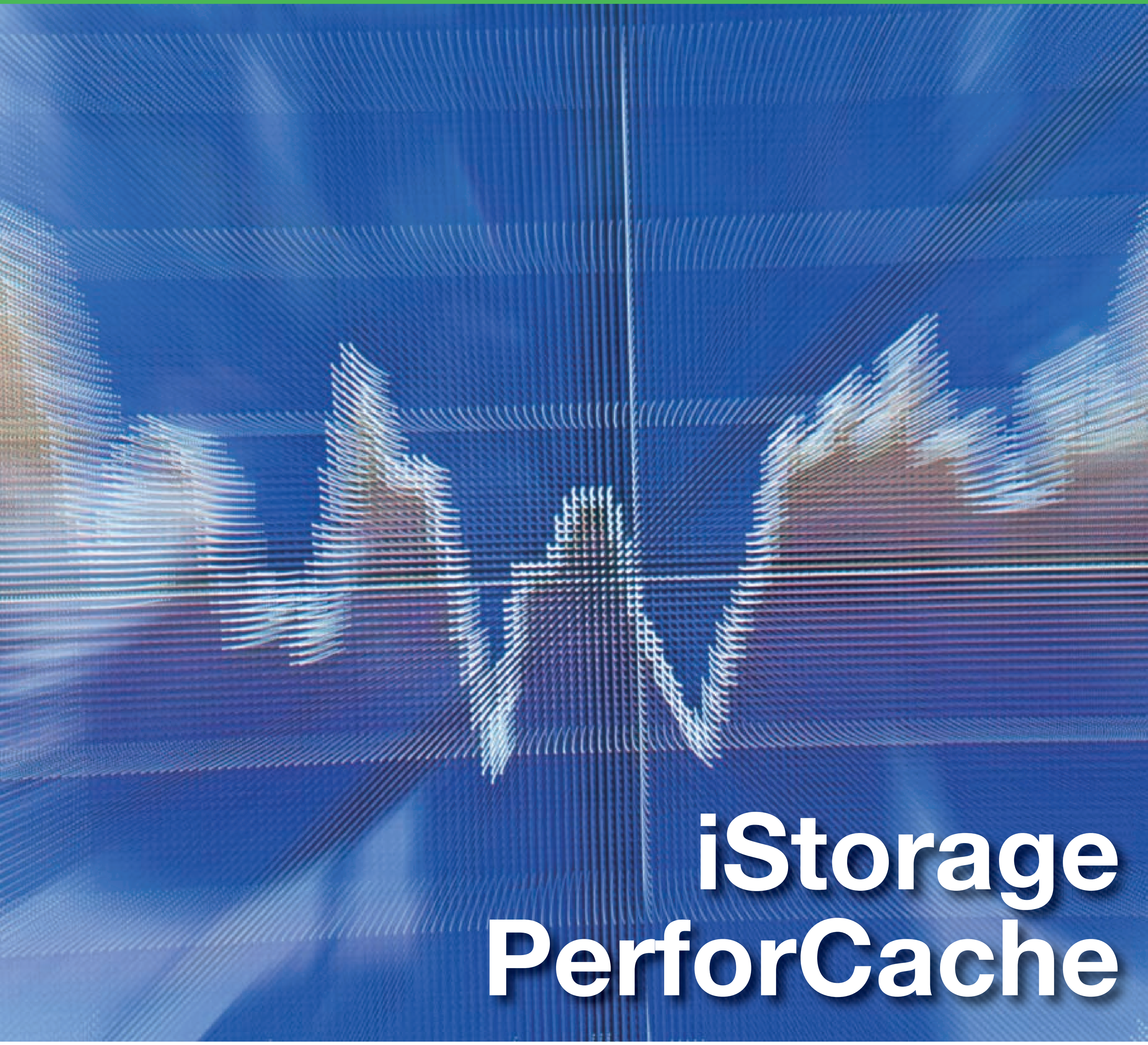
White Paper

Empowered by Innovation NEC

予測不可能なホットスポットを解消する SSD二次キャッシュ機能



iStorage

http://www.istorage.jp/

1. はじめに

「ビッグデータ」と呼ばれるデータ容量の巨大化は、「データをどこに保存し、利用するか」という従来の課題をよりいっそう複雑にしました。また、仮想環境の提供など、求められるサービスの内容にも変化が起きています。
こうしたシステム運用における現在の背景と課題について、次のようにまとめました。

1.1. SSD の登場

物理的な動作部分がないため消費電力が低く、ランダムアクセスに強いSSDは、ストレージ設計時における有力な選択肢となりました。
しかし、予測を超えたペースで年々等比級数的に増加するビッグデータの保存には、SAS、NL-SASといった容量あたりのコストの低い物理ディスクを視野に入れる必要があります。
I/O性能の高いSSD、コストパフォーマンスの高いSAS HDD（以降SASと表記）、ニアラインSAS HDD（以降NL-SASと表記）の中で、最適なものはどれか。SSDの登場は、ストレージ設計時における新たな難問を生み出しました。

1.2. 性能設計の複雑化

近年、システムは複雑化の一途をたどっています。ファイルサーバ、データベース、仮想化環境など、用途によってストレージに求められる性能は異なります。各物理ディスクにリソースをどのように分配するか、性能設計が必要です。
例えば、マルチテナント下で仮想化環境を提供するサービス形態を考えてみましょう。事前に各クライアントの利用状況が予測できない中で、適切な性能設計をすることは容易ではないでしょう。
さらに納期が短縮され、設計にかける経費が削減される傾向にある中で、最適な性能設計を行うことは困難を極めます。

1.3. システムの安定稼働の必要性

適切に設計され、導入されたシステムにおいても、ある時間帯にストレージへのアクセスが集中し、予測不可能な「ホットスポット」が発生してレスポンスの低下が起こることがあります。性能低下などの問題が発生すれば、その原因を特定するためには新たにエンジニアの人件費などのコストがかかります。さらにサーバの無停止運用が当然となりつつあり、対応は無停止で行う必要があります。
また、例えば仮想マシンを一斉に起動させる際の一時的な負荷の増加や、ビッグデータへの大量のデータ参照など、従来の想定を超えるような状況も出てきており、一次キャッシュの追加など従来の対応だけでは限界がきつつあります。

1.4. 求められる可用性の高まり

ストレージのコントローラは冗長化されており、障害発生時にも滞りなく業務をこなせるような設計になっています。
しかし、万が一の障害発生時には、データの保全性を確保するため、正常動作中のキャッシュに書き込むと同時に物理ディスクにも書き込みを行う「ライトスルー動作」に切り替わるため、運用は継続できるものの書き込み要求へのレスポンスが大幅に低下します。
また、サーバに対して24時間365日の稼働が求められている現在、メンテナンス時にもレスポンスの低下は許されません。
いまやメンテナンス時や障害発生時においても、運用継続はもちろん、通常時と同様のパフォーマンスが求められているのです。

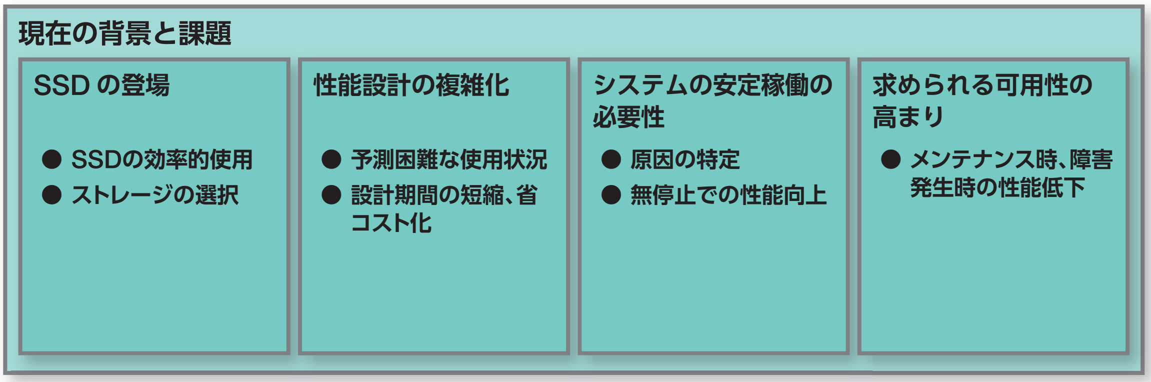


図 1：現在の背景と課題

© 2012 NEC Corporation　NEC and the NEC logo are registered trademarks of NEC Corporation.

2. NEC が提案する解決策

NECは、これらへの解決策としてSSD二次キャッシュ機能「iStorage PerforCache」とデータ最適配置機能「iStorage PerforOptimizer」をご用意いたしました。
そのうち、本稿では、「iStorage Mシリーズ」のSSD二次キャッシュ機能「iStorage PerforCache」について説明します。

2.1. SSD を効率的に利用

I/O性能の高いSSDを、一次キャッシュの働きを補完する「二次キャッシュ」として使用することで、すでに一次キャッシュを最大まで増設しているシステムにおいてもヒット率を向上させ、システム全体のパフォーマンスが向上します。
利用状況に応じて「リード専用」、「リード/ライト兼用」いずれか一方を選択可能なため、SSDを効率的に使用することができます。

2.2. 性能最適設計が不要

二次キャッシュは、すでに稼働しているシステムにも簡単操作で無停止で導入できます。ストレージとして使用しているSSDを転用すれば、導入コストの削減も可能です。
論理ディスク単位での二次キャッシュのオン･オフ切り替え、容量の追加などの操作も業務無停止で実行でき、性能設計時には予測できなかった性能低下にも対応します。

2.3. システムの安定稼働を実現

すでに一次キャッシュを最大まで増設しているシステムにおいても、二次キャッシュ追加によりヒット率を向上させ、ホットスポットの発生を抑制します。一度設定してしまえば二次キャッシュとして自動的に動作するため、通常運用での操作は不要です。
また、WebSAM Storage PerforMate Suiteを導入すれば、二次キャッシュの効果や使用状況を可視化して確認できます。

2.4. 業務の可用性向上

「パーシステントライト」は、ストレージのコントローラに障害が発生した場合に、二次キャッシュによりキャッシュの冗長性を確保する機能です。
この機能により、万が一の障害発生時にもライトスルー動作を避け、レスポンス低下を最小限に抑制することができます。

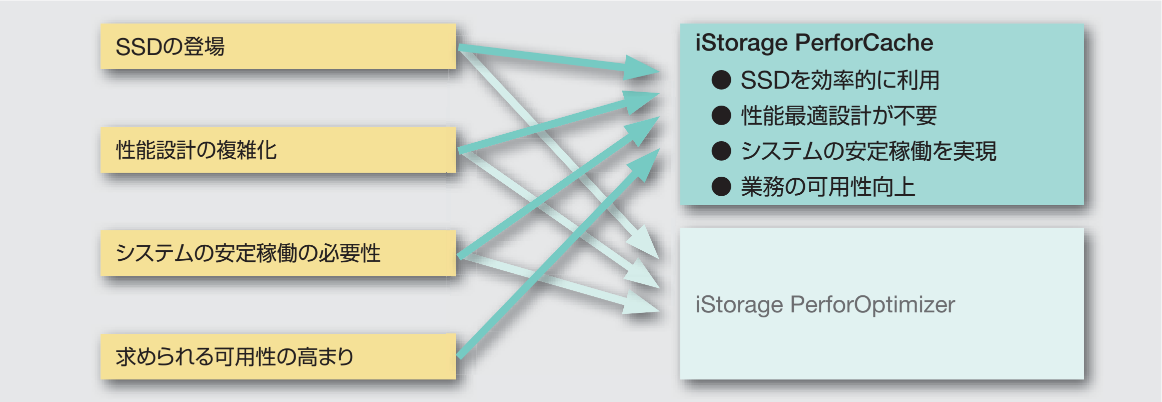


図 2：NEC が提案する解決策

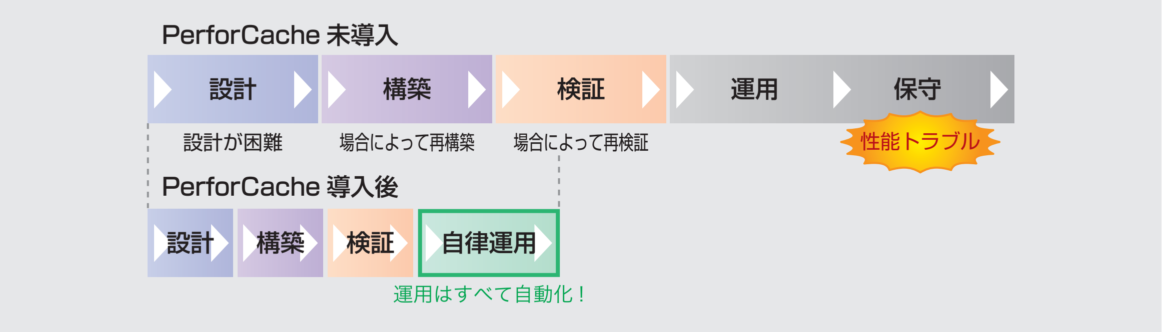


図 3：iStorage PerforCache の導入効果

© 2012 NEC Corporation　NEC and the NEC logo are registered trademarks of NEC Corporation.

3. iStorage PerforCache の機能

iStorage PerforCacheには、リードキャッシュ、ライトキャッシュ、パーシステントライトの3つの機能があります。「リード専用」、「リード/ライト兼用」の2種類が設定できます。また、障害発生時には自動的にパーシステントライト機能が働きます。

3.1. リードキャッシュ機能

図4はリードキャッシュ機能の概念図です。導入後には、一次キャッシュでヒットしなかった参照要求を二次キャッシュでヒットさせ、全体でのヒット率を向上させます。
結果としてHDDへのアクセスが減少し、ホットスポットの発生を抑制します。

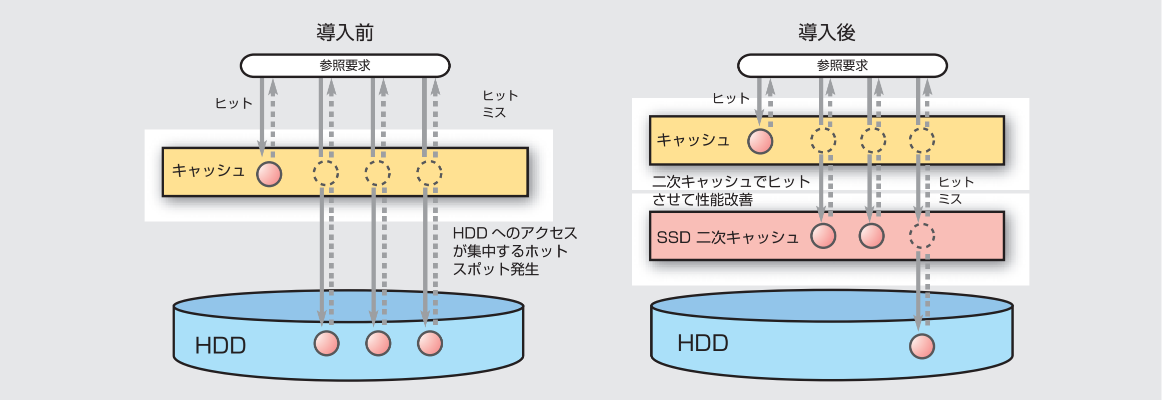


図 4：リードキャッシュ機能の概念図

3.2. ライトキャッシュ機能

図5はライト機能の概念図です。導入前では、一次キャッシュでヒットした更新要求はキャッシュ上で処理され、一定時間経過後にHDDに書き込まれます。ヒットしなかった場合はHDD上のデータが直接更新されるため、ホットスポットが発生しやすい状況です。
二次キャッシュ導入後では、頻繁に更新されるデータを二次キャッシュに乗せてヒット率を向上させます。また、二次キャッシュ上で更新されたデータも遅延書き込みすることによりHDDへのアクセスを減少させ、ホットスポットの発生を抑制します。

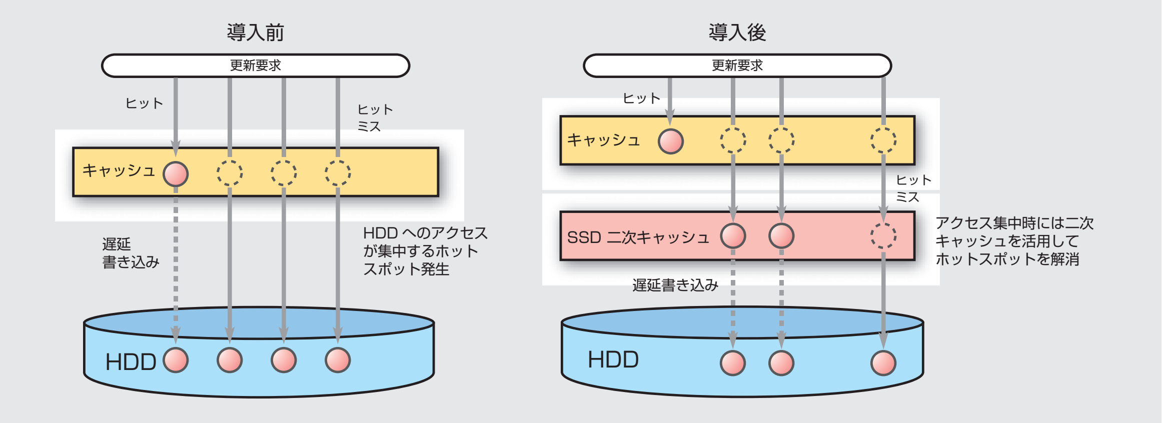


図 5：ライトキャッシュ機能の概念図

3.3. パーシステントライト機能

万一のコントローラ障害または保守交換などで一時的にコントローラの冗長性が失われた場合には、二次キャッシュが一次キャッシュの代替として働き、性能低下を最小限に抑えます。

図6はパーシステントライト機能の概念図です。導入前では一次キャッシュに障害が発生した場合にキャッシュにデータを書き込むと同時に物理ディスクにもデータを書き込むライトスルー動作により大きな性能低下が発生するのに対して、導入後は書き込み速度の速いSSDがキャッシュの代替機能を果たすことにより性能低下を最小限に抑えます。

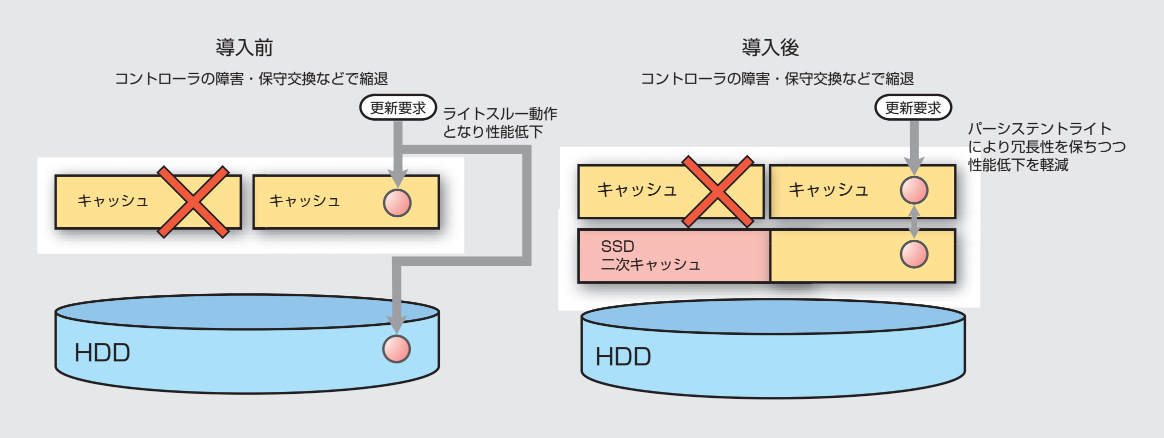


図 6：パーシステントライト機能の概念図

© 2012 NEC Corporation NEC and the NEC logo are registered trademarks of NEC Corporation.

4. iStorage PerforCache の導入

iStorage PerforCacheは、稼働中のシステムにも無停止で簡単に導入できます。また、導入後の二次キャッシュの管理は自律的に行われるため、通常運用では操作･設定変更とも不要。運用コストを削減できます。

効果を確認したい場合には、オプション製品の性能監視ソフトウェアWebSAM Storage PerforMate Suiteを導入し、性能情報を可視化して分析することも可能です。各員結果を踏まえて二次キャッシュの容量拡張、キャッシュ種別の設定変更などを行えば、二次キャッシュをさらに効率的に運用できるでしょう。

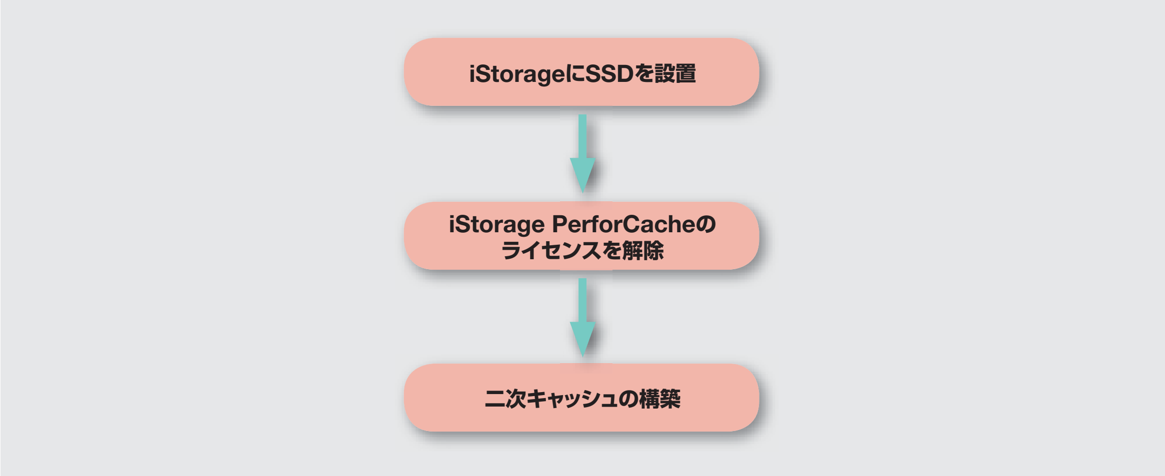


図 7：iStorage PerforCache の導入フロー

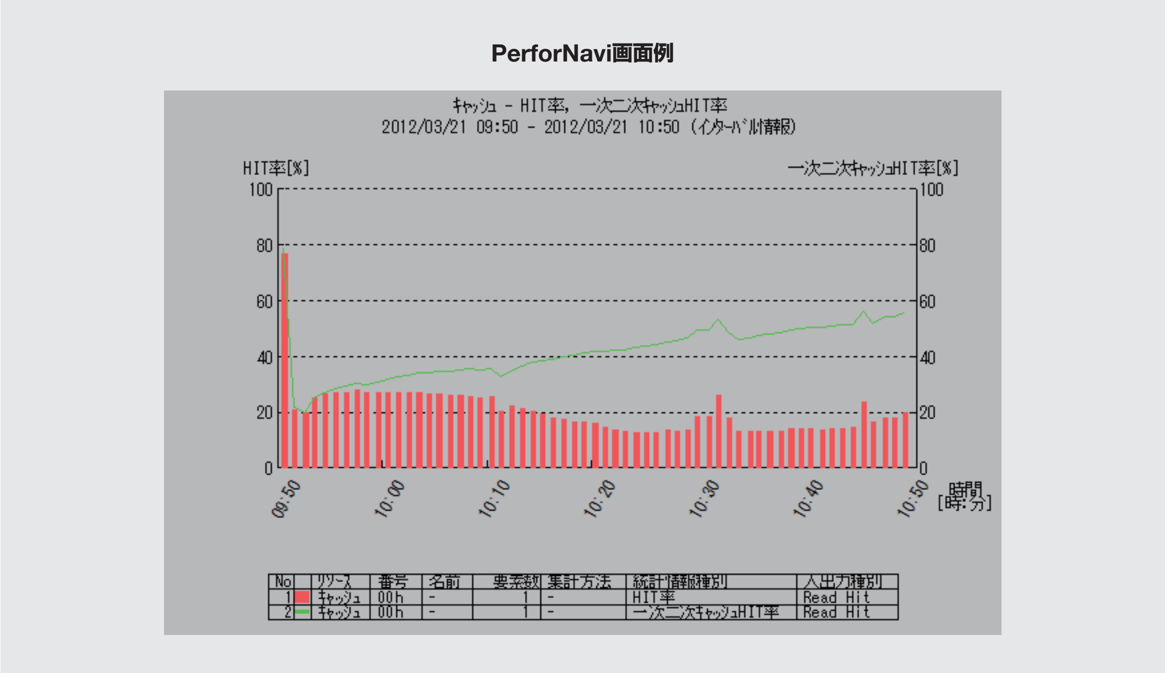


図 8：性能情報の可視化の例

5. 検証

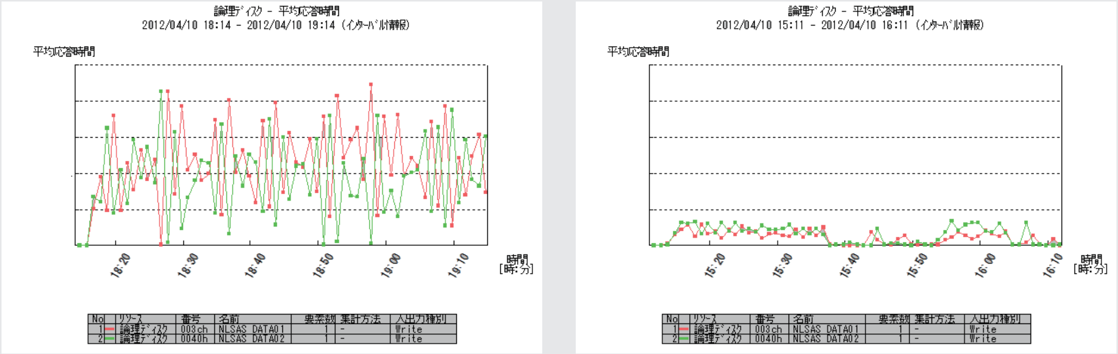
では、二次キャッシュを導入した場合、実際にどのような効果があるのかを、具体的な数値を挙げて説明します。

5.1. ストレージの動作の検証

● 平均応答時間の変化

図9は導入前後での論理ディスクの平均応答時間を比較したグラフです。導入前にはキャッシュのヒット率が低く数値がばらついていますが、導入後は平準化されて低下しており、ホットスポットが解消されていることがわかります。

<導入前>　　**<導入後>**



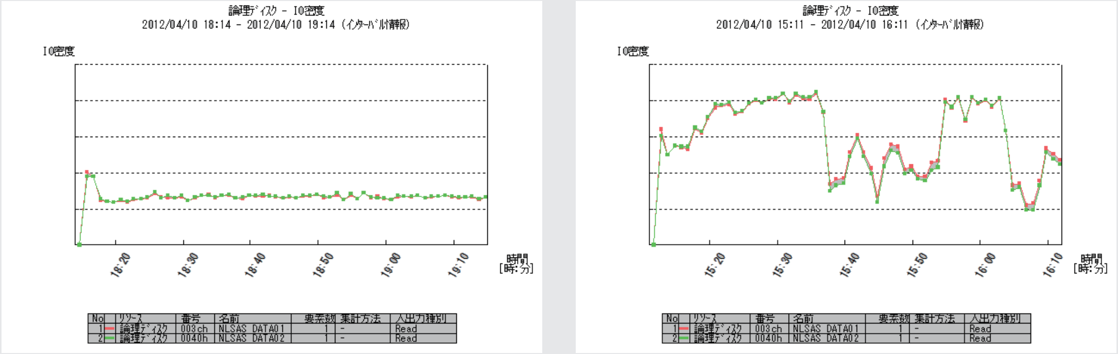
測定条件: iStorage M300/RAID6(NL-SAS 6台)×2/
一次キャッシュ 6GB×2/二次キャッシュ SSD 100GB×2
測定条件: 銀行業務を想定した負荷を一時間かけたときのトランザクション数をOLTPベンチマークツール(TPC)で測定

図 9：平均応答時間

● I/O密度の変化

図10は導入前後でのI/O密度を比較したグラフです。導入前にはキャッシュのヒット率が低くI/O密度は平均して低くなっていますが、導入後は要求に応じてI/O密度が上昇し、レスポンスが向上していることがわかります。

<導入前>　　**<導入後>**



測定条件: iStorage M300/RAID6(NL-SAS 6台)×2/
一次キャッシュ 6GB×2/二次キャッシュ SSD 100GB×2
測定条件: 銀行業務を想定した負荷を一時間かけたときのトランザクション数をOLTPベンチマークツール(TPC)で測定

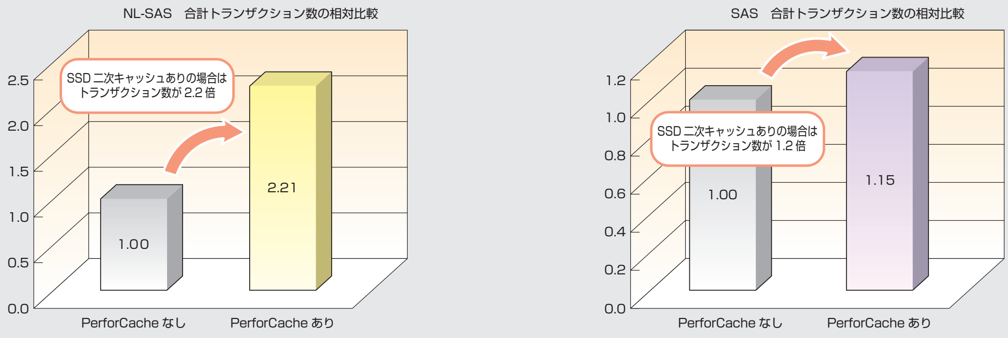
図 10：I/O 密度

© 2012 NEC Corporation　　NEC and the NEC logo are registered trademarks of NEC Corporation.

5.2. 実務環境における検証

● トランザクション数の変化

図11はOracle Database 11g環境におけるSSD二次キャッシュ導入前後でのトランザクション数を相対比較したグラフです。NL-SAS、SASのどちらにおいても、合計トランザクション数が増加していることがわかります。



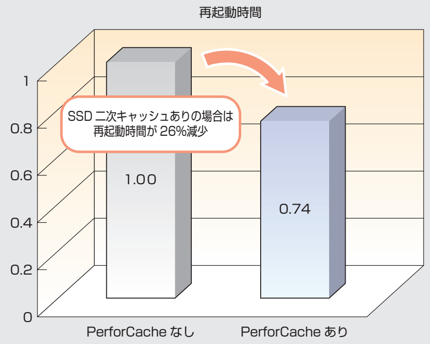
測定環境: iStorage M300/RAID6(NL-SAS 6台)×2/RAID10(SAS 6台)×2/
一次キャッシュ 6GB×2/二次キャッシュ SSD 100GB×2(リード/ライト兼用)

測定条件: 銀行業務を想定した負荷を一時間かけたときのトランザクション数をOLTPベンチマークツール(TPC)で測定

図 11：トランザクション数

● 仮想マシン起動時間の変化

図12はVMware環境でのゲストOS同時再起動までの時間を、SSD二次キャッシュ導入前後で比較したグラフです。導入後は起動までの時間が26%削減されていることがわかります。



測定環境: iStorage M300/RAID6(NL-SAS 5台)×2/
一次キャッシュ 8GB×2/二次キャッシュ SSD 400GB×2(リード/ライト兼用)

測定条件: ゲストOS 32台(実マシン4台)を同時再起動

図 12：仮想マシン起動時間

お問い合わせは、下記へ

NEC プラットフォーム販売本部(ストレージお問い合わせ)

〒108-8424 東京都港区芝五丁目33番8号(第一田町ビル)
TEL: 03(3798)9740

問い合わせURL　http://www.nec.co.jp/contact
国内向け製品URL　http://www.istorage.jp

●本ホワイトペーパー内の社名、商品名は各社の登録商標または商標です。
●製品の仕様は、都合により予告なしに変更することがあります。

日本電気株式会社　〒108-8001 東京都港区芝五丁目 7-1 (NEC 本社ビル)　2012年12月現在　Cat.No. E06-12120718J