

## 京都工芸繊維大学における 10Gbps ネットワークインフラの導入について

梶 田 秀 夫<sup>†1</sup> 村 田 和 義<sup>†1</sup>  
渋谷 雄<sup>†1</sup> 若 杉 耕 一 郎<sup>†1</sup>

京都工芸繊維大学では 2009 年 3 月に、11 年以上使用してきた ATM ネットワークを更新し、10GBASE-X を基本とするネットワークインフラ KITnet4 を導入した。KITnet4 は、Juniper Networks 社のボックス型 Layer3 スイッチである EX4200 を、Virtual Chassis(VC) と呼ばれる多数のスイッチを仮想的に一台に見せる技術を使用して構築した。本報告では、本構成を導入するにあたっての経緯や、利点・欠点などについて述べる。

### 10Gbps Campus Network Infrastructure at Kyoto Institute of Technology

HIDEO MASUDA,<sup>†1</sup> KAZUYOSHI MURATA,<sup>†1</sup> YU SHIBUYA<sup>†1</sup>  
and KOICHIRO WAKASUGI<sup>†1</sup>

In our university, we have updated our network infrastructure from ATM network system to 10-Giga Ethernet system in March 2009. New system consist of EX4200 ethernet switches and is configured with Virtual Chassis (VC) technology of Juniper Networks Inc. In this paper, we show merits and demerits of our system.

#### 1. はじめに

京都工芸繊維大学情報科学センターでは 2009 年 3 月に、11 年以上使用してきた ATM

ネットワークを更新し、10GBASE-X を基本とするネットワークインフラ KITnet4 を導入した。

本報告では、本構成を導入するにあたっての経緯や、利点・欠点などについて述べる。

KITnet4 は、Juniper Networks 社のボックス型 Layer3 スイッチである EX4200 を合計 29 台使用して大学キャンパス内に分散配置し、Virtual Chassis(VC) と呼ばれる多数のスイッチを仮想的に一台に見せる技術を使用して構築していることが特徴である。

#### 2. システム更新の目標

##### 2.1 大学におけるネットワーク整備の現状

ネットワークインフラは、教育、研究から事務作業に至るまで、いまやなくてはならないものとなってきている。しかしながら、大学においては、定期的に更新が可能なレンタルもしくはリース型のシステムとして導入されているところは少数派であり、買い取り型の調達であることが多い。本学でも、1994 年に FDDI (Fiber-Distributed Data Interface) を基盤とするネットワーク (KITnet1) を整備し、1996 年には ATM (Asynchronous Transfer Mode) ネットワーク (KITnet2)、2001 年に整備されたギガビットイーサネットワーク (KITnet3) を、いずれも補正予算による買い取り型で整備している。

ネットワークシステムは、計算機システムと比較し、多くの場合に故障頻度が典型的なバスタブ型、つまり初期導入時のトラブルが収束すればほとんど故障することなく、数年経過した後から急激に故障頻度 (特に電源装置や冷却ファンなどの故障) が高まる状況になることが経験的に知られている。買い取り型のシステムでは、定常的な保守費用を積極的に工面しつづけることが必要となるが、導入後 2,3 年はほとんど故障に見舞われない為に費用対効果が疑問視されやすく、5,6 年ほど経過してからでは保守費用を新たに工面することはなかなか難しい。さらに稼働年数が経過すると、機器自体が保守期限切れとなり、故障対応ができなくなることもしばしば発生する。事実、本学においても、KITnet2 は 13 年目に突入していたため故障修理対応ができず、稼働機器を縮退させながら、KITnet3 に移行のできない部分では動作するモジュール同士を組み合わせた機器を作成して運用を続けていた。

##### 2.2 本学における更新の計画

情報科学センターとしては、本学のネットワークシステムは重要なインフラであり、不定期の買い取り型の調達は今後も続けていくことは重大なリスクであると主張し続けている。また、レンタル型であっても買い取り型であっても、いつかは物理的な機器の入れ替え作業が発生し、ネットワークの一時的な停止は避けられない。しかし更新にあたって、ネット

<sup>†1</sup> 京都工芸繊維大学 情報科学センター  
Center for Information Science, Kyoto Institute of Technology

ワークシステムのすべての機器を一気に入れ替えるような構成となっていては、長期に渡る停止となってしまう、インフラとしては大きな問題（リスク）を抱えてしまうことになる。

従って、更新の計画を以下のように立て、設備更新プランに入れるように要求している。

- ネットワークインフラの更新経費を毎年積み立てる。
- ネットワークインフラは、半分ずつ更新できる構成にし、併用期間中に徐々に移行できる構成にする。

### 3. 更新計画

2008年度予算で大容量基幹情報ネットワークシステム一式の調達を開始した。この調達は買い取り型であった。

#### 3.1 技術的要件

買い取り型によるネットワークシステムの場合、調達にあたっては以下の点に着目して選定を行う必要があると考えられる。

- (1) 導入時点でできる限り最先端技術を採用する。  
定期的に更新するとしても、6年から8年程度の期間は使い続ける必要があるため、できるだけ新しい技術を選択するべきである。ただし、ネットワークインフラの性格上、規格が固まっていない技術の採用は避けるべきである。
- (2) できるだけオープンな技術を採用する。  
特定のベンダーのプロトコルに強く依存したシステムになっていると、更新の際の足枷になってしまう。少なくとも、相互接続部分についてはIEEEやIETFで規格化されているプロトコルを使用することを仕様書などで規定するべきである。
- (3) 最低限の冗長化技術を採用する。  
買い取り型の調達においては、24時間運用保守などの保守体制の費用を工面し続けることが難しい。従って、二重化のような最低限の冗長化技術を採用し、機器の単体故障時にスポット修理で対応せざるを得ない場合であってもインフラの停止期間を最低限に留められる構成にするべきである。
- (4) 性格の異なるシステム構成を組み合わせた構成を採用する。  
半分ずつのシステムの更新に耐えるためには、複数のネットワークシステムを重ね合わせるような構成にし、インフラへの接続拠点はいずれのシステムにも接続可能な状況にしなければならない。性格の異なるシステム構成を組み合わせることで、冗長性を高めることも可能となる。

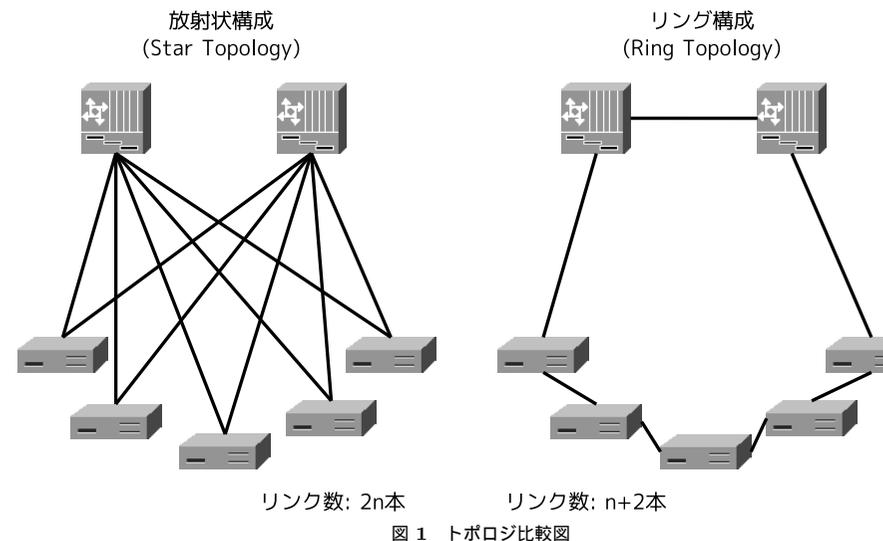


図1 トポロジ比較図

- (5) できるだけ同一の機器で揃える。  
買い取り型の調達の場合、特に導入後5,6年が経過すると機器の故障を避けることができない。このとき導入する機器を揃えておくと、タイミング的にすぐに切り替えなどの対応が難しい機器が故障した場合に、正常動作中の機器と入れ替える、といった対応が採り易くなる。

#### 3.2 使用可能な技術

##### 3.2.1 リング状トポロジのイーサネット

キャンパス規模のネットワークインフラの構築にあたり、冗長性をもったトポロジとしてよく使われるものが、放射状トポロジである(図1左)。このトポロジは、情報センターのような基幹(コア)となる建物に多くの機器が集約されている環境に適している。利点として、維持管理をコアスイッチに注力することが可能であり、末端(エッジ)の装置はそのエッジスイッチの配置された場所にしか影響がないため、障害の範囲を限定し易い。しかしながら、コアスイッチにエッジスイッチの数だけのネットワークポートが必要となるため、通常シャーシ型の大型スイッチが必要となり導入コストだけでなく、保守コストやスポット修理コストが高額となる傾向にある。

これに対し、各エッジスイッチをリング状に接続したリング状トポロジ(図1右)をとることができる製品が提供されつつある。利点としては、コアスイッチに必要となるエッジスイッチを収容するためのネットワークポート数がリング数の2倍で済むことから、エッジスイッチの数に対して少なくできるため、小型のスイッチ装置でも対応でき比較的安価に実現ができる。しかしながら、エッジスイッチからコアスイッチに至るまでに複数のエッジスイッチを経由する必要があるために、障害の範囲が比較的広範囲になる場合がある。それでもリング状であることから最低限2つの経路があるので冗長性は保たれる。

リング状トポロジは、古くはIBM Token Ring や FDDI として使われていたが、ネットワークメディアの高速化が遅れ、ネットワークのデータリンクとしては、イーサネットに取って代わられているのが現状である。イーサネットはブロードキャストを基盤とするデータリンクであるので、リング状トポロジを形成する為に、一般的にはSTP (Spanning Tree Protocol) を使用し、リングの一箇所を blocking port とすることで実現する。しかし、STP は収束に時間がかかることや最大7ノードまでの構成でしか動作しないこと、複数のリングを構成した場合にどこを優先的に blocking port になるように設計すべきか、といった運用上の困難さが残る。

近年、イーサネットをリング状トポロジで運用できるように工夫されたプロトコルを有する機器が出てきている。

日立電線(株)社は、MMRP (Multi Master Ring Protocol)<sup>2)</sup> と呼ばれるプロトコルを有している。MMRP は、Layer2 のイーサネット高速冗長の Ring プロトコルであり、リンク切断時に最速 19ms で切り替わりを実現していることが特徴である。

Foundry Networks 社<sup>\*1</sup>は、MRP (Metro Ring Protocol)<sup>3)</sup> と呼ばれるリング構成に適した高速なプロトコルを有している。リンク切断時に 100ms 程度での切り替わりを実現していることや、複数のリングを有するトポロジにおいて、複数のリングを収容する物理リンクが存在していても構築できる (MRP Phase 2) 機能を有していることが特徴である。

H3C Technologies 社は、RRPP (Rapid Ring Protection Protocol)<sup>4)</sup> と呼ばれるプロトコルを有している。リングの構成に関して比較的自由度が高く、Single ring, Tangent ring, Intersectant ring, Dual homed ring といった柔軟な構成がとれることが特徴である。

Juniper Networks 社は、EX4200 シリーズのスイッチに対して VC (Virtual Chassis)<sup>5)</sup> とよばれる仮想化技術を有している。リング構成に特化したプロトコルではないが、最大

10 台までのスイッチをリング状に接続することが可能である。

### 3.2.2 仮想化技術

キャンパス規模のネットワークインフラの構築にあたり、通常多くのエッジスイッチを配置する必要がある。また、冗長化の為にコアスイッチを二重化し、VRRP (Virtual Router Redundancy Protocol) などを利用した冗長構成を取るために複数のスイッチを運用管理することは避けることができない。複数のスイッチ機器を運用する場合、設定が矛盾しないようにしなければならないが、スイッチ機器の増加に従い手間は増大する。特に、冗長構成機器の間では設定を同期させる必要があり、それがコアスイッチであればより慎重な操作が必要となる。

近年、これらに対応するために、仮想化技術が採用されつつある。仮想化技術は、物理的な制約を隠蔽することができることから、ネットワークサーバなどに使用されてきているが、これをネットワークデバイスにまで適用し、サーバからネットワークまで一体的に考えられるように、という動向と考えられる。

Cisco Systems 社は、VSS (Virtual Switching System)<sup>6)</sup> と呼ばれる仮想化技術を有している。VSS は、Catalyst6500 シリーズに採用されており、2 台のコアスイッチを仮想的に 1 台のスイッチに見せることができる。放射状トポロジにおける冗長化構成のために 2 台のコアスイッチを置いた場合、STP/MSTP などを用いたり VRRP を用いたりする構成では、それぞれのコアスイッチの設定を慎重に揃える必要があり運用の手間が大きいのが、VSS により 1 台のスイッチの操作になるので、同期化を意識する必要がなくなることが特徴である。

Juniper Networks 社は、EX4200 シリーズのスイッチに対して Virtual Chassis<sup>5)</sup> とよばれる仮想化技術を有している。VC は、最大 10 台までのスイッチを 10GBASE-X もしくは専用スタックポートで接続することができ、単一のスイッチのように扱えることが特徴である。VSS はコアスイッチへの適用が考えられるが、VC はコアスイッチからエッジスイッチまでを一体として捉えることが可能となる。

## 4. システムの構成

3.1 節に挙げた要件にあわせて、本システムでは以下の技術を用いてシステムを構成した。

- (1) 10GbE を基盤とするバックボーンとする。
- (2) リング状トポロジとする。
- (3) Juniper Networks 社の EX4200 を用いた Virtual Chassis を採用する。

\*1 現在は Brocade Communications Systems 社

ネットワーク全体構成図

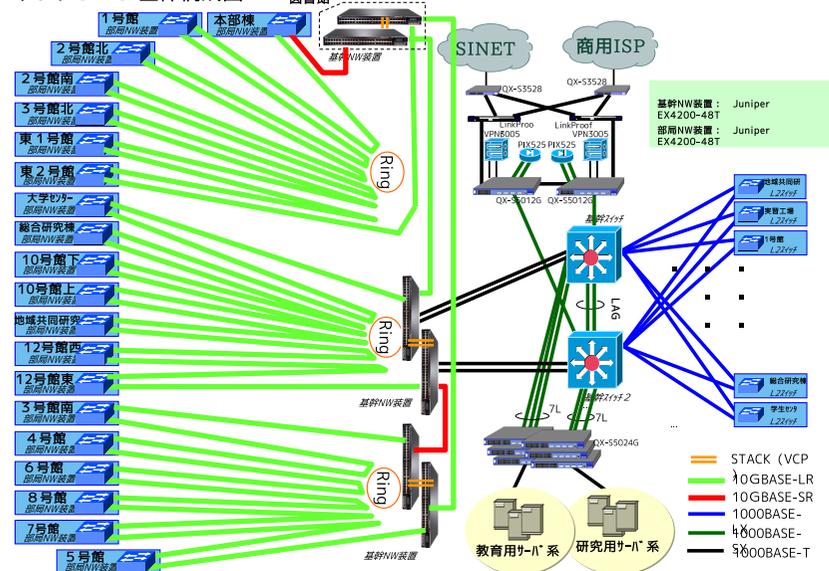


図2 システム構成概略図

## RING接続グループ Campus Map

松ヶ崎キャンパス

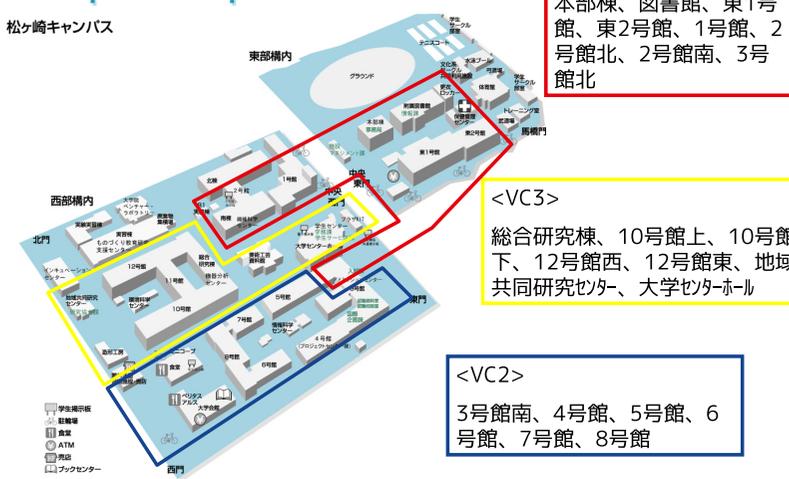


図3 システム配置図

(4) KITnet3とはOSPFを用いて相互接続を実施する。

図2は、本システムの構成概略図である。

KITnet3はCisco Systems社のCatalyst6506をコアスイッチに配置し、エッジスイッチとしてCatalyst3508とCatalyst3512をスタック接続したものを1000base-Xで接続配置した放射状トポロジであり、Cisco社独自プロトコルであるPVST+(Per-VLAN Spanning Tree Protocol plus)とHSRP(Hot Standby Routing Protocol)による冗長化を行っていた。そこで、KITnet4では、異なる技術的な構成を目指し、10GBASE-Xと共にリング状トポロジを採用した。実際にも、10GBASE-Xのポート単価はまだまだ高価であり、放射状トポロジを採用することは予算的にも厳しかった。

また、VCを採用したことにより、コアスイッチからエッジスイッチまで、同一機種(EX4200-48T)を使用することができたため、年数が経過した後、機器本体の修理ができない状況に陥ったとしても、VC(リング)を構成する機器を減らしたり移動させたりして運用を継続す

ることができると考えられる。通常高価となるコアスイッチもエッジスイッチと同一価格帯となるため、保守費用も抑えられると考えられる。

VCは1リング構成あたり最大10台までという制約があるため、本学では22箇所のエッジスイッチ配置場所にあわせて3つのリングに分割してVCを形成している(図3)。VC間の接続は、コアスイッチとしての配置を2から3台としてスタック接続することで10GBASE-Xのポートに余裕を持たせ、3つのVCをトライアングル状に10GBASE-Xで接続することで冗長性を持たせている(図2)。

さらに、リング状トポロジではあるが、物理的な光ファイバの配線は図2のように、基幹となるセンターですべて折り返しのパッチにより接続されている。これにより、エッジスイッチの障害や光ファイバの断絶などにより一部の機器(群)を切り離す必要が生じた場合でも、基幹側でパッチの接続替えのみで対応が可能となっている。光ファイバの芯数が余分に必要になるかのようにも見えるが、放射状トポロジでは基幹拠点からエッジ拠点に2系

統 4 芯<sup>\*1</sup>が必要になるため、結果的に差はなく、将来トポロジをリング状から放射状に変更する際に光ファイバの敷設を変更する必要もない。

## 5. 現在の運用状況と考察

2009 年 3 月末に稼働開始であるため、まだ運用管理上の諸問題にあまり遭遇していないが、現時点での所感は以下の通りである。

### 5.0.2.1 優 位 点

- 10GBASE-X を基幹とするネットワークとしては非常に低コストで導入できた。  
コアスイッチ部分に、シャーシ型のコアスイッチに比べて 1 桁以上低価格の機器を用いて実現できることが大きい。
- VC を用いることにより、複数の離れた拠点を 400 ポートもある単一のスイッチ 3 台分に対して運用管理するだけで済むようにできた。  
22 箇所の拠点に加えて 2 台のコアスイッチを運用管理することに比べれば格段に管理対象が少ない。
- 機器障害時に機器の交換手順が簡略化された。  
VC を組んでいることから、機器の交換の際に、新しい機器は、Factory Default の状態で繋ぎ、VC に参加させる、という手順を実施するだけで、VC のマスターノード (RE) から設定情報が反映されて VC に組み入れられる、という動作になる。通常であれば、機器交換あたっては、以前の設定をきちんと保存し、その設定を流し込んでから接続する必要があったことに比べれば非常に簡略化できている。

### 5.0.2.2 懸念, 要改善点

- VC が 1 つあたり 10 台までという制約  
EX4200 が、10GBASE-X ポートを 2 つしか持っていないため、これらのポートをリング構成の為に使用する (VCE 接続と呼ばれる) と、それ以外への接続に 10GbE 接続が使用できない。そのため、基幹に配置する EX4200 は、VCP と呼ばれる専用スタックケーブルによる接続を併用することで 10GBASE-X ポートに余裕を持たせる必要がある。しかし、そうすると拠点に配置できる EX4200 の数が減り、さらに VC の数を増やす必要が生じ、それに伴ってコストパフォーマンスが悪くなってしまふ。1000base-T ポートに対して 8 本までの Link Aggregation が可能でありこれを代替えとすることも

できるが、適用規模に制約があると考えられる。

- “Virtual” であるが故の障害対応の困難さ  
ネットワークインフラ部分は比較的低いレイヤー、とりわけレイヤー 1 は物理的なデバイスが絡む部分であり、物理デバイスに関する障害検知機能は重要な役割を果たす。一方仮想化は、いわば物理デバイスの差異を吸収しそれを見せなくする技術であるため、物理デバイスの情報は隠蔽されがちであり、要求が相反してしまう部分が存在する。さらに、Juniper 社の JUNOS は BSD をベースとする安定性を重視し、多くのノウハウが蓄積されたネットワークデバイス用の OS と言えるが、VC 自体がまだまだ新しい技術であるため、ノウハウの蓄積が不十分であるところも否めない。  
例えば、VC を構成する機器との接続に使用されるポートでの障害が、VC 自体の安定性に強く依存してしまうため、単なるポート不良よりも障害判別が難しくなる場合があったり、VC を構築した機器数に応じて設定の反映 (commit) にかかる時間が無視出来ない程度 (2 分程度) になることが報告されている。

## 6. おわりに

低コストで 10GbE を基盤とするネットワークインフラとして、Juniper Networks 社の VC 技術を用いた EX4200 を採用して構築を行った。仮想化技術は、これからもネットワークインフラやサーバ環境、ストレージに至るまで適用されていくと考えられることから、今後も活用されていくことが予想される。

今後の課題としては、定常的な運用を続けノウハウの蓄積をすることが挙げられる。

## 参 考 文 献

- 1) Juniper Networks <http://www.juniper.co.jp/>.
- 2) 日立電線 (株): 高速 Layer2 リング MMRP,  
<http://www.apresia.jp/products/func/function/mmrp.html>.
- 3) Foundry Networks: Metro Ring Protocol,  
<http://www.foundrynetworks.co.jp/services/documentation/sribcg/Metro.html>.
- 4) H3C Technologies: Rapid Ring Protection Protocol,  
[http://www.h3c.jp/jp/Products\\_\\_\\_Solutions/Technology/LAN/RRPP/](http://www.h3c.jp/jp/Products___Solutions/Technology/LAN/RRPP/).
- 5) Juniper Networks: EX4200 Switch - Virtual Chassis,  
<http://www.juniper.net/us/en/products-services/switching/ex-series/>.
- 6) Cisco Systems: Cisco Virtual Switching System,  
<http://www.cisco.com/web/JP/event/campaign/fy08q4cat65kvss/index.html>.

\*1 10GBASE-BX のように WDM 技術を応用して 1 芯でリンクするデータリンクもあるがより高価である。