

HPC環境統合運用ミドルウェア ShareTaskのご紹介

(株) アンクル 齊藤 隆之

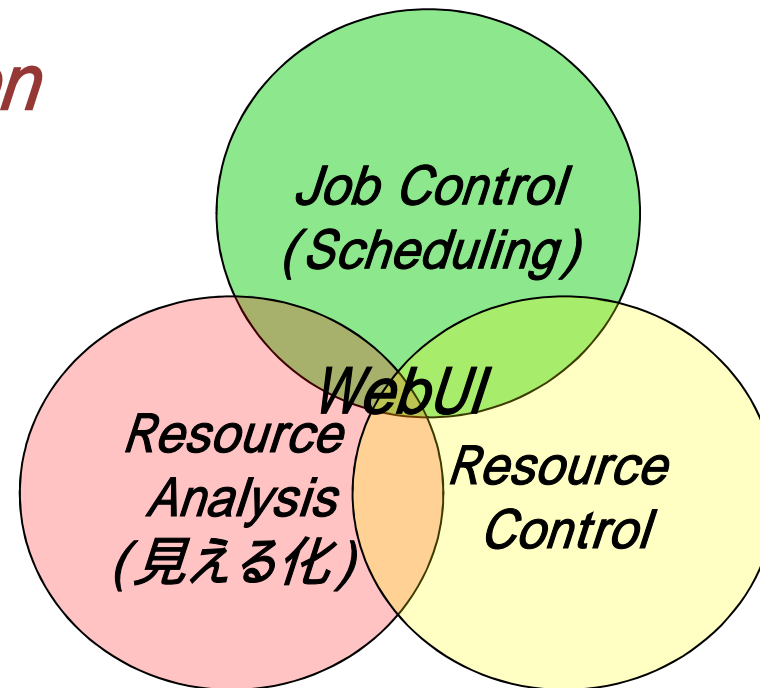
2013年6月

ShareTask (シェアタスク)とは

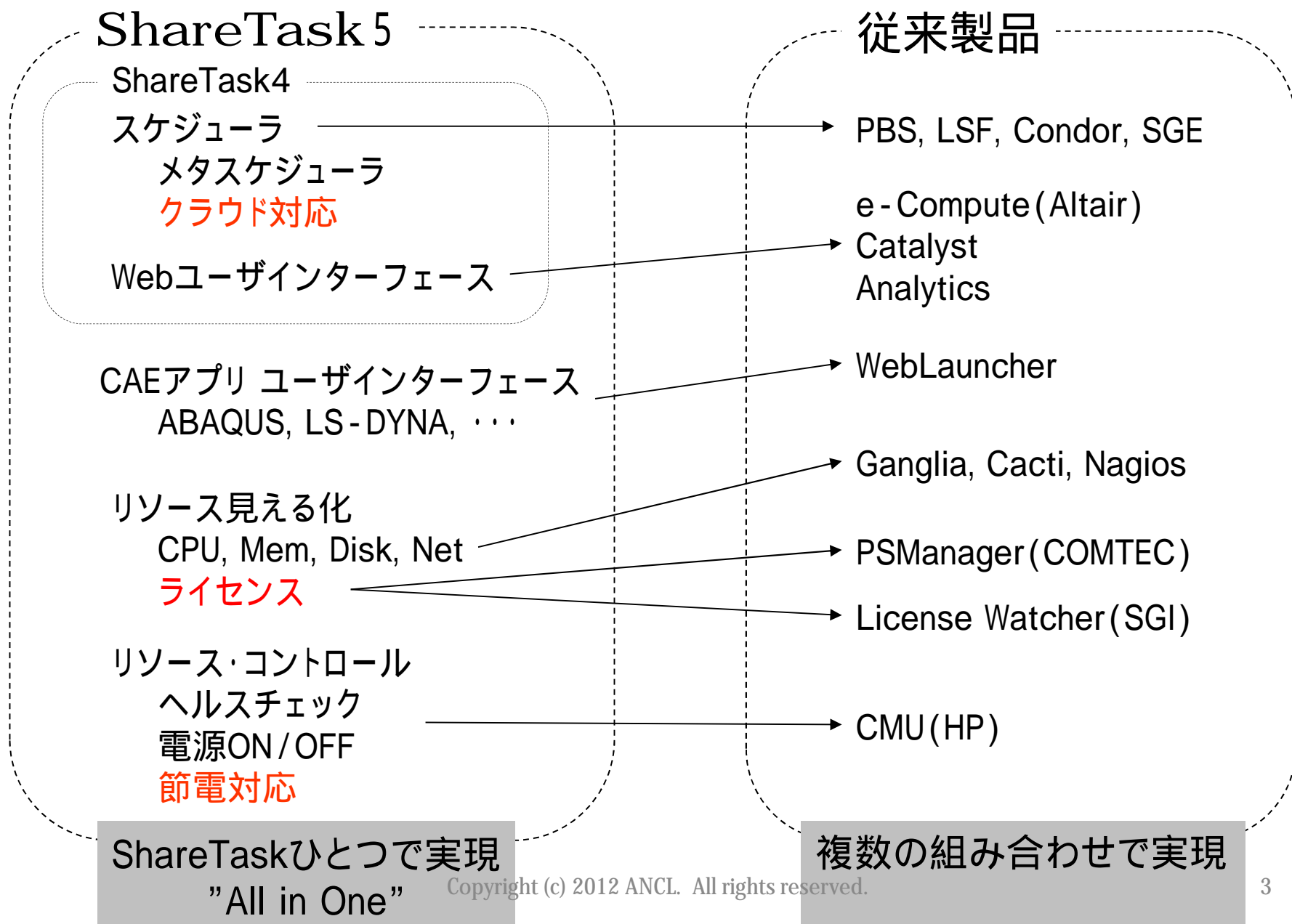
- HPCのための統合運用ツール

- *All-in-One Solution*

- *Web based*

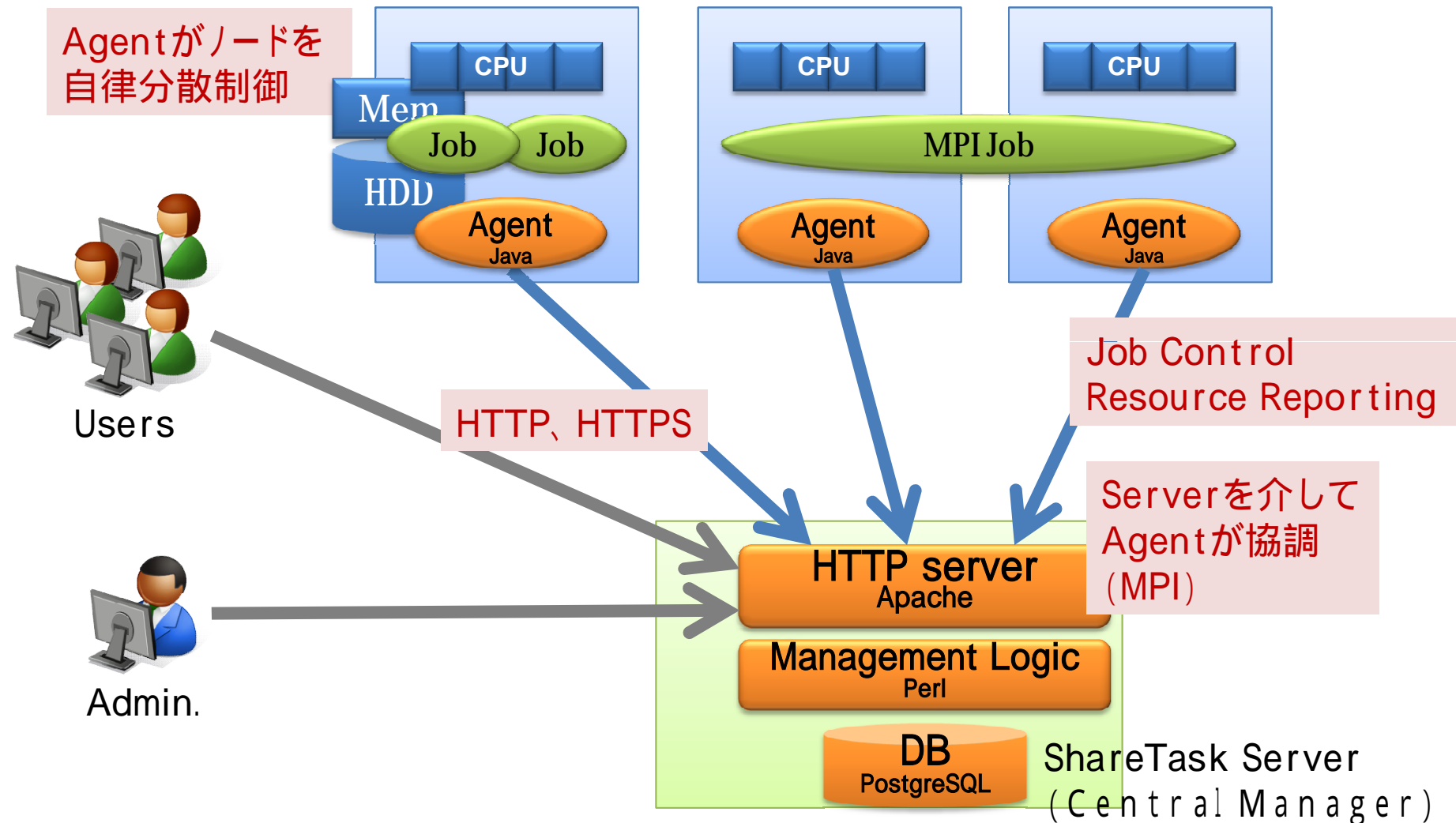


ShareTask5: All-in-One



ShareTask: pure Web application

Nodes(inc. Login Node, File Server)



ShareTask: High Availability

- 自律分散制御なので
 - 障害につよく、
障害箇所を自動的に切り離して業務継続
- Central ManagerもHA構成が可能
 - Webサーバ、DBサーバのHA技術を利用

Job Control(Scheduling)

- Resource Mapping
 - *Job - Hardware - License - Priority Mapping*
- CAE application support
 - *Submit and In/Output file management*
 - *ABAQUS, (LS-DYNA), (DYTRAN), (NASTRAN)*
 - *(PowerFlow), (STREAM), (OpenFOAM)*
 - *FieldView, ParaView*
- Job status monitoring
 - *execution/waiting time estimation*

Job Control (cont.)

CAE application support

ジョブの登録(ABAQUS)

メインメニュー | ジョブリスト | CPU時間集計

ようこそ saito (admin) さん [LOGOUT]

☒ 登録完了後に入力を残す

登録 キューステータス

プログラム名 *	abaqus_6.10-2	実行ディレクトリ	参照	saiko/models/model1/
キュー *	large3	入力ファイル(*.inp)	<input type="checkbox"/> ローカルのファイルを送信	
オプション	cpus= 8 <input type="checkbox"/> datacheck <input type="checkbox"/> recover <input checked="" type="checkbox"/> double <input type="checkbox"/> compress その他追加オプション	job=	case2.inp	
割り込み優先度	NORMAL(5)	リスタートファイル(*.res)	<input type="checkbox"/> ローカルのファイルを送信	
強制打ち切り時間	10h 例: 1h 2d 1w	oldjob=	case1.res	
コメント	contact test model	モデルファイル(*.odb)	<input type="checkbox"/> ローカルのファイルを送信	
		globalmodel=		
		ユーザープログラム(*.f)	<input type="checkbox"/> ローカルのファイルを送信	
		user=		

▼詳細オプション

登録

input file
recognition

ABAQUS
options

5.00-13-demo (C)2012 ANCL Inc.

Job Control (cont.)

Job status monitoring

メインメニュー | キューステータス | CPU時間集計

ジョブリスト

ようこそ saito (admin) さん [LOGOUT]

Job List

新規ジョブ登録 新規ジョブ登録(ABAQUS) 一括再実行 一括強制終了 一括削除 一括保留 [?]

ジョブ名	アプリケーション名	キュー	コメント	ユーザー
				saito (admin)
サブミット日時	計算開始日時	計算終了日時		

更新 10 Sec フィルター リセット

全選択 全解除 ☒ 待機中 ☒ 実行準備中 ☒ 実行中 ☒ 正常終了 ☒ 異常終了 ☒ 強制終了中 ☒ 強制終了 ☒ 一時停止中 ☒ 保留中 ☒ LOST

< << 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 >> > (19499/19499)

ID ▼ ▲	ジョブ名 ▼ ▲	アプリケーション名 ▼ ▲	キュー ▼ ▲	待ち予測	CPU	状態 ▼ ▲	登録 ▼ ▲	開始 ▼ ▲	終了 ▼ ▲	コメント
28607	TP8	gamess	large	-	8/8	正常終了	05/08 06:14	05/08 17:39	05/08 17:47	
28606	TP4	gamess	small	-	8/8	正常終了	05/08 06:13	05/08 17:38	05/08 17:40	
28605	TP8	gamess	large	-	8/8	正常終了	05/08 06:12	05/08 17:37	05/08 17:44	
28604	TP4	gamess	small	-	8/8	正常終了	05/08 06:12	05/08 17:37	05/08 17:38	

Job Details

28603

TP8

gamess

28602

TP4

gamess

28601

TP8

gamess

28600

TP4

gamess

28599

TP8

gamess

28598

TP4

gamess

28597

TP8

gamess

28596

TP4

gamess

28595

TP8

gamess

28594

TP4

gamess

28593

TP8

gamess

28592

TP4

gamess

28591

TP8

gamess

28590

TP4

gamess

28589

TP8

gamess

28588

TP4

gamess

ジョブ詳細 28607

ジョブID

28607 [owner:saito(501)]

正常終了

更新

再実行

削除

NCPU

8/8

HOST LIST

-

サブミット日時

2012-05-08 06:14:04

計算開始日時

2012-05-08 17:39:55

状態更新日時

2012-05-08 17:47:56

計算終了日時

2012-05-08 17:47:56

ジョブ名

TP8

アプリケーション名

gamess

キュー

large

引数

TP8.inp

コメント

計算ノード

domU-12-31-39-0F-1C-11/107.20.52.65(104)

計算結果詳細

exit status = 0; host = domU-12-31-39-0F-1C-11

通知メール

-

必須メモリ・ディスク

-

開始時間指定

-

割り込み優先度

LOW(0)

平均CPU使用率

87%

入力ファイル

TP8.inp(2.61KB)

[Download]

TP8.F05(2.61KB)

[Download]

TP8.F15(4.01MB)

[Download]

Down in /outp

Download in/output files

Resource Analysis(見える化) Visualization

- Graph tools
 - *CPU, Memory, Disk, Network*
 - *Application License*
 - *Utilization analysis*

Resource Analysis(見える化) (cont.)

Graph tool : CPU, Memory, Disk, Network

ようこそ saito (admin) さん [LOGOUT]

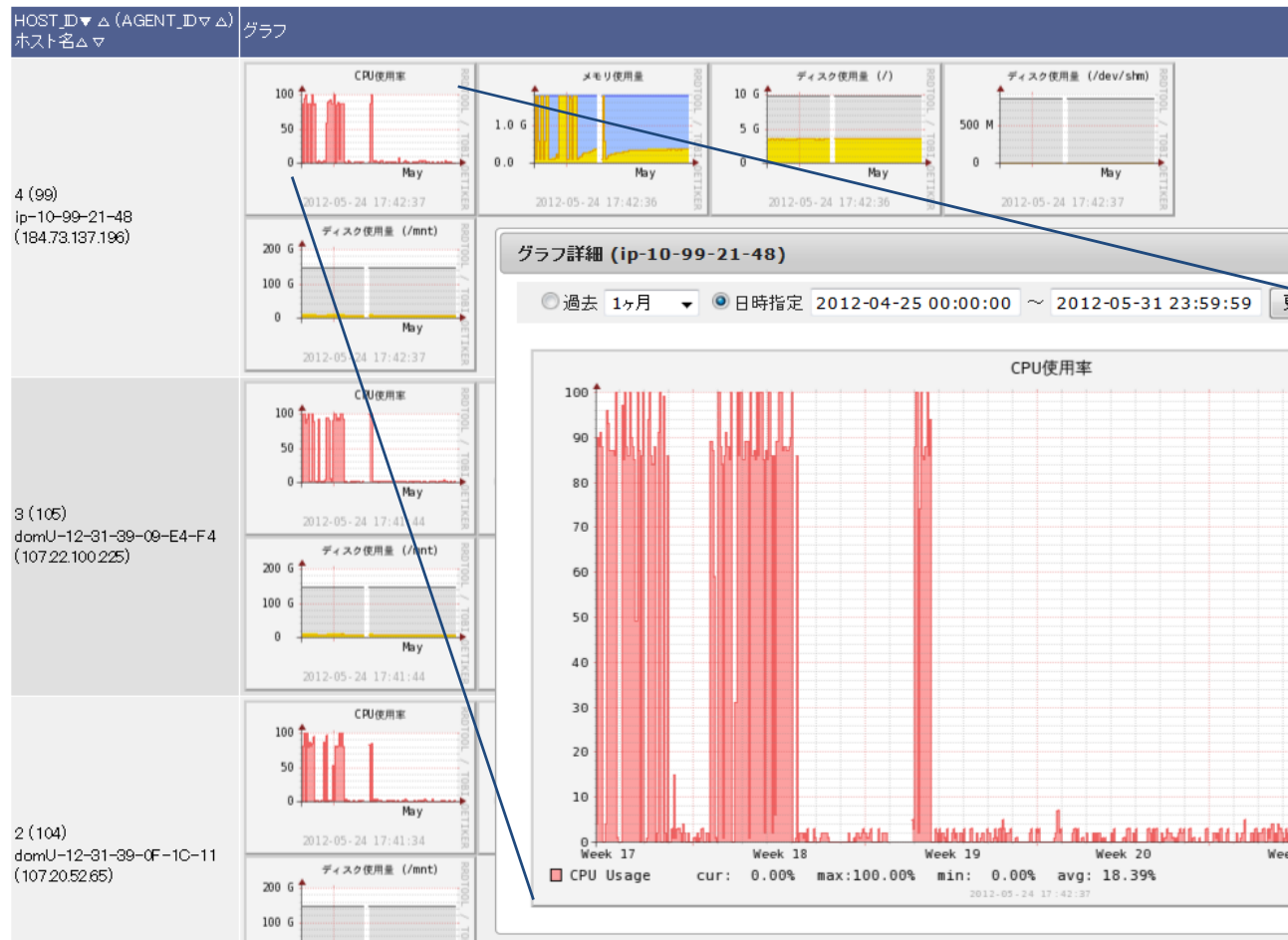
ID	エージェントID	ノードグループ	ホスト名	OS種別

グラフ表示範囲 ☒ 過去 1ヶ月 ☐ 日時指定 ~

☒ CPU使用率 ☒ メモリ使用量 ☒ ロードアベレージ ☒ ディスク使用量 ☒ ディスクI/O ☒ トラフィック ☒ プロセス

更新 300 Sec
フィルター リセット

|< << 1 >> >| (4/4)



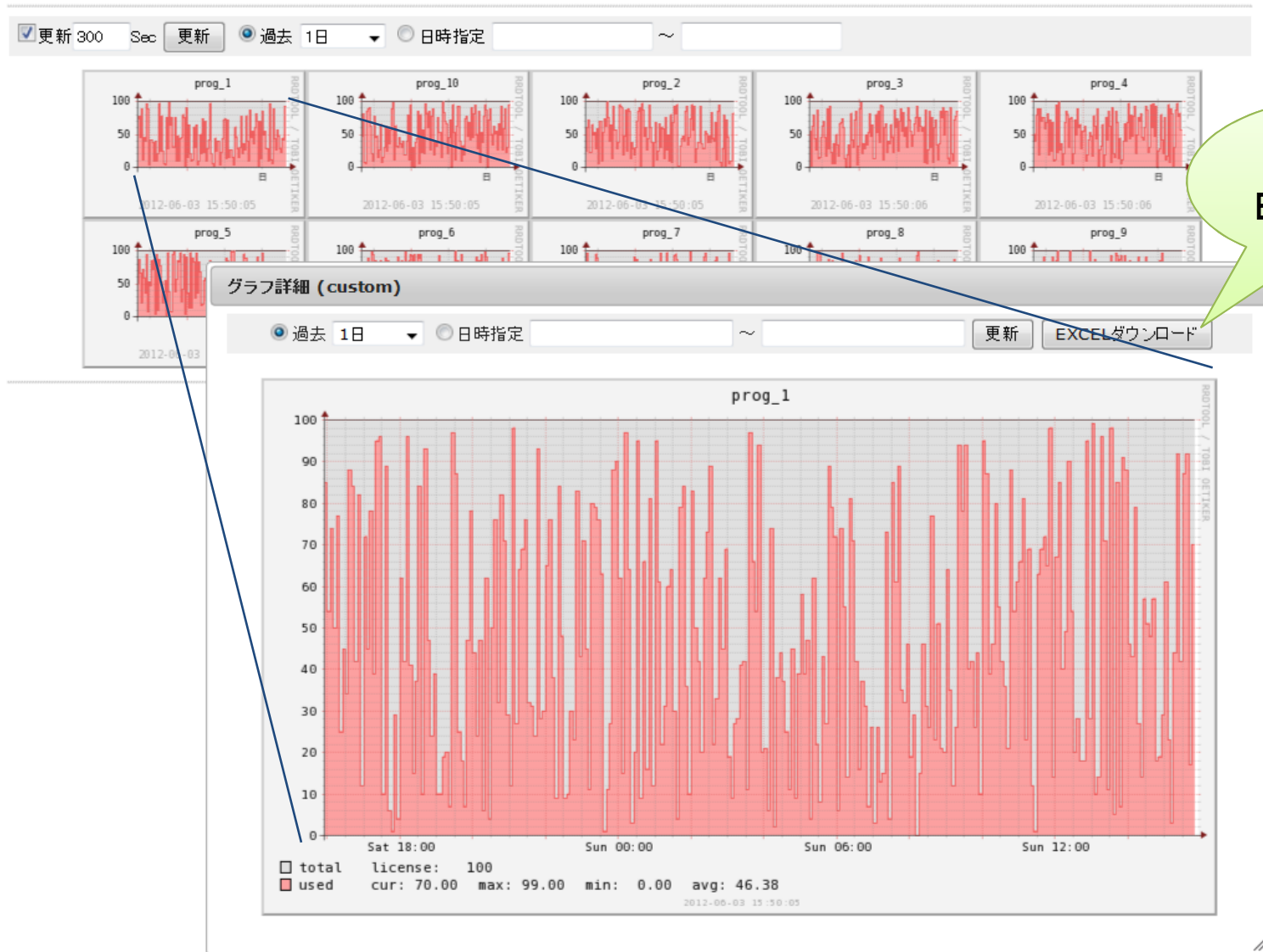
download
EXCEL file

Resource Analysis(見える化) (cont.)

Graph tool : Application License Utilization

メインメニュー | ジョブリスト | エージェント コントロール | キューステータス | ノードステータス | CPU時間集計

ようこそ saito (admin) さん [LOGOUT]



Resource Analysis(見える化) (cont.)

Graph tool : Utilization Analysis

稼働履歴

[メインメニュー](#) | [ジョブリスト](#) | [エージェントコントロール](#) | [キューステータス](#) | [ノードステータス](#) |

ようこそ saito (admin) さん [[LOGOUT](#)]

download
EXCEL file

EXCELファイルダウンロード

集計期間

2012-03-01

～

2012-06-01

キュー

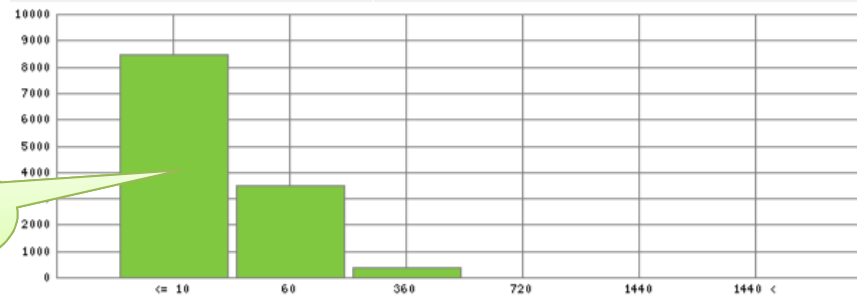
ホスト名

グラフ表示 * 時間は分単位

▼ 実行時間(X):ジョブ数(Y)

X軸メモリ 10,60,360,720,1440

(カンマ区切りで数値)



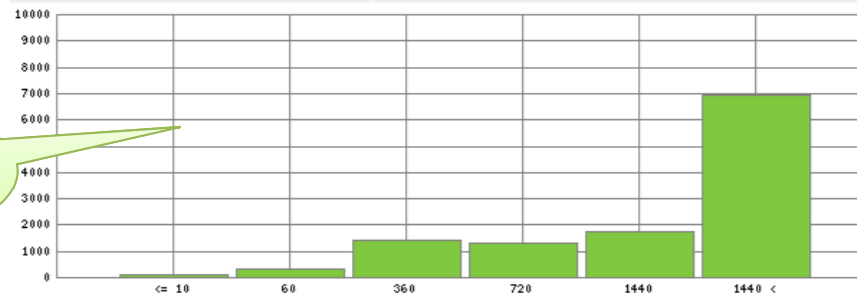
実行時間(分)	ジョブ数
<= 10	8460
60	3507
360	382
720	0
1440	0
1440 <	1
合計	12350

Histogram
Execution
time

▼ 待ち時間(X):ジョブ数(Y)

X軸メモリ 10,60,360,720,1440

(カンマ区切りで数値)



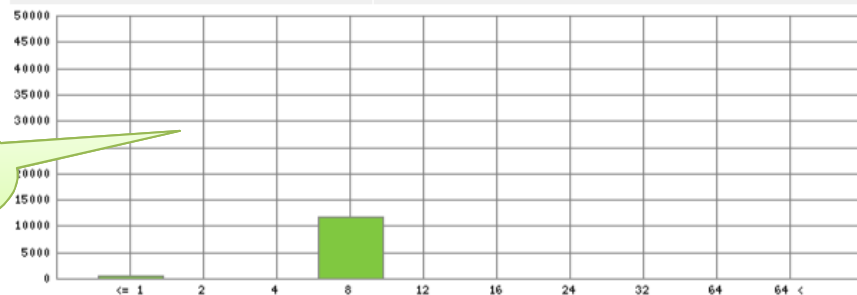
待ち時間(分)	ジョブ数
<= 10	115
60	331
360	1429
720	1323
1440	1731
1440 <	6955
合計	11884

Histogram
Waiting time

▼ CPU数(X):ジョブ数(Y)

X軸メモリ 1,2,4,8,12,16,24,32,64

(カンマ区切りで数値)



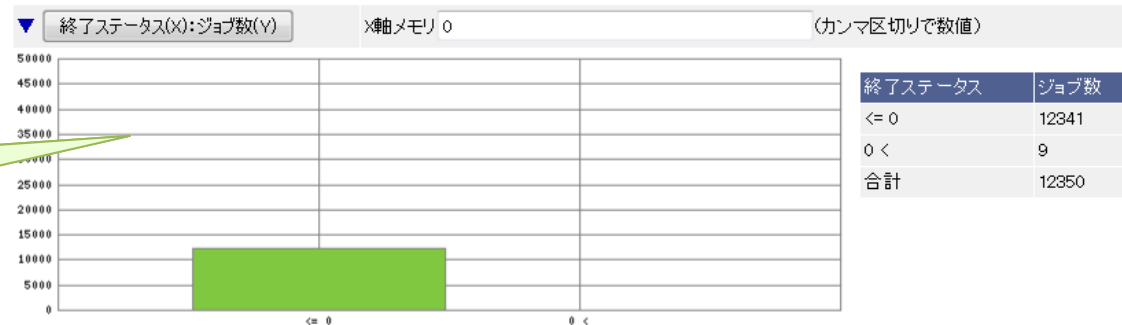
CPU数	ジョブ数
<= 1	487
2	0
4	0
8	11767
12	0
16	8
24	0
32	88

Histogram
CPU

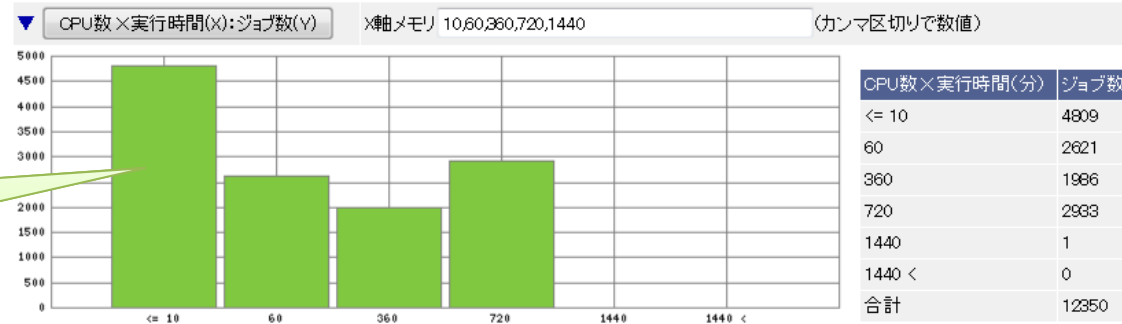
Resource Analysis(見える化) (cont.)

Graph tool : Utilization Analysis (cont.)

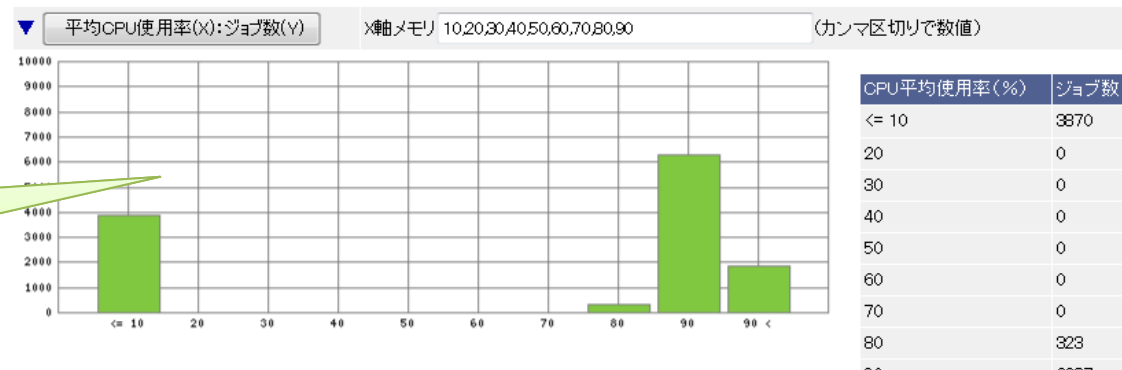
Histogram
exit status



Histogram
#CPU x time



Histogram
CPU
utilization



Resource Analysis(見える化) (cont.)

- Graph tools

- *CPU, Memory, Disk, Network*
- *Application License*
- *Utilization analysis*



- *Find idling*
- *Find shortage*
- *Budget planning*

Resource Control

- Health check
 - *active or dead*
 - *capacity management -> Lustre*
 - *syslog database -> trouble analysis*
- Power on / off
 - *auto shutdown / auto boot by WoL / IPMI*
 - *emergency shutdown (temp. sensor, etc)*
- Power saving
 - *節電対応 (shutdown idling node)*

Resource Control (cont.)

Health check: active or dead

エージェント コントロール

メインメニュー | ジョブリスト | ホストコントロール | キューステータス | ノードステータス | CPU時間集計 ようこそ saito (admin) さん [LOGOUT]

(接続グループに所属しているエージェントは重複して表示されます)

ID	ノードグループ	エージェント名	ホスト名	実行中ジョブ	エージェントコントロール
					start stop halt suspend
ステータス	コマンド	設定ファイル	接続先	<input checked="" type="checkbox"/> 更新 10 Sec	設定ファイル
				フィルター リセット	set view new/edit delete

「*」を入力するか、マウスのクリックで設定ファイルの全検索を参照できます

|< << 1 | 2 | 3 | 4 | 5 >> >| (93/93)

ID ▼ ▲	ノードグループ名 ▼ ▲	ホスト名 ▼ ▲ (エージェント名 ▼ ▲)	実行中ジョブ	ステータス ▼ ▲	コマンド ▼ ▲	設定ファイル ▼ ▲	接続先 ▼ ▲	アクセス日時 ▼ ▲
105		domU-12-31-39-09-E4-F4 (107.22.100.225)		shutdown	start	sample1.conf	sh-demo	2012-05-24 17:45:17
104		domU-12-31-39-0F-1C-11 (107.20.52.65)		shutdown	start	sample1.conf	sh-demo	2012-05-24 17:45:17
103		ip-10-244-141-220 (184.72.175.12)		operational	start	sample1.conf	sh-demo	2012-06-03 16:35:00
99		ip-10-99-21-48 (184.73.137.196)		shutdown	start	sample1.conf	sh-demo	2012-05-24 17:45:18
98	demo	ip-10-36-15-188 (184.73.141.161)		shutdown	start	sample1.conf	sh-demo	2011-12-24 23:33:26
97		domU-12-31-39-09-B5-42 (184.73.146.208)		lost	stop	sample1.conf	sh-demo	2011-12-13 23:04:45
96		extaic01 (211.126.199.28)		shutdown	stop	sample1.conf	secanclab.com	2011-12-02 00:57:48
92	demo	(001)ip-10-245-114-115 (184.72.142.173)		shutdown	stop	sample1.conf	secanclab.com	2011-12-07 16:01:04
91	demo	(002)ip-10-70-70-252 (107.22.19.151)		shutdown	stop	sample1.conf	secanclab.com	2011-12-07 16:02:13
90	demo	(003)domU-12-31-39-00-C4-B2 (107.20.62.233)		shutdown	stop	sample1.conf	secanclab.com	2011-12-07 16:24:43
89	demo	(004)ip-10-46-63-26 (204.236.241.59)		shutdown	stop	sample1.conf	secanclab.com	2011-12-07 16:30:05

Resource Control (cont.)

Health check: syslog database -> trouble analysis

ようこそ saito (admin) さん [LOGOUT]

表示 30 行 | RELOAD 0 Sec

(ALL ON|ALL OFF) info .. emerge

☒ kern ☒ user ☒ mail ☒ daemon ☒ auth
☒ authpriv ☒ cron ☒ syslog ☒ news ☒ uucp
☒ lpr ☒ ftp ☒ local0 ☒ local1 ☒ local2
☒ local3 ☒ local4 ☒ local5 ☒ local6 ☒ local7

日付: < > フィルター リセット ダウンロード

ホスト: RegE 除外: RegE
プログラム: RegE 除外: RegE
メッセージ: RegE 除外: RegE

< < 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 >> > (261546/4668225)

日時▼▲	ホスト ▼▲ (プログラム▼▲)	Fac▼▲ (Pri▼▲)	メッセージ	2012-06-03 00:00:00 .. 2012-06-03 23:59:59
12-06-03 17:02:04	mpi (ShareTask)	local1 (info)	[3227] [mpi_calc_task_get.cgi] NO MPI-JOB for AGENT:726(NCPU:8) PROG_V>>	
12-06-03 17:02:04	mpi (ShareTask)	local1 (info)	[3227] [mpi_calc_task_get.cgi] AUTH SUCCESS AGENT:726	
12-06-03 17:02:04	mpi (ShareTask)	local1 (info)	[3227] [mpi_calc_task_get.cgi] START AGENT:726(NCPU:8) PROG_VER:small >>	
12-06-03 17:01:54	mpi (ShareTask)	local1 (info)	[3221] [calc_get_agent_command.cgi] 726 start sample1.conf	
12-06-03 17:01:54	mpi (ShareTask)	local1 (info)	[3221] [calc_get_agent_command.cgi] AUTH SUCCESS AGENT:726	
12-06-03 17:01:54	mpi (ShareTask)	local1 (info)	[3221] [calc_get_agent_command.cgi] START /stask/scripts-kk/calc_get_a>>	
12-06-03 17:01:50	demo (ShareTask)	local0 (info)	[27326] [mpi_calc_task_get.cgi] END (0.20 sec)	
12-06-03 17:01:50	demo (ShareTask)	local0 (info)	[27326] [mpi_calc_task_get.cgi] NO MPI-JOB for AGENT:103(NCPU:8) PROG>>	
12-06-03 17:01:50	demo (ShareTask)	local0 (info)	[27326] [mpi_calc_task_get.cgi] AUTH SUCCESS AGENT:103	
12-06-03 17:01:50	demo (ShareTask)	local0 (info)	[27326] [mpi_calc_task_get.cgi] START AGENT:103(NCPU:8) PROG_VER:small>>	
12-06-03 17:01:48	demo (ShareTask)	local0 (info)	[27322] [calc_get_agent_command.cgi] 103 start sample1.conf	
12-06-03 17:01:48	demo (ShareTask)	local0 (info)	[27322] [calc_get_agent_command.cgi] AUTH SUCCESS AGENT:103	
12-06-03 17:01:48	demo (ShareTask)	local0 (info)	[27322] [calc_get_agent_command.cgi] START /scripts-kk/calc_get_agent>>	
12-06-03 17:01:46	mpi (ShareTask)	local1 (info)	[3216] [mpi_calc_task_get.cgi] END (0.19 sec)	
12-06-03 17:01:46	mpi (ShareTask)	local1 (info)	[3216] [mpi_calc_task_get.cgi] NO MPI-JOB for AGENT:726(NCPU:8) PROG_V>>	
12-06-03 17:01:46	mpi (ShareTask)	local1 (info)	[3216] [mpi_calc_task_get.cgi] AUTH SUCCESS AGENT:726	
12-06-03 17:01:46	mpi (ShareTask)	local1 (info)	[3216] [mpi_calc_task_get.cgi] START AGENT:726(NCPU:8) PROG_VER:small >>	
12-06-03 17:01:45	mpi (sharetaskd)	local0 (info)	[3430] MAIN ALIVE	
12-06-03 17:01:32	mpi (sharetaskd-Director)	local0 (info)	[3432] Director PROCESS END (4.6237 sec) (sleep:55.3763)	
12-06-03 17:01:32	mpi (sharetaskd-Director/Host)	local0 (info)	[3171] Director/Host EXIT	
12-06-03 17:01:32	mpi (sharetaskd-Power)	local0 (error)	[3201] D/A STATUS 35:domU-12-31-39-07-8D-A5(23.20.118.89)/23.20.118.89>>	
12-06-03 17:01:32	mpi (sharetaskd-Power)	local0 (error)	[3195] D/A STATUS 41:TEST(172.25.0.35)/172.25.0.35 to [lost]	
12-06-03 17:01:32	demo (ShareTask)	local0 (info)	[27304] [mpi_calc_task_get.cgi] END (0.20 sec)	
12-06-03 17:01:31	demo (ShareTask)	local0 (info)	[27304] [mpi_calc_task_get.cgi] NO MPI-JOB for AGENT:103(NCPU:8) PROG>>	
12-06-03 17:01:31	mpi (sharetaskd-Power)	local0 (error)	[3197] D/A STATUS 38:domU-12-31-39-0F-F1-73(204.236.252.40)/204.236.252.40>>	

Resource Control (cont.)

Power on / off

*auto shutdown / auto boot by WoL / IPMI
emergency shutdown (temp. sensor, etc)*

Power saving

節電対応 (shutdown idling node)

ホスト コントロール

メインメニュー | ジョブリスト | エージェント コントロール | キューステータス | ノードステータス | CPU時間集計

ようこそ saito (admin) さん [LOGOUT]

ID	エージェントID	ノードグループ	ホスト名	OS種別	ブート種別
ログ種別	グラフ種別	ステータス	コマンド	<input checked="" type="checkbox"/> 更新 10 Sec フィルター リセット	パワーコントロール boot shutdown

< << 1 >> > (4/4)

ID ▼ ▲	CID ▼ ▲	ホスト名 ▼ ▲	AID ▼ ▲	ノードグループ ▼ ▲	OS/ブート ▼ ▲	ログ/グラフ ▼ ▲	ステータス ▼ ▲	コマンド ▼ ▲	アクセス日時 ▼ ▲
4	1	ip-10-99-21-48 (184.73.137.196)	99		linux/wol	none/agent	lost	boot	2012-06-03 16:37:59
3	1	domU-12-31-39-09-E4-F4 (107.22.100.225)	105		linux/wol	none/agent	lost	boot	2012-06-03 16:37:59
2	1	domU-12-31-39-0F-1C-11 (107.20.52.65)	104		linux/wol	none/agent	lost	boot	2012-06-03 16:37:59
1	1	ip-10-244-141-220 (184.72.175.12)	103		linux/wol	none/agent	booting	boot	2012-06-03 16:37:59

< << 1 >> > (4/4)

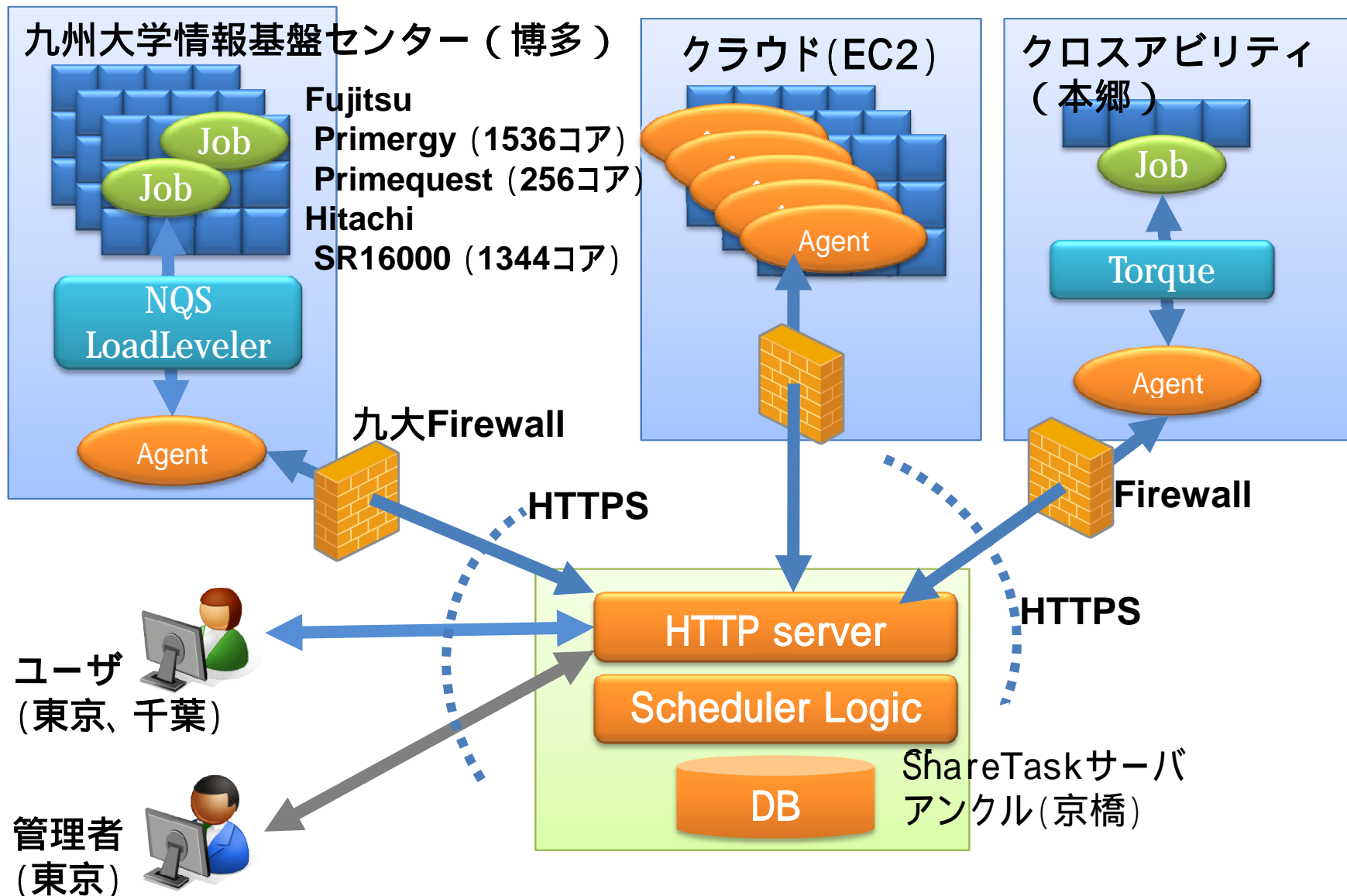
5.00-13-demo (C)2012 ANCL Inc.

導入実績

	組織数	ライセンス数(同時実行ジョブ数)
製造業	27	1260
研究機関	2	52
大学	2	90

2013年5月末現在

実証実験 (九大・東京・US間メタスケジューリング)

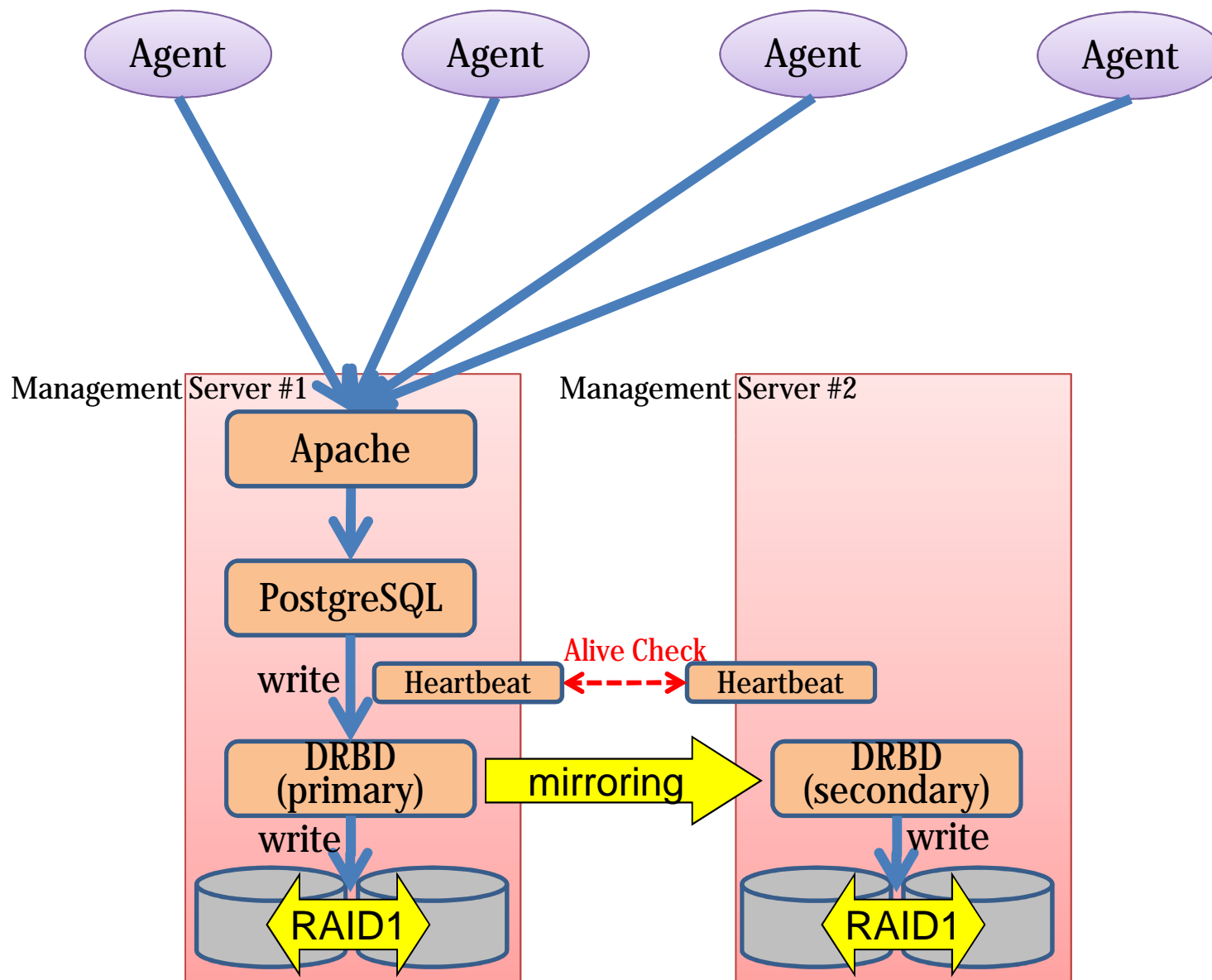


ShareTaskさえあれば、クラスターの運用管理は完璧です。

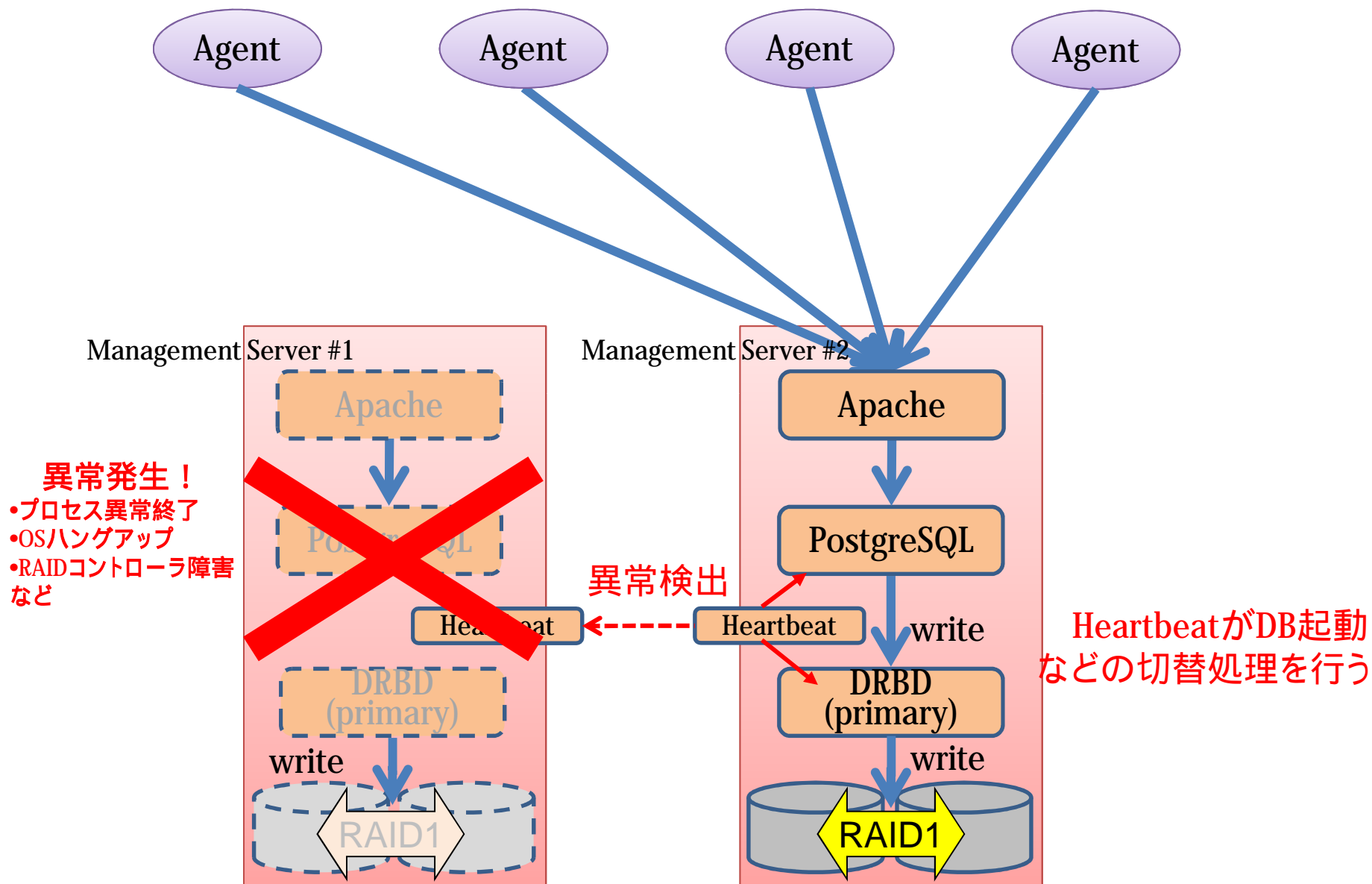
ありがとうございました。

お問い合わせ (株) アンクル
ShareTaskチームリーダー：齊藤
03-3562-7817
saito@anclab.com
<http://www.anclab.com/>

付録：サーバ冗長化構成 (HA)



障害発生時のフェイルオーバー



冗長構成回復処理

