

第30部

WIDEネットワークの現状

樋山 寛章、堀場 勝広、上野 幸杜、井上 博之、前田 朋孝、岡部 寿男、Glenn Mansfield Keeni、齋藤 武夫、土井 一夫、松本 智、高橋 航平、畠山 元也、近藤 賢郎、宮下 山斗、関口 貴久、川口 慎司、嶋村 孔明、清水 倫人、森 康祐、中村 遼、山本 成一

本ドキュメントでは、WIDE backboneと各NOCの現状について述べる。

開発している研究者、開発者らの新技術の運用実験の場としても頻繁に活用されている。

第1章 はじめに

WIDEバックボーンネットワークは国内はもとよりSan Fransico, Losangels, Bangkokなど海外にも拠点(NOC, Network Operation Center)を持つ広大なレイヤー 2およびレイヤー 3ネットワークである。WIDEバックボーンネットワークは各接続組織の対外接続ネットワークとして活用されるだけでなく、インターネットの新技術を

WIDEバックボーンネットワークの運用はTwoワーキンググループに参加する各NOCの運用者による定常的な運用に支えられている。本年度のTwoワーキンググループの活動報告として、WIDEバックボーンネットワークの運用報告を行い、合わせて、レイヤー 2接続のオーバーレイ化を目指したVXLANの運用実験についても報告する。最後に今後のWIDEバックボーン運用についての展望を述べる。

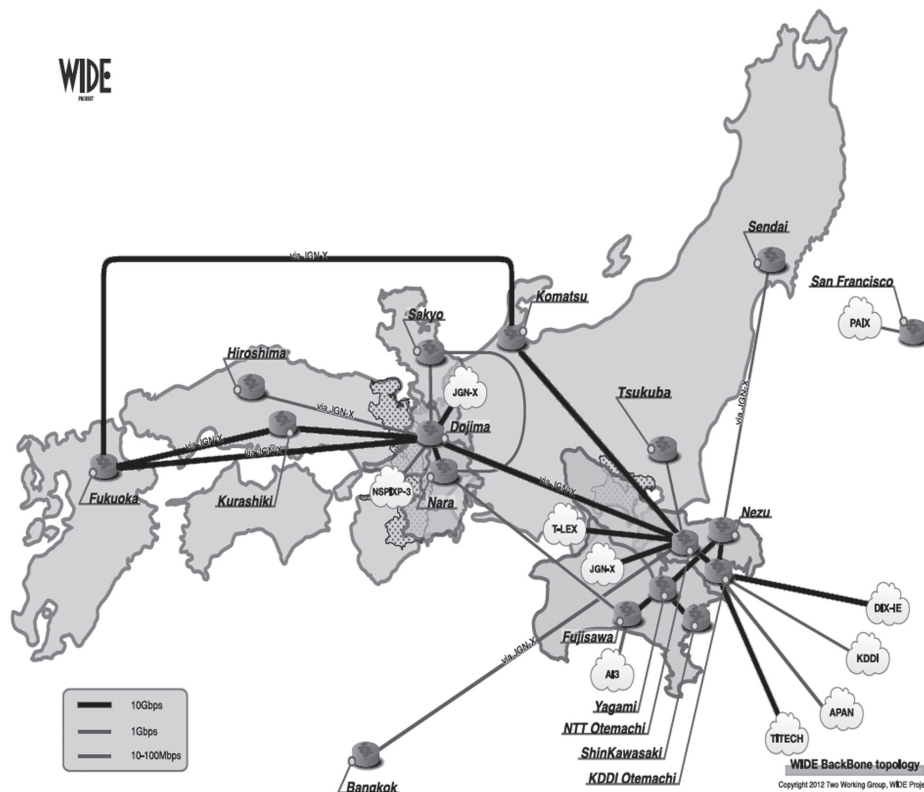


図2.1 WIDEバックボーントポロジ

第2章 WIDEバックボーンの運用

本節では、WIDEバックボーンの各拠点での2012年12月31日から2013年12月31日までの運用報告と2013年12月31日現在のWIDEバックボーンのネットワーク構成を報告する。図2.1は2013年12月31日現在のWIDEバックボーンの概略図である。

2.1 San Francisco

サンフランシスコNOC(sanfrancisco)は、2004年4月からそれまでのsanjoseに代わり稼働した新しいNOCで、Los AngelesからOC-3により接続されていた。その後OC-3から100M Ethernetに変更された。主な接続先は、PAIXや

ISCである。

2010年9月のLos Angeles NOC撤収にともない、2010年10月にLos AngelesとSan Francisco 間の回線も廃止され、専用線による接続の無い独立NOCとして存在する。

2013年はM-ROOT関連の機材更新があったが、WIDE SFO NOCとしての構成変更は無かった。

- ・(2013/06/19) データセンターの電源故障による電源断発生
- ・(2013/11/15) pc5.sfoディスク障害によりHDD交換
- ・(2013/12/16) NOC作業中の電源ショートによる電源断発生

As of 2014/01/08

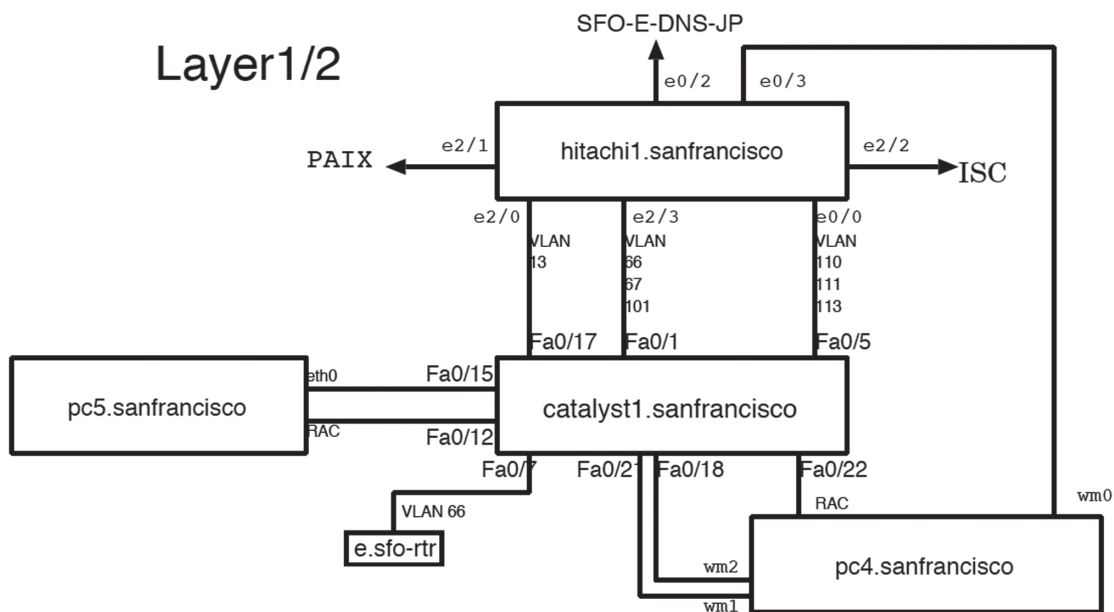


図2.2 San Francisco NOC

2.2 仙台

仙台NOCは仙台周辺の拠点を収容するNOCとして運用されている。接続回線の障害や停電もなく、安定して運用された。トポロジーに変更はなく、FreeBSDが稼働するpc5, pc10, pc13を物理マシンからVMware上の仮想マシンへと移行した。

- (2013/07/27) pc5, pc13をVMware 上の仮想マシンに移行
- (2013/07/29) pc10をVMware上の仮想マシンに移行

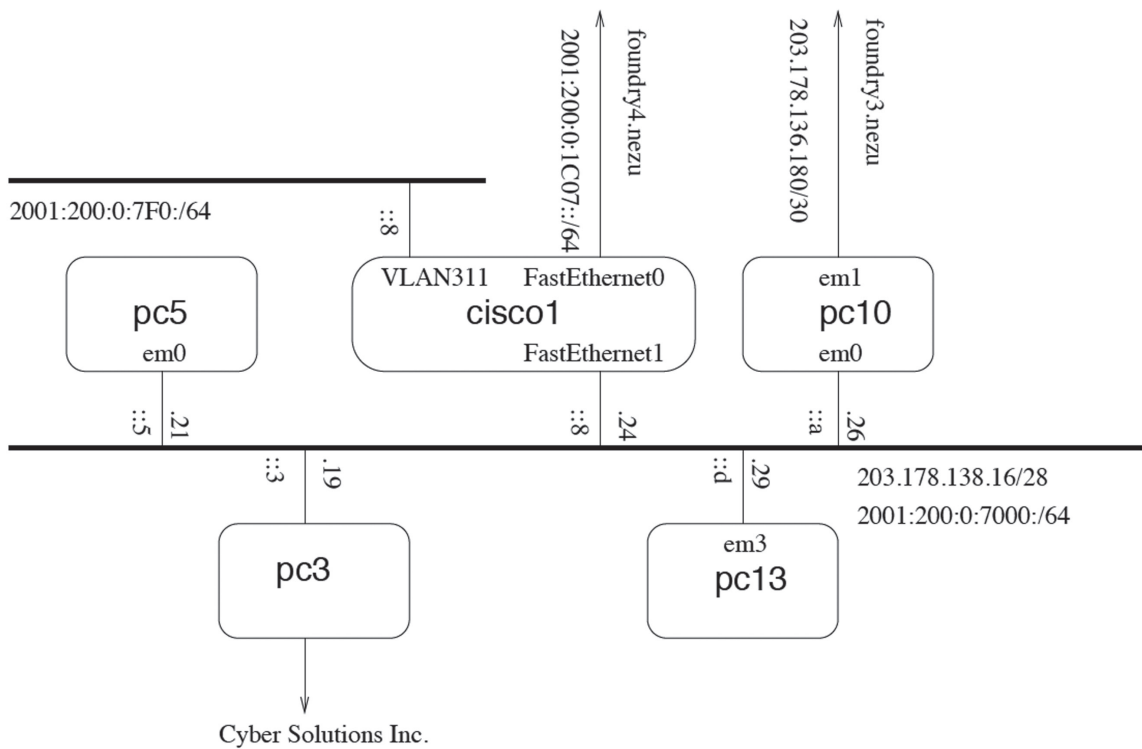


図2.3 仙台NOC

2.3 筑波

筑波NOCは筑波大学学術情報メディアセンター内に設置されている、システム情報工学研究科産学間連携推進室をはじめとする周辺の研究組織を収容している。

株式会社ソフトイーサと共同で、グローバル・固定IPv6アドレス割当型トンネル接続実験サービス(v6ip.tsukuba.wide.ad.jp)を運用しており、2012年にはDNS64/NAT64

によるIPv4ネットワークとの相互接続の試験運用も開始した。

- (2013/10/26) 電気事業法に基づく電気設備の定期点検のため停止
- (2013/10/26) FTPサーバ(ftp.tsukuba.wide.ad.jp)のハードウェア強化
- (2013/10/27) 同上

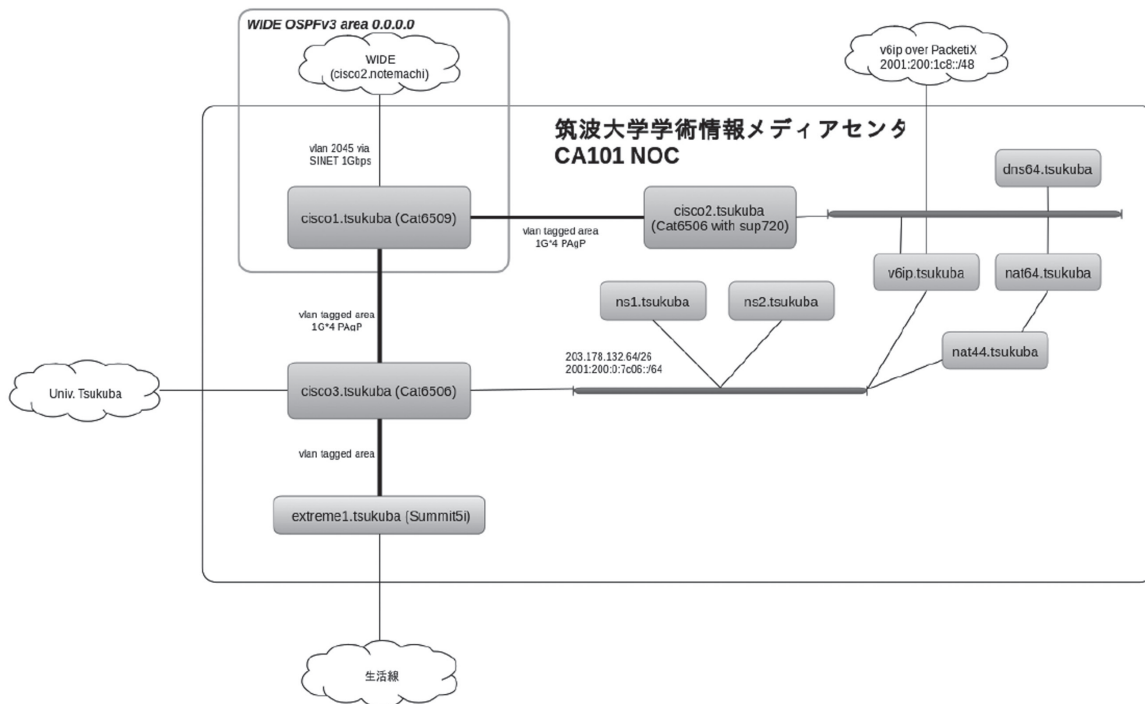


図2.4 筑波NOC

2.4 根津

根津NOCは、WIDE関東地区の重要な接続拠点として、東京大学、JGN-X等との接続を行っている。またWIDEクラウドの拠点としても重要な機器が設置されている。

- (2013/01/17) callmanager2故障
- (2013/09/22) 東京大学法定点検のため停電(ダウン無し)
- (2013/10/27) 東京大学法定点検のため停電(ダウン無し)
- (2013/12) foundry2.nezu引退

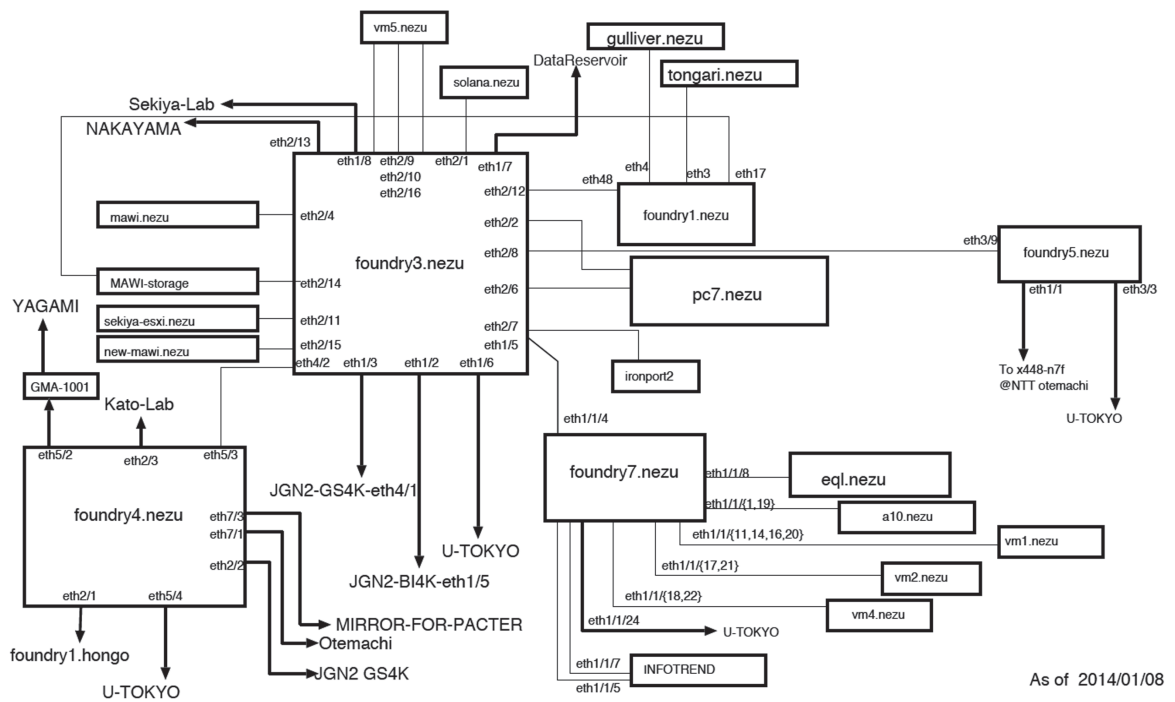


図2.5 根津NOC

2.5 NTT大手町

NTT大手町NOC(notemachi)は、1999年終りから稼働した比較的新しいNOCで、現在、関西方面、北陸方面へのL2網、JGN-X、APAN-JPの接続拠点として重要な立場にある。また、日本のインターネットトラフィック交換の1拠点として、DIX-IE、T-LEXを設置しISPおよび学術研究NWを収容している。

- (2013/10/15) Notemachi-JAIST間回線借用
- (2013/08/20) pc1.notemachi機材更新
- (2013/07/02) juniper2.notemachi(MX80)設置

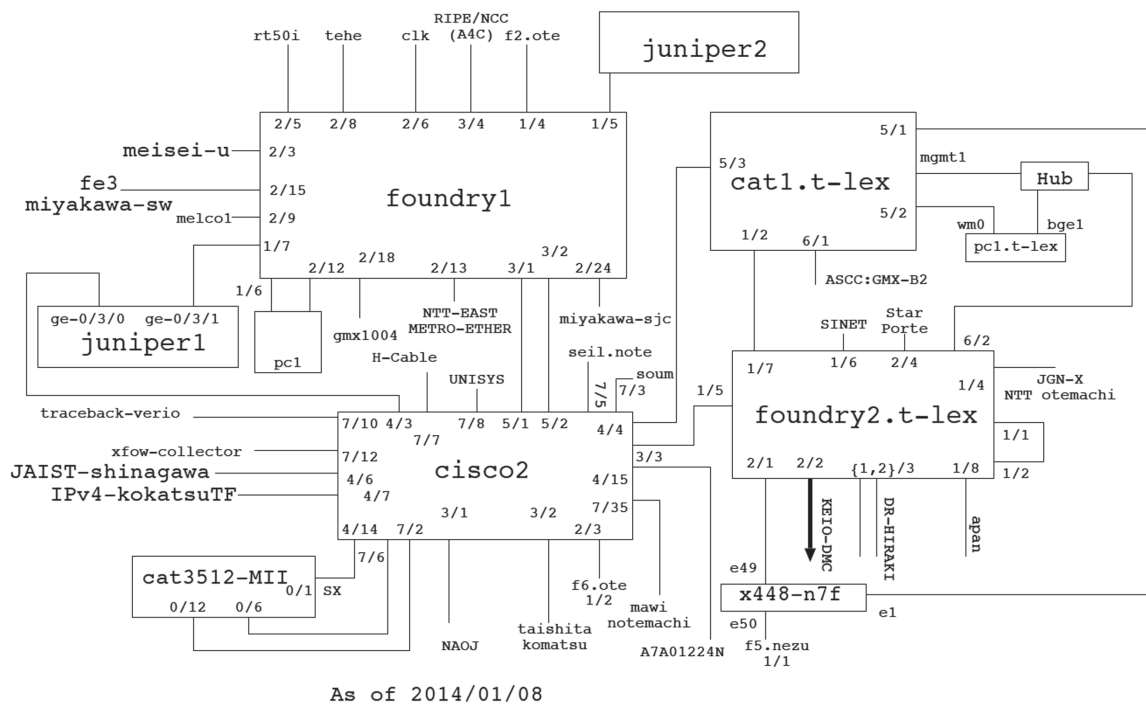


図2.6 NTT大手町NOC

2.6 KDDI大手町

KDDI大手町NOCはWIDEバックボーンの中でも中核を担う重要なNOCとなっており、外部組織接続が最も多いNOCとなっている。10GbEによるバックボーンが導入され、NTT大手町NOCとの連携がより強まり、WIDEからDIX-IEへの接続拠点となっている。

- (2013/06) cisco1.otemachi ハードウェア故障
- (2013/07) foundry6.otemachi 一時的なルーティング障害
- (2013/08) foundry2.otemachi Syn Attackにより一時的な過負荷
- (2013/12) foundry6.otemachi DATAHOTEL (旧称 livedoor)とのBGP peer-ing廃止

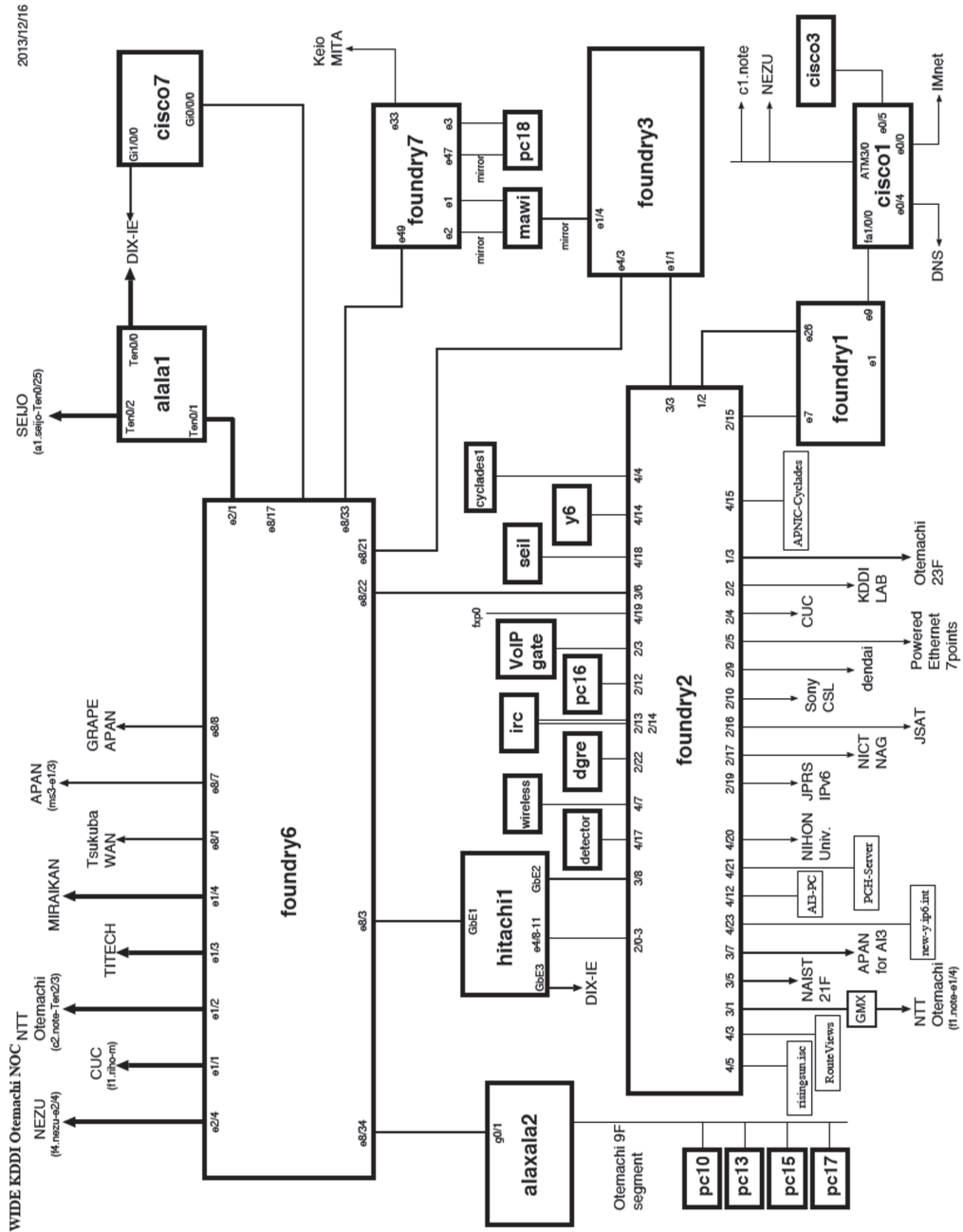


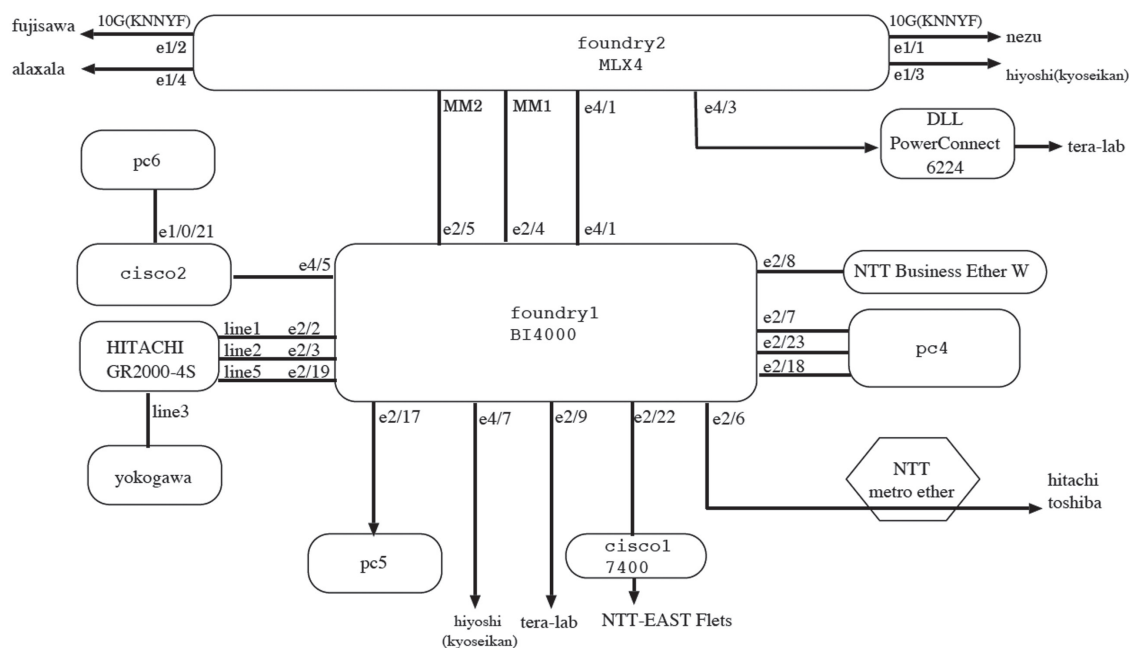
図2.7 KDDI大手町NOC

2.7 矢上

矢上NOCは慶應義塾大学工学部矢上キャンパス構内にあり、同大学工学部情報工学科および周辺の研究組織を収容すると共に慶應DMC を介してJGN-X, CineGridとの接続を行っている。

- (2013/07/11) 藤沢==矢上間回線借用
- (2013/08/17) 定期保安点検による停電
- (2013/10/30) 矢上==アラクサラ間回線借用

YAGAMI NOC TOPOLOGY (Layer1)



\$Id: yagami-layer1.eps,v 1.1 2013/12/15 21:26:22 latte Exp \$+9hours(JST)

図2.8 矢上NOC Layer-1 トポロジ

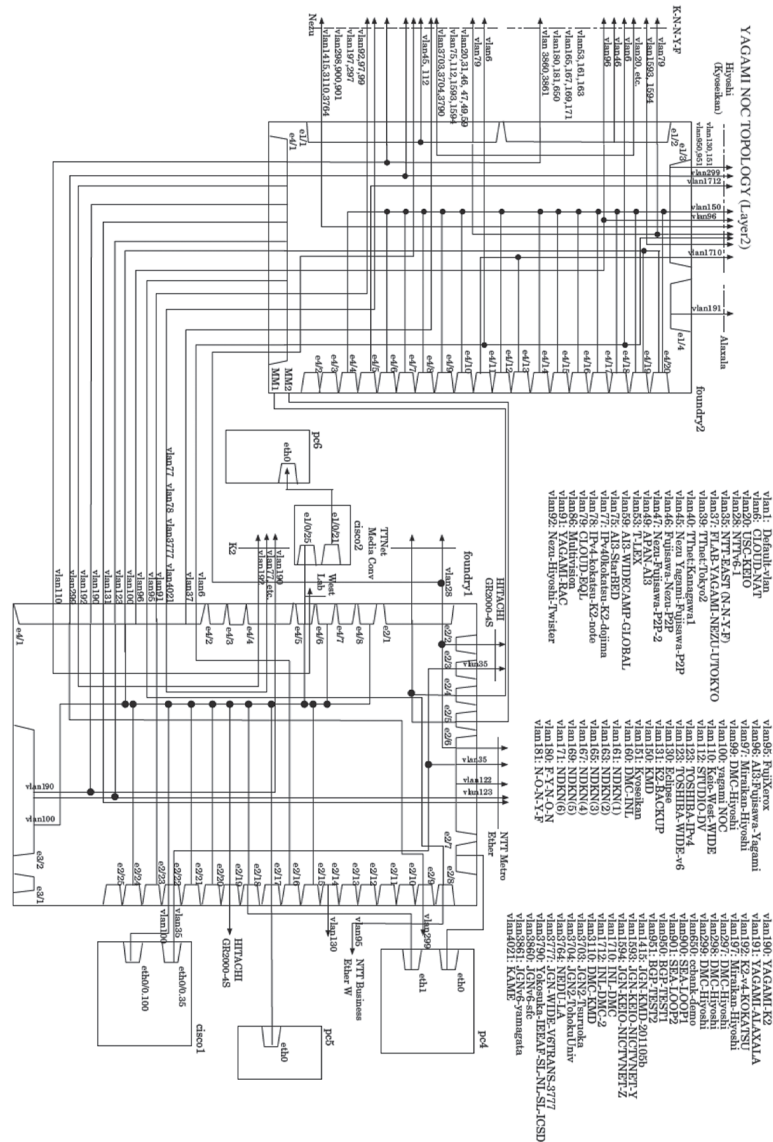


図2.9 矢上NOC Layer-2 トポロジ

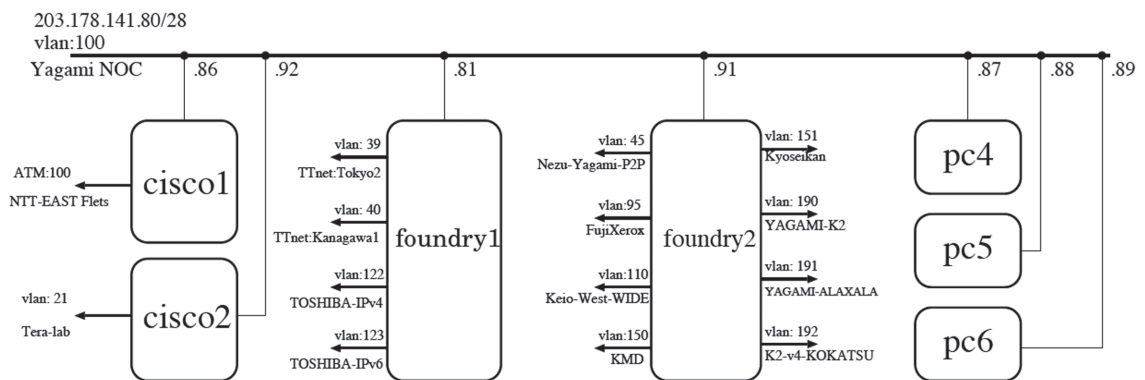


図2.10 矢上NOC Layer-3 トポロジ

2.8 新川崎

新川崎NOCは、K2タウンキャンパス内の村井研究室を拠点としたNOCである。K2タウンキャンパス村井研究室はこれまで矢上NOCの下部組織として運用されてきたが、リーフ組織への回線提供を行うため、2005年後半よりNOCとして運用していた。2008年度の構成変更にて、リーフ組織であったアラクサラの接続先が矢上NOCへと変更となったため、一時NOCではなくなったが、2009

年度にIPv4枯渇TFがリーフ組織として接続したため、再びNOCとしての機能を担うこととなった。2013年に新川崎K2タウンキャンパスからの撤収にともない、新川崎NOCも廃止された。

- (2013/03/31) 新川崎K2タウンキャンパス撤収に伴うNOC廃止

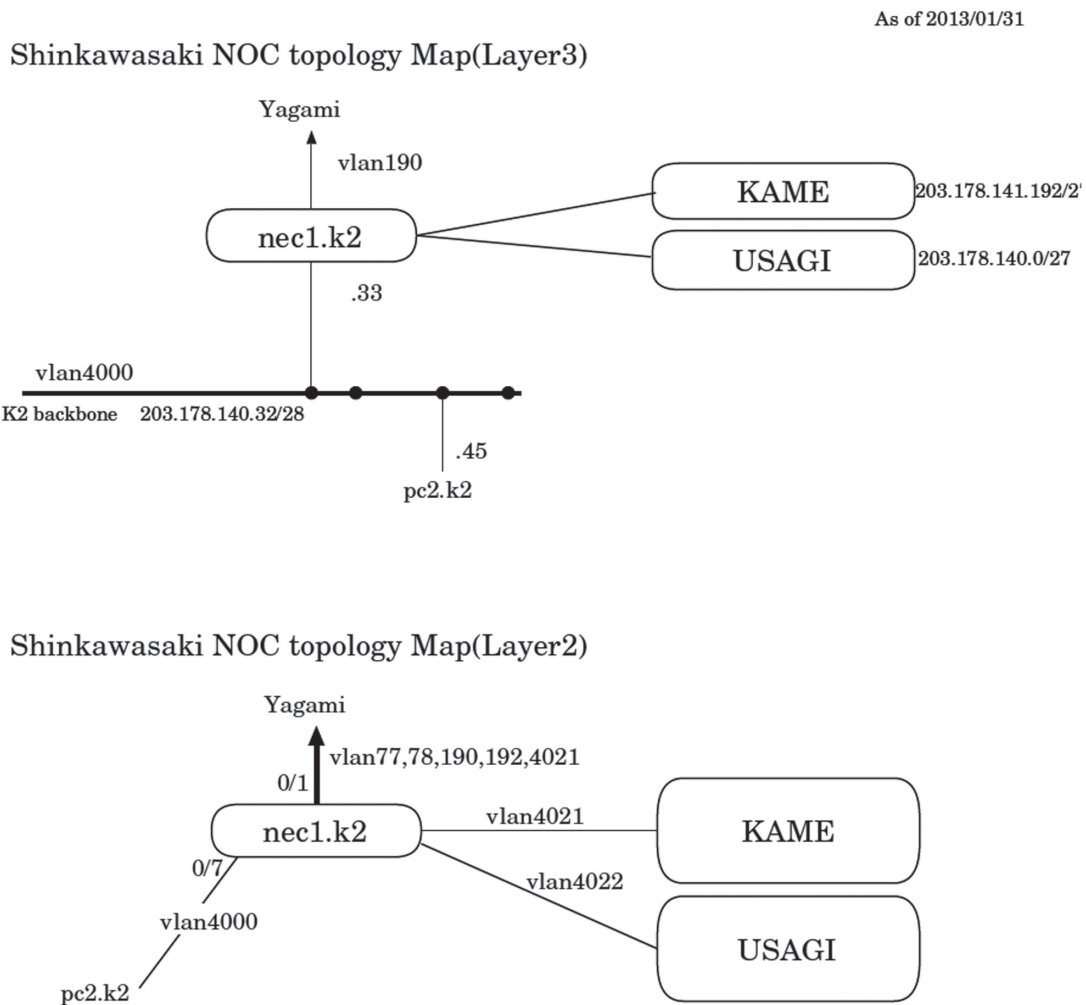


図2.11 新川崎NOC

2.9 藤沢

藤沢NOCは慶應義塾大学湘南藤沢キャンパス内にあり、慶應義塾大学や村井研究室の他、周辺の研究組織を収容している。同時にW3CやAI3との接続、VoIP関連サービス、外部研究組織のトラフィック計測サーバの設置及び接続

性の提供などを行っている。

- ・ (2013/01/27) IPv4経路数の増加による不調のため、基幹ルータをリプレース
- ・ (2013/02/01) W3Cへの接続性提供を開始

, height,

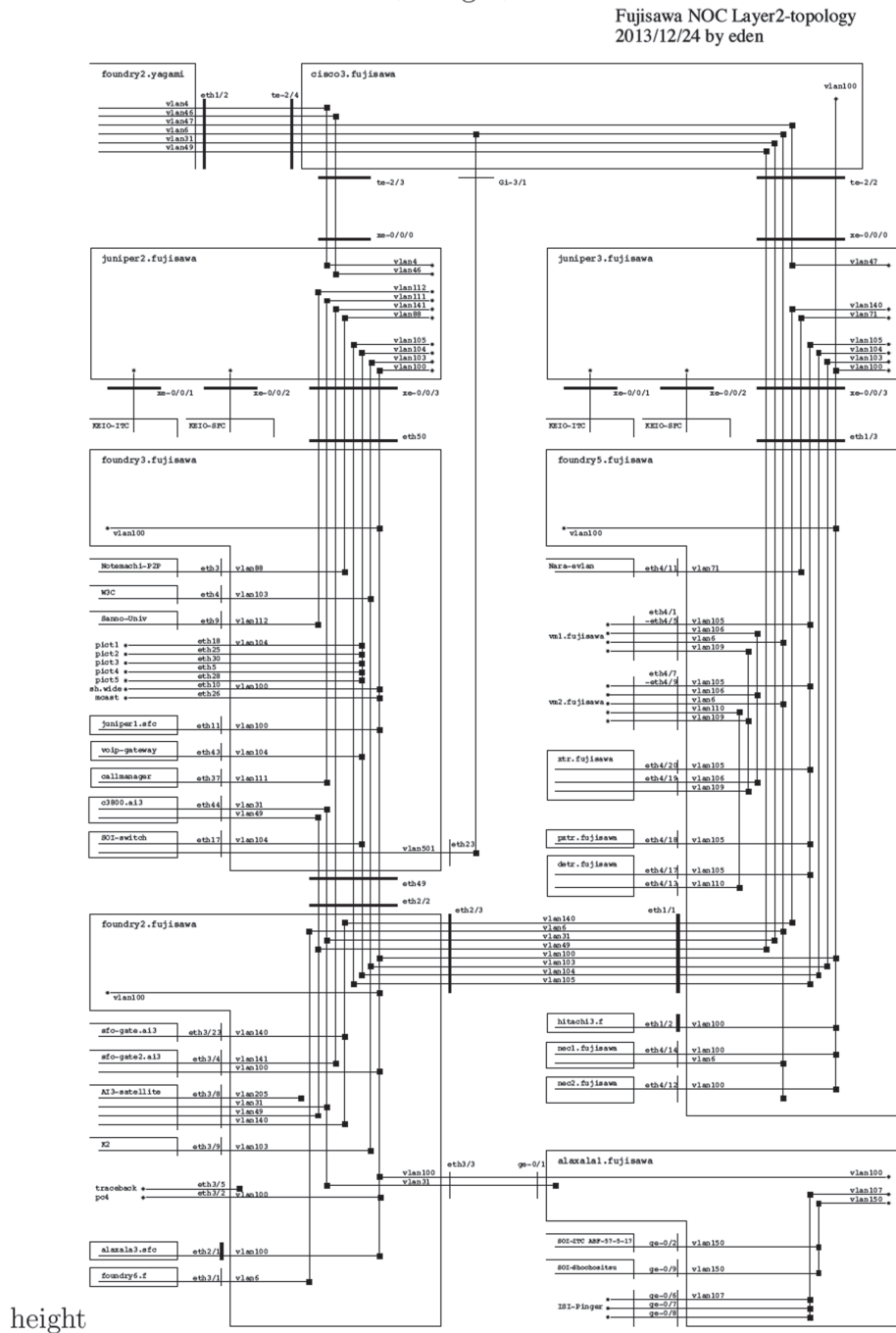


図2.12 藤沢NOC Layer-2トポロジ図

- (2013/03/27) K2の藤沢NOC移設に伴い接続性提供を開始
- (2013/03/27) 産業能率大学の収容回線変更
- (2013/03/27) 奈良・左京・藤沢間専用線サービス切り替え工事
- (2013/04/01) XCAST用サーバをVMに移行
- (2013/04/15) 奈良・左京・藤沢間専用線サービスの収容変更
- (2013/11/07) 大手町-藤沢間バックアップ線の利用開始
- (2013/12/01) SFC構内全域の変電設備の定期保安点検による構内停電

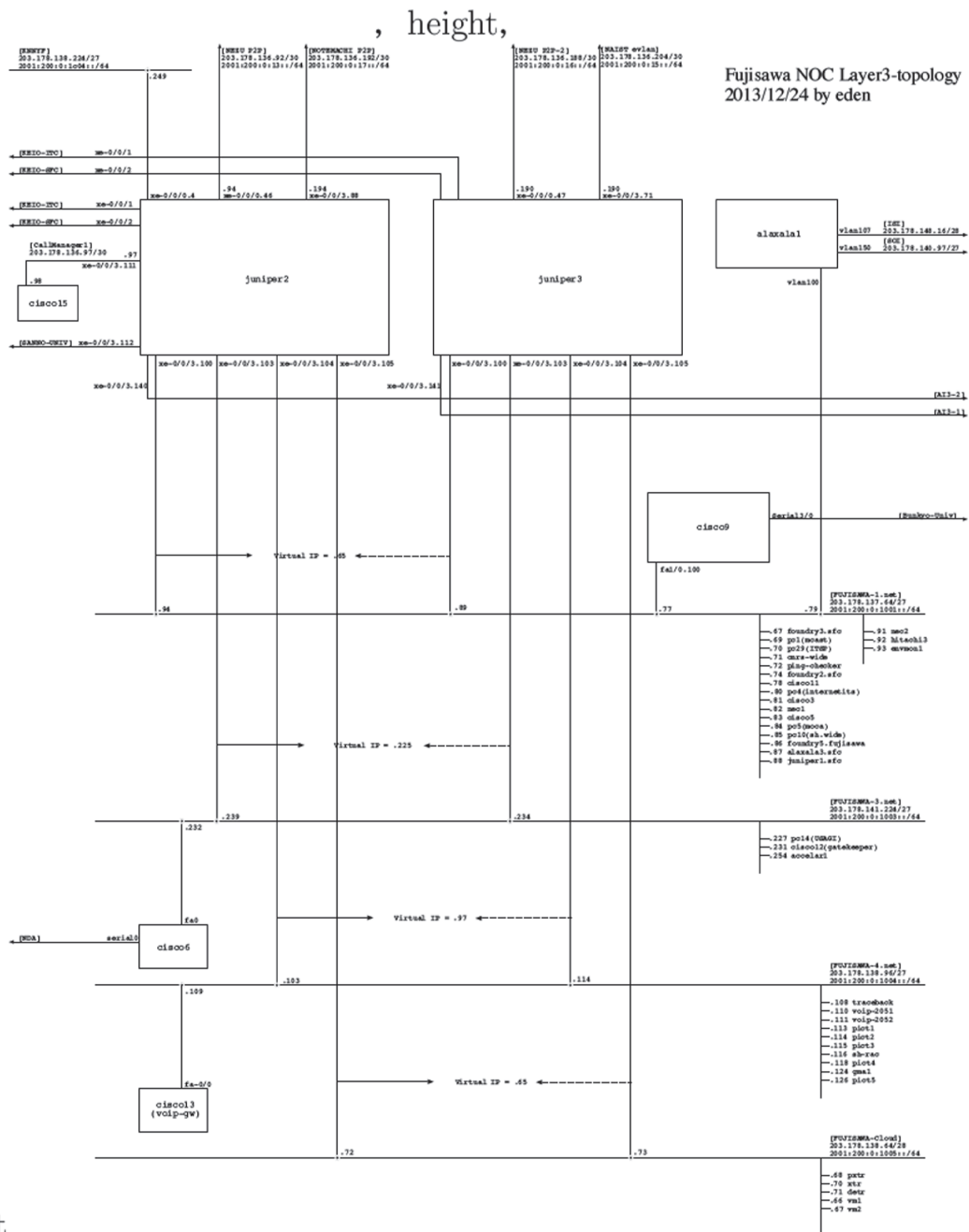


図2.13 藤沢NOC Layer-3トポロジ図

2.10 小松

小松NOCは北陸先端科学技術大学院大学(JAIST / 石川県能美市)内に設置されたNOCであり, 同大学, NICT北陸StarBED技術センター (通称:StarBED)等への接続を収容している. NOC間接続として関東および関西方面に対し複

数のリンクを持ち, 東阪間リンク障害時の迂回経路としての役割も担っている.

- (2013/03/23) 08:00{17:00 JAIST全学停電に伴うサービス停止.

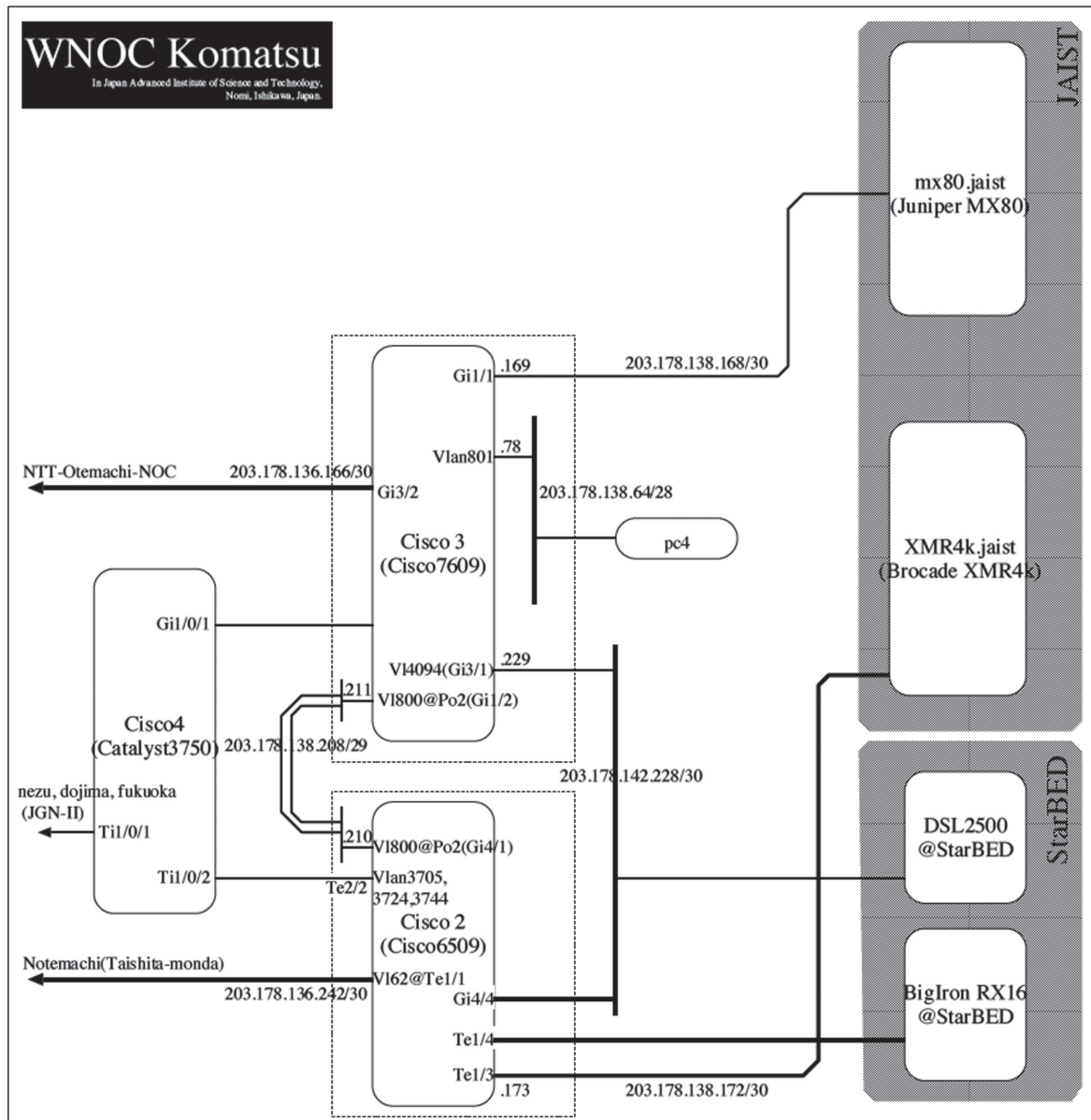


図2.14 小松NOC

2.11 堂島

堂島NOCは、WIDEプロジェクトのネットワークにおける西日本のコア拠点となっている。NTTテレパーク堂島第1ビルと第3ビルに拠点を構え、NTT大手町NOCとともに10GigabitEthernetバックボーンを担ったり、大阪における学術IX(NSPIX3)拠点を担ったりしているNOCである。また、第3ビル内においてJGNやSINETとも接続し、西日本方面の多数のNOCとリーフサイトを収容している。主にNSPIX3とWIDEバックボーン接続を担っているcisco2.dojimaの老朽化が顕著となっており、近い将来に向けたリプレースの検討が進んでいる。

- (2013/1/17) cisco2.dojimaにおいてIGP経路再計算におけるCPU負荷増大のため再起動(のちにbrocade1.komatsuとのアドレス重複によるものと発覚)
- (2013/5/2) cisco2.dojimaのWIDEバックボーン側インターフェースの1000Base-SXモジュール故障に伴い、1000Base-Tへの収容変更
- (2013/12/3) JGN-Xより借り受けている仮想ルータcrs1.dojimaのファームウェアをアップデート

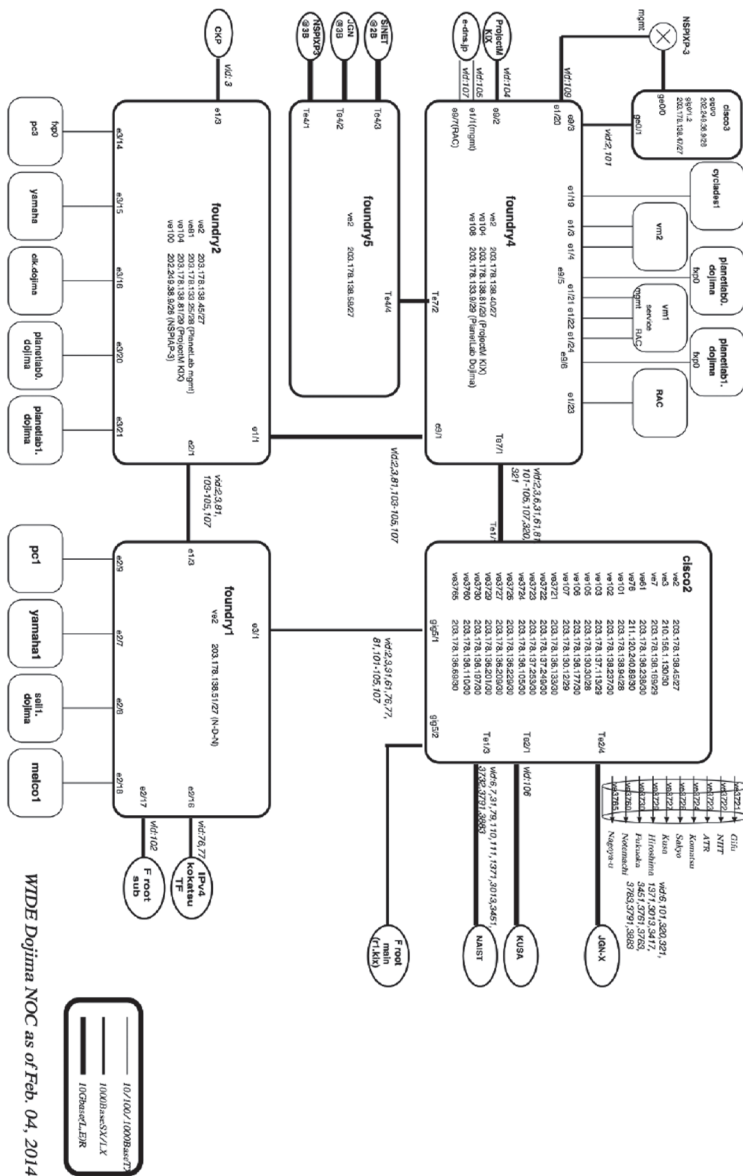


図2.15 堂島NOCトポロジ

2.12 奈良

奈良NOCは奈良先端科学技術大学院大学内にあり、大学およびNOC周辺の研究組織を収容するとともにAI3と接続している。また、Debian JP等の公式ミラーを始めとする10以上のミラーを提供するFTPミラー (ftp.nara.wide.ad.jp)をサービスしている。

- ・(2013/4/4) 奈良・左京・藤沢線をケイ・オプティコムイーサネットVPNからNTTcommunications Universal Oneに切り替え
- ・(2013/5/18) 奈良-堂島線接続ルータをhitachi2.naraからjuniper6.naraに移行

NARA NOC L2 Topology (Dec. 2013)

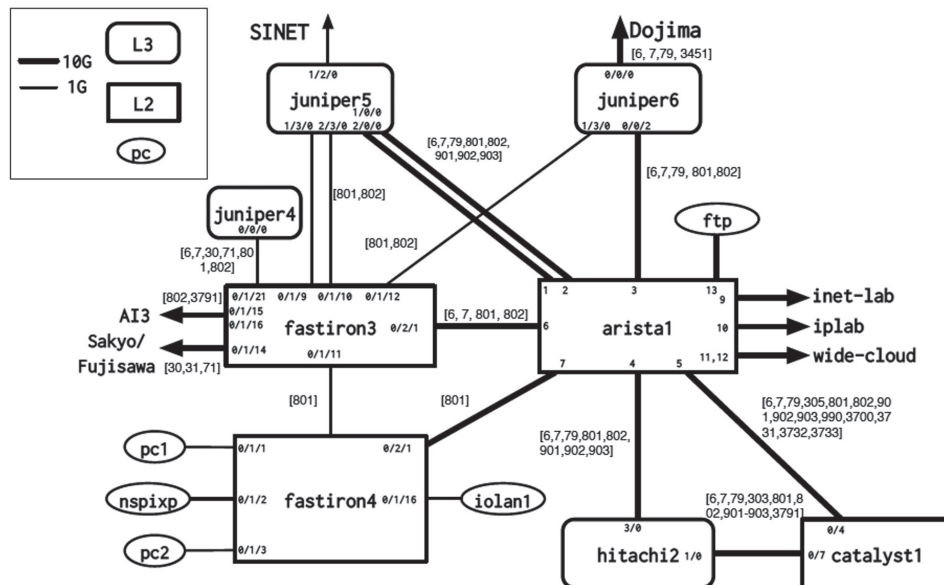


図2.16 奈良NOC Layer-2 トポロジ

NARA NOC L3 Topology (Dec. 2013)

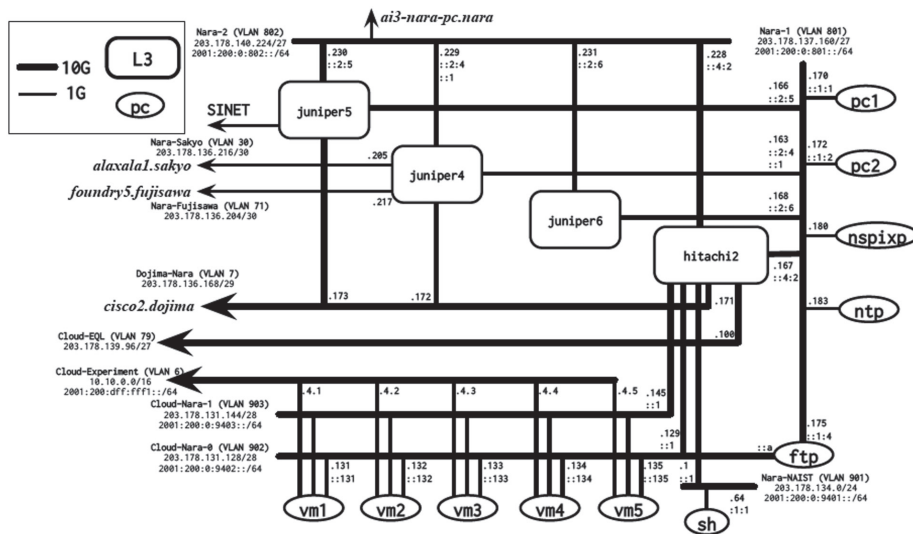


図2.17 奈良NOC Layer-3 トポロジ

2.13 左京

左京NOCは京都およびその周辺に存在する組織に対する接続拠点であり京都大学に設置されている。

- ・ (2013/10/25) 対奈良NOC回線の切り替え

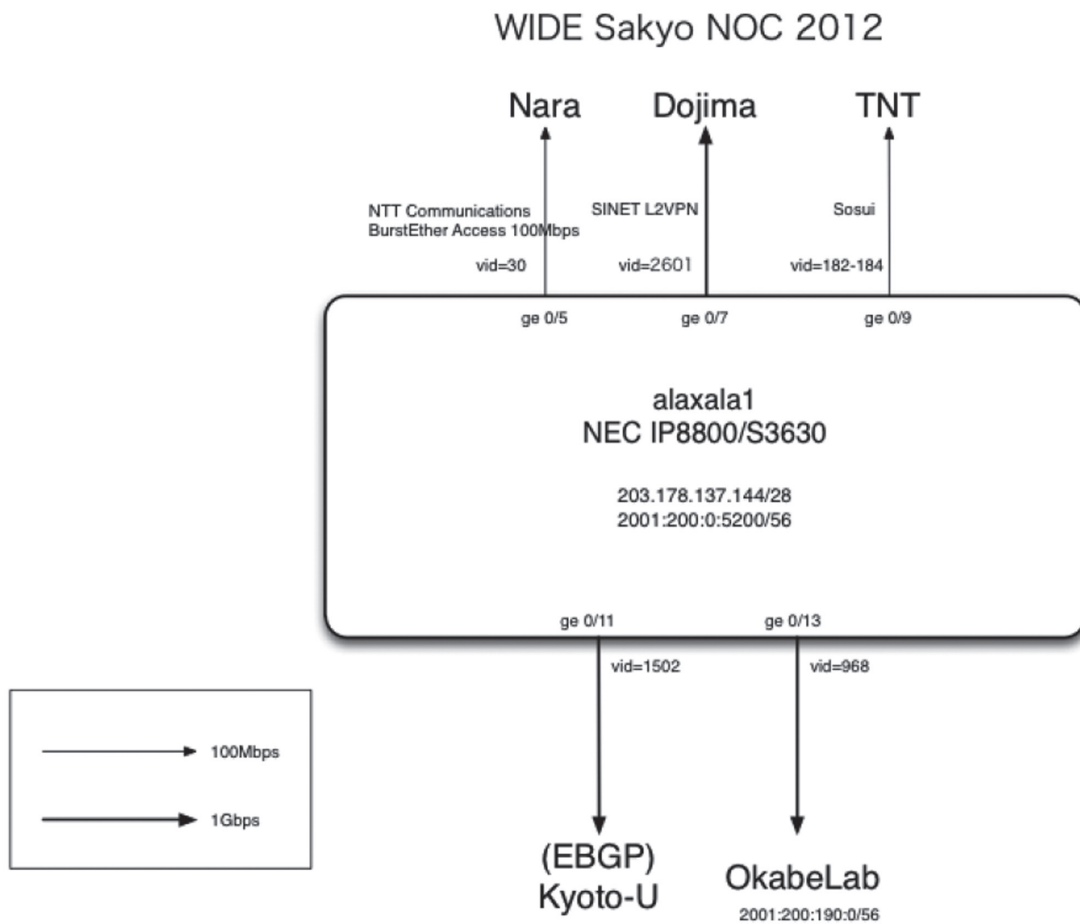


図2.18 左京NOC

2.14 倉敷

倉敷NOCは、平成24年度に学内ネットワーク機器の更新にともないNOC機器の更新など全体構成の変更を実施した。外部接続回線を収容していたGS4KはASR9Kに更新し、対外的なL3ルータのうちGR4KをJuniper MX80に更

新した。基本的な接続設定は、従前の装置の設定を踏襲している。また、倉敷NOC機器については、死活監視ツールで監視を行っているが特に大きな問題もなく運用されている。

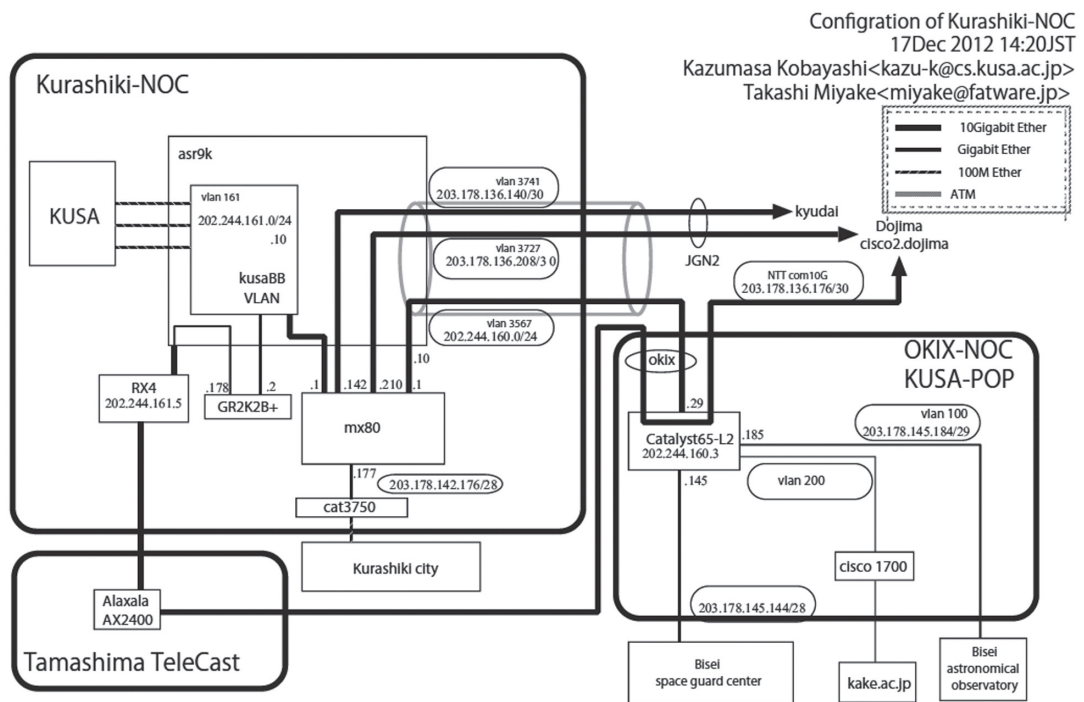


図2.19 倉敷NOC

2.15 広島

広島NOCは、2012年4月から広島市立大学にNOCの設置場所と管理担当者を変更し、また、VMとして動作するVyatta Routerというハイパーバイザとソフトウェアルータによる運用を行っている。サービス用のLinuxサーバも、同じVMとして動作させている。

数年前から接続性が失われていた福岡NOCとは、2013

年3月につながり、本来の冗長構成のトポロジとなった。ソフトウェアルータを使用しているが、安定性やパフォーマンスについても特に問題は発生していない。

- (2013/03/14) 福岡NOCとの接続確認
- (2013/03/25) 大学内でのサーバ室間でルータを移動
- (2013/09/01) 法令点検による計画停電

WIDE Hiroshima NOC

updated: 2013/03/26 hinoue@hiroshima-cu.ac.jp

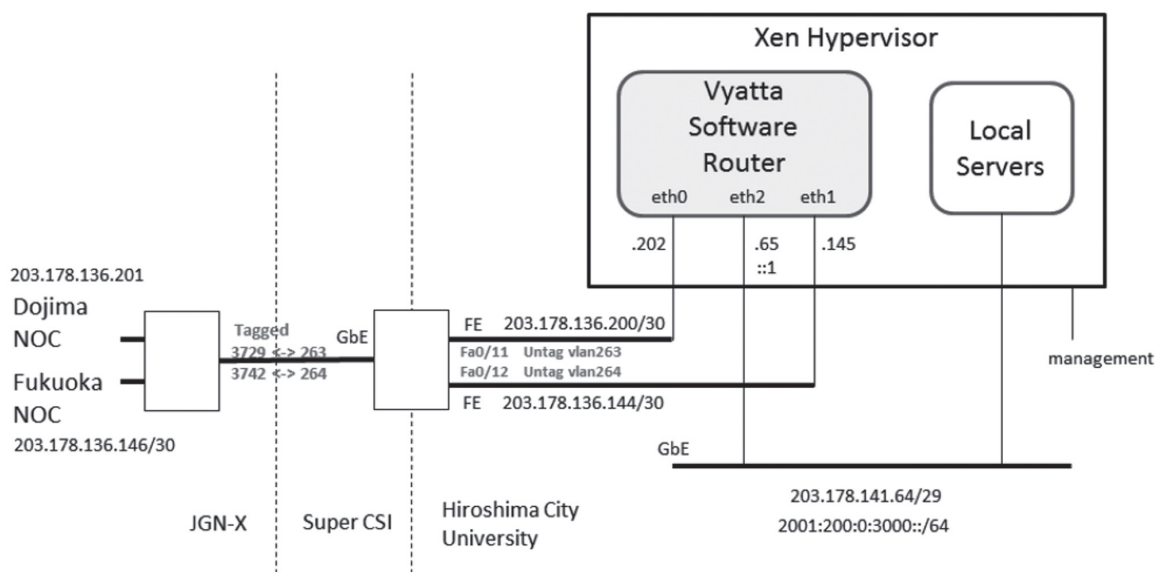


図2.20 広島NOC

2.16 福岡

福岡NOCでは、日立GR2000をコアルータとして運用を行っている。支線は2つあり、それぞれ帯域を必要としないローカル実験用の100Mbpsのセグメントと、グローバル実験用の1Gbpsのセグメントである。ローカル実験用の

経路情報は現在、インターネットには広告していない。

トポロジーに変更、接続回線の障害や停電もなく、安定して運用された。

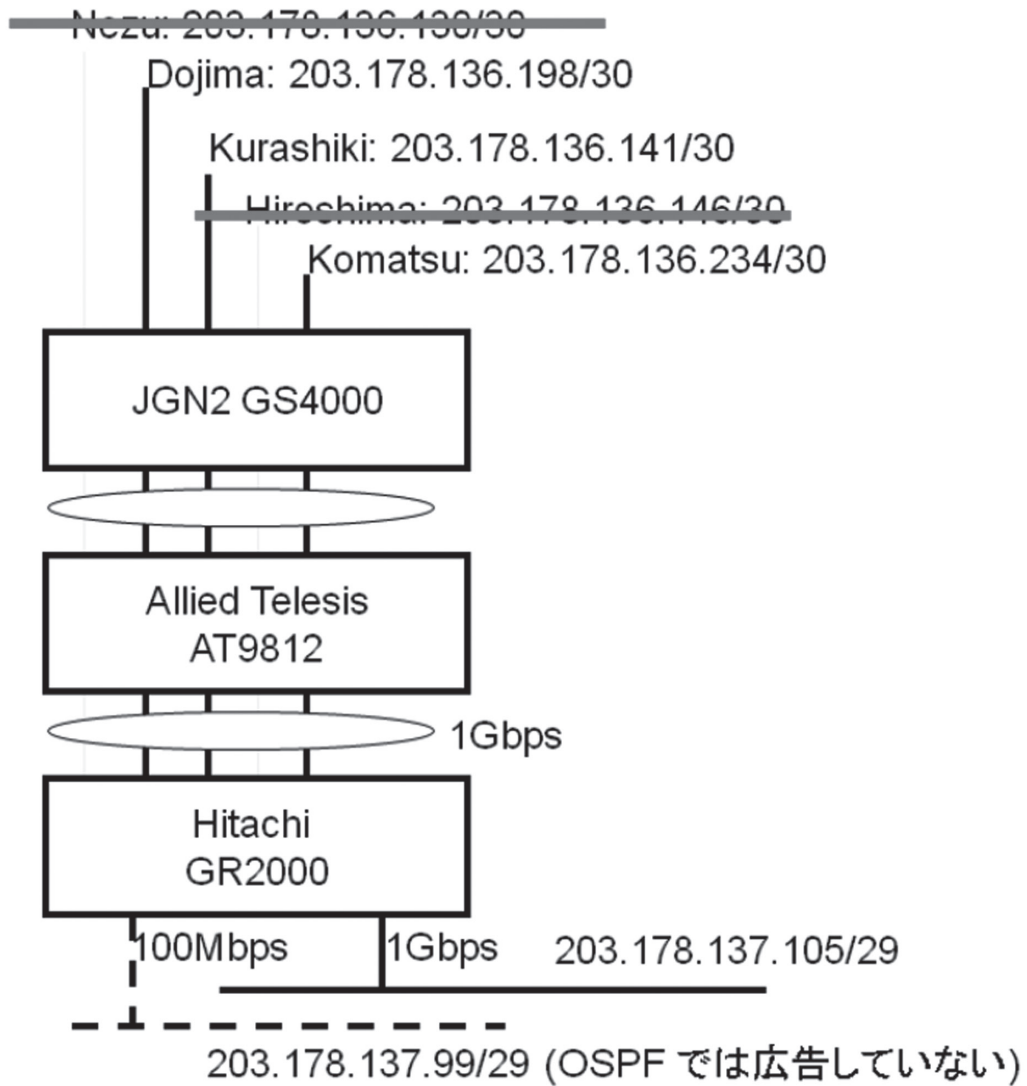


図2.21 福岡NOC

2.17 バンコク

2007年5月15日に設置されたバンコクNOCは、NECTECやUniNETといったタイの学術研究組織との研究活動強化を目的に設立された。今年度も引き続き、WIDEプロジェクトとしての独自の回線は存在しないが、JGN-Xの回線を利用し、VLANを用いてWIDEインターネットをバンコクまで延長し、IPv4、およびIPv6の接続性を提供している。バンコクNOCは、JGN-Xバンコク回線を収容しているNECTECと同じ建物に存在し、そこからUTPケーブルを延伸し、バンコクNOCが存在する部屋にネットワークを引いている。バンコクNOCの主な利用者は、バンコクを中心に活動しているSOI AsiaプロジェクトのメンバーであるPatcharee Basu、および関係者になる。

2013年1月に故障したルータ(pc1.bangkok)の代替とし

て、2013年10月にyamaha rtx810 が設置され、接続性が回復した。

- 2013年11月BTT(バンコクタイタワー)10Fの改修が終了。ThaiSARNおよびJGN-Xラックが10F内の最終設置場所に移動
- 2013年10月ルータ(RTX810)設置を行い、ユーザセグメントの接続性が回復される
- 2013年1月BTT(バンコクタイタワー)10F改修のため、ThaiSARNおよびJGN-Xラックが10F内の一時設置場所に移動
- 2013年1月WIDEバンコクオフィスが10Fから12Fに移動
- 2013年1月pc1.bangkok(ルータ)故障し、ユーザセグメントの接続性が無くなる

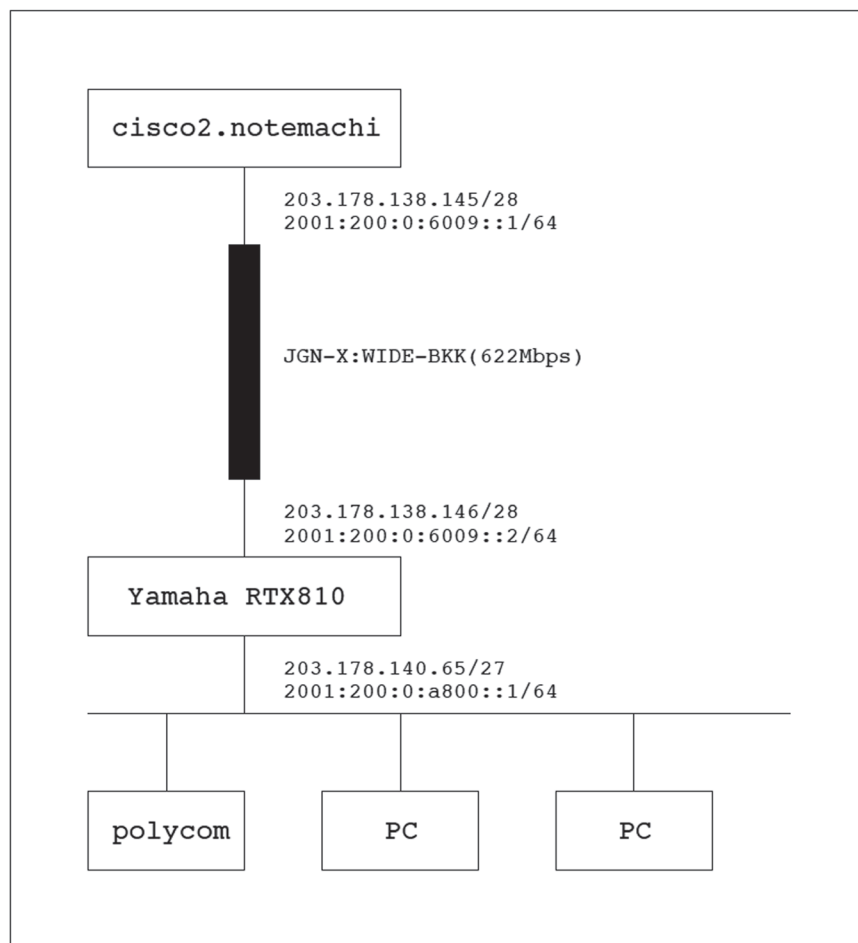


図2.22 Bangkok NOC

第3章 VXLANによるバックボーン貫通型VLANの構成変更

WIDEプロジェクトでは高精細な映像ストリーミングをはじめとした広帯域トラフィック伝送実験や、接続組織に対する仮想的な回線サービスに対して、バックボーン貫通型のVLANサービスを提供してきた。しかしながら、WIDEバックボーンのように自律分散協調的に運用される学術ネットワークにおいては、VLAN番号と端点の管理、ループ対策のプロトコル、ループに対する運用ルール等を統一的に扱うのは難しい。結果的に拠点のトポロジ変更や障害時のサービス継続性において問題が発生していた。

今日ではこう言った広域Ethernet的なサービスはレイヤ3ネットワーク上にオーバーレイネットワークを構成し、Ethernetをはじめとしたレイヤ2パケットをカプセル化した配送方法が主流になってきている。Ethernetパケットをレイヤ3上にオーバーレイネットワークを構成して配送する技術としては、L2TPやMPLSと言った技術が一般的であるが、接続ポイントの管理やネットワーク機器やそのOSでのインターオペラビリティの問題等で、WIDEバックボーンの構成機器と運用スタイルには合致しなかった。

近年その需要に対してVXLANがデータセンタ向けのマルチテナンシーを実現するために着目されている。WIDEバックボーンではVXLANを一つのデータセンタ内ではなく、バックボーン全体でレイヤ2セグメントの拡張に利用する。2013年はその下準備として藤沢NOC、根津NOCにて接続試験を行い、関東一円のバックボーンにてIPマルチキャストの疎通性を確保した。また、VXLANのVNIとIPマルチキャストアドレスのマッピング方法等を含めた運用方式について検討を行った。

今回接続試験に用いた実装はLinuxカーネル上での実装である。(Linuxカーネルはバージョン3.7よりVXLANをサポートしている)本実装はiproute2によってユーザランドからvxlanに対応したネットワークインターフェースを提供する形態をとっている。iproute2のリンクタイ

プvxlanではvxlanにおけるEthernetの仮想化識別子であるVNI、その識別子の中でBUM(BroadcastUnknownunicastMulticast)の配送に用いられるIPマルチキャストアドレス、マルチキャストパケットのTTLを指定し、Linuxのネットワークインターフェースとして抽象化されるデバイスに結びつけられる。下記にvxlan1と言う名前のネットワークインターフェースにVNIを1、マルチキャストアドレスを239.0.0.1、TTLを16、マップされる物理インターフェースをeth0とした場合のサンプルコマンドを示す。

```
ip link add vxlan 1 type vxlan id 1 group 239.0.0.1 ttl 16 dev eth0
```

LinuxカーネルサポートのVXLANを利用するに当たって注意すべき点は、対応する物理ネットワークインターフェース(IPエンキャプレーションを行うネットワークインターフェース)で、VXLANのヘッダサイズ分を考慮したMTUサイズを指定する必要がある点である。(今回試しては居ないが、他の実装においても同様の注意が想定される)

これらの実装を加味し、既存のVLANの管理からVXLANを用いた実装へシームレスに切り替え、かつVLAN IDの管理・運用をスマートに行う方法を議論した。議論の結果、まずは既存のVLANは1つのIDを1つのVNIとして運用し、各VNIは独立したIPマルチキャストアドレスとして運用することとした。その際の懸案としては利用するIPマルチキャストアドレスの数だけ途中経路のルータにIPマルチキャストの経路表エントリを維持する必要がある点、当然ながら同エントリをRPで管理しなくてはならない点が挙げられる。VXLANのInternet Draftにも記載されているように、IPマルチキャストの経路制御にPIM-SMだけでなくBidirectional PIMを用いて双方向のRPツリーによってトポロジの単純化等も検討に値することは事実である。(しかし、実際にBidirectional PIMをサポートするネットワークベンダは著者の知る限りではCiscoとJuniperのみである。)

現段階では実験としてEthernetのブロードキャストドメインがVXLANを用いて、バックボーンを渡るオーバーレイネットワークで延伸可能である点まで検証をおこなっ

た。2014年には広域なL2セグメントをVXLANを用いて構成し、バックボーン全体をシンプルなレイヤ3ネットワークに置き換える予定である。また、2014年の報告書では他の実装との比較や性能測定について詳しく記載する予定である。

第4章 おわりに

本年度もWIDEバックボーンネットワークの安定運用を行ってきた。来年度は、AS対外接続地点への計測ノードの増設、バックボーンルータやCall Managerの入れ替えを予定している。ほか、VXLANの本格運用やOSPF計測の再開など広域運用環境を用いた実験を精力的に行っていく予定である。

第5章 CopyRight

©2013 WIDE Project Two Working Group