

Japan Lattice Data Grid 運用開始について

○概要

筑波大学計算科学研究センター、高エネルギー加速器研究機構計算科学センター、京都大学基礎物理学研究所、大阪大学核物理研究センター、広島大学理学部物理学科、金沢大学自然科学研究科は、2008年5月23日、Japan Lattice Data Grid (JLDG) の運用を開始しました。JLDG は、国内の格子 QCD (Lattice QCD) をはじめとする計算素粒子物理の研究グループが、スーパーコンピュータを駆使して得られる貴重な大規模データを大域的かつ効率的に共有することを可能にする、広域分散型の次世代ファイル共有システムです。各拠点のスーパーコンピュータで生成した計算結果を JLDG に置くことによって、それを任意の拠点からあたかも拠点内にあるデータのように取り出し、その拠点のスーパーコンピュータで解析するといった作業を、効率よく行うことができます。また、国内の研究者にとって有用と思われるデータを一般に公開することもできます。JLDG は計算素粒子物理分野だけではなく、様々な分野の e-Science を促進する新たなインフラストラクチャとなることが期待されています。

○詳細

計算素粒子物理の研究では、基礎データの生成に膨大な計算資源を必要としますが、生成された基礎データを元に素粒子の様々な性質を計算する事ができます。このため、計算素粒子物理では、異なる拠点に所属する複数の研究者が一つの研究プロジェクトを組織し、複数拠点のスーパーコンピュータを用いて基礎データを生成するとともに、様々な観点からのデータ解析を分担しておこなうのが通例です。JLDG に接続する6つの機関は国内の計算素粒子物理学の主要な研究拠点であり、計算素粒子物理のなかでも最も基礎的な、格子 QCD の研究を推進しています。各機関は、国立情報学研究所が提供する SINET 上の Hepnet-J/sc というプライベートネットワークに接続されており、2002年から、データを複数拠点間で複製することにより、データの共有を行ってきました。各拠点に置かれたファイルサーバは、一般に、比較的小さな容量のパーティションに区切られており、計算素粒子物理で扱う一纏まりのデータはパーティションの容量より遥かに大きいので、どのデータが何処に置かれ、その複製が何処にあるかを管理するのは容易ではありませんでした。

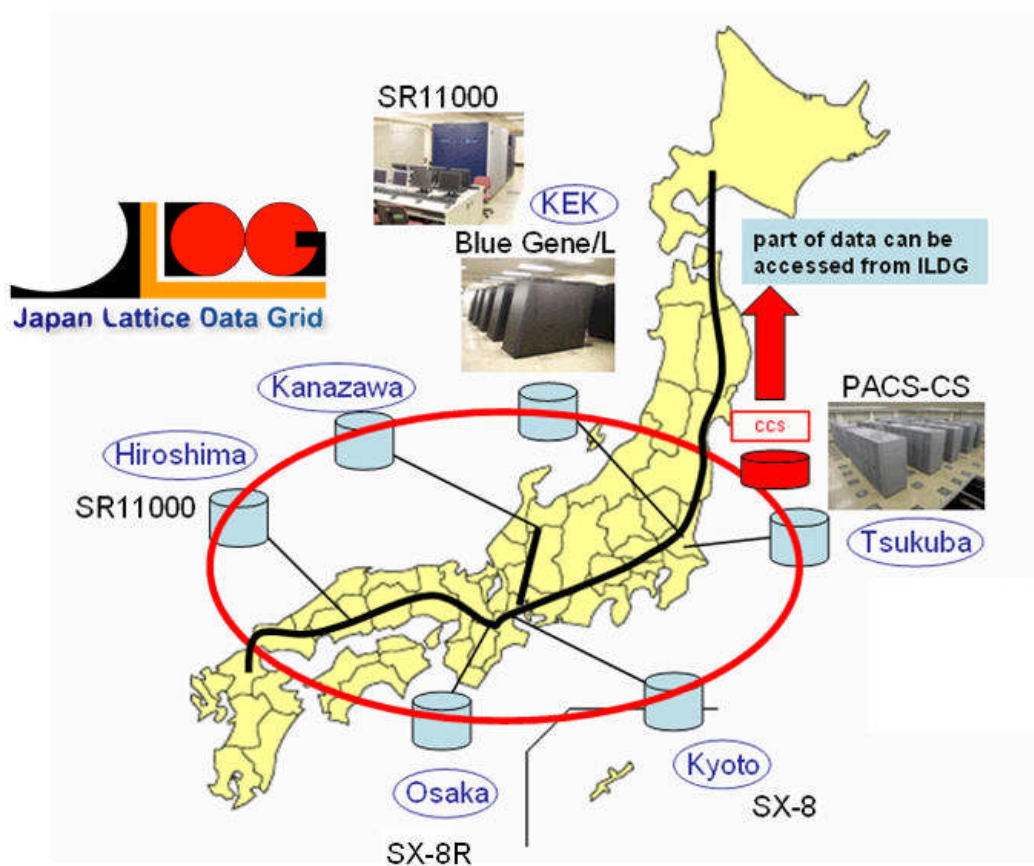
この困難を解決するため、6つの拠点の計算素粒子物理の研究者と筑波大学の計算機工学の研究者が緊密に協力し、2005年11月から JLDG の開発を行ってきました。JLDG では、

Hepnet-J/sc 上に分散して置かれたファイルサーバを、産業技術総合研究所・筑波大学で開発された Gfarm システムにより束ね、仮想的な単一のフラットなファイル共有システムを実現しました。ユーザーは、データが実際にどの拠点のどのサーバ上にあるかを意識することなく、ファイルの読み書きを行うことができます。JLDG は、データ量の増大に応じてファイルサーバを追加することによって、事実上、容量の制限のないファイル共有システムを提供します。

JLDG は、国立情報学研究所の NAREGI プロジェクトの開発した認証局ソフトウェア NAREGI-CA と、EDG (European Data Grid) プロジェクトの開発した仮想組織管理ソフトウェア VOMS (Virtual Organization Management System) を用いて、研究者の実際の所属によらない研究グループを管理しています。ユーザーはユーザー証明書を取得し、何れかの研究グループに所属し、その権限で JLDG にアクセスします。

JLDG は 筑波大学計算科学研究センターにて International Lattice Data Grid (ILDG) と接続しています。ILDG は国際規模での格子 QCD の基礎データ「配位」を共有する為に構築されたデータグリッドです。ILDG は 2002 年にその構想が提案され、各地域グリッドの構築を経て、2007 年 7 月に相互運用を開始しました。現在、イギリス、アメリカ、ドイツ・イタリア・フランス、オーストラリア、日本、の 5 つの地域グリッドが接続され、世界の研究者に配位を提供しています。国内の研究グループが ILDG に公開する配位は、全て JLDG 上に置かれます。この意味で、JLDG は ILDG の一つの地域グリッドの役割を担っています。

JLDG の開発の一部は、国立情報学研究所 CSI 委託事業「グリッド・認証技術による大規模データ計算資源の連携基盤の構築」および「e-Science 研究分野の振興を支援する CSI 委託事業」の研究課題「計算素粒子物理学の高度データ共有基盤 JLDG の構築」によるものです。



JLDG の概念図

○関連情報

JLDG ホームページ <http://www.jldg.org/>

○問い合わせ先

- 佐藤 三久 筑波大学計算科学研究センター長
システム情報工学研究科
email: msato@cs.tsukuba.ac.jp
- 建部 修見 筑波大学システム情報工学研究科 (計算科学研究センター)
email: tatebe@cs.tsukuba.ac.jp
- 吉江 友照 筑波大学数理物質科学研究科 (計算科学研究センター)
email: yoshie@ccs.tsukuba.ac.jp
- 松古 栄夫 高エネルギー加速器研究機構計算科学センター
email: hideo.matsufuru@kek.jp

大野木 哲也 京都大学基礎物理学研究所
email: onogi@yukawa.kyoto-u.ac.jp

外川 浩章 大阪大学核物理研究センター
email: togawa@rcnp.osaka-u.ac.jp

石川 健一 広島大学理学部物理科学科
email: ishikawa@theo.phys.sci.hiroshima-u.ac.jp

出淵 卓 金沢大学自然科学研究科
email: izubuchi@hep.s.kanazawa-u.ac.jp

○用語解説

* 格子 QCD (Lattice QCD)

「QCD」は、陽子やパイ中間子などハドロンと呼ばれる素粒子の構成要素であるクォーク・グルーオンの強い相互作用の基礎理論。QCD を離散化した時空上で定義した格子 QCD の数値シミュレーションによれば、第一原理である QCD から素粒子の様々な性質を解明できる。

* e-Science

ネットワークを介して多数の科学者の協調やデータの共有によって進められる科学。

* SINET (Science Information Network, 学術情報ネットワーク)

国立情報学研究所が運営する、国内の学術研究用ネットワーク。国内の主要な大学・研究所が接続している。

* Gfarm

産業技術総合研究所・筑波大学を主体に研究開発が進められるオープンソースの次世代型広域分散ファイルシステム。ファイル複製による遠隔でのファイル共有の効率化、高い信頼性に特徴がある。

<http://sf.net/projects/gfarm/>

* NAREGI-CA

国立情報学研究所で研究開発が進められる NAREGI グリッドミドルウェアにおける認証局を構築するためのソフトウェア。シングルサインオンでそれぞれの拠点にあるコンピュータ等の設備やデータを利用するための証明書の発行、管理を行う。

* 仮想組織

組織を超えたプロジェクトやグループ等、実際の組織とは関係なく自由に構成する仮想的な組織。

* VOMS(Virtual Organization Management System)

EDG (European Data Grid)で開発された仮想組織のメンバを管理するソフトウェア。メンバに対し、仮想組織で利用可能なコンピュータ等の設備が自動的に利用可能となる仕組みを備える。

* ILDG(International Lattice Data Grid)

格子 QCD の基礎データ「配位」の国際規模での共有をサポートするためのデータグリッド。地域グリッドを共通のインターフェースで束ねた、いわゆるグリッドのグリッドとして運用されている。

<http://www.lqcd.org/ildg>

* CSI(Cyber Science Infrastructure)

大学や研究機関等が有しているコンピュータ等の設備，基盤的ソフトウェア，コンテンツ及びデータベース，人材，研究グループそのものを超高速ネットワーク上で共有するための最先端学術情報基盤。