

## Yahoo! JAPANが 中核サービスを支えるRDB基盤の データ保護システムを刷新

HPE StoreOnce System重複排除テクノロジーが  
大規模データ保護のコストパフォーマンスを変えた

“半期に一度Oracle Data Guard環境を増築するの  
と同等のコストで、HPE  
StoreOnce Systemによる  
『バックアップ／災害対策  
システム』を構築することが  
できました”

—ヤフー株式会社  
データ&サイエンスソリューション統括本部  
データプラットフォーム本部  
開発3部 DBMS技術 リーダー  
桑島 宏明氏




### 目的

Yahoo! JAPANの中核的なサービスを支えるOracle データベース基盤の「バックアップ／災害対策システム」の刷新。データ保護の信頼性を維持しながら、増大し続けるデータ量を圧縮し、復旧に要する時間を短縮させる。

### アプローチ

従来の遠隔地へのバックアップ手段として利用していたOracle Data Guardによるデータ保護を、バックアップアプライアンスによる重複排除とレプリケーション機能に置き換えることで、データ保護にかかる諸問題を解決し、運用面・機材面でのコストも抑制する。

### ITの効果

- Oracle RMANによるバックアップ環境にバックアップアプライアンス「HPE StoreOnce System」を採用し、既存のバックアップデータ保存用として利用していたNAS製品を段階的に撤廃
- Oracle RMANに対応するプラグイン「Catalyst Plug-in for RMAN」により、サーバー側で重複排除を実行しネットワーク帯域の逼迫を解消
- 遠隔地バックアップ運用のみが必要だった環境に対して、Oracle Data Guardによる遠隔地レプリケーションを「HPE StoreOnce System」の重複排除とレプリケーション機能に置き換えて運用を容易にし、データ保護の信頼性も向上

### ビジネスの効果

- 実機テストを行った他社製品との比較でRMANバックアップ時間を1/3、RMANリカバリ時間を60%短縮
- データ保護の運用手順を標準化しOracle製品に精通した技術者不在でも安全なデータ保護・復旧が可能に
- 同等バックアップアプライアンス製品との比較で導入および保守コストを25%削減
- Oracle Data GuardとNAS製品による従来環境の1年分の拡張コストで、HPE StoreOnce Systemの新規導入を実現



ヤフー株式会社  
データ&サイエンスソリューション統括本部  
データプラットフォーム本部  
開発3部 DBMS技術  
リーダー 桑島 宏明氏



ヤフー株式会社  
データ&サイエンスソリューション統括本部  
データプラットフォーム本部  
開発3部 DBMS技術  
佐藤 誠氏



ヤフー株式会社  
データ&サイエンスソリューション統括本部  
データプラットフォーム本部  
開発3部 DBMS技術  
井上 和博氏

ヤフオク!、Yahoo!ショッピング、Yahoo!ウォレットなど——Yahoo! JAPANの中核的なサービスを支えるデータベース基盤の「バックアップ/災害対策システム」が刷新された。増大し続けるバックアップデータを抑制するとともに、複数データセンター間でのレプリケーション時間を大幅に短縮したのは、最先端の重複排除テクノロジーを備えたHPE StoreOnce Systemである。Yahoo! JAPANが構築した本システムは、大規模データ保護におけるコストパフォーマンスの常識を変えるものだ。

## チャレンジ

### 創業20年、インターネットの進化とともに

2016年4月、Yahoo! JAPANが日本語での情報検索サービスを開始して20年を迎えた。月間PV数699億という国内最大規模のポータルから提供されるサービスの数は、現在ゆうに100を超えている。ヤフオク!、Yahoo!ショッピング、Yahoo!ウォレットといった多彩なサービス群は、私たちの暮らしを豊かに、快適にするだけでなく、社会インフラとしての使命を高めている。データ&サイエンスソリューション統括本部データプラットフォーム本部 開発3部 DBMS技術 リーダーの桑島宏明氏は、次のように話す。

「Yahoo! JAPANでは、創業20周年に際して『UPDATE JAPAN』というビジョンを掲げました。情報技術で今ある課題を解決するとともに、未来志向で日本をより良くしていきたいという決意を込めたものです」

データ&サイエンスソリューション統括本部は、多種多様なサービスを横断するビッグデータの分析によって、顧客体験の改善やサービス価値の向上を目指している。同時に、サービス提供や分析の実行基盤であるデータベース環境の設計・構築・運用も担う。同統括本部で扱うデータ管理システムは、商用・OSSのリレーショナルデータベースから、オブジェクトストレージ、Hadoopまで幅広い。

「決済系、顧客管理系といったミッションクリティカルな領域には、商用リレーショナルデータベース製品Oracle 11g/12cを採用しています。2016年7月、このデータベース基盤のバックアップおよび災害対策のためのシステムを刷新しました。従来のOracle Recovery Manager (RMAN)によるバックアップ、Oracle Data Guardによる遠隔地レプリケーションを見直したのです」

Yahoo! JAPANは国内屈指の大規模Oracleユーザーとして知られる。技術とノウハウの蓄積も豊富だ。バックアップ/災害対策システムの見直しを決断した理由はどこにあったのか。

「データ量の増大に伴うバックアップ時間の伸張、遠隔地レプリケーションの運用負荷の高まり、データ保護にかかるトータルコストの増大が問題となっていました。これらを解決するため

に、重複排除機能を使えるバックアップアプライアンスに着目したのです」(桑島氏)

Yahoo! JAPANがデータベース基盤のバックアップ/災害対策システムに新たに採用したのは、最先端の重複排除テクノロジーを備えた「HPE StoreOnce System」である。

## ソリューション

### HPE StoreOnce Systemの 重複排除テクノロジーを利用

HPE StoreOnce Systemは、複数サーバーの統合的なバックアップや災害対策に最適なデータ保護ソリューションとして急速に支持を伸ばしている。「StoreOnce Catalyst」と呼ばれる重複排除テクノロジーを搭載し、4Kという細かな粒度でブロックデータを比較して高い重複排除率を実現する。最大の特長は、HPEが「連携型重複排除」と呼ぶ機能にある。バックアップ対象となるサーバーとHPE StoreOnce System本体が連携し、重複排除の分散処理を効果的に実行するものだ。

「データベース環境のバックアップで最も大きな課題は所要時間の伸張です。データ増によりネットワーク帯域が逼迫し、1インスタンスあたり数時間を要している状況でした。バックアップに必要な容量も右肩上がりが増えていました。重複排除によりデータ量を削減すれば、こうした状況を大きく改善できると考えたのです」と同部の佐藤誠氏は話す。

バックアップの対象となるデータベースはおよそ230インスタンス、データの総量は50TBに達する。従来はOracle Recovery Manager (RMAN)を利用し、バックアップサーバーを介してNASストレージにバックアップを取得してきた。これに対してStoreOnce Catalystでは、RMANと連携して、Oracle DBの稼働するサーバーからダイレクトにHPE StoreOnce Systemへのバックアップが可能だ。

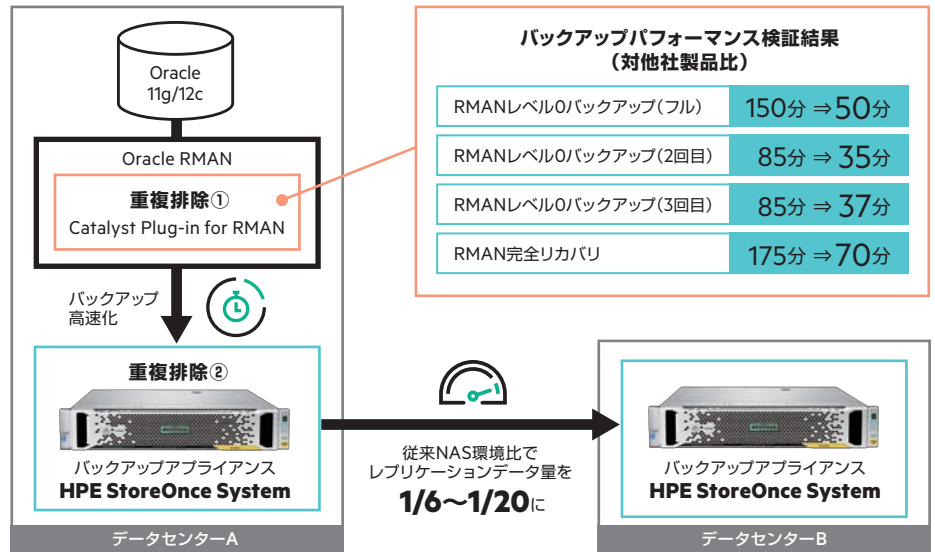
「サーバー側で重複排除済みのデータを送りますので、ネットワーク帯域を圧迫することなく高速なバックアップが可能になると期待しました。バックアップサーバーとNASストレージをこれ以上増やさなくていいというメリットも大きなものです」(佐藤氏)



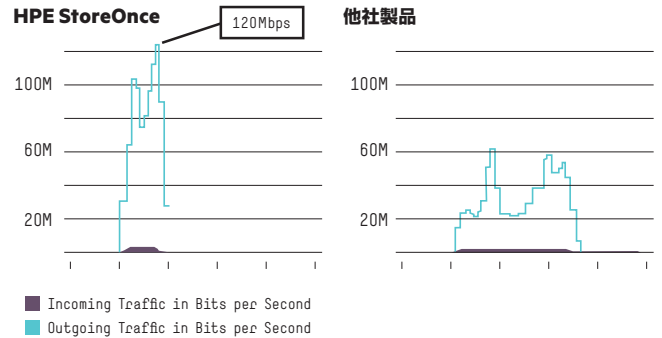
ヤフー株式会社  
データ&サイエンスソリューション統括本部  
データプラットフォーム本部  
開発3部 DBMS技術  
山本 秀平氏



ヤフー株式会社  
システム統括本部  
サイトオペレーション本部  
インフラ技術4部ストレージ  
熱田 啓保氏



Yahoo! JAPANが行った検証によると、HPE StoreOnce(左)が最大120メガビット/秒を発揮しているのに対し、他社製品(右)は1/2程度の性能にとどまっている。この性能差はバックアップ時間に直結する。



Yahoo! JAPANでは、数100km離れた複数のデータセンター間で相互にデータを二重化しているが、「Oracle Data Guardによるデータレプリケーションにも課題があった」と佐藤氏は続ける。

「Oracle DBのアーカイブログをリアルタイム転送し保護してきました。しかし、データ量の増大によって所要時間が伸長するとともに、高負荷が影響して動作が不安定になっていました。レプリケーションが正常に完了できなかったときなどに、Oracle Data Guardに精通した技術者にしか復旧できないことも問題でした」

桑島氏は、「Oracle Data Guardの運用は私たちにあって大きな負担でした。これをHPE StoreOnce Systemのレプリケーションに置き換えれば、重複排除によって転送時間を短縮できるだけでなく、属人的な運用体制からも脱却できると考えました」と狙いを話す。

### 他社製品比で 約3倍のバックアップ性能を実証

Yahoo! JAPANは複数のバックアップアプライアンス製品を比較検討した。HPE StoreOnce Systemが選ばれたのは、実際に重複排除率とパフォーマンスを測定した結果、明らかな優位性が確認できたからである。同部の井上和博氏は、テスト結果を次のように説明する。

「RMAN レベル0バックアップ(完全バックアップ)、同じく2回目、3回目の増分バックアップの実行時間と重複排除率を測定しました。さらに、RMAN完全リカバリの所要時間も測定しています。その結果、HPE StoreOnce Systemでは他の製品との比較において、バックアップ時間を1/3~1/2に短縮し、リストア時間を1/2以下に抑えられることがわかったのです」

比較検討した製品はいずれもサーバー側での重複排除が可能だったが、実効性能に大きな差があることが明らかになった形だ。バックアップアプライアンスを評価する視点には大きく2つ。ひとつは「重複排除率」であり、転送するデータの絶対量を抑制するものだ。もうひとつは「転送速度」だ。テストでは、HPE StoreOnce Systemは最大120メガビット/秒のパフォーマンスが測定されており、これは他社製品の約2倍に達する。

同部の山本秀平氏は、「私たちの目的はバックアップ所要時間の短縮です。重複排除率では大きな差はありませんでしたが、所要時間に最大で3倍もの差が出たことに驚いています。重複排除処理に要する時間の差が、この結果につながったものと考えています」と分析する。

システム統括本部 サイトオペレーション本部  
インフラ技術4部ストレージの熱田啓保氏は、ス

## ソリューション概略

### 導入ハードウェア

- HPE StoreOnce System

“データ保護の信頼性の向上、重複排除による転送時間の短縮、自社の標準化方針への合致——すべての要件を満たしながら、他社製品の3倍の性能と導入コスト1/2を実現することが明らかになりました。その結果、HPE StoreOnce Systemを採用することとなりました”

ヤフー株式会社 データ&サイエンスソリューション統括本部  
データプラットフォーム本部 開発3部 DBMS技術 リーダー 桑島 宏明氏

ストレージ製品を技術視点から評価する立場から次のように話す。

「Yahoo! JAPANでは8万台規模のサーバーを運用していますので、インフラ機器と運用の標準化は重要なテーマです。たとえば、標準ラックにムリなく収容できるか、メンテナンス時に他の機器に干渉しないかは仕様書から読み取れないことも多く、実機での確認が不可欠です。HPE StoreOnce Systemは、ハードウェアの仕様、ソフトウェアの品質、パフォーマンス、可用性のすべてにおいて私たちの要求に合致していました」

同じ実効容量で本体を比較すると、HPE StoreOnce Systemは高さ4U、奥行きは一般的なラックマウント型サーバーと同じ750ミリだ。これに対して、比較対象となった機種は高さ10U、奥行きは901.7ミリとおよそ150ミリ長く、標準ラックへの収容が難しいという問題もあったという。

「データ保護の信頼性の向上、重複排除による転送時間の短縮、自社の標準化方針への合致——すべての要件を満たしながら、他社製品の3倍の性能と導入コスト1/2を実現することが明らかになりました。その結果、HPE StoreOnce Systemを採用することとなりました」(桑島氏)

## ベネフィット

### 段階的に従来のNAS環境を撤廃予定

HPE StoreOnce Systemは、データ量の増大に伴うバックアップ/災害対策の課題に、シンプルかつ投資対効果に優れた解決をもたらした。段階的に既存のNAS製品を撤廃していくことで、さらなるコスト削減効果も期待できる。

「従来のNAS環境と比較すると、データ容量は1/6程度になっています。データベースは重複

排除が効きやすい特性があるため、最大1/20まで改善できるものと期待しています。運用が進むにつれ、HPE StoreOnce Systemのコスト削減効果はさらに大きくなっていくはずです」(山本氏)

「従来は半期に一度、サーバーとNASを購入してOracle Data Guard環境を増築してきましたが、上限なく拡大していくシステムとコストに歯止めをかけることができませんでした。Oracle製品に精通した技術者不在でも安全なデータ保護・復旧が可能になったこと、バックアップ用ネットワークを増強せずに高速化できたことも大きいですね」(井上氏)

バックアップアプライアンスを検討する際には、導入・運用コストの構成要素を知っておくことも重要だ。StoreOnce Catalystはライセンスコスト不要で利用できることに注目すべきだろう。サーバーの台数やバックアップデータが増えようともコストへの影響が心配ない。他社製品ではボリュームライセンスが適用されることもあるので注意が必要だ。桑島氏が次のように語って締めくくった。

「重複排除ライセンスの有無は、バックアップアプライアンスの導入・運用コストに大きく影響します。結果として私たちは、半期に一度Oracle Data Guard環境を増築するのと同様のコストで、HPE StoreOnce Systemによる『バックアップ/災害対策システム』を構築することができました。本システムは、大規模データ保護におけるコストパフォーマンスの常識を変える好例となるに違いありません」

詳しい情報

**HPE StoreOnce Systemについてはこちら**  
[www.hpe.com/jp/ja/storage](http://www.hpe.com/jp/ja/storage)



日本ヒューレット・パッカード  
公式ソーシャルメディア

[facebook.com/HPEJapan](https://facebook.com/HPEJapan)

[twitter.com/HPEJapan](https://twitter.com/HPEJapan)

[youtube.com/HPEJapan](https://youtube.com/HPEJapan)

