

サーバの仮想化とネットワークの動的構成変更による データセンタ省電力運用管理手順の実装

荒井大輔 吉原 貴仁 井戸上 彰

株式会社 KDDI 研究所

1. はじめに

近年のトラフィック量増大に伴い、IT 機器の数も増加し、その消費電力量が 2025 年には 2006 年比で 5 倍になると推計[1]されている。そのため、IT 機器の消費電力量の削減は、トラフィック量の増大を支える上で、重要な課題である。IT 機器全体の消費電力量のうち、とくに、ルータやスイッチなどのネットワーク機器(以下、ネットワークを NW、ネットワーク機器を NW 機器と呼ぶ)とサーバが約半分を占めており[1]、それらが集中して設置されるデータセンタ(以下、DC と呼ぶ)の省電力化は急務である。

これまでに著者らは、サーバの仮想化と、NW の動的構成変更による DC 省電力運用管理手順を提案[2]している。本稿では、提案手順の実装概要を述べる。

2. サーバの仮想化と NW の動的構成変更による DC 省電力運用管理手順[2]

2.1. 想定環境

複数の NW セグメントに配置された NW 機器やサーバと、それらの安定稼働のための空調機が稼働する DC を想定する。また、サーバは仮想化により、物理的なサーバ(以下、物理マシンと呼ぶ)の CPU やメモリ、NW インタフェースなどの資源を論理的に分割し、1 台以上の仮想的なサーバ(以下、仮想マシンと呼ぶ)に、分割した資源を割当て、1 台の物理マシン上で複数の仮想マシンを同時に実行可能とする。仮想マシンによる資源の使用率は、利用者からの要求に応じて変化し、その要求は時々刻々と変化する。

2.2. DC 省電力運用管理手順[2]の概要と処理手順

本手順では、資源の使用率に応じて、1 台の物理マシン上で実行する仮想マシンの数と、それら仮想マシンを実行する物理マシンを NW セグメントの制限なく変更し、省電力化を図る。このため、省電力管理サーバを導入し、仮想マシンに割当てた資源の使用率を定期監視する。さらに、省電力管理サーバは、監視結果に基づき、次の 2 つの手順を NW セグメントの数だけ繰り返し実行する。(手順 1)1 つの NW セグメント(例えば、図 1 中 NW セグメント 1)に注目し、1 台の物理マシン上で実行する仮想マシンの数を変更する。

例えば、休日夜間など、資源の使用率が低くなる場合(図 1 左から右)、1 台の物理マシンの資源を分割し、より多くの仮想マシンが 1 台の物理マシン上(仮想マシン 1 と 2 を物理マシン 1)で実行できるように変更する。

(手順 2)DC 全体に注目し、NW の構成変更により、NW セグメントにある物理マシンを変更する。

従来、利用者との通信を維持するため、手順 1 の変更は、同一の NW セグメントにある物理マシンの資源間に限られていた。本手順では、手順 1 に加え、NW 構成を変更(例えば、図 1 中スイッチ 1 の VLAN 設定の変更により、NW セグメント 2 に物理マシン 2 があるように)することにより、この制限を無くしている。

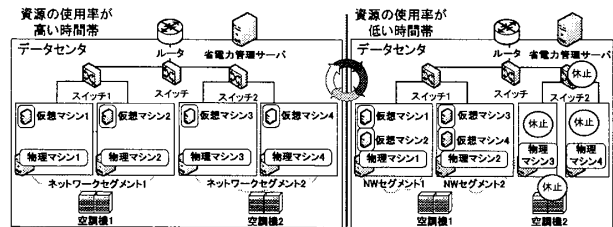


図 1: DC 省電力運用管理手順[2]の概要

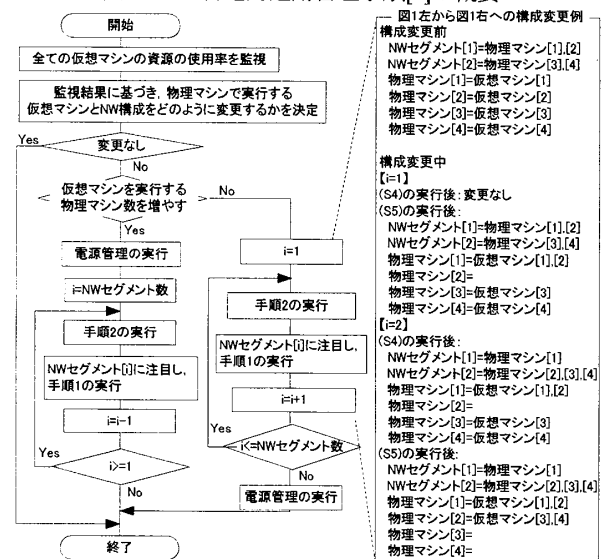


図 2: 省電力管理サーバの処理手順

さらに、省電力管理サーバは、手順 1 と手順 2 に加え、物理マシンと NW 機器、空調機の電源を管理する。仮想マシンの実行に不要となった物理マシン(図 1 左から右の例では、物理マシン 3 と 4)がある場合、物理マシンと、それら物理マシンにのみ接続された NW 機器(スイッチ 2)と、それらを冷却する空調機(空調機 2)を休止する。逆に休止中の機器で、仮想マシンの実行に必要な機器がある場合(図 1 右から左の例では、物理マシン 3 と 4、スイッチ 2、空調機 2)、それらを起動する。

以上の手順により、資源の使用率に応じて、物理マシンに加え、スイッチと空調機の電源管理(利用率が低い時間帯におけるスイッチ 2 と空調機 2 の休止)を可能とすることで、本手順は DC の省電力化を実現する。

省電力管理サーバの処理手順を図 2 左に、資源の使用率が高い時間帯から低い時間帯の構成に変更する場合を例に、NW セグメントと物理マシン、仮想マシンの状態の時系列変化を図 2 右にそれぞれ示す。

3. DC 省電力運用管理手順の実装

提案手順の有効性評価のため、図 3 に示す環境を構築し、省電力管理サーバを実装する。サーバの仮想化に、VMware 社の VMware ESX Server 3.5 と、VMware ESX Server 3.5 上で実行される仮想マシンの一元管理、運用機

能を提供する VMware Virtual Center Management Server (VCMSV)を用いる。物理マシン上で実行する仮想マシン数を変更するには、VMware ESX Server 3.5の仕様により、外部ストレージが必要となる。このため、レイヤ 2 スイッチを介して各物理マシンと接続する。NW 機器と空調機の電源休止については今後実装予定である。

3.1. 省電力管理サーバのソフトウェア構成

まず、次の 5 つの情報をそれぞれ管理するテーブルを省電力管理サーバに設け、運用者等が値を事前登録する。

- (1)NW 機器情報(NW 機器毎)：IP アドレス、ログイン ID、パスワードの 3 組(図 3(T1))。
- (2)サーバ情報(物理マシン毎)：仮想マシンの割当て(後述の機能 3)に用いる優先度、自身が属する NW セグメント、自身の上で実行中の仮想マシンの 3 組(図 3(T2))。
- (3)トポロジ情報(NW 機器毎)：NW 機器と物理マシンの接続関係(図 3(T3))。
- (4)割当率(DC に 1 つ)：仮想マシンに割当てる資源の量の算出(後述の機能 4)に用いる(図 3(T4))。
- (5)更新間隔(DC に 1 つ)：提案手順の繰返し周期に用いる(図 3(T5))。

次いで、図 2 の処理手順を実現するため、以下の 7 つの機能を実装する。機能 1, 5, および 6 の実装には、VCMSV が提供する管理インターフェースを用いる。

機能 1：資源使用率把握機能(図 3(F1))。VCMSV に各仮想マシンの、CPU、メモリ、NW インタフェースの資源毎の最新の使用率(%)を要求し、取得する。

機能 2：割当て資源量算出機能(図 3(F2))。機能 1 により取得した使用率を受取り、各仮想マシンに割当てる資源の量(使用率の上昇に備えた余剰分を含む)を、資源毎に次式により算出する。“仮想マシンに割当てられた資源量×使用率(機能 1 の結果)×割当率(図 3(T4))”。

機能 3：物理マシン割当て決定機能(図 3(F3))。機能 2 により算出した各仮想マシンの資源毎の割当て量を受取り、仮想マシンをどの物理マシンの資源に割当てるかを決定する。決定の際、サーバ情報(図 3(T2))に含まれる優先度に従い、次の 2 つの条件を満たす限り、仮想マシンを順に割当てることにより決定する。

(条件 1)資源毎に仮想マシンに割当ててる資源の量を求め、全ての資源について、物理マシンの資源の量以下であること。

(条件 2)異なる NW セグメント上で実行していた 2 つの仮想マシンは異なる物理マシンに割当てること。

機能 4：NW 構成変更機能(図 3(F4))。図 2 の手順 2 に対応する。機能 3 で決定した割当てを受取り、NW 構成の変更を行う。NW 構成の変更は、1 つの NW セグメントに注目し、異なる NW セグメントにある物理マシンに、当該 NW セグメントで実行していた仮想マシンを割当ての際、次の 3 つのステップにより、上記物理マシンが当該 NW セグメントに含まれるように NW 構成を変更する。

- (1)トポロジ情報(図 3(T3))から物理マシンが接続されたスイッチを求める。
- (2)NW 機器情報(図 3(T1))から当該スイッチのログイン ID とパスワードを得る。
- (3)当該スイッチにログインし、VLAN 設定を変更する。また、サーバ情報(図 3(T2))を更新する。

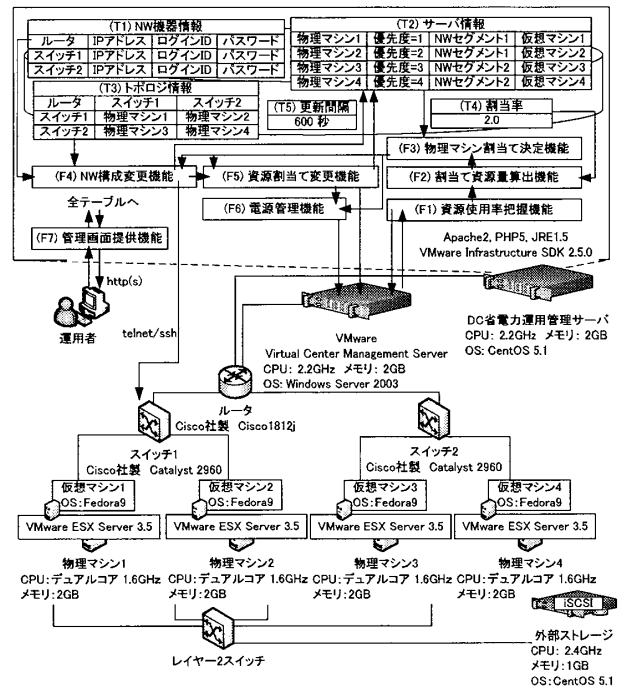


図 3：提案手順の実装概要

機能 5：資源割当て変更機能(図 3(F5))。図 2 の手順 1 に対応する。機能 4 より NW 構成変更の完了通知を受取ると、機能 3 の結果に従い、VCMSV に仮想マシンの物理マシンへの割当てを要求する。また、サーバ情報(図 3(T2))を更新する。

機能 6：電源管理機能(図 3(F6))。機能 3 で決定した割当てを受取り、決定に従い、物理マシンを休止、起動する。

機能 7：管理画面提供機能(図 3(F7))。運用者が NW 機器情報、サーバ情報、トポロジ情報、割当率、更新間隔の各情報を登録するための管理用ブラウザ画面を提供する(図 3(F7))。

3.2. 実装システムの動作確認

提案手順の有効性評価の一環として、図 3 の実装環境において、省電力管理サーバを用いて図 1 に例示の変更(1 台の物理マシン上で、仮想マシン 2 台を実行し、不要な物理マシンを休止する変更)を 10 回行った。その際、クライアントから仮想マシンへの通信断なく変更可能であることを確認した。

4. おわりに

本稿では、著者が先に提案した、データセンタ省電力運用管理手順の実装概要を述べた。本手順では、サーバの仮想化とネットワークの動的構成変更により、ネットワークセグメントの制限なく、仮想化されたサーバへの資源の再割当てを可能とすることで、サーバに加え、ネットワーク機器および空調機の電源管理によるデータセンタの省電力化が期待できる。有効性評価が今後の課題である。最後に、日頃ご指導いただく株式会社 KDDI 研究所秋葉所長に感謝する。

[1] 経済産業省，“グリーン IT イニシアティブ”，http://www.meti.go.jp/press/20071207005/03_G_IT_ini.pdf。

[2] 荒井他，“ネットワーク機器およびサーバの動的構成変更によるデータセンタ省電力運用管理手順の提案”，2008 信学ソ大。