



# バイオグリッドの国際連携

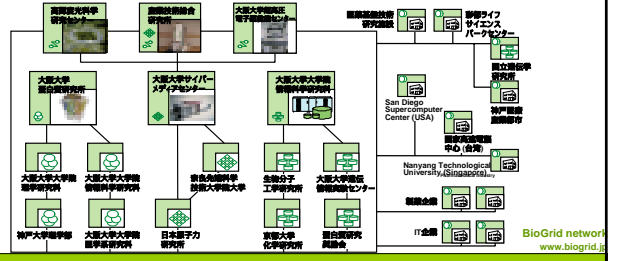
文部科学省  
「アジアグリッドイニシアティブ」

下條真司 大阪大学サイバーメディアセンター  
伊達 進 大阪大学大学院情報科学研究科

[www.biogrid.jp](http://www.biogrid.jp)

## スーパーコンピュータネットワークの構築 (BioGrid Project)

- 文部科学省 ITプログラム「e-scienceプロジェクト」の一つ
- 生命科学 (医学・生物学など) 分野におけるIT応用研究を推進するため、大阪大学サイバーメディアセンターを核として、高度な計算資源とネットワークを利用する基盤環境とソフトウェアをグリッド上に構築する。
- 生命科学とIT分野を結び架け橋となる人材を育成する。
- 国立研究期間、民間企業、国際機関との連携を通じて、スピードの速い開発を行う。
- これらの基盤環境とソフトウェアにより関西における関連機関の連携を強化し、in silico創薬の可能性を開くとともに我が国の国際競争力の強化をねらう。
- コンピューティンググリッドとデータグリッドの融合を生命科学分野に特化して行う



## 研究開発体制

プロジェクトリーダー 下條真司 (大阪大学)	外部助成委員会
運営委員会	プロジェクト管理 千里国際情報事業財団
<b>グリッド基盤グループ</b>	リーダー: 伊達 進 (大阪大学)
大阪大学, NECシステムテクノロジー, 三井情報開発, 千里国際情報事業財団	
<b>データグリッドグループ</b>	リーダー: 松田 秀雄 (大阪大学)
大阪大学, 京都大学, 日立ソフトウェアエンジニアリング, 蛋白質研究奨励会, 日本ヒューレット・パッカー, 富士通九州システムエンジニアリング, 三菱スペースソフトウェア, アズテックシステム	
<b>コンピューティンググリッドグループ</b>	リーダー: 中村 善木 (大阪大学)
大阪大学, NEC基礎研究所, 日立製作所, 生物分子工学研究所	
<b>HTCグループ</b>	リーダー: 藤川 和利 (奈良先端科学技術大学院大学)
大阪大学, 神戸大学, 奈良先端科学技術大学院大学, 日本原子力研究所	
<b>Telescienceグループ</b>	リーダー: 八木 重人 (高輝度光科学研究センター)
大阪大学, 高輝度光科学研究センター, 産業技術総合研究所	

## 世界の関連プロジェクト

- ITBL
- BIRN (BioMedical Information Research Network)
- [NC Biogrid](#)
- BII@SINGAPORE
- 東京ゲノムベイ/OBIGRID
- 大阪大学COE「ネットワーク共生環境」
- UK E-Science Project
- GGF/Semantic Grid RG

## 国内外での連携

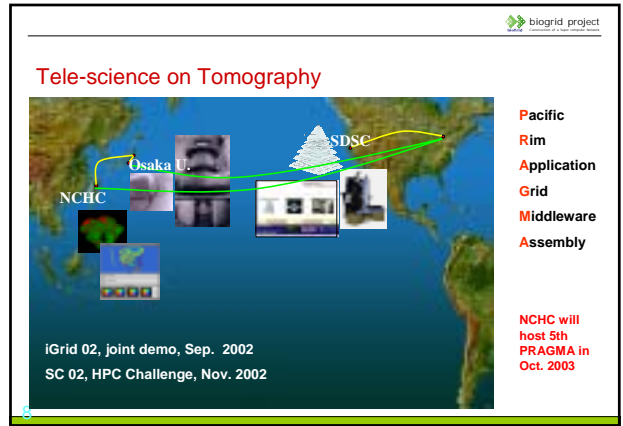
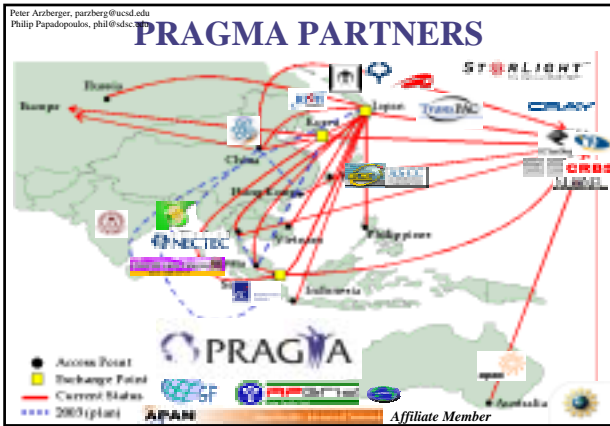
- 産総研ライフエレクトロニクスラボ  
- MEGプロジェクト
- PRAGMA (Pacific Rim Application and Grid Middleware Assembly)
- BIRN
- BII
- CAS (中国科学技術院)
- GGF/Life Science WG



<http://www.pragma-grid.net>

## Founding Motivations

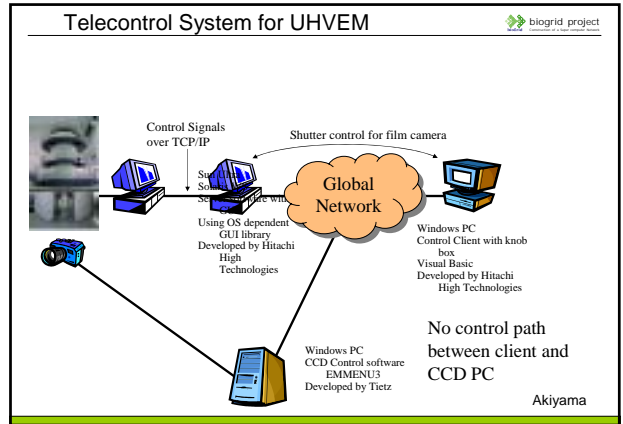
- Science is an intrinsically global activity
- The grid is transforming computing and collaboration
- The problem remains that the grid is too hard to use
- Middleware software needs to interoperate



Goal

- Improve throughput of observation facility providing the following functions by the introduction of Grid:
  - Data analysis function
  - Knowledge sharing function
  - Telecontrol function

Over International Collaboration



Data analysis function

- Image data analysis on UHVEM and High resolution CT (SPring-8)
  - Real-time analysis of large scale data obtained from special devices
    - Dataflow optimization
  - Adaptive visualization system for handling various environments

Copyrighted by Dr. Naoki Yamada (NCHC)

biogrid project

## Progresses

- GT3 enabled new architecture of telescience portal (NCMIR, SDSC)
- Telecontrol and analysis system for UHVEM (Cybermedia Center, NCHC)
- Pipelined image processing method (Cybermedia Center)
- New HDTV codec using JPEG2000 (KDDI R&D Laboratory)

Akiyama

## Demonstration



- Telecontrol and image processing system prototype
  - HDTV transfer system (demonstrated at *NPACI* booth)
  - Image analysis web portal (demonstrated at *BioGrid* booth)



## HDTV codec using JPEG2000



- Previous codec was using MPEG2
- JPEG2000 enables us to ...
  - Show clear image
  - Control bandwidth to be required

## SC2003



## Collaborations



MoU among NCMIR (UCSD), CMC, UHVEM, (Osaka Univ), KDDI, KBSI

## Our Goals



- Develop telescience portal environment among Japan, Korea, US and other.
- International collaboration for Agriculture & Ecogrid
  - Taiwan, US, Japan, Thailand, Australia, China and other
- Integration of Data and Computation on Grid
  - GGF10 RG may be proposed for instruments on the Grid – from synchrotrons to sensors

## 中国との農業を通じたグリッド連携

-黄土高原生態文化回復へむけて -

大阪大学大学院情報科学研究科  
バイオ情報工学専攻 ゲノム情報工学講座  
助手 伊達 進

## 講演内容

- バイオグリッド成果(の復習): 中国科学院とのテストベッド
  - GUIDE (Grid User Interface to the Distributed Environment)
  - GSI-SFS ( A Secure Filesystem on the Grid)
- 今後の方針: 黄土高原生態文化回復プロジェクト
  - プロジェクト概要とその意義
  - 導入される技術の紹介

## Collaboration with CAS

- China is a megadiversity country that is said to have 70%+ species.
- CAS is moving towards the construction and integration of databases of specimens and names (SDB, Scientific database project, Kai Nan, CAS)



Osaka University and CAS team (currently, Juncai Ma's team) started building a proto-type BioGrid using Osaka's Grid technologies and CAS's database technologies and data resource of importance.

## GUIDE概要

- 創業研究過程において、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へのポータルの提供
  - 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

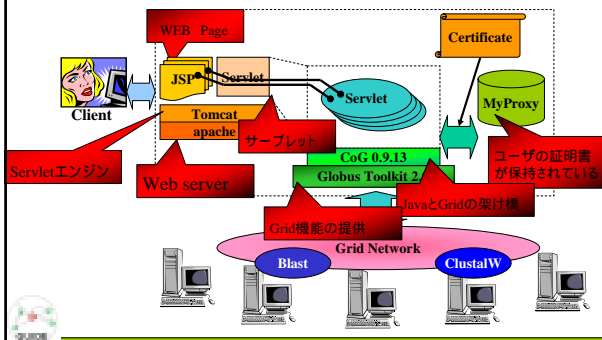
### Grid User Interface to the Distributed Environment



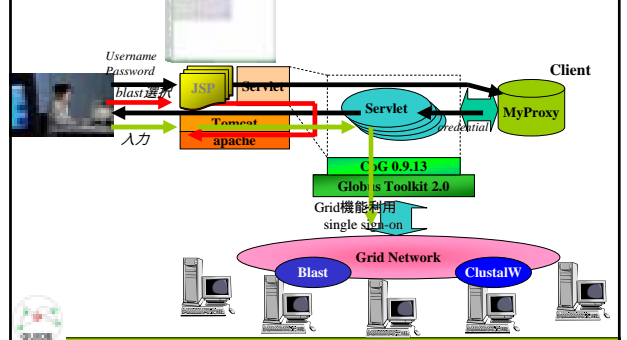
## Technical view of GUIDE (1)

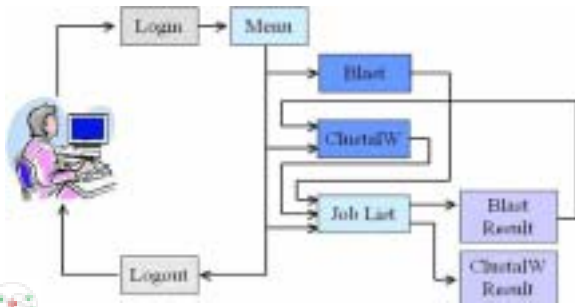
- Web technologies
  - JSP (JavaServer Pages) – 動的ウェブページ生成
  - Tomcat – Servlet エンジン
  - Apache – Web サーバ
- Grid technologies
  - Java CoG Kit – Javaインタフェースの提供
  - Globus Toolkit ver2.0 – Grid機能の提供
  - MyProxy – online credential repository

## Technical view of GUIDE (2)



## GUIDE in Action(1)





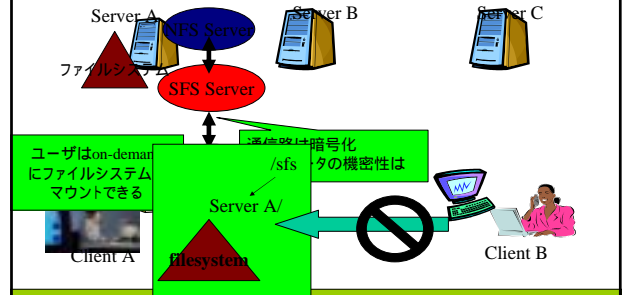
**ユーザの要求:**

- 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

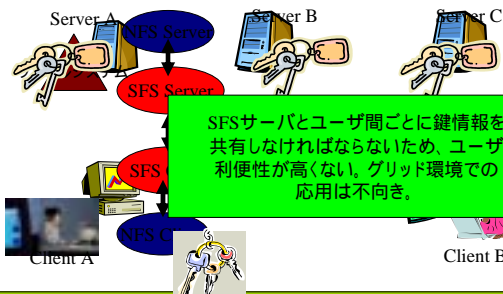
要求される機能

- **SDI (Single Disk Image):** ユーザはNFSのようにファイル空間を一元的に扱いたい
- **Single Sign-on:** ファイルにアクセスする際に必要となる認証の回数を最小限にしたい
- **On-demand Remote Filesystem Access:** アクセスの必要性に応じてファイルにアクセスしたい
- **データ機密性を確保:** アクセスしているデータに関して、他人に知られたくない
- **ユーザ利便性:** データセキュリティのために不便になるのは避けたい

SFS (Self-Certifying Filesystem)

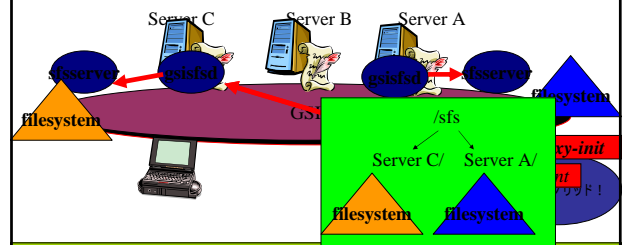


SFS (Self-Certifying Filesystem)



GSI-SFS

- GSI (Grid Security Infrastructure)とSFS (Self-Certifying Filesystem)の機能を統合したセキュアファイルシステム

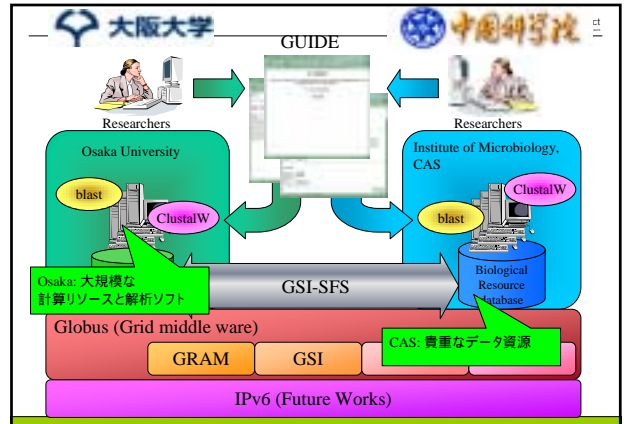


## GSI-SFSの特徴

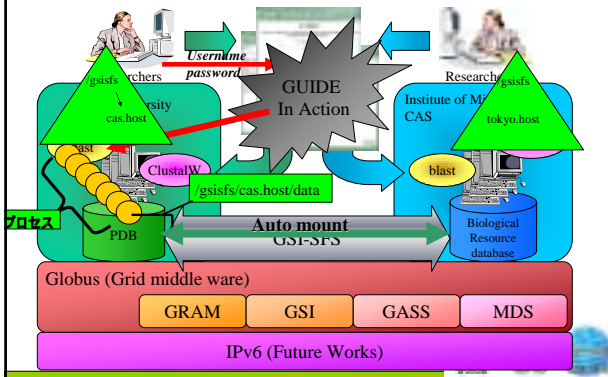


- **SDI (Single Disk Image)**
- **Single Sign-on**
- **On-demand Remote Filesystem Access**
- **データ機密性を確保**
- **ユーザ利便性**

グリッド環境上で機密データの共有を実現!



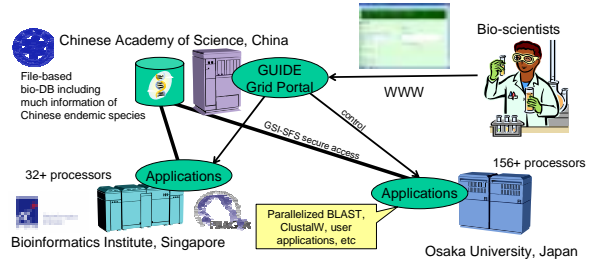
## 中国科学院-阪大リサーチテストベッド動作



## A testbed on bioinformatics on sc2003



- **Connected data and computational resources**
  - Chinese valuable biological database
  - Large computational resources provided by our cluster and PRAGMA collaborators



## Research Testbed のまとめ



- 中国CAS, シンガポールBIIのデータ資源と計算資源を統合
- GSI-SFS + GUIDEにより、データ、計算資源の位置に依存しないBLAST, ClustalWが実現
- 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

じゃあ今後の方針は？この環境がなににかされるのか？

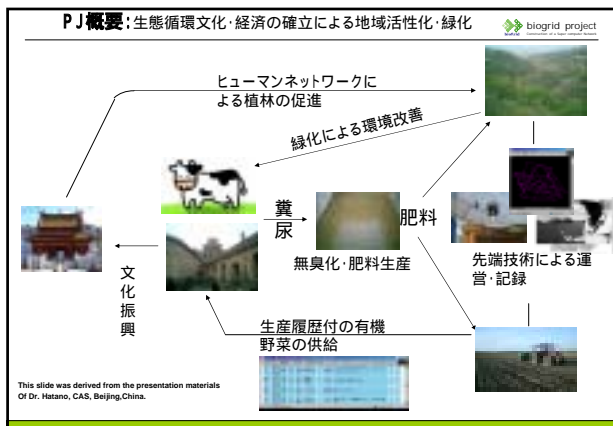
## 黄土高原の砂漠化



都会の発展  
高原の民は貧しい。



砂漠化の影響



**導入される技術**

- ・北海道自然学研究所・青木電器・DGC総研の環境浄化技術(北海道小清水方式)
- ・未来農業集団(北海道芽室町)のIT農業
- ・中央農研・三重大学生物資源学部の遠隔監視技術
- ・中国科学院微生物研究所・北海道農研の土壤・生物資源管理技術
- ・大阪大学サイバーメディアセンターの次世代インターネット技術

This slide was derived from the presentation materials Of Dr. Hatano, CAS, Beijing,China.

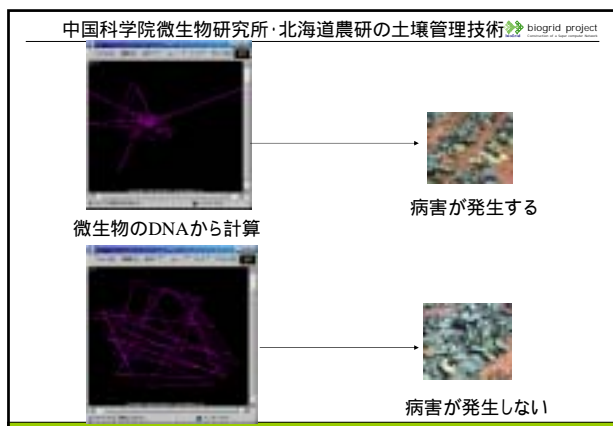
**中央農研・三重大学生物資源学部の遠隔監視技術**

・温度、湿度、現場映像をモニタリングするフィールドサーバー

モニター画面(Web)

インターネット接続型気象ロボット

全世界同時収集 & 配信



**大阪大学サイバーメディアセンターの次世代インターネット技術 6Grid**

**【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れ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)

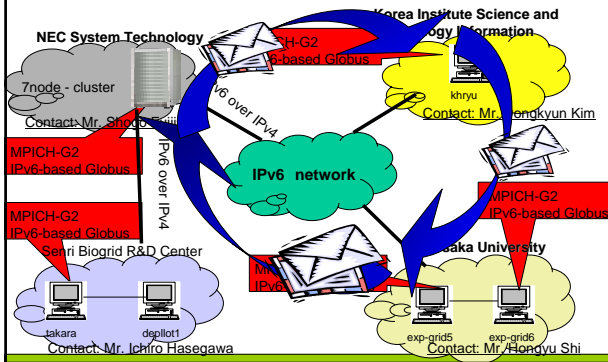
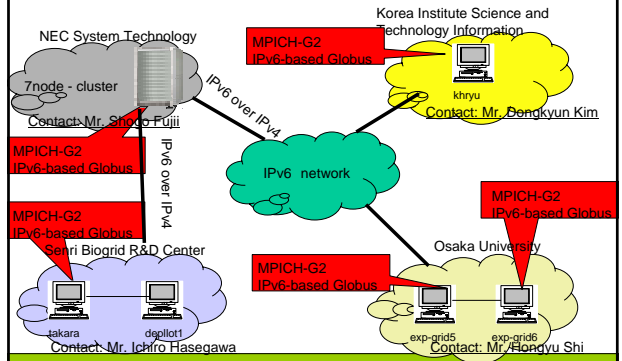
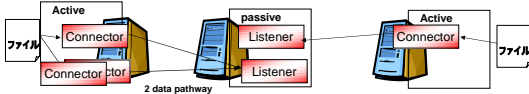
**【これまでの関連成果】**

- [-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

- 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)

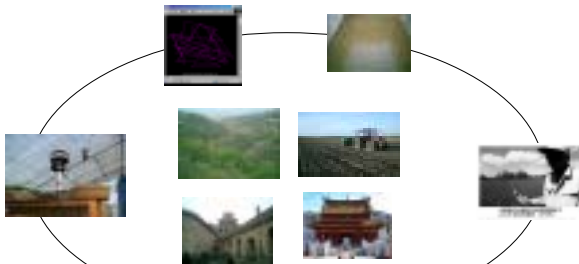


- Striped Transfer**: 複数計算機間で並列的にデータを転送 (実装中)

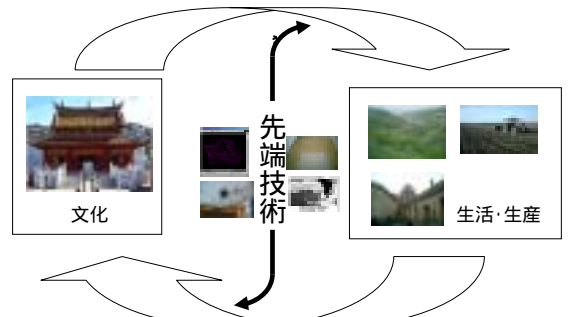


- 創薬研究過程において、比較的使用頻度が高く、共通性の高いソフトウェア、ツールのグリッド化
  - BLAST (相同性検索)
  - ClustalW (多重整列と系統樹作成)
- “Gridified” toolへのポータル提供
  - ユーザフレンドリなGUIにより複雑なグリッドのメカニズムを隠蔽

Grid User Interface to the Distributed Environment



さまざまな研究機関、大学などで開発された要素技術を、大阪大学サイバーメディアセンターのグリッド技術によって総合情報化していくことにより、農業、文化への技術のフィードバックをおこなっていく



- APAN Honoluluデモンストレーションにむけ奮闘中。
  - 黄土高原 - (tun6to4 in openvpn) IMCAS - (tun6to4 in openvpn) KDD <導通確認>
  - バイオグリッド6Grid solutionのdeployと、阪大へのIPv6接続 <現在も作業中>

大阪大学サイバーメディアセンターでは、台湾、韓国、中国の研究機関と密な連携をとり、大阪大学サイバーメディアセンターのグリッド技術を利用した共同研究開発を行い、今後の成果への胎動が感じられる。

**[アジアグリッドイニシアティブの目的]**

本課題ではアジア地域におけるグリッドの先導的研究者を集め、アジア地域におけるネットワークおよびグリッド技術の協調的研究開発を促進する

- 本スライドの一部は、中国科学院微生物研究所信息网络中心 馬俊才氏よりご提供いただいたものを利用させて頂きました。利用させて頂きました資料作成にかかわっております皆様方に感謝いたします。