

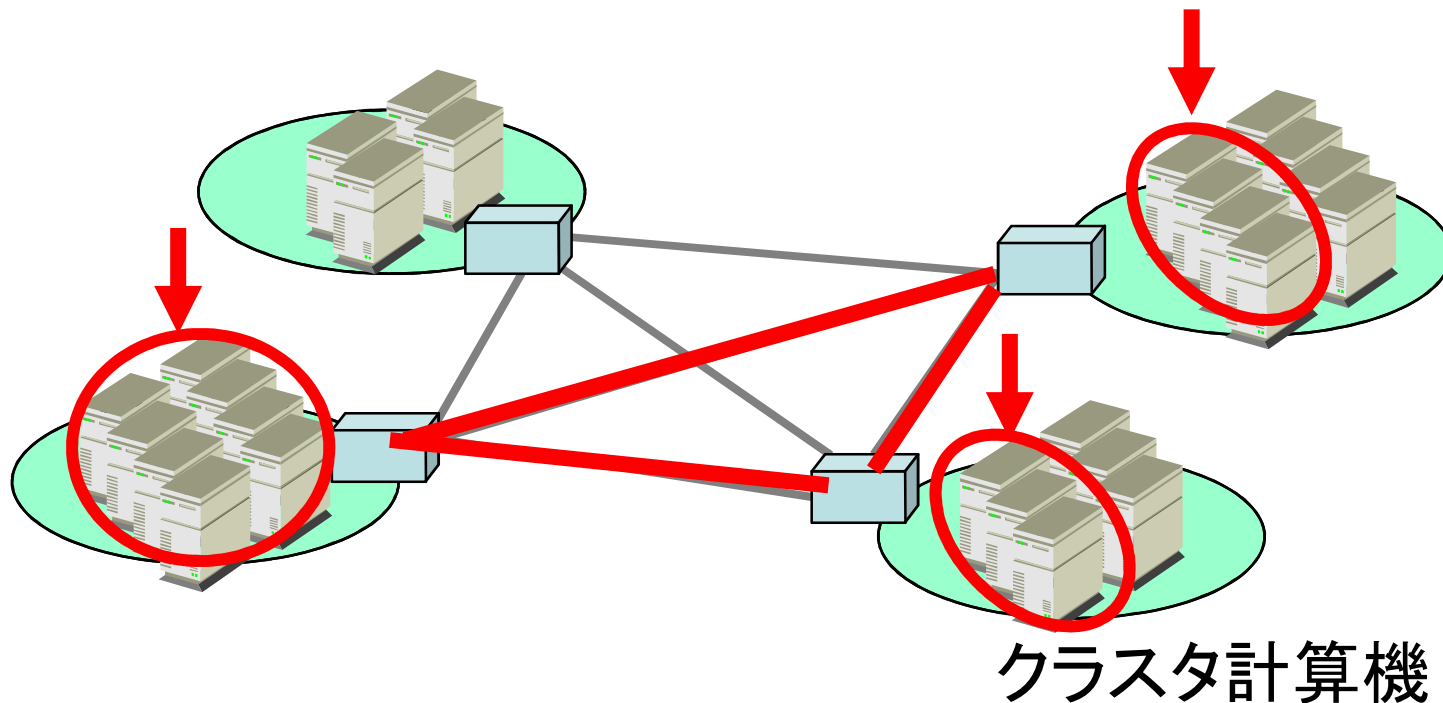
グリッドにおける計算資源と光パスネットワーク資源のコアロケーション実験

竹房あつ子¹, 林通秋², 長津尚英³, 中田秀基¹,
工藤知宏¹, 宮本崇弘², 大谷朋広², 田中英明²,
鮫島康則³, 今宿互³, 神野正彦³, 滝川好比郎³,
岡本修一⁴, 田中良夫¹, 関口智嗣¹

¹産総研グリッド研究センター, ²株式会社KDDI研究所,
³日本電信電話株式会社, ⁴情報通信研究機構

グリッドにおけるメタコンピューティング

- 異なる組織から提供される分散した計算資源を同時に利用
- 計算機の性能・負荷とクラスタ間のネットワークの通信遅延・帯域がアプリケーションの実効性能に影響



グリッド資源のコアロケーション

- グリッド上のネットワーク資源と計算資源のコアロケーション(同時確保)が重要
 - インターネットではベストエフォート型の通信を行うため、通信性能を保証できない
 - 光ネットワークでは高品質通信が提供可能だが、必要なネットワークの構成・提供は自動化されていない
 - スーパースケジューラはコアロケーションを実現するものであるが、実用段階に至っていない
- ユーザアプリケーションの性能保証が困難
- 資源を利用する際に人為的手続きが必要

事前予約によるコアロケーションの実現

- 通信性能の保証とネットワークの自動制御のため,
GMPLSベース光パスネットワークを利用
- 事前予約による計算・ネットワーク資源のコアロケーションシステムの構築
 - グリッド資源スケジューラ(GRS)
WSRFに基づき, NRMと連携してユーザの要求する計算・ネットワーク資源を事前予約でコアロケーション
 - ネットワーク資源管理システム(NRM)
ウェブサービスを介してGMPLSで制御される光パスネットワーク資源を提供
 - G-lambdaプロジェクトでGRS-NRM間のウェブサービス
インタフェースGNS-WSI (Grid Network Services / Web Services Interface)を策定
- iGrid2005で提案システムの実証実験

発表内容

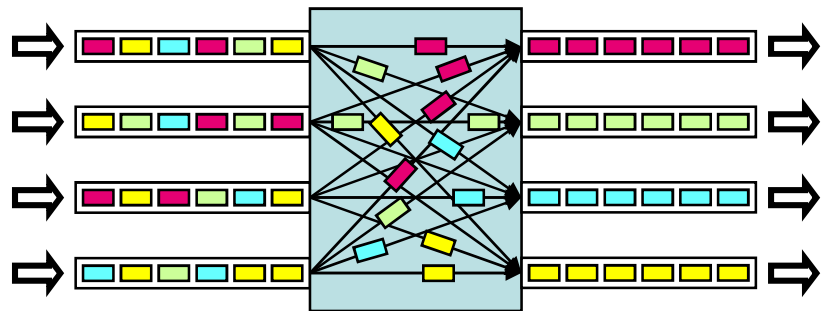
- 光パスネットワークとGMPLS
- 事前予約に基づくコアロケーションシステムの概要
 - グリッド資源スケジューラ(GRS)
 - ネットワーク資源管理システム(NRM)
 - GRS-NRM間インタフェースGNS-WSI
 - G-lambdaプロジェクト
- iGrid2005におけるコアロケーション実験
 - デモ動画
 - 議論
- 関連研究
- まとめと今後の課題

光パスネットワークとGMPLS

- **光パスネットワークで通信品質が保証可能**
 - 送信側と受信側を光パスで接続
 - データが破棄される可能性のある装置を中間に持たない

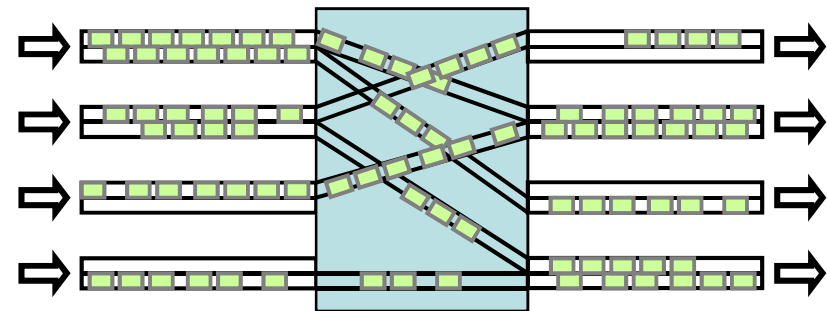
パケットネットワークスイッチング

→オーバーヘッド大, パケットロスあり



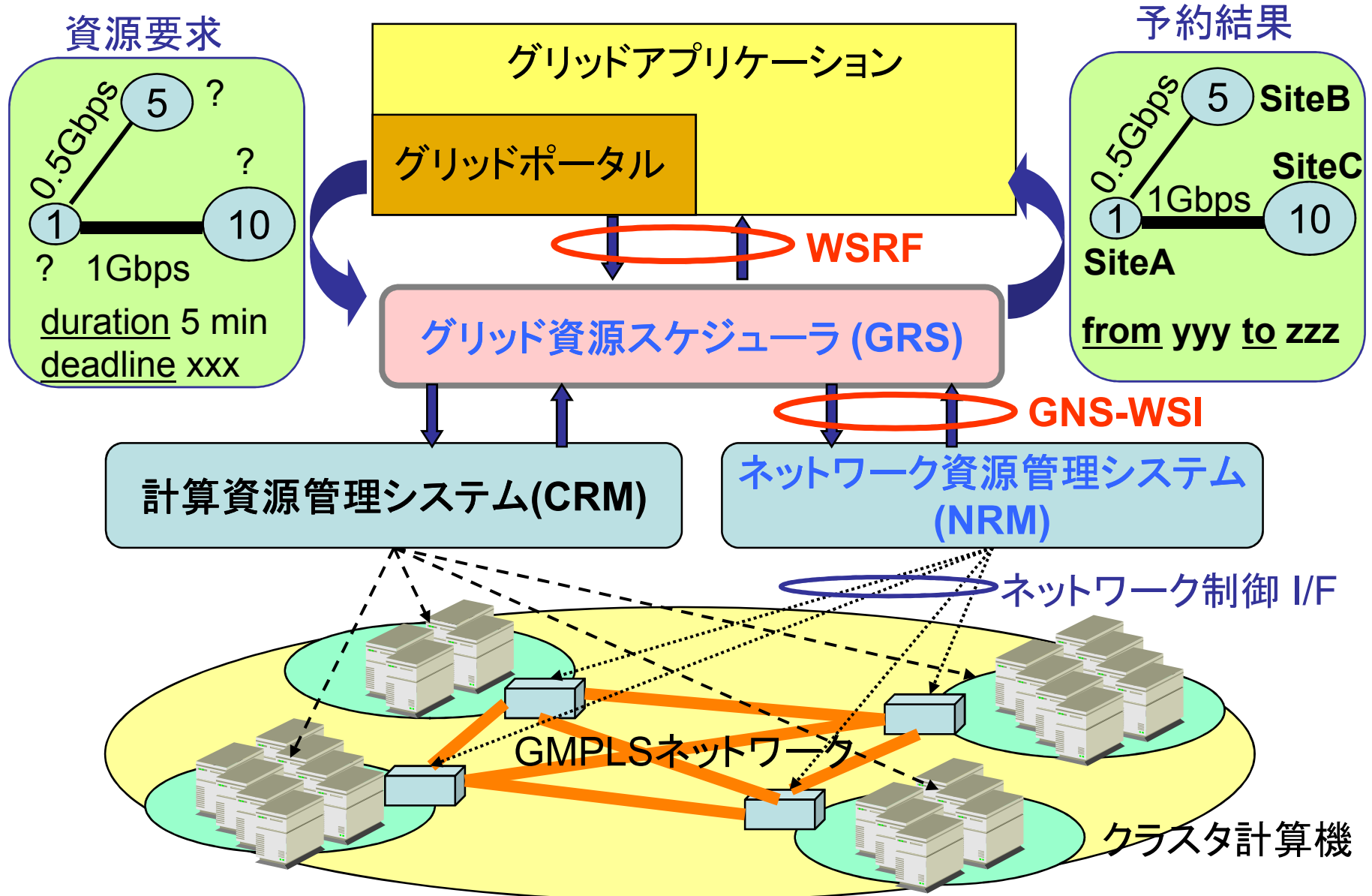
光パスネットワークスイッチング (光クロスコネクタ(OXC))

→オーバーヘッド小, パケットロスなし



- **GMPLS (Generalized Multi-Protocol Level Switching)**
 - 光パスネットワークを制御するためのプロトコル
 - **自動制御**が可能

事前予約ベーススコアロケーションシステム



グリッド資源スケジューラ (GRS)

- 産総研が開発
- WSRFに基づく実装
 - WSRF: ウェブサービスにおけるステートフルリソースのモデリングとアクセスのための仕様でOASISで標準化
 - 参照実装の1つであるGT4 (Globus Toolkit 4)を用いてGRSを実装
- ユーザの要求を満たす資源を選択し, 事前予約
 - ユーザ資源要求: クラスタ数, 各クラスタのCPU数, クラスタ間バンド幅, 予約時間, デッドラインなど
 - NRMと各サイトのCRMと連携し, ユーザの要求を満たす計算資源・ネットワーク資源(光パス)を選択・事前予約

ネットワーク資源管理システム (NRM)

- KDDI研究所が開発
- GNS-WSIを介してGRSのリクエストにレスポンス
- ネットワーク資源の仮想化
 - 物理パスの詳細を隠蔽
 - エンドポイント間のパスとして提供
 - e.g. パス: AKB-TKB, バンド幅: 1Gbps
- 光パスの自動制御
 - GMPLSプロトコルを用いて予約開始時刻にパスを利用可能にし, 予約終了時刻にパスを解放

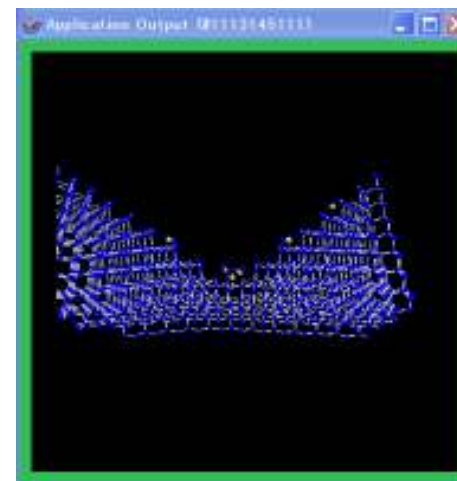
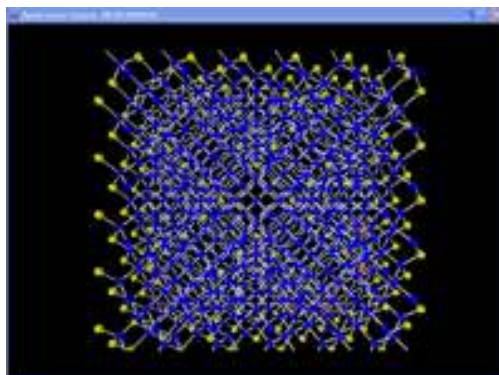
GNS-WSI (Grid Network Services / Web Services Interface)

- **GRS-NRM間のウェブサービスインタフェースとして**
G-lambdaプロジェクトで規定
 - 産総研, KDDI研, NTTの共同プロジェクト(2004.12~)
<http://www.g-lambda.net/>
 - 初期インタフェース(GNS-WSI ver. 1)を規定 → **実装**
- **ポーリングベースのオペレーションを提供**
 - エンドポイント間のパスの事前予約
 - 予約の修正 (予約時刻や予約時間の変更など)
 - 予約状況に関する問い合わせ
 - 予約のキャンセル

ユーザアプリケーションプログラムの概要

- Ninf-G2とGT2(Globus Toolkit 2)を用いて実装された分子動力学シミュレーション
 - Ninf-G2 はGGF (Global Grid Forum) で規定された標準Grid RPC programming APIを提供
 - GT2のジョブ投入・通信機能を利用
 - 広域ではマスタ・ワーカ, クラスタ内ではデータ並列処理
- シミュレーションシナリオ
 - シリコン原子のエネルギー伝播
 - シリコンの水分子による腐食過程

Start



議論 (1/2)

- GRS-NRM間のSOAP通信遅延
 - 資源探索時にGRS-NRM間で複数回メッセージを交換
 - GRS-NRM間の通信遅延, SOAPオーバーヘッドが応答時間に大きく影響
 - 通信回数を減らすよう, 探索手法, GNS-WSIを改良
- 耐故障性とエラー通知機能
 - 予約時刻にパスが利用可能にならないケースあり
 - Ninf-G2とGT2で実装されたプログラムはエラーが検知できなかつたため, 実行開始時刻を遅らせたり, GRSにモニタ機能を付加
 - 耐故障性の強化とユーザへの通知や保証が必要

議論 (2/2)

- 光パスネットワークのデバッグ
 - GMPLS光パスネットワークのデバッグ手法が未熟
 - 本実験ではネットワークの予約状況地図画面が有効
 - 光パスネットワークレベルのデバッグ手法の改良・標準化が重要
- 光パスネットワークのルーティング
 - 動的な経路設定では指定された遅延条件が満たせない
 - 動的ルーティング機能の有効利用のため、通信遅延に関する要求条件の明確化が必要

関連研究

- CANARIE
 - カナダにおいて教育目的の光パスネットワークCA*net4を構築
 - ウェブサービスインタフェースで光パスネットワークを直接操作可能にするUCLP(User Controlled LightPaths)ソフトウェアを開発
 - 事前予約, 計算機との連携はなし
- VIOLA
 - UNICOREベースのグリッド環境でGGFのWS-Negotiation/-Agreement仕様に基づくメタスケジューラを開発
 - 実装の詳細は明らかでない
- グリッドスーパースケジューラ
 - Silver, CSF, NAREGIスーパースケジューラなど
 - 実用段階になく, 計算資源の割当に着目

まとめ

- GMPLS光パスネットワークを用いて**高品質通信を保証するネットワークの提供**
- **事前予約による計算・ネットワーク資源のコアロケーションの実現**
 - GRSがNRMと連携し、計算・ネットワーク資源を事前予約でコアロケーション
 - NRMによりGMPLSで光パスネットワーク資源を制御し、ウェブサービスを介して提供
 - G-lambdaプロジェクトで規定したGRS-NRM間のウェブサービスインタフェースGNS-WSIを実装
- iGrid2005での実証実験から、**提案システムで適宜高品質計算環境が提供可能**であることを示した

今後の課題

- GRS
 - 事前予約機能付ローカルスケジューラとの連携
 - スケジューリングアルゴリズムの改良(プライオリティ, フェアネス)
 - 資源予約の耐故障性
 - スケーラビリティ (複数GRS, 複数NRM)
- NRM
 - 複数NRMの連携
 - QoSと耐故障性
- GNS-WSIの改良・拡張→標準化へ

謝辞

実証実験にご協力いただいた方々に深く感謝いたします

産業技術総合研究所

谷村勇輔氏

武宮博氏

岡崎史裕氏

KDDI研究所

鈴木正敏氏

鶴沢宗文氏

NTT未来ねっと研究所

大原拓也氏

築島幸男氏

NICT 大阪JGN II RC

下條真司氏

秋山豊和氏

東京大学

小林正朋氏

G- lambda project

<http://www.g-lambda.net/>