

データ保存に最適な、優れた圧縮効率と高速性を両立するiStorage HSシリーズ

廣永 龍治 川名部 正純

要旨

ICTシステムで取り扱うデータ量が爆発的に増加していくビッグデータ時代には、ストレージは予測困難なデータ増加に対応していく必要があります。iStorage HSシリーズ (HYDRAsstor) は、革新的なグリッドアーキテクチャの採用により、高い性能、拡張性、信頼性、及び運用管理の省力化を実現した、この要望に応えられる世界でも最高レベルの性能を誇るスケールアウトストレージです。

本稿では、HYDRAsstorの特長と、NEC Cloud IaaSの要件を取り込み実装した新機能について紹介します。



オブジェクトストレージ/帯域制限/ビッグデータ/重複排除/スケールアウト/高信頼性

1. まえがき

NECでは、増大している企業データ、メールアーカイブ、画像データなどのビッグデータを高い信頼性で格納し、データ量の増加に応じて拡張が可能な、バックアップ・アーカイブ向けスケールアウトストレージ基盤「iStorage HSシリーズ (HYDRAsstor)」を提供しています。

HYDRAsstorは、業界最高レベルの性能を持ち、ビッグデータの格納に適した、次のような特徴を持っています。

- (1) 重複排除による高効率なデータ格納
- (2) 性能・容量の動的拡張と自動最適化
- (3) RAIDを超えた高信頼性
- (4) ノード交換による長期データ保存

今回、クラウド基盤サービス「NEC Cloud IaaS」の要件に合わせて、オブジェクトストレージインタフェース及び帯域制限機能を、新たにHYDRAsstorでサポートしました。

次章から、HYDRAsstorの特長と新機能について具体的に紹介していきます。

2. HYDRAsstorの特長

2.1 重複排除による高効率なデータ格納

HYDRAsstor独自の重複排除技術 (DataRedux) と物理圧縮により、データの格納効率を格段に高め、格納するデータ容量に対する実際の物理ディスク容量が大幅に縮小されます。書き込まれたデータは、DataReduxにより重複がチェックされ、既にストレージ内に書き込み済みのデータと重複するデータが排除されます。更に、残った重複していないデータは、物理圧縮してから格納されます。重複チェック処理・書き込み処理は、システム内の全ノードのリソースを利用して分散処理されるため、極めて高速に行われます。

DataReduxは、既存データとの重複を最大限に検出するよう知的にデータを可変長に分割します。これより、固定長のデータ分割手法では検出できなかったデータの重複を最大限に検出することが可能となります。図1は、格納されている元ファイルの一部にデータを挿入したものを格納した場合に、DataReduxを使用すると挿入後のブロックも重複を認識できる様子を示しています。

この重複排除技術により、ディスクへのデータ転送量と、保存するデータに対する実際の物理ディスク容量が大幅に縮小されます。このため、日々のディスクへのデータ書き込みを

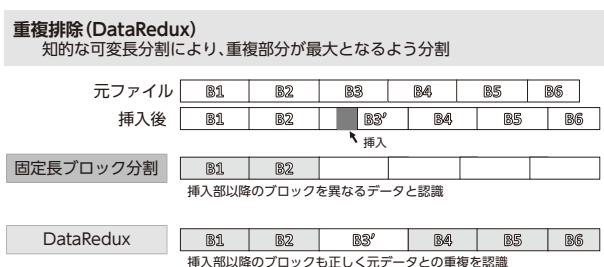


図1 重複排除 (DataRedux)

高速かつ低コストで実現することが可能です。また、この技術を遠隔レプリケーションにも適用することにより、転送データを更に圧縮することが可能になります。これにより、遠隔サイトへのデータ転送量を飛躍的に小さくでき、帯域の小さい低速回線による遠隔レプリケーションを実現できます。

2.2 性能、容量の動的拡張と自動最適化

HYDRAsstorで採用している独自のグリッドアーキテクチャでは、ノードを追加していくことで性能と容量を動的に拡張していくことができます。また、複数ノードに分散されたデータは、仮想化された単一のストレージプールと見えるため、プロビジョニングを必要としません。

HYDRAsstorは、データリクエストの処理を行うアクセラレータノードと、データブロックを実際に保存するストレージノード、アクセラレータノードとストレージノードの両方の機能を併せ持つハイブリッドノードの3種類のノードから構成されます。アクセラレータノード、ストレージノード、ハイブリッドノードを追加することにより、性能と容量を動的に拡張することができます。

性能は、1ノード当たり最高40TB/h (OpenStorage高速重複排除オプション使用時)であり、外部ネットワーク環境の条件が許す範囲でニアに性能を向上させることができます。また、容量は、156TBから37.4PBまで拡張することができます (20倍圧縮時の論理容量)。ノード追加は、データ格納位置を考慮することなく、必要に応じて実施可能です。更に自動最適構成技術 (Dynamic Topology) により、追加ノードは自動的に既設システムに認識され、容量、信頼性及び性能が最大となるよう、データは自律的に最適な構成に分散再配置されます。

Dynamic Topologyにより、従来は極めて複雑であった、以下の運用管理が簡便となり、管理コストを大幅に削減できます。

- ・ 保存データ量の増大に伴う容量の拡張
- ・ データ転送量の増大に伴う性能の拡張
- ・ 性能ボトルネックの改善
- ・ 障害発生時のノード交換

2.3 RAIDを超えた高信頼性

第2章1節で説明した重複排除技術を用いると、1つのデータブロックを複数のデータが共有します。このとき、1つのデータブロックが消失した場合の影響は、そのデータブロックを参照しているすべてのデータに及ぶため、影響が広範囲にわたる恐れがあります。そこで、HYDRAsstorでは、従来のRAID (Redundant Array of Independent Disks) よりも更に強固な信頼性を実現するため、データの分散冗長配置 (Distributed Resilient Data) を行います。Distributed Resilient Dataでは、保存するデータブロックを分割し、冗長コードを付加して複数のストレージノードへデータを分散格納することにより、信頼性を向上させています。

図2は、元のデータブロックを9つに分割し、3つの冗長コードを付加する場合の例です。この例では、1から12までの分割データ (フラグメント) が4台のストレージノードに分散して配置されています。このとき、12のフラグメントのうち同時に3つまで失っても、元のデータを復元することが可能です。このときの信頼性は、一般的に最大で2台のHDD同時故障にまで耐えることで知られるRAID6よりも優れています。更にこの冗長度は、保存データの重要度などに応じて自由に設定することができ、管理者は、柔軟にシステムを

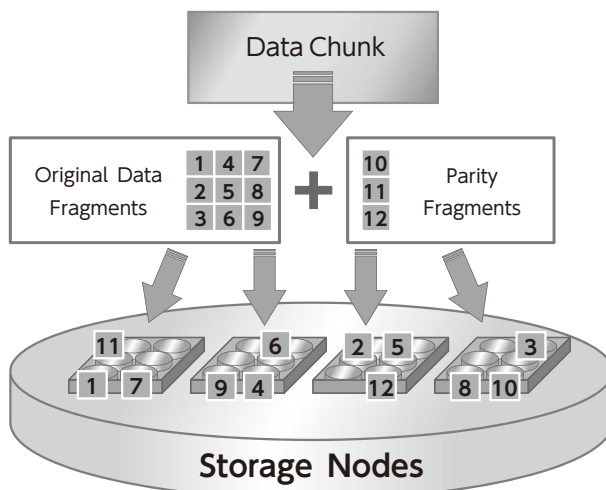


図2 分散冗長データ配置

構築・管理することができます。

HYDRAsstorは万一の故障の際、故障部分を自動的に検出し、バックグラウンドで再構成の処理を実施します。そのため、一般に管理者に必要とされる面倒な管理業務を必要としません。データの復旧はシステム内の利用可能なHDD/ノードを使用して行われるため、いわゆるRAIDのホットスワップディスクなどを準備しておく必要はありません。また、1つの障害HDD上のデータは、論理的に有効な部分のみを複数のHDD上に復旧するため、RAID技術による復旧よりも数倍から数十倍高速に行われます。更に、この再構成処理は十分な処理能力を持った複数のストレージノードで、他に実行中の処理を妨げるようなオーバヘッドをかけずに行われます。

図3は、故障により失われたフラグメント2、5、12が、即座に検知されて自動的に他のストレージノードへ再構成・再配置される様子を示しています。

この分散冗長配置技術により、既存ディスクストレージ製品を大幅に上回る信頼性と、障害発生時の管理コストの削減を実現しています。

2.4 ノード交換による長期データ保存

HYDRAsstorの自動最適化技術により、システムのハードウェアを更新する際に、データ移行を行わなくても、データを保持したまま古いノードを新しいノードに置き換えることが可能です。

新しいノードをHYDRAsstorに追加すると、新しいノード

へデータが自動的に再配置されます。次に、古いノードをシステムから削除すると、古いノードにあったデータは、他の残されたノードに、システム全体でバランスをとれるように自動的に再配置されます。

このようにしてノード交換を計画的に行い、順次古いノードを新しいノードに交換していくことで、データ移行なしに、システムを徐々に新しいハードウェアに置換していくことが可能となり、長期間大容量のデータを保管することができるようになります。

3. HYDRAsstorの新機能

HYDRAsstorはデータアクセスプロトコルとして、NFS/CIFS及びSymantec NetBackupとの専用プロトコルであるOST (OpenStorage Technology) をサポートしています。それに加え、NEC Cloud IaaSの要件に応え、新たにオブジェクトストレージインタフェースをサポートします。

オブジェクトストレージインタフェースは、Webアクセスに使用されているHTTPプロトコルに準拠した、REST APIを使用しており、インターネットとの親和性が高く、Unix、LinuxやWindowsなど幅広いプラットフォームから使用できます。一方で、書き込まれたデータの一部の書き換えやランダムアクセスや、頻繁に更新されるデータの格納には適しておらず、保存した後で変更されることの少ない、映像データ、バックアップファイル及びアーカイブデータなどの保存に適しています。

HYDRAsstorは、オブジェクトストレージインタフェースとして、パブリッククラウドでデファクトスタンダードとなっているAmazon Simple Storage Service (Amazon S3) API及び、OpenStack環境でオブジェクトストレージとして使用されるSwiftのAPIと互換性を持つインタフェースをサポートします。

NEC Cloud IaaSでは、HYDRAsstorの本インタフェースを使用して、Amazon S3互換のオブジェクトストレージサービスを提供します。

また、NEC Cloud IaaSの提供するオブジェクトストレージサービスでは、複数テナントのユーザーから同時にアクセスが行われます。この際、特定のテナントから大量のアクセスが行われると、他のテナントのユーザーに対するレスポンスが悪化する可能性があります。

そこで、HYDRAsstorでは、テナントごとにread/writeの

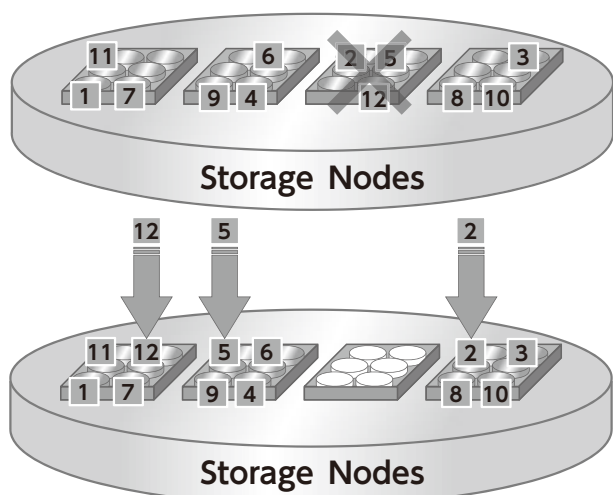


図3 自律データ復旧

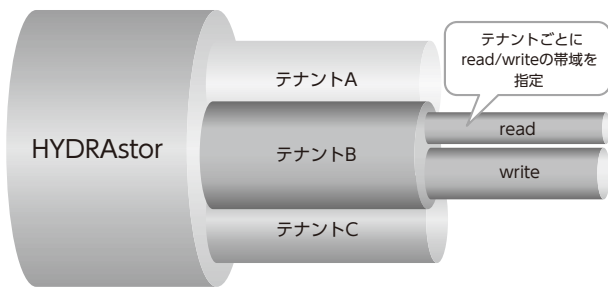


図4 帯域制限

上限帯域を指定し、その帯域を超えたアクセスを抑止することができる帯域制限機能をサポートします(図4)。

これにより、特定のテナントのユーザーが大量のアクセスを行った場合に、そのアクセス量を制限し、他のテナントのユーザーのアクセスに与える影響を抑えることができます。

4. おわりに

NEC Cloud IaaSのオブジェクトストレージサービスを実現するために、新たにオブジェクトストレージインタフェース及び帯域制限機能を実装したHYDRAsstorについて紹介しました。2015年度は更にAmazon S3互換性の向上などの機能強化を予定しています。HYDRAsstorは、今後もビッグデータ格納に最適なストレージ基盤として、進化を続けていきます。

* NetBackupは、米国およびその他の国におけるSymantec Corporationまたはその関連会社の商標または登録商標です。

* UNIXは、The Open Groupの米国およびその他の国における登録商標です。

* Linuxは、Linus Torvalds氏の米国およびその他の国における登録商標です。

* Windowsは、米国およびその他の国における米国Microsoft Corporationの登録商標です。

* Amazon Simple Storage Service、Amazon S3は、米国その他の諸国における、Amazon.com, Inc.またはその関連会社の商標です。

* OpenStackは、OpenStack Foundationの登録商標または商標です。

執筆者プロフィール

廣永 龍治

ITプラットフォーム事業部
ストレージ統括部
マネージャー

川名部 正純

ITプラットフォーム事業部
ストレージ統括部
マネージャー

関連URL

iStorageシリーズ製品情報

<http://www.istorage.jp>

NEC 技報のご案内

NEC 技報の論文をご覧くださいありがとうございます。
ご興味がありましたら、関連する他の論文もご覧ください。

NEC技報WEBサイトはこちら

NEC技報(日本語)

NEC Technical Journal(英語)

Vol.67 No.2 ICTシステムを担うこれからのクラウド基盤特集

ICTシステムを担うこれからのクラウド基盤特集よせて
NECのクラウド基盤への取り組み

◇ 特集論文

NEC C&Cクラウド基盤 NEC Cloud IaaSのサービス

マルチ環境統合を実現するポータルサービス
多用途環境に対応するハイブリッド型サーバサービス
多様なネットワーク環境を提供するネットワークサービス
内部統制手法を活用した堅牢なセキュリティサービス
クラウド基盤を支えるデータセンターサービス

NEC C&Cクラウド基盤を支える製品、最新技術

運用の自動化によりトータルコストを最適化する [WebSAM vDC Automation]
運用自動化により効率的な管理を実現する統合運用管理基盤
データセンターのTCO削減に貢献するマイクロモジュラーサーバ及び相変化冷却機構
クラウド環境に適した高信頼基盤を提供する iStorage M5000
データ保存に最適な、優れた圧縮効率と高速性を両立する iStorage HSシリーズ
大規模データセンターの管理自動化をサポートする SDN対応製品 UNIVERGE PFシリーズ
省電力を実現する相変化冷却技術・熱輸送技術

NEC C&Cクラウド基盤の将来技術

低コスト・省電力・低フットプリントを実現するアクセラレータ活用技術
スケールアップにより多種多様なコンピューティングを実現する Resource Disaggregated Platform
クラウド環境を対象にしたモデルベース設計支援技術
モデルベースでのサイジングと構成管理によりクラウド上のSIを効率化するクラウド型SI
ビッグデータ分析とクラウド ～異常を見抜くインバリアント分析技術～

導入事例

クラウドで遠隔監視保守システムの安定稼働を実現 全国約1,100基のタワーパークの安全を支える
ビジネスの中核を担うシステムを NEC Cloud IaaSへ移行 NECのトータルサポート力を評価
クラウド基盤サービスでグループのIT環境を共通化 ITガバナンスのさらなる強化を目指す

◇ NEC Information

C&Cユーザーフォーラム &iExpo2014

Orchestrating a brighter world 世界の想いを、未来へつなげる。

基調講演
展示会報告

NEWS

2014年度C&C賞表彰式典開催



Vol.67 No.2
(2015年3月)

特集TOP