

VeloBit HyperCache

高性能SSDキャッシュソフトウェア

目 次

VeloBit HyperCacheの紹介	1
概要	1
SSDの採用検討	1
第一の課題	1
第二の課題	1
VeloBit HyperCacheソリューション	1
VeloBit HyperCacheによるソリューションの内容	3
Line-Speed Delta Compression	3
Write - cache	4
Scan-Resistant Caching	5
結び	5

VeloBit HyperCacheの紹介

概要

VeloBit HyperCacheはソリッドステートディスク(SSD)とRAMのコンビネーションにより、透過的なアプリケーションアクセラレーションレイヤーを構成し、ハードウェアのボトルネックを削減する、プラグアンドプレイのキャッシュソフトウェアです。VeloBit HyperCacheはSSDを通常のソリューションより高性能で経済的な独自の方法で使用し、システムにSSDを追加した場合の複雑な問題を削減します。

SSDの採用検討

VeloBit HyperCacheはより高いアプリケーションの性能に関する要求に対応します。機械的なディスクスピードではムーアの法則のペースには対応できませんので、殆どのアプリケーションのボトルネックはストレージのIO性能によります。このボトルネックを解消する方法として、機械的なディスクよりはるかに低い消費電力で、フラッシュベースのソリッドステートディスクへのニーズが高まっています。このようなSSDベースには2つの課題があり、ソリューションを難しくしています。

第一の課題

第一の課題としてエンタープライズ仕様のSSDのコストがあげられます。現行のフラッシュ技術はデータの読み出しには最適ですが、データ書き込みは比較的遅く、加えて上書きを実行するためにフラッシュデバイスのページ(4K)ブロックを消去してから、書き込みが可能になります。通常、一回の処理では64ページを消去します。この消去、書き込みのサイクルは性能上の影響に加え、ドライブ寿命に影響を与えます。フラッシュのこの消去、書き込みの問題を解消する技術も有りますが、実装されている製品は非常に高額になってしまいます。

第二の課題

第二の課題はSSDをシステムに搭載する際に、ストレージの管理に問題が発生することです。殆どのユーザーは高位性能なストレージを選択し、安全にデータが保存され、オペレーションの複雑性とコストを最小限にするようにします。ここにSSDを持ち込む場合、以下の3つの選択があります。

- ・ SSDをプライマリストレージデバイスとして使用する。
- ・ SSDをアプリケーション専用のデバイスとして使用する。
- ・ SSDを新たな階層化ストレージとして使用し、データのTiering(階層化管理)を行なう。

以上の選択肢にはデータ保護とデータの管理システム、ポリシー設定、処理、更には、ストレージ管理者の負担増とコスト増の問題がつきまといます。

VeloBit HyperCacheソリューション

VeloBit HyperCacheは以上のSSDに関する二つの課題を、幾つかの鍵により克服することができます。

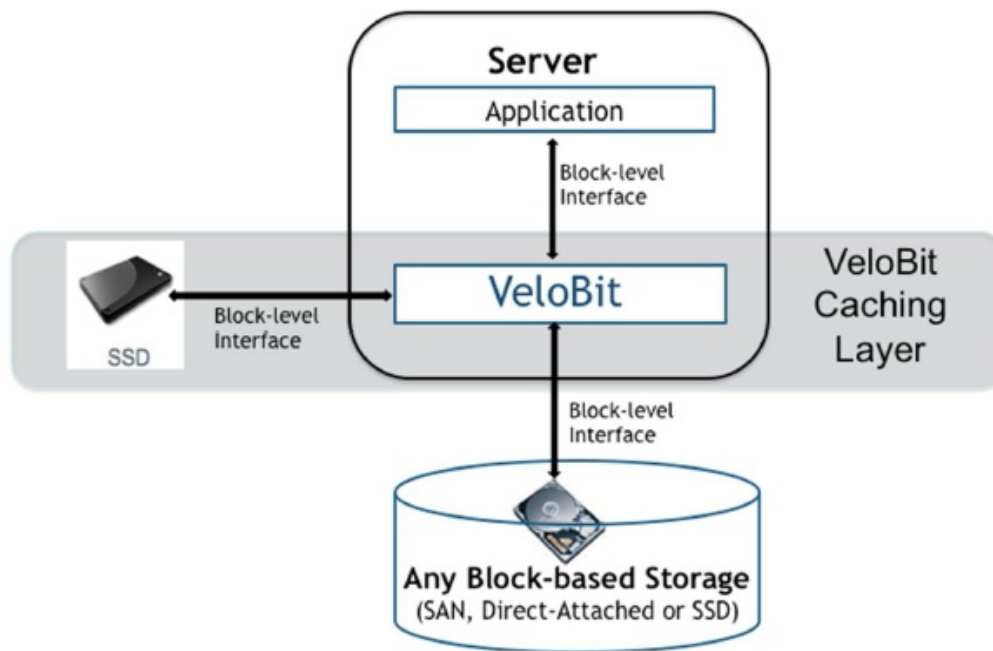
- ・ フラッシュベースのメモリーはアンバランスな性能を持っています。高速なリード性能の一方、ライトの負荷に対しては、コストの問題と信頼性の問題が浮上します。
- ・ 異なるメディアを使用する従来からの垂直的なストレージ階層技術ではSSDは使用されません。垂直的な階層では性能重視のデータ(Hot)は最も高性能で、高額なストレージに直接保存されますが、それほどHotではないデータも一旦はストレージ階層の高い部分に書かれ、やがて、低速なストレージに移動します。結果的に階層の高い部分への負荷は増えることになります。このような方法は、SSDを使用する場合にはあまり好ましくないソリューションです。また、関連する問題として、ストレージの階層化にはデータの仕分け、移動という時間が掛かり、複雑な作業がつきまといます。



VeloBit HyperCacheはサーバー内のメモリーとSSDを使用して、キャッシュレイヤーを構成します。このレイヤーはアプリケーションやプライマリストレージからは認識することはできません。VeloBit HyperCacheは”horizonttal” アーキテクチャーという技術を使い、SSDをメモリーの一部として取扱い、キャッシュの領域を拡張させます。このSSDが実現するキャッシュ領域の拡張にはデータのコピーを保持し、最終的にはプライマリストレージに保存されます。SSDはホストの一部として、ストレージシステムに付加されます。この技術は、データがSSDをバイパスし、ストレージに直接I/Oすることも可能で、SSDへの書き込みの回数を削減することができます。

キャッシュレイヤーに書かれたデータは圧縮されます。この圧縮で、SSDの書き込み負荷は同時に軽減されます。更に、データ圧縮は、VeloBit HyperCacheのメモリーとSSDの機能をより効果的にします。キャッシュから直接サービスされるI/Oリクエストの数を増やし、ストレージの負荷を軽減することで、ストレージのサービスレートを劇的に向上させます。

VeloBit HyperCacheのこれらの機能はSSDの書き込み負荷を軽減し、SSDの性能を向上させ、SSDの消耗を軽減します。則ち、課題#1を解決することができます。また、VeloBit HyperCacheはアプリケーションやプライマリストレージに何らの変更も必要としないので、データマネジメントシステムやプロセスの変更が必要なく、第二の課題を解決します。



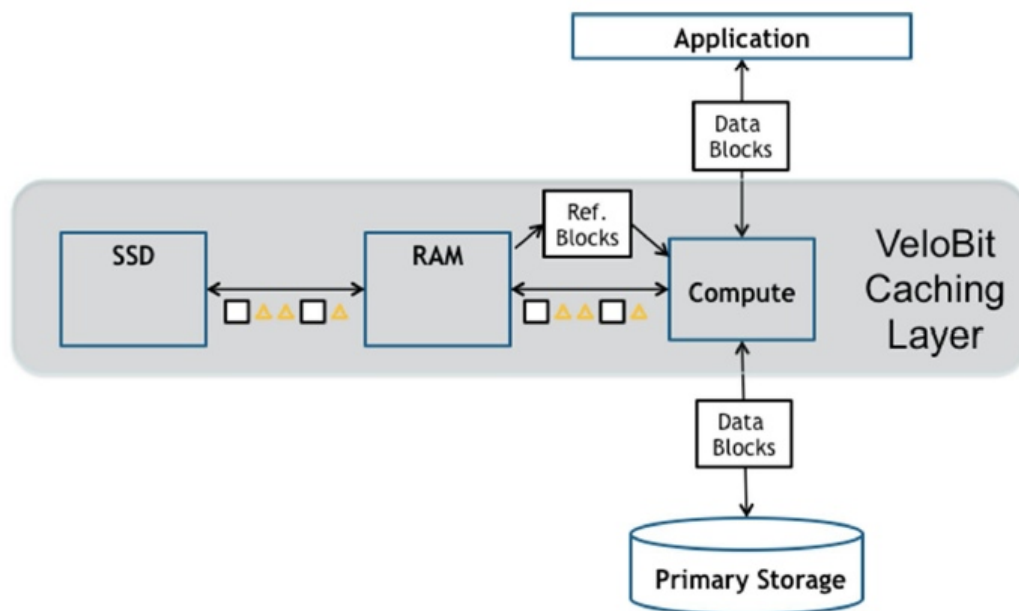
VeloBit HyperCacheによるソリューションの内容

Line-Speed Delta Compression

VeloBit HyperCacheはインラインでのデータ圧縮を使用し、データの特定されたブロックを圧縮します。これは、総てのソフトウェアプロダクトは独自のアルゴリズムにより、類似性の高いブロックを持つ特性を利用しています。一方、従来からのキャッシュアルゴリズムはリード確率の高いデータを一時的にキャッシュに入れておくか、あるデータの読み出しを契機に連続するデータを先読みし、キャッシュに保存する方法が一般的でした。第三の方法として研究開発されたアルゴリズムをContent Localityとよばれます。これはVeloBit社が独自開発したデータを区分する新たなアルゴリズムです。

次の図はVeloBit HyperCacheがどのような様に稼動するかを表しています。Content Localityの区分はデータのブロックを”Hot”、”Cold”と云うようには見ません。VeloBit HyperCacheはそれぞれのブロックがどの程度その他のブロックに類似しているかを判断します。この結果、以下の3つのブロックのタイプが存在します。

- ・ Reference block : Reference blockとは他のブロックと類似性がある内容を含むブロックです。この様なブロックのコンテンツは”popular”と呼ばれます。
- ・ Associated block : 類似性の高いカテゴリーとしてブロックのコンテンツが区分された場合、新しいブロックは、そのReference blockの”associated”になります。Associated blockはReference blockへのポインターとそれらの差分データを持ちます。
- ・ Independent block : あるブロックが共通性のあるコンテンツを持っていない場合、Reference blockにはなりません。もし、このブロックがReferenceでもAssociatedでもない場合は”Independent”ブロックとして扱われ、圧縮されません。



ラインスピードでのブロックの差分データの圧縮プロセスは、ReferenceのブロックとAssociatedのブロックを比較することから始まります。それらに差がある場合、その差分を圧縮し、キャッシュに書き込みます。これをDeltaと呼びます。複数のDeltaブロックは一塊にされ、一回の書き込み動作でSSDに保存されますので、SSDの消耗を軽減します。Reference “セット”はVeloBit HyperCacheが新しいブロックを見つけた場合にそのセットにこのブロックを追加します。Associated blockによって使われなくなったReference blockはReference セットから削除され、Independent blockになります。新しいAssociated blockは古いAssociated blockに代わります。アプリケーションがVeloBit HyperCacheからブロックを読み出す場合、システムは最初に、キャッシュにこのブロックが無いかチェックします。もし、ブロックがキャッシュにある場合、VeloBit HyperCacheはそのタイプを判断します。ReferenceかIndependentブロックの場合、オリジナルのフォームでデータを返します。また、Associated ブロックの場合、delta データと、関連するReference ブロックを再結合して、アプリケーションに返されます。

Write - cache

リードキャッシュと同様、VeloBit HyperCacheはオプションのライトキャッシュを使用することができます。一時的にメモリーやSSD内に保存し、書き込まれたデータがプライマリストレージに保存される前に、アプリケーションに対して処理が完了したことを伝えます。この通知によりアプリケーションからストレージへの書き込みを高速化することができます。

Write-cachingはWrite-Cache depthパラメータでコントロールされます。これは、VeloBit HyperCacheがデータを受け取った時点とソフトウェアがプライマリストレージにデータを書き出しを開始する時間を意味します。注意していただきたいのは、データはこの時間内にストレージに書き込まれるとは限りません。VeloBit HyperCacheのレイヤー内にあるデータを“ダーティー”と呼びますが、以下のパラメータによりデータはプライマリストレージにフラッシュされます。:

- ・ Write-cache-depthパラメータが定める時間により、VeloBit HyperCacheがレイヤー内に保持されたブロックをプライマリストレージに書き出す。
- ・ VeloBit HyperCacheドライバーがプライマリストレージや、SSDがアイドル状態であると判断した場合、データの書き出しがプライマリストレージに開始されます。

プライマリストレージに書かれるブロックは、その前に解凍され、元のフォームで転送されます。ライトキャッシュがオフになっている場合は総てのデータは直接プライマリストレージに書き込まれます。

Write-cache-depthを“0”にすることで、この“Write Through”モードに設定することができます。

Scan-Resistant Caching

VeloBit HyperCacheはScan Resistant Cachingという機能を持っています。これは、大きなシーケンシャルデータをキャッシュに保存することを防ぎます。SSDはランダムアクセスに優れていますので、VeloBit HyperCacheの第一義的なゴールは高頻度にアクセスされるデータで、ストレージ内にはランダムに配置され、アクセスに時間がかかるデータをキャッシュに保存することです。シーケンシャルデータの場合は、むしろ高性能なプライマリストレージの方が優れた性能を発揮する可能性があります。

更に、大容量のシーケンシャルデータをディスクに保存する場合、もしキャッシュしてしまうと他のランダムで高頻度なデータへのアクセスを阻害します。例えば、バックアップアプリケーションはディスク上のデータを一度総て読み出します。しかし、このデータはアプリケーションから利用されることはありません。従って、このデータはキャッシュされるべきではありません。

Scan-Resistant Cachingは、Sequential Filteringとも呼ばれ、VeloBit HyperCacheではデフォルトでenableになっています。

結び

VeloBit HyperCacheはSSDの書き込み負荷を軽減します。アプリケーションとプライマリストレージの間にキャッシュレイヤーを構成することで、SSDの性能を向上させ、消耗を軽減します。