

NAREGI本格運用による e-サイエンス/CSIの実現及び 研究開発

東京工業大学
学術国際情報センター

実施内容

1. 実運用システムへのグリッドミドルウェア導入 & 運用

- 東工大スパコンTSUBAMEへのNAREGIグリッドミドルウェア導入 & 運用

2. 広域分散ファイルシステムの構築 & グリッド連携

- 多拠点配備したRENKEI-POPストレージサーバにGfarmを導入

3. グリッドの動作整合性を監視するシステムを開発

- NAREGIグリッド、Gfarm分散ファイルシステムを監視

1の計算環境、2のデータ共有環境からなるe-Science/CSI
インフラの管理・運用を、3の監視システムで支援する

実運用システムへの グリッドミドルウェア導入 & 運用

- 目的: 9大学 + NIIで構築を進めている学術グリッドへのスパコンリソース提供による本格運用
 - スパコンユーザ・グリッドユーザのアカウント、課金情報の統合管理
 - スパコンサービスが提供する、大規模並列、多数のアプリ等をグリッド利用者にも開放
- 達成: TSUBAMEのプロダクションキューへのNAREGIグリッドジョブ投入
 - 最大並列度: 118ノード、1888CPUコア
 - TSUBAME用GridVMによる、TSUBAME計算サービスへのフルアクセス

グリッド配備・運用TF
メンバー



NII

北海道大学

東北大学

筑波大学

東京大学

東京工業大学

名古屋大学

京都大学

大阪大学

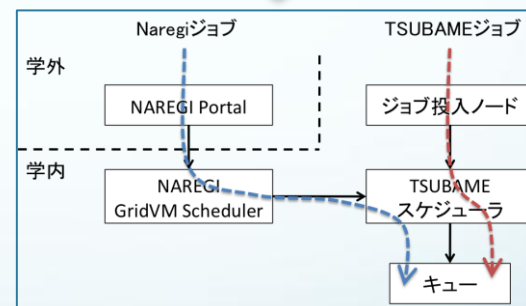
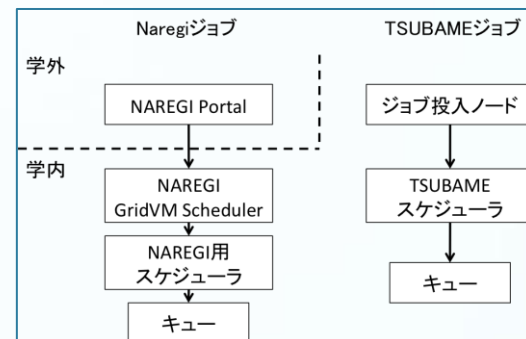
九州大学

実運用システムへの グリッドミドルウェア導入 & 運用

- 目標達成までの作業

2009/2 ~ 2009/7	TSUBAME用NAREGI GridVMを独立環境で実行
2009/8 ~ 2010/1	TSUBAME上のテストキューにNAREGIジョブが投入されるようGridVMを構成
2010/2 ~	TSUBAMEの1プロダクションキューにNAREGIジョブが投入されるようGridVMを構成

- TSUBAME計算ノードにGridVM Engineを導入せずに運用
- キューにDisableノードが1つでもあるとキューが認識されないGridVMの問題を修正
- 通年で28名の学内外グリッド運用評価用アカウント発行
 - グリッド配備・運用TFで策定したグリッドパックによる



広域分散ファイルシステムの構築 & グリッド連携

- 目的: 高速SINET網を活用したセンター間データ共有基盤の構築
 - RENKEIプロジェクト(文科省e-Science委託事業)と連携
- ストレージサーバRENKEI-PoP (Point of Presence)
 - 大容量、高速IO性能を備えたサーバアプライアンス



CPU	Core i7 975 Extreme (3.33 GHz)
Memory	12GB (DDR3 PC3-10600 , 2GB*6)
NIC	10GbE (without TCP/IP Offload Engine)
System Disk	500GB HDD
SSD RAID	30TB (RAID 5, 2TB HDD x 16)

- 全9拠点に配備、110TBのGfarmストレージとして利用可能

東京工業大学	大阪大学
国立情報学研究所	高エネルギー加速器研究機構
名古屋大学	筑波大学
産業技術総合研究所	



広域分散ファイルシステムの構築 & グリッド連携

- 学術グリッドの認証基盤の利用が可能
 - グリッドパックにより取得した学術グリッド用アカウントをRENKEI-POP Gfarmでも利用可能
- RENKEI-POPとTSUBAMEストレージ、テープライブラリを用いたデータインテンシブコンピューティング利用者の開拓が進行中
 - 例: 大規模ゲノム情報生産・高度情報解析プロジェクト
 - 新型DNAシーケンサによる小容量(数MB ~ 数百MB)大量データのバースト的な生成
 - データベース化し、遺伝研、東大柏、宮崎大学、東工大で共有

学外へ



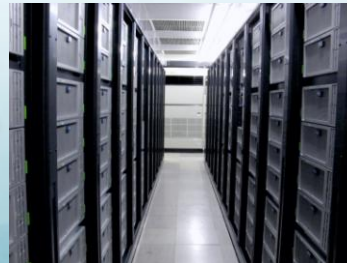
RENKEI-POP



ステージングのためのコピー



TSUBAMEストレージ



アーカイブ & リストア



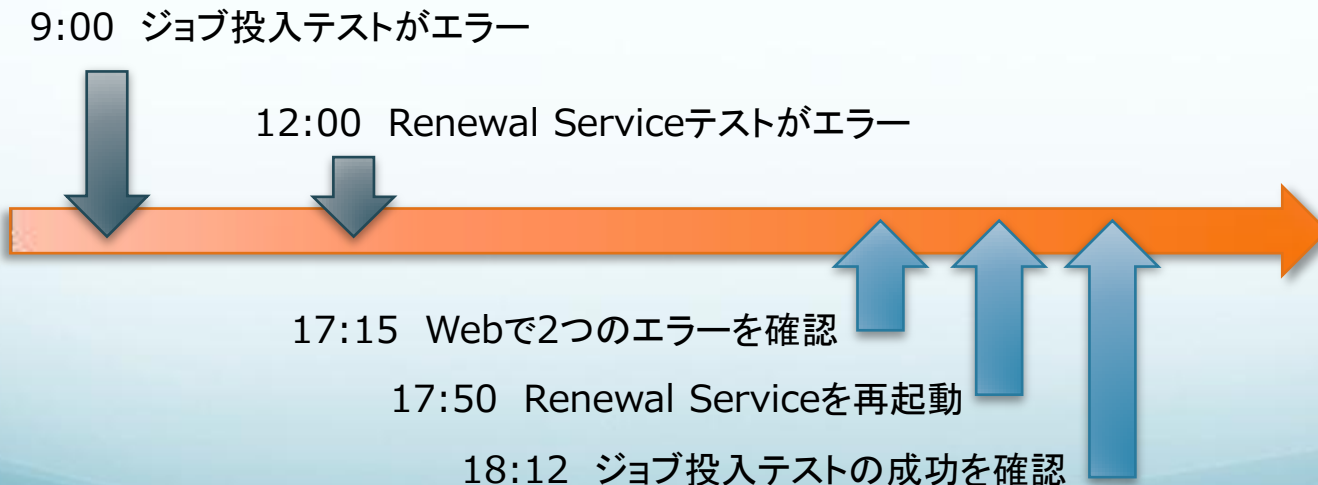
テープライブラリ



グリッド動作整合性監視

- 活用事例
 - 東工大学内NAREGI VO監視に使用
 - 学術グリッドへ提供している東工大内NAREGI関連サーバ (ISCDAS、GridVM Scheduler) の監視に使用

2009/5/21の障害発見から復旧までのタイムライン



※ 当時はメール通知が実装されておらず、障害発見までに時間がかかっている

まとめ

- NAREGIグリッドミドルウェア、Gfarm分散ファイルシステムを用いたe-Science/CSIインフラの構築と、実運用に向けた体制整備を行った
 - ゲノム解析プロジェクトなどのユーザ開拓
 - 監視システムによる運用支援
- H22年度はTSUBAME 2.0上へe-Science/CSIインフラを展開
 - TSUBAMEでの経験を活かしたNAREGIグリッドサービス構成
 - TSUBAME 2.0大容量ストレージとRENKEI-POPストレージの連携、グリッドからの利用
 - TSUBAME 2.0 ポータル、Shibboleth IdPを用いたグリッドアカウント発行の簡易化