

5.3 Bファクトリー計算機システム

鈴木 次郎、真鍋 篤

5.3.1 概要

計算科学センターでは、Bファクトリー(Belle)実験のデータ解析/アーカイブ環境としてBファクトリー計算機システムの運用を行っている。このシステムは、2006年3月に稼働を開始し、2009年2月にシステム増強を行い、2012年2月22日に運用を終了した。Belleグループの計算環境は、計算科学センター運用の中央計算機システムに統合され、2012年度初頭から運用が開始された。1985年に導入されたトリスタン計算機システムに起源をもつトリスタン/Bファクトリー計算機システムは2011年度末で27年間にわたる運用を終え、中央計算機への引き継ぎを行った。

Bファクトリー計算機は、導入当初は1208ノード(2416CPU)の計算サーバ、1PBの磁気ディスク、3.5PBのテープライブラリで構成されるHSMと呼ばれるストレージシステムで構成されていた(HSM = hierarchical storage management)。2009年のシステム増強で560ノード(4480CPU)の計算サーバ、1.5PBの磁気ディスクに増強された。図1にシステムの概要を示す。図2に計算サーバ、テープライブラリを示す。

ユーザーは、WG(ワークグループサーバ)にログインし、プログラム開発やデータアーカイブなどの作業を行った上で、LSFを使ってバッチジョブの投入を行い、計算サーバ(SC)でデータ解析などのジョブの実行を行う。

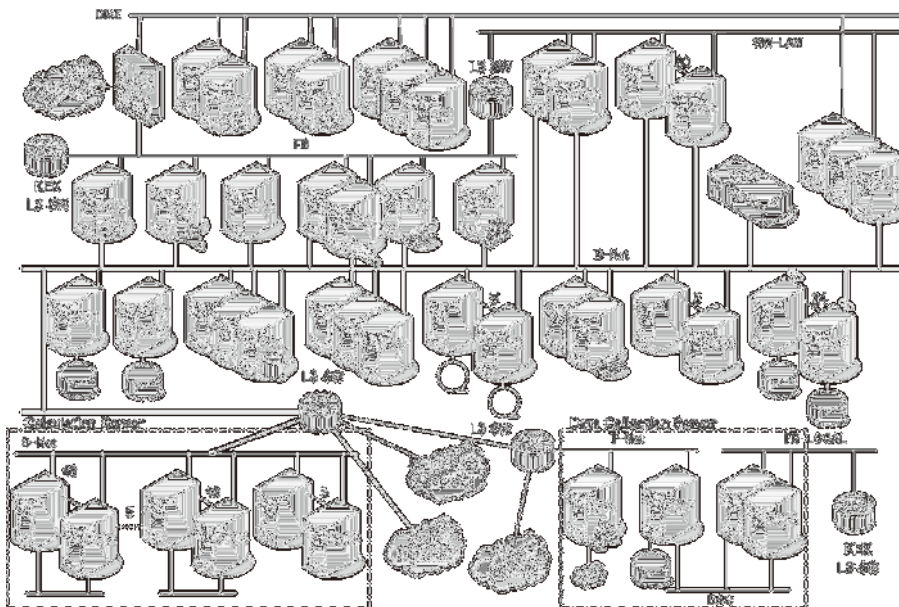


図1 Bファクトリー計算機システムの構成図

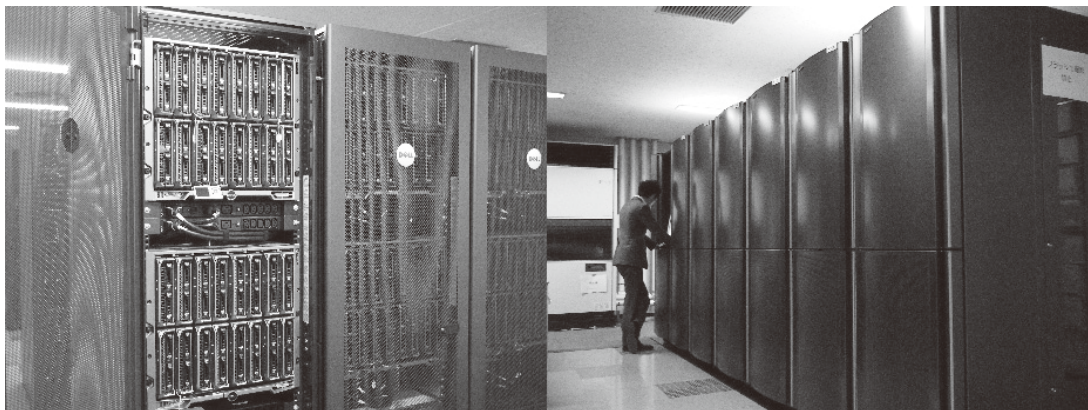


図 2 左図: 計算サーバ, 右図: HSM システム(テープライブラリ)。
2012 年 2 月 22 日(運用最終日)に撮影。

5.3.2 運用状況

2011 年度の運用は、システム自体に損傷は受けなかったものの 2011 年 3 月の震災に非常に大きな影響を受けた。年度後半は、ストレージシステムにアーカイブされているデータの共通計算機への転送(中央計算機への引き継ぎ)を運用中に行うこととなった。

2011 年度は、地震による影響を除けば、重大な問題点、セキュリティインシデントは発生しなかった。日常的なシステムメンテナンスは、5 名の常駐保守員が対応し、月に数台の故障した計算サーバ(ブレード)、HSM のディスクとテープ、ネットワークスイッチ、各種電源ユニットなどの交換であるが、多くの場合は冗長構成のため無停止で対処を行っている。システムの停止が伴う保守作業は、年末年始の数日と夏期停電時の前後にまとめて行いユーザーへの影響を最小限にした。一方で緊急性の高いセキュリティパッチ適用などは、ユーザーへ連絡の後に速やかに作業を行い、問題発生を未然に防ぐようにしている。

2010 年 6 月に KEK B ファクトリー加速器は Belle-II への改造のために運転停止した。停止によって新たな実験データの生成はなくなったが、それまでに蓄積された実験データの解析やモンテカルロ計算などに使われ続けた。図 3 に 2011 年度 計算サーバ利用統計を示す。おおむね 40-50%の CPU 使用率であった。加速器からの新規な実験データが生成されなかったために利用率が低下傾向にあった上(昨年度は 50%程度)に、2011 年 3 月の地震のために 4-5 月は運用停止/縮退運転を余儀なくされたことが CPU 利用率に大きな影響を与えた。そして年度後半は、共通計算機へのデータ転送のために、HSM システム上のデータアクセスを一部制限しながら運用したため CPU 使用率を押し下げる結果となった。HSM 磁気ディスクシステムは、2009 年度で最大容量まで書き込みがなされ、2011 年度末で容量の 9 割以上の書き込みが行われた(昨年度のアクティビティレポート参照)。

B ファクトリー計算機システムの 6 年間の運用全体では、CPU は 50%以上、ストレージ

ジは最大容量を使い切ったことから、おおむね順調に運用を行えたと考えている。一方で、LSF のパラメータの調整など解決しきれなかった問題もあり、現行の中央計算機の運用に反映させる方法を検討している。

震災対応は、昨年度のアクティビティレポートに示したので詳細は省略する。

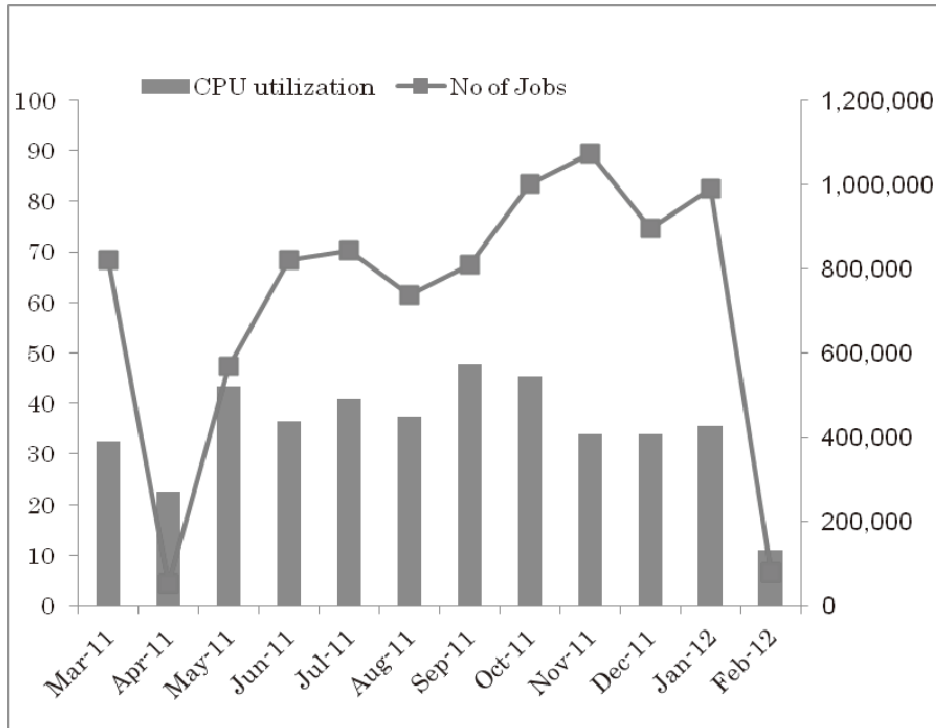


図 3 2011 年度 計算サーバ利用統計 (CPU 利用率, Job 投入数)