

広域分散アプリケーション特論

2005前期 月曜 3時限

場所: 情報基盤センター3F 多目的講義室

担当 青柳 睦

aoyagi@cc.kyushu-u.ac.jp

5月30日(月)

講義の内容, 成績評価方針(server-500.cc.kyu...)

サイエンスGrid NAREGI

Globus, Unicoreの現状

Globus共通基盤層(続): MDS, GRAM

Unicore基本層: イントロダクション



講義の内容

■ グリッドの概要

- Gridコンピューティングとは
 - サイエンス分野での利用
 - ビジネス分野での利用

■ 計算科学の概論

- 主要なシミュレーション手法

■ サイエンスGrid NAREGI

- Globus, Unicoreの現状
- NAREGIミドルウェア概要
- 連成計算とその類型化

■ テーマ 考え中

GT4の動向に依存・・・

講義資料はWebで公開
server-500.cc.kyushu-u.ac.jp



資源情報管理（MDSの概要）

- MDS (Monitoring and Discovery Service)
 - Gridの運用やアプリケーションの構成には、システムの情報が必要不可欠
 - どのリソースが利用可能か？
 - Gridの“状態”は？
 - リソースの利用をいかに最適化するか
 - これらの疑問に答えるためには汎用の情報インフラが必要
 - スケーラビリティ、汎用性

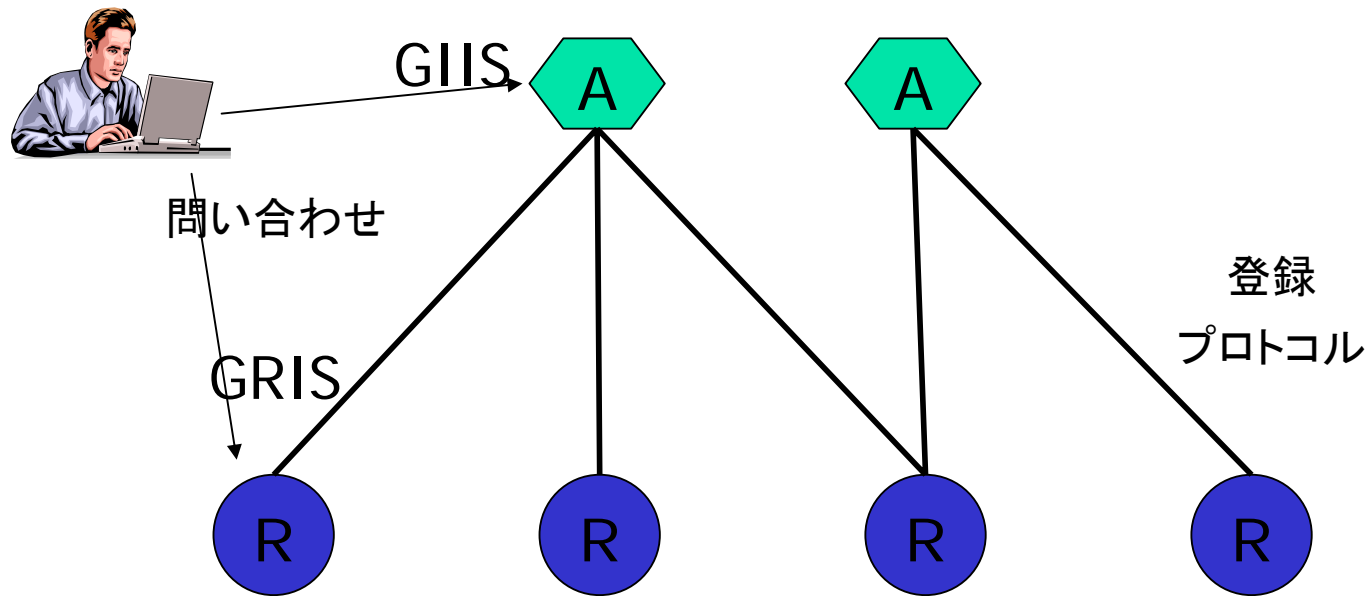


管理情報の例

- 計算リソースの特徴
 - IP アドレス、使用可能なソフトウェア、システム管理者、接続しているネットワーク、OSバージョン、負荷状態
- ネットワークの特徴
 - バンド幅, レイテンシ、プロトコル、論理的なトポロジ
- Globusインフラストラクチャの特徴
 - ホスト、リソースマネージャ

MDS のアーキテクチャ

カスタマイズされた集合ディレクトリ

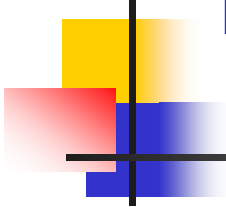


各サーバで稼動するリソースサービス

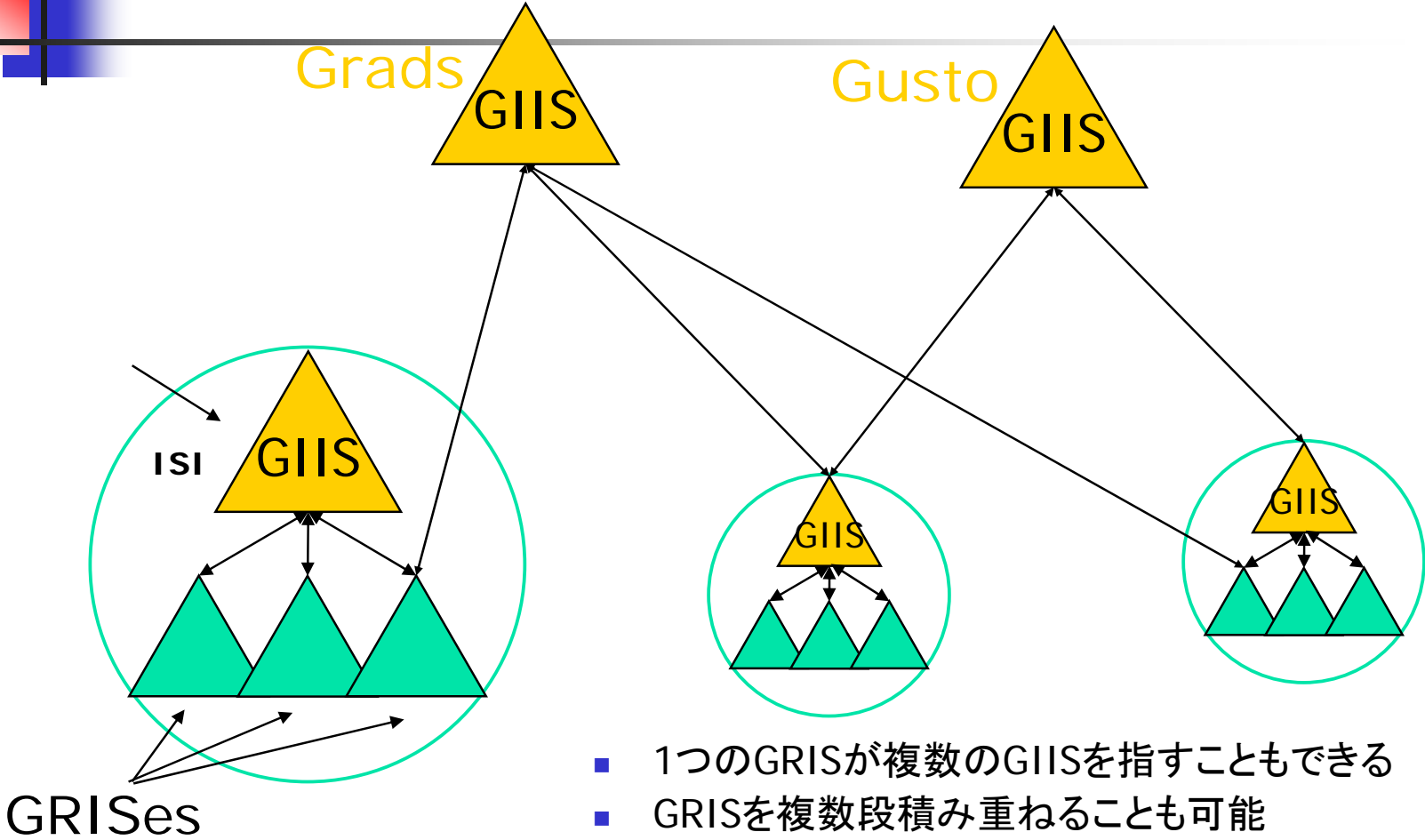


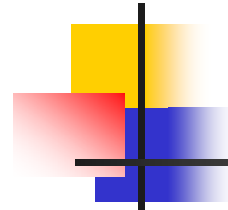
2階層のMDS サーバ

- Grid Resource Information Service (GRIS)
 - 特定のリソースに関する情報を提供
 - 複数の情報プロバイダをサポートするように設定可能
 - 問い合わせにはLDAP プロトコルを用いる
- Grid Index Information Service (GIIS)
 - 複数のGRISサーバで集めた情報を提供
 - 複数のGRISサーバに分散した情報を効率的に問い合わせることを支援
 - 問い合わせにはLDAP プロトコルを用いる



MDS の論理的な配置





LDAP

- Lightweight Directory Access Protocol
 - IETF 標準
 - X.500 DAP プロトコルを軽量化したもの
 - 分散ストレージへの分散アクセスをサポート(リフェラル)
 - 認証とアクセス制御をサポート
- LDAPは以下を定義する:
 - ディレクトリの内容へアクセスするネットワークプロトコル
 - 情報の形を与える情報モデル
 - 情報の参照、構造化に用いる名前空間



Example: CPU負荷値の検索

- 計算資源のCPU負荷のフィールドを取得

```
% grid-info-search "(ObjectClass=MdsCpuFree)" dn Mds-Cpu-Free-1minX100
```

```
# processors, brain-n.a02.aist.go.jp, local, grid  
dn: Mds-Device-Group-name=processors, Mds-Host-hn=brain-  
n.a02.aist.go.jp,Mds-Vo-name=local,o=grid  
Mds-Cpu-Free-1minX100: 078
```



資源情報管理のまとめ

- MDS = GIIS + GRIS
 - 階層化することで比較的スケーラブルに
- プロトコルはLDAP
 - 条件を指定して絞込みをすることができる

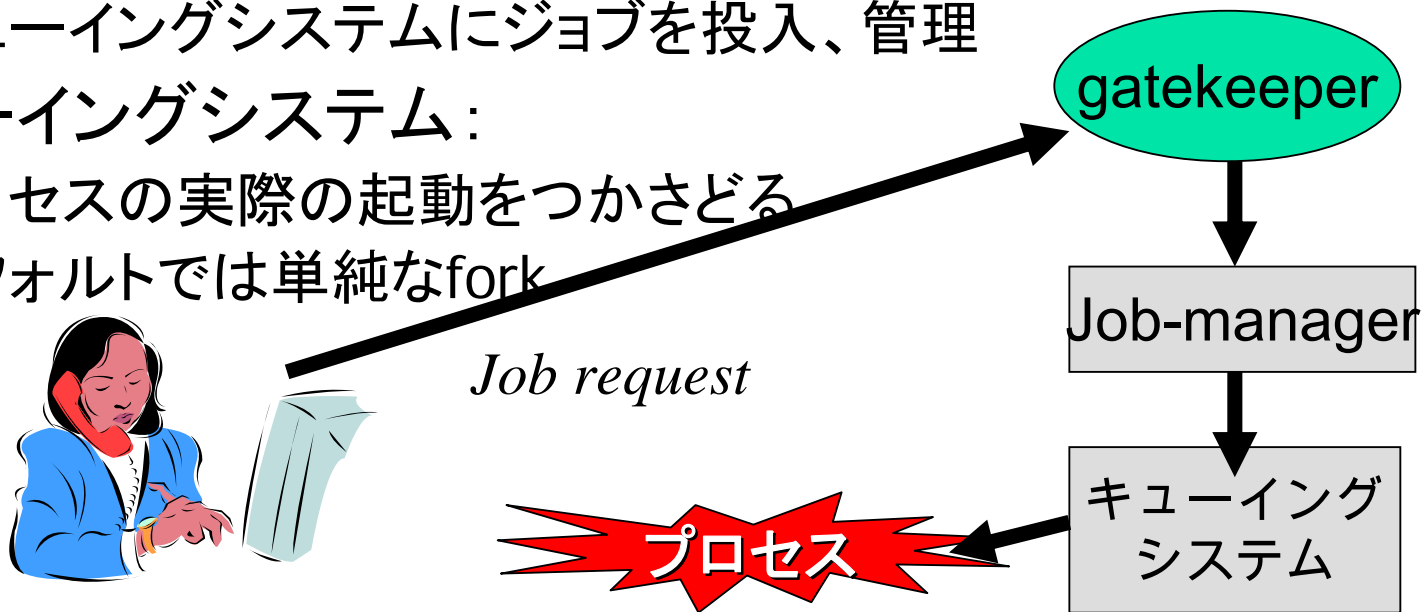


リモートホスト上でのプログラムの起動

- GRAM (Grid Resource Allocation Manager)
 - リモートサーバ上でプログラムを安全に実行
 - GSIを用いた認証、マップファイルを用いた認可
 - シングルサインオン
 - Unixアカウントと連動した認可
 - GASSと連動したプログラムの転送と入出力のリダイレクト
 - 実行ファイルや引数ファイルを自動的にステー징

GRAMの構成

- Gatekeeper:
 - 計算機に常駐し、クライアントからのジョブ要求を待つ
- Job manager:
 - キューイングシステムにジョブを投入、管理
- キューイングシステム:
 - プロセスの実際の起動をつかさどる
 - デフォルトでは単純なfork





Gatekeeper の役割

- リクエストを受け付ける
 - GSIを用いた認証
 - マップファイルに基づく認可
- ユーザの権限でJob Managerを起動
 - Job Managerが実際にプロセスを起動
- Gatekeeper自体は inetd, xinetd などから起動されるようにするのが一般的

grid-mapfile

■ Global Idからlocal accountへのマッピング

"/C=US/O=Globus/O=Electrotechnical Laboratory/CN=Yoshio Tanaka"	yoshio
"/C=JP/O=AIST GTRC/CN=Sugree Phatanapherom/Email=g4465027@ku.ac.th"	sugree
"/O=Grid/O=Globus/OU=apgrid/CN=Hidemoto Nakada"	nakada
"/O=Grid/O=Globus/OU=apgrid.org/CN=Osamu Tatebe"	tatebe
"/O=Grid/O=Globus/OU=a02.aist.go.jp/CN=Ninf Demonstrator"	ninfdemo
"/O=Grid/O=Globus/OU=itbrain.co.jp/CN=Hirofumi Miyamoto"	gcomp
.....	

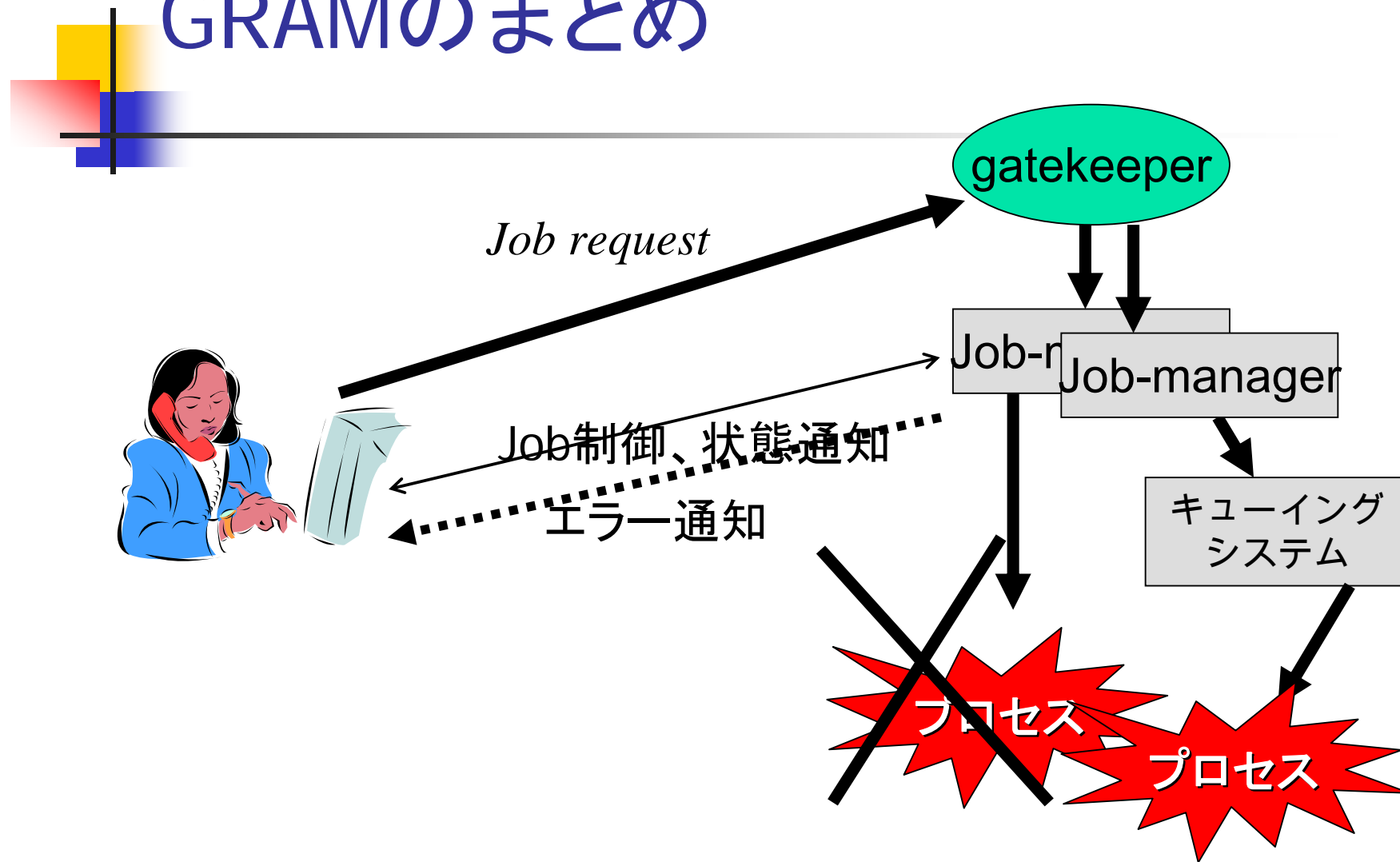
- このファイルにかかれていなければGRAM経由でプロセスを起動することはできない



Job Managerの役割

- プロセスの状態を監視、管理
 - クライアントからのリクエストで状態を報告
 - API上は、自動的にコールバック関数が呼ばれる
 - キューに入っている、アクティブ、失敗、終了
 - プロセスの制御
 - 強制終了
- キューイングシステムを抽象化
 - 同じインターフェイスでさまざまなキューイングシステムが使用できる
 - Fork, PBS, Condor,

GRAMのまとめ





SSHとの相違

- GSIによるシングルサインオン
 - さらに別のサーバに接続することが可能 proxy
 - ホストとユーザ名の対応をユーザが管理する必要がない
- サーバ側でローカルなスケジューラの使用が可能
- リモートジョブの詳細な管理
 - 状態の検出、ジョブの破棄など。
- クライアントとの間にストリームができるわけではない
 - Globus-IOなどでクライアントにコールバック接続する必要がある
 - ファイアウォールとの親和性が低い原因のひとつ



ホスト間のデータ転送

- 要請
 - セキュリティ – シングルサインオン
 - 高レイテンシ、高スループット環境下での効率的な転送
- GridFTP
 - GSI+FTP 、GGFで標準化
 - ストライピングによる高速転送、第3者転送をサポート
- GASS
 - GSI+HTTPベースの古いモジュール
 - GRAMとの連動以外ではあまり使用されなくなりつつある



GridFTP: 基本的なアプローチ

- FTP プロトコルの標準は、幾つかのIETFのRFCによって規定されている
- もっとも良く用いられているサブセットからはじめる
 - get/put など, サードパーティ転送
- 標準だがあまり用いられていない機能を実装
 - GSSバインディング, 拡張ディレクトリのリスト, 簡単な(シンプル) リスタート
- 既存のサーバとの相互運用性を保ちつつ、様々な方向性に拡張
 - ストライピング/並列データチャンネル、部分ファイルアクセス、自動/手動 TCPバッファの調整、転送状態のモニタリング、拡張リスタート



GridFTP プロトコル仕様

- 既存の標準
 - RFC 949: File Transfer Protocol
 - RFC 2228: FTP Security Extensions (セキュリティ拡張)
 - RFC 2389: Feature Negotiation for the File Transfer Protocol
 - Draft: FTP Extensions
- 新しいドラフト
 - GridFTP: グリッド向けのFTPのプロトコル拡張
 - グローバルグリッドフォーラム (Global Grid Forum)の Data Working Groupにて議論・標準化



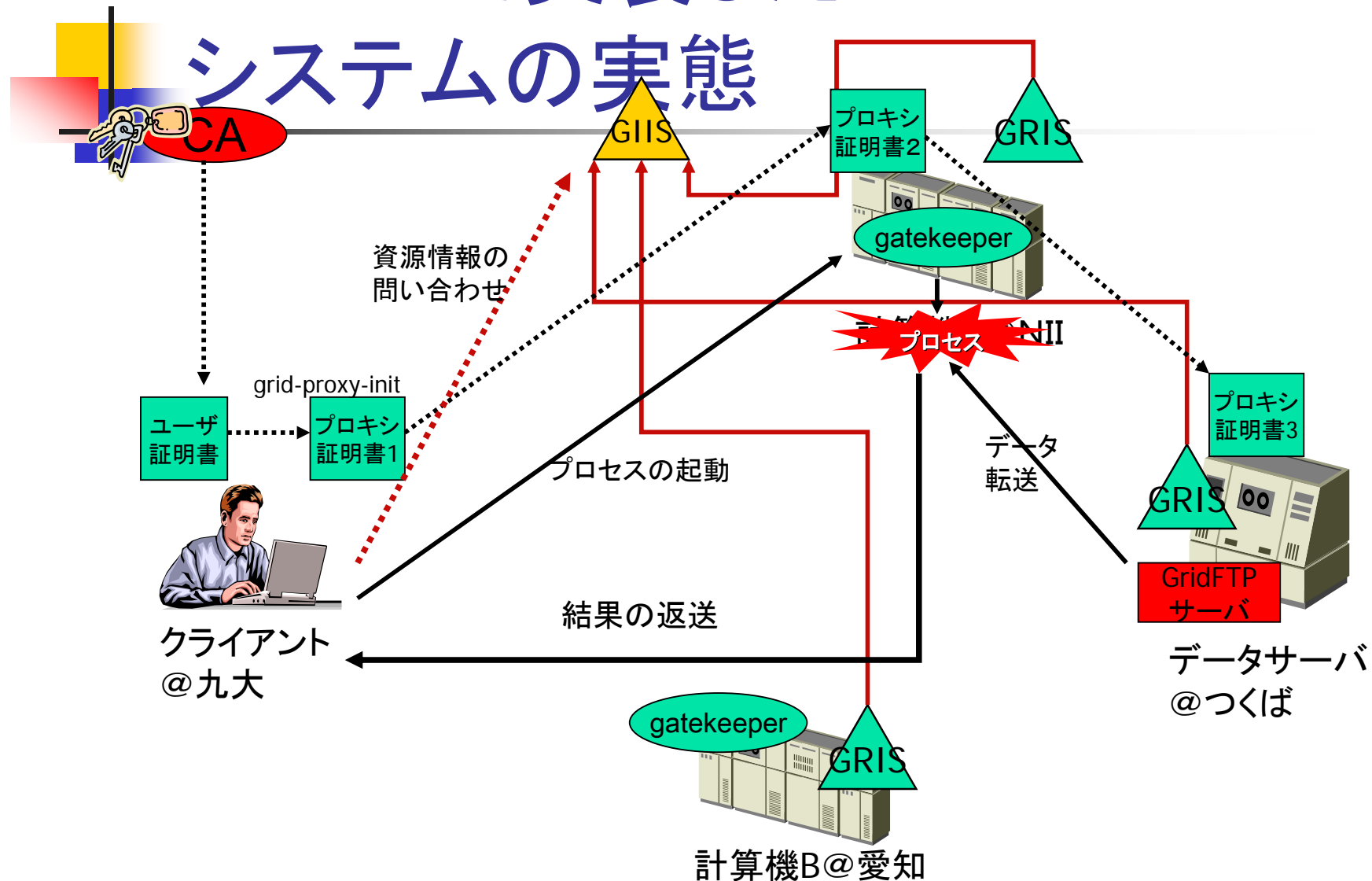
GridFTP のツール群

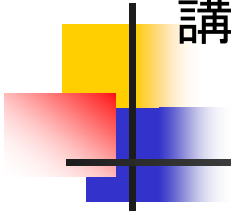
- 既存のFTPプログラムコードへのパッチ群
 - 既存のFTPクライアントやサーバに対するGSI化するためのパッチ。特に、高品質のプロダクションコードが必要な場合のため。
 - GridFTPの一部のプロトコルのサポートのみ
- 新たに開発したライブラリ群
 - フルセットのGridFTPプロトコルの実装。高性能でカスタム化されたアプリケーションをターゲット
- 新たに開発したツール群
 - 新たな機能を持ち高性能なGridFTPサーバやクライアント

GASS (Globus Access to Secondary Storage)

- ファイル転送コンポーネントのひとつ
- GRAMとうまく協調動作
 - ファイルや実行のステー징ングやI/Oのリダイレクションをサポート
- GSI化されたHTTPプロトコルを使用
 - 必要に応じてデータ転送を行うキャッシングのアルゴリズムを実装している。
- `globus_gass`, `globus_gass_transfer`, `globus_gass_cache` などのAPIがある

Globusで実装した システムの実態





次週(6/6)だけ 工学部防音講義室201にて
講義を行います. 6/13日以降はまたここ. 基盤センター3F

次週は

- Unicore基本層の概要
 - NAREGIミドルウェアの概要
 - Globus, Unicore の実演
- を予定しています.

レポート課題2 (2005.5.23 出題)の×切は6月6日(月)ですので忘れないように.