統合ディスクレスネットワーク基盤システム

- 次世代の Linux ネットワークブートテクノロジ -

1. 背景

　教育用計算機システムなどでは、まったく同じ設定のコンピュータ環境を多数用意する必要がある。また、科学技術計算においては、PC クラスタが必要不可欠な技術となっている。このような環境では何十、何百という数のコンピュータが一度に導入されるが、多数のコンピュータの管理とネットワークの構築には、多大な労力と時間が必要であった。

　また、コンピュータシステムは常に更新していく必要がある。しかしそのためには、既存のシステムを一度撤去しなければならない。そのため新しいシステムは確実に動作するという保証を得る必要があるが、その作業も時間的なコストと地道な労力が必要になる。結果として古いシステムが更新されないことも多い。

2. 目的

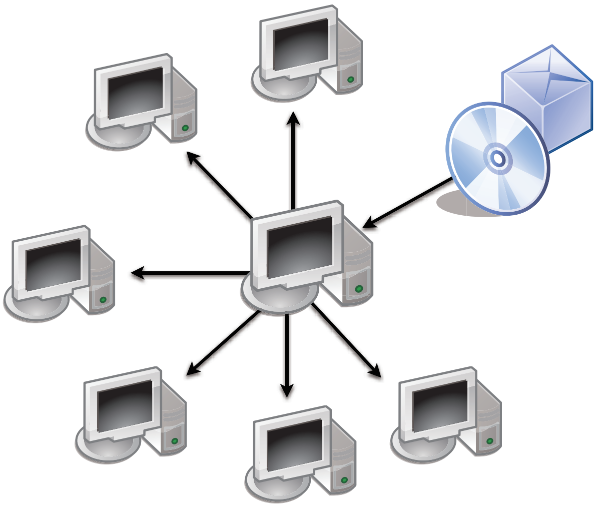
　多数のコンピュータの扱いが難しい原因は、大部分は同じだが一部だけ異なる作業を何度も何度も行わなければならない点にある。同じ設定作業は一度行えば済むようにすることで、労力を軽減できる。新しいシステムの導入の際に既存のシステムを撤去しなくてはならないという問題は、システムをディスクレスで動作させることで解決できると考えた。ディスクレスシステムは既存のシステムを残したまま使用できる。さらにシステム全体を CD や USB メモリなどの小さいメディアに納めてしまえば、システムの更新は手軽にメディアを入れ替えるだけで完了する。このような解決策を実現するため、一台のコンピュータを CD や USB メモリなどのメディアから起動し、続いて他のコンピュータを次々にネットワークブートすることで、瞬時にディスクレスネットワークを構築するシステムを開発する。

3. 開発の内容

ディスクレスブート

　一台のコンピュータを CD や USB メモリなどのメディアから起動し、続いて他のコンピュータを次々にネットワークブートする。

　ブートメディアには、あらかじめ圧縮した Linux システムを格納しておく。ネットワークブート時にはこの圧縮されたディス



ディスクレスブート

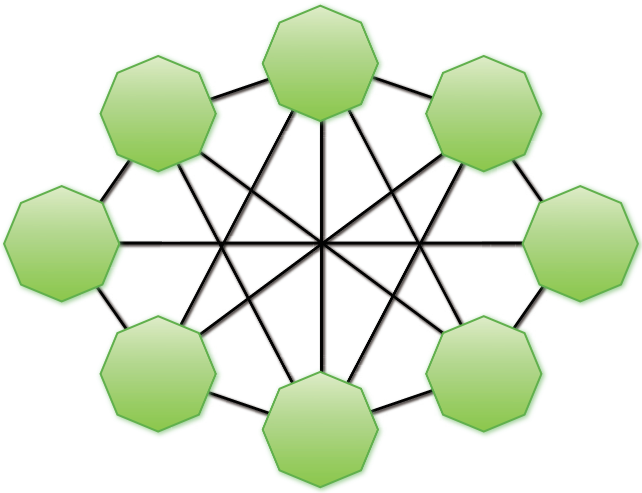
クイメージをそのまま共有することで、ネットワークトラフィックと負荷を大幅に減らすことができた。

起動時にハードウェアの自動検出を行う機能を実装した。ディスクドライブや NIC を自動的に発見・設定・使用する。グラフィックカードを検出し、X11 の設定を自動的に行う機能も実装した。

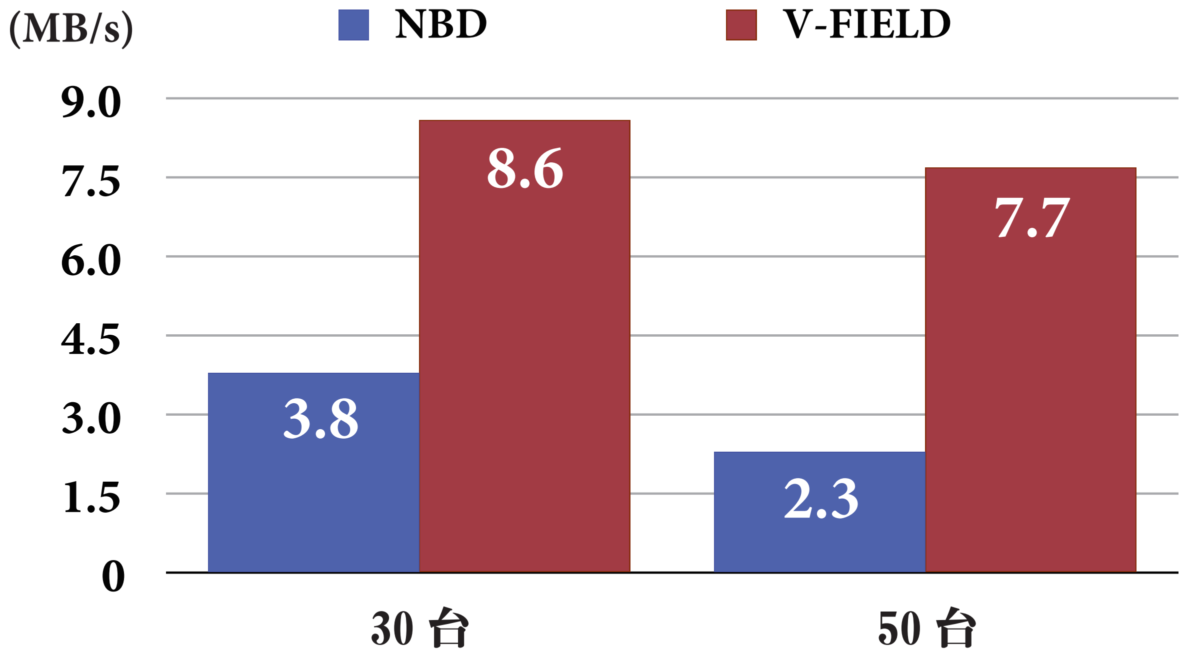
また、HDD を一切使わずに起動できる。既に別の OS がインストールされている環境でも、そのまま利用できる。HDD（あるいは USB メモリ）が接続されていれば、そこにデータを保存することもできる。この際に HDD を利用するかどうかを、HDD 上に保存されているファイル名を使って指示する機能を実装した。この機能によって、起動時に一切指示を与えられなくても自動的に HDD を発見し、使用する。

分散ディスク共有

ネットワークブートには通常 NFS が用いられるが、NFS ではサーバーがダウンするとすべてのクライアントがダウンしてしまう。本プロジェクトでは、共有するディスクイメージを多数のコンピュータに重複させながら分散保持させることで、ネットワーク中の数台のコンピュータがダウンしてもネットワーク全体がダウンしないようにするプログラムを開発した。加えて



V-FIELD



性能比較

アクセス負荷を分散させ、コンピュータの台数が増えても高い性能を保つ。

性能を比較したグラフは、本プロジェクトで開発したプログラム「V-FIELD」と、従来のクライアント・サーバー・モデルに基づくプログラムである「NBD」 を比較したものである。540MB のランダムなデータが書き込まれたファイルを RAM ディスク(tmpfs)上に置いて共有し、すべてのコンピュータで一斉にシーケンシャルリードを行い、平均の速度を計測した。検証には IBM Blade Center HS20（CPU Xeon 3.6GHz × 2way, Memory 1GB, NIC Broadcom BCM5704s）を使用した。

ネットワークサービスの構築

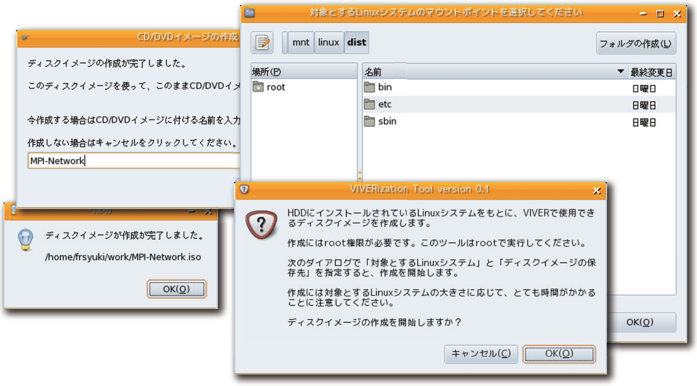
単一のディスクイメージから複数のコンピュータを起動するため、そのままでは完全に同じ設定のシステムが起動してしまう。そこでプラグイン型のアーキテクチャによる機能拡張システムを開発し、ホスト名や IP アドレスなどの割り振りを自動で行うようにした。

ディスクレスブートには標準規格である「PXE」を用いるが、その際に必要となる DHCP サーバーや TFTP サーバーの設定も、このシステムで自動的に行う。

プラグインを追加することで、NFS を使ったファイル共有システムや並列コンパイル環境も構築できた。

ディスクイメージ作成ツール

HDD にインストールされている Linux システムを、本システムで利用できるディスクイメージに変換するためのツールを開発した。非常に簡単な手順で Linux システムをディスクレスブートさせることができる。



ディスクイメージ作成ツール

4. 従来の技術（または機能）との相違

本システムのディスクレスブートプログラムは、特定の Linux ディストリビューションに依存しない。このため HDD にインストールされている Linux を、ディストリビューションを問わず、ディスクイメージ作成ツールを使ってディスクレスブータブルにすることができる。従来 CD からブートする Linux ディストリビューションは存在したが、このようなシステムでは環境に合わせたカスタマイズには制約がある。

また本システムは、1 枚のブートメディアから 1 台のコンピュータを起動するのではなく、1 枚のブートメディアから 1 つのネットワークを起動する。このため、メディア 1 枚の中に並列処理計算ネットワークをインストールしておくなどということが可能になる。利用するときはコンピュータを次々とネットワークブートしていくだけで良い。

分散ディスク共有システムは、従来の NFS を利用するネットワークブート方法よりも

スケーラブルであり、さらに自動的に多重化を行う。

5. 期待される効果

Linux システムをネットワークブートするシステムとして、冗長性を持ち、多重化構成は自動的に行い、コンピュータの台数が増えても高い性能を保ち、そして非常に柔軟なカスタマイズができ、それでいて非常に簡単に使えるシステムである。また本システムを利用して作成したシステムは、ポータブルメディアに入れて手軽に扱える。

6. 普及（または活用）の見通し

OS のブートという低レイヤーの技術であるから、そのままでの活用よりも、他のソフトウェアとの組み合わせによって効果を発揮する。本プロジェクトで開発したソースコードは Web サイト（http://viver.sourceforge.jp/）で公開している。また、さらに継続して本システムを活用して動作するソフトウェアを開発している。

本システムを利用して完成されたネットワークスイートは、ネットワークブートによる動的なネットワークの構築が適した小規模な企業や、多数のコンピュータを一度に使用する研究組織、同じ設定のコンピュータを多数利用する教育機関で利用できる。

7. 開発者名（所属）

古橋貞之　（筑波大学第三学群情報学類）

（参考）
公開 URL: http://viver.sourceforge.jp/