Grid技術の動向と動作実績
-Grid Computing検討WG報告-

2009年11月25日

WGまとめ役　内田啓一郎

報告内容

- WG活動目標
- 活動メンバー
- 活動経過
- Grid Computing技術動向調査
- NAREGI関連報告
- NAREGIなどGrid適用実験・事例
- まとめ

当初のWG活動目標

- 広くGrid技術の動向調査と情報共有
- NAREGIなどの Grid Middlewareを実際に動作させ、その経験に基づいてGrid技術のあり方を検討し、センター等におけるGridのメリットについて検討

活動メンバー

担当幹事	岩宮 敏幸	宇宙航空研究開発機構
担当幹事	水本 好彦	国立天文台天文データセンター
まとめ役	内田 啓一郎	神奈川大学メディア教育・情報システムセンター
	井町 智彦	金沢大学総合メディア基盤センター
	南里 豪志	九州大学情報基盤研究開発センター
	義久 智樹	京都大学学術情報メディアセンター
	金澤　正憲	京都大学学術情報メディアセンター
	佐々木 節	高エネルギー加速器研究機構計算科学センター
	宇佐見 仁英	国立情報学研究所リサーチグリッド研究開発センター
	白崎 祐治	国立天文台天文データセンター
	津田 知子	名古屋大学情報連携基盤センター
	近堂 徹	広島大学情報メディア教育研究センター
	黒川 原佳	理化学研究所
まとめ役	藤崎 正英	TCソリューション事業本部
	井口 裕次	次世代テクニカルコンピューティング開発本部
	楢橋 幹一郎	文教ソリューション事業本部
	畑瀬 圭太	文教ソリューション事業本部
	北村 正和	文教ソリューション事業本部
	石原 康秀	TCソリューション事業本部
	荒井 雄二	TCソリューション事業本部
	坂口 吉生	TCソリューション事業本部
	岸本 光弘	富士通研究所 ITシステム研究所
話題提供	友石 正彦	国立情報学研究所リサーチグリッド研究開発センター
	建部 修見	筑波大学
	東田　学	大阪大学
	青柳　睦	九州大学情報基盤研究開発センター
	槌本　裕一	富士通研究所 ITシステム研究所
	三浦 謙一	国立情報学研究所リサーチグリッド研究開発センター
	合田 憲人	国立情報学研究所リサーチグリッド研究開発センター

活動経過

	日　時	場　所	活動内容
第1回	2007年10月31日(水) 14:30～17:30	富士通本社	・　WG活動内容の検討
第2回	2008年 2月22日(金) 14:30～17:45	富士通本社	・各種話題提供 ・　マシンの仮想化とGridとの関連 ・　NAREGIの現状と今後 ・　NAREGIの認証、PKIの話題 ・成果報告書における著作物等使用許諾について
第3回	2008年 6月 9日(月) 14:30～17:30	富士通本社	・2008年度の活動について ・各種話題提供 ・ 高エネルギー分野でのGrid適用事例 ・ データグリッド最新の話題 ・ NAREGIミドルウェアによる6拠点連携実験の話題 ・成果報告書の方向性について
第4回	2008年 9月18日(木) 14:30～17:40	富士通本社	・各種話題提供 ・クラウド・コンピューティング ・天文学研究におけるデータグリッド ・NAREGIのこれまでと今後 ・成果報告書のまとめ方
第5回	2009年 1月21日(水) 14:30～17:30	富士通本社	・各種話題提供 ・NAREGIの状況報告(九大) ・NAREGIの状況報告(名大) ・研究コミュニティ形成のための資源連携技術に関する研究報告 ・成果報告書のまとめ方
第6回	2009年 8月18日(火) 14:00～17:30	富士通本社	・各種話題提供 ・NAREGIの状況報告(九大) ・NAREGIの状況報告(名大) ・NAREGIの今後 ・成果報告書のまとめ方

Grid Computing技術動向調査

- マシンの仮想化とGridとの関連
  ーOGFやDMTFの標準化動向ー
- データグリッド
  ーRNSとGfarm v2ファイルシステムー
- 天文学研究におけるデータグリッド
- クラウド・コンピューティング

Grid Computing 検討WG 第2回会合

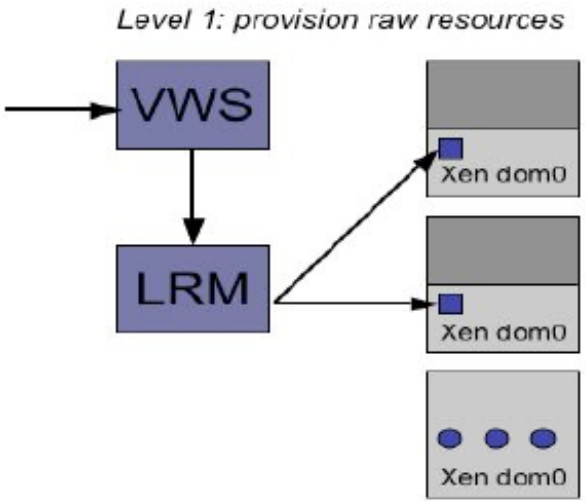
マシンの仮想化とグリッドとの関連
～OGFやDMTFでの標準化動向を中心に～

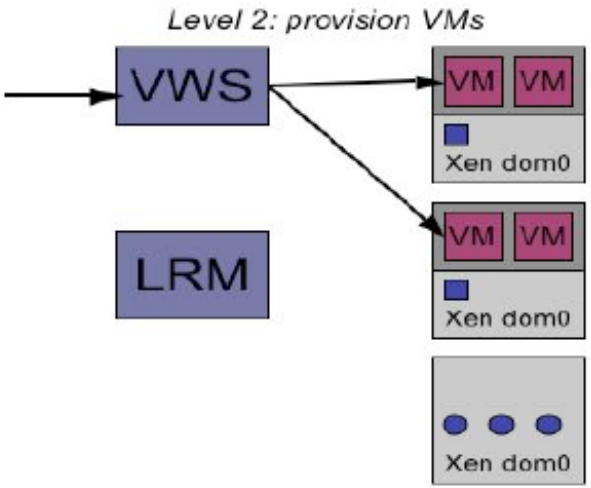
岸本光弘
（株）富士通研究所
2008.2.22

GT4 virtual Workplaces

FUJITSU

- Globus incubation projectのひとつ
  - Leader: Kate Keahey (ANL)
  - 最新リリースは v1.3.1 (2008.2.14)
- ユーザの作業環境提供に仮想マシンを活用
  - ユーザが、所望のOS, ソフトスタック, アクセスポリシを備えた実行環境を作成可能
  - ユーザが、VMイメージの開始・停止・再開・マイグレーションを制御可能
  - 資源割り当て量を厳密に履行可能
  - ユーザ間の完全な分離可能





GridVirt-WG

FUJITSU

- OGFで仮想サーバに関する標準化活動を行うため、OGF19(2007/1)にトップダウンで創設
  - 議長：Erol Bozak (SAP), Wolfgang Reichert (IBM)
- 目的(案)
  1. Verification that within existing Grid standards the specifications are neutral to virtualized systems and resources
     - The request for “resources” may be satisfied either by “virtualized resources” or “physical resources”
     - Is virtualization “transparent to the current Grid standards”?
  2. Explore how virtualization technologies can be exploited to better support Grid use cases
     - Define the use cases / scenarios wherein the Grid infrastructure is seen interacting with system virtualization platforms and making use of its capabilities
     - Exploit existing concepts and work, e.g.
- 標準仕様の開発を目指しているわけではないし、活動も始まっていない

System Virtualization WG

FUJITSU

- System Virtualization Partitioning and Clustering (SVPC) Workgroup
  - クラスタシステムや、LPARのCIMモデルを検討していた昔からあるWG
  - VM以前の活動は、あまり活発ではなかった
- Work ongoing to produce CIM model (CIM profile and associated CIM schema changes) for virtual systems and the virtual resources which compose them.
- Virtual Machine Image Format
  - Open Virtual Machine Format Specification (OVF)

まとめ

FUJITSU

- グリッドと仮想技術は、相互補完的
- 仮想化技術を使うことで、グリッドの適用範囲が広がり有用性が増す
  - 例： Amazon EC2やGlobus Virtual Workspaces
- OGFのGridVirt-WGは、グリッドユースケースを議論
- DMTFのSVPC-WGは、仮想サーバの標準仕様を作成
  - Virtual System Profile
  - System Virtualization Profile
  - Resource Allocation Profile, Allocation Capabilities Profile
  - Open Virtual Machine Format Specification

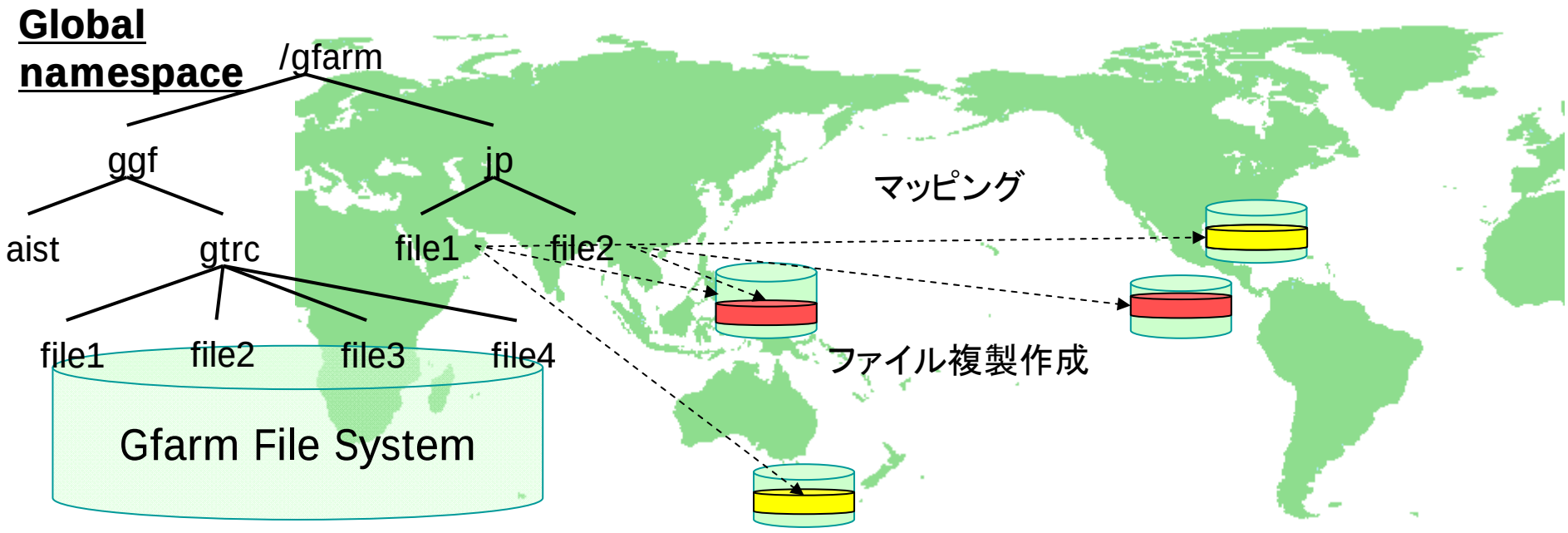
SS研Grid WG
2008年6月9日＠汐留

データグリッド最新の話題
- RNSとGfarm v2ファイルシステム

建部　修見
筑波大学
tatebe@cs.tsukuba.ac.jp

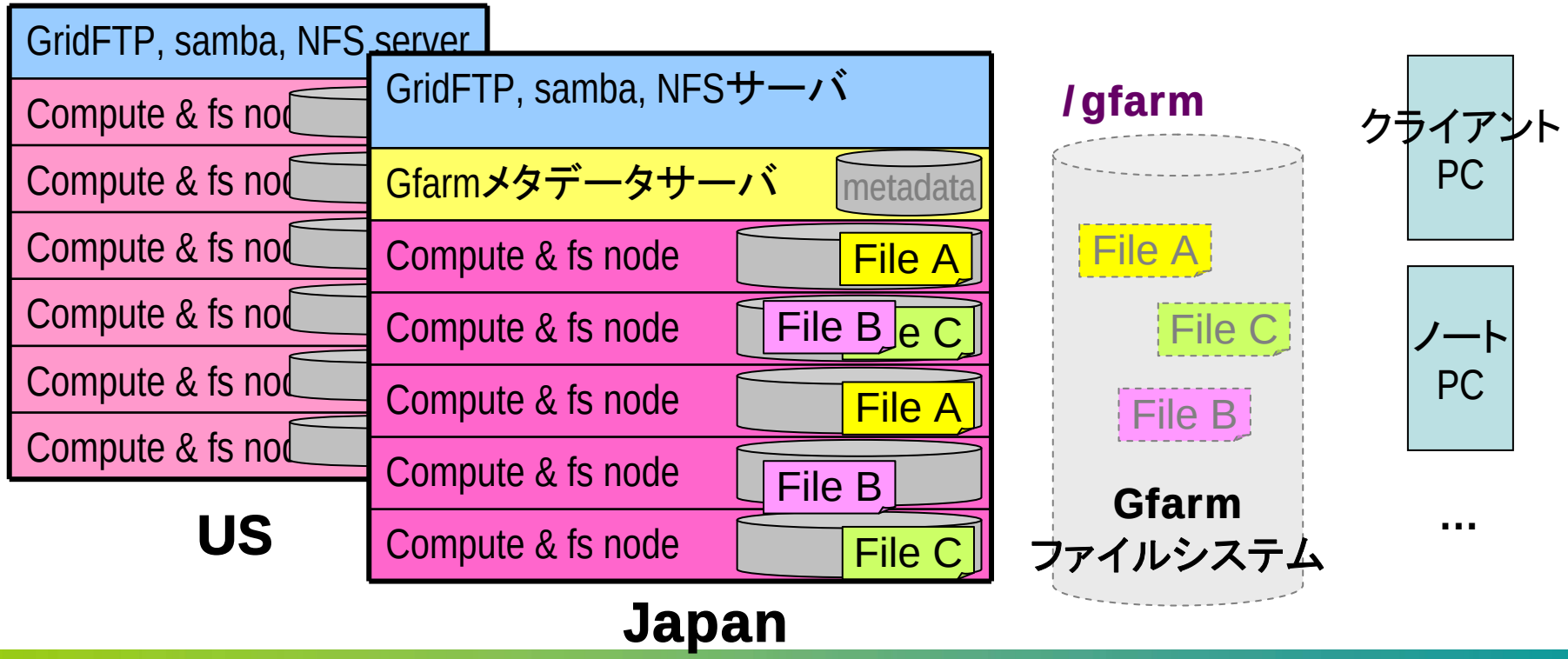
Gfarmグリッドファイルシステム（1）

- ストレージを束ねた分散ファイルシステム
- マウントして世界中から利用可能。アプリケーションを選ばない
- 利用者、プログラムの並列性に応じたスケーラブルなI/O性能
- 自動ファイル複製選択による耐故障性の実現とアクセス集中の回避



Gfarmグリッドファイルシステム（2）

- ファイルはどこからでも共有可能
- 実際には、ファイルは複製が作成され、どこかのファイルシステムノードに格納される
- アプリケーションは格納場所を意識することなくアクセス可能



まとめ

- **Gfarm v2ファイルシステム**
  - ファイルシステム一貫性の向上，セキュリティ向上，アクセス速度向上
  - 広域のクライアントによるスケーラブルなI/O性能
    - 日米の22クライアントから1,433 MB/sec
  - ファイル操作はメタデータサーバまでのRTTの2~4倍
    - ローカルクラスタではNFS(async)とほぼ同等
    - Gfarm v1と比べ最大23倍の高速化
  - ./configureはNFS(async)とほぼ同等，Gfarm v1より2.2倍高速
- **オープンソース開発**
  - 2008年5月27日にGfarm version 2.1.0をリリース
  - http://sf.net/projects/gfarm

Nominate this project
SOURCEFORGE.NET® 2008
COMMUNITY CHOICE AWARDS

天文学研究における
データグリッド

国立天文台　天文データセンター
白崎裕治

Data Resources in NAOJ

- Subaru 8.2m Optical-Infrared Telescope
- Kiso 105cm Schmidt Camera
- Okayama 188cm Optical Telescope
- Nobeyama 45m Radio Telescope
- Nobeyama Millimeter Array
- Nobeyama Radioheliograph
- VSOP
- VERA
- *ALMA*

Subaru

Data Resources in the World

- ESO VLT @ Paranal
- ESO telescopes @ La Silla
- Telescopes in Mauna Kea
  Subaru, Keck, JCMT, CSO, SMA,,,
- Telescopes on board satellites
  HST, GALEX,,,,

International Virtual Observatory Aliance

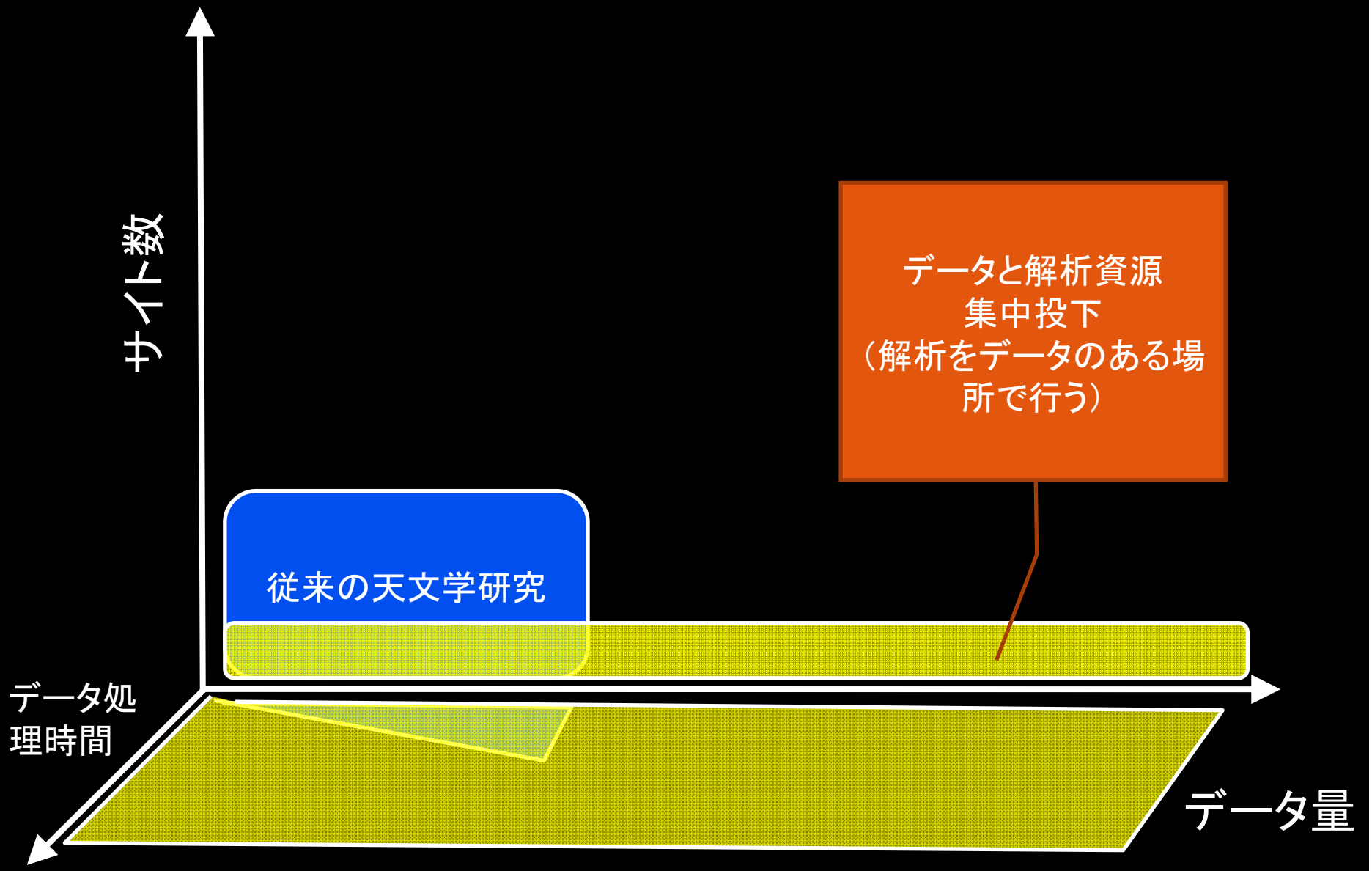
- 世界各国の 16 VO プロジェクトが参加
- 天文データの共有をより容易で効率的に行うための基盤整備
- データベースアクセス方式の標準化
- 検索言語や検索結果フォーマットの標準化

天文コミュニティー内で閉じたデータグリッドの実現

- 分散ファイル共有を行うための仕様
- シングルサインオン方式
- 非同期検索実行方式

国立天文台では Japanese Virtual Observatory (JVO) を開発 (ポータル)

グリッド解析システムで解決される領域



まとめ

- 多サイトの天文データベースを連携する基本的仕組みはほぼ完成した。(バーチャル天文台)
  - 小規模データに限定。
  - 単純なデータ検索
- バーチャル天文台の仕様拡張が現在進行中
  - 中規模データにも対応。
  - より高度な検索。
  - オントロジーにもとづくデータ発見など
- 大規模データの利用はデータアーカイブと並列計算システムをセットとしたシステムを日本各地に分散配置して対応したい。

THE POSSIBILITIES ARE INFINITE FUJITSU

クラウド・コンピューティング

2008年9月18日

株式会社 富士通研究所
ITシステム研究所
システムミドルウェア研究部
槌本 裕一

Copyright 2008 FUJITSU LABORATORIES LIMITED

クラウドコンピューティングとは

FUJITSU

- データもプログラムも雲（クラウド）の上に置かれている。ネットに接続するブラウザがあれば、どんな端末からでも雲に届く
  - by Eric Schmidt, Google CEO

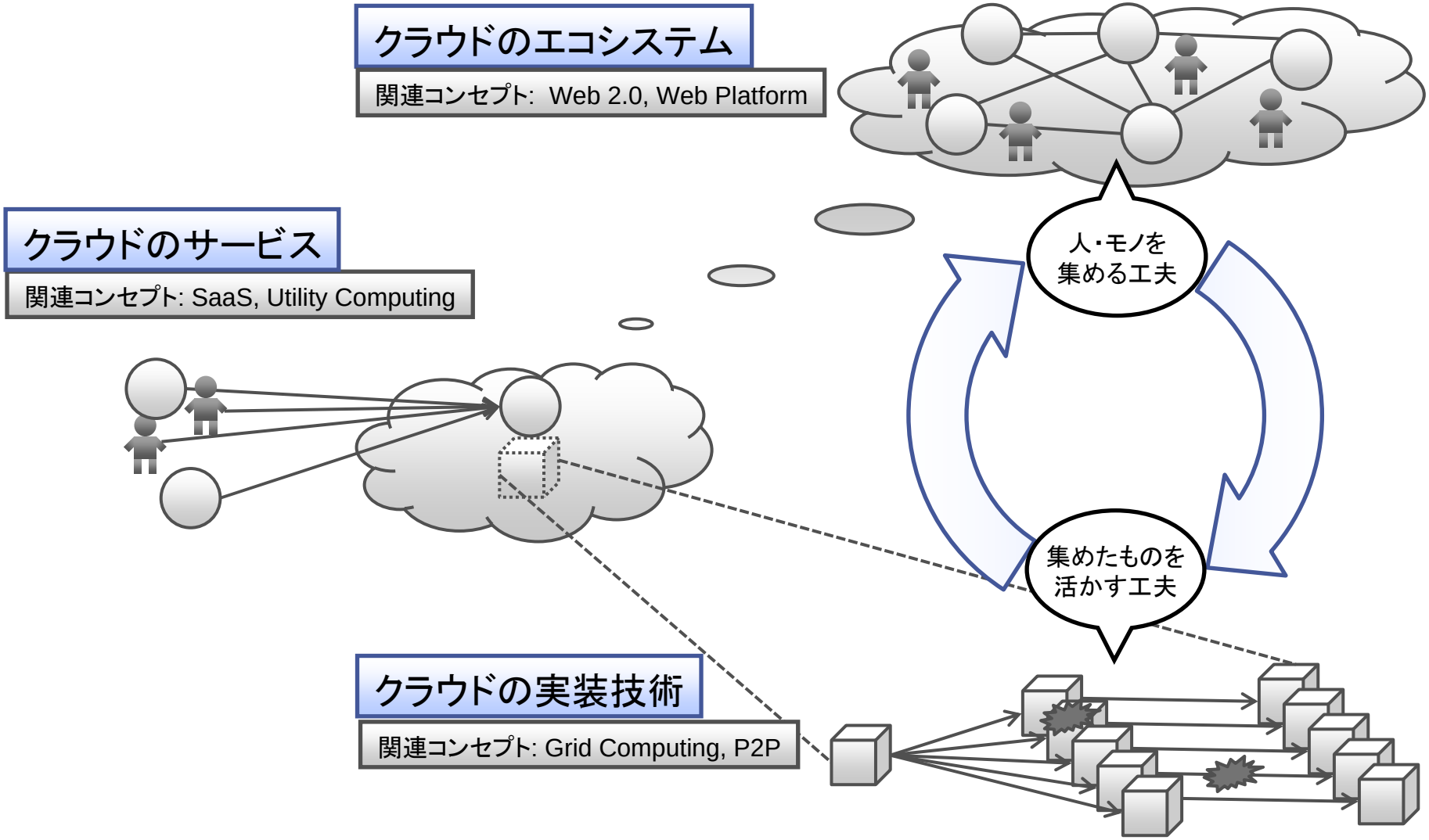
- 一連のITケーパビリティをサービスとして外部に提供すること
  - by Gartner

- インターネットという「雲 = クラウド」の向こう側に、サービスを提供するサーバなどがあるが、ユーザーはサーバの場所も台数も構成も認識できず、単にサービスを受け取る
  - by Wikipedia

ここではクラウド・コンピューティングの基本定義を
「ネット経由でよそのITケーパビリティを使うこと」
とする

Copyright 2008 FUJITSU LABORATORIES LIMITED

3つの観点の概念図

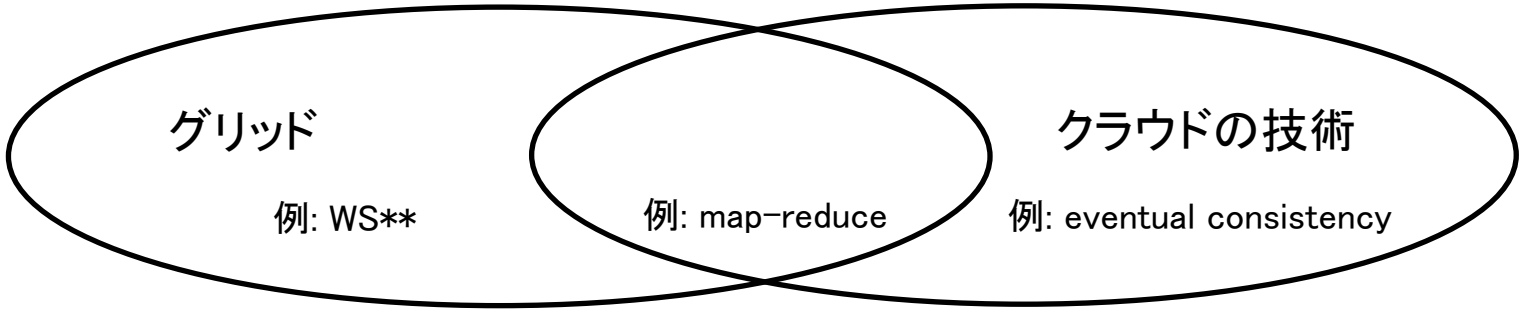


Copyright 2008 FUJITSU LABORATORIES LIMITED

グリッドとクラウド（私の理解）

FUJITSU

- グリッド
  - 出発点は、主に科学技術計算分野での現実の課題と、それに対する技術的ビジョン
  - 「技術的な理想」を表すバズワードへ
  - ITリソースの現状は不可避な前提条件という立場。主に三つの異なる技術的課題にフォーカス
    - 大量の不揃いなリソースをどう活用するか？ → クラスタのヘテロ版・クラスタのクラスタ
    - データボトルネックをどう克服するか？ → オンメモリ分散データ共有・大規模データの分散管理など
    - 複数の管理主体の下にあるリソースをどう扱うか？ → 標準化
- クラウド
  - 出発点は、主にコンシューマや中小企業向けのタダ・安価な魅力的サービス群の成功と、その提供者の思惑
  - 「IT利用環境の変化」を表すバズワードへ
  - 技術についてはニュートラル。ITリソースの在り方も含めて、単純に手段として優劣が問われる。その背景:
    - サービス提供者（インフラ屋ではなくアプリ屋）が単独で設計
    - 規模の経済を極限まで追求
    - 利用者からはサービスの<u>実装を隠蔽し垂直に全体最適化</u>



Copyright 2008 FUJITSU LABORATORIES LIMITED

企業システムへの影響

The Big Switch (by Nicholas Carr)*

FUJITSU

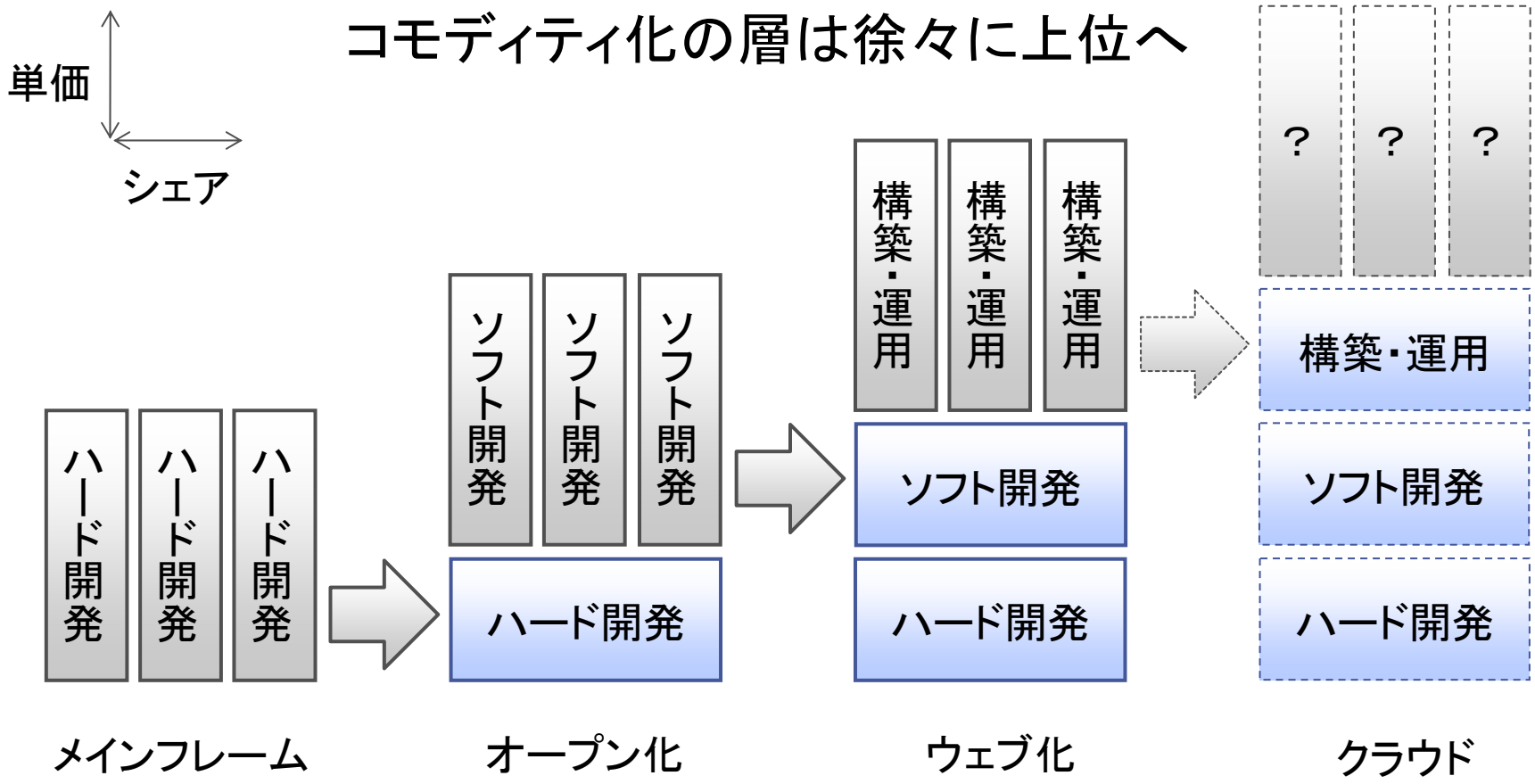
- 電力とITは非常に広範に利用できるので、供給を集約することができれば、**汎用技術は規模の経済の利点を享受**できる
- ユーティリティコンピューティングの興隆は、これまでのビジネスモデルを破壊する脅威である。数千から数億のユーザが使えるようなスケーラブルなサービスは、**企業やSMB向けソリューションの特徴とコストを激変**させる (Bill Gates)
- 中小企業は完全ユーティリティモデルにより、大幅な経済的恩恵を受けられる。一方、**大部分の大企業は、過去のインハウスへのIT投資と、ユーティリティの利点を慎重にバランス**しなければいけない
- ユーティリティコンピューティングが**成熟するには、数年**が必要である。もっとも困難なのは、企業が所有するシステムをあきらめさせ、データセンタを使うよう仕向ける点である
- 従来のハードウェア(・ソフトウェア)ビジネスの多くが、単純に失われるだろう

＊クラウド化する世界(ビジネスモデルの大転換)(内田コメント)

Copyright 2008 FUJITSU LABORATORIES LIMITED

企業ITシステムのコモディティ化

FUJITSU



クラウド化 ＝ 構築・運用のコモディティ化

Copyright 2008 FUJITSU LABORATORIES LIMITED

まとめ

FUJITSU

- 企業システムも含めクラウドへの移行は進む
- クラウド利用の形態は様々な可能性がある
- 必要とされている実装技術は幅広い
- 主戦場・技術・プレーヤーが一変する可能性がある

Copyright 2008 FUJITSU LABORATORIES LIMITED

NAREGI関連報告

- NAREGIの現状と今後
- NAREGIのこれまでと今後
- 研究コミュニティ形成のための資源連携技術に関する研究報告

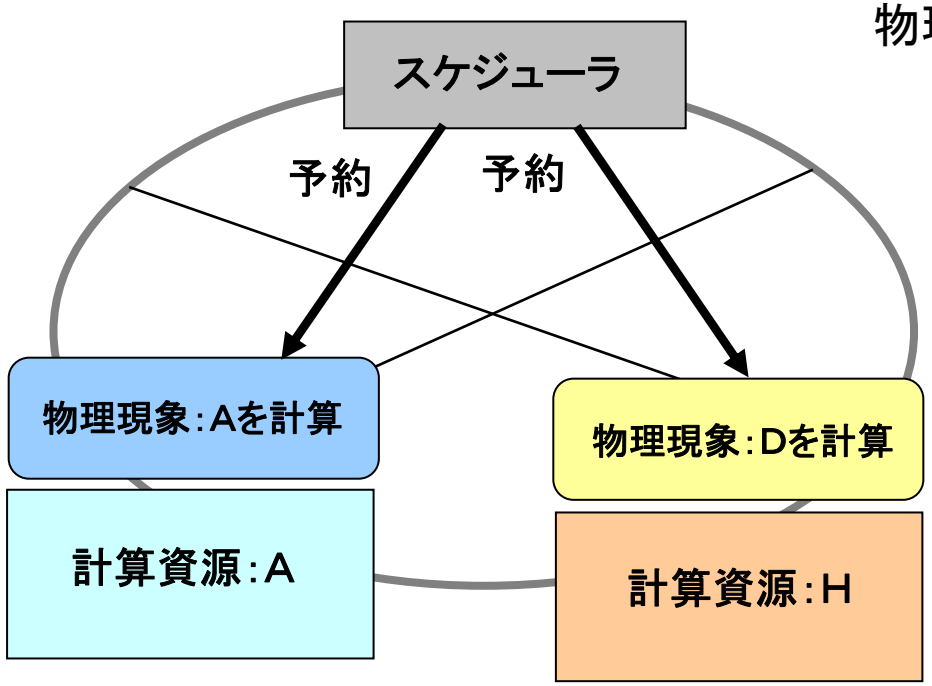
NAREGIの現状と今後
～ Gridのトピックス的な話題 ～

2008年2月22日

国立情報学研究所
リサーチグリッド研究開発センター
宇佐見 仁英

連成解析処理を実現する予約方式

これからの計算科学では様々な物理現象を連携させ、
より現実に近い状態の解析をする連成解析が重要になる



条　件

物理現象Aの計算に適した計算資源Aタイプ
物理現象Dの計算に適した計算資源Hタイプ

1）スケジューラはネットワークに繋がった
計算資源の中からAタイプ、Hタイプを探す。

2）計算資源の予定を調べ同時に使える
時間に各計算資源に予約を入れる。

2）それぞれの計算資源に物理現象を計算
するプログラムをインストールする。

3）時間が来ると予約された各計算資源が
起動し、計算結果のデータをプログラム
に従いやり取りして、計算処理を進める。

グリッドは連成解析の重要な解である

NAREGIミドルで「何が」できる?

- NAREGIミドルは様々な機能を持つが。。。
  - *ユーザ側:*仮想研究組織(VO)の構成、VO毎の仮想化された計算やデータなどの資源(=グリッドインフラ)の透過な利用(利用手順の標準化)、VO毎の研究用Web Servicesの立上げ、大規模なデータ共有、国際VOとの連携、など
  - *センター側:* 計算機やストレッジによるグリッドインフラの管理運用(利用の平準化)、国際グリッドインフラとの連携、VO管理のホスティング、など

ではVOとは正確には何？CSIやグリッドとどう関係がある？センターはどう対応する?

まとめ

1）NAREGIグリッドミドルウェアはサイバーサイエンスインフラ（CSI)構想におけるインフラの構築、コミュニティー形成、計算機利用環境構築のための重要な要素技術となるものです

2）NAREGI Ver.1.0では情報基盤センターでの運用、分野別VOの運用を目指し、信頼性の高いグリッドミドルウェア開発を進めます

3）さらに次世代スパコンプロジェクトで開発されるシステムも含めた階層的な計算研究環境の実現を目指したグリッドミドルウェアの開発を進めます

4）GOC（Grid Operation Center）を設置し、グリッド利用者のための認証局の運用を推進します

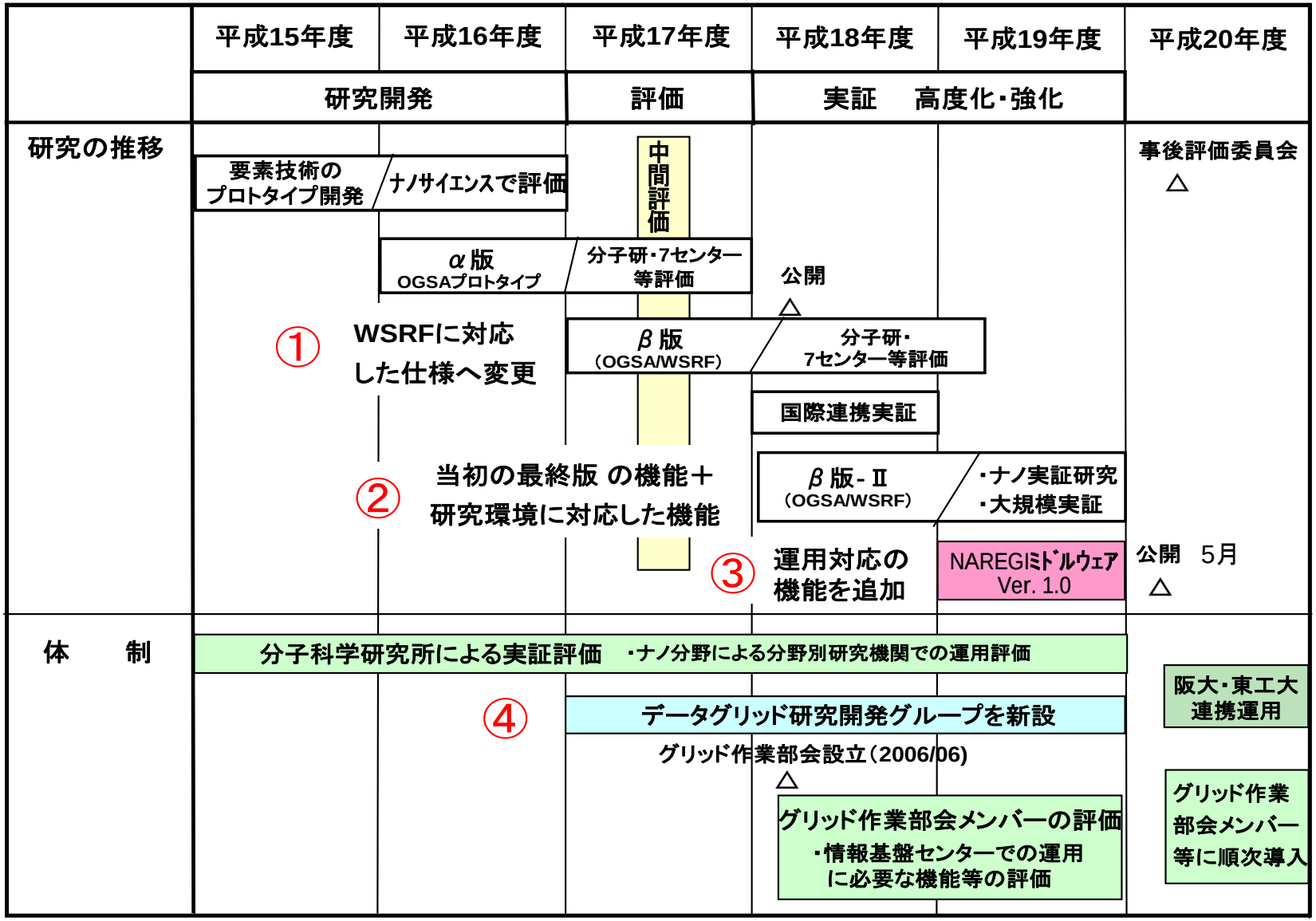
5）さらにNAREGIグリッドミドルウェアの導入支援、利用及び技術支援（ヘルプデスクの設置）を実施します

NAREGIに関する話題提供

- NAREGIのこれまでと今後-

九州大学
情報基盤研究開発センター
青柳　睦

修正後の研究開発経過


平成15年度
平成16年度
平成17年度
平成18年度
平成19年度
平成20年度
研究開発
評価
実証　高度化・強化
研究の推移
要素技術の
プロトタイプ開発
ナノサイエンスで評価
中間評価
α版
OGSAプロトタイプ
分子研・7センター
等評価
公開
①
WSRFに対応
した仕様へ変更
β版
(OGSA/WSRF)
分子研・
7センター等評価
国際連携実証
②
当初の最終版 の機能＋
研究環境に対応した機能
β版-Ⅱ
(OGSA/WSRF)
・ナノ実証研究
・大規模実証
③
運用対応の
機能を追加
NAREGIミドルウェア
Ver. 1.0
事後評価委員会
公開　5月
体　制
分子科学研究所による実証評価　・ナノ分野による分野別研究機関での運用評価
④
データグリッド研究開発グループを新設
グリッド作業部会設立(2006/06)
グリッド作業部会メンバーの評価
・情報基盤センターでの運用
に必要な機能等の評価
阪大・東工大
連携運用
グリッド作業
部会メンバー
等に順次導入

その後の展開

7センターのNAREGIミドルウェア導入へ向けた動き

・平成20年7月17日臨時センター長会議（東大） グリッド研究会の再開を決定
・平成20年8月20日グリッド研究会再開（京大） グリッド配備・運用TFを設置
・平成20年9月17日グリッド配備・運用タスクフォース開催（国情研）
　９演算＋１管理サイトsの「ミニマム構成」「GOC CA」「SINET3 L3/VPN+Com」

グリッド認証局の本格運用へ向けた動き

・OGF24 (平成20年9月15日-) LRA運用を含むIGTFへの提案が認められた

NAREGグリッドミドルウェアV1.1のソース凍結を正式決定

平成20年9月5日グリッド研究開発拠点センター会議（PSL会議）
今後は、
・オープンソースとしての健全な発展を支援する体制の検討
・グリッド研究開発センターも1つの重要なコントリビュータ
・今後、開かれた組織によるV1.2 仕様策定と開発体制、成果認定プロセスの整備
Who is NAREGIアーキテクト?

まとめ

- NAREGIプロジェクトの成果物としてのNAREGIグリッドミドルウェア第１版（Ver.1.0)の開発とオープンソースによる一般公開

  V1.1（平成20年10月公開予定)をもって凍結→オープンコミュニティーにおける発展

- これからの計算研究環境に不可欠な機能である仮想組織（VO) の開発によるサイバーサイエンスインフラへの足掛かりを提案
- 大規模実証研究により、サイエンスグリッドとして情報基盤センター向け機能と分野別研究機関（例：分子科学研究所、KEK,国立天文台）向けの機能を有することを確認→今後、運用へ。
- 国際標準の準拠と貢献、海外の主要なグリッドプロジェクトとの連携による国際的なグリッドミドルウェアとして認知される。
- 認証を含む運用グリッドに向けての方向付けを確固たるものに

  IGTF準拠のGOC－CAの実運用

- 北大・東北大・筑波大・東大・東工大・名大・京大・阪大・九大・国情研、連携による全国共同利用環境への展開と少なくとも3つの検討課題

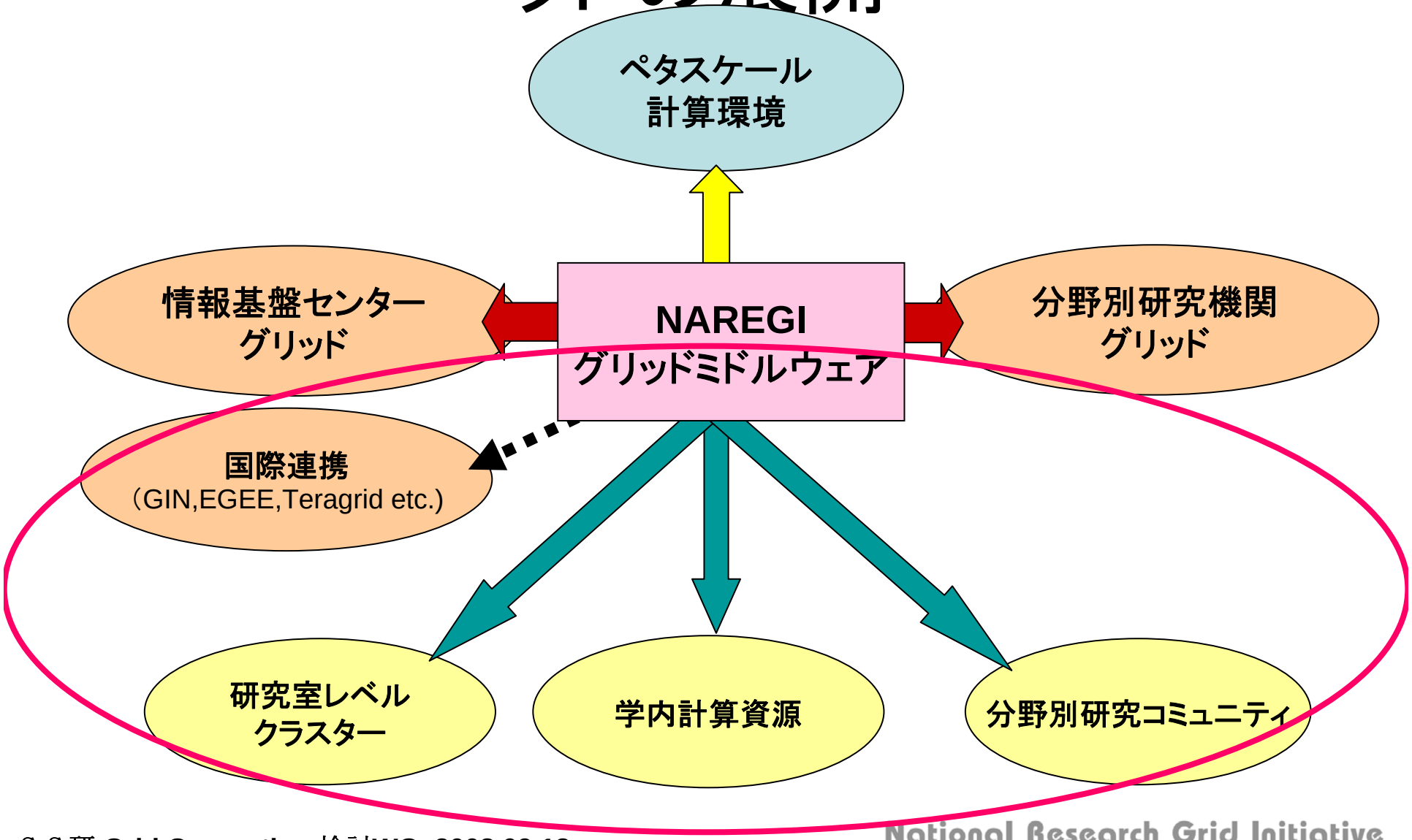
研究コミュニティ形成のための資源連携技術に関する研究

2009年1月21日

研究代表者

国立情報学研究所
リサーチグリッド研究開発センター
教授　三浦　謙一

次世代計算研究環境としてのグリッドの展開



開発の目標

- エンドユーザが研究室のPCから、データ検索、データ解析、シミュレーションなどをローカルな資源とグリッド上の資源を問わず利用できる。
  ここで資源とは、各種計算機、データ格納装置、データ類、データベースなど
- どのグリッドミドルウエアも同じユーザ環境(アプリケーション開発ユーザ向け)で利用できる。
- 複数のグリッドミドルウエア下にある資源も必要に応じて、同時に利用できる。
- 大学研究室では、グリッドミドルウエアに関するノウハウや保守を必要としない。
- OGFで、既に標準化を目指しているSAGAをベースに開発し、足りない機能(資源情報、RNS参照など)は独自に開発する。
  SAGA: Simple API for Grid Applications
- ファイルカタログ情報は、RNSをベースに検索する。
- アプリケーション開発ユーザには、Python環境で提供する。

NAREGIなど適用実験・運用

- NAREGI PKIをめぐる周辺
  ー連携実証実験での事例ー
- 高エネルギー分野でのGrid適用事例
- NAREGIミドルウェアによる6拠点連携実験
- 九州大学導入報告
- 名古屋大学導入事例
- ９大学情報基盤センターへのグリッド配備・運用

NAREGI PKIをめぐる周辺
～ 連携実証実験での事例～

2008年2月22日

国立情報学研究所
リサーチグリッド研究開発センター
友石 正彦

NAREGIミドルウェア 証明書の種類

- ホスト証明書
- ユーザ証明書
- VO(属性)証明書
  - ユーザ証明書にVOの属性情報を追加
  - VOMSのホスト証明書を利用
- (委譲)プロキシ証明書 – 実際に使われる
- コンテナ証明書
  - ホスト証明書を利用

NAREGIミドルウェア ユーザ証明書

- ユーザ証明書 + VO(属性)
  - 計算リソース毎に受け入れVO情報
  - 計算リソースの gridmapファイルで
    ユーザDNとローカルアカウントとの対応情報
  - 上位/ポータル側に伝達
- 各ノードにルート証明書

- 認証局が発行
- UMSに保存
- 各ノードでプロキシ証明書として利用

異なるポリシー(発行)

- (何といっても)発行するとき… 対面確認は
  - いる？ - 世界(APGrid)にあわせる
  - いらない？ - 共通アカウントからの実績
- 発行した証明書、鍵のあつかい
  - ユーザ本人 + ソフトウェア
  - ユーザ本人 + ハードウェア(トークンの中のみ)
  - UMSなどシステムで管理
  - システムから出さない(発行以上の規則)
- (計算機の利用だけなら、そもそも証明書が面倒!)

大阪大学認証局 ユーザ証明書

- 大阪大学認証局の発行したユーザ証明書は、
  - すべてUMSへ、and 出さない CP
  - ユーザ証明書、鍵を厳密に管理
    - 証明書を意識させないポリシーでは、かなり納得。
- 東工大ポータルでのユーザ証明書は別に必要
  - ホスト証明書は大阪大学
  - ユーザ証明書はNAREGI

Phase 2.2 PKIドメイン

- 計算資源、ポータル
  - 受け入れるジョブの証明書発行認証局のドメインに所属
    - ホスト証明書と異なる証明書受け入れ。
    - つまり、ポータルも複数のドメインに所属
- (受入れ条件は？)

- 異なるドメインのユーザ証明書の受け入れも可能？
  (複数のドメインを意識した運用?)

- 実験は現在は、東工大でジョブが動いたところまで。

まとめ

- 組織(学内)PKIは進展中
  - Gridは大きなアプリケーション
  - これまでのユーザレベルで理解されているとは限らない
  - ((ユーザ)秘密キーをちゃんと管理してもらえるか。。)
  - (UMSは大きな「セキュリティトークン」?)
- 発行ポリシーは一つとはかぎらない
  - APGrid vs ...
  - (専門家は厳しく、現場は(ユーザに)やさしくなりがちな感じ)
  - ユーザ毎に違う範囲の機械を使うことになるかも
  - 決定打はない?
- 柔軟な受入れポリシーの検討を
  - 相手CP/CPSがあれば、明示的に検討可能
- 実証実験では、阪大発行証明書とNAREGI発行証明書
  - 計算ノード以外での双方の ルートCAでの運用も

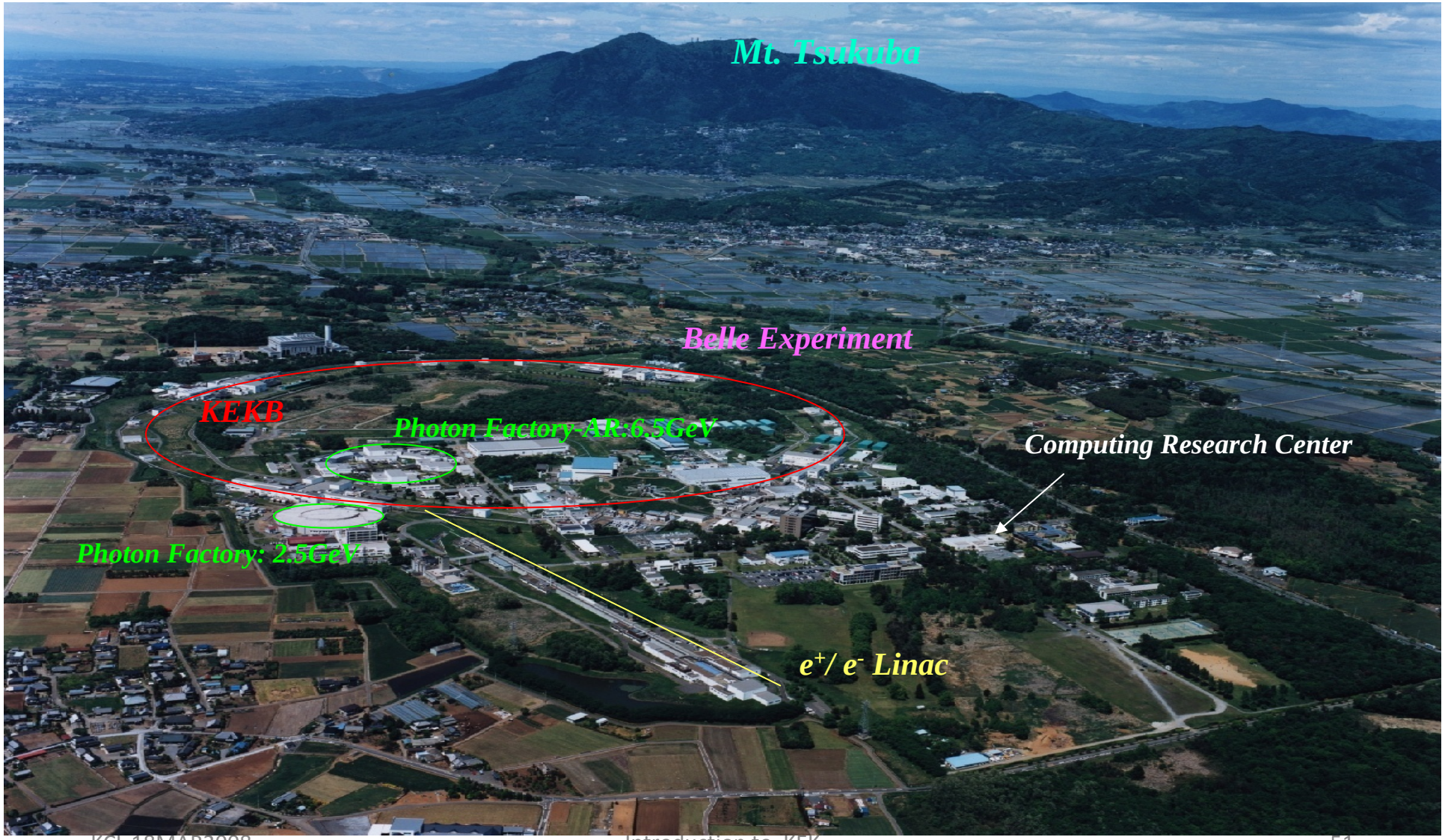
GRID activities and future plans at KEK

Takashi Sasaki
KEK Computing Research Center
KEK: Inter University Research Corporation
High Energy Accelerator Research Organization

Basic strategy

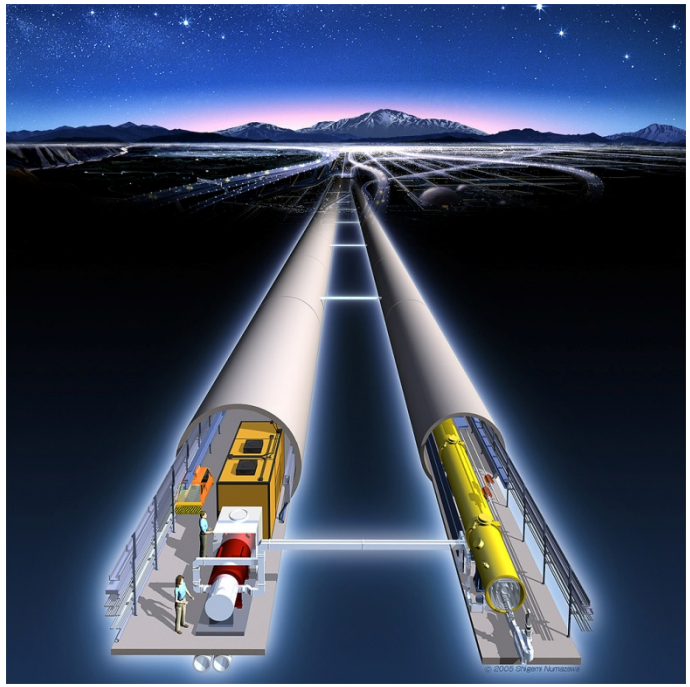
- GRID deployment and operations for projects at KEK
  - Belle, J-PARC, ILC and so on
- LHC regional center
  - Univ. of Tokyo is operating tier-2, but no tier-1 in Japan
  - Preparing for future tier-1 level GRID center operation at KEK toward ILC
    - Depends on the decision of the society
    - Technical assessments and staff training
- University support
  - Help for deployment and operations of LCG at universities out side of WLCG
- NAREGI
  - Our human resource is very limited and we want to depend on the operation infrastructure of NAGEI if possible
  - Would help university people and non-HEP users

KEK (Tsukuba site)



ILC

- International Linear Collider
- 素粒子の質量の起源、時空の対称性、宇宙創成などの根源的な疑問を解明する為に国際的な協力で建設する計画の最高エネルギー電子・陽電子衝突の線形相互衝突型加速器
- 日本への誘致を目指している
- http://www.linear-collider.org



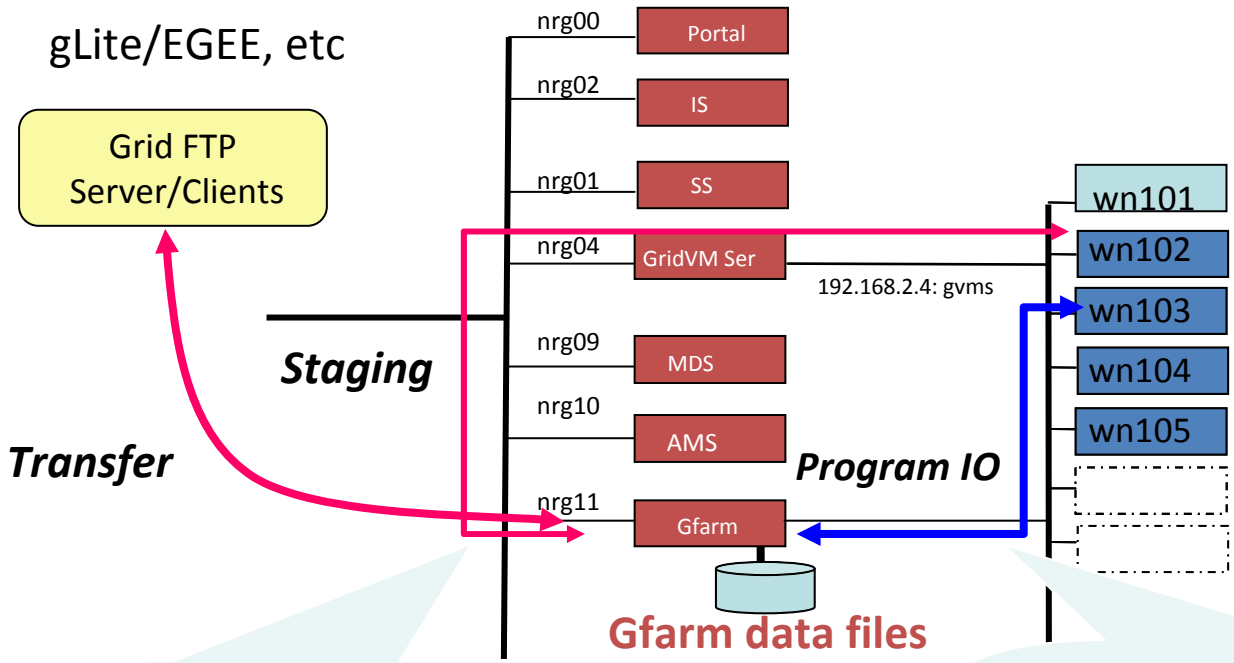
Why GRID?

- 国際共同研究
  - 手元にデータがほしい
    - 研究効率の改善
- 処理すべきデータ量、計算量が膨大になりすぎ、分業する必要に迫られている
  - 解析そのものよりも、シミュレーションに必要な計算量が膨大
    - 収集する事象数以上は最低でも必要

NAREGI-β at KEK

- Testbed: 9 server nodes + 5 compute nodes
- Middleware installation
  - NAREGI-beta 1.0.1 : Jun. 2006 - Nov. 2006
    - Manual installation for all the steps
    - Confirmed functionalities of Information Service, PSE, WFT, GVS
  - NAREGI-beta 1.0.2 : Feb 2007
    - DG comprehensive installation manual was released in Jan. 2007
  - NAREGI-beta 2.0.0: Oct. 2007 apt-rpm installation
  - NAREGI-beta 2.0.1: Dec. 2007
- Site federation test
  - KEK-NAREGI/NII :Oct. 2007
  - KEK-National Astronomy Observatory (NAO): March 2008

DataGrid:Gfarm File access



- Data Staging
- File transfer
  - Grid FTP
- Program IO
  - with Gfarm Fuse
  - Data analysis program can read/write data files without any modification

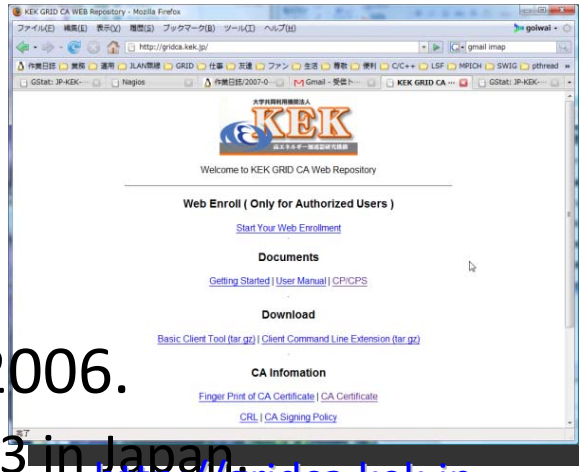
Transfer URL through Gfarm-DSI:
gsiftp://nrg11.cc.kek.jp/gfarm/user01/data1-22.dat

Program access name space for VO:
/gfarm/user01/data1-22.dat

Summary

- Many kinds of GRID middleware have been deployed and in operation
  - Related R&D is also going on
- Future strategy for the regional center will be decided by the Japanese HEP society
  - Super B, T2K, ILC, LHC and etc.
- In our experience, human network is much more important than computer network.

KEK Grid CA



http://gridca.kek.jp

- KEK Grid CA has been started since Jan 2006.
  - 75 CAs are in production all over the world, 3 in Japan.
    - AIST, NAREGI, and KEK
  - accredited as an IGTF (International Grid Trust Federation) compliant CA
- KEK Grid CA has been audited by Yoshio Tanaka (chair of APGridPMA), AIST on May 2007

KEK Grid CA: Statistics of Issued Certificates

	JFY2006 Apr 2006 - Mar 2007	JFY2007 Apr 2007 -
Globus Client Certificate (Personal cert.)	68	119
Globus Server Certificate (Host cert.)	139	238
Web Server Certificate	4	0

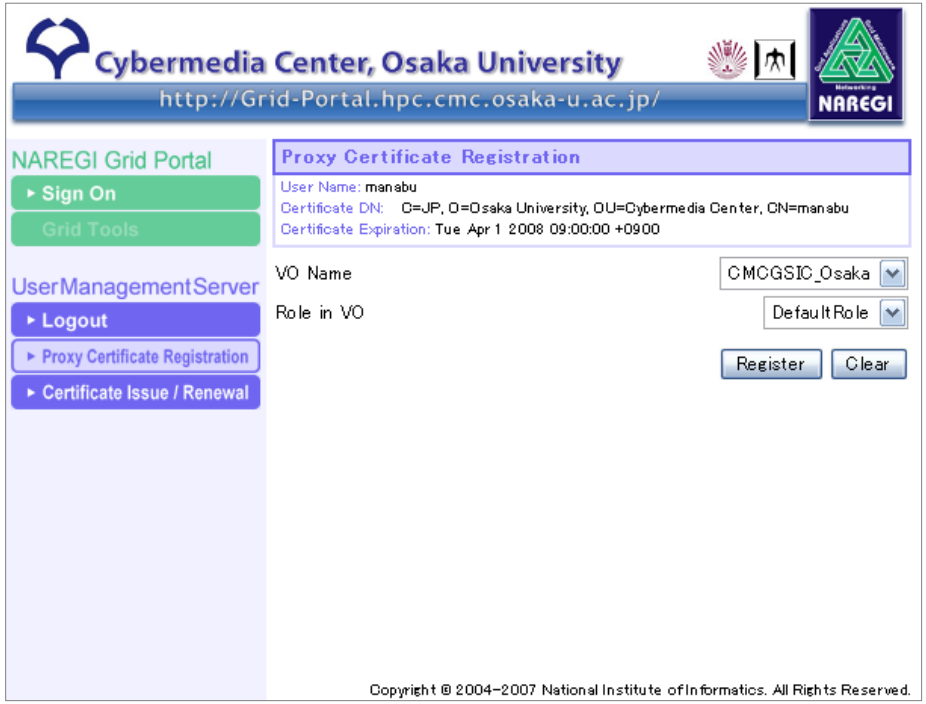
NAREGIミドルウェアによる 6拠点連携実証実験

Manabu Higashida
manabu@cmc.osaka-u.ac.jp

2008/06/09

SC07 Demonstration at Booth#3059

- ✓ すべての共同利用ユーザにGrid証明書を発行します!
- ✓ すべての計算機資源をGridに提供します!
- ✓ 東工大GSIC-TSUBAMEと連携しています!



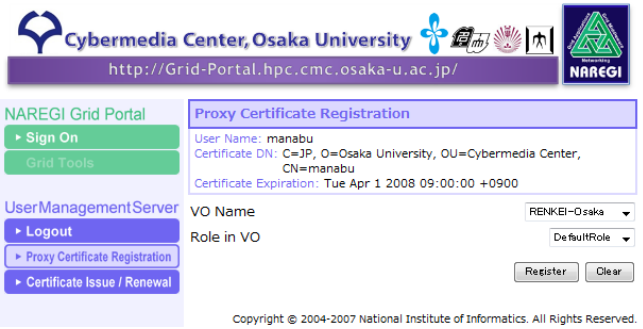
“NAREGI 100T Project” (仮称) の課題: SC07の反省

✓ とにかく一度はきちんと「仕様どおり」に動かしてみたい

- GridVMのことはかなりよくわかった
  - ローカル・スケジューラのことを熟知しているから
- IS (Information Service) も何故かよくわかった
  - OGSA-DAI≒SQLだから
- SS (Super Scheduler) のダークサイドぶりが際だつ
  - まるでちゃんと動いている気がしない･･･
    - “Missing” ：落ちるとジョブが消失、ただしローカル・スケジューラにはジョブが残ったまま･･･
    - “Exception” : 何が原因で落ちたか一切不明･･･
  - このままでは何がどう「仕様どおりでないか」すらフィードバックできない･･･

✓ 再挑戦するなら「連携実証」の名に相応しい連携をやってみたい

- ポリコム分配器が許す限り多くの拠点で連携したい
  - 阪大、東工大、九大、国情研、分子研、名大
- ミドルウェアの連携機能を実証したい
  - IS 連携
  - SS連携 (V1の機能を先出し導入)



なぜ “NAREGI” の評判は良くないのか!?

- ✓ そもそも認証局がないと評価できない
- ✓ とにかく安定して動かせない
  - 設定が煩雑
    - 度重なる仕様変更
      - 設定ファイルの記述の更新が追い付いていない･･･
      - 最後の最後でSSクライアントの仕様変更で、それまで無事動いていたGridRPCが動かなくなる･･･
  - サーバが不安定
    - 短期間の統合試験で動作検証を成功させるための甘い実装が「2000年問題」化している
    - 統合試験では検証できない複雑な構成を組むと落ちる:「分子研問題」

認証ポリシーが異なるセンター間の相互連携

- 複数の認証局が発行した証明書を利用できるNAREGI計算機資源環境を構築する
- 実際に運用中の計算機センターの大規模資源に対して、NAREGIミドルウェアからジョブ投入できる環境を構築する

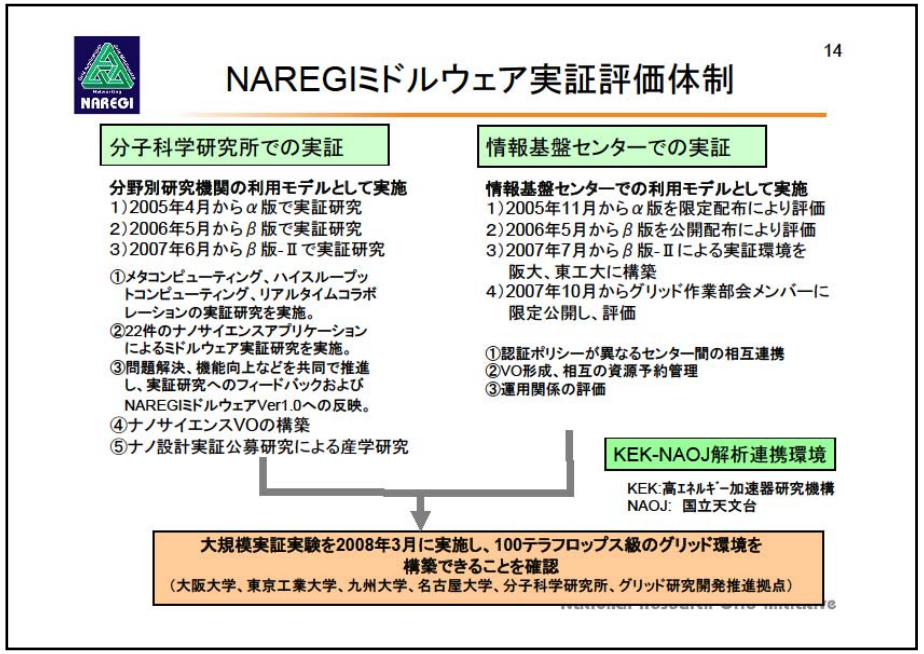
VO形成、相互の資源予約管理

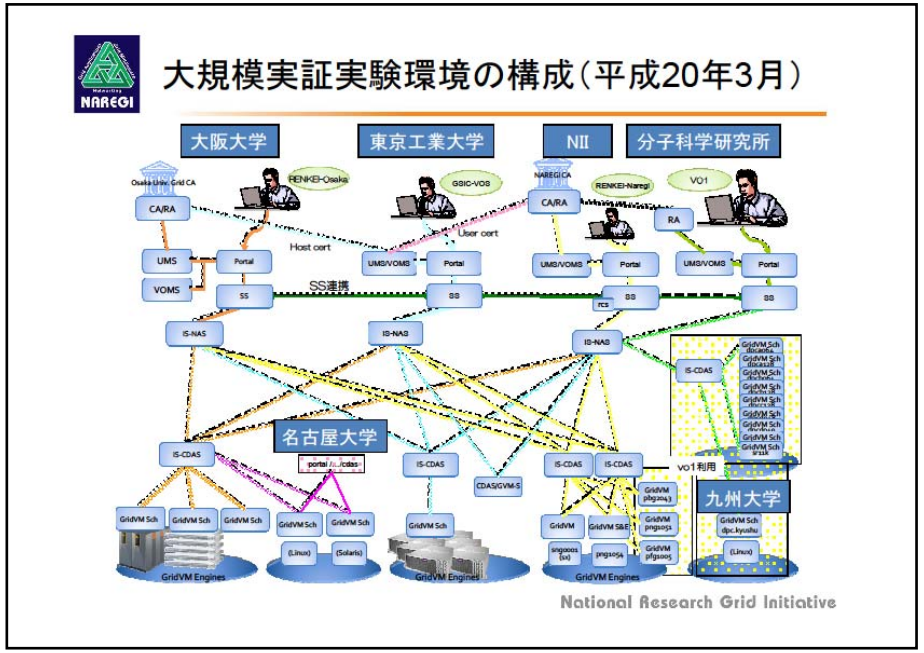
- 各拠点から提供された計算機資源のAUP (利用規定) に対応するVO (仮想組織) を形成する
- 複数のメタスケジューラが他スケジューラの資源予約の状況を反映した資源予約を行ったうえで、実アプリケーションによるジョブ投入ができる環境を構築する

運用関係の評価

- 実運用環境に展開するに先だって支援体制の実地評価を行う
  - GOC (Grid Operation Center)
  - PERT (Performance Enhancement and Response Team)

NAREGIミドルウェアで100TFLOPS級のグリッド環境を構築できるか!?





動き出したサイエンスグリッドNAREGI
—研究リソース共有の世界を広げるミドルウェアを公開—
平成20年5月9日付けプレスリリースより抜粋

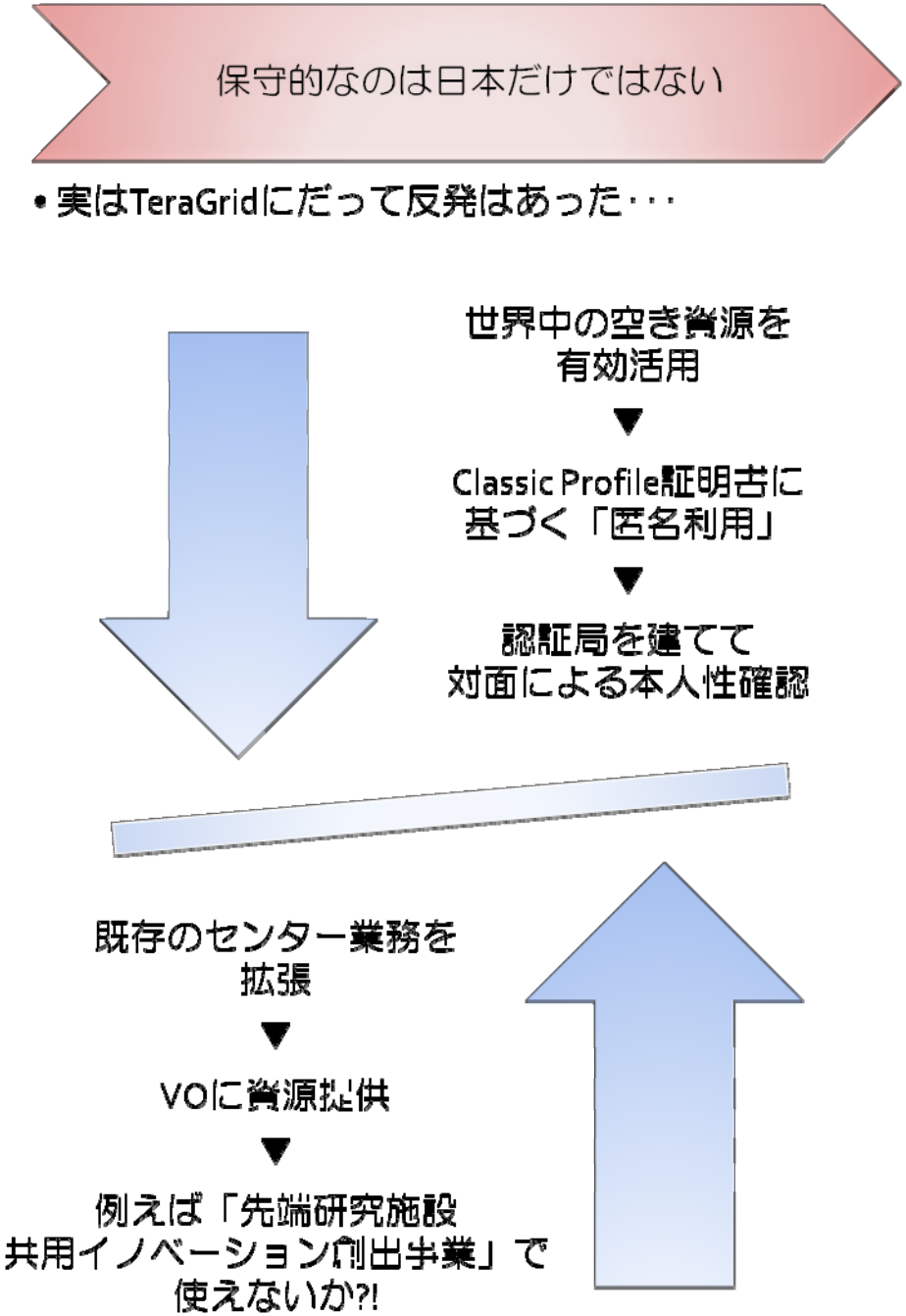
グリッド認証局

- ✓ NII/NAREGI認証局 (次年度はNII/GOC認証局)
  - APGrid PMAプロダクション・レベル
    - クラッシック・プロファイル
- ✓ 阪大CMCグリッド認証局
  - APGrid PMA実験レベル
    - 限りなくプロダクション・レベルにある
      - 登録局での対面による本人確認を要求していない (運用上無理だから)
    - APGrid PMAにMICSプロファイルの導入と適用を交渉中・・・
  - 阪大CMCの全国共同利用登録者を対象に "1-Click"でグリッド証明書を発行
    - 既存の全国共同利用の窓口が登録局として機能
      - 煩雑な業務処理を自動化し、事務員からも利用者からも隠蔽
    - 抜け駆けしたわけではないけれど、気がつけば基盤センターではうちだけ?!
      - CA/RA室の維持、人員の配置 (最低6名) など相応のコストが発生・・・

原因の一端は、国内の計算機センター組織の構造にもある。米国でも欧州でも、一般に大学組織内でのグリッド管理は大学内の計算機センターが担う。しかし、計算機センターの性格が、国内と欧米とでは多少異なるように思える。欧米の大学計算機センターはそれ自身が独立した研究組織として機能しており、学生やポスドクを含め、多くの計算科学者を内部に抱え、新しい技術の開発や導入を独自に行う地力を備える。これに対し、（すくなくともこれまでの）国内の計算機センターの多くは、研究者を内部に持たず、ベンダーから派遣されるSEに依存し、**既存の計算資源を既存の方法で運用するだけ**の硬直した機関になっていたのではないか。

科学技術計算の需要は今後も増え続けることが予想される。これに対して有限の資源で答えるためには、グリッド的な技術を用いて、**資源を適切に配分**することで、有効に活用していくしかない。そのための大学での受け皿は、やはり各計算機センター以外にはないだろう。各計算機センターに適切な拡充を行い、グリッド運用の積極的な主体とすることが、今後の計算科学の発展に必要であると考える。

中田秀基，「グリッドのこれまでとこれから」
京都大学学術情報メディアセンター全国共同利用版［広報］
Vol.6 No.2 (Oct. 2007)



大学間NAREGI運用実験

九州大学　南里豪志

NAREGI運用実験：　経緯

- グリッドコンピューティング研究会において
  7センターへのNAREGI配備・運用に向けた検討開始
  - 2008年7月

- グリッドタスクフォース発足
  - グリッドコンピューティング研究会の中
    - 主査：青柳グリッドコンピューティング研究会主査

  - 参加大学：　7センター＋東京工業大＋筑波大＋NII
    - 東京工業大：大阪大学との連携運用デモを実施済み
    - 筑波大：T2Kの枠組みでグリッドミドルウェアの使用を検討中

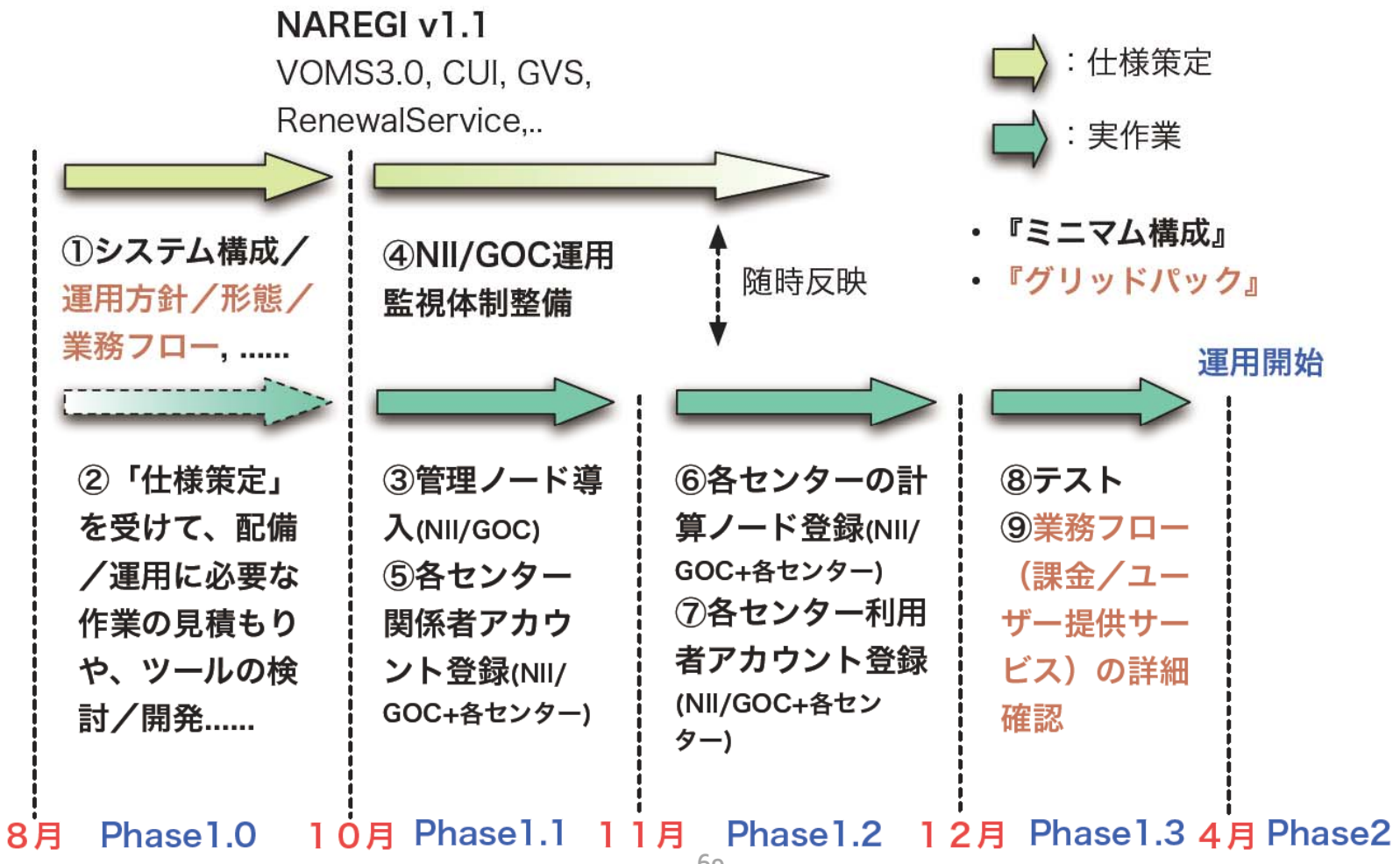
NAREGI運用実験：　目標

- 平成20年度末までに、各参加大学の計算機環境にNAREGIミドルウェアを配備し、運用の準備を整える。
  - 「ミニマム構成」による運用実験の実施
    - NAREGIの基本機能を用いた最もシンプルな構成を選択
  - 「グリッドパック」の策定
    - 利用者の利用申請から証明書発行、利用開始を経て利用終了までの業務フローと実施体制

各大学の計算機資源

	機種名	OS	スケジューラ	分離型／一体運用型	演算性能 (GFLOPS)
北海道大学	HITACHI HA8000 DELL PowerEdge R200	CentOS 5	PBS Pro9	分離型 30ノード(約60CPU)	638.4
東北大学	NEC SX-9	SUPER-UX ver18.1	NQSⅡ ver2.50 Job Manipulator ver1.20	一体運用型 最大4ノード(64CPU)	6,553.6
筑波大学	Appro XtremeServer-X3	RHEL 5	Grid Engine	一体運用型 最大32ノード(512CPU)	4,710.4
東京大学	HITACHI HA8000-tc/RS425	RHEL 5	日立OSCNQS または独自インストール	分離型 8ノード(128CPU)	1,177.6
東京工業大学	Sun Fire X4600	SLES 9 (2008/10 SLES 10 SP2 に移行)	Sun N1 Grid Engine 6.1 東工大カスタマイズ版	一体運用型 未定	
名古屋大学	Fujitsu PRIMERGY RX200 S3/S2 Fujitsu PRIMEPOWER HPC2500	CentOS 5.2 Solaris 8	PBS Pro9.1 Parallelnavi 2.3	6ノード(12CPU) 2ノード(64CPU)	490.8
京都大学	Fujitsu HX600	RHEL AS 4	Parallelnavi 3	分離型 4ノード(64CPU)	588.8
大阪大学	NEC SX-8R NEC SX-9R NEC Express 5800/120Rg-1	SUPER-UX ver17.1 SUPER-UX ver18.1 SLES 10.1	NQSⅡ ver2.40 Job Manipulator NQSⅡ ver2.50 Job Manipulator NQSⅡ ver2.40 Job Manipulator	一体運用型 最大8ノード(64CPU) 一体運用型 最大8ノード(128CPU) 一体運用型 最大32ノード(128CPU)	16,896.0
九州大学	Fujitsu PRIMERGY RX200 S3	RHEL WS 4	PBS Pro9.2.2	分離型 12ノード(48CPU)	576.0
				合計	31,631.6

スケジュール


NAREGI v1.1
VOMS3.0, CUI, GVS,
RenewalService,..
：仕様策定
：実作業
①システム構成／運用方針／形態／業務フロー, ...
④NII/GOC運用監視体制整備
随時反映
• 『ミニマム構成』
• 『グリッドパック』
運用開始
②「仕様策定」を受けて、配備／運用に必要な作業の見積もりや、ツールの検討／開発.....
③管理ノード導入(NII/GOC)
⑤各センター関係者アカウント登録(NII/GOC+各センター)
⑥各センターの計算ノード登録(NII/GOC+各センター)
⑦各センター利用者アカウント登録(NII/GOC+各センター)
⑧テスト
⑨業務フロー（課金／ユーザー提供サービス）の詳細確認
8月 Phase1.0
１０月 Phase1.1
１１月 Phase1.2
１２月 Phase1.3
４月 Phase2

ミニマム構成

- NAREGI v1.x の必要最小限なコンポーネントを選択
  - 今回，データグリッドは保留
- VOは 1つのみ
  - Portal, IS, SS等の管理ノード群も全体で 1組
- 各センターは演算ノードのみを提供
  - 管理ノード群は GOC(NII)で運用
  - 演算ノードは実運用から切り出しても良い
- Firewall問題回避のため，L3-VPNによる専用線を利用

グリッドパック

- グリッド運用に向けた業務フローの策定
  - 利用申請〜利用開始〜利用終了

- 基本方針
  - 今までのセンターでの業務フローを出来るだけ踏襲
  - 利用者は1回の手続きで利用可能に
  - GOC: センター間に跨がる事項への対応を担当
  - 課金については平成21年度以降に検討
    - 平成21年度は課金無しで試験運用

導入時の主なトラブル

- PBS Proでジョブが実行されない
  - PBS Proのライセンスをホスト名で登録していた
    - NAREGIの全ノードは hostnameとして FQDNを設定するため、PBS Proのライセンスを認識できなかった
  - ライセンスファイル中でFQDNに書き換えて解決
    - ライセンスの再取得は不要
- NATが使えない
  - IPフォワーディングが無効になっていた
    - RHELのデフォルト設定
- あとは、ルーティングの設定ミス程度

まとめ

- NAREGI運用実験
  - 「ミニマム構成」を選定
  - 各大学で作業中
- グリッドパック
  - 実運用に向けたワークフローの策定
- 九大センターでの実験環境構築
  - インストーラにより、ほぼ順調に進捗

九州大学における NAREGI実験環境

2009年8月18日　富士通SS研グリッドWG

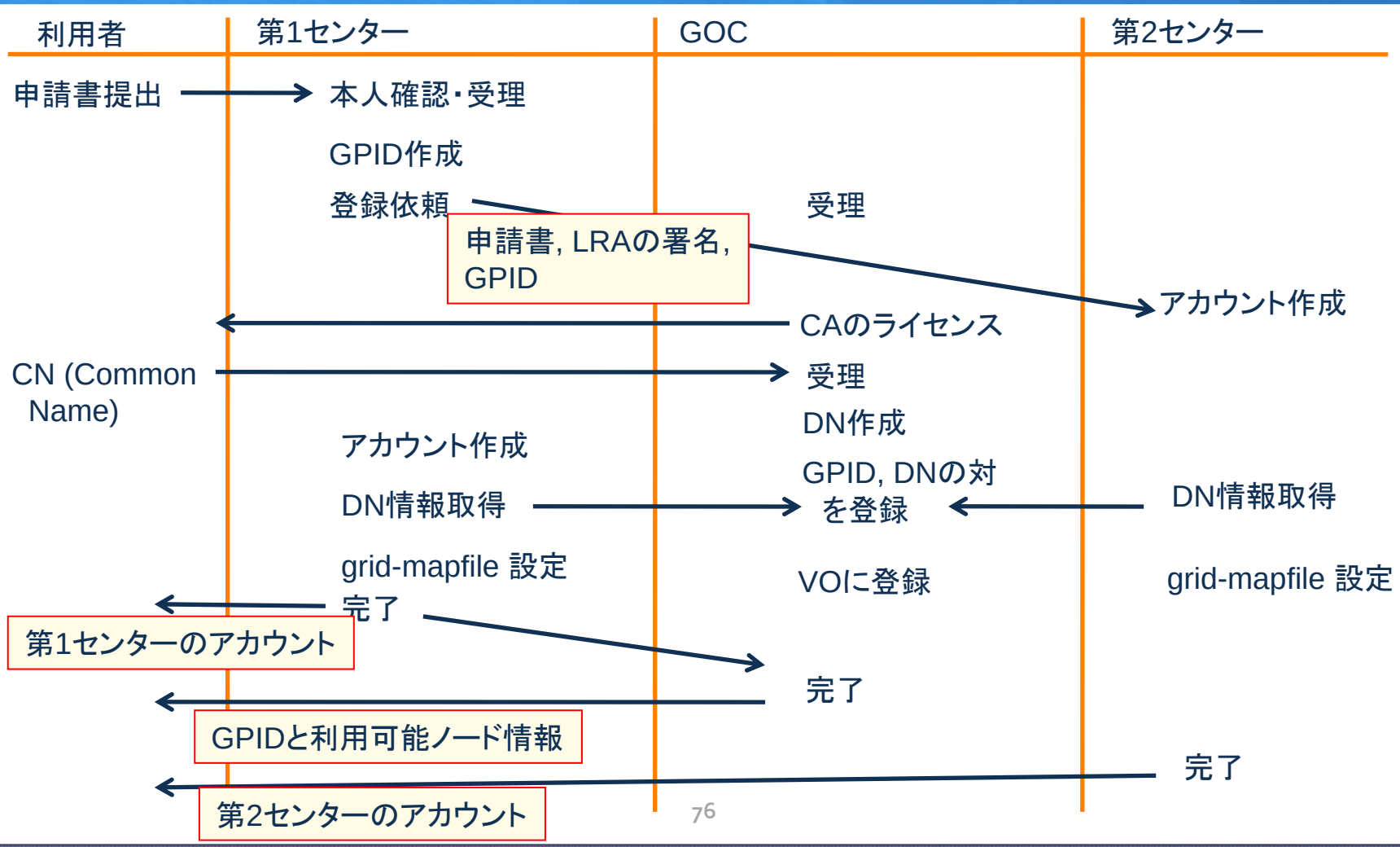
九州大学　南里豪志

NAREGIの改良

- 九大の環境（RedHat Enterprise Linux 4 x86_64）のサポート追加
  - NAREGIの全コンポーネントに対して追加
    （Portal, IS, SS, GVMS, GVME etc.）

- 現バージョンNAREGI-1.1.3のサポート対象：
  - CentOS5 (i386, x86_64(32bitで動作))
  - RedHat Enterprise Linux 5 (i386, x86_64(32bitで動作))
  - RedHat Enterprise Linux 4 (x86_64)
  - openSuSE10 (x86_64)

新規利用申請の流れ


利用者
第1センター
GOC
第2センター
申請書提出
本人確認・受理
GPID作成
登録依頼
受理
申請書, LRAの署名, GPID
アカウント作成
CAのライセンス
CN (Common Name)
受理
DN作成
アカウント作成
GPID, DNの対を登録
DN情報取得
DN情報取得
grid-mapfile 設定
VOに登録
grid-mapfile 設定
完了
第1センターのアカウント
完了
GPIDと利用可能ノード情報
完了
第2センターのアカウント

DN更新情報取得ツール (gmap_tool1)と grid-mapfile更新ツール(gmap_tool2)

- 1日1回程度実行し，GOCにおける利用者情報の変更を九大のシステムに反映させる
- 基本的に主導で起動し，毎回パスフレーズを入力
  - 別途，安全な接続が確保できれば cronによる自動化も可
- GOCとのDN情報の共有にはバージョン管理システムSVN(Subversion) を利用

まとめと今後の予定

- NAREGIミニマム構成実験
  - 基本的な動作を確認
  - いくつかの便利なツールを開発
    - インストールツール
    - DN情報更新ツール
- 今後のグリッドタスクフォース全体の予定：
  実運用に向けた課題の抽出とその解決
  - 課金回り
- 九州大学の予定：
  - SR16000の対応
    - 主にインストールツールの調整

【SS研Grid Computing検討WG ‘09.1.21】

NAREGIグリッドミドルウェアによる グリッド環境の構築

名古屋大学情報連携基盤センター

津田　知子

- NAREGIグリッドミドルウェアV1による実験環境の構築と問題点
- １０機関によるグリッド配備／運用計画と接続での問題点

NAREGIグリッドミドルウェアでの問題点（その１）
＜NAREGI V1.0＞
--- 2008.8 ～ 2008.10---

項番	内容	原因と対処
１	PSEのメニューでAction Status Browse the action statusを選択したときエラーとなる。	pse_dbというデータベースが存在していないため。（未解決）
２	Register Applicationでプログラムの登録ができない。	同上
３	WFTにてWorking Directoryを指定するとジョブがException。	GridVM管理ノードのファイル（自動生成）に「” 」が2つあり、grid-mapfileの内容と異なったため。 DNにスペースを含んでいたため。 O=Nagoya University
４	ISのCPUグラフが表示されない。	grid-mapfileの記述をemailAddress=からE=に変更
５	ISでCPUグラフ表示をし、Sign outした後、再度グラフを表示させると、HTTPステータス500のエラー。	V1.1では解決していた。

NAREGIグリッドミドルウェアでの問題点（その２）

＜NAREGI V1.1＞
--- 2008.11 ～

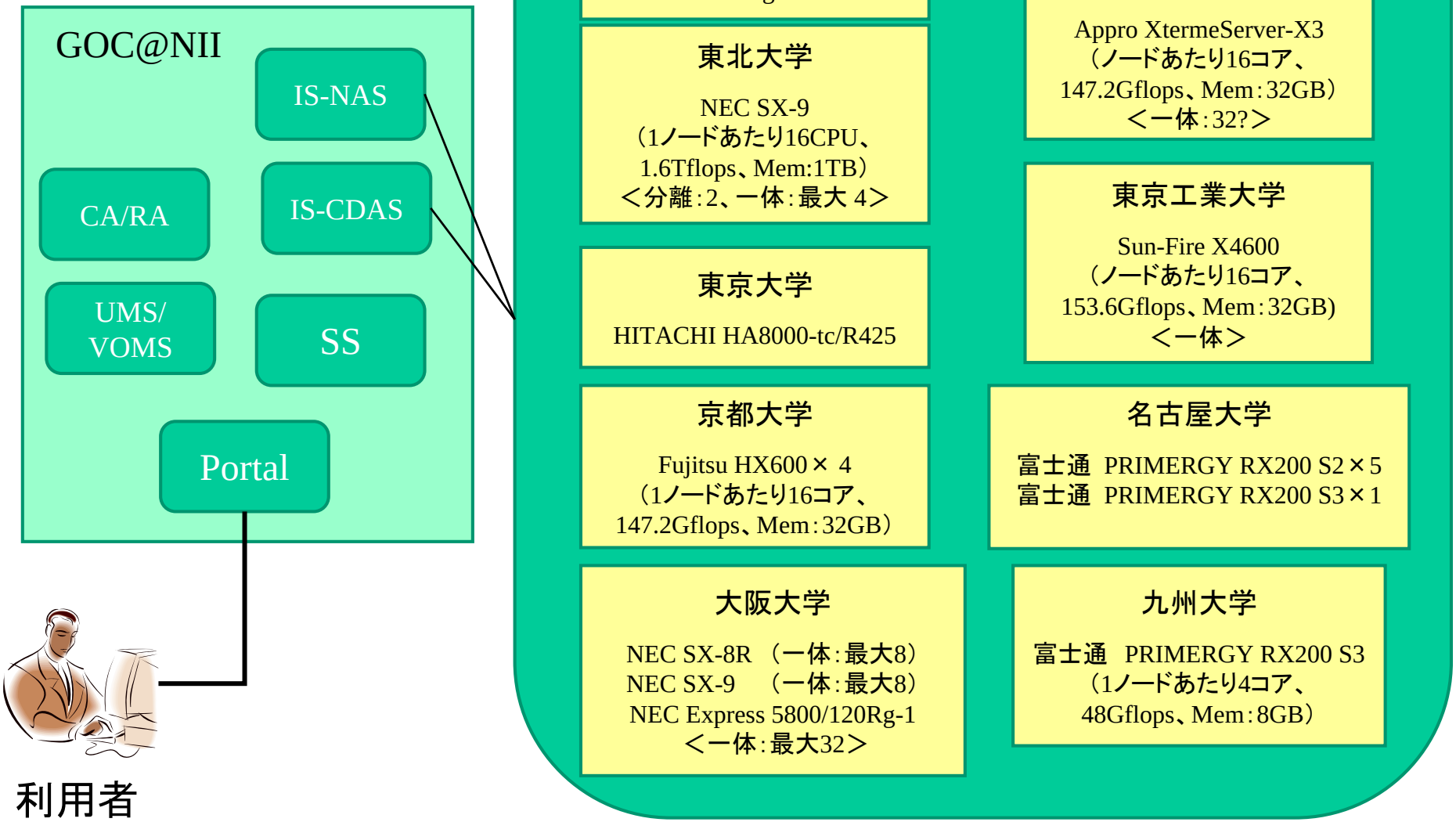
項番	内容	原因と対処
１	ジョブをsubmitすると”Cannot get proxy Certification from MyProxy,”のエラー	V1.1のUMSは証明書にE-Mailの項目が含まれていると、正常に動作しない。UMSノードが動作しているScientific Linuxと他のノードが動作しているCentOSのOpensslのバージョンの違いが原因。 E=を指定しない証明書を用いることで対処。
２	PSEのRegister Application の Resource Requirement でShow Serversを行うとエラー。また、Compile Applicationでも同様のエラー。	証明書のDNの項目にスペースが含まれる場合と最新のPESを使用している場合に発生。 PSEのアップデートとISハンドルの再作成で対処。

項番	内容	原因と対処
３	NIIのホスト証明書を取得時に“SSL Send alert:[Fatal alert] BAD_CERTIFICATEのエラー	証明書取得にあたってNIIのRA証明書をインポートすべきところをCA証明書をインポートしていたため
４	ISでHTTPステータス500のエラー	メモリ不足が原因か？ Portalで一度ヒープメモリ不足の現象が発生すると、回復手段なし。 サービスの再起動で対処。 Portalノードの推奨メモリは2GB以上。名大センターのPortalのメモリは2GB。

＜Data Grid＞

項番	内容	原因と対処
1	1GB のファイルのアップロードでエラー。	DGAMS設定ファイルのファイルのアップロード時のファイルサイズ上限値のパラメータにて対処。 1GB → 2GB
2	1GB のファイルのダウンロードで約25%のところで必ず失敗。	・メモリリークの可能性のあるモジュールの修正を行ったが未解決。 ・GridFTP Clientにてタイムアウトエラー発生のため。モジュールを置き換えて解決。
3	Datagrid Access Managemnt Systemにアクセスできない。	dguftノードのファイルシステムが100%になっていたため。 データグリッドで転送したファイルがゴミとして残っていたため（修正を検討中）。

グリッド配備／運用計画
（平成21年度実施予定）



第1回グリッド配備・運用タスクフォース会議（‘08.9.17)資料より

【SS研Grid Computing検討WG ‘09.8.18 】

名大での
グリッド配備の現状

名古屋大学情報基盤センター
津田　知子

アップデートの手順

- ◆ 全てのノードでOSのupdateをコマンドベースで実施
- ◆ インスローラによる演算ノードの追加を実施

↓

インストーラによる演算ノード群のインストール処理が飛ばされた。

↓

GVMSの設定ファイルのパラメータが空白となった。

↓

IS-NASノードのデータベースが情報を正常に取得できなかった。

↓

- □ NII@GOCへの接続不可
- □ 名大のNAREGI管理システムで不具合発生
  - ➢ VOのリストが初期化された。
  - ➢ Portalノード上のデータグリッドに関係したファイルが初期化された。
  - ➢ WFTでユーザが作成したWFT用のファイルが初期化された。
  - ➢ 名大の演算ノードでGridMPIの実行不可。

NII@GOCの管理ノードへの接続での問題点

＜前回のWGでの報告＞
基盤センターVO(csi-grid-cc)に名大の演算ノード群追加。

- NII@GOCから名大の演算ノードが見えない（LRPSも認識しない）。

（原因１） 名大のGridVM管理ノードがフィルタリングしていた。

（原因２） インストーラによるアップデートの手順に問題があったことにより、GVMSの設定ファイルのパラメータが空白となり、IS-NASノードの　データベースが情報を正常に取得できなかった。

【グリッドパック申請処理】

通常のアカウント登録処理に以下の処理を追加

- LRA担当者による本人確認
- Gridユーザ情報の管理
  グリッドパック整理番号（GPID）とローカルアカウントの対応付け
- GOCへの書類発送
- 第二センターへの申請書類の送付
- GOCのDBからのDN情報の取得(gmap_tool1)
  GOCではDN情報はsvnで管理
- grid-mapfileの編集(gmap_tool2)
- GOCへの設定完了通知の送付

（ここでの書類送付は、メールにて実施）

九大の小林先生が作成されたツールをセンター用にカスタマイズして使用

【名大へのグリッドパック申請状況】（8/17現在）

第1センター登録： 2
第2センター登録： 20
grid-map登録完： 21

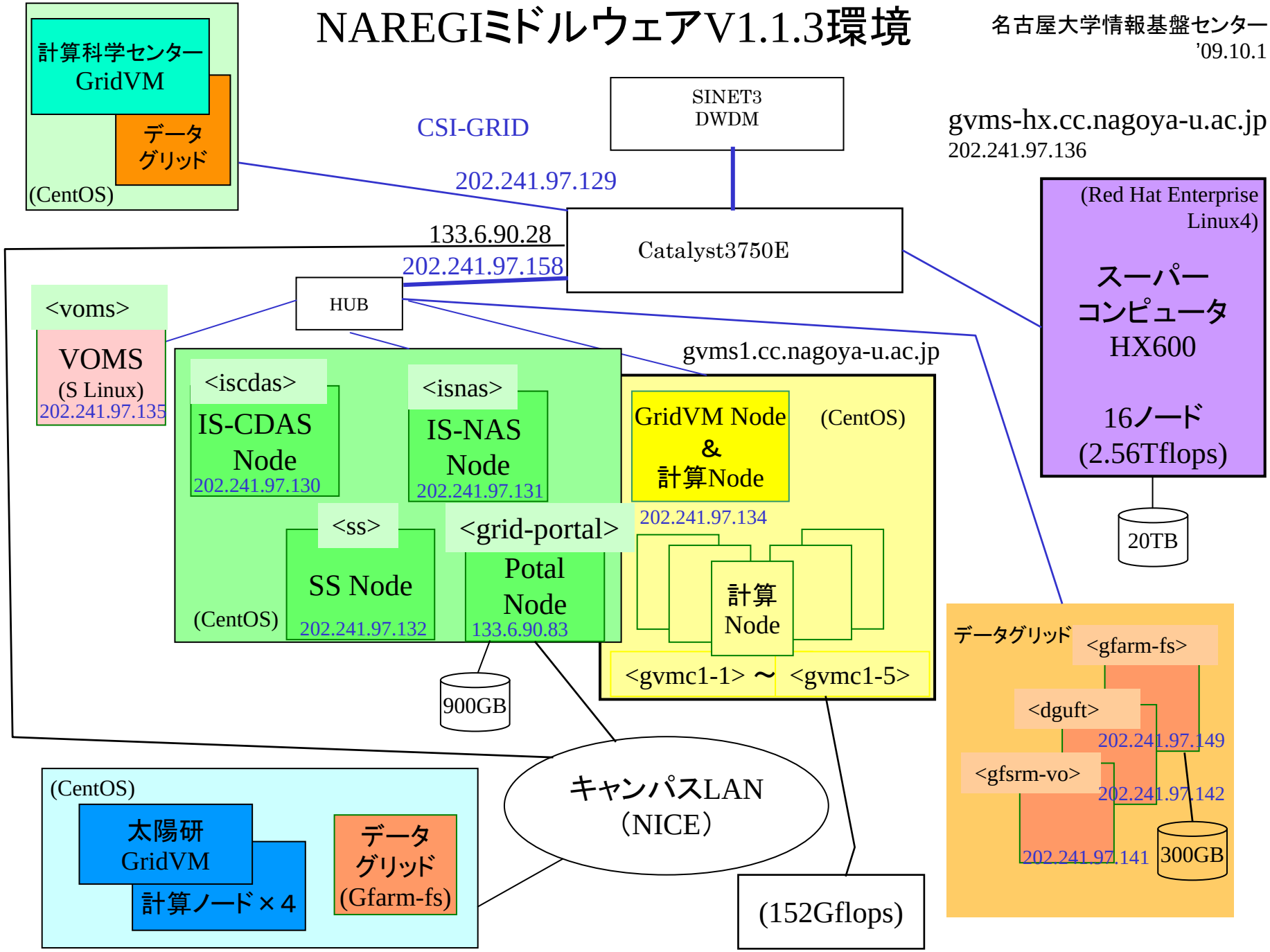
データグリッドでの使いにくいところ

名古屋大学太陽地球環境研究所ジオスペース研究センター
河野　貴久さんの報告より
（’09.3.31まで名古屋大学情報連携基盤センター）

機能面	Portal 画面で複数ファイルやディレクトリを選択してのインポートが不可。（複数のファイルを扱う場合は、アーカイブにしてファイル数を減らして対処した。）
操作性	Portal 画面でメタデータにコメントを付加することができるが、データが大量にある場合には入力が面倒。
実際の動き	データグリッドでファイルを削除しても画面の情報が自動的に更新されない（手動でReloadボタンを押す必要あり）
	使用環境によっては、画面のレイアウトがおかしくなる。（Linux のFirefox）

NAREGIミドルウェアV1.1.3環境

名古屋大学情報基盤センター
'09.10.1



Ⅳ．おわりに

- 基盤センターVOの利用環境はほぼ整った。
- グリッド環境をどのように利用促進していくのか？

９大学情報基盤センターへのグリッド配備・運用

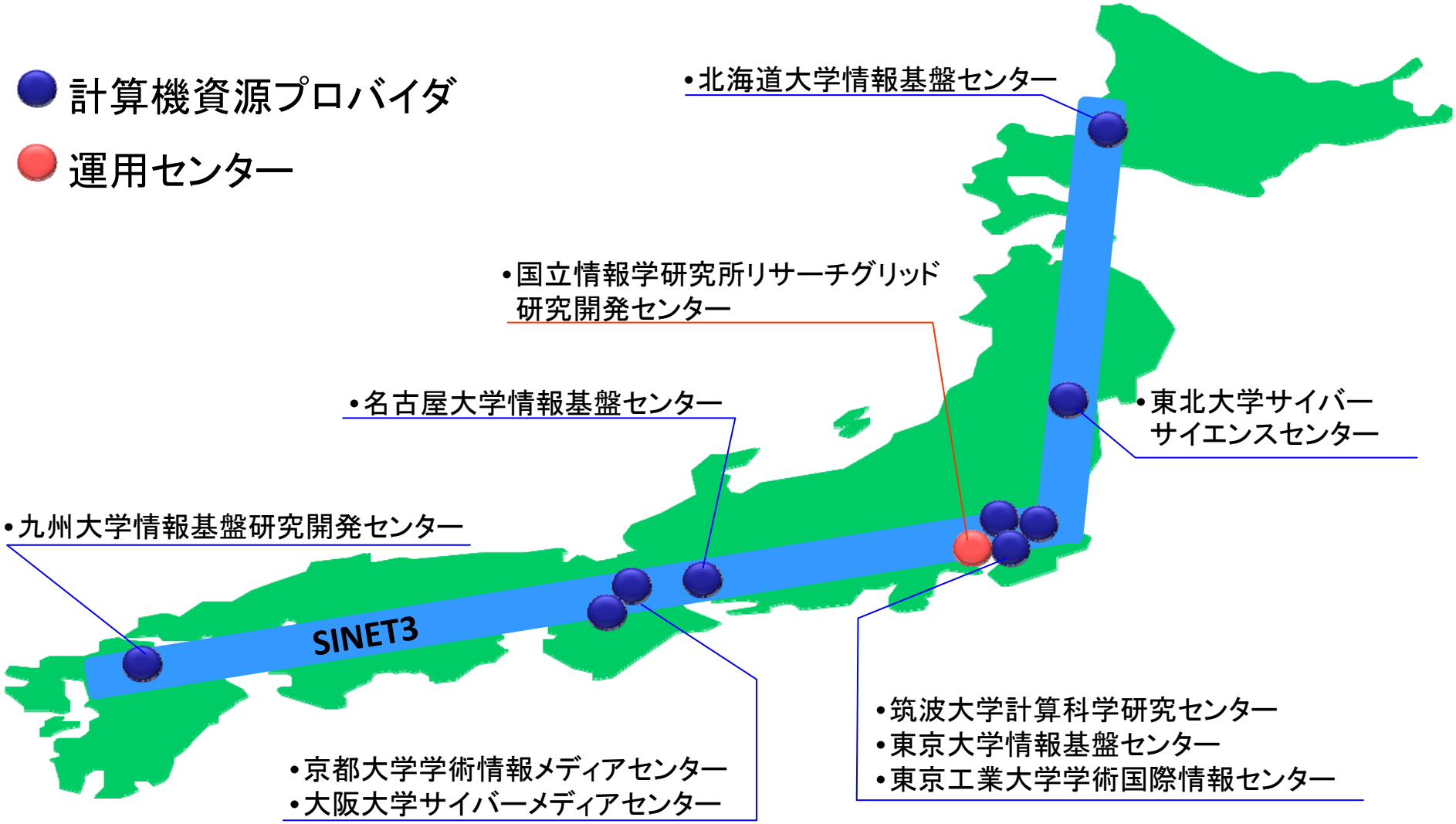
合田 憲人
国立情報学研究所

NII

9大学情報基盤センターグリッド

- CSIにおけるグリッド本格運用に向けて
  - 2008年秋　キックオフ
  - 2009年春　試験運用開始
- ミドルウェア
  - NAREGI middleware V.1.1.3
    - ✓ミドルウェアの機能・品質を実運用の視点で評価する.
    - ✓運用者・ユーザからの声を収集
    - ✓品質向上
- 運用
  - 情報基盤センターにおけるグリッド運用のための課題検討
    - ✓ユーザ管理，課金，…

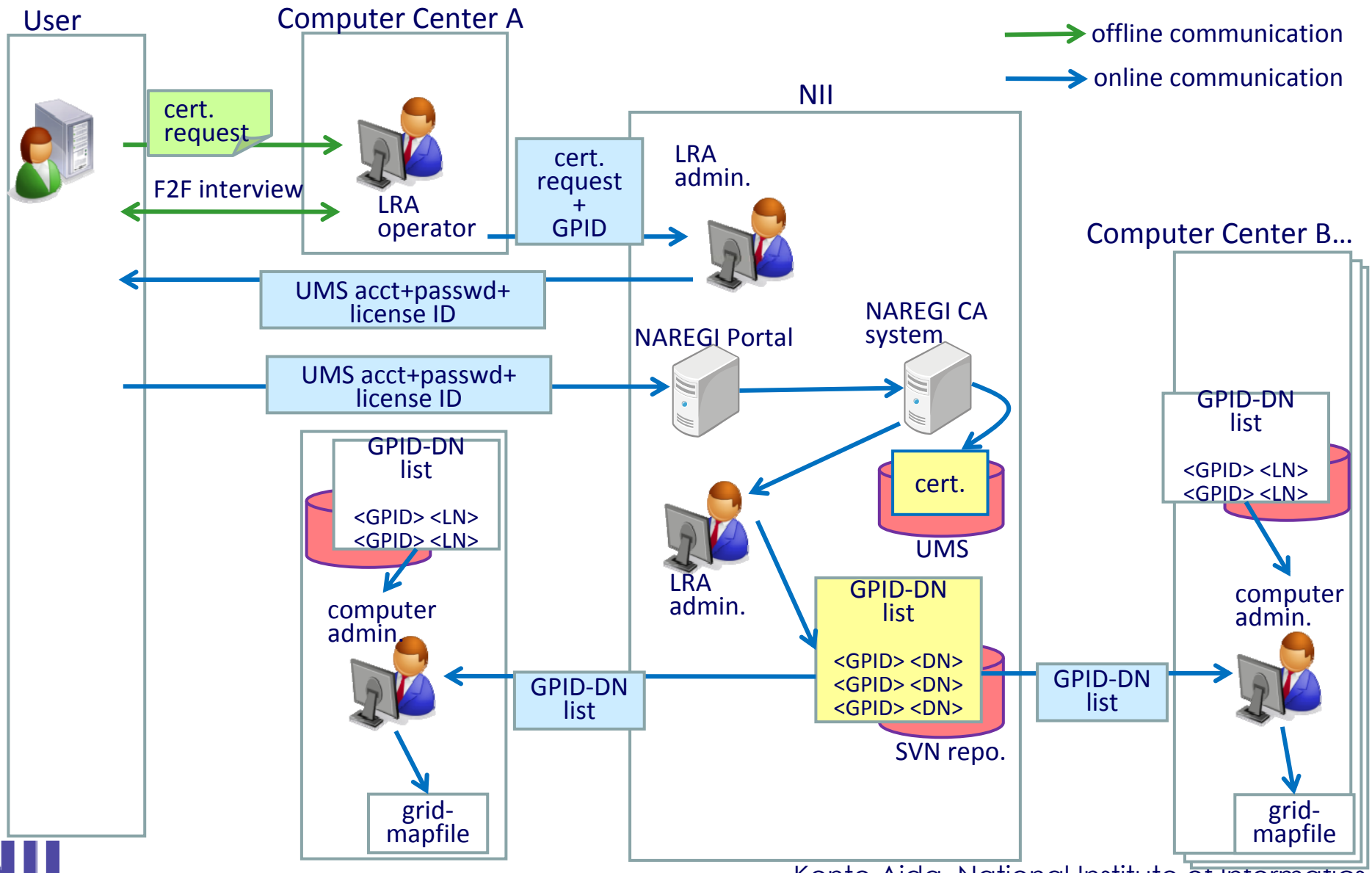
参加組織


計算機資源プロバイダ
運用センター
•北海道大学情報基盤センター
•国立情報学研究所リサーチグリッド
研究開発センター
•名古屋大学情報基盤センター
•東北大学サイバー
サイエンスセンター
•九州大学情報基盤研究開発センター
SINET3
•京都大学学術情報メディアセンター
•大阪大学サイバーメディアセンター
•筑波大学計算科学研究センター
•東京大学情報基盤センター
•東京工業大学学術国際情報センター

NAREGI認証局

- CSI参加組織（含ユーザ）向けのグリッド認証局運用
  - クライアント（ユーザ）証明書，サーバ証明書
- Asia Pacific Grid Policy Management Authority (APGrid PMA)認定
- 情報基盤センターとの連携
  - 情報基盤センターや研究機関でのLRA（ローカル登録局）の運用

グリッド証明書とgrid-mapfile



スケジュール

<table>
<tr><th colspan="2">09/2008</th><th colspan="2">05/2009</th><th colspan="2">04/2010</th></tr>
<tr><td colspan="2"></td><td colspan="2"></td><td colspan="2"></td></tr>
<tr><td>ミドルウェア
開発</td><td colspan="5">ミドルウェアの機能・品質向上</td></tr>
<tr><td rowspan="3"></td><td colspan="2">9大学情報基盤センターへの展開</td><td colspan="2"></td><td rowspan="3"></td></tr>
<tr><td colspan="2">試験運用準備
✓ユーザ管理方針
✓ツール</td><td colspan="2">試験運用（ユーザ限定）
✓ミドルウェア評価
✓評価結果のフィードバック
✓無料</td></tr>
<tr><td colspan="2"></td><td colspan="2">本格運用に向けた検討
✓ツール
✓課金方法</td><td>運用拡大</td></tr>
</table>

まとめ

- NIIでは，最先端学術情報基盤（CSI）の構築を進行中.
  - ネットワーク，認証，グリッド
- 9大学情報基盤センターでは，NAREGIミドルウェアを用いたグリッド試験運用を開始.
  - グリッドの実運用の視点でのNAREGIミドルウェアの評価
  - 情報基盤センターにおけるグリッド運用上の課題検討.

まとめ

- 2年3か月のWG活動により、Grid及びその周辺技術の動向を把握
- NAREGIプロジェクトの成果を実際に運用して、その評価と問題点を指摘し、プロジェクトへフィードバックした。Grid Middleware開発プロジェクトへの有効な示唆として利用していただけると幸いである。
- 当初のWG目標は達成
- 今後のセンターのあり方、Grid技術の適用の可否
  十分な議論は今後に期待
- Grid技術はCloud ComputingというITビジネスモデルへの技術基盤として存在