

## 第 9 部

# マルチキャスト通信



# 第 1 章

## はじめに

MC WG では、JP Mbone コミュニティーと係し、日本国内の MBONE の運用に携わり、トラフィック監視やツールの開発を行い、安定運用に貢献している。ここでは、JP Mbone の現状を述べ、MC WG が提供している統計収集技術とその情報を報告する。また、WIDE 大手町 NOC で行なっている multicast exchange project について報告する。

## 第 2 章

### 国内の MBONE の現状

まず、日本国内の MBONE アクティビティを統計情報で示す。表 2.1 は、1999 年 5 月 7 日現在の、JP MBONE に接続されているルータと組織の数である。

表 2.1: JP MBONE ルータと組織数

	合計	ac.jp	ad.jp	co.jp	go.jp	gr.jp	or.jp	ne.jp	etcJP	notJP	NA
組織数	81	41	15	11	7	0	3	0	1	2	1
(active)	48	26	11	3	3	0	3	0	0	1	1
ルータ数	252	149	24	12	7	0	4	0	1	2	53
(active)	137	103	17	4	3	0	3	0	0	1	6

(etcJP = その他の jp、notJP = jp 以外、NA = DNS 未登録)

この調査は、DVMRP\_ASK\_NEIGHBORS2 の IGMP を用いて `mbone.otemachi.wide.ad.jp` から順に `threshold` が 64 未満の範囲をたどっていくことで得られるデータを元に算出した。

次に、JP MBone で使われているマルチキャストルータの種類を表 2.2 に示す。

現在 JP MBone において使われているマルチキャストルータの種類は二分されており、`mrouterd3.8` および `mrouterd3.9` といった Unix ホストをベースとした場合と、Cisco IOS version 11 系の場合に分けることが可能である。表 2.2 中の version 3.255 は、そのルータが DVMRP version3 [109] に従っていることを示す。現在この値を用いているのは `mrouterd 3.9` のみである。昨年の調査では、1:5 の割合でしか Cisco はマルチキャストルータとして用いられていなかったが、現時点ではほぼ同数のルータが稼働していることがわかった。また、MBone 全体で調査したルータ数とルータの種類を表 2.4 と表 2.6 に示す。

JP MBone における経路問題の解決および現状の把握をするため、JP MBone 以外とのリンクやループ構造を持つ部分に関する国内主要ルータについては接続トポロジ図を常に最新の状態にあわせ維持している。この 1 年の中での主な変更点は、SINET のハブルルータである `totoro.sinet.ad.jp` へのトラフィックの集中を避けるために各所に冗長経路を設定した点があげられる。また一部主要ルータ間の接続が変更され経路が変わるなどの変化があった。このようすを表 2.3 に示す。

表 2.2: JP Mbone ルータの種類 (1998 年 4 月 27 日)

ルータの種類	数
11.0PM	14
11.1PM	8
11.2PM	35
11.3PM	7
12.0PM	2
3.255PGM	50
3.6PGM	1
3.8PGM	20
not active	115
合計	252

表 2.3: JP MBONE のトポロジ (1998 年 4 月 27 日)

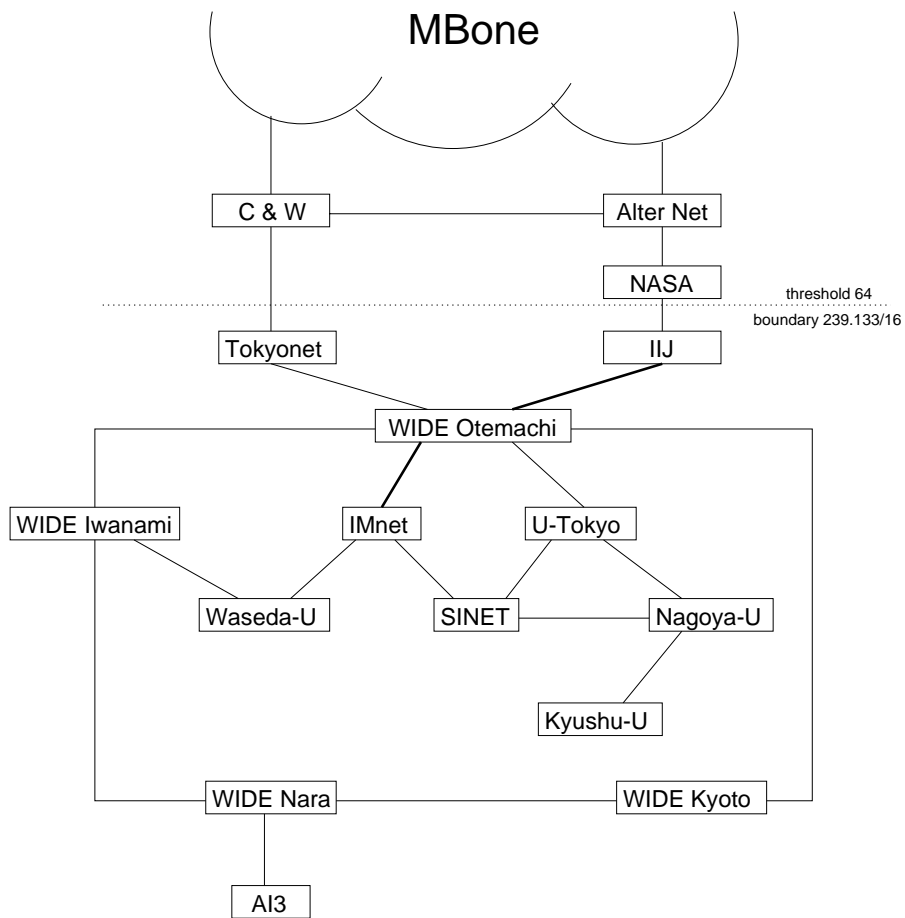


表 2.4: MBone ルータと組織数

ドメイン	ルータ数	ルータ数 (active)	組織数	組織数 (active)
at	11	5	1	1
au	24	20	3	3
be	22	15	2	2
ca	37	23	20	12
ch	32	25	8	6
cl	1	0	1	0
com	676	462	92	53
cz	36	32	10	10
de	9	5	3	2
dk	25	14	11	7
edu	415	308	79	59
ee	1	0	1	0
es	83	46	31	22
fi	109	88	21	20
fr	107	64	50	35
gb	1	1	1	1
gov	88	62	8	8
gr	62	56	14	13
hu	10	8	6	5
ie	2	1	1	1
in	1	0	1	0
int	1	0	1	0
it	35	27	18	15
jp	234	162	92	56
kr	7	7	1	1
lu	4	4	2	2
mil	88	58	10	8
my	2	1	2	1
net	1795	1281	139	120
nl	31	15	16	11
no	103	77	12	10
org	29	19	19	13
pk	1	1	1	1
pl	2	1	2	1

表 2.5: MBone ルータと組織数 (続き)

ドメイン	ルータ数	ルータ数 (active)	組織数	組織数 (active)
pt	3	1	3	1
se	41	27	9	8
sg	8	6	3	2
si	10	9	5	5
th	5	3	3	2
tw	6	4	1	1
uk	128	84	5	3
us	44	37	6	4
NG	545	169	1	1
合計	4874	3228	715	526

表 2.6: MBone ルータの種類

ルータの種類	数
1.0	7
10.2	6
10.3	9
11.0PM	46
11.1PM	943
11.2PM	623
11.3PM	256
12.0PM	832
2.0	1
3.255PGM	243
3.38PGM	5
3.6PGM	9
3.8PGM	248
NG	1646
合計	4874



## 第 3 章

### 経路情報総数の推移

JP MBone では、現状では DVMRP を用いて経路制御を行っている。また、海外を含めてもルータ数において、その半数は DVMRP を用いて経路制御を行っていることがわかっている。このように、距離ベクトル型プロトコルを広域で使用することはさまざまな問題をひき起こすことがわかっている。それを、箇条書にしてしめすと以下ようになる。すなわち、

- 経路の安定性
- 経路の収束性
- 経路情報量

である。現状を把握するために、経路数の推移を収集している。WIDE 神保町ルータ (sun1.tokyo.wide.ad.jp) において収集した経路数の推移を図 3.1 に示す。ここでいう経路数とは、(ソースネットワーク, グループアドレス) の対であり、この全て情報が全ルータの間で共有されることが DVMRP 経路制御に於いては必須である。

経路数の時系列推移をみてわかることは、ほぼ 4000 経路の上下を推移しており、増加よりも減少の傾向がみられるということである。MBone 参加サイトは増加していることを考えると、参加組織数に比べて経路が減少しているということが出来る。この要因としては、全 MBone ルータの半数を占める Cisco 製ルータにある”経路の集約機能”が用いられていることが考えられる。

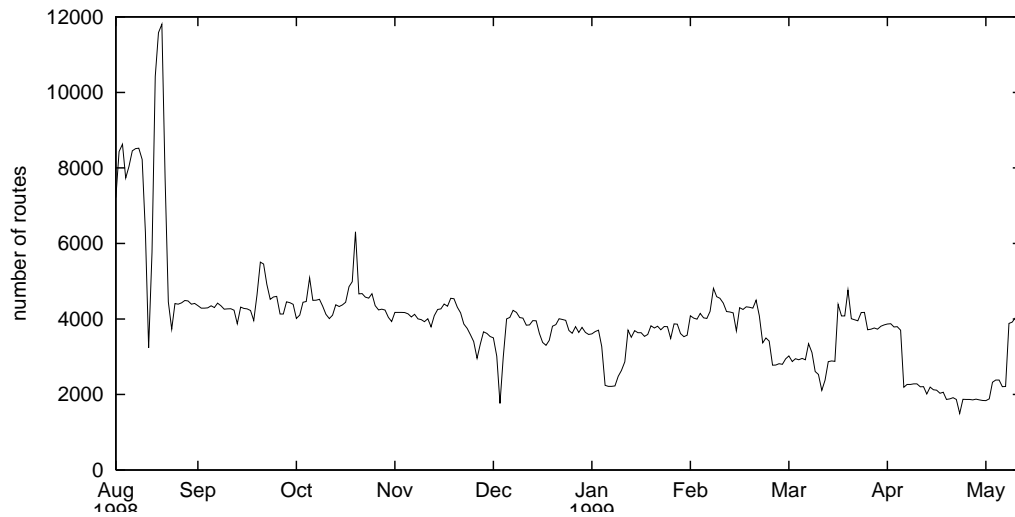


図 3.1: MBone における DVMRP 経路数

## 第 4 章

# Multicast Exchange Project

### 4.1 phase I

WIDE 及び JP MBone の接続トポロジ整理とともに当時増えつつあった商用 ISP の接続を整理するため、1997 年 3 月に WIDE 大手町 NOC に `mbone.otemachi.wide.ad.jp` を設置してトポロジの整理及び NSPIX2 を通しての ISP からの新規接続を受け付けていった結果、1998 年 3 月には 11 本のトンネルを抱えることになった。

これにより JP MBone 接続トポロジの整理をすることができたが、これら多数のトンネルを抱える `mbone.otemachi.wide.ad.jp` を中心とした放射型の接続トポロジとなっているため、次のような構造上の基本的な問題があった。

- 各 ISP 同士のやりとりが必ず WIDE 経由の間接やりとりになっている。
- 常に `mbone.otemachi.wide.ad.jp` を通過するため、そこが過負荷になる。
- もしも `mbone.otemachi.wide.ad.jp` が落ちると、すべて切断されてしまう。
- すべての接続がトンネル利用であり、また、`mbone.otemachi.wide.ad.jp` はユニキャスト的には各 ISP に対してルータでないため、すべてのトンネリングの packets が出入り口である 1 つの物理インタフェースに集中し、トンネル数が増えるとともに比例する形で通過するトラフィックが増えてしまう。

### 4.2 phase II

phase I における問題を解決するため、`mbone.otemachi.wide.ad.jp` のおかれている WIDE 大手町 NOC の近くにマルチアクセスのメディアのセグメントを一つ新たに用意し、各 ISP のマルチキャストルータをそのセグメントにすべて接続するという phase II 構想を発表した。[131].

具体的には 100BaseTX のセグメントを新規に設置し、このセグメントにおいて複数のマルチキャストルータが直接マルチキャスト packets をやりとりするようにする。これは、物理層レベルでのマルチキャストを使用することで複数のルータへ渡す時の packets の複

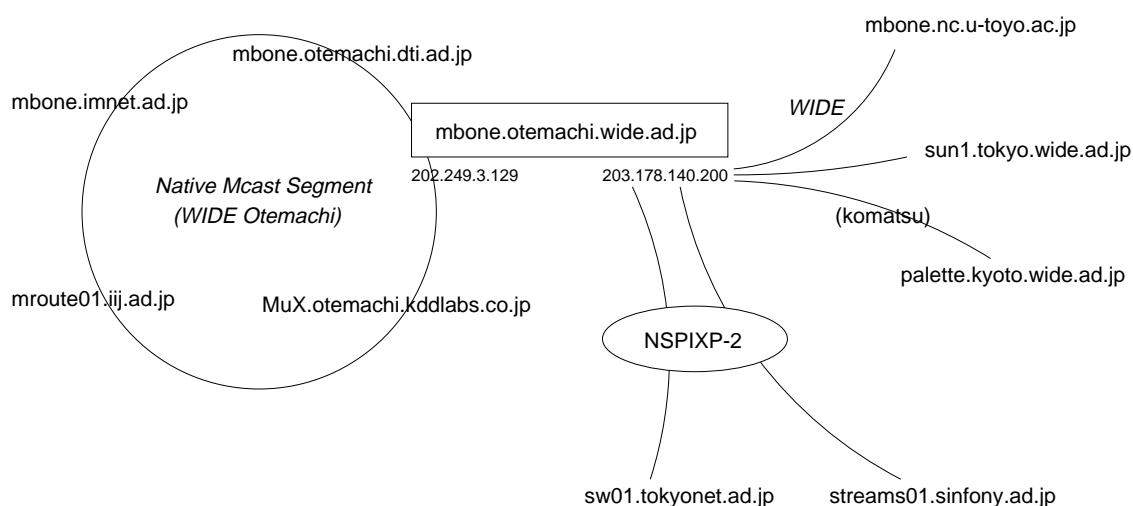


図 4.1: Multicast Exchange の接続図

製を避け、トラフィックの削減、ルータ負荷の低減を狙ったものである。今回 100BaseTX を採用した理由は、多くの ISP のマルチキャストルータが接続して本格的利用が始まると 10Base ではすぐに飽和する可能性があること、および、現在において PC などで 100Mbps 級を構成するのにもっとも安価なものは 100BaseTX とされるためである。

この計画に wide, apan, iij, dti, ttnet, ibmnet, imnet, jpix, sinfony, mesh, spin, psi, kddlabs, tokyonet, sphere の 15 の ISP が参加表明し、1998 年 7 月より実際の接続を開始した。具体的には WIDE 大手町 NOC に Fast Ethernet のノンスイッチングハブを設け、参加 ISP は各自のハウジングにある自組織のマルチキャストルータの 1 インタフェースをそこに接続する、という形をとった。このセグメントには 202.249.3.128/27 が割り当てられており、現時点で以下の組織のマルチキャストルータが接続されている。

- WIDE (mbone.otemachi.wide.ad.jp)
- IJ (mroute01.iij.ad.jp)
- IMNET (mbone.imnet.ad.jp)
- DTI (mbone.otemachi.dti.ad.jp)
- KDDLABS (MuX.otemachi.kddlabs.co.jp)

接続状況を図 4.1に示す。

### 4.3 稼働状況

従来、トンネルを流れるトラフィックは、BSD におけるトンネリング仮想インタフェースである vif に付与されたカウンターから SNMP エージェントが値を取得し、それを MRTG によって定期的に収集することで得られていたが、ネットワーク層(今回の場合、Ethernet)のマルチキャストを用いた場合、従来の方が使えない。vif レベルでは、特定の Ethernet インタフェース(これにも vif が付与される)に対してどれだけの量のパケットの出入りがあったかのみしか記載されない。

これを解決するために、libpcap [85] を用いて Ethernet のソースアドレス別にトラフィックを集計するソフトウェア mtraffic を実装した。

これにより、Ethernet インタフェースから出ていったトラフィックをどのルータが実際にフォワードしたかは(機構上)分からないながらも、どのルータからフォワードされてきたトラフィックであるかを判別することが可能となった。Phase II 開始から現在に至るまでの、主要なルータから exchange に対して発せられたトラフィックの量を図に示す。図 4.2 には WIDE 大手町ルータ (mbone.otemachi.wide.ad.jp) をソースとするトラフィック、図 4.3 には IJ ルータ (mrout01.ij.ad.jp) をソースとするトラフィック、図 4.4 には IMnet ルータ (mbone.imnet.ad.jp) をソースとするトラフィックが示されている。

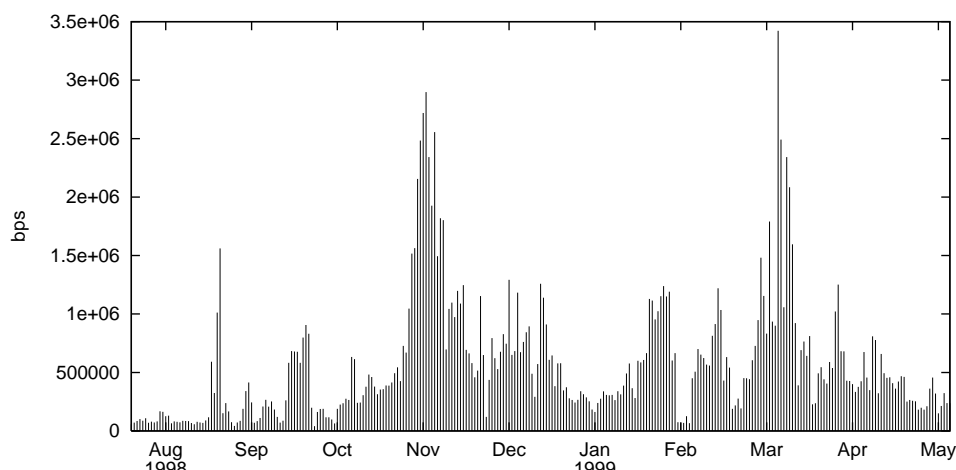


図 4.2: WIDE mbone.otemachi.wide.ad.jp)

従来と変わらずトンネルによって接続が継続されているサイトも残っている。主要なサイトである、東京大学 (mbone.nc.u-tokyo.ac.jp)、Sinfony (streams01.sinfony.ad.jp)、WIDE 東京 NOC (sun1.tokyo.wide.ad.jp)、Tokyonet (sw01.tokyonet.ad.jp)、IBMnet (mbone.ibmnet.ad.jp)、WIDE 京都 NOC (palette.kyoto.wide.ad.jp) に関して、一年のトラフィックを図 4.5 から図 4.10 に示す。

## 4.4 phase III

phase II では経路制御プロトコルとして従来と同じく DVMRP を用いているが、DVMRP の利用を続けていくには以下のような問題がある。

- 経路情報交換の安定性
- 経路数増大に対する耐性
- 政策的経路制御の可能性

特に ISP 間の経路情報交換においてはユニキャストの時と同様に経路選択の自由度を高めるために政策的経路制御ができることが好ましい。このため、phase II で設置したマルチキャスト交換のためのセグメントやその他の地点での接続においても ISP 間のマルチキャスト経路情報交換について脱 DVMRP の体制を構築することを目標にする phase III を今後進めていく。

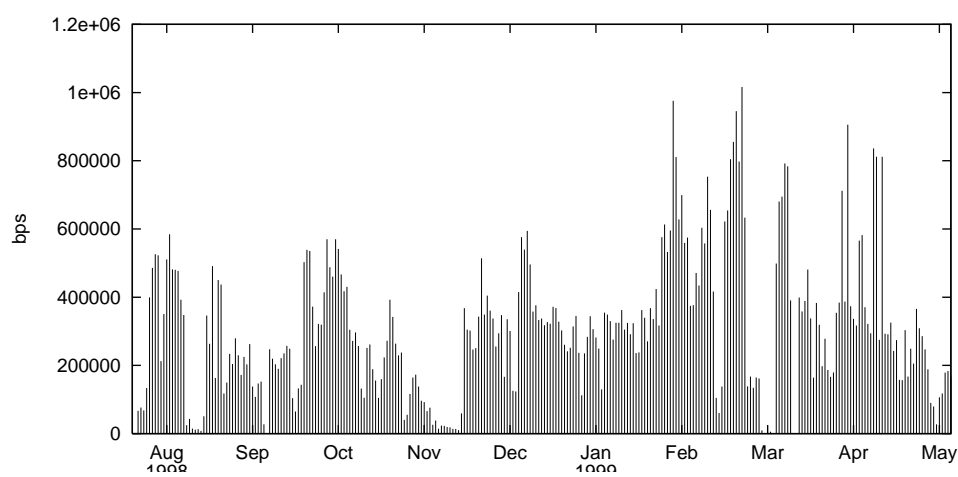


図 4.3: IIJ(mroute01.iij.ad.jp)

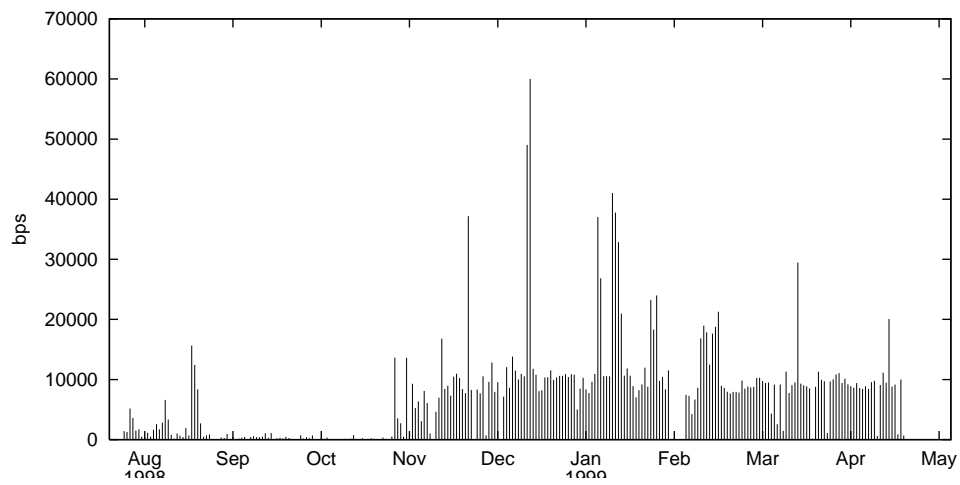


図 4.4: IMnet(mbone.imnet.ad.jp)



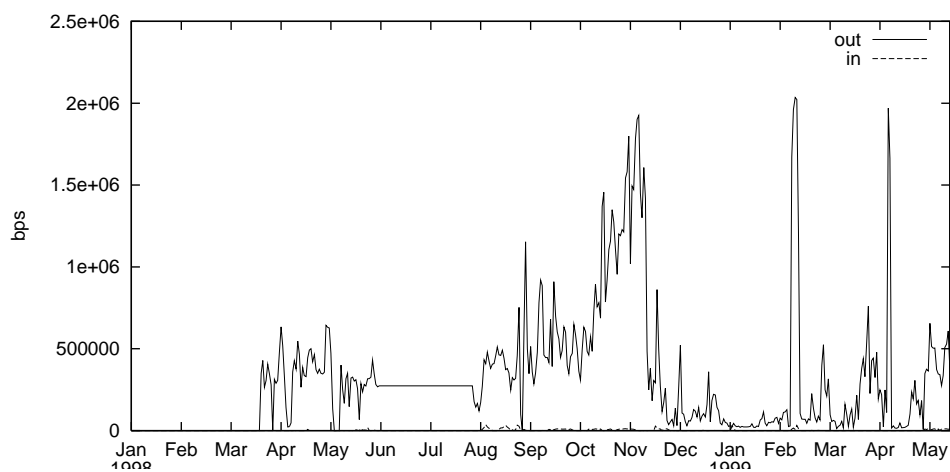


図 4.5: MIX-東京大学間 (mbone.nc.u-tokyo.ac.jp)

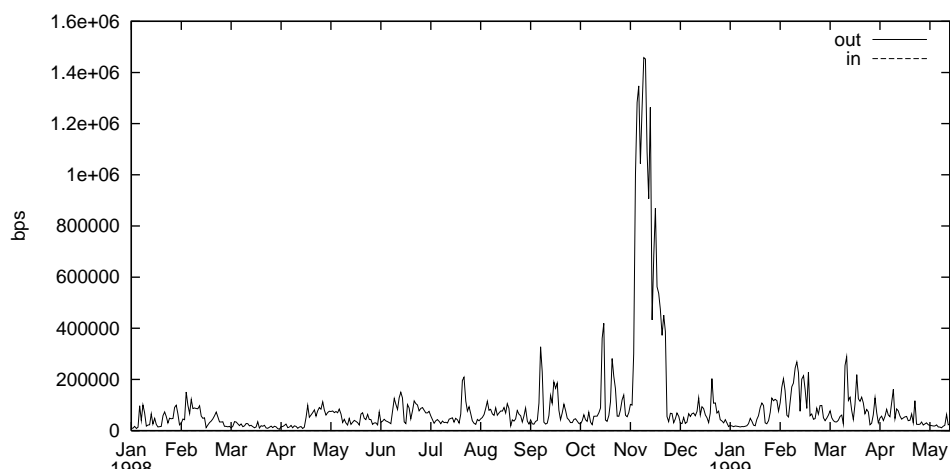


図 4.6: MIX-Sinfony 間 (streams01.sinfony.ad.jp)

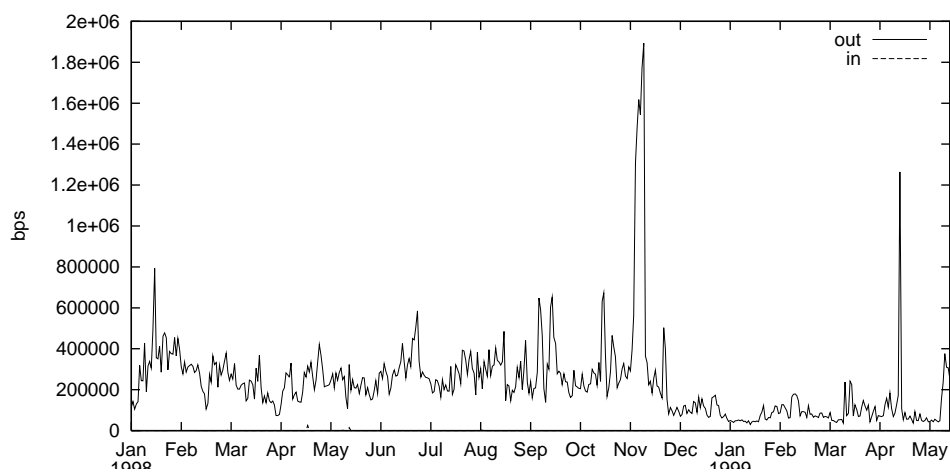


図 4.7: MIX-WIDE 神保町 NOC 間 (sun1.tokyo.wide.ad.jp)

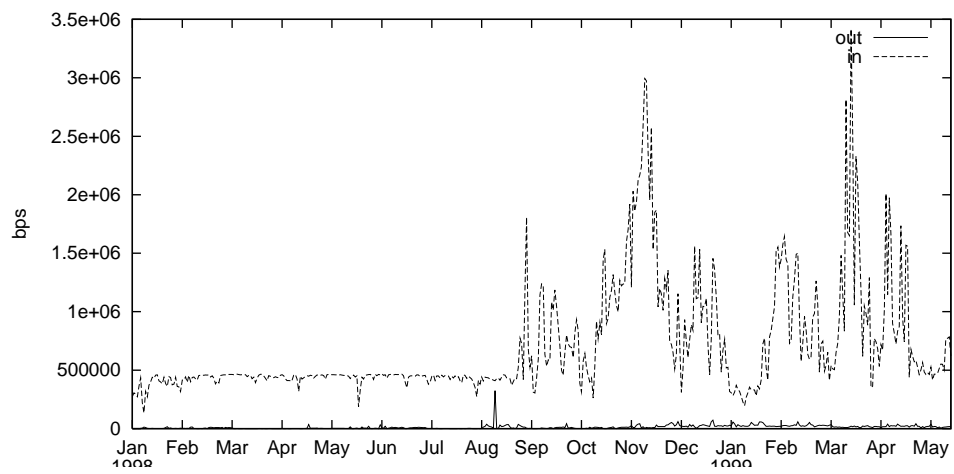


図 4.8: MIX-東京ネット間 (sw01.tokyonet.ad.jp)

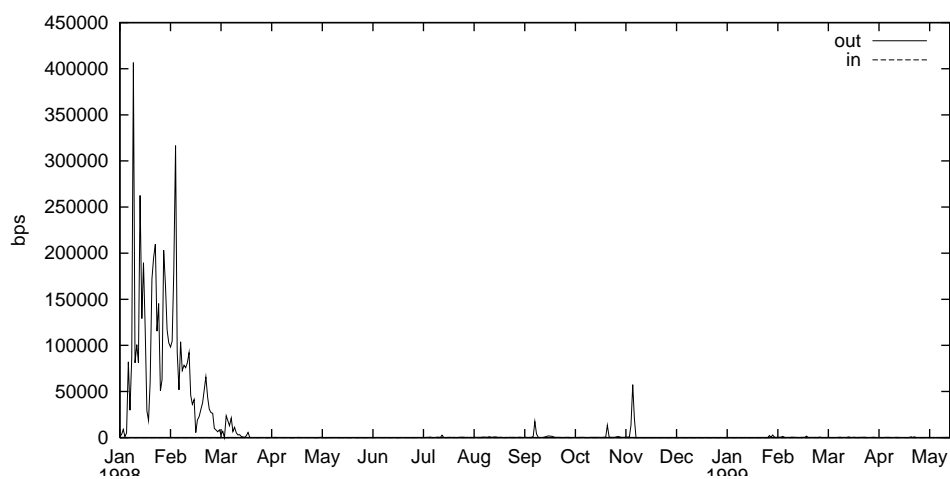


図 4.9: MIX-IBM ネット間 (mbone.ibmnet.ad.jp)

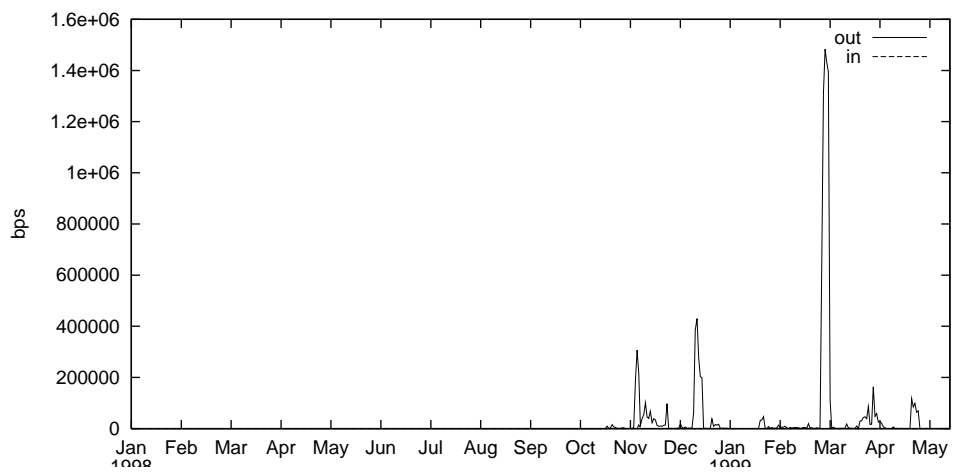


図 4.10: MIX-WIDE 京都 NOC 間 (palette.kyoto.wide.ad.jp)