

仮想サーバシステムのための環境管理支援ツールの構築

佐藤 幸紀[†], 上埜 元嗣[‡], 宇多 仁[†], 井口 寧[†], 敷田 幹文[†], 松澤 照男[†]

北陸先端科学技術大学院大学 情報科学センター[†]

北陸先端科学技術大学院大学 技術サービス部[‡]

近年、新しい計算機システムの構築運用手法として仮想化技術が注目されている。仮想化技術により大学内で運用されるサーバの大幅な削減が可能となるため、近年のサーバ数の増加に伴う様々なコストの削減が期待できる。本論文では、学内に分散されているサーバを情報科学センターの管理する仮想サーバに集約し様々なコストを削減することを目的としてオープンメインフレームを用いて高信頼な仮想サーバシステムを構築したことを報告する。さらに、この仮想サーバシステムの管理コストを低減しユーザの利便性を高めるためには必要不可欠な仮想サーバ管理支援ツールを構築し、その動作評価により有効性を確認した。

Building a Management Tool for Virtualized Server System

Yukinori Sato[†], Mototsugu Ueno[‡], Satoshi Uda[†],

Yasushi Inoguchi[†], Mikifumi Shikida[†] and Teruo Matsuzawa[†]

[†]Center for Information Science, Japan Advanced Institute of Science and Technology

[‡]Technical Service Department, Japan Advanced Institute of Science and Technology

Recently, virtualization technologies are emerging as effective solutions for the resource management of computer systems. In this paper, we attempt to consolidate the large number of servers distributed among all over the university to the virtualized server system in our information science center and reduce various costs arisen from the management for the servers. We build a virtualized system using mainframe infrastructure to increase its reliability and performance. We also present a management tool for virtualized system which is essential for increasing its convenience for all of users and effectiveness of the system.

1 はじめに

北陸先端科学技術大学院大学 (JAIST) では開学当初から情報科学センターが一元的な情報環境を提供してきた [1, 2, 3]。情報環境を一元化している理由は大規模で高機能かつ高信頼なシステムを導入し、集中的に管理することで安定したサービスと管理コストの低減をはかることにある。現在提供しているサービスの例を挙げると、電子メールサービス、ファイルサーバシステム、情報環境端末、www サーバ、並列計算機の提供など多岐にわたる。提供するサービスの増加に付随して必要となるサーバ数も増加する。JAIST 情報科学センターで管理・運用をおこなっているサーバ数は、2004 年には 232 台であったが 2007 年には 353 台に達しているように、年々増加する傾向が続いている。

サーバ数の増加はサーバ運用・管理・保守、電気使用量、空調設備、設置場所といった様々な面でのコストの増加につながっている。その一方で、

各サーバの CPU 利用率は常に高いわけではないため、CPU 利用率の低いサーバの統合をはかりサーバ数を減少させることにより、計算機リソースを有効利用することが望まれている。計算機リソースの有効利用はサーバ数の増加に伴う各種コストを削減する。しかしながら、サーバの統合を行う場合、異なるアプリケーションスタックや異なるオペレーティングシステムを混在させる場合、運用上必要なセキュリティーを確保・維持できるかを慎重に判断しなければならない。また、1 台のサーバにサービスを集約した場合、ハードウェア障害によるサービス停止の範囲が広がることも懸念される点である。

近年、新しい計算機システムの構築運用技術として仮想化技術が注目を集めている [4, 5, 6, 7]。仮想化技術により物理的には 1 つのサーバのハードウェアリソースは複数の仮想サーバとして共有され、そのハードウェアリソースは効率よく利用される。仮想サーバ間の独立性やセキュリティーも、仮想サーバシステムが古くからメインフレームで培

われきたため十分な実績がある。さらに、これまではシステム固有の OS に限定されていた仮想化技術が、近年は Linux 等の汎用 OS を用いて利用できるようになり、ユーザの利便性やシステムの可搬性が格段に向上している。

しかしながら、サーバの仮想化による集約が目されている一方で、サーバシステムの集約化には運用上の課題もある。JAIST 情報科学センターで管理しているサーバは、一般利用のユーザに管理権限を与えていないため、アプリケーションのインストール、設定の変更などは情報科学センターと相談の上で行うことになっている。しかしながら、管理上の管理者権限が必要という理由から独自に研究室などで独自にサーバを設置するケースが増えており、情報科学センターへ申請のあったもので約 50 のサーバが設置されている。各研究室が独自に設置するサーバはハードウェアの購入費を削減するために低価格なハードウェアを選択するケースが多く見られ、障害や保守などにより長時間サービスを停止するケースも少なくない。従って、仮想化技術を用いてサーバ数の削減を進める場合においても、集約によるサーバ管理の問題点を克服することが必要である。

本論文では、JAIST 情報科学センターにおいて仮想サーバシステムを構築し試験的な運用を行っている事例について述べる。情報科学センターにおける仮想サーバシステムの用途として、情報科学センターが提供するサービスの集約化、学内の各研究室で独立して構築しているサーバの集約化、各種実験のためのシステムを自由に構成できる環境の提供、サーバ構築やネットワーク構築手順を実習する環境の提供などを想定している。これらの用途として仮想サーバシステムを開放するためには、システム全体の管理者権限のないユーザが仮想サーバの構築・運用を行えることと、仮想サーバの構築・運用のために必要となる設定や管理の利便性が高いことが必要である。そこで、これらの課題を解決する仮想サーバの構築・運用手法に必要な要件を考察した結果、学内の多くのユーザが仮想サーバを管理するためには利便性の高い仮想サーバ管理支援ツールが必要不可欠であるという知見を得た。この知見に基づき、仮想サーバの環境をインタラクティブに構築する Web クローニングツールを開発を行った。

本論文は以下のような構成である。2 章では現状の仮想サーバシステムの技術の概要を述べる。3

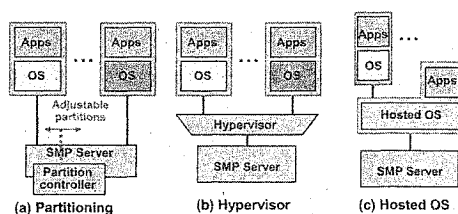


図 1: 仮想サーバ構築法。

章では分散したサーバの集約を行うためにオープンメインフレーム [8] を用いた仮想サーバシステムの構築について述べる。4 章では利便的な仮想サーバの構築運用には Web クローニングツールが必要となる理由と、Web クローニングツールの動作を確認したことについて述べる。5 章では関連研究を述べる。6 章は結論である。

2 仮想サーバシステム

計算機サーバ分野における仮想化は実リソースを管理する手法によりいくつかの技術に分類される [4]。図 1 に既存の仮想サーバ構築法を示す。仮想サーバは共有メモリを持つ対称マルチプロセッサ (SMP) 上に構築される。

パーティショニング型は SMP のリソースを主に空間分割により管理する手法である。SunDomains などの物理パーティショニング [9]、IBM pSeries などの LPAR という論理パーティショニングがある [10]。リソースのパーティションはファームウェアのような低水準のハードウェアにより調整される。性能は高いがリソース管理の柔軟性が低いという傾向がある。

ハイパーバイザ型はハイパーバイザという OS が持つさまざまな機能のうちリソースを管理する機能のみを有するソフトウェアを用いてリソースを管理する手法である [4]。IBM z/VM、VMware ESX server、Xen などに見られる。仮想サーバ間で高い独立性を保ちつつ高い性能を達成可能であるが、多くの場合ファームウェアのような低水準ハードウェアでのサポートを必要とする。ハードウェアと密接にかかわっているために、デバイスドライバの対応が必要となったり、管理が難しい傾向がある。

ホスト OS 型はホスト OS がリソースを管理をおこなう手法である。Microsoft Virtual Server、VMware Server などに見られる。OS が混在する環境を汎用

的なハードウェアを用いて構築できるが、すべての処理にホスト OS の影響が介在してしまうため性能はあまり期待できないとされている。

3 仮想サーバシステムの構築

仮想サーバシステムを選定し構築していく上で実際の運用の形態を十分に考慮することが重要である。そこで、JAIST 情報科学センターの情報環境におけるサーバへの要求を分析し、サーバの運用上の問題点を踏まえ必要条件を以下のようにまとめた。

- 仮想化を行ってもサービスの品質を保つシステム性能を持つこと。
- 信頼性の高いハードウェア構成であること。
- 一般的な汎用 OS を使用できること。

サーバの集約を行いつつもサービスのクオリティを保つためにはホスト OS 型による仮想化では性能面で不十分であると考えられるため、ハイパーバイザまたはパーティショニングにより仮想化を行っている製品から選ぶこととした。信頼性の点では、長年の実績がある大型の UNIX サーバやメインフレームといった製品が適していると考えられる。UNIX サーバは汎用的な OS を使用している一方で、I/O の仮想化をソフトウェアで行っているため I/O 性能が低く I/O 性能がシステム性能のボトルネックとなってしまう可能性が高いという欠点がある。一方で、メインフレームは I/O 性能は UNIX サーバと比べて強化されているが、専用 OS の場合が多く一般のユーザが使い慣れた汎用的な OS ではないという欠点がある。

近年、メインフレーム上で専用 OS のほかに汎用 OS である Linux を仮想的に動作させる仮想サーバシステムが登場した。メインフレームでありながら汎用性の高い OS を搭載できる仮想サーバが最も望ましいと考えられるため、我々は IBM eServer zSeries 890 (以下 z890) というメインフレームを導入し Linux を仮想的に動作させるという仮想サーバシステムを構築した [8]。

IBM z890 は高い冗長性および信頼性により圧倒的な連続可用性 (z890 本体: 平均故障間隔 35 年) を実現している。また、仮想化技術としてパーティショニング (LPAR) とハイパーバイザ (z/VM) を

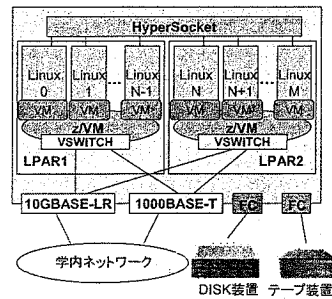


図 2: z890 の構成。

採用しているため、性能の高い仮想サーバの構築を可能としている。I/O デバイスおよびネットワークもハードウェアレベルでの仮想化を行うため、I/O の性能が高く I/O がシステム性能のボトルネックとなりにくい。OS として一般的な汎用 OS である Novell SUSE Linux Enterprise Server9 または 10 を利用可能であり、ユーザがアプリケーションを構築し運用する上で利便性が高い。その一方で、他の仮想サーバシステムと比べて z890 上で動作可能な OS は必ずしも多いわけではないことと、標準の仮想サーバ管理ユーザインターフェースとして採用されている CMS (Conversational Monitor System) が使いにくいという不利な点もある。

図 2 に JAIST 情報科学センターに導入された IBM z890 による仮想サーバシステムの概要を示す。仮想サーバシステムは z890 本体 (CPU 数 2, メモリ 8GB)、DISK 装置 DS6800 (約 2TB)、テープ装置 LTO 3582-L23 から構成される。学内 LAN への接続は 10GBASE-LR および 1000BASE-T インタフェースで接続しており、冗長化構成となっている。IBM z890 システムは、論理パーティション (LPAR)、および、ハイパーバイザ (z/VM) という 2 つの仮想化技術をサポートするため、目的により使い分けできるようにした。LPAR は論理パーティション間で性能面で互いに影響を及ぼさないように設定できるため、安定した運用を目的とした LPAR1、および一時的もしくは実験などでの使用を目的とした LPAR2 の 2 つのシステムに分割し運用することとした。各 LPAR には CPU 数 1、メモリ 2.6GB を割り当て、各パーティションの CPU リソースが空いているときはその空いている CPU リソースを使用することも出来るように設定した。

ハイパーバイザである z/VM は各 LPAR に独立して存在し、LPAR 上に仮想サーバに対応する Virtual Machine (VM) を任意の数だけ作成することがで

きる。VMは互いに独立しており、セキュリティー上また管理上他のVMに依存したり影響を受けることはない。このVM上においてOSとしてLinuxをブートしたものが仮想サーバとして提供される。仮想サーバ間の通信はHyperSocketというメモリ用の高速なバスを介して行われる。また、z/VMにおいては独立して仮想ネットワーク(VSWITCH)を構築することが出来る。今回はそれぞれのLPARに1つのVSWITCHを用意しVM上の仮想的なネットワークを集約するように設定した。

構築した仮想サーバシステムを用いて、実際にサーバの集約が同等のサービスのクオリティを保ちつつ可能となるかの検証を行うために、いくつかの分散したサーバが行っていたサービスを仮想サーバに移行した。現在までに2台のDNSサーバをLPAR1に集約することが完了し、問題なく動いていることを確認した。また、1つのLinuxがブートしている状態で必要なメモリ量を測定した結果、267MBであった。従って、2.6GBが割り当てられている各LPAR上のメモリ上に10程度のLinuxを展開し問題なく稼働すると予想される。

4 仮想サーバ管理支援ツール

4.1 仮想サーバ管理の問題

サーバ仮想化技術により学内の様々なシステムがシステムの独立性を保ちつつ情報科学センターの仮想サーバへと集約することが可能である。しかしながら、仮想サーバシステム管理者以外のユーザが自由に仮想サーバを構築することができないという問題がある。この問題は、従来からの仮想サーバの構築・運用はz890システムの構成を熟知している仮想サーバシステム管理者権限を持つ者が行うことを前提していたため既存の仮想サーバ管理ツールでは仮想サーバシステム管理者権限のみがシステムのリソースを更新可能であるという制限があったことに由来する。しかしながら、仮想サーバシステムを学内に開放し、研究室等のサーバ運用や各種実験のために自由に仮想サーバを構成できる環境として多くのユーザに利用してもらうためには、仮想サーバシステム管理者以外のユーザでも自由に仮想サーバを構築できる必要がある。

管理者権限の問題に加えて、z890システムに標準の仮想サーバ管理ユーザインターフェースであ

るCMSでは仮想サーバの設定・管理には時間がかかるという問題もある。この問題は仮想化を行うほど実リソースと仮想リソースの対応が複雑となり、リソース管理を正確に行うことが難しくなることに由来する。そのため、仮想サーバシステム管理者であっても仮想サーバであるVMを作成しLinuxをインストール設定、起動するまでの行程をミスなく完了するために1時間程度必要とする。従って、インタラクティブなシステムのリソース管理を行いながら仮想サーバの環境を構築するためのデータベースを構築する必要がある。

4.2 Web クローニングツールとクローニング管理サーバ

既存の仮想サーバ管理ツールにおける管理者権限の問題とリソース管理の問題を解消するため、インタラクティブなWebユーザインターフェースの操作だけで簡単に仮想サーバを管理できるWebクローニングツールを構築した。Webクローニングツールは、z890上に構築されるクローニング管理サーバを用いて仮想サーバであるVMの作成(ハードウェアリソースの割り当てを含む)、Linuxのインストール、VMの起動停止、Linux上のネットワークの設定、およびVMの一般ユーザへの割り当て管理を学内全ユーザが使用可能なWebベースのユーザインターフェースにて行う。

図3にz890仮想サーバシステム上の仮想サーバにLinuxを構築するためのフローを示す。仮想サーバであるVMの追加・削除・変更やLinuxのインストールや起動は、ユーザの要求をクローニング管理サーバが受け取ることから始まる。仮想サーバであるVMはクローニング管理サーバがVM管理機構(DIRMAINT)を操作することにより作成される。既存の管理ツールではVM管理機構は仮想サーバシステム管理者のみしかアクセスできなかったが、TCP/IPのソケットをインターフェースとすることにより仮想サーバシステムの管理者以外でもVM管理機構を操作できるように工夫した。これにより、仮想サーバシステムの管理者以外にも仮想サーバを自由に構築可能となる。

Linuxのインストールは仮想サーバを構築した後に始まる。本Webクローニングツールでは事前に雛形となるインストールするLinuxを用意し、新規の仮想サーバにLinux雛形を割り当てるとい

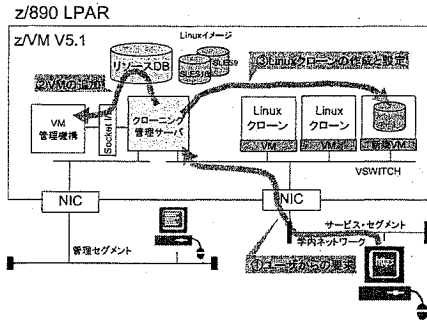


図 3: 仮想サーバ上での Linux 構築

う手法をとる。雛形を用いた DISK クローニング手法は個別に Linux をインストールする方法と比べて Linux 構築の時間を短縮できるという利点がある。管理サーバ上のデータベースを用いてユーザの指定する OS がディスク上のどの領域に存在するかを把握し、DISK クローニング機能により雛形を仮想サーバに割り当てられる予定の DISK 領域へとコピーする。次に、コピーを行った領域を管理サーバがマウントし、ユーザの指定する設定を主に/etc 以下の設定ファイルへと反映させる。最後に、管理サーバは設定を行った領域をアンマウントすることによりインストールは完了し、仮想サーバが Linux をブートすることにより仮想サーバ上の Linux が起動する仕組みである。

4.3 動作評価

Web クローニングツールおよびクローニング管理サーバの作成は日本 IBM (株) に依頼し、z890 上に作成した。クローニング管理サーバを構築する OS には SUSE Linux Enterprise Server 9 を使用し、アプリケーションとして Apache、PHP、MySQL を利用した。クローニング管理サーバには Web ベースのユーザの認証・管理に必要なユーザ情報やハードウェアリソースのデータベースが一元管理される。

図 4 に完成した Web クローニングツールの画面の一部を示す。Web クローニングツールによる仮想サーバ環境のインタラクティブな構築の手順は以下ようになる。はじめに、仮想サーバを利用したいユーザは Web ブラウザから管理サーバにアクセスし、ログインする。ユーザの認証には学内統一されたアカウント (可能であれば電子証明書) を使用する。管理サーバにログインすると仮

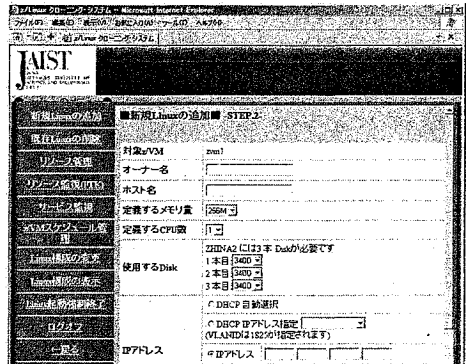


図 4: Web クローニングツール

想サーバの追加・管理、Linux のインストールや Linux の起動・停止がメニュー形式で操作できるトップページが現れる。仮想サーバの追加を選択すると、仮想サーバのホスト名、使用する CPU 数、メモリ量、DISK、IP アドレスや VLANID をインタラクティブに指定することにより、仮想サーバが構築される。最後に Linux の起動を選択し、仮想サーバ上の Linux を稼動する。

今回構築した Web クローニングツールでは、仮想サーバシステム管理者権限を持つ者が管理サーバにアクセスした際に表示されるトップページで選択できる内容に新規 Linux の追加、既存 Linux の削除の項目が追加されるようにした。この機能を用いて、仮想サーバシステム管理者権限を持つ者が Linux の雛形を管理することにより仮想サーバシステム管理者権限以外の利用者がクローニングすることのできる OS の雛形を管理することができる。

Web クローニングツールの効果の評価として、VM の作成から Linux をインストールしネットワークの設定までを行う時間を測定した。Web クローニングツールを用いない場合、およそ 20 項目にも及ぶ設定を手動で行う必要があり、この一連の作業にかかる時間は約 1 時間を要した。しかし、Web クローニングツールを用いた場合、必要な項目は VM の追加と Linux のクローニングの 2 つに削減され、かつ、Web 上の入力により実際のコマンド操作が自動化されるため、作業時間は約 10 分間となった。また、仮想化により煩雑となった物理的なリソースとの対応関係もリソース管理のデータベースにより正確に把握可能となるため、設定時のミスを低減することも可能となる。従って、仮想サーバ管理支援ツールにより、仮想サーバシス

テム管理者以外のユーザであっても短時間で容易に仮想サーバ環境の構築が可能となることから、今回構築した Web クローニングツールの有用性は極めて高いといえる。

5 関連研究

近年、新しい計算機システムの構築運用技術として仮想化技術が多く研究者により研究されている。金田らは単一システムイメージを提供するための仮想マシンモニタを構築し計算機クラスタを効率的に運用しようということを試みている [11]。しかしながら、分散されたサーバを集約するための手段としての仮想化技術に着目している本論文とは目的が異なる。

丸山らは [12] 大規模教育用計算機システムにおいて仮想化技術を用いて Windows と Unix 系の OS が混在する環境を構築した。しかしながら、教育用計算機システムの仮想化技術は、性能やセキュリティの維持の点で本論文のような基幹サービスを提供するサーバの仮想化に適応できないと考えられる。

今回構築した仮想サーバ管理支援ツールは、既存のシステムにおける管理ツールでは不可能であった仮想サーバシステム管理者以外のユーザが自由に仮想サーバを構築可能とする点で独創的な研究であり、仮想サーバシステムの適応領域をいっそう広げることが期待されている。

6 まとめ

本論文では、学内に分散するサーバを情報科学センターの仮想サーバシステムに集約しコストの大幅な削減や信頼性を達成しつつユーザの利便性を維持することを目的として仮想サーバ管理支援ツールを構築した。仮想サーバ管理支援ツールである Web クローニングツールは仮想サーバシステム管理者以外でも自由に仮想サーバを構築できる点、また、仮想化されるリソースの管理が容易になる点で仮想化システムの管理コストの低減に有効である。

今後の課題として、現在、試験運用中の仮想サーバシステムを実運用させて評価することが挙げられる。導入から運用にいたるまでのコスト削減を分析し、仮想サーバシステムへの集約によりコストの削減をどの程度達成したか検証していく予定

である。さらに、実運用を通して仮想化に適する計算機サービスの分析等も行う必要があると考えている。

謝辞

本仮想サーバシステムの構築および「Web クローニングツール」を作成いただいた日本 IBM (株) ならびに (株) 理経の担当の方々に感謝の意を表す。

参考文献

- [1] 敷田幹文, 井口寧, 丹康雄, 松澤照男. 大規模分散システムの集中運用管理における効率化技術の提案. 情報処理学会 分散システム/インターネット運用技術シンポジウム, pp. 75-80, Feb. 1999.
- [2] 敷田幹文, 井口寧, 三輪信介, 丹康雄, 松澤照男. 大規模高可用性サーバの設計と運用. 情報処理学会 分散システム/インターネット運用技術シンポジウム, pp. 57-62, Feb. 2001.
- [3] 敷田幹文, 井口寧, 藤枝和宏, 松澤照男. 高可用性システム統合監視機構の提案. 情報処理学会 分散システム/インターネット運用技術シンポジウム, pp. 51-56, 2002. DSM2002.
- [4] James E. Smith and Ravi Nain. *Virtual Machines: Versatile Platforms For Systems And Processes*. The Morgan Kaufmann Publishers, 2005.
- [5] Renato Figueiredo, Peter A. Dinda, and JosAc Fortes. Resource virtualization renaissance. *Computer*, Vol. 38, No. 05, pp. 28-31, 2005.
- [6] 大町一彦. 仮想マシン道しるべ: 仮想マシン草創期. 情報処理学会誌, Vol. 48, No. 8, pp. 903-905, 2007.
- [7] G.R.McClain 編, 川添良幸, 早川美徳, 小野木之隆監訳. *VM アプリケーションハンドブック*. 共立出版, 1992.
- [8] IBM System z. <http://www-06.ibm.com/systems/jp/z>.
- [9] Sun Microsystems. Sun enterprise 10000: Dynamic systems domains. Technical report, Technical White Paper, 1999.
- [10] T. L. Borden, J. P. Hennessy, and J. W. Rymarczyk. Multiple operating systems on one processor complex. *IBM Systems Journal*, Vol. 28, No. 1, pp. 104-123, 1989.
- [11] 金田憲二, 大山恵弘, 米澤明憲. 単一システムイメージを提供するための仮想マシンモニタ. 情報処理学会論文誌 ACS 13, pp. 27-39, 2006.
- [12] 丸山伸, 最田健一, 小塚真啓, 石橋由子, 池田心, 森幹彦, 喜多一. Virtual machine を活用した大規模教育用計算機システムの構築技術と考察. 情報処理学会論文誌, Vol. 46, No. 4, pp. 949-964, 2005.