

核融合実験で生成される膨大なデータのアーカイビングにCloudStation dSSを採用

概要

CloudStation dSSを導入し、多種のデータで構成される核融合実験の大容量・長期アーカイビングを柔軟で運用に手間のかからないKVSシステムに切り替えることで、ハードウェアに依存しない容量追加やレプリカ遅延の解消といった導入効果を高いコストはかけずに実現。

自然科学研究機構 核融合科学研究所 <http://www.nifs.ac.jp/>

核融合科学研究所は、世界最大の超伝導プラズマ閉じ込め実験装置である大型ヘリカル装置 (LHD) などを擁し、プラズマ・核融合分野の大学共同利用機関として、国内や海外の大学・研究機関とともに双方向の活発な研究協

力を進める。優れた人材の育成や社会連携に基づいた基礎的研究・教育の強力な推進も行う。

【課題と導入効果】

課題	効果
逐次、合計で14GBほども生成される100種類ほどの様々なデータを、遅延も手間もなく確実にレプリケーションしたい	データ保存後数秒で自動的に完了するレプリケーション機能により、バックアップリスクと管理者のリソース消費を解消
ハードウェアに依存せず自由に容量追加を行えるクラウド型ストレージを実験データの中長期アーカイビングに導入したい	中長期アーカイブの運用、容量追加に絶大な柔軟性を発揮するKVSを、H/W一体型のアプライアンス形態で余計なコストをかけずに導入
過去のものを含めると膨大なサイズとなる大容量データアーカイブを楽に管理したい	管理GUIのビジュアライザで複数ノードの管理情報をまとめて確認でき、起動停止や書き込み禁止モード切り替えなど設定変更も一括で可能に

核融合科学研究所とは

太陽エネルギーの源でもある核融合プラズマの研究は、化石燃料のように二酸化炭素 (CO2) を発生せず、海水中に燃料となる物質が全て含まれていること等から、恒久的なエネルギー源として世界で最重要視される研究課題である。これによるエネルギー生産が実現すれば、人類は将来の長きにわたり安定して環境負荷の少ない新エネルギー源を手に入れることになる。

一億度にもいたる超高温・高密度の核融合プラズマは、物理学や電気工学、超伝導工学、材料工学、情報工学等の幅広い分野を包括した研究対象であり、こうした世界中の研究者の知の結節点として機能するのが、自然科学研究機構 核融合科学研究所である。



核融合科学研究所が所有するLHDの真空容器

核融合プラズマ実験で生まれる膨大なデータアーカイブ

核融合科学研究所は、世界最大の超伝導プラズマ閉じ込め実験装置「大型ヘリカル装置」(LHD) を有し、そこで生成される核融合プラズマから、温度や密度など膨大な各種データの取得・アーカイビングを行う。



自然科学研究機構
核融合科学研究所
高温プラズマ物理研究系
准教授 中西秀哉博士

「LHDには100種類ほどの計測器が備えられ、色々な波形や画像等のデータを取得します。現在は14GB程度のデータが数秒で出てくる1回の実験(ショットと呼ぶ)を180秒周期で1日に180回ほど繰り返し、それを圧縮してアーカイブ化します。こうした核融合研究の究極の目標は、プラズマ中で水素が核融合反応を起こしヘリウムになる時に放出される核融合エネルギーを取り出すことです。」

(自然科学研究機構 核融合科学研究所 高温プラズマ物理研究系 准教授 中西秀哉博士)

参考:実験ごとに生成される主な計測データ(1GB超のもの)

マイクロ波イメージング反射計(MIR)	2.8GB
燃料ペレット撮像カメラ	2.8GB
遠赤外/赤外線レーザー干渉計	1.9GB
ECEイメージング(ECEI)	1.1GB
マイクロ波反射計	1.1GB
協同トムソン散乱	1.0GB

上記の他、電流・エネルギー、プラズマ位置・形状、赤道面電子温度・密度分布、プラズマ密度分布、イオン温度分布、不純物、周辺プラズマ温度等々、多岐にわたる目的に応じた種々の計測器から、14GB超/回のデータが取得される。各計測器は、直径13.5mにもなる大型ヘリカル装置(LHD)の周囲に設置されている。

核融合プラズマ生成実験はこれまで約11万回にもおよび、取得された膨大なデータは全国の研究者によりアクセスされる。そのデータ量の総和は1PB弱になるという。

柔軟性の求められる運用、レプリカ作成での課題

アーカイブのストレージの第一世代は1997年に導入された。中央ストレージを持たず、データ収集ホストが各々ローカルストレージを備え、アプリケーションサーバーも兼ねるという構成であった。それが、2001年頃からは、中央ストレージの共有ディスク (NFS/CIFS) に対して収集ホストからデータを転送する方式となり、収集ホストがローカルストレージを持つという形態から徐々に脱却してゆくこととなった。当時のストレージとしては、ぶらっとホームのTrusRAID (SCSI DAS) が導入されている。

「その頃は、データ収集はいわゆるネットワークファイルシステム (NFS/CIFS) で賄っていましたが、計測するデータの種類は年々10種類くらいずつ増加してゆきます。これに伴い、接続が50セッションを超えたあたりからNFSサーバーが過負荷になっていきました。そこで、データ移送をFTPによるセッション・オリエンテッドな方式に切り替え、サーバー側の負荷低減を図りました。」(中西氏)

こうして2006年ころには、新たに導入されたFC RAIDでSANを構成し、そこにFTPでデータを移送する運用形態に移行した。なおデータ配布は当初よりアプリケーションサーバーを介する方式をとっており、現在まで変更は行われていない。以後、FC SAN環境が長らく使用されていたが、それは必ずしも安心して使えるものではなかったという。

「データを失ったことはないのですが、何度か重大なストップがあり、実験の途中でシステムが止まってしまうことがありました。また、ストレージシステムを拡張する際も、クラスタ型ストレージだと全面的にシステムを止めてから行う必要があります。」(中西氏)

加えて、データ量増加に伴い、レプリカ生成でも問題が生じていた。

「データがどんどん出てくるのでレプリカを作成するタイミングが難しい。データ生成は都度一瞬ですが、I/O速度やサイズがばらばらなので保存タイミングは皆違います。それらデータの一つずつ複製するとジョブが100ほど必要になるし、全部まとめてとなるとタイミングの見極めが難しく、アプリケーションレベルでの吸収も困難でした。そこで、もともとレプリカ機能の備わったストレージシステムがあれば、と考えていました。」(中西氏)

当初用いていたRAID-1による対称形ミラー構成も、瞬停等でミラーリングが外れた際の再構成に何十時間もかかるなど、早期に破綻をきたした。このためFC SAN導入以後もレプリカ作成については試行錯誤が続けられつつ、長らく課題として残っていた。

こうした中で中西氏は、より自由度が高く、管理する人間への運用負荷が少ないストレージシステムに移行したいという思いを2006年ころから抱くようになっていたという。



実験で取得されたデータは多くの研究者により参照される

各社の製品をいくつも比較検討

その結果、中西氏が検討を始めたのが、ハードウェアに依存せず容量追加やバックアップに高い柔軟性を持つストレージの導入である。大手数社を含めたいくつかの製品を「取っ替え引っ替え評価・検討した」という。

「IT展示会などで目に付くものすべて話を聞きましたが、まず金額的に非常に高くつくものがほとんどでした。ミニマムの三台構成で最初から数万円の見積もりというオーダー。またそうしたベンダでは、ライセンス料などの経常費用も高コストでした。」(中西氏)

そんななかで行き当たったのが、ぶらっとホームのKVSアプライアンス、CloudStation dSSであった。

「CloudStation dSSは、非常にリーズナブルに収まっている。予算的な移行のしやすさもあつたし、なによりまず、膨大なデータのアーカイブを行ううえで必要な機能を全て満たしていました。」(中西氏)

優れた運用性、システム拡張でも高い柔軟性を実現

膨大なデータをシンプルに、フラットに管理できるキーバリューストア (KVS) でありながら、従来にない導入しやすいコスト、そして諸々の要求を満たす仕様で、CloudStation dSSは中西氏の導入検討をクリアした唯一の製品となったCloudStation dSSは、その製品コンセプトにおいて「要求に過不足なく合致している」と中西氏は語る。

「最初は8台、そして今年度に入ってからさらに追加で1台導入しました。追加に関しても、ケーブルをただ接続するだけで、何も問題なく済みましたね。」(中西氏)



CloudStation dSSを収納するアーカイブストレージ用マウントラック

「今後ストレージの拡張を行うときも、やはりCloudStation dSSに切り替えたことは効いてくると思います。データ共有クラスタだと、必ずどこかの設定を忘れそうになったりしていたのですが、そういうことはなくなるでしょうね。」

(大学共同利用機関法人 自然科学研究機構 核融合科学研究所 技術部 小嶋護氏)

高い柔軟性を誇るKVSだから享受できるこうした導入性の高さも、中西氏を満足させる要因となったという。もちろん、膨大なデータの管理性も満足のいくものであるようだ。

「管理GUIツールのビジュアルライザも使い易いですね。エラー情報や、複数ノードの管理情報をまとめて見られる。また、起動停止もビジュアルライザから一発でできるし、書き込み禁止モードへの切り替えなども一括してできます。以前のSAN環境では、コマンドを駆使して全体の状況を把握する必要があったり、設定ファイルの書き換えなどもわかりづらかった。それに比べると、管理する側としては、一目で全体がどうなっているのかわかるので、非常にありがたいですね。」(小嶋氏)



大学共同利用機関法人
自然科学研究機構
核融合科学研究所 技術部
小嶋護氏

また、懸念であったレプリカ作成についても、高い導入効果を得ている。

「レプリケーションを必要以上に気にしなくてもよいのは実に大きいです。昨年などはデータ量の増加によって、2~3ショット (のプラズマ生成実験) 分、レプリカ作成が遅延してしまうこともあり、とても気を遣いました。それが今では、数秒でもうレプリカが作成されているので、安心してみられます。」(中西氏)

「このような研究所のストレージがニッチな要求を抱えているうえで、対応のしっかりしたぶらっとホームと協業でき、結果として要求したものをに入れてもらえた。大変感謝しています。」(中西氏)