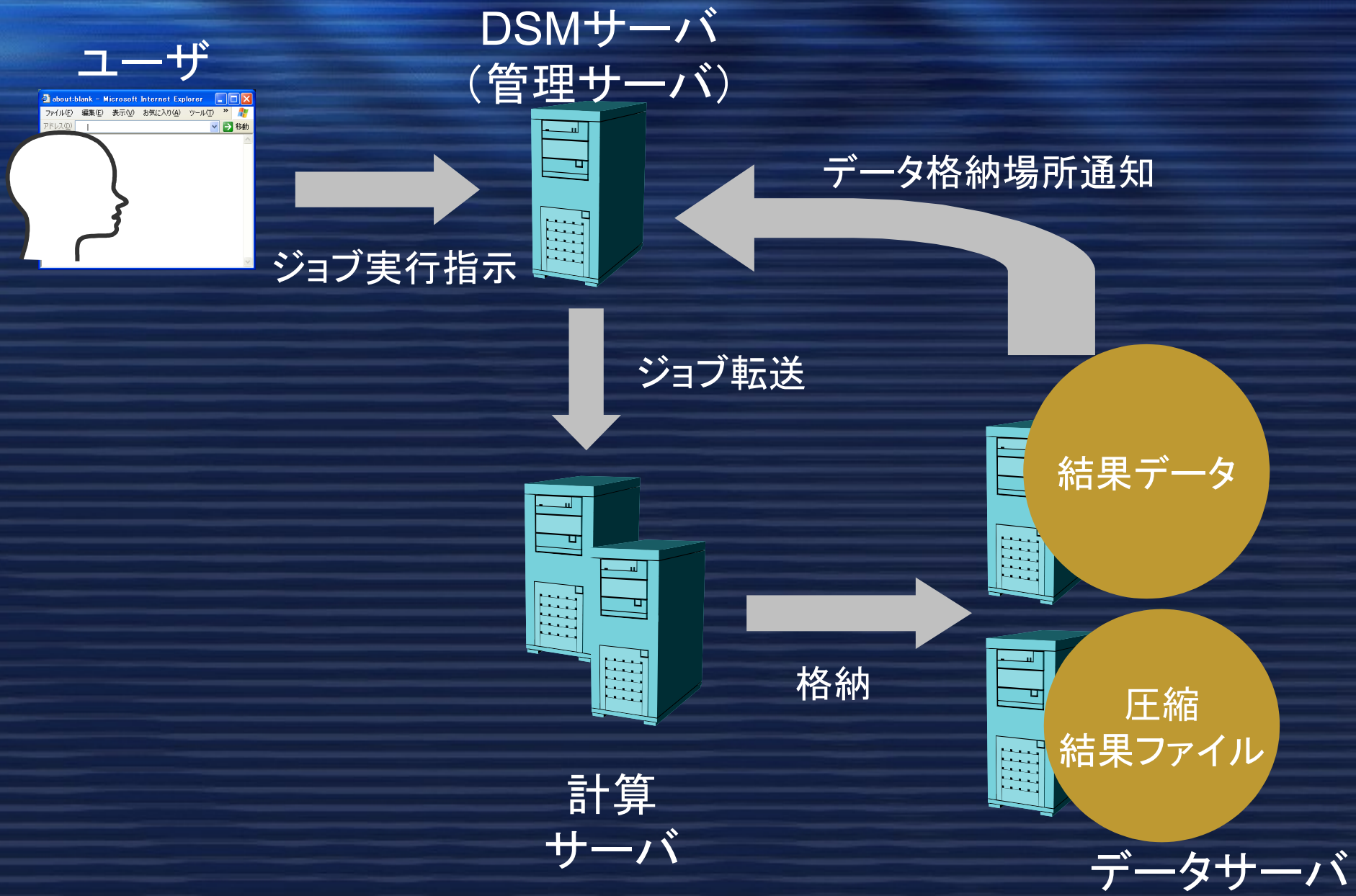
サーバ群

本システムの導入

- 分散資源を把握
 - Webページにアクセスするのみで把握可能
- ジョブ設定
 - 各サーバへ投入するジョブの設定をWebベースで可能

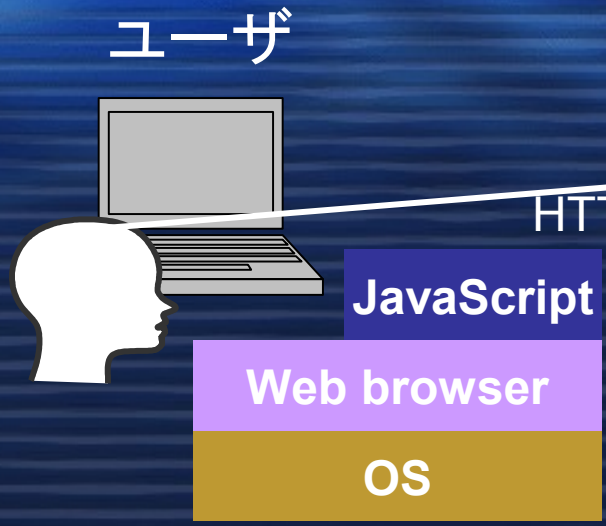


ジョブ実行概念図

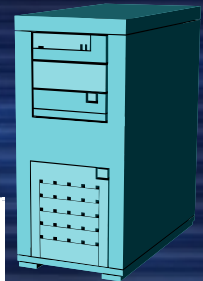
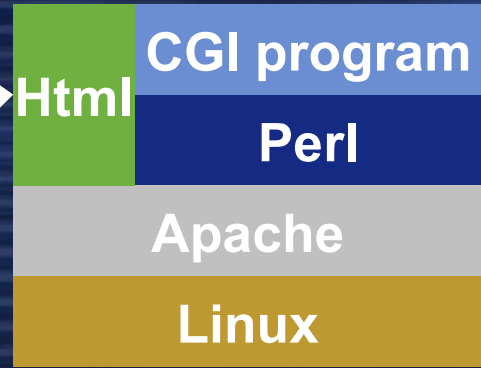


システム環境

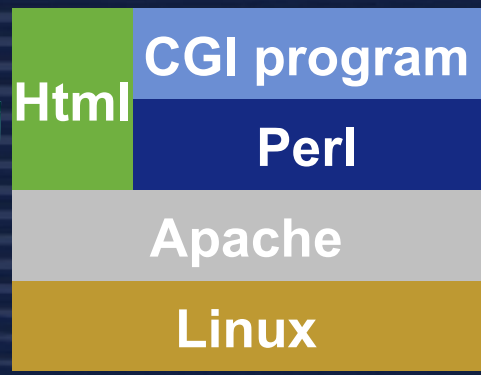
DSMサーバ



HTTPS or HTTP



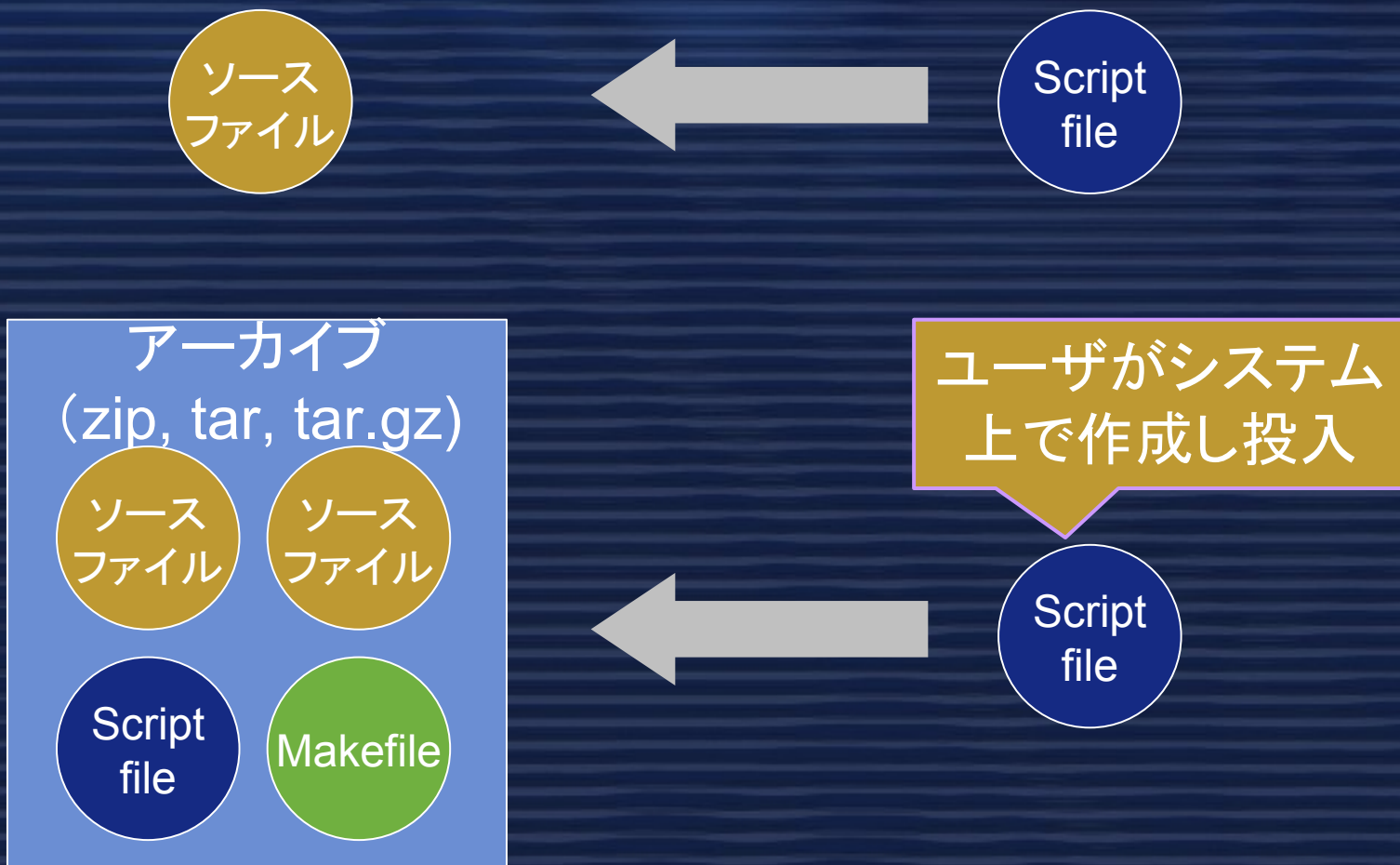
計算サーバ群



HTTPS or HTTP

ジョブ設定(1)

- ジョブ投入ファイル



ジョブ設定(1)

ジョブ投入ファイル

ソース
ファイル

アーカイブ
(zip, tar, tar.gz)

ソース
ファイル

ソース
ファイル

Script
file

Makefile

The screenshot shows the JES System web interface in Mozilla Firefox. The page title is 'Job実行支援システム - Mozilla Firefox'. The main content area displays configuration for two jobs: 'test.c' and 'NewPICcode.tar.gz'.

実行前処理選択 (Pre-execution selection): A callout points to the radio button options for each job. For 'test.c', the selected option is 'コンパイル or makeコマンド' (Compile or make command). For 'NewPICcode.tar.gz', the selected option is 'アーカイブ内のシェルスクリプト名' (Shell script name in archive).

実行コマンド (Execution command): A callout points to the '実行パス' (Execution path) field. For 'test.c', the path is '/a.out'. For 'NewPICcode.tar.gz', the path is '/a.out'.

Other visible fields include 'コメント' (Comment) and '新規スクリプト作成' (Create new script) links.

At the bottom of the page, it says '対象サーバの選択' (Select target server) and '完了' (Completed).

ジョブ設定(1)

ジョブ投入ファイル

ソース
ファイル

アーカイブ
(zip, tar, tar.gz)

ソース
ファイル

ソース
ファイル

Script
file

Makefile

スクリプト記述可能

実行前

test2.c用スクリプト

```
#!/bin/sh
mkdir DataStrage/
gcc test2.c
```

保存して開じる

完了

ジョブ設定(2)

Job実行支援システム - Mozilla Firefox

ファイル(F) 編集(E) 表示(V) 移動(G) ブックマーク(B) ツール(T) ヘルプ(H)

https://skk-s1/~jas/cgi-bin/management/ 移動

Logout
Top

JES-System

ジョブ実行支援システム

投入対象サーバの選択

対象サーバの選択

No.	check	Host name	CPU	All memory	Use memory	Free memory	Load average	HDD free capacity
1	<input checked="" type="checkbox"/>	skk-c5	Pentium4--2.387GHz	1035.468MB	444.404MB	591.064MB	0.00	175GB
2	<input checked="" type="checkbox"/>	itsc-9	Pentium4--2.386GHz	1035.016MB	1013.312MB	21.704MB	0.00	179GB
3	<input checked="" type="checkbox"/>	skk-s1	Pentium4--2.387GHz	1035.016MB	1000.188MB	34.828MB	0.05	163GB
4	<input checked="" type="checkbox"/>	gridlab0	Celeron--2.02GHz	483.208MB	394.336MB	88.872MB	0.07	3.9GB
5	<input checked="" type="checkbox"/>	waste01	PentiumIII--999MHz	1035.528MB	963.576MB	71.952MB	0.00	215GB
6	<input type="checkbox"/>	gridlab1	Not available					
7	<input type="checkbox"/>	gridlab2	Not available					

結果データ保存方法の選択

保存先の設定

システムに委任
 保存先を指定

skk-c5 デレクトリ: 空欄可 : 絶対パス
 デレクトリを指定する場合、パーミッションを777にしておく必要があります

設定し実行

完了

skk-s1

MPIジョブ設定

Job実行支援システム - Mozilla Firefox

http://k2/fuji/cgi-bin/management/index.cgi

Logout IES System

コンパイル or makeコマンド:

 宿フォルダ内のシェルスクリプト名:

 用に新規シェルスクリプトを作成: [新規スクリプト作成](#)

 実行パス:

 コメント:

サーバ管理

Content

- ジョブ実行へ
- ジョブデータ管理
- サーバー一覧
- パスワード変更

管理者メニュー

- 動的管理サーバ機能
- 構築支援機能
- サーバ登録と削除
- ユーザ登録
- システムプログラム管理

Other

- help file
- KawataLab

結果データ出力サーバを選択

ID	Master	Slave	Data Output Server	Host name		All memory	Use memory	Free memory	Load average	HDD free capacity
gridGroup										
	<input type="radio"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	grid3	Xeon--2.392GHz	2064.4MB	820.06MB	1244.34MB	1.00	5.8GB
	<input type="radio"/>	<input type="checkbox"/>	<input type="checkbox"/>	grid4	Xeon--2.392GHz	2055.46MB	1697.76MB	357.7MB	0.02	7.0GB
	<input checked="" type="radio"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	grid1	Xeon--2.392GHz	2064.4MB	1480.556MB	583.844MB	0.00	4.6GB
	<input type="radio"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	grid2	Xeon--2.392GHz	2064.4MB	1440.016MB	624.384MB	0.00	3.4GB
kGroup										
	<input type="radio"/>	<input type="checkbox"/>	<input type="checkbox"/>	k0	Xeon--3.19GHz	4091.58MB	4081.316MB	10.264MB	1.00	60GB
	<input type="radio"/>	<input type="checkbox"/>	<input type="checkbox"/>	k1	Xeon--3.189GHz	4090.68MB	2588.036MB	1502.644MB	0.06	20GB
	<input type="radio"/>	<input type="checkbox"/>	<input type="checkbox"/>	k2	Xeon--3.19GHz	4091.58MB	470.624MB	3620.956MB	0.00	73GB

完了

ジョブ管理

Job実行支援システム - Mozilla Firefox

ファイル(F) 編集(E) 表示(V) 移動(G) ブックマーク(B) ツール(T) ヘルプ(H)

http://k2/~fuju/cgi-bin/management/index.cgi

Logout
Top

JES-System

ジョブ実行支援システム

Project list

情報更新

Project	Project server	Main result	Sub result	Date	Delete
2005.12.07.21.20	k2	k2	k0	2005/12/07/21:21	delete
2005.12.07.21.35	k2	k2	k0	2005/12/07/21:35	delete
2005.12.07.21.46	k2	k2, k2	k0, k0	2005/12/07/21:46	delete
2005.12.08.16.03	k2	f1	k2	2005/12/08/16:04	delete
2005.12.08.16.06	k2	f1	k2	2005/12/08/16:06	delete
2005.12.09.16.23	k2			2005/12/09/16:24	delete

Content

- ジョブ実行へ
- ジョブデータ管理
- サーバー一覧
- パスワード変更

管理者メニュー

- 動的管理サーバ機能
- 構築支援機能
- サーバ登録と削除
- ユーザ登録
- システムプログラム管理

詳細参照リンク

プロジェクト削除リンク

ジョブ管理

ジョブ実行中の場合

Job実行支援システム - Mozilla Firefox

ファイル(F) 編集(E) 表示(V) 移動(G) ブックマーク(B) ツール(T) ヘルプ(H)

http://k2/fuju/cgi-bin/management/index.cgi

Logout
Top

JES-System

ジョブ実行支援システム

情報更新

ファイル名:	MPIPaxial.tar.gz
プロジェクト保存先:	2005.12.09.16.23ディレクトリ
結果データ保存先:	
実行前処理:	run.sh
実行コマンド:	mpirun -np 3 ./a.out
ジョブ投入対象サーバ:	grid1(master), grid4, grid2
結果保存方法:	システム依存
コメント:	MPIジョブ 3CPU
実行サーバ:	grid1
	8929
	MPIジョブモード
	XML参照
ジョブ実行ディレクトリを見る:	クリック
ジョブのキル:	kill Job

KawataLab

完了

ジョブ実行ディレクトリ参照

ジョブキル

ジョブ管理

ジョブ終了後の場合

ジョブキ

Job実行支援システム - Mozilla Firefox

ファイル(F) 編集(E) 表示(V) 移動(G) ブックマーク(B) ツール(T) ヘルプ(H)

http://k2/~fuju/cgi-bin/management/index.cgi

Logout
Top

JES-System

ジョブ実行支援システム

情報更新

Content

ジョブ
ジョブ
理
サーバー監
パスワード変更

管理者メ
ニュー

動的管理サー
バ機能
構築支援機能
サーバ登録と
削除
ユーザ登録
システムプログ
ラム管理

Other

help file
KawataLab

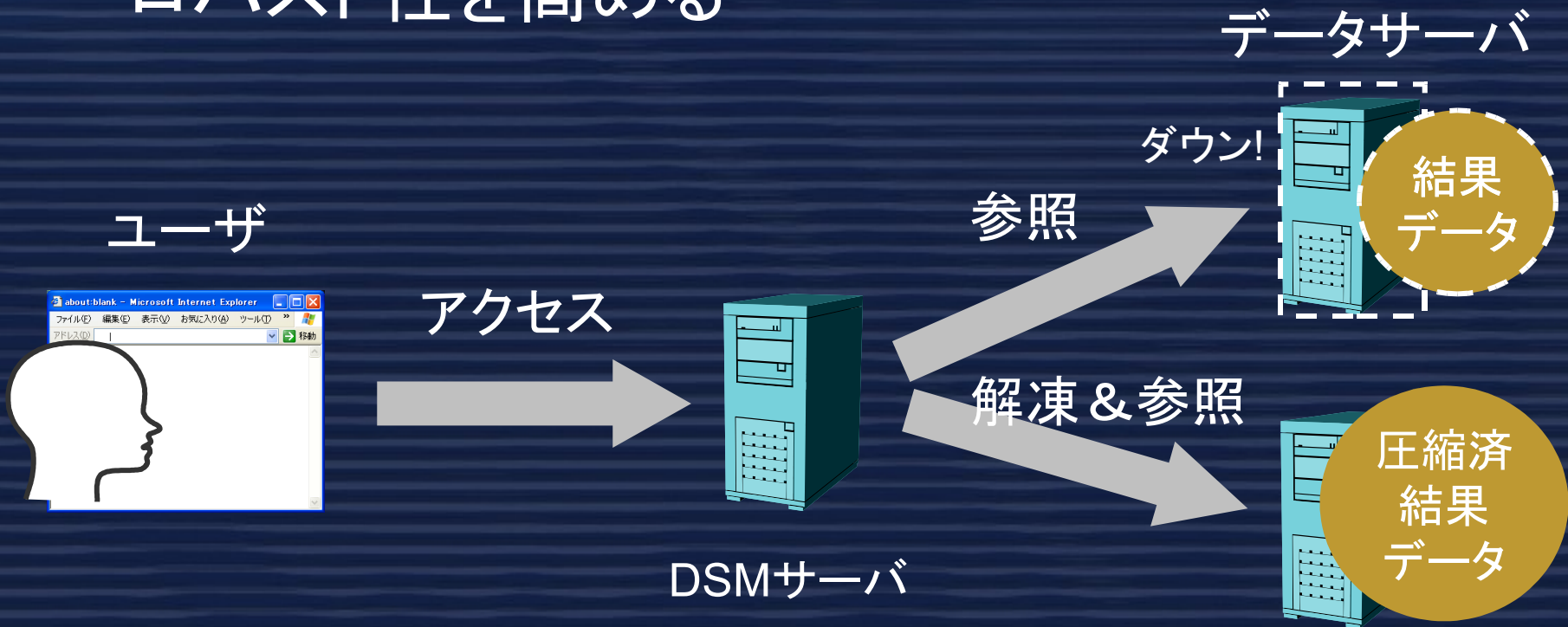
結果データ保存先リンク

プロジェクト保存先:	MPIPaxial.tar.gz 2005.12.09.16.23ディレクトリ
結果データ保存先:	f1(main) k2(sub)
実行前処理:	run.sh
実行コマンド:	mpirun -np 3 ./a.out
ジョブ投入対象サーバ:	grid1(master), grid4, grid2
結果保存方法:	システム依存
コメント:	MPIジョブ 3CPU
実行サーバ:	grid1
ジョブPID:	8929
MPI:	MPIジョブモード
ジョブに関するXMLを参照:	XML参照
ジョブRetry:	<input type="button" value="リトライ"/> <input type="button" value="編集後リトライ"/>

完了

結果データ2重化

- ・ DSMサーバを通じて結果データを取得
 - データの場所を気にする必要がない
- ・ 結果データの2重化は、データ取得においてロバスト性を高める



システム機能

- ・ 動的システム管理
- ・ システム構築支援
- ・ ジョブスケジューリング
- ・ ジョブエラー通知機能
 - －コンパイルエラーなど
- ・ ジョブキル機能
- ・ ジョブリトライ機能
 - －内容をWeb上で編集
- ・ システムアップデート機能
 - －システム更新ファイルを全サーバへ波及

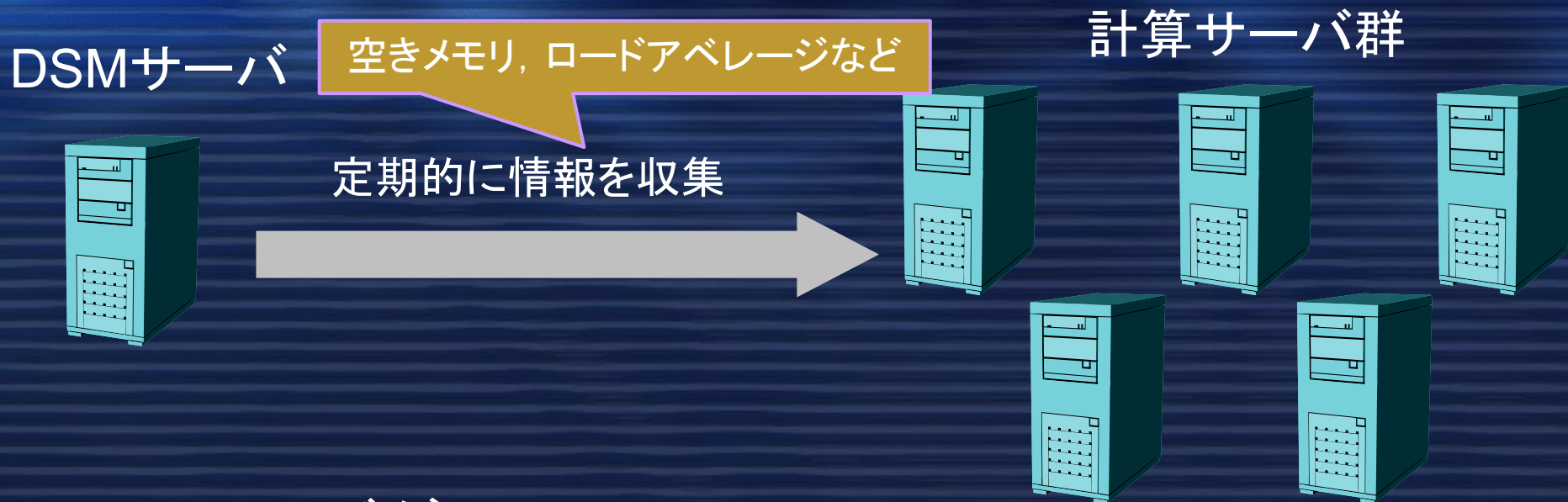
動的システム管理

- DSMサーバはシステム全体を管理
 - DAサーバによって管理は2重化
 - DSMサーバでトラブルが発生した場合, DAサーバがシステム管理を再構築
- 2サーバは自立的かつ動的にその機能を移動する



管理サーバの2重化は
システムに対しロバスト性をもたらす

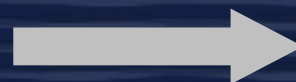
動的システムコントロール(1)



DSMサーバが

- ビジー状態
- ジョブから開放されたよりハイパフォーマンスなサーバを見つけた時

現在のDSMサーバ



New DSMサーバ

動的システムコントロール(2)

DSMサーバ



仮変更リクエスト

決定リクエスト

確認リクエスト

管理機能
停止

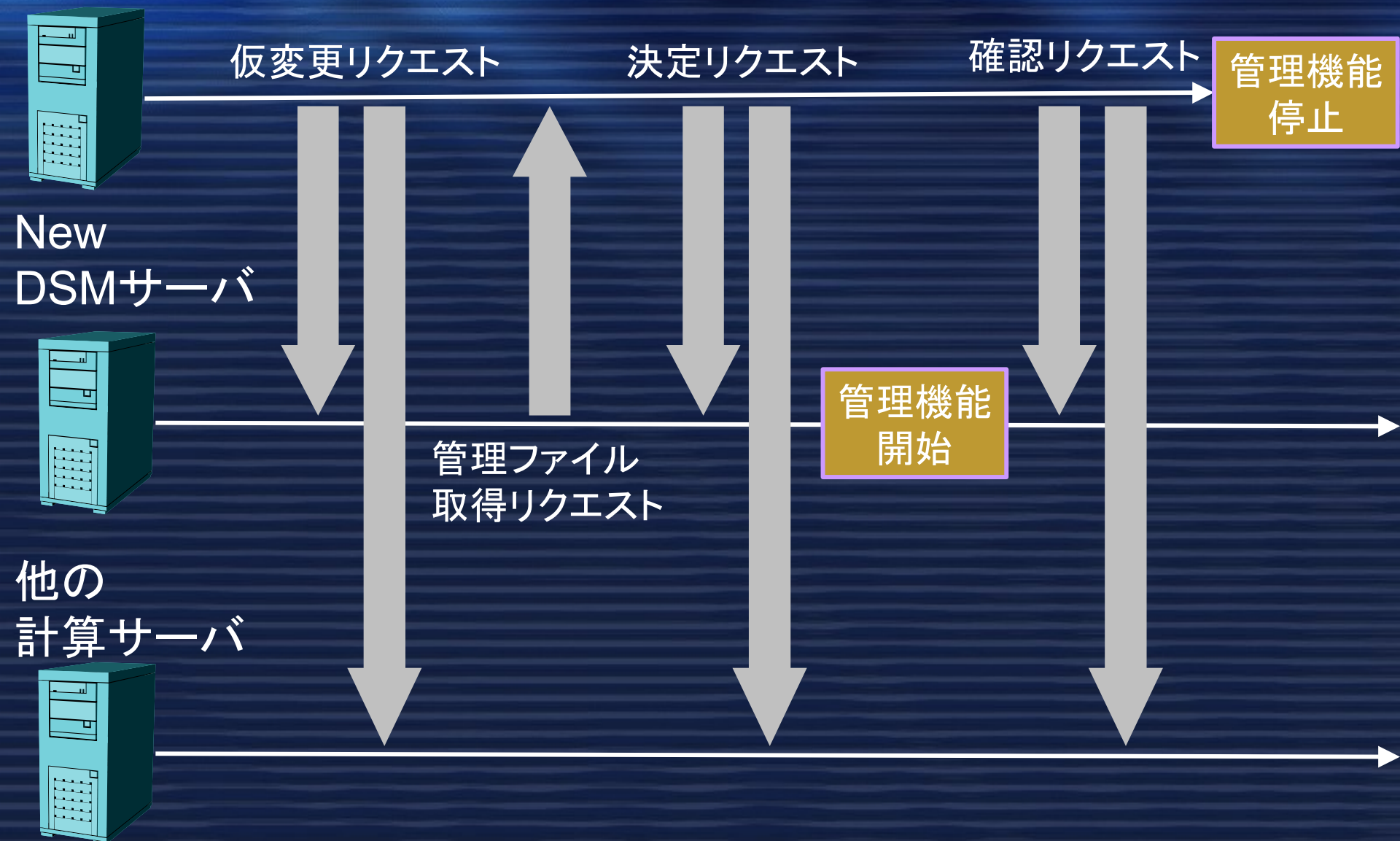
New
DSMサーバ



管理ファイル
取得リクエスト

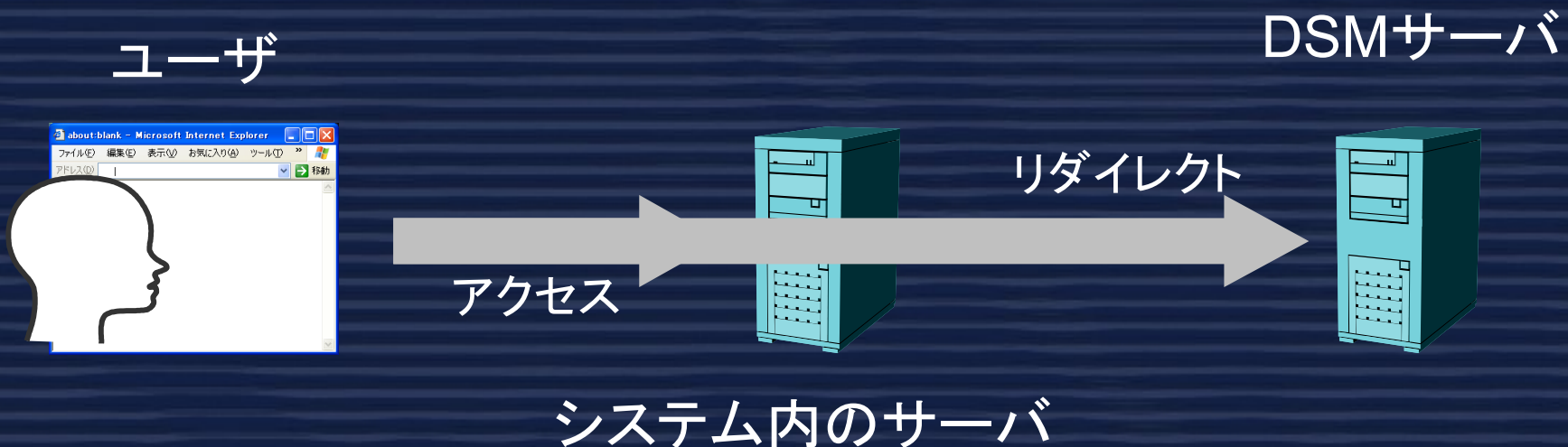
管理機能
開始

他の
計算サーバ



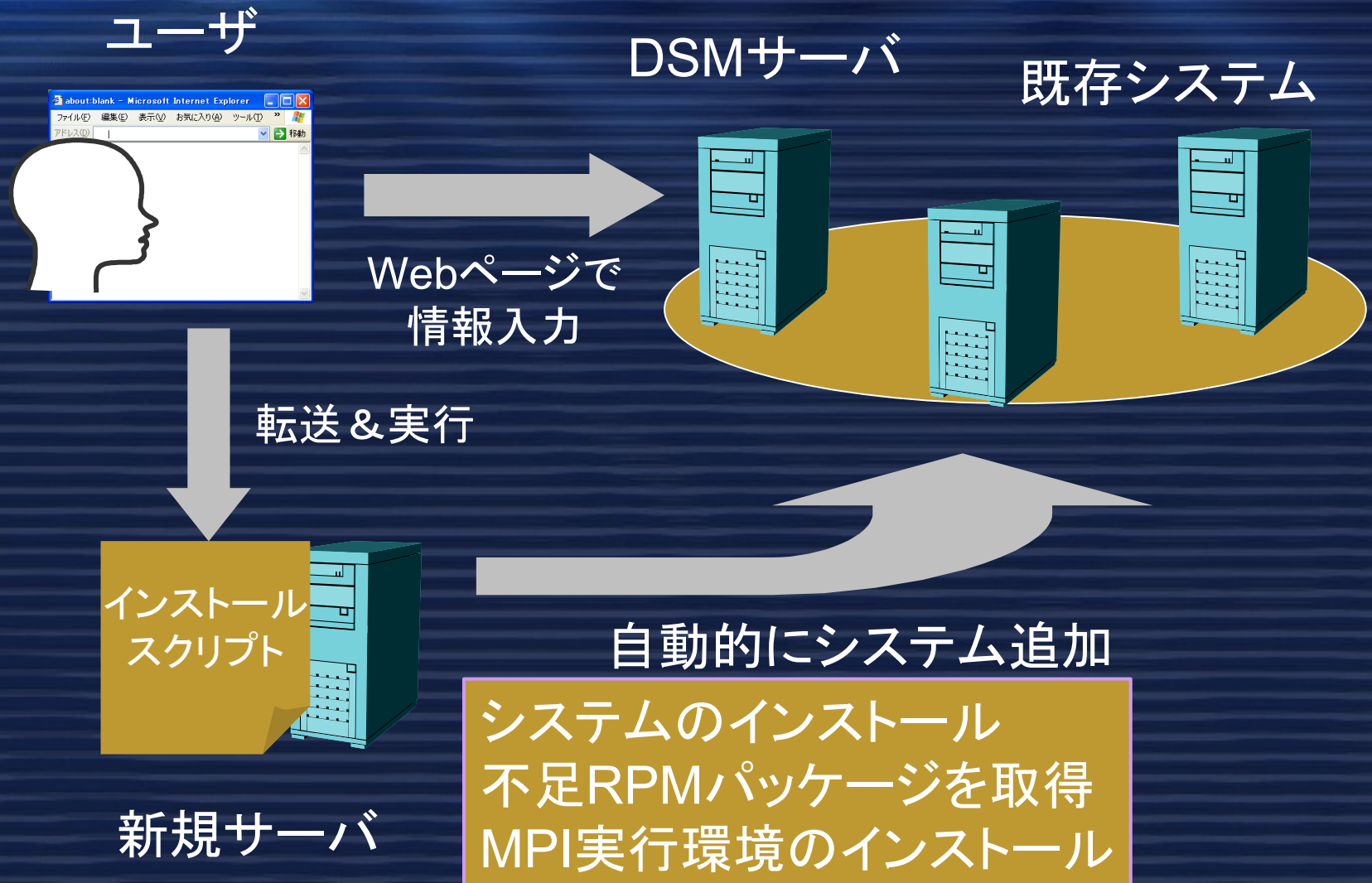
リダイレクト

- ・ DSMサーバへユーザを誘導
 - CGIによるHTTPリダイレクトを利用
- ・ 容易なアクセス
 - ユーザは, システム内のどれかのサーバのみを知っていればよい



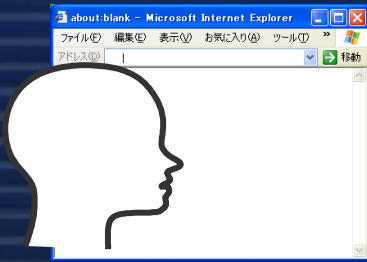
システム構築支援

既存システムへサーバを追加する場合



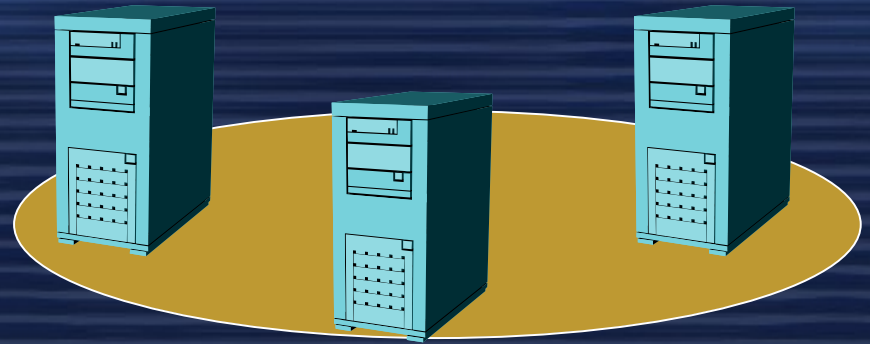
システム構築支援

ユーザ



DSMサーバ

既存システム



新規環境構築も可能

インストールCD



新規サーバ

自動的にシステム追加

システムのインストール
不足RPMパッケージを取得
MPI実行環境のインストール

システム構築支援(2)

システムWebページでのスクリプト生成

Job実行支援システム - Mozilla Firefox

ファイル(F) 編集(E) 表示(V) 移動(G) ブックマーク(B) ツール(T) ヘルプ(H)

http://skk-c5/~fuju/cgi-bin/management/index.cgi

Logout
Top

JES-System

ジョブ実行支援システム

本ページは、本システムにサーバを追加するためのスクリプトを生成するページです。
必要事項を入力し、生成されたスクリプトを新規サーバにて実行することで、
ジョブ実行支援システムに組み込むことができます。

項目	入力領域
サーバ名	skk-c4
あなたのメールアドレス	mt043225@cc.utsur
システムの配備場所(例:/home/fuju)	/home/fuju
MPIインストール	<input checked="" type="radio"/> Install <input type="radio"/> Pass
MPI配備場所(推奨:/usr/local)	/usr/local
MPI相互通信サーバ	skk-c5 itsc-9
	skk-s1
必要パッケージ不足時の動作	<input type="radio"/> Install <input checked="" type="radio"/> Pass
RPM取得コマンド	Yum apt-get
Proxyサーバ名(空欄可)	utsunomiya-u.ac.jp
Proxyポート番号(空欄可)	8080
	決定

完了

システム構築支援(2)

インストールCDを用いたインストール

```
root@skk-c4:/home/fuju/setupMpi
ファイル(E) 編集(E) 表示(V) 端末(T) タブ(T) ヘルプ(H)
[fuju@skk-c4 ~]$ su -
Password:
```

```
[root@skk-c4 ~]# cd /home/fuju/
[root@skk-c4 setupMpi]# perl se
本プログラムは対話でMPIをインス
(C)2005 Hideaki Fuju. All Right
```

注意: 実行ユーザはROOTである必要

ユーザ名を調べています...

本プログラムの実行が許可されま

MPIをインストールするディレクト

(本システムでは/usr/localを推奨

/usr/local

/usr/localでよろしいですか?(y

y

/usr/localにインストールを開始し

しばらくお待ちください...

作業を開始します...

一時フォルダを作成します 1/5

configureを実行します 2/5

```
root@skk-c4:/home/fuju/setupMpi
ファイル(E) 編集(E) 表示(V) 端末(T) タブ(T) ヘルプ(H)
```

```
for the <X.h> header for C++ includes, or <iostream> instead of the deprecated h
eader <iostream.h>. To disable this warning use -Wno-deprecated.
```

```
/usr/local/mpich/bin/mpicxx -o hello++ hello++.o
```

```
/usr/local/mpich/bin/mpicc -c cpilog.c
```

```
/usr/local/mpich/bin/mpicc -o cpilog.cpi
```

```
rm -f *.o *~ Pl* cpi pi3 simpleio hello
```

```
rm -rf SunWS_cache ii_files pi3f90.f pi3f
```

```
installed MPICH in /usr/local/mpich
```

```
/usr/local/mpich/sbin/mpiuninstall may be
```

```
インストール完了
```

```
一時フォルダを削除します 5/5
```

```
*****
```

```
xinetdのインストール有無をチェックします
```

```
xinetdがインストールされていません
```

```
*****
```

```
yumまたはapt-getコマンドを使ってxinetdをインストールします
```

```
使用するアップデートを選択してください(yum or apt-get)
```

```
yum
```

```
Proxyの設定が必要ですか(y or n)
```

```
y
```

```
Proxyサーバ名・ポート番号を入力してください
```

```
Proxy Server:uumes.cc.utsunomiya-u.ac.jp
```

```
Proxy Port:8080
```

```
Proxyサーバ:uumes.cc.utsunomiya-u.ac.jp, ポート:8080でよろしいですか(y or n)
```

yum, apt-getを利用
したインストール

まとめ

- ジョブ実行支援システムの開発
 - ジョブの投入から結果データの取得までを支援
 - 動的システム管理, 結果データ2重化によりロバスト性を高めた
 - システム構築支援により一般ユーザでも比較的容易に構築可能
- 今後の課題
 - 結果データのインテリジェンスな格納と整理

研究業績

- Hideaki FUJU, Hideaki SUGIURA, Yuichi SAITOH, Yoshikazu HAYASE, Hitohide USAMI, Motohiro YAMADA, Yutaka MIYAHARA, Hiroyuki KANAZAWA, Takashi KIKUCHI, Shigeo KAWATA : Job Execution Service System in a distributed computer environment, 第8回問題解決環境ワークショップ/第3回グリッドセミナー論文集, pp.52-57, 2005
- 杉浦秀明, 藤生英昭, 齊藤祐一, 早勢欣和, 菊池崇志, 川田重夫 : 分散型問題解決環境“D-NCAS”の構築, 第8回問題解決環境ワークショップ/第3回グリッドセミナー論文集, pp.12-15, 2005
- Shigeo KAWATA, Masumi INABA, Hideaki FUJU, Hideaki SUGIURA, Yuichi SAITOH, Takashi KIKUCHI : Computer-Assisted Liaison among Modules in a Distributed Problem Solving Environment (PSE) for Partial Differential Equation Based Problems, Trans. of the Japan Society for Computational Engineering and Science, Paper No.20050029, 2005
- 杉浦秀明, 藤生英昭, 稲葉真純, 早勢欣和, 宇佐見仁英, 菊池崇志, 川田重夫 : 分散計算機環境におけるジョブマネージメントシステムの開発, 計算工学講演会論文集 第10巻 第2号, pp.739-742, 2005
- 藤生英昭, 稲葉真純, 齊藤祐一, 杉浦秀明, 早勢欣和, 菊池崇志, 川田重夫 : 分散型問題解決環境“D-NCAS”におけるモジュール連携, 計算工学講演会論文集 第10巻 第2号, pp.735-738, 2005
- 稲葉真純, 藤生英昭, 齊藤祐一, 早勢欣和, 菊池崇志, 川田重夫 : 分散型PSE“D-NCAS”におけるモジュール連携支援, 第7回問題解決環境ワークショップ/第2回グリッドセミナー論文集, pp.13-18, 2004
- 杉浦秀明, 藤生英昭, 稲葉真純, 早勢欣和, 宇佐見仁英, 菊池崇志, 川田重夫 : 分散環境におけるデータの分散処理システムの構築, 第7回問題解決環境ワークショップ/第2回グリッドセミナー論文集, pp.99-104, 2004
- 藤生英昭, 稲葉真純, 杉浦秀明, 早勢欣和, 宇佐見仁英, 菊池崇志, 川田重夫 : 分散コンピュータ環境におけるジョブ実行支援システムに関する研究, 第7回問題解決環境ワークショップ/第2回グリッドセミナー論文集, pp.117-122, 2004
- Inaba, M., Fuju, H., Kitamuki, R., Kawata, S., Kikuchi, T. : “Computer-Assisted Documentation in a Problem Solving Environment (PSE) for Partial Differential Equation Based Problems”, Trans. of the Japan Society for Computational Engineering and Science, Paper No.20040025, 2004
- 稲葉真純, 藤生英昭, 北向亮治, 早勢欣和, 菊池崇志, 川田重夫 : 分散型問題解決環境“D-NCAS”の開発, 計算工学会講演会論文集 第9巻 第2号, pp.689-692, 2004
- 藤生英昭, 稲葉真純, 北向亮治, 早勢欣和, 宇佐見仁英, 菊池崇志, 川田重夫 : 分散コンピュータ環境におけるジョブ実行支援PSEの構築, 計算工学会講演会論文集 第9巻 第2号, pp.693-696, 2004
- Kawata, S., Usami, H., Hayase, Y., Miyahara, Y., Yamada, M., Fujisaki, M., Numata, Y., Nakamura, S., Ohi, N., Matsumoto, M., Teramoto, T., Inaba, M., Kitamuki, R., Fuju, H., Senda, Y., Tago, Y. and Umetani, Y. : “A Problem Solving Environment (PSE) for Distributed Computing”, accepted by Int. J. High Performance Computing and Network, 2004