MiFにおける仮想化技術の利用

Implementation of Virtualization Technology for MiF

立　花　幸　治，浅　井　保　行

要　約　日本ユニシスは，サービス型ビジネスのiDC基盤に対する顧客の要求は提供のスピードとコスト，そして運用だと考えている．これらの要求を実現するために統合化，仮想化，自動化した次世代iDC基盤がMiF（Modeled iDC Farm）である．

従来の基盤システム設計，調達，構築では1ヶ月以上の時間を要していたが，MiFでは最短5日で基盤システムを提供できるようになった．

短期間でのリソース提供を可能にするため，MiFを構成するサーバ，ストレージでは物理層と論理層の分離による仮想化の技術を採用することで統合化，仮想化を実現している．また，ストレージにおいてはシン・プロビジョニングによってリソースを効率的に利用する．

ネットワークにおいては，L2レベルでのVLAN，L3レベルでのルータ，ファイアウォール，L4レベルでのロードバランサーの仮想化機能を採用することで統合化，仮想化を実現している．

Abstract　Nihon Unisys believes the requirements of customers for iDC infrastructure of service business are speedy deployment, cost and operation. MiF (Modeled iDC Farm) as the next-generation iDC infrastructure is built by the integration, virtualization, and automation technology to meet those requirements.

Although it took over a month to perform the infrastructure system design, procurement, and implementation in traditional model, MiF makes the infrastructure system operational in 5 days. To facilitate to provide computing resources short period of time, servers and storages of MiF are implemented through the integration and virtualization by separating physical and logical layers. In addition, the storage resource is effectively used by the thin provisioning.

As for the networking, the integration and virtualization is implemented by adopting virtualization functions of VLAN in layer 2, the router in layer 3, the firewall in layer 4, and the load balancer in layer 4.

1. はじめに

ITインフラにおける仮想化技術は，必要なときに必要なだけ，かつ短期間にコンピューティングリソースを提供するというオンデマンドコンピューティングを実現するための要素技術であるだけでなく，リソースの余剰を削減して利用効率を上げ，消費電力や設置面積を削減してグリーンITを推進するといった，今後のIT環境における重要な技術である．

本稿では，短期間での基盤システム設計，調達，構築を可能とする日本ユニシスの次世代iDC基盤“MiF（Modeled iDC Farm）”における仮想化技術に関し，2章ではサーバとストレージの，3章ではネットワークの実現方法について解説する．

2. MiFにおけるサーバとストレージの『仮想化』

この章では，MiFのコンセプトのひとつである『仮想化』で，「サーバ」と「ストレージ」

を要素として，どのような仮想化技術を用いて MiF の目的を実現しているかを説明する．

2. 1　MiF で必要なサーバとストレージの仮想化技術

MiF の目的のひとつは，スピーディにプロビジョニングで必要なリソースを切り出すことである．この節では，MiF の物理インフラ群を構成する「サーバ」および「ストレージ」で必要な仮想化技術について説明する．

2. 1. 1　サーバ，ストレージのリソースとは

サーバおよびストレージにおけるリソースとは，ハードウェア・リソースのことである．「サーバ」で言えば，CPU のコア数であり，搭載メモリー量であり，NIC の数である．「ストレージ」で言えば，容量である．ハードウェア・リソースはソフトウェアと違い物理的なものであり，すでに固定されているものでもある．あるシステムを構築するうえで，システム内の各役割を担うサーバは，使用する用途（例えば Web サーバとして使用されるのかデータベース・サーバとして使用されるのか）で CPU のコア数やメモリー量などが異なる．また，使用するソフトウェア，OS やデータベースの種類などでも要求されるサーバのリソースは異なる．システム規模の大小によってデータベースの大きさが異なるので，使用するストレージの容量も異なる．このように，通常システムを構築しようとすると，そのシステムの規模，使用するソフトウェアなどでサーバおよびストレージのリソースを決定していくが，要求される要素が変わるとサーバやストレージに「必要な」リソースも変わってしまう．そのため，「スピーディ」に，リソースを確保することができない．

2. 1. 2　サーバ，ストレージの仮想化技術 ― 物理層と論理層

MiF のコンセプトのひとつである「スピーディにプロビジョニングで必要なリソースを切り出す」ことを実現するためには，物理層と論理層の分離が必要であり，分離するためにサーバおよびストレージの仮想化技術を用いることが必要となる．

今までの，あるシステムのリソースを決める工程では，そのシステムが要求するサーバ，ストレージのリソースを，ハードウェア・リソースと捉えていたため，そのリソース要求に合ったハードウェアを選択する必要があった．また，そのハードウェア・リソースがコスト面から妥当であるかどうかの判断も必要であった．

MiF は，今までのハードウェア・リソースとシステムが要求するリソースを分離して考えることで，「スピーディにプロビジョニングで必要なリソースを切り出す」ことを実現している．MiF では，ハードウェア・リソースを「物理層」と考え，システムが要求するリソースを「論理層」として捉えている．

2. 2　MiF におけるサーバでの仮想化

MiF で「物理層」と「論理層」に分けることを具体的にサーバではどのように実現しているかを述べる．

2. 2. 1　サーバの仮想化 ― 共有リソース・プール

今までのサーバの仮想化とは，図 1 に示すように一般的に物理サーバを統合することで説明

されることが多い．企業でサーバを仮想化する大きな理由のひとつとして，現在あるサーバの利用率（CPU 使用率など）に無駄なサーバ・リソースがあるとの観点から，サーバを仮想化して統合することでサーバの利用率を向上させ，サーバの台数を減らし，コストを低減させることがあるからである．

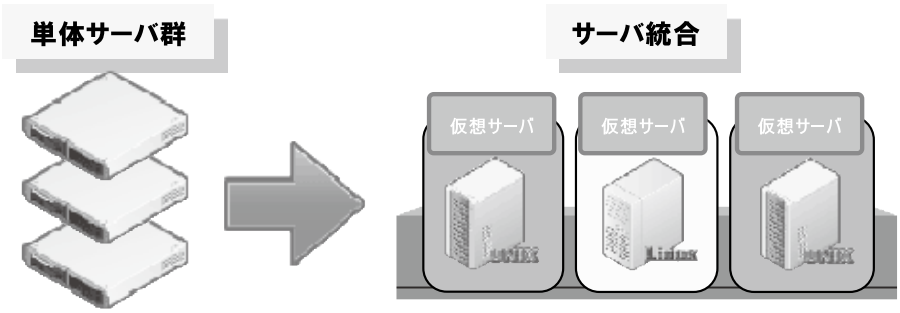

図 1　今までの仮想化

それに対し MiF におけるサーバの仮想化とは，図 2 に示すように物理サーバのハードウェア・リソースを論理的なインフラ・リソースに置き換えることである．それによって，個々の物理サーバのハードウェア・リソースは，物理サーバに依存する形から共有インフラ・リソースとなり，必要なときに必要なリソースをスピーディに切り出して，サーバを提供できるようになる．

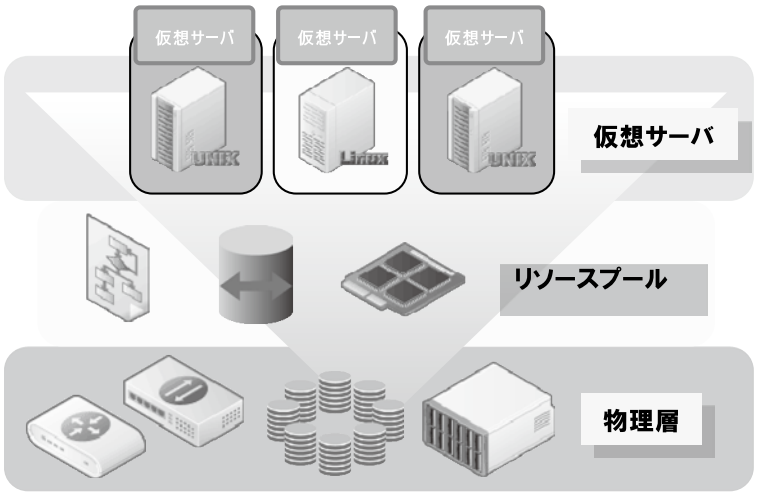

図 2　MiF の仮想化

MiF は，必要なサーバを 24 時間で提供することを目標にしている．そのために，必要なリソースを常にリソース・プールとして保持している．

リソース・プールは，それぞれが同じスペックを持つ物理サーバで構成されている．それにより，大きな単一のリソース・プールを作成することができ，プロビジョニングで必要なリソースを容易にリソース・プールから切り出すことが可能となった．

また，サーバを仮想化することで，物理サーバのハードウェア・リソースが分かれて論理的なリソースとなり，リソース配分も自由となる．これにより，最初はCPUコア，メモリーの少ないサーバから使用を始めて，不足していれば容易にリソースを追加することが可能となる．

2.2.2　サーバの仮想化 ― 障害対応

MiFでのサーバの障害対応について述べる．物理サーバで構成されている場合，障害対応としてハードウェア障害とソフトウェア障害の両方を合わせて考慮する必要があるが，サーバを仮想化した場合は，障害対応もハードウェア障害とソフトウェア障害を切り離して考えることができる．

障害への対応としては一般的にはクラスタシステムを採用することになる．クラスタシステムはホットスタンバイ形態であるため，サービスを実施していない待機（スタンバイ）状態のサーバが必要となる．これは，サービスとして利用していないリソースが存在することとなり，仮想化技術によりリソースの余剰を削減して利用効率を上げることにはならない．MiFでは，ハードウェア障害はN＋1での自動切り替えを採用し，サーバのソフトウェア障害はクラスタシステムのようなホットスタンバイの形態を採用しなくても，仮想サーバの再起動だけで十分に対応できると考えている．

クラスタシステムに比べると，復旧時間はかかるが，コスト面では待機サーバを用意する必要がないので有利である．

2.3　MiFにおけるストレージでの仮想化

本節では，MiFの「仮想化」コンセプトをストレージの観点でどのように実装したかについて述べる．

2.3.1　ストレージの仮想化 ― 物理層＝ハードウェアRAID構成層と論理層＝データ層

ストレージでの物理層は実際のハードウェアRAID構成層であり，論理層はその上のデータ層となる．使用者から見ると，管理したいのはデータであって，ハードウェアRAID構成ではない．ところが，一般的なストレージ・システムでは，データ管理＝ストレージ管理となっており，物理層と論理層，ハードウェアRAID構成層とデータ層が切り離されていない．

一般的なストレージ・システムでは，データ領域はハードウェアRAID構成を構築し，そのハードウェアRAID構成の容量がデータ容量となる．使用者は，データ領域を必要とすると，そのデータ容量と同じかそれ以上の容量を持つようにハードウェアRAIDを構成する作業を行う．作業としては，使用する物理ディスクを選択し，RAID構成を作成し，そのRAIDの初期化を行う．これでは，スピーディなリソース変更の要求に応えることはできない．また，物理的なディスクの容量に依存するため必要以上の容量を持つことにもなる．

MiFにおけるストレージでは，図3に示すように物理層と論理層を切り離し，使用者には論理層からデータ領域のみを切り出して提供することで仮想化を実装している．

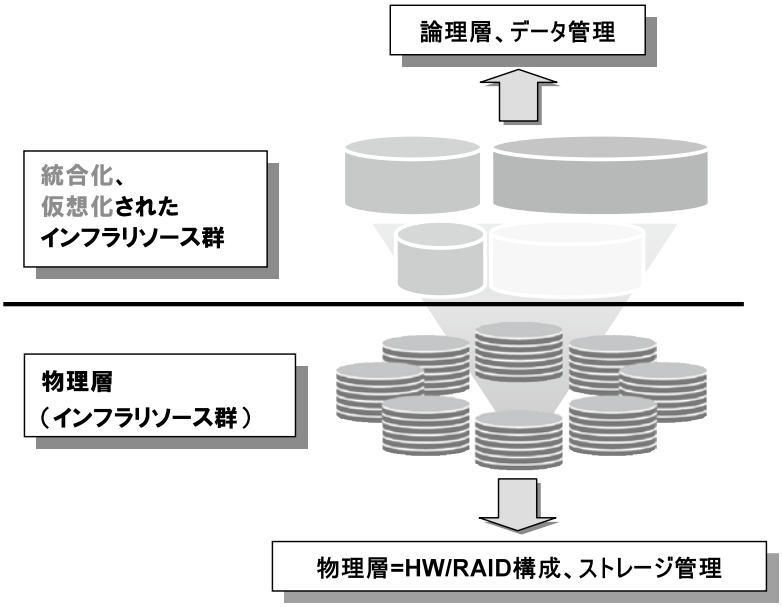

図 3　物理層と論理層

使用者が必要とするリソースは，データを格納する領域である．ストレージから見ると，そのデータを格納する領域はボリュームである．使用者が要求するのは，そのデータを格納する領域の容量である．つまり，使用者が要求するのはストレージのボリュームの容量である．ストレージに要求されるプロビジョニングは，使用者に必要とされるデータ容量をスピーディに切り出すことである．

物理層と論理層を分けることで，データ管理をストレージ管理から切り離すことができ，使用者が必要なデータ領域を切り出すことができる．このデータ領域は，データの容量を表しており，それはボリュームのサイズとなる．プロビジョニングにおいては，このボリュームのサイズのみを考慮すればよいことになり，すぐにデータ領域を切り出してサービスを提供することができる．また，仮想化することによりデータ容量の増量についても即座に対応可能となる．そして，データ容量を減らすこともでき，非常にフレシキブルなプロビジョニングを可能にしてくれる．

2. 3. 2　ストレージの仮想化 ― ハードウェア障害

現在のストレージでは，RAID システムを採用することで可用性を確保している．MiF のストレージでは，可用性の観点からも物理層と論理層を切り離すことで，ディスク障害が発生しても論理層に影響しないようになっている．通常の RAID システムは，可用性・冗長性とコスト面からパリティ付きストライピングの RAID4，もしくは RAID5 が採用されることが多い．RAID4，または RAID5 では 1 本のディスク障害には対処できるが，2 本同時の障害には対処できない．MiF のストレージでは，最新の RAID システムである RAID6 を実装することで 2 本同時の障害でも論理層に影響しないようになっている．つまり，同時に 2 本のディスクに障害が起きても稼働できる RAID 構成にすることにより，物理層のハードウェア RAID で障害が発生しても，論理層には一切の影響は無く，業務を継続することが可能である．

2.3.3　ストレージの仮想化 ― シン・プロビジョニング

仮想化されていないストレージの場合，そのハードウェア RAID 構成の容量がデータ容量となるため，余分に容量を持つ必要が生じる．ストレージの仮想化により，データ領域である論理層が物理層＝ハードウェア RAID 構成から分かれるため，必要なデータ領域，容量のみを切り出すことが可能となる．また，仮想化技術によって「シン・プロビジョニング」と呼ばれる機能を実装することで，さらにストレージの利用効率を上げることが可能となる．

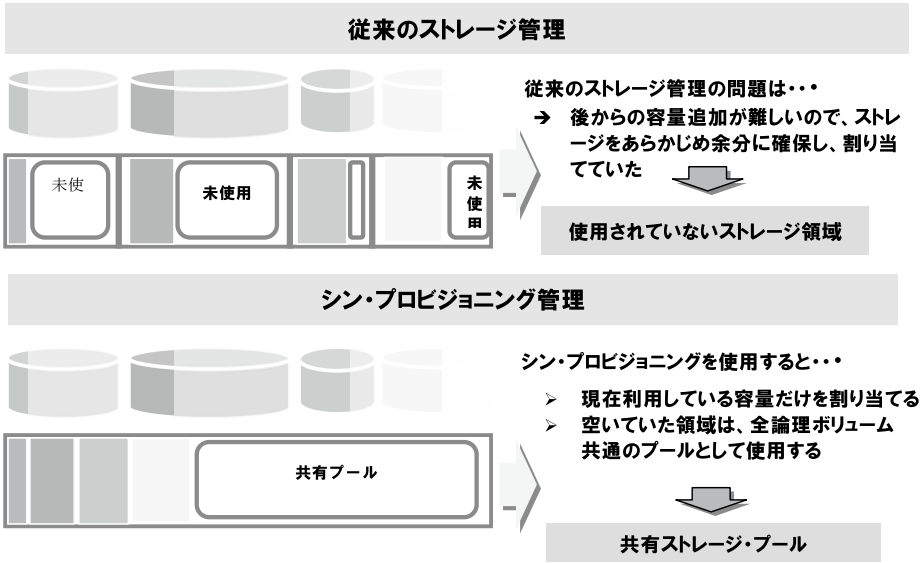

図 4　シン・プロビジョニングの説明

「シン・プロビジョニング」とは，図 4 に示すように，各システムで使用するデータ領域に対し，それぞれが現在利用している容量だけを割り当てるものである．物理的なストレージ・システム上で，どの論理ボリュームにもまだ割り当てられていないストレージ容量を空き容量としてプールし，必要に応じて各システムに割り当てる機能である．

シン・プロビジョニング機能を実装することで，従来のようにシステム単位でストレージ容量を多めに見積りあらかじめ確保しておくことから生じる無駄は，ほとんど完全に回避できるようになる．無駄なストレージ領域を無くして，必要な容量を確保していくことができる．

ストレージの容量が「共有ストレージ・プール」となり，ストレージ容量を意識することなく，使用者が必要とするデータ容量をスピーディに切り出すことが可能となる．

3.　ネットワーク仮想化の考え方

ネットワークの仮想化はサーバとは逆に，一台の物理デバイス上で複数の仮想デバイスを構成する．この技術により，物理デバイスを用意することなく，顧客の要望に応じたネットワーク構成を短期間に提供することが可能となる．本章では，ネットワークにおける仮想化技術と仮想化ネットワークの設計手法について説明する．

3. 1　ネットワークの仮想化技術

現在実用化されている代表的なネットワーク仮想化技術について以下に紹介する．

3. 1. 1　ルータ

ルータにおける仮想化技術は，VRF（Virtual Routing Forwarding）という技術が一般的である．これはルーティングテーブルを複数保持できる機能を指し，一台のルータで IP アドレスが重複したルーティングテーブルを保持できる．この技術を採用することにより，一本の物理回線で複数のネットワークを互いの IP アドレスを意識することなく共用することができる．通信事業者の IP-VPN（InternetProtocol-VirtualPrivateNetwork）網において利用されている場合がある．

3. 1. 2　スイッチ

スイッチにおける仮想化技術は VLAN（Virtual LAN）が代表的であるが，スイッチを仮想デバイスに分割して L2（Layer2）および L3（Layer3）機能を互いの仮想デバイスに影響することなく実現する技術や，複数のスイッチを一つの仮想スイッチとして扱う技術が実用化されている．スイッチを分割する場合は，一台の仮想スイッチの VLAN，ルーティング，セキュリティなどの設定情報やトラフィックが他の仮想スイッチに全く影響しないという利点がある．一方，複数スイッチをマージする場合は，例えば二台の物理スイッチ間でのチーミング（複数のインターフェイスを一つの論理インターフェイスとして扱う）が可能となるという利点がある．

3. 1. 3　ファイアウォール

仮想ファイアウォールは，一台の物理デバイスで複数のファイアウォールポリシーを保持し，互いのポリシーに影響を与えず独立して動作する．サーバセグメントなどとは VLAN で接続して仮想ネットワークを構成する．

3. 1. 4　ロードバランサー

ロードバランサーには回線負荷分散をするリンクロードバランサーとサーバへのトラフィックを負荷分散するサーバロードバランサーがあるが，本稿ではサーバロードバランサーを対象とする．仮想ロードバランサーでは，一台の物理デバイスで複数のロードバランスポリシーを保持し，互いのポリシーに影響を与えず独立して動作する．

サーバセグメントなどとは VLAN で接続して仮想ネットワークを構成する．

3. 1. 5　VPN（Virtual Private Network）

本稿では，VPN はインターネットなどのパブリックネットワーク上で，企業内ネットワークなどのプライベートネットワークをセキュアに構築する技術を指す．通信経路は暗号化され，プライベート IP アドレスでの通信が可能となる．VPN 終端装置では，アクティブ・スタンバイの HA（High Availability）構成だけでなく，アクティブ・アクティブのクラスタ構成にも対応している機器もある．

3.2　物理ネットワーク設計と仮想ネットワーク設計

仮想化ネットワークを設計する場合は，物理ネットワーク設計と仮想ネットワーク設計に分けて考える必要がある．検討に際しては，仮想ネットワーク設計から考えると分かりやすい．仮想ネットワーク設計は，通常のネットワーク設計における物理ネットワーク設計をイメージすればよいからである．図5に仮想ネットワーク構成図を示す．

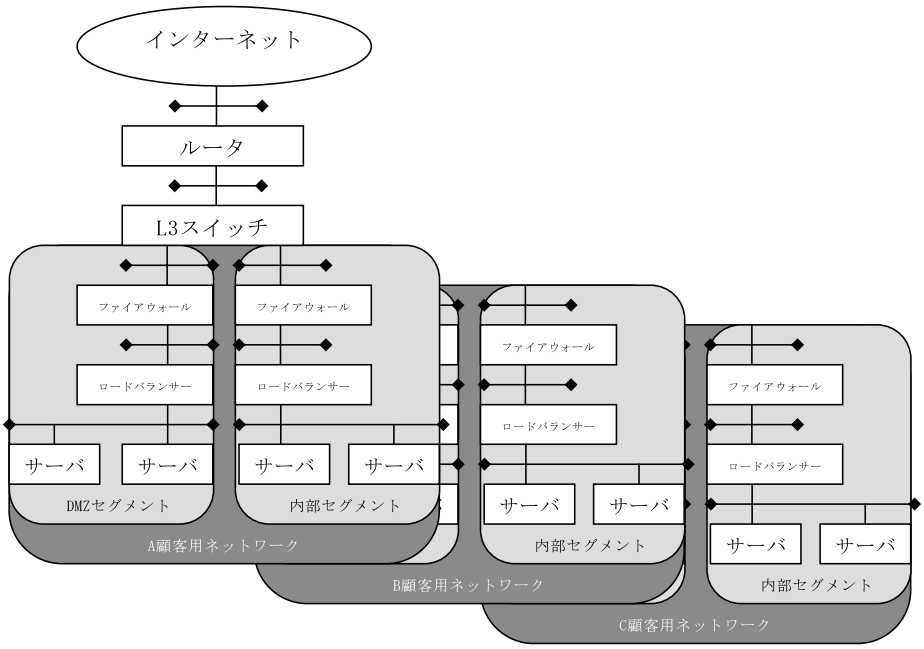

図5　仮想ネットワーク構成図

図5は物理ネットワーク構成図のように見えるが，仮想ネットワーク構成図である．仮想化コンポーネントは基本的に独立して動作し，物理デバイスと機能も同等なので通常のネットワーク設計をすればよい．ファイアウォール，DMZセグメントおよび内部セグメントを構成するロードバランサー，サーバは仮想デバイスであり，ホスティング顧客毎に用意する．

図5の構成は，イーサネットケーブルはVLANに置き換えられており，ファイアウォール，ロードバランサー，サーバも全て仮想デバイスである．

物理ネットワーク設計においては，帯域と拡張性に注意が必要である．サーバ，ファイアウォール，ロードバランサーといったリソースをハードウェアに集約するため，物理デバイスがボトルネックになりやすい．MiFでは，物理的なサーバおよびネットワーク機器の接続は，全て複数の10GEthernetで接続した．サーバ収容スイッチには，4Tbpsのスループット，100ポート超の10GEthernetを収容可能な高性能機種を選択した．仮想デバイス（ファイアウォールとロードバランサー）を収容するスイッチにおいても，ハードウェアが追加可能なスロット数と高スループットを有する機種を選択した．

3.3　外部接続

MiFと外部環境との接続は大きく二種類に分けられる．一つは，他拠点と通信事業者の回線で接続する形態やデータセンター内での構内回線接続といった回線接続型である．今回

MiF で構築したネットワークは SaaS（Software as a Service）やホスティングといった共用型サービス基盤であるため，インターネット接続については BGP（Border Gateway Protocol）を利用した複数 ISP（Internet Services Provider）とのマルチホーム環境とし，企業内イントラとの接続には顧客専用ファイアウォール（外部接続ファイアウォール）を介した接続インターフェースを用意している．企業内イントラとの接続の場合はプライベート IP アドレスが競合する可能性があるので，外部接続ファイアウォールにて NAT している．

もう一つは，仮想化非対応機器との接続である．仮想ネットワークの構築は，基本的に仮想化対応機器のみで構成されるため，それ以外の機器との接続は一種の外部接続と考えたほうがよい．BOX 型アプライアンスや VOIP ゲートウェイなどについては外部接続用 L2 スイッチを介してポート VLAN で収容することで仮想ネットワークの VLAN に接続する．

3.4　セキュリティ

セキュリティおいては，セグメント構成と IPS 等のセキュリティ機器の接続について検討した．セグメント構成については，顧客がアクセス可能なサービス側と MiF 基盤運用管理システムのみアクセス可能な MiF 運用管理セグメントという大きく二つのゾーンに分けている．サービス側では DMZ セグメントと内部セグメントを提供している．

MiF 運用管理セグメントは MiF 環境全体の管理セグメントのため，サービス側からは切り離している．この MiF 運用管理セグメントには監視システム，自動運用システム，セキュリティシステムなど，MiF 基盤としての統合管理サーバのみとの通信を許可し，MiF 運用管理セグメントを通して他の顧客のホスティングサーバ間がアクセスできない仕様にしている．

セキュリティ機器については，インライン型の IPS やアンチウィルスアプライアンスを想定した接続環境を用意している．ワンアーム型の Web セキュリティゲートウェイやメールセキュリティゲートウェイなどは，仮想化非対応機器との接続の考え方で収容可能である．

3.5　運用性

セグメント設計，ファイアウォールのポリシー，ロードバランサーのポリシーなど，仮想ネットワーク設計は一般的なネットワークの知識があれば可能であるが，仮想デバイスへの実装は特別なスキルを要する．安全，迅速なプロビジョニングも今回の MiF 基盤構築においては重要な要件であったため，運用自動化ツールを最大限活用した．具体的には，プロビジョニングの構成をテンプレート化し，実機への設定コマンドをスクリプト化し，さらにコマンドに対するパラメータは XML で定義した．これによって特別なスキルを持ったネットワークエンジニアでなくとも仮想環境におけるサーバおよびネットワークのプロビジョニングが可能となった．このテンプレートとスクリプトを増やしていくことで，全てのプロビジョニングを自動化することを目指している．

4.　お　わ　り　に

今回の取り組みにより，通常は最低 1 ヶ月以上を要する基盤システム構築を数日間で実現した．また，物理デバイス数を劇的に削減できたため，グリーン効果も非常に大きい．今後，他のネットワーク機能やセキュリティ機能が仮想化対応していく事が予想されるため，積極的に新技術を取り込んでいく計画である．

現在の MiF では，必要なサーバを 24 時間で提供することを目標としたため，サーバに関しては同じスペックの物理サーバを用意した．ただ，顧客のニーズも多様化しているため，今後は異なるスペック（高いスペック，低いスペック）の物理サーバでも「スピーディにプロビジョニングで必要なリソースを切り出す」ことが必要となる．また，ストレージとしては圧縮技術の採用などで，より利用効率を上げることが課題である．加えて，IPv6 や NGN（Next Generation Network）といった新しいネットワーク環境の具体的な利用も早急に取り組むべき課題である．

参考文献
［1］ WIKIPEDIA，http://wikipedia.org/
［2］ IETF，http://www.ietf.org/
［3］ Cisco Systems 合同会社，http://www.cisco.com/jp/
［4］ NTT グループ，http://www.ntt.co.jp/
［5］ ネットアップ株式会社，http://www.netapp.com/jp/

執筆者紹介

立　花　幸　治（Koji Tachibana）
1985 年青山学院大学経営学部卒業．同年日本ユニバック（株）入社．ユニバック系 OS の導入・提供およびサポートに従事．2001 年ユニアデックスに転籍．2002 年よりストレージ設計，構築に従事．現在 ICT サービス本部に所属．

浅　井　保　行（Yasuyuki Asai）
1992 年東京電機大学工学部卒業．同年日本ユニシス（株）入社．ネットワーク設計，構築に従事．2001 年ユニアデックスに転籍．現在 ICT サービス本部に所属．