

## 5.2 グリッドシステム

飯田好美、岩井剛、河井裕、佐々木節、鈴木聡、渡瀬芳行、八代茂夫

### 5.2.1 概要

計算科学センターではグリッドコンピューティングの仕組みを利用し、本機構と共同研究を行う研究機関との間で、データ及び計算資源を共有するためにグリッドシステムを構築し、運用を行っている。

混乱を避けるため、先に本節で用いられるグリッドシステムという用語に対する解釈を与える。図 1 は実際にインストールされた計算科学センターに於いて運用中の要素をサービスに着目し、まとめたものである。図 1 に示されるように多くの場合、グリッドに於けるサービスとは一つ以上の小サービスがお互いに作用することにより形成され、複数のサービスにより、グリッドに於ける管理単位に相当する「サイト」が形成される。これはグリッドの「複数機関に跨った計算基盤を利用する」という設計思想に基づく。このサイトを形成・構築するためのソフトウェアを特に区別してミドルウェアと呼ぶが、計算科学センターでは複数のミドルウェア (gLite、NAREGI、iRODS、SRB 等) を使用し、複数サイトを運用している。本節では計算科学センターで運用される、これら複数のサイトを全て含め「グリッドシステム」として扱うこととする。

次節以降でグリッドシステムを構成しているサイトの利用状況と、これらにより提供される計算基盤上での昨年度活動内容について報告する。

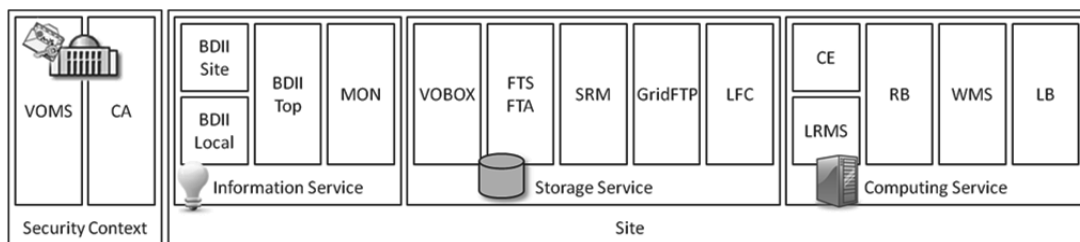


図 1 計算科学センターにて実運用中の Grid 計算基盤の構成。実際には複数のミドルウェア (NAREGI, SRB, iRODS, gLite 等) により複数のサイトが構成されているが、混乱を避けるため、図中には gLite ミドルウェアにより構成された LCG (LHC Computing Grid) 計算基盤のみを示した。

## 5.2.2 KEK Grid CA (Certificate Authority)

KEK Grid CA は世界に 80 程度ある IGF<sup>1</sup>の策定する規格に準拠・承認された認証局の一つである。2006 年 2 月に正式運用を開始して以来、主に国内のグリッド利用者、及びサービス提供ホストに対して電子証明書を発行している。

図 2 に KEK Grid CA が 2010 年度に発行した個人証明書とホスト証明書の発行部数を示す。2010 年度は 267 部 (JFY2009 +35) の個人証明書と 358 部 (JFY2009 +50) のホスト証明書が発行され、2006 年の認証局運用開始以来の累計発行部数はそれぞれ 907 部、1553 部に達した。

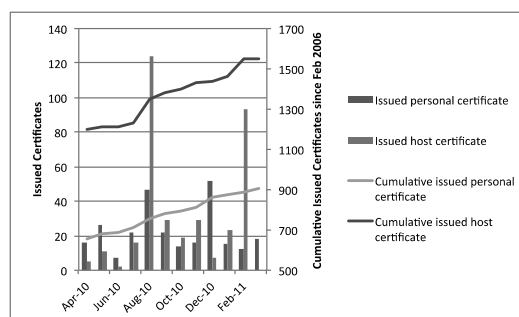


図 2 2010 年度の電子証明書発行部数

## 5.2.3 LCG 計算基盤

JP-KEK-CRC-01 及び JP-KEK-CRC-02 (以降 KEK-1 及び KEK-2 と表記) は EGEE<sup>2</sup>が開発を推進した gLite ミドルウェアにより構築された計算資源を提供するサイトである。約 50 カ国・200 機関から 300 程度のサイトが参加することによって、LCG と呼ばれる全世界規模の計算基盤が展開され、2010 年度中、当センターではそのうちの 2 サイトである KEK-1 及び KEK-2 が運用中であった。

KEK-1 は研究開発用途に、KEK-2 は KEK-1 で培われた経験に基づき、より安定したサービスを提供するために分離して運用がなされている。以降は特に注釈のない限り、KEK-2 のみに関して報告する。

### サービスの品質

図 3 に 2010 年度中に発生したダウンタイム、およびサービス稼働率を示す。2010 年度は 200 時間程度 (JFY2009 -100 時間) のダウンタイムが発生し、システムの年間稼働率は 88%であった。これは 8 月の計画停電によるシステム全停止、1 月の計算サーバー移行作業、および 3 月に発生した東日本大震災によるシステム全停止に依るものである。LCG に於いて定められた月間最小サービス稼働率は 70%に設定されており<sup>3</sup>、これは図 3 に見ることが出来るように 2010 年 8 月と 2011 年 3 月以外の月次は年度を通じ

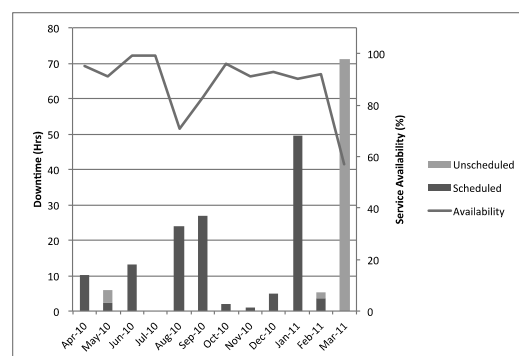


図 3 2010 年度に発生したダウンタイムとサービス稼働率

<sup>1</sup> The International Grid Trust Federation: <http://www.gridpma.org/>

<sup>2</sup> Enabling Grids for E-science: <http://public.eu-egee.org/>

<sup>3</sup> 算出式を含む完全な資料は <https://edms.cern.ch/document/860386/> から取得可能。

で満足した。なお、年間稼働率の平均は通常、95%程度である。

利用者とシステム管理者が幅広く分散する LCG インフラではコミュニケーションチャンネルのひとつとして KIT (Karlsruhe Institute Technology) が主導する GGUS (The Global Grid User Support) と呼ばれるチケットシステムが利用される。年間 1 万件以上のチケットが発行されるが、2010 年度中に当センターに関連したチケットは 45 件のうち 16 件は Belle 仮想組織 (VO) に特化した問題であった。ほとんどのチケットは 24 時間以内に初期対応が行われ、1 週間以内でチケットはクローズされた。

## 利用状況

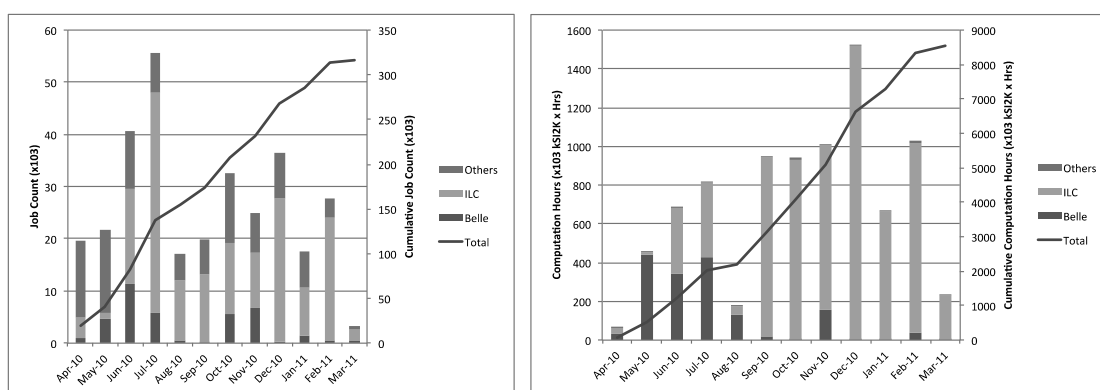


図 4 2010 年度のジョブ処理数 (左) と CPU 消費時間 (右)

図 4 に 2010 年度中に投入されたジョブ数と CPU 消費時間<sup>4</sup>を示した。投入ジョブの総数は 316,756 (JFY2009 13 倍) でそのうち 12% は Belle VO、56% は ILC VO であった。CPU 消費時間 (JFY2009 14 倍) は Belle VO が全体の 19% で ILC VO が 81% であった。

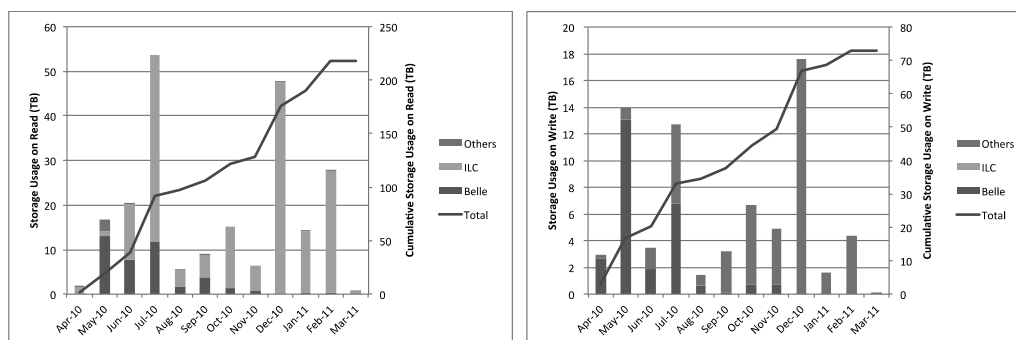


図 5 2010 年度ストレージ利用量

ストレージサービスは共通情報システム HPSS をバックエンドに VFS マウントした領域を GridFTP 及び SRM プロトコルでサービスを提供している。記憶領域の増設については、グル

<sup>4</sup> SPEC (Standard Performance Evaluation Corporation) が策定したシステムの性能評価を行うベンチマークスコアにより規格化した CPU 消費時間。高エネルギー実験分野で使用されるプログラムは SPEC Int にスケールすることが多いため、ここでは 1000 Spec Int 2000 で CPU 消費時間を規格化している。

テープ毎にテープカートリッジを購入してもらう方針で運用している。2009 年後期までは現行の共通情報システム導入時に割り当てた初期状態の領域を各 V0 で共有していたが、この領域が Belle と ILC の要求容量には見合わないため、テープを購入してもらい独立した領域を割り当てている。年間を通じて 218TB (JFY2009 109 倍) の読み出しと 73TB (JFY2009 18 倍) の書き込みが行われた。(図 5)

#### 5.2.4 ユーザー対応・他

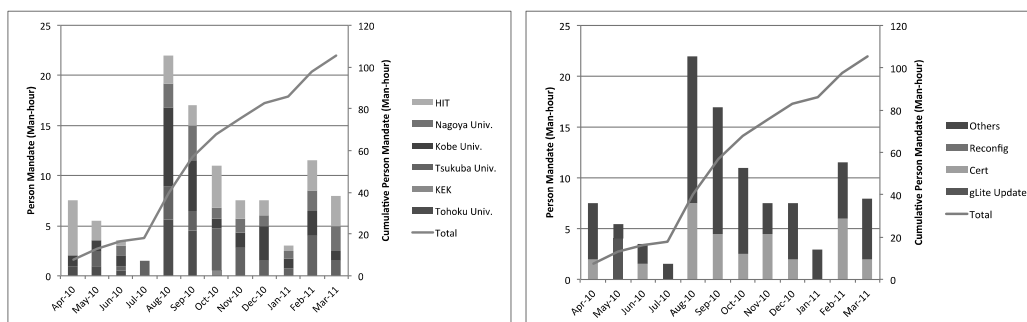


図 6 2010 年度に実施された加速器科学仮想組織に於ける運用支援作業時間

グリッドにより構成される計算基盤の運転維持には多くの人的資源が必要とされるため、大学研究室の規模で継続的に運用をまかなうことは困難である。当センターでは 2007 年 3 月に国内の共同研究機関（東北大・筑波大・名古屋大・神戸大・広工大）にシステム一式を構築し、拠点間を遠隔相互診断するための計算基盤として KEK を含む 6 拠点間で科学的目標に依存しない加速器科学仮想組織を整備した。

各拠点のソフトウェア・アップデート、QA 対応といった運用支援も継続的に行われている。2010 年度は 106 人・時間 (JFY2009 -43%) の運用支援が各拠点に対して行われた。(図 6)

機構内の研究グループ (ILC と BELLE) への支援も継続的に行われている。定期的にミーティングを開催し、性能向上・環境改善に役立てられた。

## 5.2.5 震災の影響

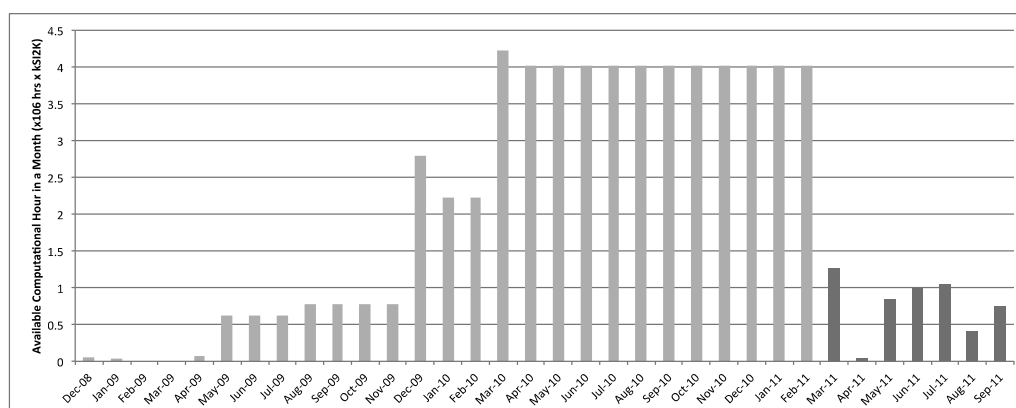


図 7 2011年3月11日に発生した東日本大震災以降の計算資源量推移

2008年12月から2011年9月までのグリッドシステムにて利用可能な計算資源量を図7に示した。2011年3月11日に発生した東日本大震災の影響により、電気使用量の制約等により、地震発生以後、約2週間にわたってシステムが完全停止状態となった。グリッドの性質上、1サイトがダウンしていても、世界に分布したグリッド全体に深刻な影響を及ぼすことはないが、当センターでは Belle V0 向けのクリティカル・サービスをいくつかホストしているため、復旧が急がれた。

スタッフの努力により、地震発生から約2週間経過した3月25日、V0の構成に必須となるサービス (VOMS, VOMRS, LFC, site BDII) の再稼働を開始した。以降は逐次、電気容量の許す範囲で資源割り当てを行い、4月末には地震発生直前の約25%程度の規模でシステムが安定的に動作することが確認された。

国内外の関係各機関の方々から温かいメッセージを多くいただきました。ここに御礼申し上げます。

## 論文

- [1] Y. Kawai, A. Hasan, G. Iwai, T. Sasaki, and Y. Watase. “An alternative method for reliably managing large files”, To appear in: 情報処理学会第73回全国大会論文集. (査読無)
- [2] Y. Kawai, G. Iwai, T. Sasaki, and Y. Watase. “SAGAによるグリッドアプリケーションとその性能評価”, 第9回情報科学技術フォーラム講演論文集 1 (1) (2010) 395-399. (査読無)

## 国際会議プロシーディングス

- [3] G. Iwai, Y. Kawai, T. Sasaki, and Y. Watase. “A development of lightweight Grid interface”, To appear in: Journal of Physics: Conference Series (JPCS),

- CHEP 2010, Taipei, Taiwan, Oct 2010. (査読有)
- [4] Y. Kawai, G. Iwai, T. Sasaki, and Y. Watase. “SAGA-based File Access Application over Multi-Filesystem Middleware”, Proceedings of the 19th ACM International Symposium on High Performance Distributed Computing (HPDC 2010) (2010) 622-626, CLADE2010 (HPDC2010), Chicago, US, Jun 2010. (査読有)
- [5] G. Iwai, Y. Kawai, T. Sasaki, and Y. Watase. “SAGA-based user environment for distributed computing resources: A universal Grid solution over multi-middleware infrastructures”, Procedia CS 1 (1) (2010) 1545-1551, ICCS2010, Amsterdam, The Netherlands, Jun 2010. (査読有)
- [6] “KEK GRID for ILC Experiments”, T. Abe, G. Iwai et al.; Proc. of the International Linear Collider Workshop (LCWS2010), 2010. (査読有)
- [7] “Computing at Belle II”, T. Hara, G. Iwai et al.; Proc. Of Advanced Computing and Analysis Techniques in Physics Research of (ACAT2010), pp. 022, 2010. (査読有)

## 会議発表

- [8] Yoshiyuki Watase. “Universal Grid User Interface - Project: RENKEI -”, OGF30, Brussels, Belgium, Oct 2010. (200名)
- [9] Go Iwai. “A development of lightweight Grid interface”, CHEP 2010, Taipei, Taiwan, Oct 2010. (800名)
- [10] Go Iwai. “SAGA-based user environment for distributed computing resources: A universal Grid solution over multi-middleware infrastructures”, ICCS2010, Amsterdam, The Netherlands, Jun 2010. (200名)
- [11] Go Iwai. “Grid Computing”, ILC 測定器研究会, KEK, Japan, Mar 2011. (30名)
- [12] Yoshimi Iida. “iRODS state at KEK”, FJKPPL workshop on GRID computing, KEK, Japan, Dec 2010. (20名)

## その他

- [13] “1<sup>st</sup> ILC gLite Training Course”, Tsukuba, Japan, Apr 2010. (10名)
- [14] “Universal Grid User Interface”, 展示・デモ, SC10, New Orleans, US, Nov 2010.
- [15] “The International Large Detector: Letter of Intent”, T. Abe, G. Iwai et al.; 2010.