



「CoreGrid: ライフサイエンス分野への グリッド基盤技術の応用」

平成15年度経過報告

大阪大学 大学院情報科学研究科
バイオ情報工学専攻 伊達 進
date@ais.cmc.osaka-u.ac.jp

<http://www.bioGRID.jp>

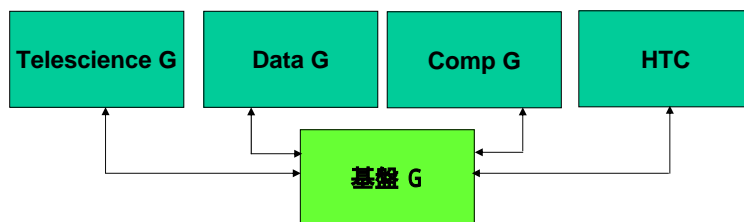
報告内容

- H15年度成果にむけた基盤Gの体制と役割
- H15年成果の紹介
 - 基盤技術の創生
 - IPv6 Grid FTP Alpha Release
 - IPv6 GSI-SFS
 - H14年度成果のフィードバック
 - 6Grid Research Environment demonstration
 - 中国科学院(CAS) - 大阪大学 バイオグリッドリサーチテストベッドの構築
 - 後方支援
 - 基盤システム整備
 - コンピューティング/データグリッド連携 (プロジェクト内研究)
 - 対外連携

- H15年度継続課題と今後の方針
 - MEGrid データ共有環境の構築
 - Virtual Laboratory for BioPfuga
 - コンポーネント化とネットワーク化にむけて
- まとめ
 - 基盤の成果の意味

[方針]他グループの後方支援に注力し、バイオグリッド全体としての成果の創造を支援するとともに、真に要求されるグリッド基盤技術を探求し、次年度以降の成果へのロードマップを定める

[体制]



[役割]他グループからの技術要求、ニーズに基いたグリッド技術・知識の供給と、基盤システムの運用保守支援

[成果]

- バイオ関連研究の効率化を実現する**新しい研究プラットフォームの実現**
- それらを構成する**真に有用なグリッド基盤技術の創造**
- バイオグリッドの成果による**新産業の導出**

• プロジェクトの成功のために



Research



Development



Strategic Management is essential.



Education



Routine Work

すべてのグループと密に連携することが求められる
基盤Gでは、この4要素のバランスを考慮した
戦略的な研究開発を実施する必要がある。

基盤Gの成果

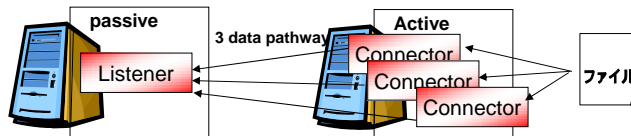
基盤技術の創生

実証実験 (成果のfeedback)

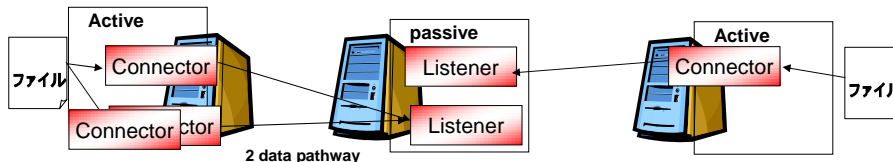
後方支援 (他グループ支援)

H15年度 基盤G成果<基盤技術の創生>

- GridFTP: サードパーティ転送、並列転送といったグリッドコンピューティング用に拡張されたファイル転送プロトコル
- Alpha Version, Release AM11:00, Nov. 14, 2003 (JST)
 - **Single Transfer** : 単一ファイルを単一ストリームで転送 (実装済み)
 - **Parallel Transfer** : 単一ファイルを並列ストリームで転送 (実装済み)



- **Striped Transfer** : 複数計算機間で並列的にデータを転送 (実装中)



[背景]

- 世界のグリッドプロジェクトにおいて、グリッド技術のIPv6化を本格的視野にいたしたプロジェクトは非常に少ない。
 - IBM and 6net in UCL Grid, e-science
- わが国の、裏づけされた高度かつ先進的なIPv6技術
 - WIDE Project, <http://www.wide.ad.jp/index.html>
- 数多くの6Gridへ期待する研究パートナー
 - SDSC, Telescience Group
 - CAS (China Academy of Sciences)
 - KISTI(Korea Institute of Science and Technology Information)

[これまでの関連成果]

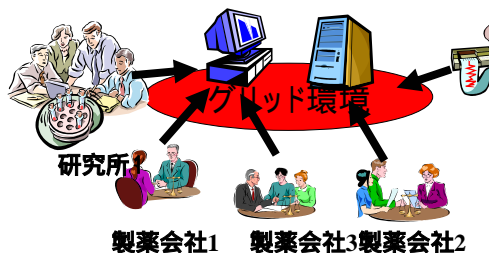
- [~H14年度] IPv6 Globus grid toolkit1.1.3, 2.2.3 Release
 - バイオ研究における必須条件であるセキュリティ対策として、IPv6/IPSecによるセキュリティを考慮



IPv6化への期待が非常に高いIPv6版GridFTPとIPv6版Globus grid toolkitの供給により、世界のグリッドプロジェクトへバイオグリッドプロジェクトの研究開発能力と存在感をアピール。



[H14年度成果 GSI-SFS: A solution for Secure Grid]



ユーザの要求:

- 1) データ共有環境はほしい
- 2) 他の製薬会社にアクセスしているデータベースについて知られたくない。
- 3) ネットワークのデータをデータ機密性を保持するため暗号化したい。
- 4) 利便性は確保しておきたい

GSI-SFSの有する機能特徴

- SDI (Single Disk Image)
- Single Sign-on
- On-demand Remote Filesystem Access
- データ機密性を確保
- ユーザ利便性

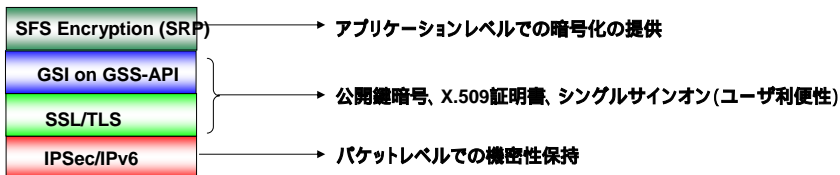
グリッドに一度ログインすれば、オンデマンドにユーザレベルでSDIを有するセキュアデータグリッド環境を実現できる。

製薬会社1 製薬会社3 製薬会社2



[現状]

- IPSecによるネットワーク層(パケットレベル)でのデータ機密性保持機能拡張と、IPv6によるアドレス空間の拡張を視野にいたした開発実施中



[H15年度成果予定]

- IPv6版 GSI-SFS Release, Dec. 31, 2003 (予定)

[ダウンロードユーザ]

(academic) 国立天文台様, 中国CAS (Chinese Academy of Sciences)様, シンガポール BII (Bioinformatics Institute)様, 韓国KISTI様、スイス <非公開> 様

(industry) 米国 N社, 国内2社 <非公開>

調査・問い合わせ: (academic) University of Vienna, Germany



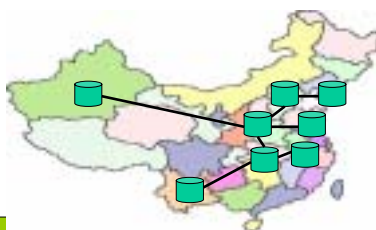
H15年度 基盤G成果 実証実験(成果のフィードバック)

中国科学院-阪大リサーチテストベッド

- 大阪大学サイバーメディアセンターbiogrid.jpと中国科学院(CAS)biogrid.cnの国際連携プロジェクト
 - 阪大BioGrid.jp <http://www.biogrid.jp>
 - 中科院 BioGrid.cn <http://www.biogrid.cn>
- H14年度バイオグリッド成果を利用したバイオグリッドリサーチプラットフォームの構築
 - GUIDE (Graphical User Interface to the Distributed Environment)
 - GSI-SFS, A Secure Filesystem on the Grid



- 中国は全体の70%程度の生物種を保持する巨大多様性国家 (Megadiversity country)である
- 巨大な生物資源を中国科学院をはじめとした研究機関がDB化 (SDB, Scientific database project led by Prof.Kai Nan)



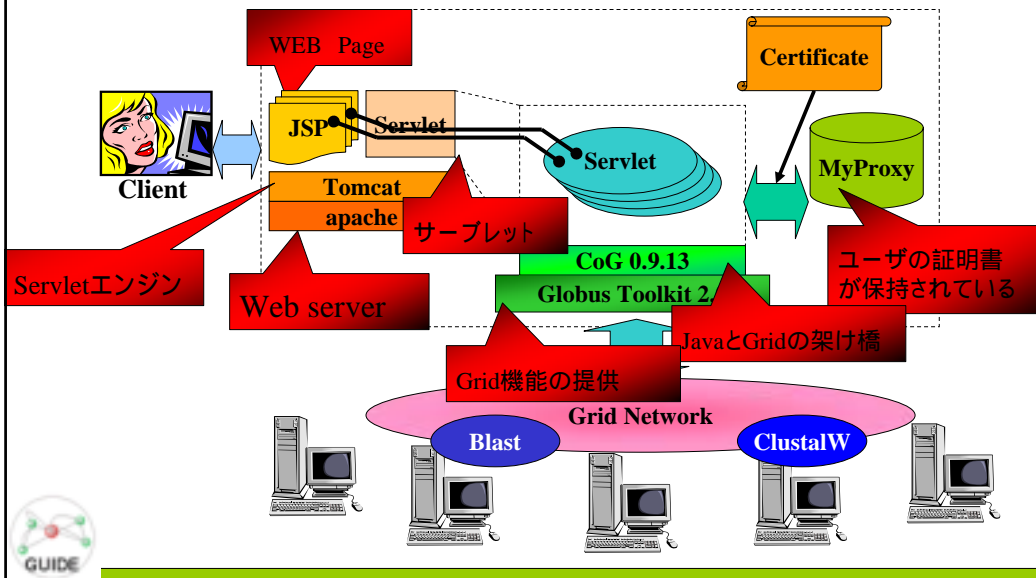
GUIDE概要

- 創薬研究過程において、比較的利用頻度が高く、共通性の高いソフトウェア、ツールのグリッド化
 - BLAST (相同性検索)
 - ClustalW (多重整列と系統樹作成)
- “Gridified” toolへのポータル提供
 - ユーザフレンドリなGUIにより複雑なグリッドのメカニズムを隠蔽

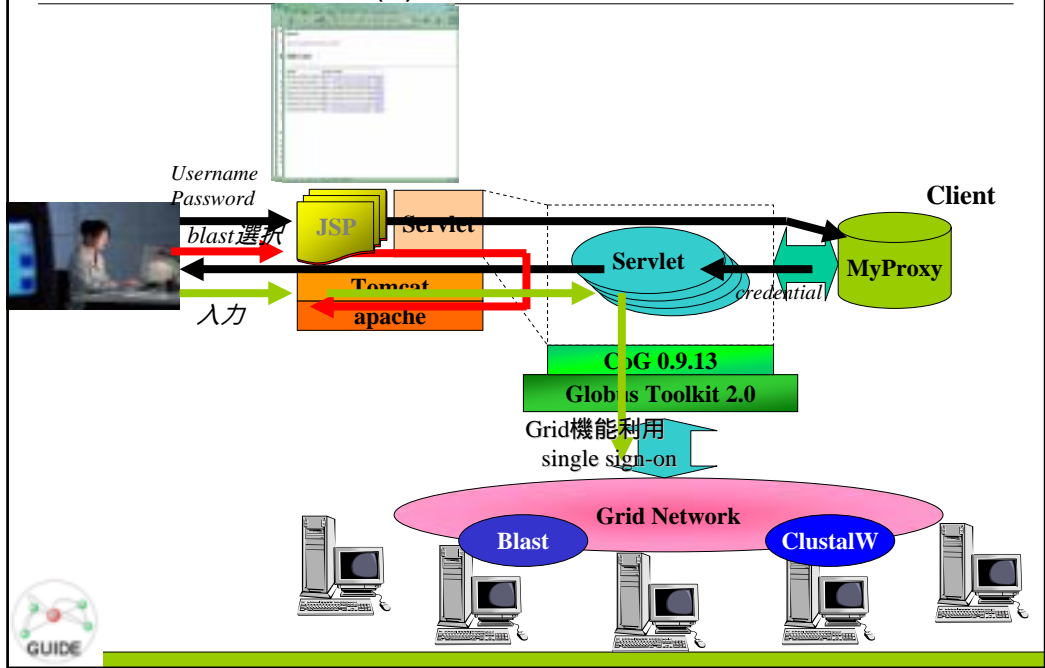
Grid User Interface to the Distributed Environment



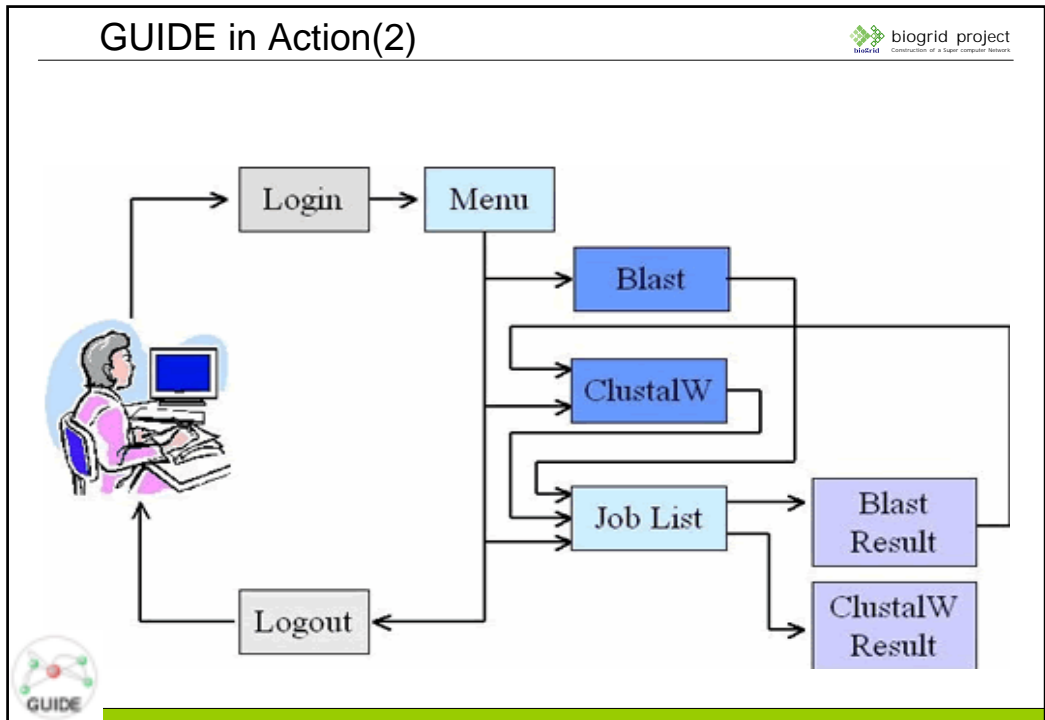
- Web technologies
 - JSP (JavaServer Pages) – 動的ウェブページ生成
 - Tomcat – Servlet エンジン
 - Apache – Web サーバ
- Grid technologies
 - Java CoG Kit – Javaインタフェースの提供
 - Globus Toolkit ver2.0 – Grid機能の提供
 - MyProxy – online credential repository

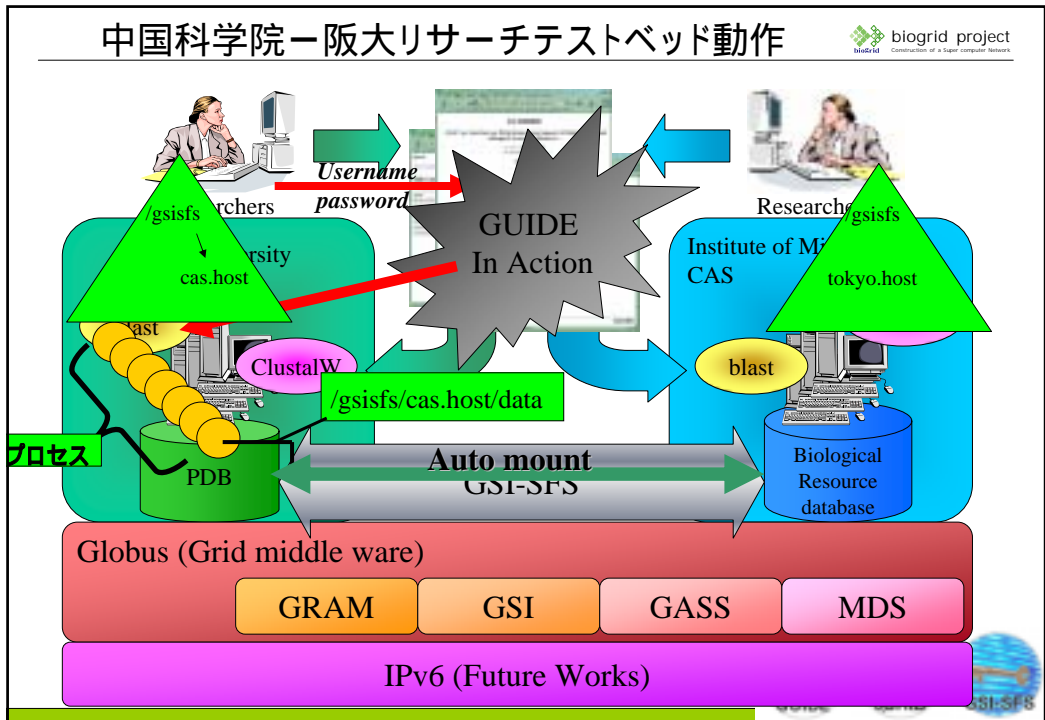
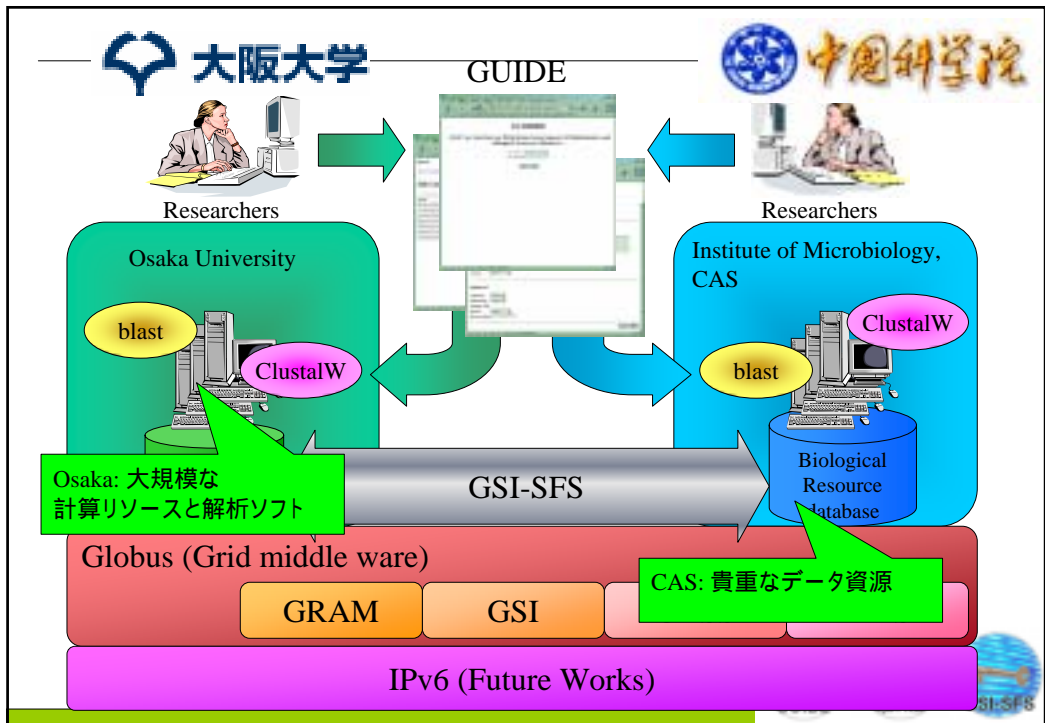


GUIDE in Action(1)



GUIDE in Action(2)





- CASの貴重なデータ資源と阪大の大規模計算システムを統合
- GSI-SFS + GUIDEを用いることにより、研究者らに位置透過性とユーザ利便性の提供
- データセキュリティとユーザ利便性を考慮したデータアクセスモデルの提供
 - 排他性
 - 通信路暗号化
 - Single sign-on



**次世代バイオリサーチプラットフォームの
1形態を示すことに成功**

マスメディアの注目

- NHK 大阪放送局 [2003/07/25]
- 朝日新聞 [2003/07/25]
- 産経新聞 [2003/07/25]
- 日刊工業新聞社 [2003/07/25]
- 時事日報社 [2003/07/25]
- 時事通信社 [2003/07/25]
- 共同通信社 [2003/07/25]
- 薬事日報 [2003/07/30]

中央農業総合研究センターからの問い合わせ

H15年度 基盤G成果 後方支援(他グループ支援)

システム整備 (1)

- 安定かつ最新状態にシステムを整備
 - 他グループからの要求への検討と対応
 - SW/HWの予防的措置と障害への迅速な対応
 - 月1回の定期保守計画作成と実施

Grid 基盤システム1 (管理1 + 計算8ノード)

Grid 基盤システム2 (管理1+計算78ノード)

ソフトウェア	version	管理	計算	ソフトウェア	version	管理	計算
Red Hat Linux	7.3			Red Hat Linux	7.3		
Score	5.4			Score	5.4		
PBS/SCore	5.4		-	PBS/SCore	5.4		-
Globus Toolkit	2.2.4		-	Globus Toolkit	2.2.4		-
Intel Fortran Compiler	7.1			Intel Fortran Compiler	7.1		
Intel C++ Compiler	7.1			Intel C++ Compiler	7.1		
Intel Math Kernel Library	6.0			Intel Math Kernel Library	6.0		
PGI Server	4.0			PGI Server	4.0		
MPI/PC-32 (SCore)	1.0.0			MPI/PC-32 (SCore)	1.0.0		

				Homology Sercher	2.2.5a		-
				Homology Parser	1.0		-
				FASTA	3.1		-
				SSEARCH	3.1		-

システム整備 (2)

- Webによる情報整備・管理
 - ユーザへの情報提供網の整備

各種マニュアル

・Howto文書

- ジョブ投入方法
 - Score
 - Globus
 - Condor
 - PBS
- バイオツール
 - Ssearch
 - Fasta
 - Homology searcher
- ソフトウェア開発環境

メインページ

システム情報提供

- PRAGMA5 Workshop(台湾)
 - Ninf-Gによる気候シミュレーション(GTRC)



[引用] <http://www.apgrid.org>で公開されている
発表スライド“Climate Simulation on the ApGrid/PRAGMA
Testbed at SC2003 (Yoshio, Nov. 2003)”

- SC2003 (Phoenix, 米国)
 - Ninf-Gによる気候シミュレーション
 - Nimrod-GによるGAMESS (Monash Univ., AU)
 - IPv6 GridFTP導入支援(SDSC+telescienceグループ)
 - HPC Challenge “Global Data-Intensive Grid Collaboration” (Univ. of Melbourne)など

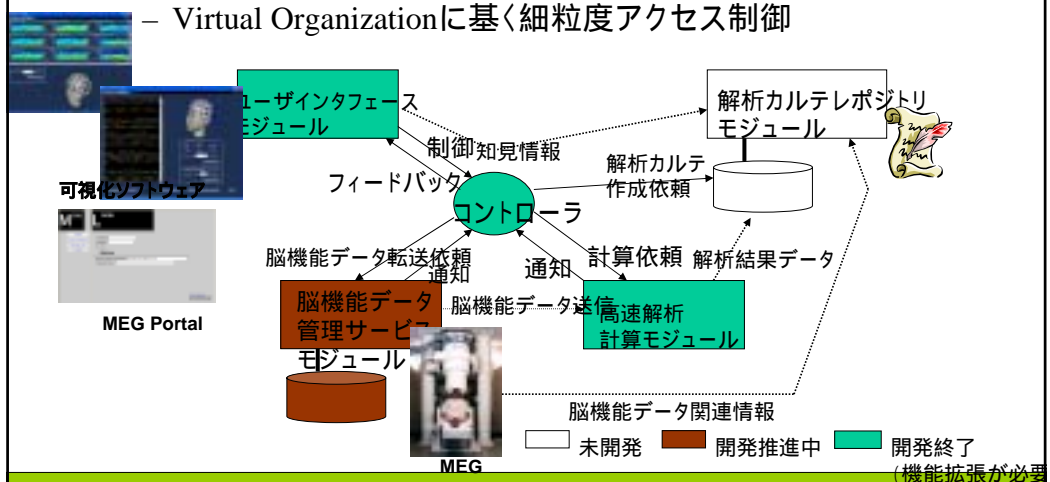
- データグリッドG連携
 - OGSA-DAI, OGSAに関する知識・技術供与
 - Web Securityに関する調査・実現可能性テスト
- コンピューティングG連携
 - グリッド上での動的プロセス生成に関する調査支援
- テレサイエンスG連携
 - IPv6GridFTP導入支援 (SC2003)

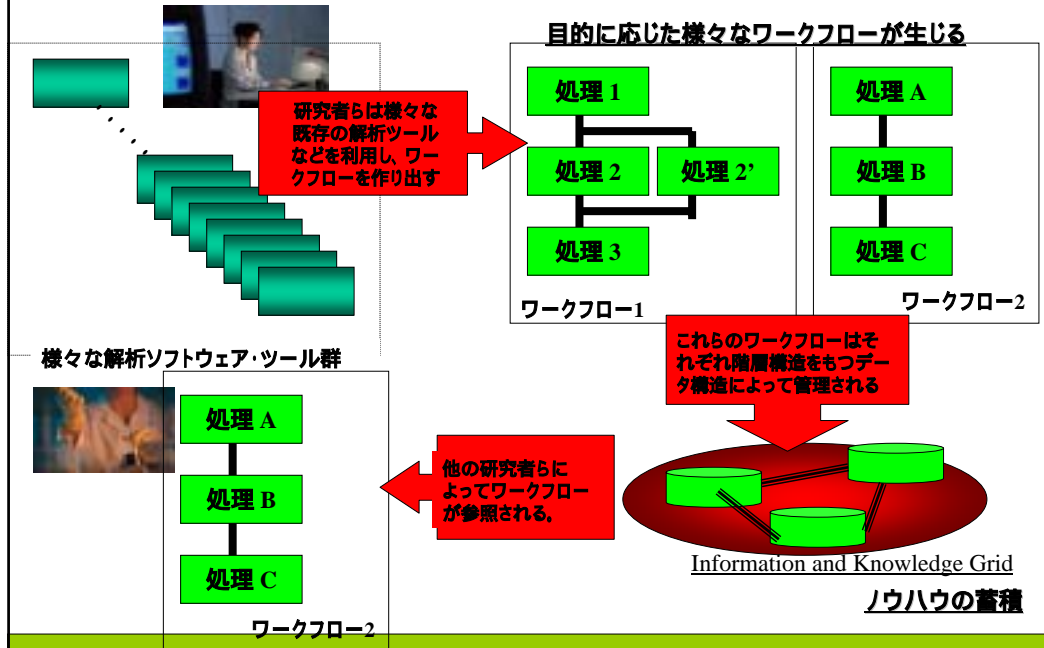
H15年度継続課題と今後の方針

継続中の研究課題 高次脳機能解析プラットフォーム

- 脳機能解析のための“解析カルテ”による知見共有環境の構築

- H14年度開発のMEGPortal環境との統合
- Virtual Organizationに基く細粒度アクセス制御





今後の方針

- データグリッドG R & D後方支援の強化
 - 関連セキュリティ技術調査
 - Semantic Web/Gridを用いたメタデータ管理手法
 - メタデータ生成
 - ワークフロー記述
 - 知見共有
- 成果公開へ向けた積極的な活動
 - GGF LSG/GFS
 - PRAGMA BioScience W.G など

- H15年度基盤Gはバイオグリッド全体の成果を重視し、
後方支援に注力するとともに、真に要求されるグリッド
技術を他グループとの密連携により探求



- 基盤Gの成果は以下の3種に分類
 - **グリッド基盤技術の創生**
 - **実証実験 (成果のフィードバック)**
 - **後方支援**