splunk> listen to your data™

CASE STUDY

独立行政法人 理化学研究所 計算科学研究機構

1日あたり数百GBのログを効果的に蓄積 「京」のシステム運用に安心感を提供

和光研究所のログ管理に導入された実績を評価してSplunkを採用。
「京」、ネットワーク、HPCIサーバ群の個別および横断的なログ管理を実現

理化学研究所

「Splunkは、RDBのような容易な操作性で、高速にログが検索できるので非常に便利だと感じました。ちょっと調べたいときには、非常に便利なツールです。一方、「京」のような巨大なシステムは、何が起きるか想定できません。しかしSplunkでログを蓄積していることで、何か起きても迅速に対応できるという安心感があります」

運用技術部門 副部門長
(兼) システム運用技術チーム チームヘッド
博士(理学) **庄司 文由氏**

OVERVIEW

業種
- 研究開発

課題／背景
- システムから出力される膨大なログの効率的な管理
- システムごとのログ管理と横断的なログ管理の両立
- 効率的なログの蓄積と高速な検索、柔軟な分析機能

ソリューション
- Splunkを採用したシステム横断的なログ監視・管理システムの構築

導入効果
- 高い操作性で検索が非常に容易
- ログを蓄積する場所を分散でき、一括検索可能
- Splunkの利用者の権限設定が容易
- REST APIによる容易なアプリ連携
- 圧縮機能やスケールアウトの拡張性

データソース
- システム稼働状況ログ

(※1) TOP500:世界最速のコンピュータシステム上位500位を定期的にランク付けするプロジェクト
(※2) HPCI:High Performance Computing Infrastructure

世界一を獲得した「京」を運用する計算科学研究機構

1917年に財団法人として創設され、株式会社や特殊法人を経て、2003年10月に文部科学省所轄の独立行政法人として再発足した独立行政法人 理化学研究所(以下、理研)は、日本で唯一の自然科学の総合研究所として、物理学、工学、化学、生物学、医科学など、幅広い分野の研究を推進しています。コンピュータシミュレーションにより未来を科学的に見通す「予測の科学」の確立を目指す理研計算科学研究機構では、スーパーコンピュータ(スパコン)「京(けい)」の運用が重要なミッションの1つ。「京」は、「TOP500(※1)」において、2011年6月および2011年11月に1位を獲得しています。

スパコン・システムなどの膨大なログをいかに活用するか

計算科学研究機構では、「京」を中核に、ネットワークシステム、HPCI(※2)サーバ群の3つのシステムを運用しています。この3つのシステムを効率的に運用・管理するためには、システムから出力される膨大なログをシステムごとに蓄積すると共に、システムごと、またはシステム横断的に分析・活用するための仕組みを導入することが必要でした。

専任技師 博士(情報科学)である黒川原佳氏は、「ログ管理は目立たない作業ですが、システムの状況を把握するためには不可欠です。このとき各システム管理者が担当するシステムの稼働状況ログを管理したいというニーズと、全体を統括管理する管理者がシステムログを横断的に管理したいというニーズを両立できる仕組みが必要でした」と話します。

そこでいくつかのログ分析ツールを比較検討した結果、Splunkの採用を決定します。採用を決めた理由を黒川氏は、「オープンソースソフトウェア(OSS)を検討しましたが、OSSを継続的に管理・運用するには人的リソースが足りないため見送りました。Splunkは2009年3月に、情報基盤センター(埼玉県和光市)で採用された実績があり、利用者の評価も高く、必要な要件を満たしていることを評価しました」と話しています。

Splunkの大量のログ管理と高速な検索を評価

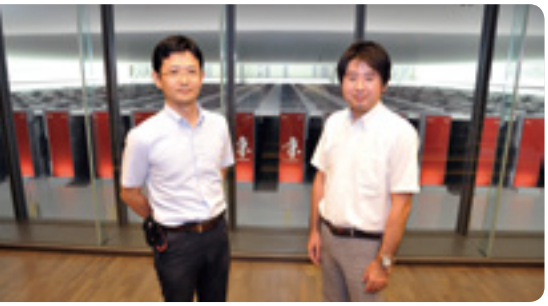
計算科学研究機構では2011年6月より、「京」、ネットワークシステム、HPCIサーバ群の3つのシステムにSplunkを導入し、ログ分析システムの運用を開始しています。黒川氏は、「1日に約10GB、ピーク時には1日に約50GBのログを蓄積しています。現在、「京」のシステムを改良していますが、改良後には1日あたり数百GBのログが生成される見込みです。その場合でも、ログ蓄積に困らない容量を確保してあります」と語ります。

Splunkを利用することで、外部からの攻撃や不正アクセスなどのセキュリティ関連のログ、ネットワーク機器やサーバの負荷状況や温度等の管理、「京」のジョブ運用状況など、機構内のあらゆるログを蓄積しています。黒川氏は、「これだけ膨大なログの量になるとログを管理するだけでも大変ですが、それに留まらずいつでも利用できるようにしておかなければ意味がありません」と話します。

運用技術部門 副部門長(兼)システム運用技術チーム チームヘッドで博士(理学)である庄司文由氏は、「どこから、いつ出力されたログなのかを探すのもたいへんだし、事象は連鎖して出てくることが多いので、Splunkのような大量のログを管理しながら高

運用技術部門 副部門長
(兼) システム運用技術チーム チームヘッド
博士 (理学) **庄司 文由氏**

理化学研究所 情報基盤センター 専任技師
(兼) 運用技術部門 システム運用技術チーム 研究員
博士 (情報科学) **黒川 原佳氏**



「京」は、生命科学や気象、防災など、約130のプロジェクトで利用されている。

速に検索できる仕組みは非常に重要でした」と話しています。

巨大システムである「京」の管理にも安心感を提供

Splunkを導入した効果を黒川氏は、次のように語ります。「Splunkは非常に高い操作性で、特に導入に関しては、まったく苦労はありませんでした。またログを蓄積しておく場所を分け、利用者権限を容易に設定できるのも便利です。さらに効率的に圧縮してくれる機能や高速な検索、スケールアウトによる拡張性なども高く評価しています」

黒川氏は、「ログの流量が急に増えたときにアラートを配信する仕組みを構築したことで、問題がすぐに分かるのも便利です。また外部のベンダーに任せているネットワーク運用が正しく行われているかの監査も行っています。さらにGUIではなく、CUIを使いたいという現場のニーズに、REST APIによる連携で対応できます」と語ります。

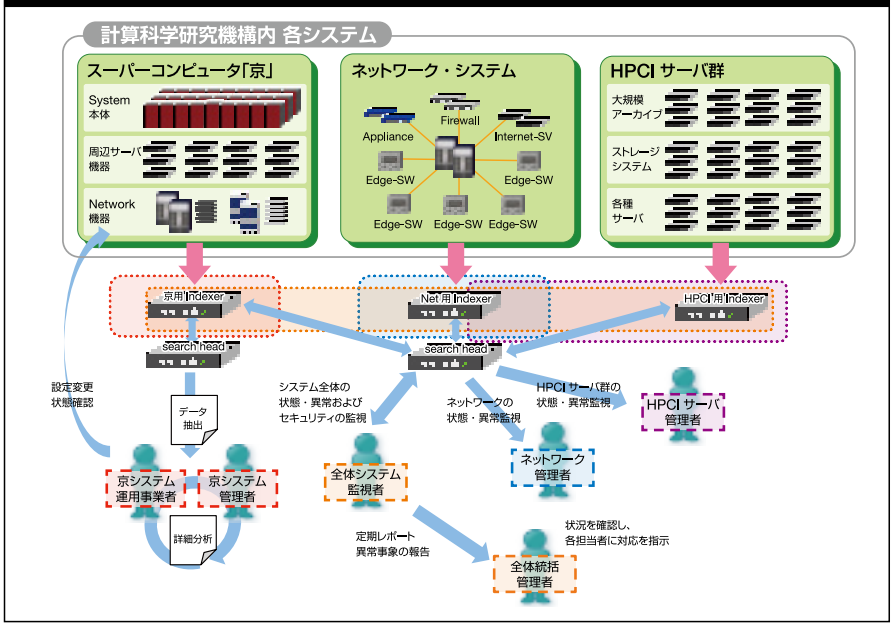
また庄司氏は、「Splunkは、RDBのような容易な操作性で、高速にログが検索できるので非常に便利です。不明点を問い合わせても、すぐに答えが返ってくるサポートも評価しています。「京」は巨大なシステムなので、何が起きるか想定できませんが、Splunkを導入したことで、何か起きても迅速に対応できる安心感があります」と話しています。

点在する事象をつないで障害の予兆を発見

今後の展望について黒川氏は、「サーバシステムのエラーからネットワークの障害を見つけだすなど、点在する事象をログ分析でつなぐことで、障害の予兆を発見できればと思っています。そのためデータ間の相関を分析して、異常検知の精度向上を目指しています。蓄積したログは宝の山で、使い方によってさまざまな効果が期待できるのですが、人手が足りないのが実情です」と話します。

一方、庄司氏は、「ジョブスケジュールに関して、蓄積されたログを分析して、最適化できないかを模索しています。また今後は、Splunkのダッシュボードで状況を可視化して利用者サービスに活用したいと思っています。そのためには、今後もSplunkは不可欠なツールです」と話しています。

計算科学研究機構内システム構成図



無料ダウンロード

Splunkは無料でダウンロードができます。1日500MBまでのデータのインデックスを作成でき、Splunk Enterpriseのあらゆる機能を60日間無料でお試しいただけます。無料期間終了後でも期間中でもいつでも、無期限のトライアルライセンスへの切り替えやEnterpriseライセンスの購入が可能です。今すぐライセンスの購入をご希望の場合は、以下のメールアドレスよりお問い合わせください。

お問い合わせ先: splunkjp@splunk.com

Splunk Services Japan合同会社　〒100-6509 東京都千代田区丸の内1-5-1新丸の内ビルディング　EGG JAPAN 日本創生ビレッジ　オフィス31　代表電話; 03-6386-0785

splunk> listen to your data™

www.splunk.com