

Company Name: 株式会社マックスサーブ

Contact Info: Tel 03-6202-9092

Fax: 03-6202-9093

Website: <http://shop.maxserve.co.jp>



IT開発現場におけるサーバ仮想化とマルチレーンの活用

嶋田 大貴 (STBBS.NET) <shimarin@stbbs.net>

IAサーバをターゲットとする昨今のIT開発現場では多様なプラットフォームへの対応が要求されます。いわゆる LAMP(Linux Apache MySQL PHP) を代表に、Ruby on Railsや J2EE、PostgreSQL、OracleやDB2を使って開発することもあるでしょう。

複数の開発プロジェクトが同時に進行する中小規模組織のオフィス内では、それぞれの開発に用いるためのサーバが乱立しがちです。壊れそうな電源タップに接続されていたり、ゴミ箱から拾ってきた10base-Tのハブが使われていたりしている光景も想像に難くないでしょう。ケーブル類は入り組んで埃だらけ。保守停電の際など、保護すべきサーバの実体が把握できておらず、結局月曜日に開発スタッフが出社した時には開発サーバがfsckで引っかかったまま起動せず大騒ぎ、などという経験のある方もおられるかと思います。

また、サーバがバラバラに配置され共有のストレージも無いような状態では計画的なバックアップなど取れようはずもありません。もし開発サーバのハードディスクがクラッシュした場合は開発スタッフ各人の端末から成果物を拾い集めるなどといった作業を強いられるでしょう。残念なことに、それでも納期の迫る速度は一定なのです。

設置場所を小さな棚に集中し、柔軟に拡張可能な共有ストレージを持つ多数のサーバを運用する。そんなことが可能でしょうか。答えのひとつとして、仮想化とマルチレーンの活用についてご紹介します。

マルチレーンの特徴

- 1本の外付用ケーブルで4台までのハードディスクを接続
- 通常のシリアルATA準拠 I/Fとハードディスクが使用可能
- ハードディスク一台ごとにシリアルATAの帯域を完全に利用可能（ポートマルチプライヤに対する優位性）
- ハードディスクを外付けすることにより本体の電源に余裕を持たせ安定性を確保
- うっかり抜けたりしにくいケーブル
- ケーブル長は2メートルまで

NASやSANは高い...

その通りだと思います。しかしながら、共有の巨大ストレージは誰もが欲する資源であることに疑う余地はありません。

少し発想を転換してみましょう。大きなハードディスクを持ったマシンをひとつ用意し、その中に必要な全てのサーバを格納してしまえば共有ストレージを実現したことにならないでしょうか。

仮想サーバ

ひとつのマシンに仮想的な複数のサーバを格納する考え方は最近に始まったことではありません。例えば、Webサーバではバーチャルホストという機能によって一台のホストで複数のサービスを運用することが出来ます。

しかしバーチャルホストでのサービス分割は所詮 Webサイトを提供するためのスペースを別々のディレクトリに分けるだけのものにすぎず、サーバの環境そのものを分割できるわけではありません。つまり、Webサーバのバーチャルホスト機能を使って CentOSと Debian GNU/Linuxを一台のマシンで同時に動かせるわけではないのです。

ご存じの方もおられるでしょうが、そのような事を可能にするソフトウェアが最近多く利用されるようになってきました。OSを丸ごと仮想環境の中で動作させてしまうようなソフトウェアとしては、VMWareをはじめ Windowsでは Virtual PC、Macでは Parallelsが有名です。これらは全て有料のソフトウェアですが、無料で利用できるものもあります。

オープンソースの仮想マシンプラットフォームである Xen(ぜん)を後ほどご紹介しましょう。

マルチレーン接続

いちどハードディスクの話に移りませぬ。ひとつのマシンにたくさんのディスクを装備すると言いましたが、ディスクをいくつも内蔵できるサーバはとても巨大で、取り回しのしにくいものです。なので、本体はなるべく(高性能

ながら)小型のものを選択し、ディスクは外付けにしたいと思うことができるでしょう。(男らしくフルタワー筐体を選択して激しく後悔したことがあるのは私だけではないはず)

ハードディスクを外付けするためのインターフェイスとして何が思い浮かびますか?まず、昔ながらの方法としてはSCSIがありますが、これを選択するとその後何をするにも高価です。関連製品の入手性が最も良いのはUSBでしょうか。しかし手元の端末で使うならまだしも、サーバにUSBハードディスクではあんまりですよね。FireWire800が良い選択ですが、いかんせん関連製品の入手性があまり良くなく、SCSIほどではないものの高価です。

転送速度が速く、関連製品の入手性が良い規格といえば今はシリアルATAです。最近ではeSATAといって外付け用のコネクタ規格もあります。ところがシリアルATAは基本的に1ポートに1台のハードディスクしか接続できませんので※、たくさんのハードディスクを外付けにしなければたくさんのケーブルを接続しなければなりません。

たくさんのケーブルをつなぐのは嫌なので、1本にまとめてしまいたくさん。それがマルチレーン接続です。シリアルATAのポート4つを、1本のケーブルにまとめてしまうことができます。

※ポートマルチプライヤを用いれば1ポートに複数のディスクを接続できます。ただドライバ側の対応が必須かつ相性もあるらしく、私はまだLinuxで利用できたことがありません。

LVM(論理ボリュームマネージャ)

マシンに何台ものハードディスクを接続すると、それらをディレクトリツリーのどこにマウントすれば良いのか悩みますね。複数のディスクをまとめてひとつのボリューム(パーティション)として扱えると便利です。とはいえ、ただひとつにまとめられた何テラバイトものボリュームが存在していても、それを果たして/varにすればいいのか、/homeにすればいいのか判断が難しく扱いに困ります。

LVMはLinuxに搭載された大変便利な機能です。LVMを使うと下記のようなことができます。

- 複数の物理ディスク(物理パーティション)をひとつのボリュームグループ(vg)にまとめる
- ボリュームグループから任意の容量を切り出して独立した論理ボリューム(lv)を作成する
- lvの容量を動的に拡張する

例えば500GBと400GBと320GBのように容量の異なるハードディスクをまとめて1.22TBのvgにしておき、100GBの領域が必要になったらその中から100GBを割り当てて新たにlv(パーティション)を作り、mkfsして好きな場所へマウントします。100GBの割当てを後で150GBへと動的に拡張することも可能です※

vgの容量が足りなくなったら新たに好きな容量のハードディスクを調達してきてvgに加えてやる事が出来ます。ディスク拡張のためにデータを移し替えたりする必要はありません。

LVMの使い方をマスターすればハードディスクの増設は自由自在です。まだご存じない方は是非調べてみてください。

※ファイルシステムによります

Xen

Xenは「仮想マシンモニタ」と呼ばれるソフトウェアです。XenはLinuxカーネルに組み込む形で利用します。

Xenの目的は、一台のマシンを複数の仮想サーバに分割にすることです。分割されたそれぞれの仮想サーバは「ドメイン」と呼ばれます。

マシンの電源を入れてXenの組み込まれたLinuxカーネルがロードされると、通常のようにLinuxが起動します。これがドメイン0と呼ばれるドメインになります。

ドメイン0は特権を持ったドメインで、他のドメインを起動したり、停止したり、設定を変更したりといった管理を行うことができます。ドメイン0以外のドメインをドメインUと言います。

先に述べた「ひとつのマシンに複数のサーバを格納する」とはすなわち、同じマシンの中にドメインUを必要なだけ作成し、同時に稼働させるということです。それぞれのドメインUは個別にIPアドレスを持ち、別のサーバとして動作します。あるドメインUで稼働しているのがubuntuで、別のドメインUではGentooが動いている、ということも出来ます。それぞれは仮想のネットワークインターフェイスでつながっているので、ストレージはNFSなりiSCSIなりで好きなように共有できます。当然ケーブルは要りません。

XenとLVMとマルチレーン

先に紹介したLVMを活用すれば、それぞれのドメインUに対してディスク容量を任意に割り振ることが出来ます。

何百万件といった文書から全文検索を行うシステムの開発を始めるときは、50GBのlvを作成してそこにドメインUを構築しましょう。

デザインの洒落た、しかしシンプルな広告とお問い合わせフォーム程度のWebサイトを作る時は3GBのlvを切り出してドメインUを構築してあげれば十分かもしれません。

いかなる目的のサーバを立ち上げるときも、ディスクの物理的な容量を意識せずvgからそれぞれ必要な容量だけのlvを切り出して、その中に仮想サーバを構築すればいいのです。これによりディスク容量の利用効率は最大限に引き上げられるでしょう（それぞれのサーバが仮想サーバでなく本物のサーバだったら、個別に搭載されるハードディスクの容量がどれくらい空きのまま無駄になるか考えてみてください）。

LVMを活用すれば1台のハードディスクでも十分な数の仮想サーバを運用できるでしょうが、冗長化のことも考えると最低2台は欲しいですし、4台あればそれに加えてストライピングによる性能向上も望めます。

マルチレーンなら、先に述べたとおり1本のケーブルで4台までのシリアルATAハードディスクが外付けできますし、それぞれのディスクに対して同時

に300メガバイト/秒の帯域でアクセスできるため性能面でも申し分ありません。

それに、安価なシリアルATAのディスクが利用できるというコストメリットが同時についてきます。これは結構お得です。

Debianのインストールは5分で

参考までに、私が新たな仮想サーバを作成する手順をご紹介します。LinuxディストリビューションはDebian GNU/Linuxです（ubuntuも同じ手順でできます）。OSのパッケージはネットワークから取得しますのでCD-ROMは必要ありません。

```
lvcreate -L 3G -n debian vg
mkfs.xfs /dev/vg/debian
mount /dev/vg/debian /mnt
debootstrap --arch i386 sid /mnt http://ftp.jp.debian.org/debian
umount /mnt
cd /etc/xen
cp template debian
vi debian
xm create debian -c
```

ここまででDebianが起動します。速いマシンと速い回線（と速い操作）なら、最初のlvcreateコマンドから時間を測ってほんの5分程度です（注:ドメインUの設定ファイルはあらかじめテンプレートを用意しています）。

5分でインストールした仮想サーバ上のDebianにOracle10g Express Editionをセットアップし、動作させることもできました。

このような処理が速く出来るのも、帯域幅に余裕のあるマルチレーン接続だからこそです。

クラスタ技術と仮想サーバ

システムのパフォーマンスを向上するため、又は耐障害性を確保するためにクラスタ技術がよく用いられます。しかしながら、アクティブ-アクティブなクラスタにしるHA構成のクラスタにし

る、ただノードを並べれば実現できるものではありません。ミドルウェアによって透過的にクラスタ処理のなされるものもありますが普通はプロプライエタリで高価なソフトウェアで、顧客の予算に合うケースばかりではありません。

オープンソースプロダクトを利用しつつ高可用性を提供するという付加価値の高いビジネスを実現するためには、a) システムの設計 b) アプリケーションの機能 c) インフラの全てをセットで考案し、自前で組み立てる必要があります。

ところで、あなたはあちこちの記事で紹介されているオープンソースのクラスタリング・ソフトウェアについてどれくらい実際の動作を知っていますか？記事で見かけたことのある、何でも解決してくれそうなあのソフトは、実はあなたがこれから作ろうとしているシステムの要件を満たすために都合のよくない制約を抱えていたりしません？

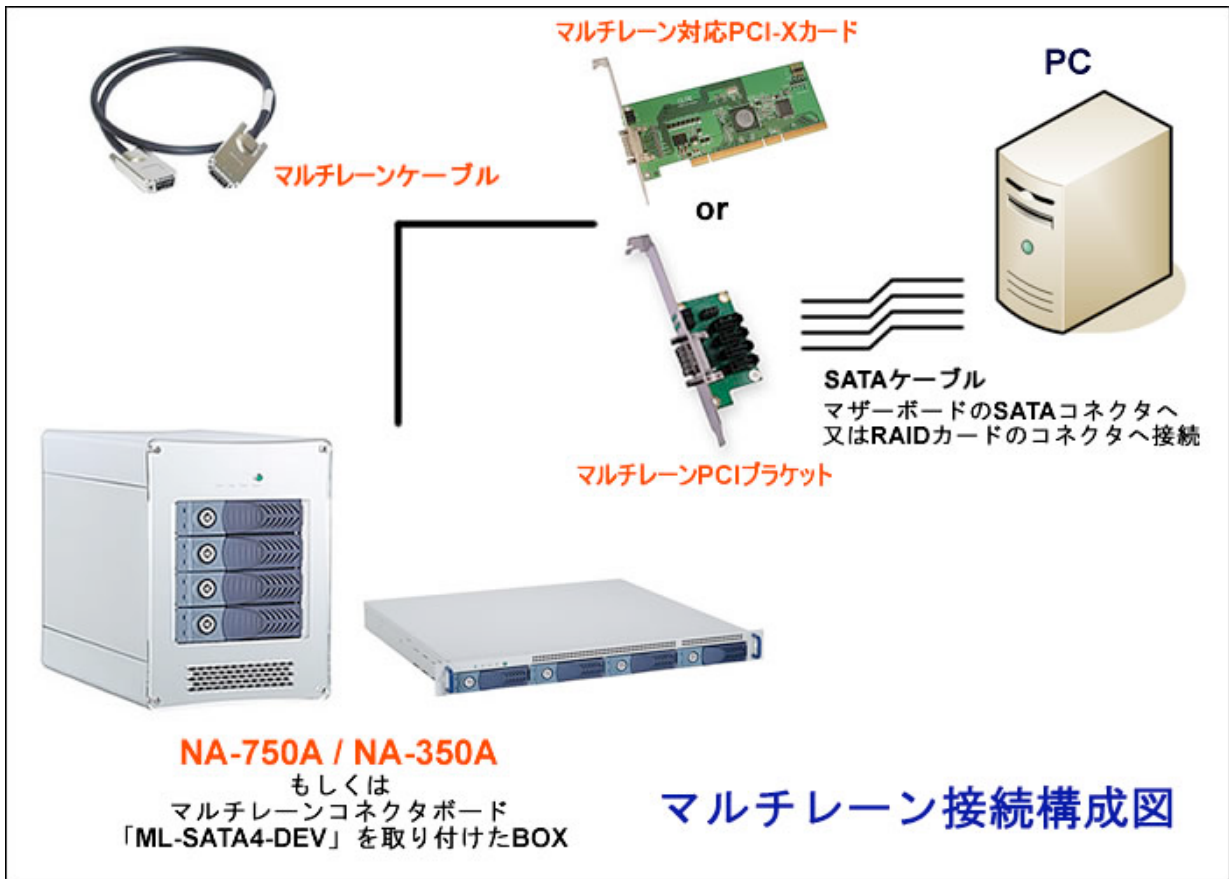
「そんなの、やってみなければわからない」

当然です。やってみましょう。ただし顧客に提案する前にです。

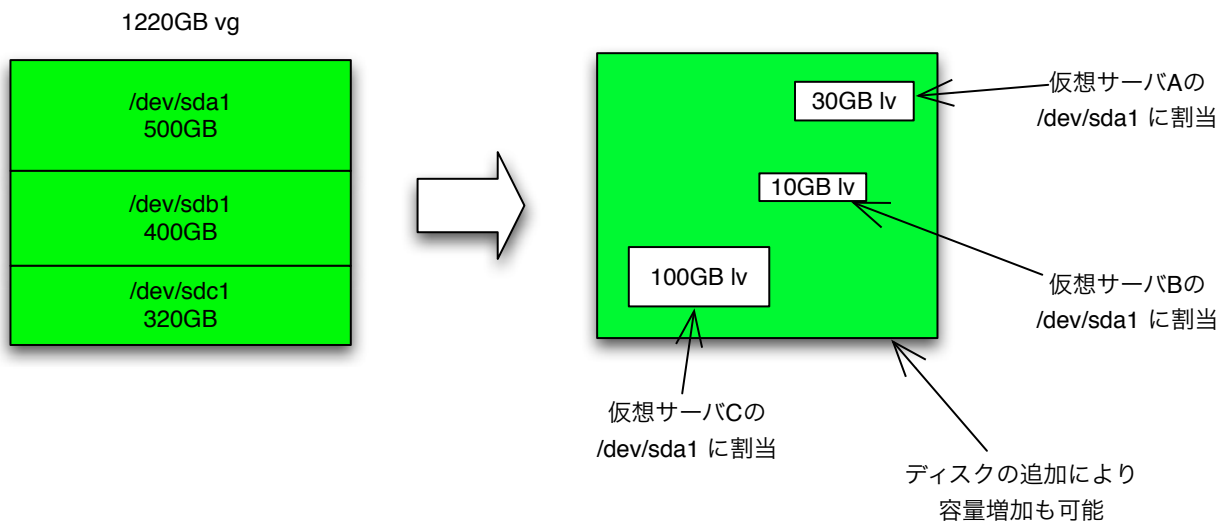
それはそうと、クラスタリングの実験をするにはサーバが複数台必要です。場合によっては2台どころではなくもつとです。検証には専用のカーネルで起動する必要があるかもしれませんが、いま稼働中のサーバはなるべく再起動したくありませんよね。

何が言いたいかわかりでしょう。仮想サーバの動作する環境を整備してあれば、クラスタ技術の実験も楽に行えます。実験が終わったら仮想サーバを停止してlvを削除すれば結構です。

仮想サーバは単に運用コストの削減をもたらすだけでなく、ビジネスに付加価値をつけるための役にだって立つかも知れません。



LVMと仮想サーバ





Multilane スターターキット

さあマルチレーンをご体感してみませんか？

ストレージを高速かつ外部へ配置する方法として今もっとも最適な方法。

それが「マルチレーン」です。

マルチレーンを用いる事によって速度を落とすことなく、HDDを外部へ出すことができます。

CPUの排熱や狭いケーブルの取り回し、本体の電源容量問題・・・

様々なPC内部の問題からストレージを解放する方法・・・それがマルチレーンです！

マルチレーンケーブル(1m)とPCリアスロット用ブラケット、ストレージデバイス側ブラケットの3つをセットにしてお届けします。

これではお好きなRAIDカードへSerial-ATAケーブルを繋いでください。

(マザーボードのSATAコネクタに繋いでも可)



ML-8000SAS

Serial-ATAディスクに加え、SCSIの次世代規格SAS(Serial Attached SCSI)ディスクも使用可能な8ベイホットスワップ対応マルチレーンエンクロージャ。

2007年3月現在、最もGB単価が安い320GBハードディスクを装着すると、約2.5TBのストレージエリアを確保出来ます。

シリアルATAコントローラ側で対応していればホットスワップも可能なので、プロジェクト毎にディスクセットを入れ替えるということも簡単にできます。



RocketRAID232

ハードウェアRAIDを低コストで。2個の外部向けminiSASコネクタで8台までのデバイスをサポート

※シリアルATAのハードディスクが使えます

LinuxにはRAID機能が含まれていますが、ソフトウェア処理ではCPU負荷が高く信頼性も十分とはいえません。

信頼性と性能を求めるなら、やはりハードウェアでRAIDを構築するのが良いでしょう。

RocketRAIDは、各主要Linuxディストリビューション向けドライバの他にOpenSourceドライバ(カーネルソースに対するパッチ、もしくはモジュールとしてコンパイル)もダウンロードにて提供しているため、特別なカーネルを使う人でも安心です。

Webベース・コマンドベース両方の管理ツールが付属。32bit/64bitとも対応です。

マックスサーブは、PCシステムの輸入・販売のほか、主に中小企業向けストレージシステムの導入・サポートサービスを行っています。またProShopMAXSERVE (<http://shop.maxserve.co.jp>)ではストレージ・ネットワーク関連商品を多数取り扱っております。

この記事は STBBS.NET が作成しました。関連する情報は STBBS.NET blog (<http://www.stbbs.net/blog/>)にも随時載せています。