

今すぐできる！ データベース高速化

I/O・ストレージの改善で業務スピードUP！



データベース性能のボトルネックとなりがちなI/Oとストレージ。
企業が今すぐに取り組むことができる、
データベース高速化の方法を紹介する。

P.2 バッチ処理が終わらない
— 課題解決のカギを握る「DB高速化」という手段

P.5 既存アプリは触りたくない！
DB高速化の本丸「Oracle RAC on SSD」を徹底解剖する

バッチ処理が終わらない

——課題解決のカギを握る「DB高速化」という手段

企業を取り巻くデータ量が爆発的に増大する中、処理時間を短縮するために必要なのがデータベースシステムの高速化だ。中でも、処理速度のボトルネックになりがちなI/Oの見直しは急務と言える。本稿では、企業が今すぐ取り組めるDB高速化の方法を紹介しよう。

企業システムで扱うデータ量が増え続ける今、IT部門には大きな課題が突き付けられている。サーバの高性能化は著しく、仮想化技術を利用して難なく複数の業務システムで共用できるまでになった。一方、処理するデータを格納するデータベース(DB)システムは、サーバの処理能力に追いついていないのが現状だ。そのため、データ量が増えれば増えるほど処理時間がかかり、「夜間バッチ」が朝までに終わらないなどの事態も発生している。

また、こうした状況では、経営層や業務部門からリアルタイムのデータ分析が求められても、実現するのは難しい。そこで急がれるのが、データ分析処理のボトルネックになっているDBシステムの高速化だ。

ボトルネックの最多原因はI/Oとストレージの性能不足

DBがボトルネックになる原因はいくつかあるが、大別するとソフトウェアが原因の場合と、ハードウェアが原因の場合に分けられる。

このうちソフトウェア側の原因としては、アプリケーションやDBの構造、SQL文、パラメータなどに問題がある

ケースが挙げられる。これを解決するには、アプリケーションやDBを見直す必要がある。具体的には、DBの構造やSQL文から性能のボトルネックになっている部分を見つけ出し、部分的に改修するか、場合によっては全面的にアプリケーションやDBを作り直すことになる。

だが現実的には、エンジニアが不足していて社内のIT部門だけで対応することが難しかったり、膨大な工数・コストがかかったりするなど、そのハードルは極めて高い。したがって、業務システムそのものを完全にリプレイスする時期などでないと取り組みづらいことが多いだろう。なお、場合によってはパラメータをチューニングするだけで高速化できることもあるが、目を見張るような効果が得られることはごくまれだ。

一方、ハードウェア側の問題としては、主にハードウェア処理能力が不足しているケースが挙げられる。

原因がCPU性能やメモリ容量の不足なら、最新CPUと大容量メモリを備えたサーバに入れ替えれば、業務システムやDBシステムに手を付けなくてもDB性能は改善される。だが実際には、DBシステムは、ピーク時の性能を想定してCPU性能やメモリ容

量に余裕を持って設計されている場合が多いため、それらがボトルネックになっているケースはそう多くない。

ハードウェア側の問題で最も多いのは、I/O帯域やストレージ性能の不足だ。DBシステムではストレージ容量は気にしても、ストレージI/O性能の設計は忘れがちである。多少気にしていたとしても、どのように設計すればいいかはっきりとした指針もないため、ログはRAID1の領域に“とりあえず”配置しておく程度の対応しかされていないのが現状である。

大量のデータを扱うDBシステムでは、HDDへの読み書きが頻繁に発生する。I/O帯域やストレージ性能が不足すると、この部分に遅延が発生するため、いくらCPUやメモリ容量があっても十分に活用できず、結果的にDBの性能が出ないのだ。

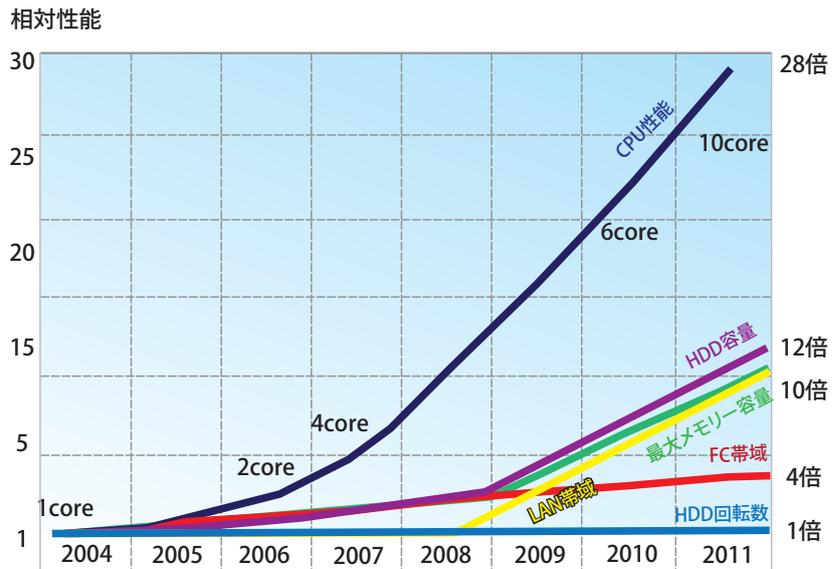
しかし、なぜボトルネックの原因がI/O帯域やストレージにある場合が多いのか。それは前述した考慮不足もあるが、CPUやメモリなどと比べ、I/Oやストレージの性能進化のペースが遅いからだ。(図1)

例えばCPUの場合、約10年前のサーバではシングルコアのプロセッサしか存在しなかったが、現在では6コア～10コアなどが当たり前であり、そ

の相対性能は約28倍にも及んでいる。それに対しI/Oでは、ファイバチャネル(FC)の帯域は1~2Gbpsから4~8Gbpsへと約4倍高速化したにすぎず、HDDの回転数に至っては最大1万5000と、全く向上していない。この相対性能の差から、DB性能のボトルネックが発生している可能性が高いのである。

この問題を改善する方法は、基本的にはハードウェアの入れ替えだ。つまり、I/Oインタフェースやストレージをより高速なものに変更するのだ。例えば、ストレージ装置を増設したり、HDDをSSDに換装したりすれば、I/Oとストレージが原因のボトルネックを解消できる可能性がある。

既存の業務システムやDBシステムにできるだけ手を加えずに、DB性能のボトルネックを解消するためには、これが最も簡単な方法と言えるだろう。(図2)



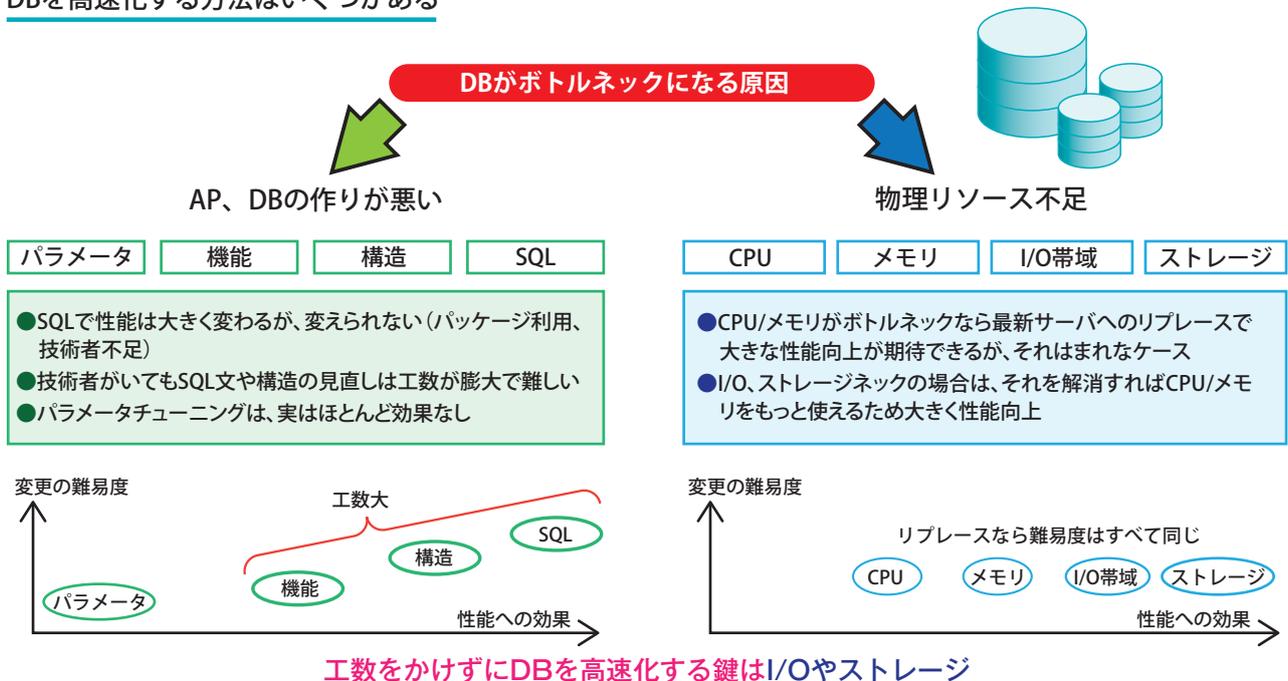
- CPUの性能の伸びに比べ、I/Oはそれ程伸びていない
- 特にストレージ性能に寄与するHDD回転数とFC帯域は最悪

➡ **ストレージ I/O の高速化を行うことにより、性能は劇的に向上する可能性がある**

※ CPU性能はSAP SDベンチマークをベースにした値

◆図1 ハードウェアの性能進化の不均衡 (出典：日立製作所)

DBを高速化する方法はいくつかある



◆図2 DBの高速化を実現する方法 (出典：日立製作所)

インメモリから分散まで、 多様なソリューションが登場

一方、DBの高速化に根本から取り組むなら、業務システムの更改時にDBシステムも合わせて刷新するのが一番だ。最近ではDB高速化の要求に応えるさまざまなソリューションが登場しているので、その一部を紹介しよう。

既存のDBシステムに近い構成が可能なのが、DB専用アプライアンス製品の導入だ。DBアプライアンスは、サーバとストレージを一体化させたものであり、構成の自由度が制限される代わりに、サーバとI/O、ストレージとの性能のアンバランスが是正されている。DBとストレージの管理を一元化したり、運用を自動化する機能が搭載されていることも利点だろう。

一方、DBシステム自体を性能重視の製品に移行するなら、インメモリDBという選択肢がある。インメモリDBとはデータストレージをメモリ上に置くもので、HDDをストレージとして利用

する一般的なDBと比較して、非常に高速かつ安定した性能を発揮する。

ただし、揮発するメモリ内にデータストレージを格納するため、可用性の面で課題が残る。また、HDDと比較するとシステムに搭載可能なメモリ容量ははるかに少なく全てのデータを入れられなかったり、入ったとしても大容量メモリを搭載することでコストが非常に高くつくという欠点もある。

データ検索の高速化を実現する別のアプローチとしては、分散DBが注目されている。その代表格と言えるのが、ビッグデータのリアルタイム処理などに活用されるようになった「Hadoop」だ。Hadoopは、独自のファイルシステム上で動作する「HBase」という分散DBの仕組みを利用し、データを分割配置してバッチ処理を並列実行することで、超巨大テーブルを高速に処理できる。

このほか、従来のDBの延長ではなく全く新しいDBを開発する取り組みも始まっている。例えば、日立製作所が製品化したデータアクセス基盤「Hitachi Advanced Data Binder

プラットフォーム」^{※1}は、「非順序型実行原理」^{※2}という仕組みを取り入れたエンジンに、サーバとSSD搭載ストレージを組み合わせることで、標準のSQLに対応しながら、従来のDBと比べて劇的な高速化を実現するという。

直近の課題を解決する 現実解は？

ここまで、DBを高速化するためのさまざまな方法を紹介したが、「今動いている業務システムのDBを高速化したい」「バッチ処理を朝までに終わりにしたい」といった直近の課題を、できる限り容易に解決したいのであれば、やはりI/Oとストレージの見直しが最も有効な策だろう。

実は、このI/Oとストレージの見直しという分野は徐々に熱い市場になりつつある。次ページからの記事ではその代表的な製品の1つとして、日立の「DB高速化ソリューション for Oracle RAC on SSD」を紹介しよう。

※1 内閣府の最先端研究開発支援プログラム「超巨大データベース時代に向けた最高速データベースエンジンの開発と当該エンジンを核とする戦略的サービスの実証・評価」(中心研究者は東京大学の喜連川優教授)の成果を利用して開発

※2 東京大学の喜連川優教授および合田和生特任准教授が考案した原理

スループットは HDD の 10 倍以上

既存アプリは触りたくない！ DB高速化の本丸「Oracle RAC on SSD」を徹底解剖する

データベース性能のボトルネックとなりがちなI/Oとストレージ。ここに焦点を当てて日立製作所が投入したのが「DB高速化ソリューション for Oracle RAC on SSD」だ。本ソリューションのもたらす威力について、日立でPFコンピテンスセンタ長を務める鬼頭昭氏に聞く

DB高速化の秘けつはI/Oと ストレージの改善にあり

企業を取り巻くデータ量の増加に伴い、データベース(DB)の処理時間は年々長くなりつつある。こうした流れが加速すれば、例えば夜間バッチが朝までに終わらないなどの事態が一層増えるはずだ。だが「DBを高速化したくても、既存の業務システムに手を加える余裕はない」というのが、多くの企業がDBに関して頭を悩ませている課題だろう。

「DBの高速化が必要な課題はさまざまですが、ベースにあるのはやはりバッチ処理を高速化したいというものです」と話すのは、日立製作所(以下、日立)の鬼頭昭氏。

「処理するデータを増やせば増やすほど、DBの負荷が高まりバッチ処理に時間がかかるという課題を抱える企業は少なくありません。それを解決するためには、データの見直しを通じて完全に新しいシステムに作り替えるという方法もありますが、多くの企業はシステムを丸ごと刷新することに抵抗があります。実際には、性能さえ問題がなければ仮想サーバ環境上に既存のシステムをそのまま移したいなどのニーズが多いのです」(鬼頭氏)

現在、DB高速化を実現するためにはさまざまな手段がある。その中で、比較的容易にDB高速化を実現できる手段として注目されているのがI/Oとストレージの見直しだ。そこに真っ向から取り組んだソリューションの1つが、日立が投入した「DB高速化ソリューション for Oracle RAC on SSD」である。

HDDの10倍以上、毎秒11G バイトのI/O処理を可能に

「I/Oとストレージを高速化するためにはHDDをSSDに換装すればよいと思われがちですが、実際にはSSDに変えただけでは大して速くなりません。劇的な高速化を実現するには、DBを稼働させた際にどれだけ多くのI/O処理を並列で実行できるか、あるいは単位時間当たりどれだけデータを読み込めるかなど、ハードウェア全体のバランスを調整する必要があります」と鬼頭氏は話す。

HDDに変わるストレージとして、不揮発性メモリを採用したSSDがエンタープライズシステムでも本格的に導入されつつある。SSDはHDDと異なり、読み書き性能が回転数によって制限されることはないが、SSDの能力を生かすにはI/Oの帯域も調整する



日立製作所
情報・通信システム社
ITプラットフォーム事業本部 開発統括本部
ソフトウェア本部
PFコンピテンスセンタ センタ長 鬼頭昭氏

必要がある。例えば、性能に応じてFC(ファイバチャネル)接続バス数を変えたり、高性能なストレージコントローラを採用するなど、I/Oの高速化に最適な構成とチューニングが求められる。

そこで日立が取り組んだのが、サーバとストレージをFCで接続する従来型DBシステムの構成は変えずにSSDを採用する、新しいDB高速化ソリューションの開発である。

「DBを高速化したい場合、データ処理を実行するサーバのCPUコア数を増やす方法もありますが、それだけではバッチなどI/Oがボトルネッ

クとなる処理は高速化できません。そのため、DB高速化ソリューション for Oracle RAC on SSDは、ストレージとI/Oインタフェースの処理能力アップを通じてI/Oのボトルネックを解消した上で、複数のサーバをクラスタ構成で運用する『Oracle Real Application Clusters』(Oracle RAC)で、DBシステム全体の性能を向上させました」

「ストレージは、PCIeスロットにフラッシュメモ리카ードを搭載する選択肢もありましたが、フラッシュメモ리카ードは障害発生時に稼働中の交換ができない問題があるため、可用性を高めるためにSSDを用いたRAID構成を採用しました。ただし当初は、SSDを採用しても期待ほどの性能が出なかったため、すぐにI/O帯域を広げて対処しました。DBへのアクセスが全てのSSDに均等に配分されるよう、Oracle Databaseのストレージ管理機能『Oracle Automatic Storage Management』(Oracle ASM)も組み合わせています」(鬼頭氏)

こうして完成したDB高速化ソリューション for Oracle RAC on SSDは、

Intel Xeonプロセッサと最大768Gバイトのメモリを搭載できるサーバ「BladeSymphony BS2000」2ブレード、SSDの搭載を前提にI/O処理能力を拡張した「Hitachi Unified Storage 130」(HUS130)2台、ブレードサーバにPCIeスロットを追加する日立独自の「I/Oスロット拡張装置」で構成されている。

これらを最適なチューニングと物理的な配置で組み合わせさせた結果、データ読み込み時には毎秒11GバイトものI/O処理を確認できたという。また、日立が検証用に用意したI/Oの多いSQLでベンチマークを計測したところ、HDDを潤沢に搭載した同等のDBシステムと比較して、10倍以上のI/Oスループットと約5分の1のレスポンスタイムを実現したという。さらに、同社が想定する6~7年前のDBシステムと比べれば、レスポンスタイムは約100分の1以下まで短くなったとしている。(図1)

独自の工夫で高速化を とことん追求

DB高速化ソリューション for Ora

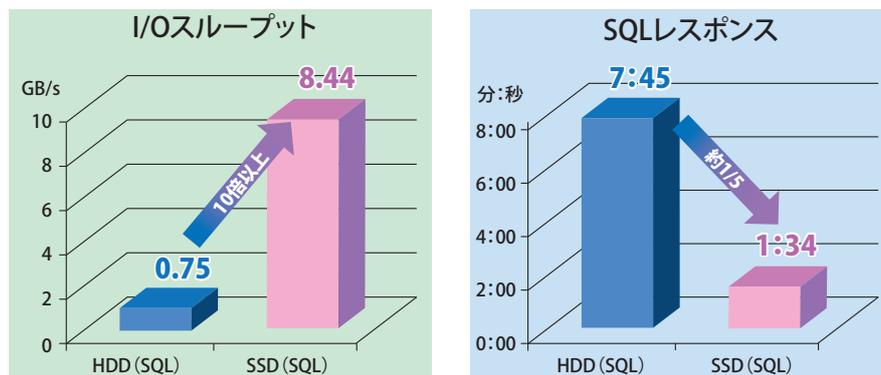
cle RAC on SSDがDBを高速化できる理由は何か。それは、ハードウェアの性能を追求し、その性能をフルに生かすさまざまなチューニングを施しているからだ。

チューニングの1つ目は、サーバからストレージへのI/O帯域の拡張だ。BS2000に接続されているI/Oスロット拡張装置は、一般的なブレードサーバが搭載しないPCI-Expressスロットを備え、FCスイッチを介さずにホストバスアダプタ(HBA)をストレージに直結できるようにする。また、BS2000は2枚のサーバブレードを搭載し、1つのブレード当たり8ポートの8Gbps FC、合計16ポートからストレージに接続する。

「ストレージI/Oの広帯域化を目指す、どうしてもある程度のFCポート数が必要です。しかし、ブレード内蔵メザニンカードの搭載枚数には限りがあるため、FCの採用を優先すると、サーバ間通信LANやパブリックLANなどの冗長化ができなくなってしまいます。そこでDB高速化ソリューション for Oracle RAC on SSDは、I/Oスロット拡張装置を使い、HBAを全てPCIeスロットに接続する方法を採用しました」と鬼頭氏は話す。(図2)

2つ目のチューニングは、サーバからSSDまでの経路を均等化し、分散I/Oの処理能力を高めていることだ。

ストレージ(HUS130)には、高い読み書き性能を持つコントローラと専用のSSDが搭載されており、これらの能力を引き出すためのチューニングが施されている。RAIDから切り出した論理ユニットと制御プロセッサを適切に割り当てたほか、4本のバックエンドパスを均等に使えるように物理的な位置も調整した。また、マルチパスFCの優先パスを効率的に分散する設定も施



◆図1 HDDを採用したシステムと比較し、10倍以上のI/Oスループットと約5分の1のレスポンスタイムを実現した(出典:日立製作所)

している。これらの構成やチューニングは全て日立のノウハウであり、従来の製品にSSDを搭載しただけの構成とは「ひと味もふた味も異なる」と鬼頭氏は胸を張る。

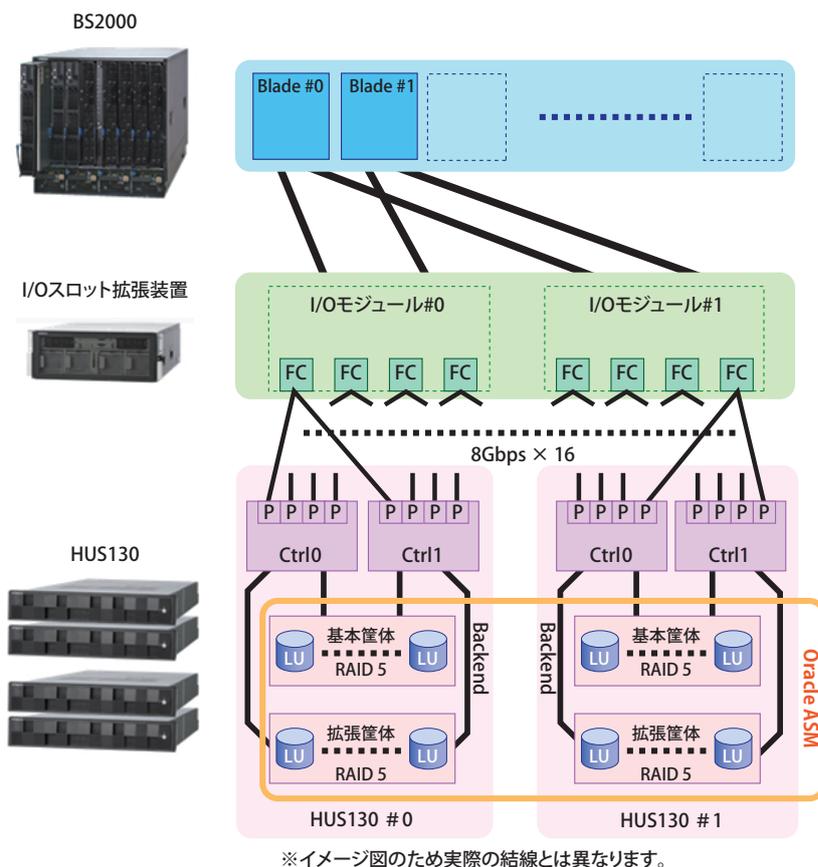
とはいえ、ハードウェアそのものは特殊なものではなく、オープンなプラットフォームとなっている。そのため、業務システムからは一般的なDBシステムに見え、それまで通りの運用方法を継承できる。なお、サーバのOSには「Red Hat Enterprise Linux」を採用し、UNIXシステムから比較的容易にリプレースできるようになっているほか、多くのミドルウェアの動作が保証されていることも心強い点と言えるだろう。

このほか、障害耐性や安定性の高さもDB高速化ソリューション for Oracle RAC on SSDの特長という。

ストレージI/Oは全てActive-Activeの冗長化構成を採用している。日立が行ったテストでは、PCIeバスの障害閉塞機能とFCのバス制御で、抜線時もDBが継続稼働するという実証結果を確認したという。

HUS130自体も、信頼性に関して日立独自の作り込みがなされている。SSDはHDDと違って書き込み回数の上限があるが、現在ではSSDの全領域を1日10回書き換えても約5年間使えるほど長寿命化している。さらにHUS130は、SSD寿命の90%（と95～98%の1%刻み）に達するとユーザーに通知を出す機能や、99%になるとスペアドライブにデータを自動複製する機能も備えている。

DB高速化ソリューション for Oracle RAC on SSDの導入が最も威力を発揮するケースは、大量データを扱うI/Oの多いバッチ処理を実行する



◆図2 DB高速化ソリューション for Oracle RAC on SSD (ミッドレンジモデル) の構成図 (出典:日立製作所)

場合だ。処理時間が短縮されるだけでなく、同時に複数のバッチを処理することも可能になる。このほか、データウェアハウスやビジネスインテリジェンスにおけるデータ抽出処理を高速化したい場合にも有効と考えられる。

サーバのCPU、メモリ容量、SSD容量は、処理内容に応じてカスタマイズできる。また、中小規模システム向けには、サーバに「HA8000/RS220」、ストレージに「BR1200」を採用した廉

価版の構成も用意している。こちらでは、最大で毎秒3.2GバイトのI/O処理を行えるという。

DBを飛躍的に高速化させたいが、既存アプリケーションには手を加えたくない——DB高速化ソリューション for Oracle RAC on SSDは、システム部門のそうした悩みに対して有効な解決策の1つと言えるだろう。

● BladeSymphony BS2000について

URL <http://www.hitachi.co.jp/products/bladesymphony/product/bs2000.html>

● Hitachi Unified Storage(HUS)について

URL <http://www.hitachi.co.jp/products/it/storage-solutions/products/hus/index.html>

※この冊子は、ITmediaエンタープライズ(<http://www.itmedia.co.jp/enterprise/>)に2013年3月に掲載されたコンテンツを再構成したものです。

URL <http://www.itmedia.co.jp/enterprise/articles/1303/18/news003.html>

HITACHI
Inspire the Next

◆ 記事に関するお問い合わせ

<http://www.hitachi.co.jp/server/solution/inq/>

株式会社 日立製作所