

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の概要 (RENKEI* Project)

2012年5月29日

研究代表者

国立情報学研究所
特任教授 三浦 謙一

*RENKEI: Resources Linkage for e-Science

背景と目的

■ e-サイエンス

- 計算機やデータ、ユーザ等からなる研究コミュニティを柔軟に形成することにより実現する新たな科学技術研究手法

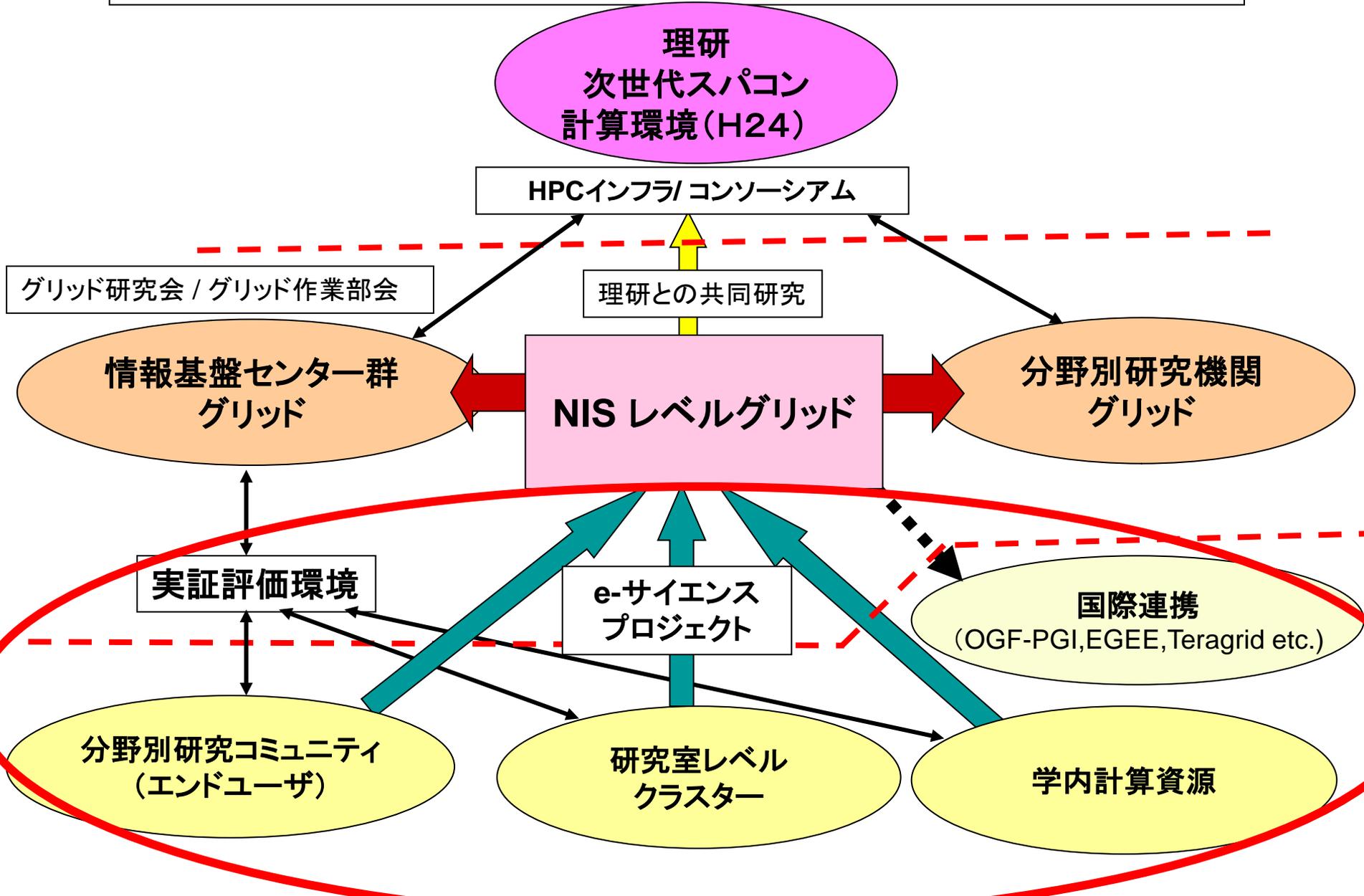
■ e-サイエンスの実現

- 単一のグリッドによる研究基盤とコミュニティの形成からさらに発展の段階へ
- 運用形態の異なる資源間の連携が新たに必要。
 - 研究室レベルの資源と情報基盤センターレベルの資源の連携
 - 異種グリッドミドルウェア間の連携

目的

研究室の資源(LLS)、情報基盤センターの高性能資源群(NIS)、異種グリッドミドルウェア上で提供される計算、データ、データベース等を柔軟に共有・連携させ、仮想研究コミュニティを形成しかつ運用するための基盤技術を確立する。

次世代計算研究環境としての展開





e-サイエンス実現のためのシステム統合・連携ソフトウェアの研究開発目標

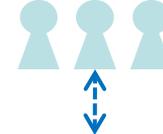
計算アプリケーションユーザ



計算/データグリッド
アプリケーションユーザ



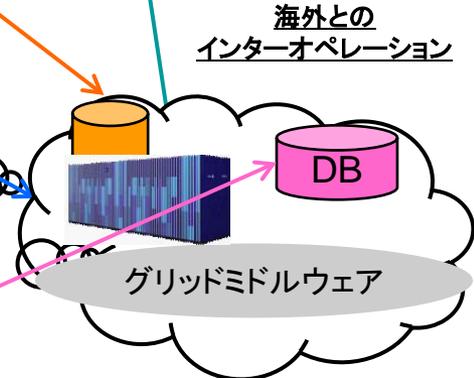
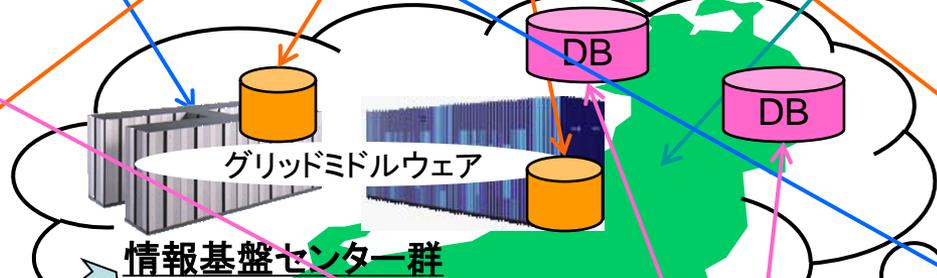
アプリケーション開発者



サブテーマ(1) 計算連携
目的: 研究室レベル・情報基盤センターレベル間でのシームレスなジョブ実行技術に関する研究

サブテーマ(2) データ共有
目的: 研究室レベル・情報基盤センターレベル間, 異種グリッド環境下でのデータ共有技術に関する研究

サブテーマ(4) アプリケーションインターフェース
目的: 異種グリッド環境下でのアプリケーション開発者向けインターフェースに関する研究



サブテーマ(5) 実証評価・ユーザ連携
目的: 実証評価基盤の構築, 情報基盤センターおよびエンドユーザと連携した実証評価

情報基盤センターユーザ

情報基盤センターユーザ等

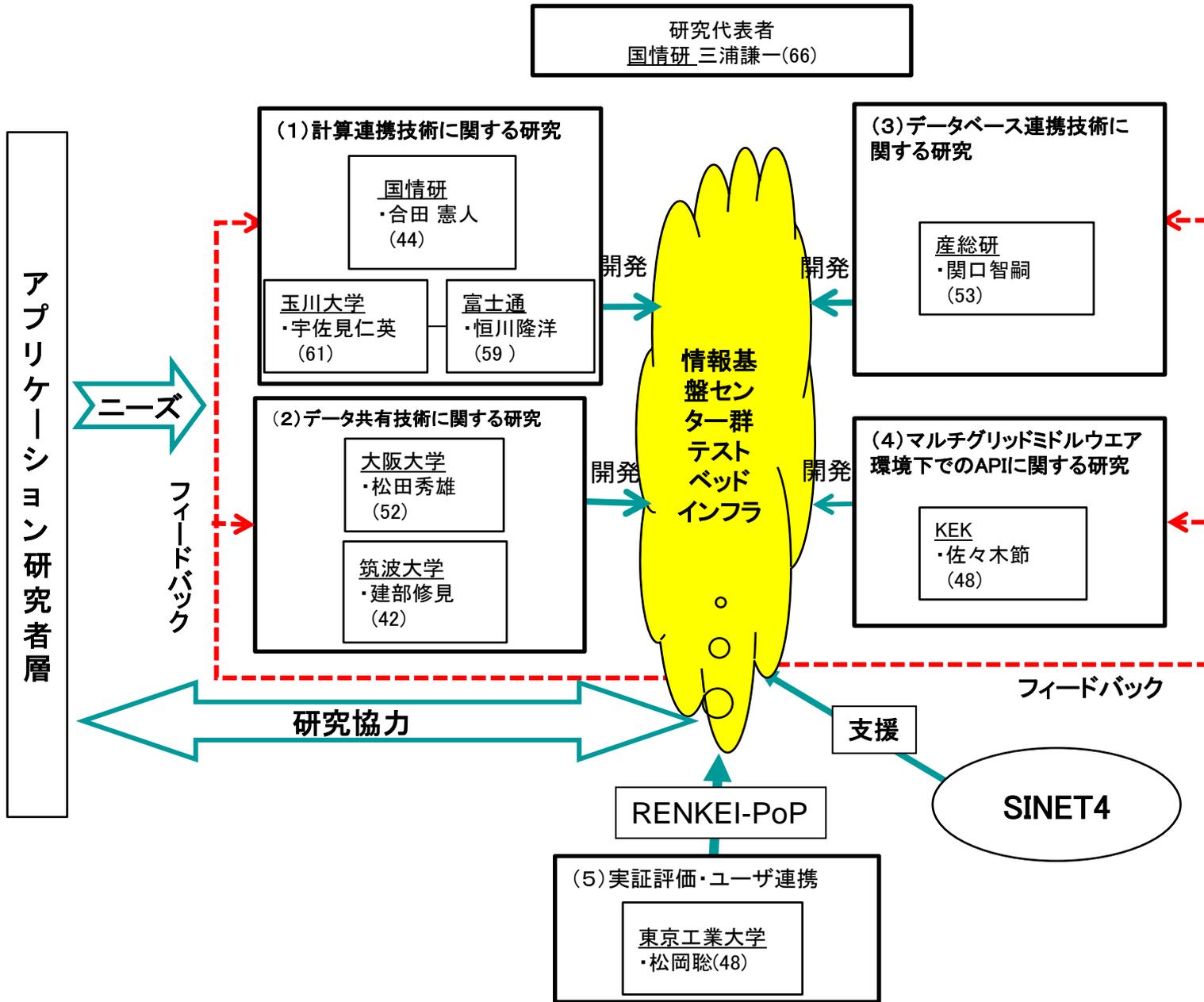
- 事例: *
- 名大太陽地球環境研究所
 - 北大インタークラウド
 - HPCI共用ストレージ

サブテーマ(3) データベース連携
目的: 異種データベースの連携・統合, ユーザ認証情報の管理技術に関する研究



データベースユーザ

研究実施体制



研究開発実施計画

研究開発項目および小項目	平成20年度	平成21年度	平成22年度	平成23年度
(1)計算連携技術に関する研究				
(1-1) ワークフローシステムに関する研究	基本設計	プロトタイプ実装・評価	プロトタイプ評価・機能強化	連携試験・評価・改良
(1-2) 異種グリッドミドルウェア間インタオペレーションに関する研究	基本設計	プロトタイプ実装・評価	プロトタイプ評価・機能強化	連携試験・評価・改良
(1-3) アプリケーション共有方式に関する研究	基本設計	プロトタイプ実装・評価	プロトタイプ評価・機能強化	連携試験・評価・改良
(2)データ共有技術に関する研究				
(2-1) 広域分散ファイルシステムに関する研究	基本設計	プロトタイプ実装・評価	プロトタイプ評価・機能強化	連携試験・評価・改良
(2-2) ファイルカタログシステムに関する研究	基本設計	プロトタイプ実装・評価	プロトタイプ評価・機能強化	連携試験・評価・改良
(3)データベース連携技術に関する研究				
(3-1) データベース連携・統合技術	外部仕様設計	内部仕様設計・プロトタイプ実装	プロトタイプ評価・機能強化	連携試験・評価・改良
(3-2) ユーザ認証情報管理技術	基本設計・プロトタイプ実装	プロトタイプ評価・機能強化	単体試験・評価・改良	連携試験・評価・改良
(4)マルチグリッド環境下でのアプリケーションインタフェースに関する研究				
(4-1) アプリケーションインタフェースに関する研究	調査・基本設計	プロトタイプ実装・評価	プロトタイプ評価・機能強化	連携試験・評価・改良
(5)実証評価・ユーザ連携				
(5-1) 研究コミュニティ構築のための実証基盤に関する研究	基盤構築・計測評価技術開発	基盤仮運用	基盤運用拡張	運用レベル実証評価
(5-2) 研究コミュニティの構築と運営支援	ユーザ開拓着手	ユーザ開拓・開発へフィードバック	ユーザ開拓・開発へフィードバック	利用促進・開発へフィードバック



サブテーマ1: 計算連携技術

● 背景

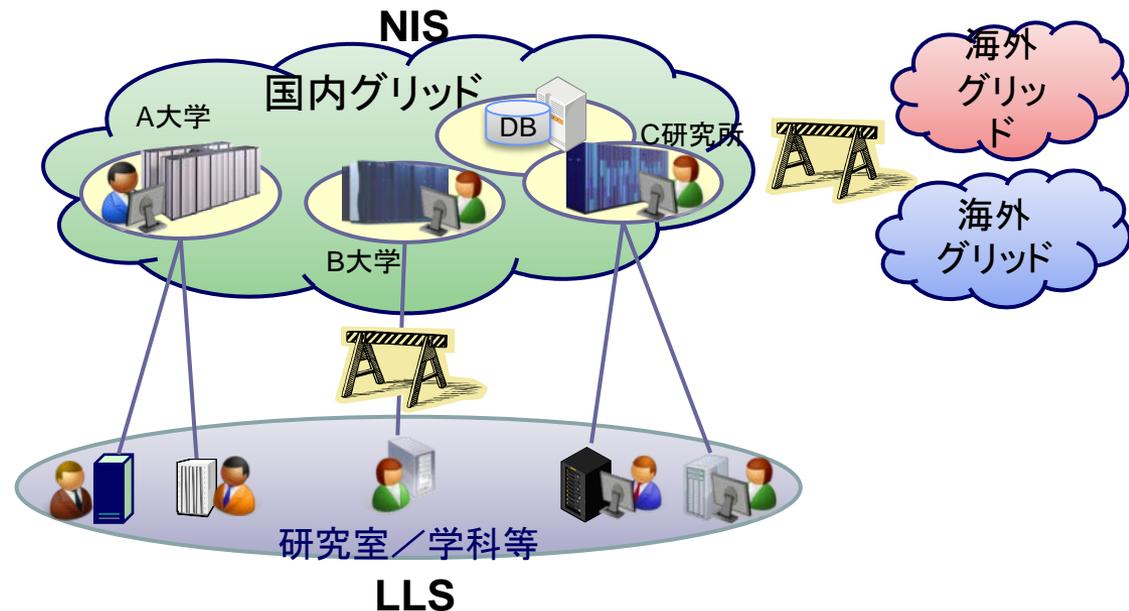
- 従来のグリッドコンピューティングの限界→異なる運用形態(研究室/基盤センター/海外グリッド)の計算資源にまたがった計算を実行できない。

● 目的

- 運用形態の異なる計算資源間のシームレスな計算(ジョブ)実行を実現し、計算資源を効率的かつ柔軟に活用する。

■ 研究開発項目

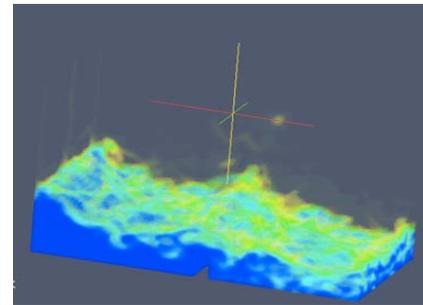
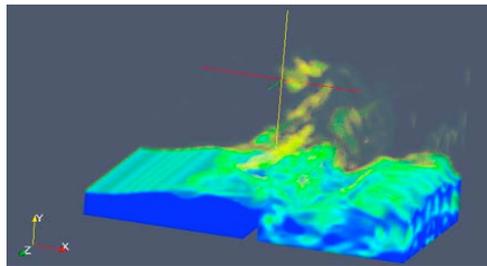
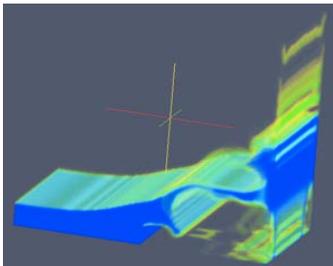
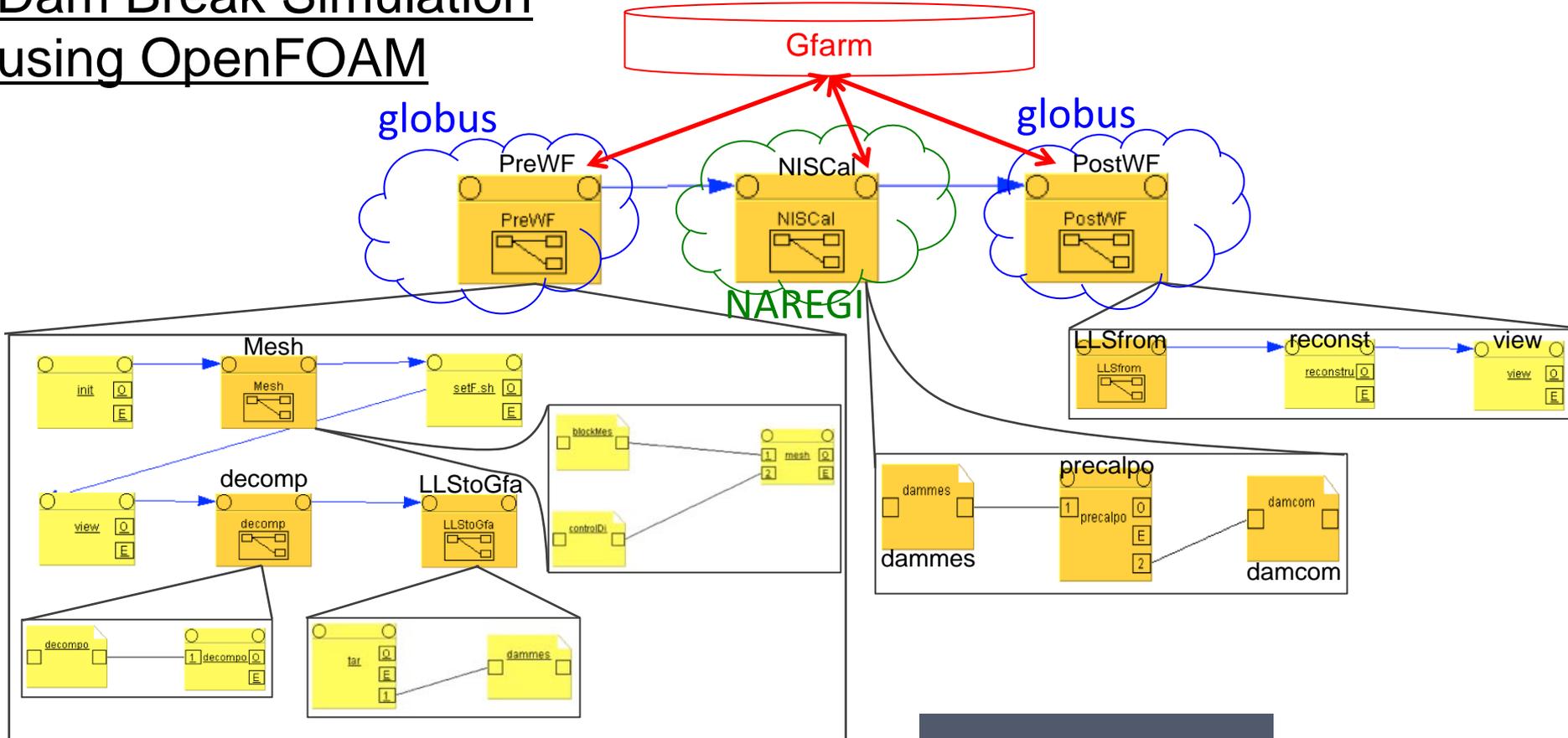
- ワークフローシステム
- 異種グリッドミドルウェア間インタオペレーション
- アプリケーション共有方式





ワークフローシステムの利用例

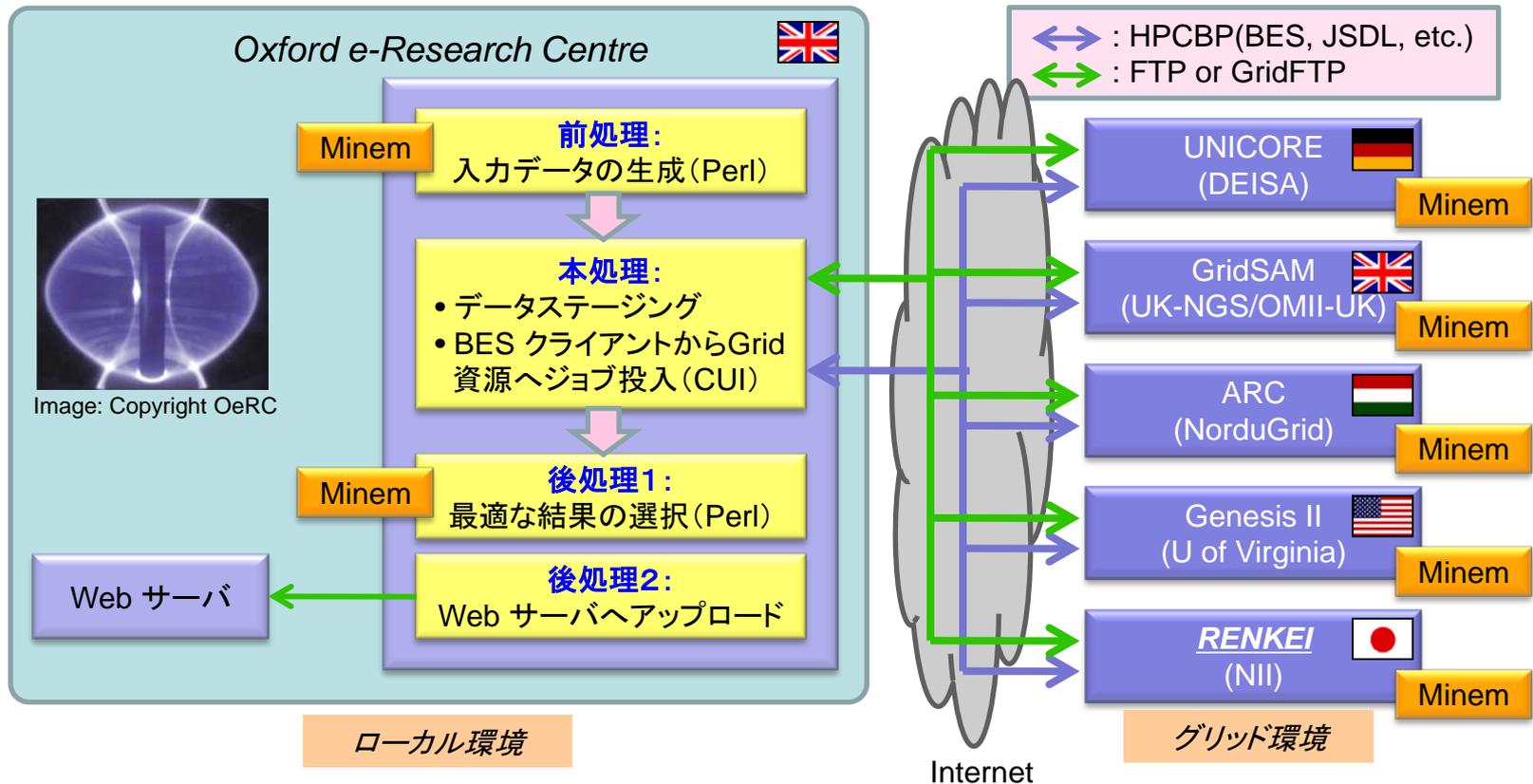
Dam Break Simulation using OpenFOAM





Interoperation Demo @ 5th IEEE eScience

- アプリケーション: Minem (Plasma Charge Minimization)
 - 球体表面の点電荷の全エネルギーの最小化
 - 前処理、後処理をローカル環境で、本処理をグリッド環境で処理
- HPCBPによる複数のグリッドへのジョブ投入のための問題点の調査と解決





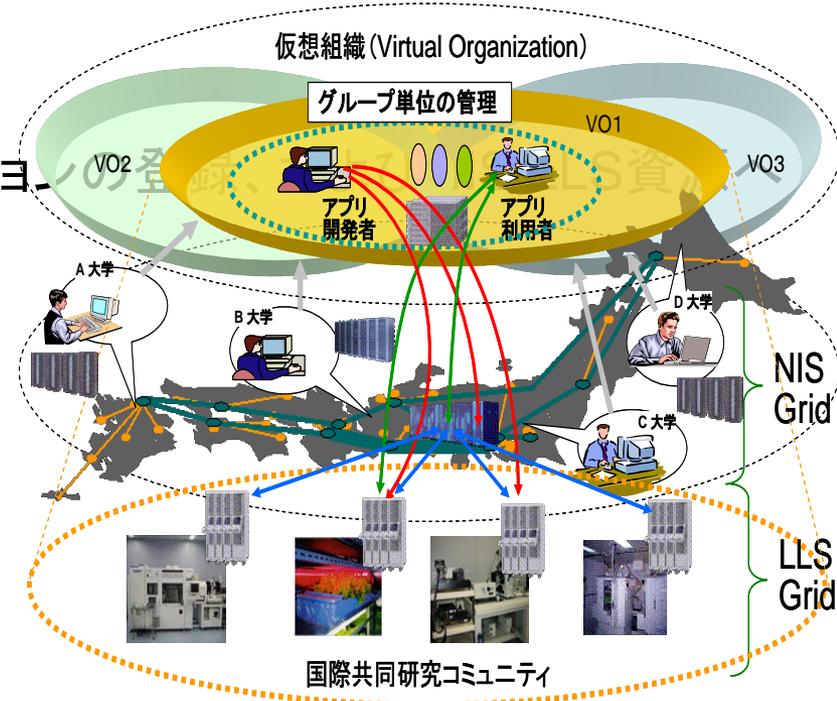
アプリケーション共有方式

● 目的

- 研究コミュニティ、あるいは研究グループで共有したい知の集積としてのアプリケーションをLLS とNIS の異なった運用環境において、統一的に共有・管理する使い勝手の良いAHS(Application Hosting Service) の方式と実証に関する研究を実施する。

■ NIS,LLSでのアプリ共有・管理

- NIS,LLS環境での統一的なアプリケーションの配置機能
- ACS(Application Contents Service)に準拠したアプリケーションおよびワークフローの管理機能
- ワークフローと連携した実行シナリオ作成支援機能



サブテーマ2: データ共有技術

● 研究の背景

実験技術やITの進歩により大量データ(“Big Data”)の発信
→ 分散しているため利用困難

● 目標

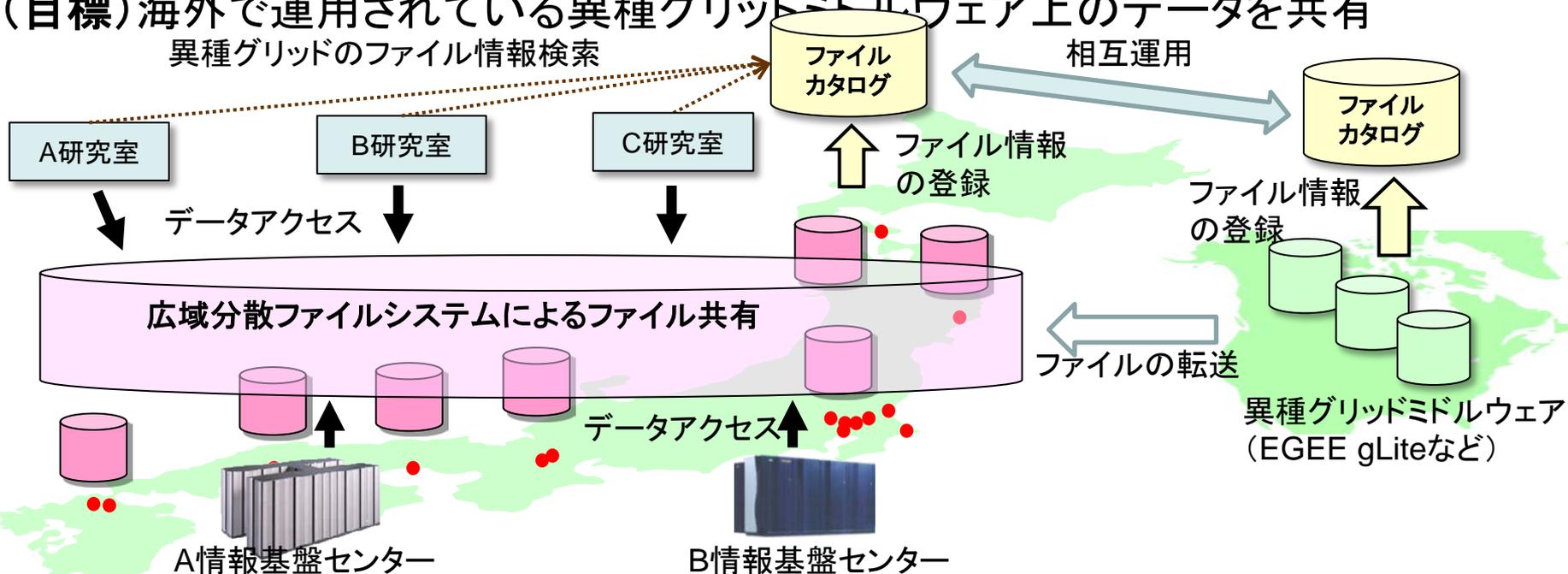
1. 広域分散ファイルシステムに関する研究 (筑波大学)

(目標) 国内に散在する研究データに対してどの場所からでも効率的にアクセス

2. ファイルカタログシステムに関する研究 (大阪大学)

(目標) 海外で運用されている異種グリッドミドルウェア上のデータを共有

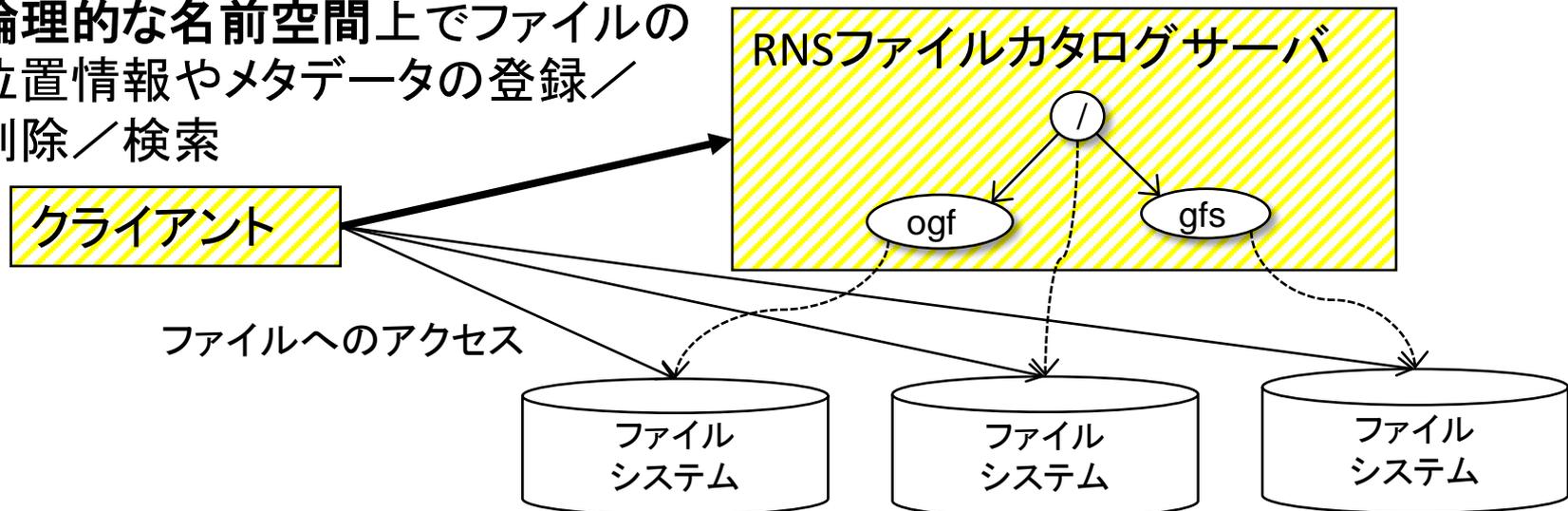
異種グリッドのファイル情報検索



RNSファイルカタログの特徴

- 国際標準仕様OGF Specification であるRNS (Resource Namespace Services)に基づいている
 - 特定のミドルウェアに依存しない
- 複数のRNSサーバを連携して、論理的に1個のファイルカタログを構成することができる
 - 大規模なファイルカタログを、特定のサーバに集中させることなく、複数のサーバに分割して負荷分散ができる
- XMLメタデータでファイル情報を検索可能

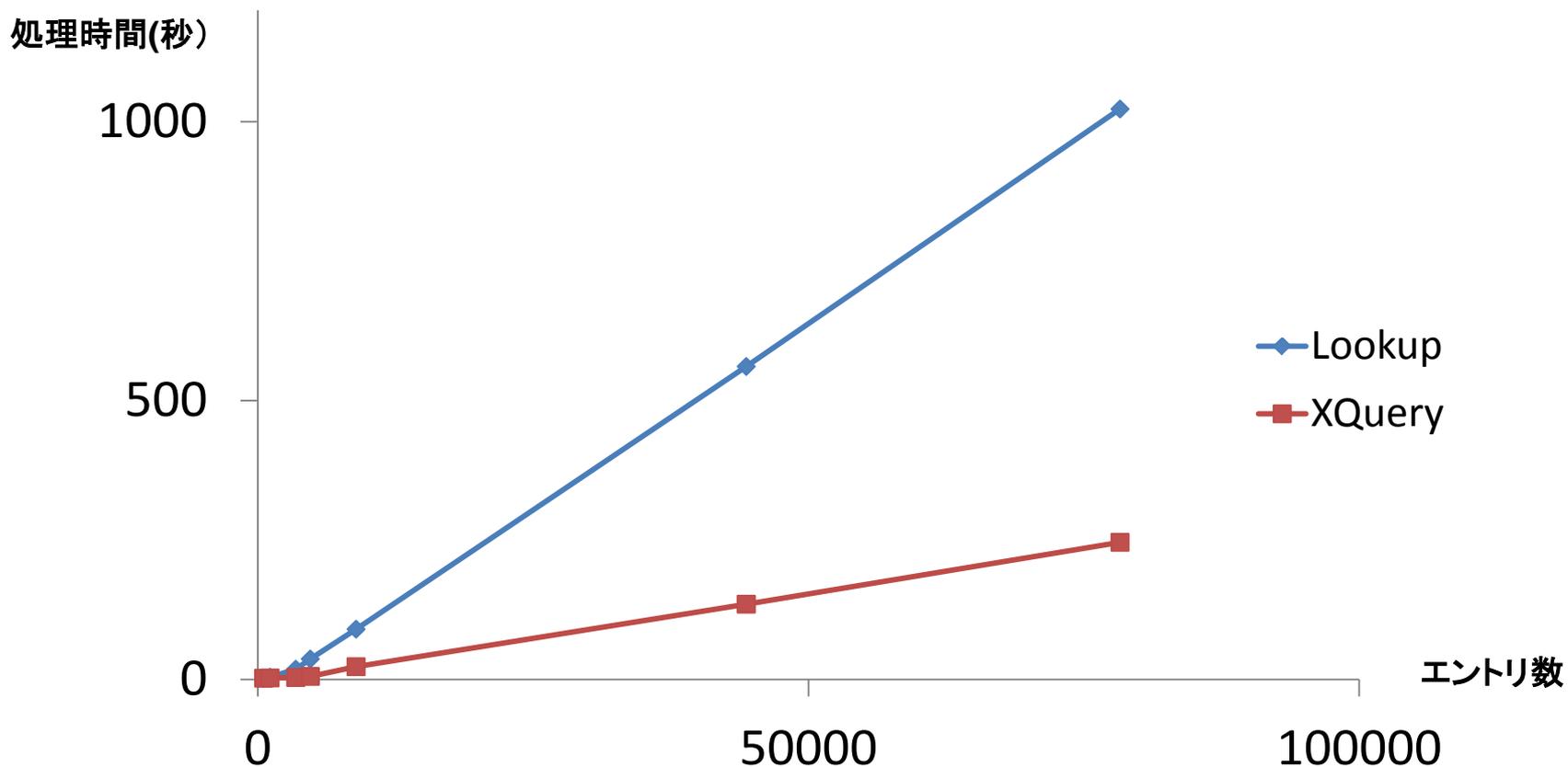
論理的な名前空間上でファイルの位置情報やメタデータの登録／削除／検索





RNS:Lookup操作とメタデータ検索の性能比較

- RENKEI-PoP環境(RPoP-NII)上でILCのファイルカタログデータをもとに処理時間を測定(3回実行したときの平均値)
- Lookup操作は1エントリごとにクライアントから操作する必要があるが、メタデータ検索は1回の操作でサーバ内で処理される



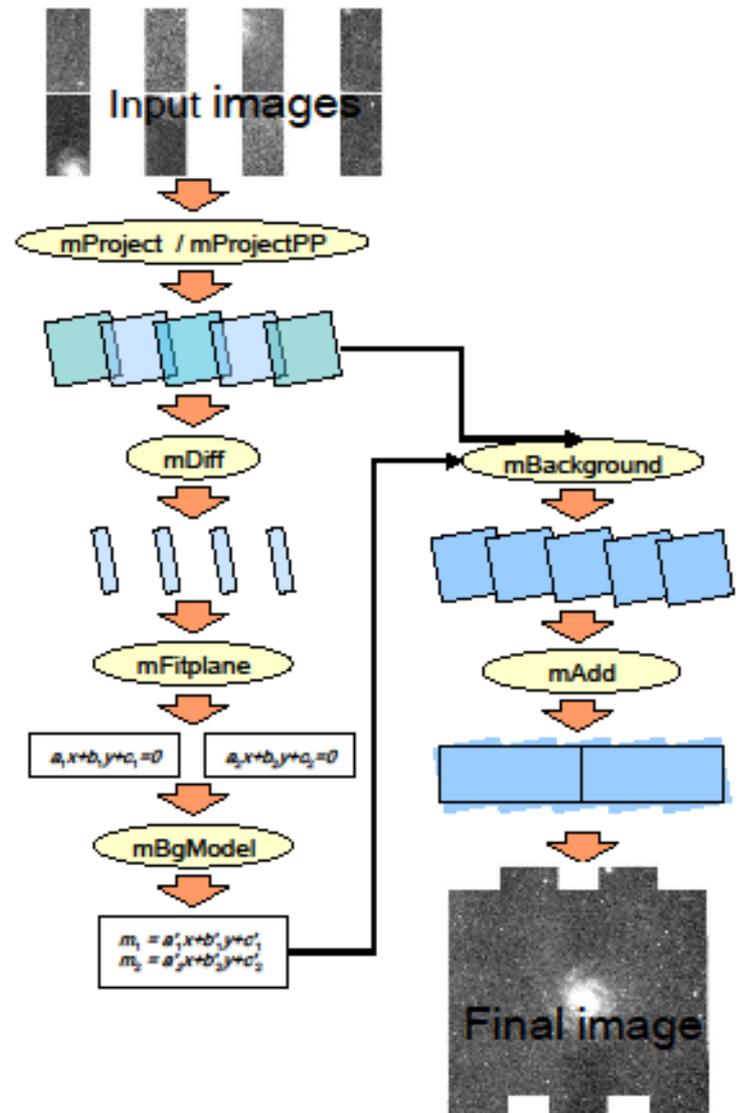
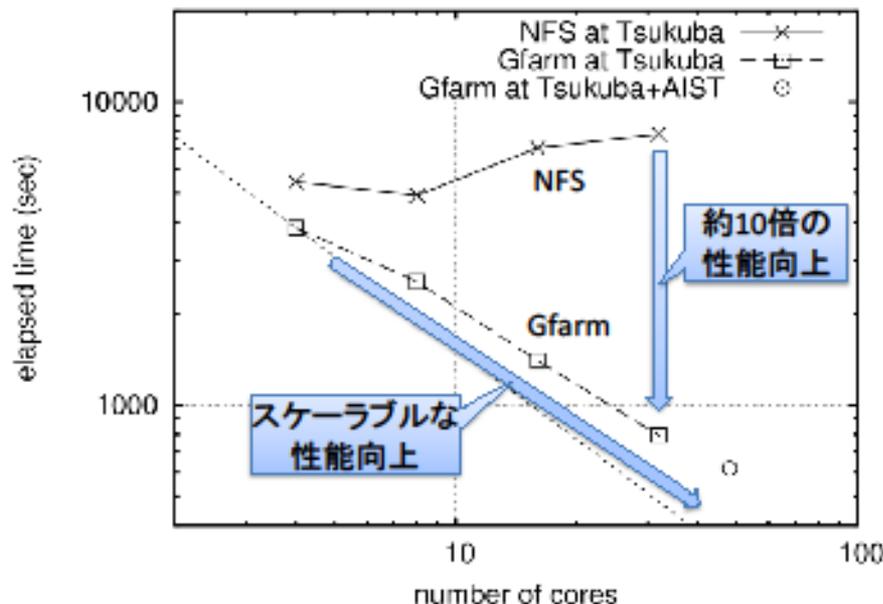


Gfarm2 ソフトウェアリリース状況

リリース日	バージョン	新機能	DL数
2009/3/26	2.2.0	数百クライアント対応, 遠隔アクセス高速化	448
2009/8/23	2.3.0	拡張属性, XML拡張属性と検索	713
2010/3/29	2.3.1	MDSのフェイルオーバー対応, クォータ機能	249
2010/7/1	2.3.2	Gfarm2fsの自動複製機能, Hadoopプラグイン	490
2010/7/22	2.4.0	更新型一貫性制御	668
2010/12/30	2.4.1	自動複製作成, 拡張属性キャッシュ, Samba	465
2011/4/23	2.4.2	拡張ACL, 分散Gfarm対応	393
2011/8/5	2.5.0	同期・非同期メタデータ複製, 読込中のgfsd 障害対応	250
2011/9/14	2.5.1	不具合修正	409
2011/12/19	2.5.2	不具合修正, ファイルトレース対応	63
2011/12/31	2.5.3	不具合修正, GSIの性能向上, MPI-IO/Gfarm	181
2012/2/25	2.5.4.1	不具合修正	105
		合計	4,434

Gfarm2: 実証実験 ～ 天文学への応用～

- 天文ソフトウェアMontageのワークフローで性能評価
- 広域ワークフロー並列実行エンジン Pwrake
 - Rakeを拡張し、並列分散実行を可能に
 - Gfarmファイルのローカリティを考慮したジョブスケジューリング
 - NFSの10倍の性能達成(32コア)
 - 2拠点でスケーラブルな性能達成





サブテーマ3: データベース連携技術

内容

1. 異種の分散データベースを連携・統合するミドルウェアの研究開発

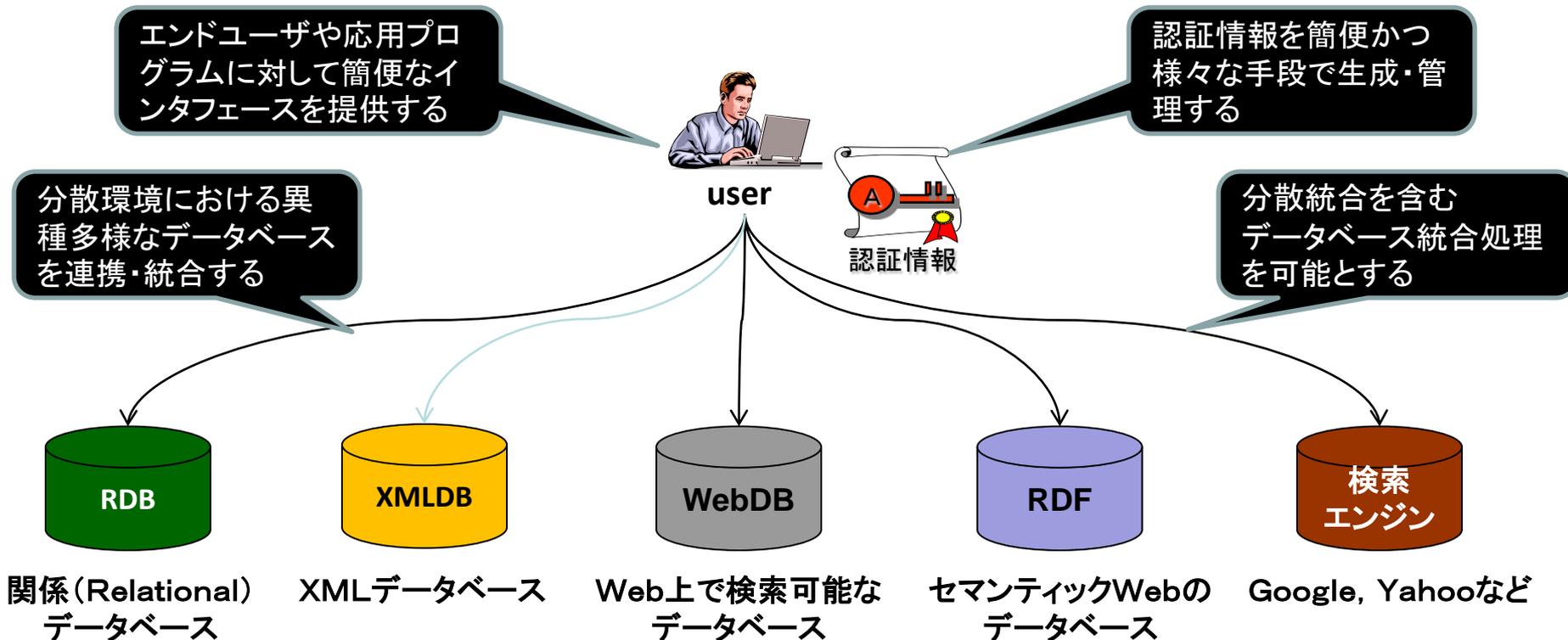
【特徴】

- ・ 分散統合を含むデータベース統合処理を可能とする。
- ・ WebDBやRDFに加え、全文検索エンジンもデータソースとして取り込むなど、異種多様なソースを結合する。
- ・ エンドユーザやアプリケーションに対して簡便なインターフェースを提供する。

2. ユーザ認証情報生成管理技術の研究

【特徴】

- ・ データベースへのアクセス制御の基礎となるユーザの認証情報を簡便かつ様々な手段で生成・管理する。
 - ・ ユーザ名・パスワードによる、OpenIDをベースにする、手持ちのユーザ証明書を用いる、など。





ユーザ認証情報生成管理技術：TSUKUBA-GAMAの開発

•Tsukuba-GAMA:バージョン1.2を公開中

- ユーザ名 & パスワード認証
- OpenIDを用いた認証
- 自己所有証明書を用いた認証
- PHPを用いたウェブポータル構築をサポート
- GridSphere Version2.3に基づいたポータル構築をサポート

様々な認証方式と
GSI-VOMSが連携可能

容易な
アプリケーション構築を
支援可能

• MyProxyに成果を反映(コードの貢献)

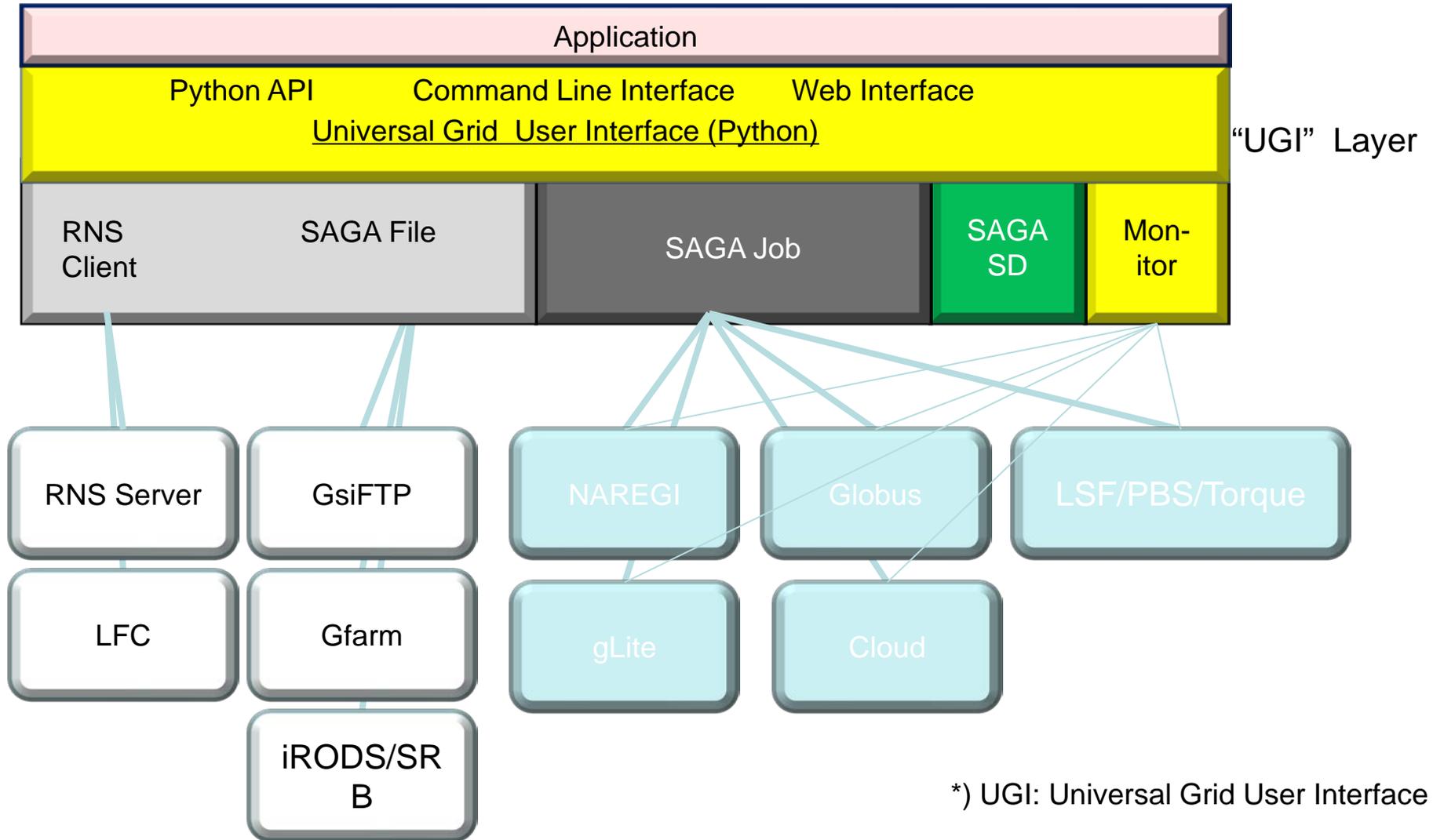
- MyProxyクライアントコマンド(myproxy-logon)へVOMSの機能を改訂. MyProxy 5.3に組み込み (1/18/2011)
- MyProxy Apache ModuleへVOMSの機能を追加. MyProxyウェブサイトで公開. (11/30/2010)
- MyProxy PHP クライアントへVOMSの機能を追加. MyProxyウェブサイトで公開. (6/11/2010)





サブテーマ4: マルチグリッド環境下でのAPI

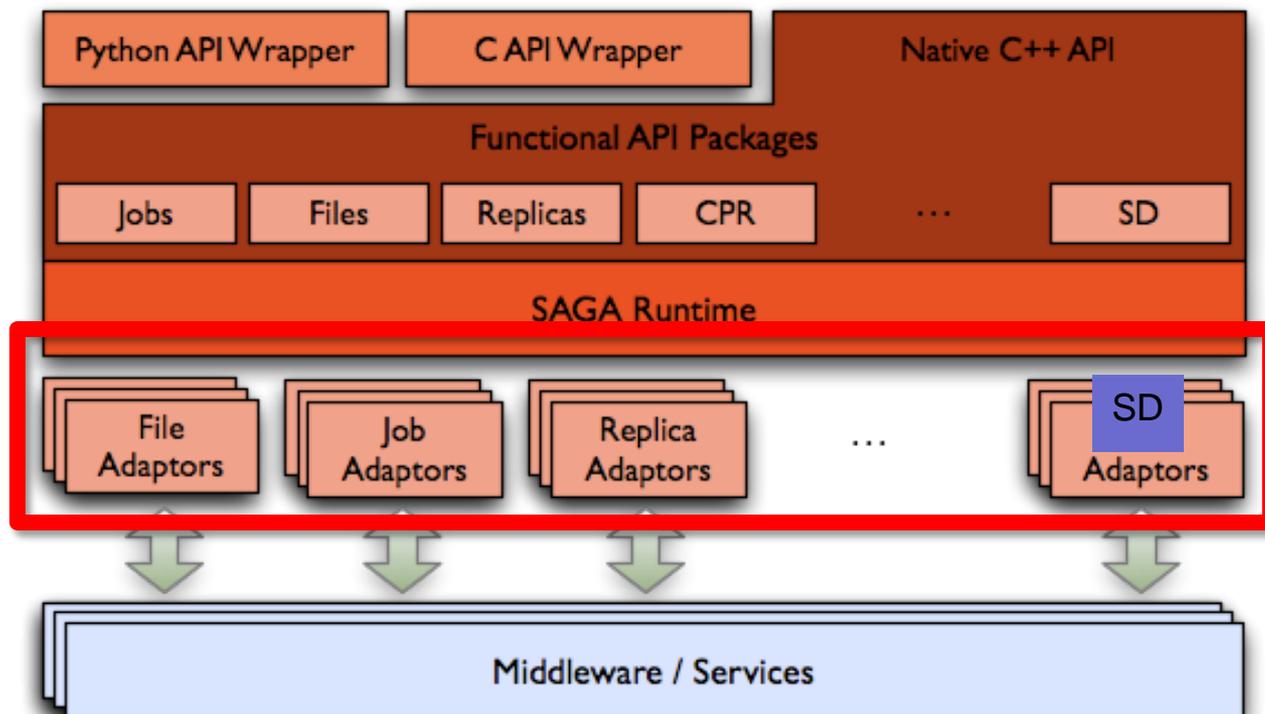
・UGI*アーキテクチャ



SAGA: A Simple API for Grid Applications

- SAGAは、OGF(Open GRID Forum)で標準化が進められているミドルウェアとユーザアプリケーションをブリッジするHigh-level programming interfaceである

C++実装とJAVAの実装があり、RENKEIでは、C++実装とそのPython wrapperを利用



アダプタの実装により新しい技術に対応

SAGAのアダプタとしては、ローカルジョブ、バッチスケジューラ、グリッド、クラウドが既に開発されており、様々な環境に適用することが可能である