



最先端の素粒子物理データ解析システム 換装で処理性能とストレージ容量を強化

NSSOLの綿密な準備で600台超の機器を3カ月でリプレース

背景

素粒子物理学の国際共同実験で取得する膨大な実験データを高速に解析する大規模計算機システムを3年単位で換装している。第4期システムの目標は実験データの増加に対応した処理性能とストレージ容量の強化である。



東京大学
素粒子物理国際研究センター
准教授
真下 哲郎氏

ソリューション

832CPU/9984コアの計算サーバーおよび10.5PBバイトと大容量のディスクアレイ装置を採用。公開入札で選ばれた新日鉄住金ソリューションズの支援によって、サーバーやストレージなど600台超の機器を3カ月で換装する。

成果

短工期ながら計画通りハードウェアを換装。ソフトウェアのインストールなどを行って運用を開始した。プロセッサコア数やハードディスク装置の容量を増やすことによって、処理性能やストレージ容量を強化している。

第4期システムでは、実験データ増加への対応をこれまで以上に重視

スイスに建設された外周27kmに及ぶ大型の素粒子物理学実験装置を使い、世界38カ国178の大学・研究機関が共同で進める「アトラス実験」。同実験では、毎秒数百Mバイトと膨大な実験データが発生するため、大規模グリッドコンピューティングプロジェクト「WLCG (Worldwide LHC Computing Grid)」により分散処理を行っている。東京大学素粒子物理国際研究センターは、アトラス実験データ解析の国内拠点であり、WLCGによる世界的な分散処理の一部も担う「アトラス地域解析センター計算機システム」を2006年末に構築し、2007年から運用を開始した。同システムでは基盤ハードウェアを3年ごとに換装して処理性能を向上させてきたが、2015年9月に調達した第4期システムでは、実験データ増加への対応をこれまで以上に重視していた。

NSSOLが綿密に準備、600台超の搬入・設置などが10日以内に終了

第4期システムの基盤ハードウェアは、計算サーバー416台で832CPU/9984プロセッサコアの計算リソースを用意し、80台のディスクアレイ装置でストレージ容量を第3期の1.5倍強、10.5PBバイトに増強する仕様だった。公開入札で選ばれた新日鉄住金ソリューションズ(以下、NSSOL)は、第1~3期のシステムを手掛けてきた実績を活かし、指定通り3カ月で第4期システムへの換装を行った。

短工期の換装は綿密なキッティングで実現した。サーバーやストレージなどの機器はキッティング場でラックに搭載して配線。性能検証までを実施して、ラック単位でセンターへ搬入・設置することなどで、作業効率を高めた。サーバーとストレージ合計で600台超の機器を扱いながら、機器の搬入・設置、配線に要する期間を10日以内に収めている。

CPUコア数増加とHDD大容量化で、処理性能とストレージ容量を強化

2015年12月末に第4期システムは基盤ハードウェアの換装を終え、ソフトウェアのインストールなどを行って運用を開始している。

成果は期待通りである。システムの処理性能は、サーバー1台当たりのプロセッサコア数を増やして向上させた。ストレージ容量は、ディスクアレイ装置で使う個々のハードディスク装置の容量を2倍に増やして拡張している。これらにより、第4期システムでは、これまで以上に膨大な量の実験データに対応することができる。

今後は、アトラス地域解析センター計算機システムと、世界各地の計算センターが参加するWLCGサイトとの接続に使うネットワーク帯域の強化などによってデータの転送速度を向上させ、さらに利便性を向上させていく予定だ。

Key to Success

第4期となるアトラス地域解析センター計算機システムの目標は、実験で取得するデータの増加に対応することである。

東京大学素粒子物理国際研究センター 准教授の真下哲郎氏は「アトラス実験では、高速の電子回路などでデータを絞り込むのですが、それでも毎秒数百Mバイトもの実験データを取得します。そのデータをさらに取捨選択して解析するのですが、データ量が膨大なため世界百数十カ所の計算センターが連携して分散処理しています。アトラス地域解析センター計算機システムは、その世界規模の分散処理の中でも規模や性能などの点で非常に重要な役割を担っています」と話す。

実験データの増加に対応するには、処理性能とストレージ容量の強化が不可欠だった。

「実験機器の改良や実験の本格化によって、これからも解析する実験データは増えていきます。第4期システムは、サーバー当たりのプロセッサコア数を増やして処理性能を高め、ストレージ容量をハードディスク装置の大容量化で増やすことにしました」(同)

600台超の機器で構成する大規模システムでありながら、換装に伴うシステムの停止期間はできるだけ短くしたいと考えた。

「今回の換装プロジェクトの時期は、実験でデータを取得している最中でした。実験データの解析も通年でを行っています。世界的な分散処理システムにおける本システムの重要性を考慮し、機器の搬入・設置、配線に要する期間は10日以内と仕様書に記載しました」(同)

入札で選定されたNSSOLは、第1~

3期システムを手掛けており、第4期システムへの換装も豊富な実績を基に効率的に行った。

NSSOLが綿密にキッティング 大規模ながら計画通り換装

「今回もNSSOLは、計画通りの確に換装を行ってくれました。計算サーバーやストレージは、学外のキッティング場でラック内の配線までを行ってから、学内のセンターへ搬入しています。また、機器の配線は仕様書で詳細に規定していますが、NSSOLは仕様以上の工夫をしてくれました」(同)

NSSOLは導入前に機器を試験的に稼働させて検証までを行っている。

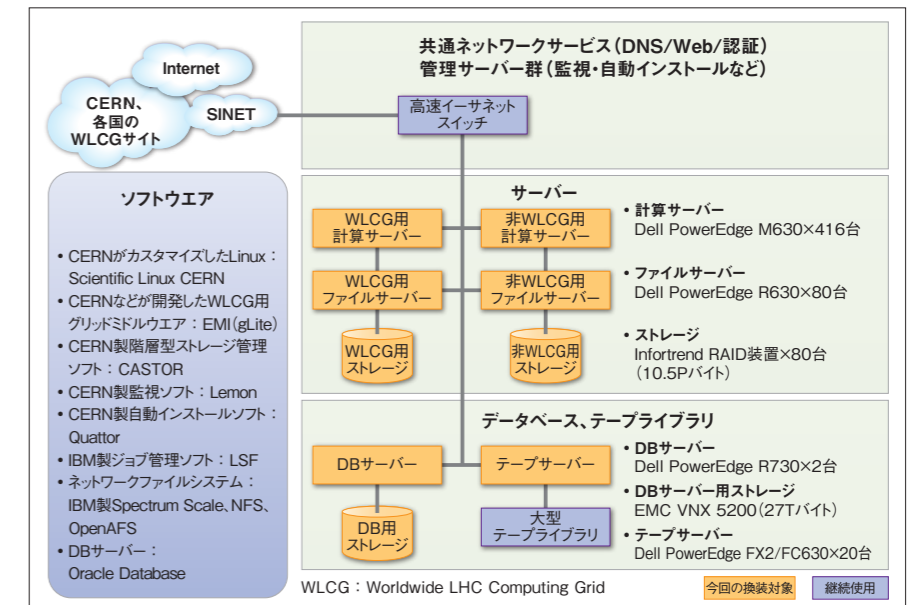
「サーバーやストレージの性能検証もキッティング場で実施しています。

システムに高い負荷をかけたとき、期待した性能が出ないことがまれにあるのですが、性能検証によって問題発生の有無を事前に調査したため、学内のセンターにラックを設置したあとは、わずかな作業でシステムを本格稼働させることができました」(同)

完成した第4期システムに対する評価は高い。

真下氏は「アトラス地域解析センター計算機システムは、世界各国のWLCGサイトの中で最も信頼性が高いシステムの一つで、扱えるデータ量もトップクラスであると評価されています。今回の換装でそのシステムの処理性能とストレージ容量がさらに増強され、利便性がこれまで以上に高まりました。今後も、世界各地のWLCGサイトとの接続ネットワークの帯域を強化することなどで、システムの利便性をより高めていく計画です」と語る。

■アトラス地域解析センター計算機システム(第4期)の概要



■コアテクノロジー

大規模なハードウェア/ソフトウェア/ネットワークをトータルで最適化するシステム構築・運用の実績と経験、HPC (ハイパフォーマンスコンピューティング)、大容量ディスクアレイ装置

■システム概要

●サーバー: PowerEdge M630×416 (計算サーバー)、PowerEdge R630×80 (ファイルサーバー、AFSサーバー)、PowerEdge R730×2 (DBサーバー) など ●ネットワーク: Brocade MLXe-32×2 (高速イーサネットスイッチ) など ●ストレージ: Infotrend ESDS3024G000F8C16D00×80 (ディスクアレイ装置) など