社団法人 情報処理学会 研究報告
IPSJ SIG Technical Reports

2007－ＤＳＭ－45（14）
2007／5／11

フロー・メディエータを用いた大容量トラフィック・コレクタの設計

小林 淳史* 松原 大典** 近藤 毅* 石橋 圭介*

*日本電信電話株式会社 情報流通プラットフォーム研究所 〒180-8585 東京都武蔵野市緑町 3-9-11
**株式会社日立製作所 中央研究所 〒185-8601 東京都国分寺市東恋ヶ窪 1-280
E-mail: akoba@nttv6.net, daisuke.matsubara.pj@hitachi.com, {kondoh.tsuyoshi, ishibashi.keisuke}@lab.ntt.co.jp

あらまし 膨大なトラフィック情報を処理する次世代バックボーン向けのトラフィック監視システムは、柔軟性とともにスケーラビリティが要求される．従来、IP ネットワーク上のトラフィック監視システムは、エクスポータから配信されるトラフィック情報を受信し処理する単一のコレクタを示すものであった．本稿では、我々が提案してきた，エクスポータとコレクタの間に介在し、トラフィック情報の集約、振り分けを実施するフロー・メディエータを用いた大容量トラフィック・コレクタの適用例を報告する．

特に、スケーラビリティが要求される大規模ネットワーク向けには、このフロー・メディエータをもちいて大容量化が可能であることを実際のトラフィックデータをもちいて確認している．その結果をもとに、フロー・メディエータを用いた大容量トラフィック・コレクタの設計手法について報告する．

キーワード IPFIX, sFlow, NetFlow, IPFIX concentrator, IPFIX Mediator, メディエータ, フロー

The Designing for Large-Scale Traffic Collector using Flow Mediator

Atsushi KOBAYASHI* Daisuke MATSUBARA** Tsuyoshi KONDOH* Keisuke ISHIBASHI*

*NTT Information Sharing Platform Laboratories 3-9-11 Midori-cho, Musashino, Tokyo, 180-8585 Japan
**Hitachi, Ltd., Central Research Laboratory 1-280 Higashi-Koigakubo, Kokubunji, Tokyo, 185-8601 Japan
E-mail:* akoba@nttv6.net, daisuke.matsubara.pj@hitachi.com, {kondoh.tsuyoshi, ishibashi.keisuke}@lab.ntt.co.jp

Abstract Recently traffic volume of Internet backbone network becomes more and more, and the diversity of networks increases, such as IPv4, IPv6 and several tunnels. To handle the huge traffic information and several kind of information, traffic monitoring system needs scalability and flexibility. Usually, traffic monitoring system is composed of multiple exporters and a collector. The exporter such as a router or switch can export the traffic information as flows, through the NetFlow, sFlow and IPFIX. On the other hand, the collector can receive them and then store and analyze.

In this paper, we propose the Flow Mediator which is intermediate node between the exporter and the collector. This Flow Mediator can concentrate traffic information and distribute them. We introduce the applicability of Flow Mediators for scalable and flexible traffic monitoring system. And also, we demonstrate the feasibility study for designing of large-scale traffic monitoring system using Flow Mediators, based on the characteristic of Internet backbone traffic.

Keyword IPFIX, sFlow, NetFlow, IPFIX concentrator, IPFIX Mediator, flow

1. はじめに

近年、ネットワークの設備計画や DDoS 攻撃などの異常トラフィックを検知する目的から、ネットワーク上に流れているトラフィックを監視する機能（sFlow[1],NetFlow[2],IPFIX[3]）に注目が集まっている．しかし、ネットワーク上に流れるトラフィック量は、年々増加傾向にあり、トラフィック測定をする際に、出力されるトラフィック情報も膨大となりつつある．現在、日本のインターネットトラフィックは、年 1.2 倍のペースで増え続けており、ブロードバンド加入者のダウンロード・トラフィック量は、523 Gbps 程度との報告がある[4].

このため、トラフィック情報を受信し処理するためのトラフィック監視システムには、拡張性を維持するための大容量化が必要とされる．また更に、IPv4, IPv6, 多種のトンネルプロトコルの使用により、流通するトラフィックも多様化が進み、これらのトラフィック情報を扱うための柔軟性も要求される．特に、MPLS 網などの VPN サービスにおいては、その必要性が増すものと考える．

従来、トラフィック監視システムは、トラフィック情報を配信するエクスポータとコレクタによって構成される．エクスポータは、主にルータ、スイッチがこれに該当する．一方、コレクタは、トラフィック情報を受信し、蓄積・分析を行うノードを指す．

本稿でその適用方法を報告するフロー・メディエータ[8][9]は、エクスポータとコレクタの間に介在し、トラフィック情報となるフロー情報を中継するノードを総称する．我々が提案を行ってきたフロー情報を集約するフロー・コンセントレータ[5][6][7]やフローの情報をもとに上位のコレクタに振り分けを行うフロー・ディストリビュータ、そしてプロキシなどもこれに包含される．ここでいうフローとは、共通の属性をもつトラフィック情報の集合体を表す．代表的なものは{送信 IP アドレス，受信 IP アドレス，送信ポート番号，受信ポート番号，プロトコル}が共通となる 5-tuple フローであるが，ここではそれらを集約した情報についても広義にフローとしている．また、ここでいう集約とは、ある単位時間に発生した共通の属性をもつフローに対して、バイト数、パケット数などを加算して 1 つのフローにすることをさす．

本稿では、フロー・メディエータを使用したトラフィック情報収集システムの大規模ネットワーク及び多様化したネットワークへの適用手法を提案するとともに、情報収集システム全体の大容量化に向けた適用性の評価を実施したので、その結果を報告する．

2 節では、フロー・メディエータを使用した大規模かつ柔軟なトラフィック情報収集システムの適用事例を紹介する．更に 3 節では実際のトラフィックデータをもちいて、集約によるフロー数の削減効果を評価した結果を示し、4 節で，3 節の結果をもとに実際の大規模ネットワークに適合した大容量トラフィック情報収集システムの設計方法を報告する．

2. フロー・メディエータを用いた適用モデル

フロー・メディエータを用いたトラフィック情報収集システムの適用モデルとして、大規模ネットワークでの適用モデルについて検討する．

2.1. フロー・メディエータ概要

フロー・メディエータはコレクタとエクスポータの間に介在し、フロー情報を仲介する以下の機能をもつ．

- 複数のコレクタからのフロー情報を収集する．
- 受信したフロー情報を蓄積する．
- 柔軟なキー属性をもとにフロー情報を集約する．
- フロー情報にもとづき、上位のコレクタにフロー情報を振り分け配信する．

これにより、1 つのフロー情報を様々な用途で活用する場合にそれぞれの上位コレクタに振り分けることや 2.2 節で示すような集約によって上位コレクタの負荷を軽減することが可能となる．以下に、フロー・メディエータの機能構成図を示す．

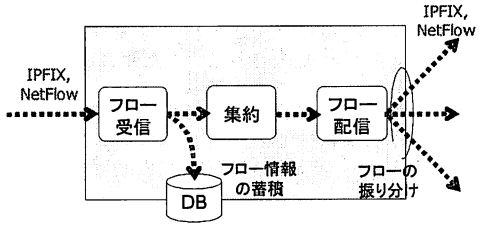


図 1. フロー・メディエータの機能構成．

2.2. 大規模ネットワークでの適用モデル

大規模ネットワーク上でのトラフィック情報収集システムの性能要件は、膨大なフロー情報を処理するためのスケーラビリティを確保することである．この一方、網全体に流れるトラフィック動向を俯瞰的に把握可能とするトラフィック交流量の測定を実現する機能も必要とされる[10]．トラフィック交流量の集計を単一コレクタで実現するため、予めエクスポータにて集約したフローを生成し、フロー数を削減した上で配信することが必要とされる．しかし，この場合、集約後のフローではインシデント発生時に原因探索のための情報が不足するという問題が発生する．これを解決するため、フロー・メディエータの以下の機能を活用し、トラフィック交流監視コレクタの負荷を軽減させつつ、詳細なフロー情報を蓄積することを可能とする．

- 受信したフローを一定期間蓄積し、詳細なトラフィック情報の検索を可能とする．
- 受信したフローを一定間隔で集約を行い、集約後のフローを更に上位のコレクタに配信する．

フロー・メディエータは、エクスポータから配信されるトラフィック情報を蓄積した上で、集約を行い、再配信することによって詳細な情報を失うことなく、トラフィックの俯瞰的な監視を可能とする．集約の手法としては、BGP Next-Hop 毎にトラフィック情報をまとめるもしくは Prefix Mask を用いてネットワークアドレス毎に情報をまとめる集約が適用される．これにより、フロー・メディエータにて蓄積した詳細なトラフィック情報をもとに原因探索を可能とする．また、これを地域拠点（PoP）単位に配置することで、管理用ネットワークのトラフィック量の削減も可能となる．この適用事例を図 2 に示す．

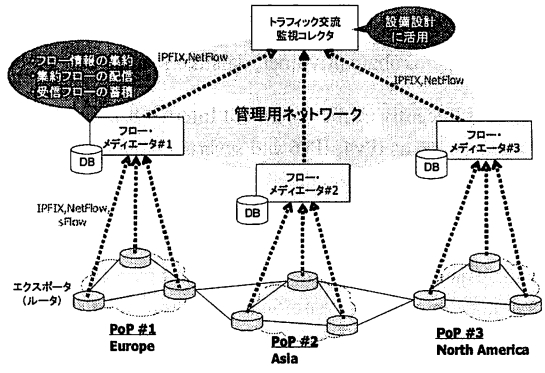


図 2. 大規模ネットワークでの適用例．

3. 集約によるフロー情報の削減効果

図 2 に示すようなフロー・メディエータの大規模ネットワークへの適用を進めるに、各集約方式によるフロー数の集約効果を考慮する必要がある．特に、受信フロー数を個々のコレクタ、メディエータの性能限界に見合う量とするため、集約及びサンプリングレートを調節することによって、扱うフロー数を削減することが必要とされる．このため、本節では、トラフィック量に応じてエクスポータから配信されるフロー数の見積もりと，そのフロー情報に対してメディエータに

て集約を行った場合のフロー数の削減効果について実際のトラフィックデータをもとに検討した結果を示す．

3.1 節では、サンプリングレートをもとにエクスポータから配信されるフロー数をモデル化する．3.2 節では、そのフロー数に対する集約後のフロー数を算出する．集約フロー数の算出は、サンプリングレート，母体トラフィック量,集約種別,集約時間間隔などの複数のパラメータが関わるが、これらは、全て集約対象とするフロー数に帰着できる．ここでは、集約対象のフロー数に基づく削減効果をモデル化している．

3.1. エクスポータから配信されるフロー数の推定

ルータであるエクスポータから配信されるフロー数（$f_{sampled}$）は、1 フローあたりのパケット数（x）の分布関数 $p(x)$もとに以下のように表される。

$$f_{sampled} = \sum_{x=1}^{\infty}\left[\left(1-\left(1-1/r\right)^{x}\right)\times p(x)\right]\times f_{all}$$

ここで f_{all} は、全トラフィックのフロー数を示し、1/r は、サンプリングレートを示す．全フロー数を示す f_{all}は、全トラフィックのパケット・レートから 1 フローあたりの平均パケット数を単純に除算することによって概算を求めることができる．

分布関数 $p(x)$を求めるために、MAWI[11]のトラフィックデータをもちいて、1 フローあたりのパケット数の分布を算出した．ここでは、ActiveTime=1 分（最長フロー持続時間）、InActiveTime=15 秒（無通信時間）として算出している．サンプルに用いたトラフィックデータと 1 フローあたりパケット数分布の結果を以下に示す．

表 1. MAWI のサンプルデータ.

	データ①	データ②	データ③
抽出日時	2007/01/07	2007/02/01	2007/02/07
全パケット数	7,524,267	20,074,663	19,917,914
全トラフィック量 (MByte)	4715.71	13214.83	11861.35
平均レート(Mb/s)	43.9	123.11	110.56
平均レート(kp/s)	8.4	22	22
平均パケット長 (Byte)	626.73	658.28	595.51
全フロー数	319538	772949	832988
1 フローあたりの平均パケット数	23.54	25.97	23.91

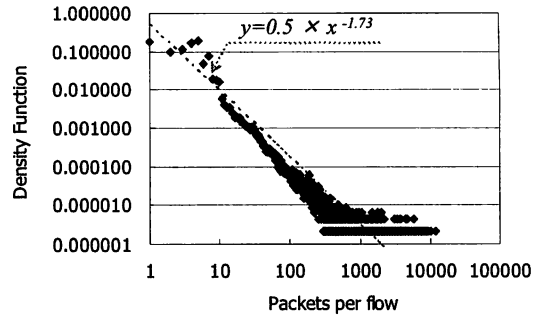


図 3. フローあたりのパケット数分布.

ここで、$p(x)$は、$0.5 \times x^{-1.7}$にてよく近似されていることがわかる．1 フロー当たりのパケット数を図 3 に示すパケット数（x）の最大値 10000 とすると、$f_{sampled}$は、以下のように表される．

$$f_{sampled} = \sum_{x=1}^{10000}\left[\left(1-\left(1-1/r\right)^{x}\right)\times 0.5\times x^{-1.73}\right]\times f_{all}$$

3.2. フロー・メディエータによる集約効果

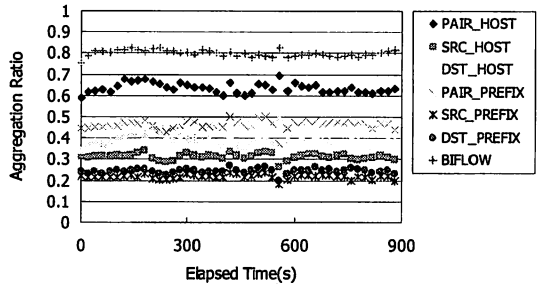
フロー・メディエータによる集約処理は、特定の時間間隔内にエクスポータから受信するフローに対して複数のフロー識別子をキー情報として柔軟な集約を実現可能とする．以下に主な集約種別を示す．

表 2. 各集約方式の具体例.

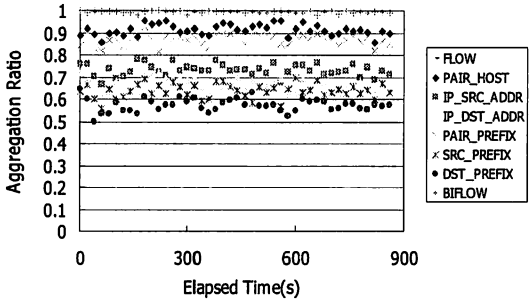
集約種別(略称)	キー情報とするフロー識別子
宛先ホスト（DST_HOST）	宛先 IP アドレス向けの同一のフロー情報を集約
送信元ホスト（SRC_HOST）	送信元 IP アドレス向けの同一フロー情報を集約
ペアホスト（PAIR_HOST）	宛先 IP アドレス,送信元 IP アドレスの対のフロー情報を集約
宛先プレフィックス（DST_PREFIX）	宛先ネットワークアドレス向けの同一のフロー情報を集約
送信元プレフィックス（SRC_PREFIX）	送信元ネットワークアドレス向けの同一のフロー情報を集約
ペアプレフィックス（PAIR_PREFIX）	宛先ネットワークアドレス，送信元ネットワークアドレスの対のフロー情報を集約
双方向フロー（BI-FLOW）	宛先 IP アドレス,送信元 IP アドレス,宛先ポート番号,送信元ポート番号,プロトコルのフロー情報の双方向の情報を 1 フローとして集約

集約後のフローには、集約前の対象フロー数などが情報要素に追加される他に、フロー開始時間、終了時間についても、集約前のフロー情報をもとに生成される．サンプリングされたフロー情報をもとにフロー・

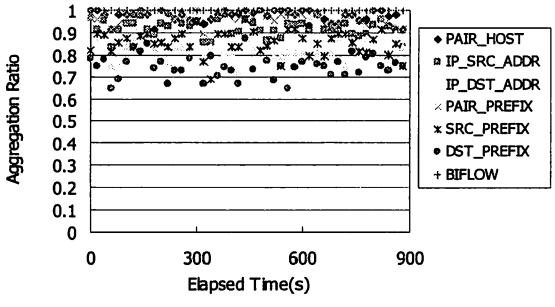
メディエータにて集約した場合の集約率の推移を算出した．前述したMAWIのトラフィックデータをもちいて、サンプリングレート=1/1, 1/128, 1/1024の場合で、それぞれ算出している．この結果を以下に示す．



(a) サンプリングレート=1/1



(b) サンプリングレート=1/128

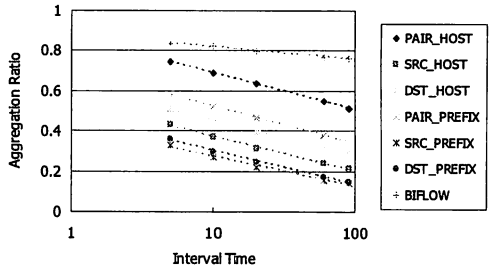


(c) サンプリングレート=1/1024

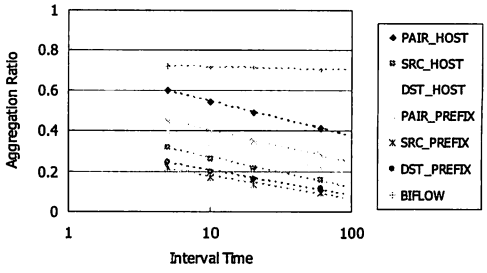
図 4. サンプリングによる集約効果の推移.

これによると、予めエクスポータによってサンプリング処理された場合は、集約対象となるフローが減少するため，集約効果が少なくなる．

次に、サンプリングレートを同一として、集約時間間隔を変化させたときの集約率の変化を評価した結果を図.5 に示す．



(a) サンプルデータ①: サンプリングレート=1/1



(b) サンプルデータ③: サンプリングレート=1/1

図 5. 集約時間間隔による集約効果の推移.

集約時間が長くなるほど，1 つのフローに集約される集約対象のトラフィック量が増加し，結果的に集約率が向上していることがわかる．また，表 1 のサンプルデータ①と③を比較すると母体のトラフィック量がサンプルデータ③の方が大きく、その分、集約効果も大きいことがわかる．以上のことから、集約による集約効果は、以下の 3 つの要因が関わっているといえる．

- 母体トラフィック量
- サンプリングレート
- 集約時間間隔

いずれの要因も集約対象とするフロー数が増えるに従い、集約効果も高まるのではないかと推測される．実際に異なるサンプリングレート,集約時間間隔で、同量のフロー数を集約した場合の集約効果を表 3 に示す．

表 3. 集約対象フロー数による集約効果への影響.

集約対象フロー数	3450	3562
集約時間間隔（s）	10	300
サンプリングレート	1/1	1/128
集約率（DST_HOST）	45%	43%
集約率（DST_PREFIX）	30%	32%

これによると、集約率は，上記の要因の複合結果である集約対象フロー数にのみ依存して決定されると仮定できる．

図 6 に、集約時間間隔内に受信するフロー数に対する集約効果の結果を以下に示す．これは、サンプルデータ①,②,③をもちいて、サンプリングレート=1/1 ～ 1/1024、集約時間間隔=5 秒 ～ 300 秒として算出した結果を示したものである．

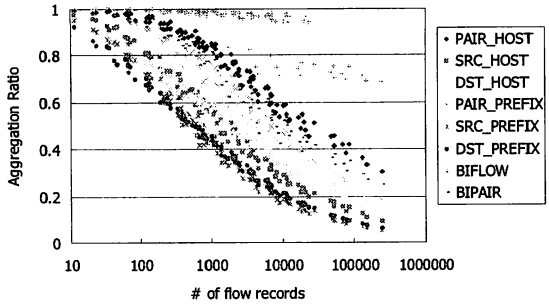


図 6. 集約対象フロー数に対する集約効果.

上記の図によると DST_HOST 集約,DST_PREFIX 集約の場合に、集約率 (R)、集約対象フロー数 (f_{total}) は以下のように近似することができる.

- DST_HOST 集約: $R=1.80\times f_{total}^{-0.18}$
- DST_PREFIX 集約: $R=2.34\times f_{total}^{-0.26}$

以上より、エクスポータから配信されるフロー数の概算とフロー・メディエータによって、集約後、再配信されるフロー数の概算を導出することが可能となる. これにより、ネットワーク規模に応じて、メディエータの配置,集約処理方式を調節することにより拡張性の高い大容量トラフィック・コレクタを構築することが可能となる.

4. 大規模ネットワークへの適用モデル

前節にて示したサンプリングレートによる配信フロー数及び集約効果の概算結果をもとに、大規模ネットワークへのフロー・メディエータを使用した大容量トラフィック・コレクタの適用モデルを検討する.

ここで、想定するネットワークは、2010 年でのシェア 20%規模プロバイダのネットワークをモデルとする. 総務省報告[4]の 2006 年 5 月時点のブロードバンド契約者ユーザのダウンロードトラヒックは、523.6 Gbps とある. これより、アップロードもダウンロード同等のトラフィック量と仮定し、今後も 1.2 倍/年のペースでトラフィックが増加していくと仮定すると想定ネットワークのトラフィック交流量は、440Gbps 相当になるもとと想定される.

$$(523.6\ \text{Gbps}\times 2)\times 1.2^4\times 0.20 = 440\ \text{Gbps}$$

ここで、440Gbps (平均パケットサイズ約 550Byte として 100Mpps 相当) のトラフィック量を全国規模で運用するネットワークモデルを対象に、フロー・メディエータを用いた大容量トラフィック監視システムの規模・集約方式を検討することとする. 図 7 に想定するネットワークの構成モデルを示す.

- 1 拠点あたりのエッジルータ : 20 台
- 1 拠点あたりのコアエッジルータ : 2 台
- ネットワーク全体 10 拠点
- ルータ台数 : コアルータ:4 台,コアエッジルータ:20 台,エッジルータ:200 台
- フローの観測ポイント:コアエッジの全 EgressIF

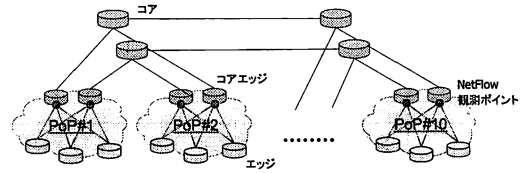


図 7. 想定ネットワーク構成モデル.

また、各拠点にフロー・メディエータを配置した場合のトラフィック監視システムの構成を図 8 に示す. ここでは、メディエータの性能を 10kf/s とし、上位ノードのトラフィック交流監視コレクタの性能を 5Kf/s として、適切な集約方式、サンプリングレートを求めることとする.

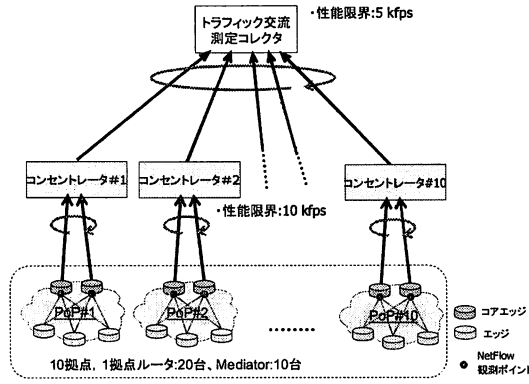


図 8. メディエータを用いたコレクタ構成.

前節 2.1 項より求めたエクスポータから配信されるフロー数の推定結果をもとに、100Mpps のトラフィック量に対してサンプリングレート=1/100, 1/1000, 1/10000 を適用した場合のフローレート (f/s) は表 4 のようになる. ここで、全フローレート (f_{all}) は、1 フロー20 パケット平均とし 5Mf/s として算出している. また、総トラフィック量は、コアエッジに均一に配分されるものと仮定している.

表 4. サンプリングによるフロー数の推移.

サンプリングレート	1/100	1/1000	1/10000
全フローレート ($f_{sampled}$)	305 kf/s	43 kf/s	5.2 kf/s
コアエッジ 1 台あたりのフローレート	15.3 kf/s	2.2 kf/s	0.3 kf/s

ここで、2.2 項に示したフロー・メディエータによる集約効果の結果から集約時間間隔を 1 分,5 分とした場合の集約後のフローレートの推定値を求めることができる. これを表 5 に示す.

表 5(a). 集約フロー数の推定（集約時間間隔 1 分）.

サンプリングレート	1/100	1/1000	1/10000
集約対象フロー総数(ftotal)	918 k フロー	132 k フロー	18 k フロー
DST_HOST	集約率:15% 45 kf/s	集約率:21% 9.0 kf/s	集約率:31% 1.6 kf/s
DST_PREFIX	集約率:7% 21 kf/s	集約率:11% 4.7 kf/s	集約率:18% 0.94 kf/s

表 5(b). 集約フロー数の推定（集約時間間隔 5 分）.

サンプリングレート	1/100	1/1000	1/10000
集約対象フロー総数(ftotal)	4.6 M フロー	660 k フロー	90 k フロー
DST_HOST	集約率:11% 34 kf/s	集約率:16% 7.0 kf/s	集約率:23% 1.2 kf/s
DST_PREFIX	集約率:4% 12 kf/s	集約率:7% 3.0 kf/s	集約率:12% 0.62 kf/s

この結果、フロー・メディエータ,トラフィック交流コレクタの性能限界に見合うフローレートと対応するサンプリングレート、集約方式を選択すると，例えば以下のようになる.

- エクスポータでのサンプリングレート:1/1000
- メディエータでの集約方式: DST_PREFIX 集約
- 集約時間間隔 1 分
- 1 メディエータの受信フローレート:4.4 kf/s
- 上位トラフィック・コレクタの受信フローレート:4.7kf/s

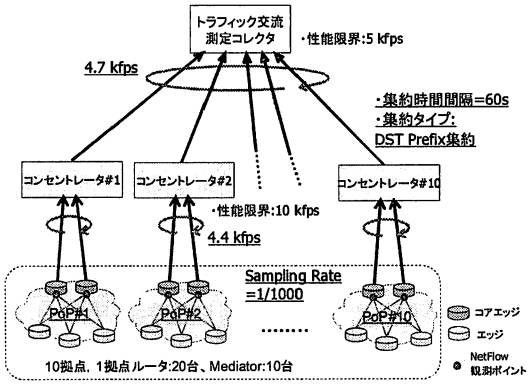


図 9. メディエータを用いたコレクタ構成の適用例.

5. まとめ

本研究では、フロー・メディエータを活用することで、トラフィック監視システムの拡張性と柔軟性を向上させるための適用モデルを検討した、

更に、大規模ネットワーク向けの大容量化に対応するため、実際のバックボーンネットワークのトラフィック特性をもとに、フロー数を推定することで、フロー・メディエータの配置・集約方式の設計例を示した．その結果、フロー・メディエータを PoP 単位に配置し、集約した情報をトラフィック交流監視コレクタに配信することで、大規模ネットワークにおける全トラフィック交流の監視を可能とするコレクタのアーキテクチャの実現可能性を示すことができた.

6. 謝辞

本稿は，総務省委託研究「次世代バックボーンに関する研究開発」による成果である.

文　献

[1] P. Phaal et al, “InMon Corporation's sFlow: A Method for Monitoring Traffic in Switched and Routed Networks”, RFC3176

[2] B. Claise, "Cisco Systems NetFlow Services Export Version 9", RFC3954

[3] B. Claise, " Specification of the IPFIX Protocol for the Exchange of IP Traffic Flow Information", draft-ietf-ipfix-protocol-24.txt(work in progress)

[4] 総務省総合通信基盤局電気通信事業部データ通信課，”我が国のインターネットにおけるトラフィック総量の把握，”2006 年 5 月 http://www.soumu.go.jp/s-news/2006/pdf/060731_1_1.pdf

[5] 小林淳史，松原 大典 他，”次世代バックボーン向け大容量トラフィック・コレクタの提案”, 2006 年 第 38 回 TM 研究会 2006 年.

[6] A. Kobayashi, K. Ishibashi, K. Yamamoto and D. Matsubara, "The reference model of IPFIX concentrators", draft-kobayashi-ipfix-concentrator-model-01.txt (work in progress)

[7] Daisuke Matsubara et al, "Development of Aggregation Device for Next Generation Backbone Networks", 2006 年 第 22 回 QAI 研究会 2006 年.

[8] 近藤毅 他,” トラフィック・フロー情報の効率的な配信手法の提案”, 2007 年電子情報通信学会総合大会.

[9] A. Kobayashi, T. Kondoh, K. Ishibashi and D. Matsubara, "IPFIX mediators", draft-kobayashi-ipfix-mediators-01.txt (work in progress)

[10] 小林淳史 他,” 次世代ネットワーク向けトラフィック交流監視コレクタの提案”, 2006 年 第 40 回 TM 研究会 2006 年.

[11] MAWI Working Group Traffic Archive http://tracer.csl.sony.co.jp/mawi/