1

NAREGI ミドルウェアの概要と サイバーサイエンスインフラ(CSI)への展開

2006年12月26日

国立情報学研究所 リサーチグリッド研究開発センター

センター長/教授

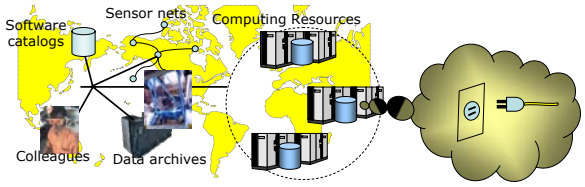
三浦 謙一

National Research Grid Initiative

2

「グリッド」とは

- 電力供給網(Electrical Power Grid)からの"アナロジー"
- 全地球的に分散した **計算資源・データ・実験装置・人的資源(研究者)**などを高速・大容量ネットワークでダイナミック(動的)につないだ仮想組織 (Virtual Organization)



National Research Grid Initiative

3

グリッドの様々な形態と利用法

- サーバグリッド
  - 資源の仮想化による ジョブの計算機やセンター間の最適な負荷分散
- メタコンピューティング
  - 分散大規模並列計算、連成計算
- データグリッド
  - 大規模データの共有と処理
- インスツルメンテーショングリッド(センサーグリッド)
  - 高エネルギー加速器、天体観測層装置、各種センサーからのデータと連携した処理、計算
- 協調グリッド
  - 研究者も含んだ使い方 (遠隔モニタリング, ステアリング)

National Research Grid Initiative

4

センターとしてのグリッド

- Resource Providerとしての計算資源/データの共有化
- 特色を持った計算資源/データベース/サービスの提供
- 研究コミュニティのVOと連携し、ホスティングすることによるユーザ層の拡大
- ユーザ登録の一元化(UPKIとの連携)

例: 米国TeraGrid (NSF)

National Research Grid Initiative

TeraGrid A National Production CI Facility

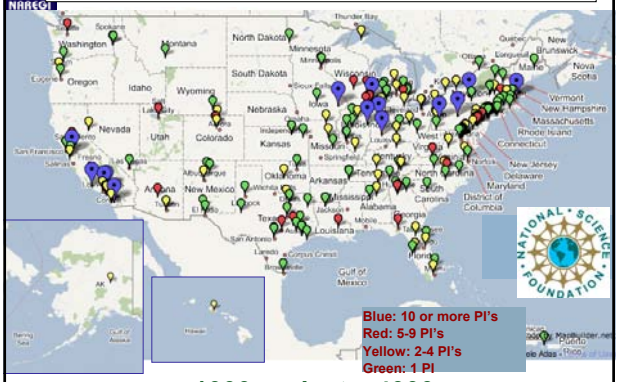
20+ Distinct Computing Resources
(>150TF today; ~600TF by 8/2007)
100+ Data Collections

Source: Dr. C.Catlett, Univ. Chicago

National Research Grid Initiative

6

TeraGrid Projects by Institution



Source: Dr. C.Catlett, Univ. Chicago

1000 projects, 4000 users

National Research Grid Initiative

7

TeraGrid Science Gateways Initiative: Community Interface to Grids

- Common Web Portal or application interfaces (database access, computation, workflow, etc), exploit standards (primarily web services)
- "Back-End" use of grid services such as computation, information management, visualization, etc.
- Standard approaches so that science gateways may readily access resources in any cooperating Grid without technical modification.

National Research Grid Initiative

8

研究コミュニティとしてのグリッド

分野別の仮想組織(Virtual Organization)

- 高エネルギー物理学
  例:EGEE, OSG
- 天文学(Virtual Observatory)
- ナノサイエンスグリッド
- バイオグリッド
- 各種ポータル/研究室環境

National Research Grid Initiative

9

NAREGIプロジェクトの概要

1)正式名称:「超高速コンピュータ網形成プロジェクト」
(英文名:National Research Grid Initiative)

2)文部科学省の「経済活性化のためのリーディングプロジェクト」の一環として推進

3) 5ヶ年計画として2003年4月にスタート(2006年、次期スパコン研究開発プロジェクトの一環として再スタート)

4) 国立情報学研究所と分子科学研究所を二拠点とし、その他の共同研究機関、大学、産業界を含む産・学・官連携研究開発体制

次世代の研究開発、製品開発に不可欠な大規模シミュレーションなどを実現するグリッド基盤用ソフトウェア(サイエンスグリッド)の研究開発を推進し、ナノ分野において計算研究環境としてのグリッドの有用性の実証を行い、産業界における製品開発のための計算環境の実現に寄与する。

10

グリッドミドルウェア研究開発体制

プロジェクトリーダ・・・・・・・・・・・・・・・・・・・・三浦謙一(国情研)

1) グリッド環境における資源管理 ・・・・・・・・・・・・ WP-1
松岡(東工大)、合田(東工大)

2) グリッドプログラミング環境 ・・・・・・・・・・・・・・・・ WP-2
関口(産総研)、石川(東大/産総研)

3) グリッドアプリケーション環境 ・・・・・・・・・・・・・・ WP-3
宇佐見(国情研)

4) データグリッド環境 ・・・・・・・・・・・・・・・・・・・・・・ WP-4
松田(阪大)

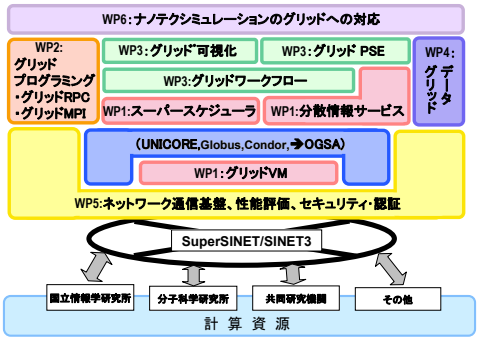
5) グリッド対応ネットワーク通信基盤 ・・・・・・・・・・ WP-5
下條(阪大)、尾家(九工大)、今瀬(阪大)

6) ナノシミュレーションのグリッド化 ・・・・・・・・・・・ WP-6
青柳(九大)

National Research Grid Initiative

11

NAREGIグリッドミドルウェア構成と研究項目



PSE: Problem Solving Environment (問題解決環境)

National Research Grid Initiative

12

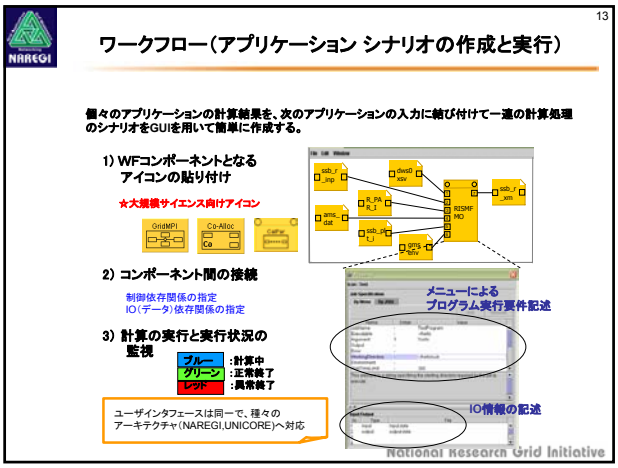
物理空間に制約されない仮想組織(VO)

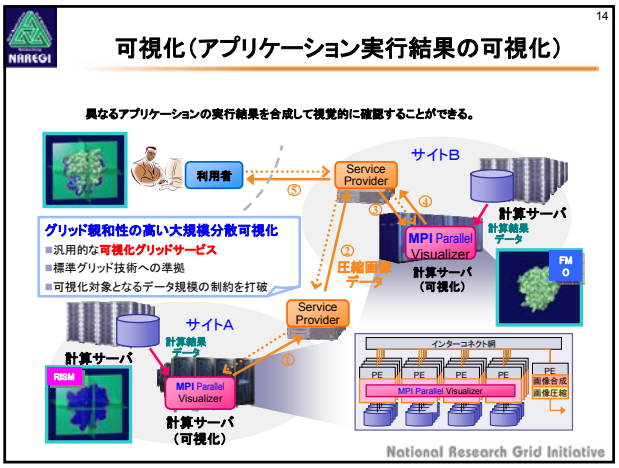
地理的・物理的に制約された環境にとらわれずに、仮想組織=VO(Virtual Organization)を形成し、共通の目的のもと各機関が連携して共同研究を進めることが可能に

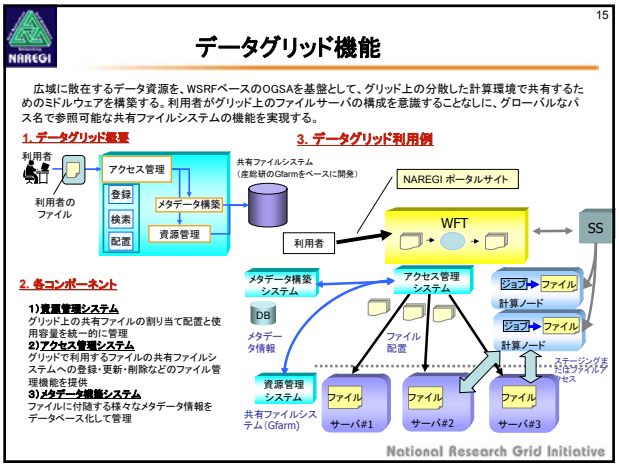
専門分野に閉じた自己完結型の研究から、複数の専門分野を集め開拓していく水平統合型の研究へ

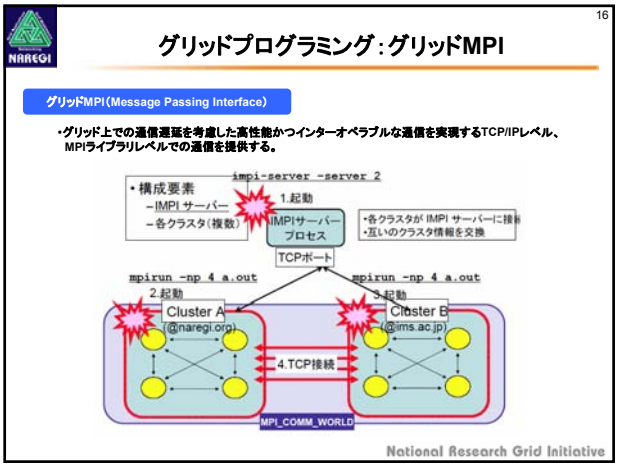
VO = 利用者+計算資源+データ

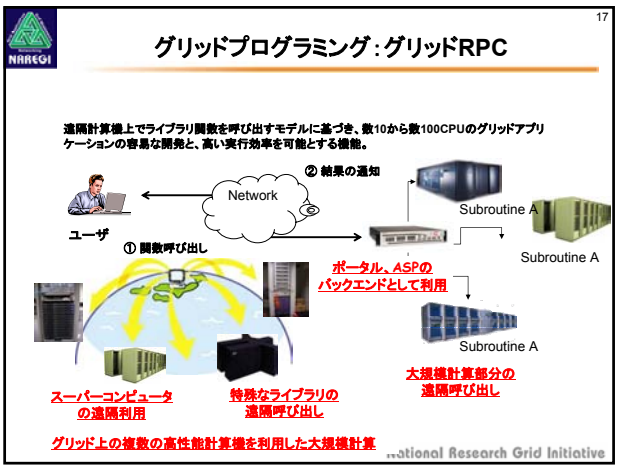
National Research Grid Initiative

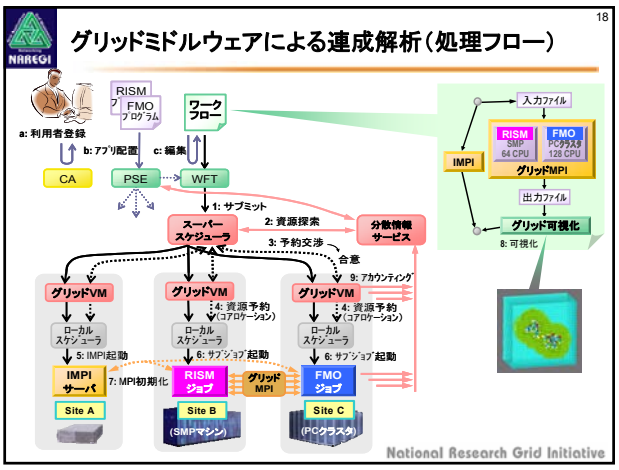

13
ワークフロー(アプリケーション シナリオの作成と実行)
個々のアプリケーションの計算結果を、次のアプリケーションの入力に結び付けて一連の計算処理のシナリオをGUIを用いて簡単に作成する。
1) WFコンポーネントとなる アイコンの貼り付け
★大規模サイエンス向けアイコン
2) コンポーネント間の接続
制御依存関係の指定
IO(データ)依存関係の指定
3) 計算の実行と実行状況の監視
ブルー :計算中
グリーン :正常終了
レッド :異常終了
ユーザインタフェースは同一で、種々のアーキテクチャ(NAREGI,UNICORE)へ対応
メニューによる プログラム実行要件記述
IO情報の記述
National Research Grid Initiative

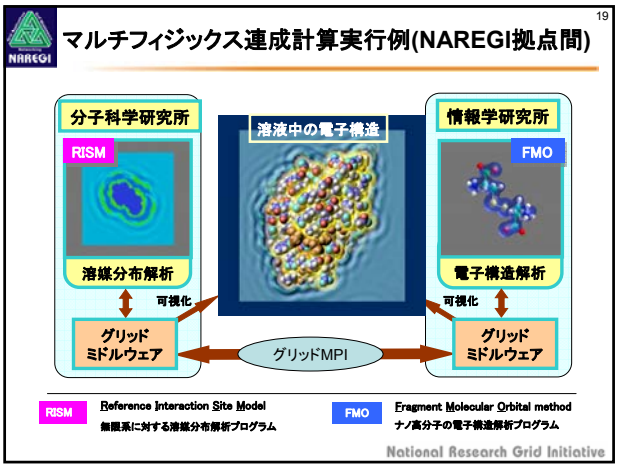

14
可視化(アプリケーション実行結果の可視化)
異なるアプリケーションの実行結果を合成して視覚的に確認することができる。
利用者
Service Provider
サイトB
計算サーバ
計算結果データ
MPI Parallel Visualizer
計算サーバ(可視化)
圧縮画像データ
グリッド親和性の高い大規模分散可視化
汎用的な可視化グリッドサービス
標準グリッド技術への準拠
可視化対象となるデータ規模の制約を打破
サイトA
RISM
インターコネクト網
PE
画像合成
画像圧縮
National Research Grid Initiative


15
データグリッド機能
広域に散在するデータ資源を、WSRFベースのOGSAを基盤として、グリッド上の分散した計算環境で共有するためのミドルウェアを構築する。利用者がグリッド上のファイルサーバの構成を意識することなしに、グローバルなパス名で参照可能な共有ファイルシステムの機能を実現する。
1. データグリッド概要
3. データグリッド利用例
利用者
利用者のファイル
アクセス管理
登録
検索
配置
メタデータ構築
資源管理
共有ファイルシステム
(産総研のGfarmをベースに開発)
NAREGI ポータルサイト
WFT
SS
2. 各コンポーネント
1)資源管理システム
グリッド上の共有ファイルの割り当て配置と使用容量を統一的に管理
2)アクセス管理システム
グリッドで利用するファイルの共有ファイルシステムへの登録・更新・削除などのファイル管理機能を提供
3)メタデータ構築システム
ファイルに付随する様々なメタデータ情報をデータベース化して管理
メタデータ構築システム
DB
メタデータ情報
アクセス管理システム
ファイル配置
ジョブ
ファイル
計算ノード
資源管理システム
共有ファイルシステム(Gfarm)
サーバ#1
サーバ#2
サーバ#3
National Research Grid Initiative

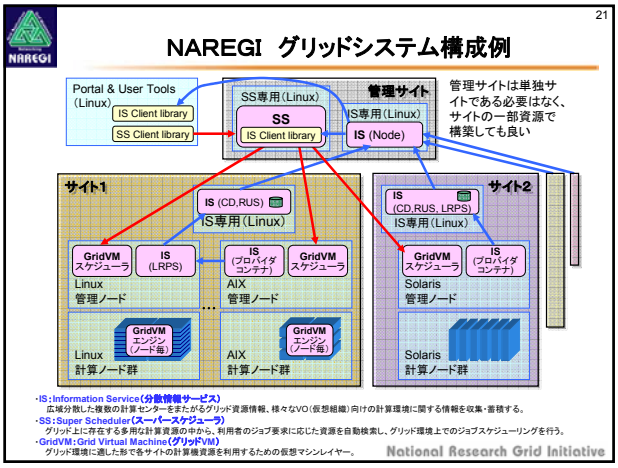

16
グリッドプログラミング:グリッドMPI
グリッドMPI(Message Passing Interface)
・グリッド上での通信遅延を考慮した高性能かつインターオペラブルな通信を実現するTCP/IPレベル、MPIライブラリレベルでの通信を提供する。
impi-server -server 2
・構成要素
-IMPI サーバー
-各クラスタ(複数)
1.起動
IMPIサーバープロセス
TCPポート
・各クラスタが IMPI サーバーに接続
・互いのクラスタ情報を交換
mpirun -np 4 a.out
2.起動
3.起動
Cluster A
Cluster B
4.TCP接続
MPI_COMM_WORLD
National Research Grid Initiative

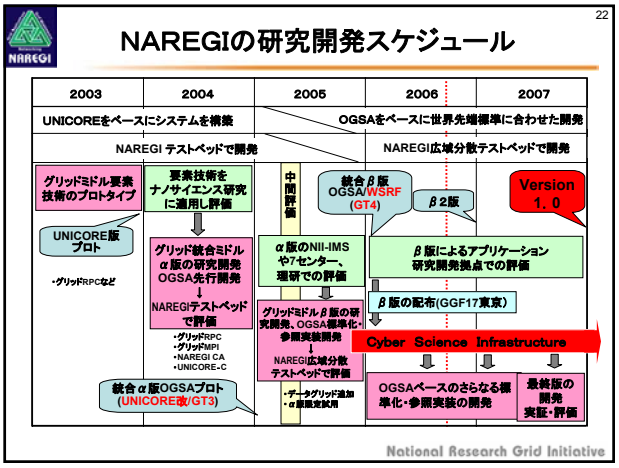

17
グリッドプログラミング:グリッドRPC
遠隔計算機上でライブラリ関数を呼び出すモデルに基づき、数10から数100CPUのグリッドアプリケーションの容易な開発と、高い実行効率を可能とする機能。
② 結果の通知
Network
ユーザ
① 関数呼び出し
Subroutine A
ポータル、ASPの バックエンドとして利用
大規模計算部分の 遠隔呼び出し
スーパーコンピュータ の遠隔利用
特殊なライブラリの 遠隔呼び出し
グリッド上の複数の高性能計算機を利用した大規模計算
National Research Grid Initiative

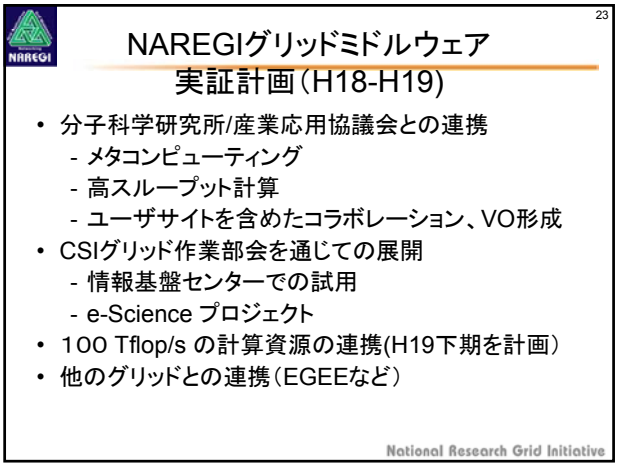

18
グリッドミドルウェアによる連成解析(処理フロー)
a: 利用者登録
b: アプリ配置
c: 編集
RISM
FMO
ワークフロー
CA
PSE
WFT
1: サブミット
スーパースケジューラ
2: 資源探索
3: 予約交渉
分散情報サービス
合意
9: アカウンティング
グリッドVM
ローカルスケジューラ
4: 資源予約
5: IMPI起動
6: サブジョブ起動
7: MPI初期化
IMPIサーバ
RISMジョブ
FMOジョブ
グリッドMPI
Site A
Site B
Site C
入力ファイル
出力ファイル
グリッド可視化
8: 可視化
National Research Grid Initiative

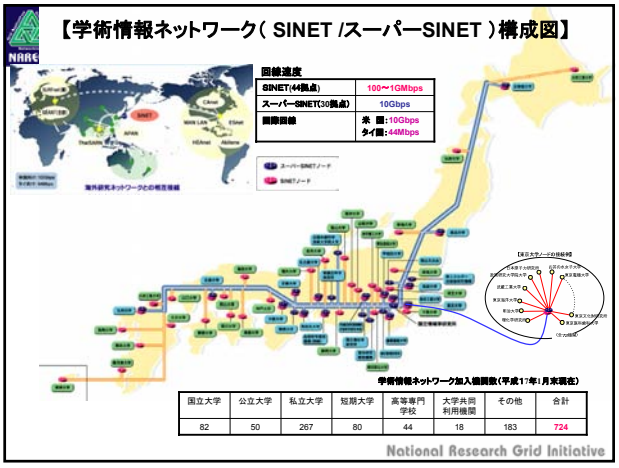

19
マルチフィジックス連成計算実行例(NAREGI拠点間)
分子科学研究所
RISM
溶媒分布解析
可視化
グリッド
ミドルウェア
溶液中の電子構造
グリッドMPI
情報学研究所
FMO
電子構造解析
可視化
グリッド
ミドルウェア
RISM Reference Interaction Site Model
無限系に対する溶媒分布解析プログラム
FMO Fragment Molecular Orbital method
ナノ高分子の電子構造解析プログラム
National Research Grid Initiative

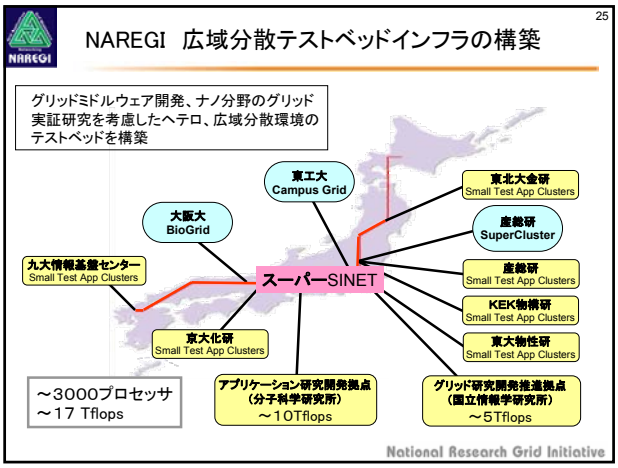
20
海外機関との相互運用性(標準化採用)
NAREGI成果のグリッド基盤としての将来性・持続性
■国際標準への準拠と貢献
ー UNICORE, Globus Toolkit ,OGSA準拠の開発方針
ー GGF(Global Grid Forum)への積極的参画
(Working Group への仕様提案、先行実装など)
■NAREGI CA(認証局)の運用
ー 第三者認証機関による身分を保証する証明書の発行
2005年9月1日 本格運用開始
2005年11月7日 APGrid PMAによるProduction Level CA
認定審査で承認(日本では産総研について2番目)
■海外のグリッドプロジェクトとの連携
ー 研究プロジェクトの国際化に伴う互換性(Interoperability)の重要性
ー 運用ツール類などの相互利活用(ポータビリティ)
ー 例: EGEE, TeraGrid, UNICORE/Unigrids等との技術交流(GIN)
■成果のオープンソース化と早期試用の促進
National Research Grid Initiative

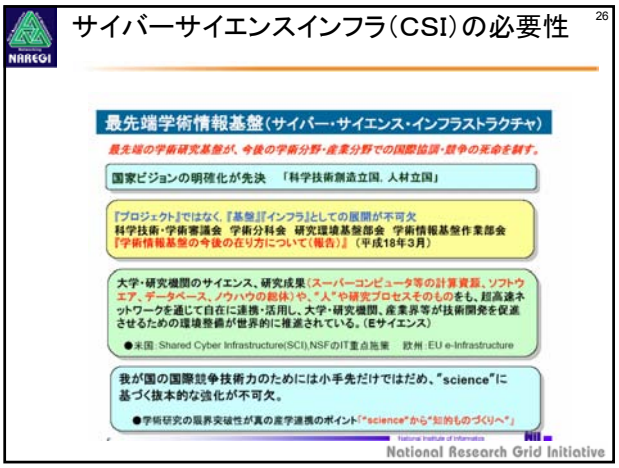

21
NAREGI グリッドシステム構成例
Portal & User Tools (Linux)
IS Client library
SS Client library
管理サイト
SS専用(Linux)
SS
IS Client library
IS専用(Linux)
IS (Node)
管理サイトは単独サイトである必要はなく、サイトの一部資源で構築しても良い
サイト1
IS (CD,RUS)
IS専用(Linux)
GridVM スケジューラ
IS (LRPS)
IS (プロバイダコンテナ)
GridVM スケジューラ
Linux 管理ノード
AIX 管理ノード
GridVM エンジン(ノード毎)
Linux 計算ノード群
AIX 計算ノード群
サイト2
IS (CD,RUS, LRPS)
IS専用(Linux)
Solaris 管理ノード
Solaris 計算ノード群
・IS:Information Service(分散情報サービス)
広域分散した複数の計算センターをまたがるグリッド資源情報、様々なVO(仮想組織)向けの計算環境に関する情報を収集・蓄積する。
・SS:Super Scheduler(スーパースケジューラ)
グリッド上に存在する多用な計算資源の中から、利用者のジョブ要求に応じた資源を自動検索し、グリッド環境上でのジョブスケジューリングを行う。
・GridVM:Grid Virtual Machine(グリッドVM)
グリッド環境に適した形で各サイトの計算機資源を利用するための仮想マシンレイヤー。
National Research Grid Initiative

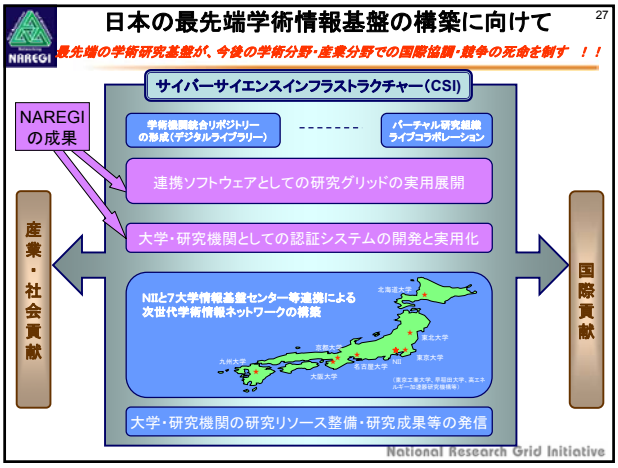

22
NAREGIの研究開発スケジュール
2003
2004
2005
2006
2007
UNICOREをベースにシステムを構築
OGSAをベースに世界先端標準に合わせた開発
NAREGI テストベッドで開発
NAREGI広域分散テストベッドで開発
グリッドミドル要素技術のプロトタイプ
UNICORE版プロト
・グリッドRPCなど
要素技術をナノサイエンス研究に適用し評価
グリッド統合ミドルα版の研究開発 OGSA先行開発
NAREGIテストベッドで評価
・グリッドRPC
・グリッドMPI
・NAREGI CA
・UNICORE-C
統合α版OGSAプロト (UNICORE改/GT3)
中間評価
α版のNII-IMSや7センター、理研での評価
グリッドミドルβ版の研究開発、OGSA標準化・参照実装開発
NAREGI広域分散テストベッドで評価
・データグリッド追加
・α版限定試用
統合β版 OGSA/WSRF (GT4)
β2版
β版によるアプリケーション研究開発拠点での評価
β版の配布(GGF17東京)
Version 1.0
Cyber Science Infrastructure
OGSAベースのさらなる標準化・参照実装の開発
最終版の開発 実証・評価
National Research Grid Initiative

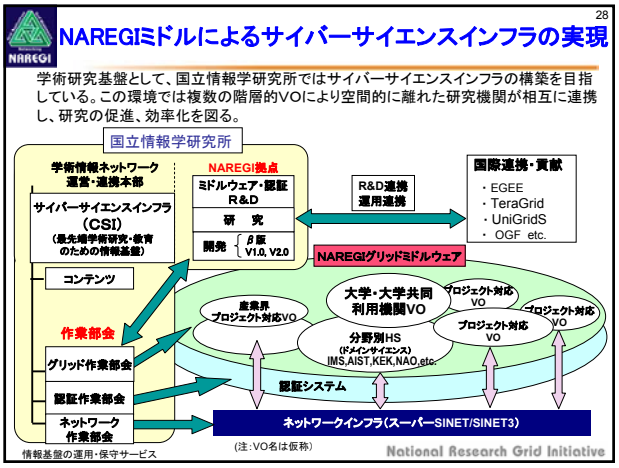

23
NAREGIグリッドミドルウェア
実証計画(H18-H19)
• 分子科学研究所/産業応用協議会との連携
- メタコンピューティング
- 高スループット計算
- ユーザサイトを含めたコラボレーション、VO形成
• CSIグリッド作業部会を通じての展開
- 情報基盤センターでの試用
- e-Science プロジェクト
• 100 Tflop/s の計算資源の連携(H19下期を計画)
• 他のグリッドとの連携(EGEEなど)
National Research Grid Initiative

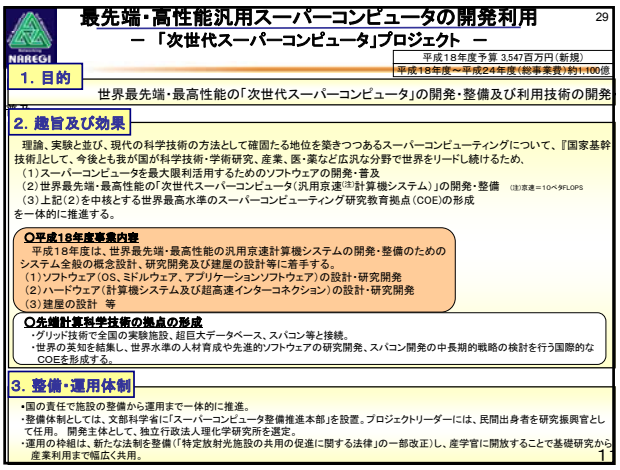

【学術情報ネットワーク(SINET /スーパーSINET)構成図】
回線速度
SINET(44拠点) 100～1GMbps
スーパーSINET(30拠点) 10Gbps
国際回線 米 国:10Gbps タイ国:44Mbps
国立大学 公立大学 私立大学 短期大学 高等専門学校 大学共同利用機関 その他 合計
82 50 267 80 44 18 183 724
National Research Grid Initiative


25
NAREGI 広域分散テストベッドインフラの構築
グリッドミドルウェア開発、ナノ分野のグリッド実証研究を考慮したヘテロ、広域分散環境のテストベッドを構築
東工大 Campus Grid
大阪大 BioGrid
九大情報基盤センター Small Test App Clusters
スーパーSINET
東北大金研 Small Test App Clusters
産総研 SuperCluster
産総研 Small Test App Clusters
KEK物構研 Small Test App Clusters
東大物性研 Small Test App Clusters
京大化研 Small Test App Clusters
～3000プロセッサ
～17 Tflops
アプリケーション研究開発拠点（分子科学研究所）～10Tflops
グリッド研究開発推進拠点（国立情報学研究所）～5Tflops
National Research Grid Initiative


26
サイバーサイエンスインフラ（CSI）の必要性
最先端学術情報基盤（サイバー・サイエンス・インフラストラクチャ）
最先端の学術研究基盤が、今後の学術分野・産業分野での国際協調・競争の死命を制す。
国家ビジョンの明確化が先決 「科学技術創造立国，人材立国」
『プロジェクト』ではなく，『基盤』『インフラ』としての展開が不可欠
科学技術・学術審議会 学術分科会 研究環境基盤部会 学術情報基盤作業部会
『学術情報基盤の今後の在り方について（報告）』（平成18年3月）
大学・研究機関のサイエンス、研究成果（スーパーコンピュータ等の計算資源、ソフトウエア、データベース、ノウハウの総体）や、"人"や研究プロセスそのものをも、超高速ネットワークを通じて自在に連携・活用し、大学・研究機関、産業界等が技術開発を促進させるための環境整備が世界的に推進されている。（Eサイエンス）
●米国：Shared Cyber Infrastructure(SCI),NSFのIT重点施策 欧州：EU e-Infrastructure
我が国の国際競争技術力のためには小手先だけではだめ、"science"に基づく抜本的な強化が不可欠。
●学術研究の最界突破性が真の産学連携のポイント「"science"から"知的ものづくり"へ」
National Research Grid Initiative


27
日本の最先端学術情報基盤の構築に向けて
最先端の学術研究基盤が、今後の学術分野・産業分野での国際協調・競争の死命を制す ！！
サイバーサイエンスインフラストラクチャー（CSI）
NAREGIの成果
学術機関統合リポジトリーの形成（デジタルライブラリー）
バーチャル研究組織ライブコラボレーション
連携ソフトウェアとしての研究グリッドの実用展開
大学・研究機関としての認証システムの開発と実用化
NIIと7大学情報基盤センター等連携による次世代学術情報ネットワークの構築
大学・研究機関の研究リソース整備・研究成果等の発信
産業・社会貢献
国際貢献
National Research Grid Initiative


28
NAREGIミドルによるサイバーサイエンスインフラの実現
学術研究基盤として、国立情報学研究所ではサイバーサイエンスインフラの構築を目指している。この環境では複数の階層的VOにより空間的に離れた研究機関が相互に連携し、研究の促進、効率化を図る。
国立情報学研究所
学術情報ネットワーク運営・連携本部
サイバーサイエンスインフラ（CSI）（最先端学術研究・教育のための情報基盤）
コンテンツ
NAREGI拠点
ミドルウェア・認証 R&D
研究
開発 β版 V1.0, V2.0
R&D連携 運用連携
国際連携・貢献
・EGEE
・TeraGrid
・UniGridS
・OGF etc.
NAREGIグリッドミドルウェア
産業界プロジェクト対応VO
大学・大学共同利用機関VO
プロジェクト対応VO
分野別HS（ドメインサイエンス）IMS,AIST,KEK,NAO,etc.
作業部会
グリッド作業部会
認証作業部会
ネットワーク作業部会
認証システム
ネットワークインフラ（スーパーSINET/SINET3）
情報基盤の運用・保守サービス
（注：VO名は仮称）
National Research Grid Initiative


29
最先端・高性能汎用スーパーコンピュータの開発利用
―「次世代スーパーコンピュータ」プロジェクト ―
平成18年度予算 3,547百万円（新規）
平成18年度～平成24年度（総事業費7約1,100億
1. 目的
世界最先端・最高性能の「次世代スーパーコンピュータ」の開発・整備及び利用技術の開発
2. 趣旨及び効果
理論、実験と並び、現代の科学技術の方法として確固たる地位を築きつつあるスーパーコンピューティングについて、『国家基幹技術』として、今後とも我が国が科学技術・学術研究、産業、医・薬など広汎な分野で世界をリードし続けるため、
（1）スーパーコンピュータを最大限利活用するためのソフトウェアの開発・普及
（2）世界最先端・最高性能の「次世代スーパーコンピュータ（汎用京速注1計算機システム）」の開発・整備 （注1京速＝10ペタFLOPS
（3）上記（2）を中核とする世界最高水準のスーパーコンピューティング研究教育拠点（COE）の形成
を一体的に推進する。
○平成18年度事業内容
平成18年度は、世界最先端・最高性能の汎用京速計算機システムの開発・整備のためのシステム全般の概念設計、研究開発及び建屋の設計等に着手する。
（1）ソフトウェア（OS、ミドルウェア、アプリケーションソフトウェア）の設計・研究開発
（2）ハードウェア（計算機システム及び超高速インターコネクション）の設計・研究開発
（3）建屋の設計 等
○先端計算科学技術の拠点の形成
・グリッド技術で全国の実験施設、超巨大データベース、スパコン等と接続。
・世界の英知を結集し、世界水準の人材育成や先進的ソフトウェアの研究開発、スパコン開発の中長期的戦略の検討を行う国際的なCOEを形成する。
3. 整備・運用体制
・国の責任で施設の整備から運用まで一体的に推進。
・整備体制としては、文部科学省に「スーパーコンピュータ整備推進本部」を設置。プロジェクトリーダーには、民間出身者を研究振興官として任用。 開発主体として、独立行政法人理化学研究所を選定。
・運用の枠組は、新たな法制を整備（「特定放射光施設の共用の促進に関する法律」の一部改正）し、産学官に開放することで基礎研究から産業利用まで幅広く共用。
11

30
広汎な分野での利活用 － 次世代スパコンが拓く世界 －
ものづくり
ナノテクノロジー
物質設計
酵素・触媒反
カーボンナノチューブ
防災
雪の解析
津波被害予測
自動車開発
ライフサイエンス
原子力
地球環境
エルニーニョ現象の影響予測
原子炉丸ごと解析
人間丸ごと解析
創薬解析
レーザー反応解析
天文・宇宙物理
銀河形成解明 惑星形成解明
ロケットエンジン設計
航空機開発
オーロラ発生解明
次世代スーパーコンピュータ
National Research Grid Initiative
12

次世代スーパコンピュータ開発スケジュール(案)

年度	平成18年度	平成19年度	平成20年度	平成21年度	平成22年度	平成23年度	平成24年度
開発項目 ＼ 評価等	★マネージメント体制 開発ターゲット、京速計算機システムの構成等 ★仕様・実装内容の判断 (概念設計内容、開発体制、立地・運用方針、採用する半導体プロセスの決定等) ★詳細なハードウェア要件、LSIの論理構成概略仕様等			研究開発状況評価★ (システム性能・機能等)		COE形成、運用評価★ (利用状況、研究成果、人材育成状況等)	
ソフトウェア：システムソフトウェア	異機種統合・グリッドミドルウェア設計・製作			評価			
ソフトウェア：グランドチャレンジアプリケーション	次世代ナノ統合シミュレーション設計・製作			評価			
	次世代生命体統合シミュレーション設計・製作					評価	
ハードウェア	設計		実装技術設計・評価		製作	システム強化	
ファイルシステム、付帯設備整備等			設計	製作		システム強化	
立地調査、建屋建設	検討	設計	建設				

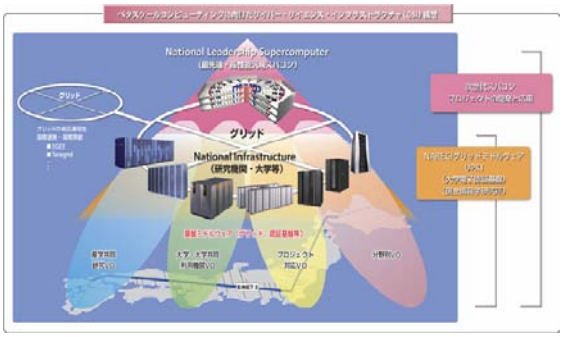
文部科学省研究振興局資料より

National Research Grid Initiative

13

次世代スパコンシステム向けミドルウェアとしてのグリッド
(最先端・高性能汎用スーパーコンピュータの開発利用プロジェクト)

32



National Research Grid Initiative

まとめ

33

1) グリッドは、21世紀の研究開発基盤であり、産業基盤です。

2) NAREGI では分子研との実証研究によって、グリッドがナノサイエンス、ナノテクノロジの分野における計算科学環境に役立つことを実証致します。

3) 開発したグリッド技術を広く産業界へも広め、計算科学の手法に立脚した産業界の国際競争力向上に寄与します。

4) さらにサイバーサイエンスインフラへグリッドの要素技術を提供することにより、計算資源利用者の裾野の拡大貢献します。

5) 次世代スーパーコンピュータシステムによるペタスケール時代に不可欠なグリッドによる計算研究環境の研究・開発を継続して行います。

6) 国際的なグリッド技術標準化に貢献します。

"Software is the Infrastructure"

- Paul Messina, NAREGI Symposium 2004 -