

国立情報学研究所 (NII) が構築・運営する研究データ基盤「NII Research Data Cloud」に「ActiveScale」を採用



ソリューション概要

- 大学共同利用機関の研究データ基盤「NII Research Data Cloud」に採用
- NII Research Data Cloudを支えるストレージの選定では、高いパフォーマンスに加え、容易なスケールアウトや安定稼働を可能にする技術を検討し、ActiveScaleを採用した。
- ActiveScaleは、Amazon S3標準オブジェクト・ストレージ用インターフェースを使用したアーキテクチャを採用しているため、各研究機関への展開や連携も容易に行うことができる。

主な利点

- APIがAmazon S3と互換
- サポート一体のアプライアンス製品ならではの容易な運用
- 24/365日の製品の保守体制が安定稼働に貢献
- 無停止でのスケールアウト、バージョンアップに対応
- 省スペース、低消費電力がエコなITシステムに貢献

システムインテグレータ

- 株式会社エヌジーシー

大学共同利用機関として、学術情報ネットワーク (SINET) の構築・運用をはじめ、クラウド基盤の整備、学術情報基盤の整備およびその活用基盤の高度化を推進している国立情報学研究所 (NII)。そのNIIが運営する研究データ基盤「NII Research Data Cloud (NII RDC)」プラットフォームに採用されたのが、クアंटムのオブジェクト・ストレージ「ActiveScale」です。GakuNin RDMは、研究プロジェクト実施に際し、研究者や研究グループが研究データや関連資料を管理するための研究データ管理基盤であり、その構成要素の1つとして、ActiveScaleが活用されています。

学術基盤の開発・運営を通じて日本のオープンサイエンス推進の礎となる

インターネットの普及、デジタル時代の本格化にともない、学術研究の活動も共同研究の機会が増加。研究に関わる様々なファイル (研究データや付帯情報、論文のドラフト等) がインターネット上で共有され、同時編集されるようになっていく。このように研究活動をより開かれたものとして行っていくというオープンサイエンスに関わる日本の政策は2015年に本格化し、さまざまな施策が推進されてきた。

「学術情報ネットワーク (SINET)」の構築・運用をはじめ、大学共同利用機関として最先端の学術情報基盤の構築・運用や学術コンテンツの提供などの事業を展開している国立情報学研究所 (NII) でも、世界的なオープンサイエンスの気運を受け、そのインフラとなる学術基盤を開発・運営するために、NII オープンサイエンス基盤研究センター (RCOS) を設置した。

RCOSの使命は「国内学術研究活動における学術論文および研究データなどを管理・公開・検索できる学術基盤の構築を通じて、日本におけるオープンサイエンス推進の礎となること」であり、「RCOSで開発・提供する学術基盤を通じて、国内の学術研究活動が円滑に行われ、またその研究成果が研究プロジェクト終了後も遺失することなく、共有・リユースされることを目標としている。」

2017年よりRCOSでは、その理念を具現化し、研究データのライフサイクルを支える基盤として、NII研究データ基盤 (NII Research Data Cloud : NII RDC) の開発を開始した。NII RDCは、研究データの管理基盤 (GakuNin RDM)、公開基盤 (WEKO3)、検索基盤 (CiNii Research) から構成し、2021年から本格運用を開始している。

研究データの管理基盤 (GakuNin RDM) にActiveScaleを採用

前述した研究データの3つの基盤の中で、研究データ管理基盤「GakuNin RDM」プラットフォームに採用されたのが、クアंटムのオブジェクト・ストレージ「ActiveScale」である。

GakuNin RDMとは、研究中のクローズド (非公開、制限共有) なデータを取り扱うことに特化した研究データ管理 (Research Data Management = RDM) サービス。

研究プロジェクト実施中に、個人の研究者あるいは研究グループが既存のストレージや研究ソフトウェアと連携し、クローズドな空間で、研究プロジェクトに関わるファイルのバージョン管理や、メンバー内でのアクセスコントロールができる。また、研究公正への対応としての研究証跡を記録する機能や、ファイルを保存する機能を備えている。

2022年1月現在、全国で40の学術機関がGakuNin RDMを利用している。2021年4月より24/365日運用体制でRDMサービスを提供しており、今後は他の研究データ基盤とも連携を予定している。

国立情報学研究所 (NII) について

NIIは大学共同利用機関として大学や研究機関、研究コミュニティと連携し、世界最高水準の高速回線で国内外をつなぐ学術情報ネットワーク (SINET) を構築・運用することをはじめ、クラウド基盤の整備、オープンアクセスやオープンサイエンスを推進する学術情報基盤の整備、および、その活用基盤の高度化を推進しています。また、教育研究の国際競争力の向上や最先端研究の加速、学際的研究の発展、効率的研究の推進、大学の機能強化に貢献しています。

NII RDCに用いるオブジェクト・ストレージシステムの調達は、2020年8月意見招請、10月入札公告、12月開札を経て、クアंटムのオブジェクト・ストレージ「ActiveScale P100」が選定された。ストレージシステム選定の過程で重視した要件について、NIIプロジェクト担当者は次のように説明する。「まず、重視したのはS3との高い互換を備えていることです。このほかの要件としては、ハードウェアとソフトウェアが一体化、運用上外付けのロードバランサ等を要しない、アプライアンス製品であること、機材と性能が高レベルでバランスできていること、低消費電力・省スペース、容量単価に優れていることなどを評価し、採用に至りました。」

最適化、柔軟な拡張、容易な運用管理などアプライアンス製品のメリットを高く評価

NII RDCは外部ストレージとの接続に、クラウドストレージのデファクトスタンダードであるAmazon S3 APIを使用している。ActiveScaleもインターフェースにAmazon S3 APIを包括したアーキテクチャであるので、Amazon S3 互換オブジェクト・ストレージとの間でのデータ複製やデータ移行も容易で、且つバックアップ/アーカイブや災害復旧 (DR) 構成での利用にも対応することができる。

また、ハード/ソフトが一体化したアプライアンス製品であるため、外部ロードバランサ (負荷分散装置) を必要とせず、高いパフォーマンスを安定して発揮することができる。システムノードとストレージノードはそれぞれ別々に増強 (スケールアウト) 可能なアーキテクチャを採用。システムノードは1台単位での増強、ストレージノードは6ノード単位で増設が可能となっている。容量の増強作業は無停止で実施でき、最大27PB (Raw) までのスケールアウトに対応、負荷分散する機能も備える。

一般的なオブジェクト・ストレージの場合、事前に定義した格納先にオブジェクトを格納する。そのため格納先のノードが故障していた場合、ノードに一旦書き込むが、復旧後にリバランスが必要になる。一方、ActiveScaleは、ストレージの増設やHDDの障害時にデータのリバランスが発生しない「BitSpread™」という独自の特許技術を採用する。これはオブジェクトが来たタイミングで動的に

格納先を決定し、故障ノードは初めから格納先の対象から動的に排除する。復旧後は動的に格納先の対象候補として追加することで、リバランスが発生せず、パフォーマンスにも影響を及ぼさない。

さらに、サポート窓口も1つであるため、万一の障害の際にも切り分けの必要がなく、迅速な対応が受けられる点もアプライアンス製品のメリットだ。

運用管理においても、ソフトウェアのバグフィックス、パッチ、アップデートが無停止で実施でき、ソフトウェアファームウェアともにローリングアップグレードにて実施することができる。

加えて、省スペース、低消費電力もActiveScaleのメリットである。

「今回の堅牢で信頼性の高いオブジェクト・ストレージの採用はNII RDCの効率的で安定した運用に貢献しています」とNIIプロジェクト担当者。

研究データ管理サービスは各研究機関が契約、所有しているクラウドやストレージとも連携する。クアंटムとしては、各研究機関に向けてActiveScaleの持つAmazon S3完全互換、容易な拡張といったメリットをアピールし、採用の拡大を目指していく方針だ。

研究データ基盤 : NII Research Data Cloud

オープンサイエンスの推進のために、3つの基盤で研究データのライフサイクルをサポート

