

間もなくIBMから登場予定のスーパーコンピュータは、アニメキャラクターと大型スポーツカーにあやかり「Roadrunner」と名付けられた。ストレージには高パフォーマンスを誇るDirectFlowを採用している。

現在、米IBMがニューメキシコ州ロスアラモスにある米国エネルギー省「Los Alamos National Laboratory」に納めるため製造中の「Roadrunner」は、1秒当たりの浮動小数点演算回数が1000兆回(1ペタフロップ)という、未曾有の性能を誇るマシンになる予定だ(関連記事)。

これがどれほどのパフォーマンスかがピンと来ない読者も多いと思うが、そのすさまじい処理能力は確実に請け合える。

エネルギー省は、米国の備蓄核兵器を安全・確実に保持し、地下核実験が再開されることのないよう核兵器保全管理計画を進めている。Roadrunnerは、そのような目的の研究に大きく貢献する次世代スーパーコンピュータだ。

Roadrunnerは、業務用に販売されているハードウェアのみを使って作られており、オペレーティングシステムにはRed Hatの「RHEL (Red Hat Enterprise Linux) 4.3」を採用している。具体的には、AMDの「Opteron」プロセッサを搭載したIBMの「System x3755」システムと、新しい強力な「Cell」チップを積んだ「BladeCenter H」システムが併用されるという。Cellチップは、もともとは高性能なビデオゲーム用に開発されたものだ。Roadrunnerは、極限レベルのコンピュータ処理にハイブリッドアーキテクチャが使用される、新たなコンピューティング時代を拓くとして、大いに期待されている。

Roadrunnerが完成したあかつきには、世界最速のコンピュータになるといっても過言ではない。

もっとも、プロセッサがより高速になり、アーキテクチャおよびI/Oは効率性を増し、ソフトウェアの性能も向上したとはいえ、重大な問題も残っている。消防ホースをいっぱい流れるほどの計算データストリームを、システムが処理できるようになるのはよいが、それだけのデータ読み書きを行うためには、どれほど大容量なストレージを構築すればよいのだろうか。これまでもストレージのI/Oは、スーパーコンピューティングにとって最大かつ最悪のボトルネックとなってきたのである。

課題を解決する「DirectFlow」とは？

だがこれからは、そうした問題に頭を悩ませなくてよさそうだ。Roadrunner開発チームは、8年前に設立され、今では125名の社員を抱えるカリフォルニア州フレモントのPanasasと協力し、同社の「ActiveScale 3.0 Storage Cluster」を新たなペタ規模スーパーコンピュータのストレージシステムに採用して、積年の課題を解決しようとしている。

Roadrunnerは、Linuxと「DirectFlow」機能を備えるPanasas Storage Clusterを活用し、きわめて複雑な科学演算を実行していくことになる。DirectFlowは、「PNFS (Parallel Network File System)」と呼ばれる完全並列型のデータパスを実現し、RoadrunnerチームのLinuxクラスタとPanasasストレージクラスタノード間における、高速かつ直接的な通信を可能にする機能だ。

従来のストレージシステムは、1つの双方向ヘッドコントローラを使って、直接的なデータ転送を行っていた。Panasasの創設者で、RAIDストレージの生みの親として世界的に知られるガス・ギブソン氏は、PNFSを実装しているのは唯一Panasasだけだと、当初から喧伝してきた。PNFSは、2つの双方向ヘッドコントローラを持っている。これはすなわち、初めから走っていた2車線の高速道路の隣に、さらにもう1つ同じものを増設したようなものだ。

The 451 Groupのアナリストであるヘンリー・バルタザール氏はeWEEKに対し、「PNFSはデータパスからのメタデータアクセスを隔離し、クライアントがNAS(Network-Attached Storage)に直接的かつ並列的にアクセスできるようにする。

標準的なSAN(Storage Area Network)もしくはNASストレージでは、特に高パフォーマンスコンピューティング(High-Performance Computing:HPC)などのパフォーマンスを重視する環境下で、ヘッドコントローラがボトルネックと化してしまう。クラスタ化ストレージシステムの主なメリットは、データロードを複数のシステムに振り分け、高速データアクセスを維持できる点である」と語った。

一方、Enterprise Management Associatesでアナリストを務めるマイク・カープ氏も、「新たなRoadrunnerエコシステムの最重要要素」はまさにこの並列ファイルシステムだと、eWEEKに話している。

「ロスアラモスの研究所が扱っているようなタイプの計算は、並列コンピューティングプロセスを必要とする複雑なレベルで処理されている。これはつまり、さまざまなCPUへデータをいっせいに送り、遅延率のごく低く、I/O率のごく高く保って、計算が同時に——すなわち並列的に——行われるよう図らねばならないことを意味している」(カープ氏)

Panasasのギブソン氏はeWEEKに、「信頼性と整合性」は同社のストレージシステムの2大特性だと述べた。

「例えば、ディスク障害の修復中にディスク読み込みエラーが発生したり、たまたまネットワークエラーが起こったりといったタイミングの悪いときでも、Panasasのシステムなら、ほんの一部のデータが利用不能になったからといって、数テラバイトの情報すべてを投げ出すようなことはない。こうした場合、Panasasのシステムは、問題のあるデータを含むファイルを自動的に囲い込み、残りの数テラバイトのデータに対するアプリケーションおよびユーザーのアクセスを確保する」(ギブソン氏)

この並列ファイルシステム構造は、いずれ企業コンピューティングにも取り入れられるのだろうか。

Evaluator Groupのアナリスト、トム・トレイナー氏は、eWEEKに次のように説明した。「Panasasは今でも、超ハイエンドスーパーコンピューティング環境における、並列I/O処理の高速化を推し進めている。こうした取り組みはある種ニッチな分野に属するもので、大半のストレージベンダーは、利益性の高い重要な事業とはとらえていない。だが実は、これこそがEMCやIBMといった企業がビジネスチャンスを逸している分野なのだ」(トレイナー氏)

また同氏は、PanasasやBlueArcなどは、スーパーコンピューティング向けの技術が企業コンピューティング分野へ徐々に波及し始めていることを、すでに意識していると話した。

「データは、驚異的なペースで生成され続けている。例えばクレジットカード会社は、クライアントのアカウント情報を瞬時に移動しなければならず、詐欺防止のためにデータの分析速度をこれまで以上に向上させなくてはならない。データセンターでもスーパーコンピュータが導入され始めているが、まるで怪物のように大量のデータを吸い込むスーパーコンピュータが普及しても、十分な能力を備えたストレージ製品を供給できるベンダーは数少ないだろう」(トレイナー氏)

これとは異なる見方を示したのが、The 451 Groupのバルタザール氏だ。「今日流通しているクラスタ技術は、(すべてが)プロプライエタリである。

しかし、PNFS仕様が策定されれば(ギブソン氏は数年前から、標準化団体に同仕様の策定を提案している)、企業もこれを採用しやすくなるはずだ。ただし現時点でのクラスタ技術は、HPCなど限定的なマーケットを対象としたものになっている」(バルタザール氏)

関連キーワード

IBM | Linux | ストレージ | スーパーコンピュータ

【Chris Preimesberger, eWEEK】

Editorial items that were originally published in the U.S. Edition of "eWEEK" are the copyrighted property of Ziff Davis Media, Inc. Copyright (c) 2008. All Rights Reserved.

ITmediaはアイティメディア株式会社の登録商標です。