Vol.2009-CSEC-46 No.34
2009/7/3

P2P ファイル交換ソフトウェアを対象としたトラヒック解析・制御システムの基本アーキテクチャの提案

大河内 一弥† 川口 信隆†

P2P（Peer to Peer）ファイル交換ソフトウェアの利用が広がり，P2P ソフトウェアの発生させるトラヒックによるネットワーク帯域の逼迫や，P2P ネットワークを介した個人情報の流出などの問題が顕在化する中，P2P ソフトウェアのトラヒック制御は重要な課題となっている．この報告ではネットワークを流れる大規模なトラヒックを観測し，P2P ソフトウェアが発生させるトラヒックを識別し，適切に制御するためのアーキテクチャを提案する．このアーキテクチャは，観測部，解析部，管理部，制御部の 4 つの機能部からなり，観測部で観測されたトラヒックを適切に解析部に割り振り，管理部は解析部から受け取った解析結果に応じて制御コマンドを制御部に対して発行する．最後に，本システムを用いて実際に P2P ソフトウェアの制御を行った実験について報告する．

Proposal of Traffic Analysis & Control System for P2P File Exchange Software

Kazuya Okochi† and Nobutaka Kawaguchi†

Many people uses P2P(Peer to Peer) file exchange software these days and problems that is occurred by using the software became serious. Therefore, analyzing and controlling P2P traffic is now an important problem. In this report, we propose a basic architecture for observing a mass traffic on a network, identifying P2P traffic, and controlling them. This system contains 4 function blocks; those are an observing block, an analyzing block, a managing block and a control block. The managing block allocate appropriate traffic at the observation block to more than one analysis modules in the analyzing block and get result from the modules and send traffic control command to the control block. Lastly we show a experimental result to show this system work appropriately.

1. はじめに

P2P（Peer to Peer）ファイル交換ソフトウェア（以下、単に P2P ソフトウェアと呼ぶ）の利用が広がる中，P2P ソフトウェアが発生させるトラヒックによる帯域の圧迫や，P2P ネットワークを介した個人情報の流出などが深刻な問題となっている（これらの問題に関する報道としては、例えば[1][2]など）．

このような状況において，P2P ソフトウェアのトラヒックの制御は，上記に挙げた問題の解決のみならず，P2P ファイル交換技術の適切な利用を促進するためにも，非常に重要な課題である．

本報告の目的は，ネットワークを流れる大規模なトラヒックの観測を行い，P2P ソフトウェアが発生させるトラヒックを識別し，適切に制御するための基本的なアーキテクチャを提案することである．

本稿で示す P2P ソフトウェアの制御システムは，観測部，解析部，管理部，制御部の 4 つの機能部からなる．本稿ではこのシステムの概要について述べた後，各機能部の機能について詳細に説明する．その後，本システムを用いて実際に P2P ソフトウェアの制御を行った実験について報告する．

既存の技術としては，P2P ソフトウェアの検知・制御を行う[3][4]などの製品があげられるが，P2P ソフトウェア毎に対応が必要なものであり，本報告のように汎用的に P2P 通信を検知することはできない．

2. トラヒック解析・制御システムの概要

第 2 節では本稿で述べるトラヒック解析・制御システムの仕様に関しての検討内容と，システムの全体概要について述べる．

2.1 仕様検討

(1) 汎用性に関する検討

本提案におけるシステムでは特定の P2P ソフトウェアに限らず，汎用的に P2P 通信を検知可能なシステムの開発を目標とする．

このため，P2P 通信を検知するための解析モジュール（以下、検知モジュールと呼ぶ）は，各モジュールが対応する P2P ソフトウェアや，検知するソフトウェアの特徴を，幅広く網羅的に設定できるよう，複数のモジュールを設置し，1 つのフローをそれぞれの検知モジュールに流して解析することができるようなアーキテクチャを提案する．

本報告におけるシステムでは，様々な特性を持つ複数の検知モジュールの出力する結果を組み合わせることにより，総合的に汎用的かつ精度の高い P2P ソフトウェアの

† （株）日立製作所
Hitachi Ltd.

ⓒ2009 Information Processing Society of Japan

検知，ひいては検知されたトラヒックの正確な制御を目指す．

(2) 大容量トラヒック対応に関する検討

本提案におけるシステムでは，ネットワークの基幹部への設置にも耐えうる性能（例えば 10Gbps 以上）を達成することを目標とする．

このトラヒック解析・制御システムはネットワークの基幹部に設置するため，解析のために通常の通信に遅延を生じさせることは許されない状況にある場合があることを考慮する必要がある．

また，検知モジュールの特性によって，トラヒックの解析を行うモジュールの対応可能な回線速度は異なることが考えられる．

これらの状況に対応するため，本報告では，観測対象とするトラヒックから実際に解析を行うトラヒックのパケットのみを遅延なくミラーリングし，対象の検知モジュールに各々のモジュールが対応可能な量のトラヒックを流して，解析を行うアーキテクチャを提案する．

2.2 システムの概要

上記の検討を踏まえたトラヒック解析・制御システムの概要を以下に示す．

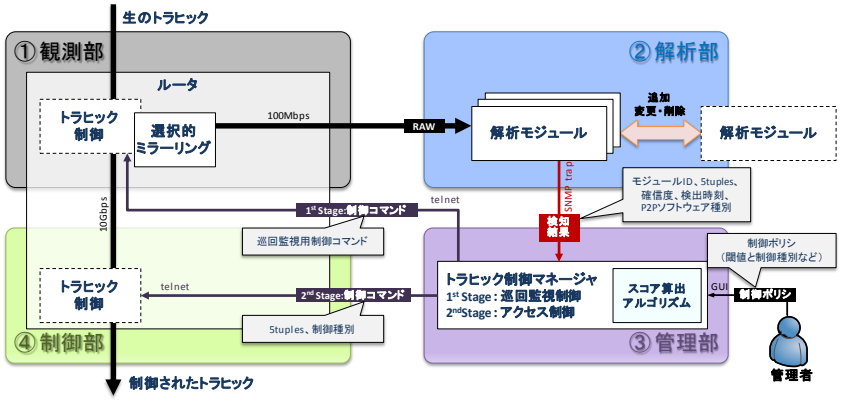

図 1：トラヒック解析・制御システムの概要

観測部から制御部にかけて流れているのが観測および制御の対象となるトラヒックである．トラヒック解析・制御システムは，観測部，解析部，制御部，管理部の大きく 4 つの機能部からなる．各機能部の概要は以下の通りである．

(1) 観測部

観測部は，観測対象となっているトラヒックから，解析対象のトラヒックをコピーする機能を持つ．この機能を実現する選択的ミラーリングのモジュールが含まれ，解析部に含まれる複数の検知モジュールに向けて個別のミラーリングを行う．

(2) 解析部

解析部は，観測部からミラーリングされたトラヒックを分析する検知モジュールが含まれる．解析結果は SNMP によって管理部に送信される．

(3) 管理部

管理部では，解析部に含まれる検知モジュールから解析結果の情報を受け取り，観測部から解析部に解析の対象となるデータを流すための制御情報の管理，および，その結果より制御部においてトラヒックの制御を行うための制御情報の管理を行う．

(4) 制御部

制御部では，管理部からの制御コマンドを受け，トラヒックの遮断・経路変更・帯域制限などの制御を行う．これらの制御をおこなう階層化シェイパのモジュールが含まれる．

以上の各機能部の詳細について第 3 章で述べる．

3. トラヒック解析・制御システムの詳細

本章では，トラヒック解析・制御システムの各機能の詳細を述べる．まず，第 3.1 節において，観測部および制御部の管理を行う管理部の機能を説明する（他の機能部を司る機能を有するため，最初に述べることとする）．第 3.2 節以降で選択的ミラーリング機能を含む観測部の機能，トラヒックの解析を行う解析部の機能，実際のトラヒック制御を行う制御部における機能について述べる．

3.1 管理部

本アーキテクチャにおいて，管理部は大きく 2 つの役割を果たす．1 つは，観測部において選択的ミラーリングモジュールに制御命令を送り，解析対象のトラヒックを各検知モジュールに流すことである．もう 1 つは制御部に対して制御命令を送り，トラヒックを適切に制御することである．

本アーキテクチャにおいては前者を第 1 ステージマネージャ，後者を第 2 ステージマネージャと呼称することとして，以下の節でそれぞれの詳細を述べる．

(1) 第 1 ステージマネージャ

第 1 ステージマネージャの概要を図 2 に示す．第 1 ステージマネージャでは，観測部におけるフロー情報生成モジュールよりフロー情報を取得し，各検知モジュールの解析可能なフロー数，トラヒック量に応じて，各検知モジュールへのトラヒックの割り振りを動的に決定し，その決定に従って選択的ミラーリングモジュールに対してミラーリングのコマンドを発行する．

ここでフローとはある IP アドレスを送信元，あるいは宛先としてもつパケットの集

ⓒ2009 Information Processing Society of Japan

合である．上記のフロー情報は，ある時点でトラヒックを流れているフロー（IP アドレス）の情報を保持している。

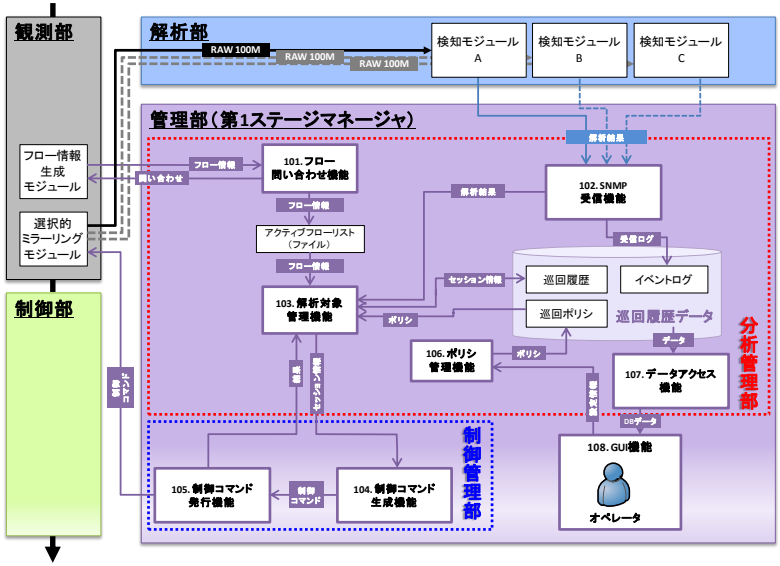

図 2：第 1 ステージマネージャの機能構成図

第 1 ステージマネージャの主要な処理を図 3 のフローチャートに示す．第 1 ステージの主要機能として，観測部よりフローリストを取得して，その時点で流れているフロー（IP アドレス）をリストアップする機能（図 2 中のモジュール 101），SNMP を受信して解析部において解析が終了したモジュールの情報を取得する機能（同モジュール 102），前記 2 つのモジュールの出力を受けて観測対象のトラヒックから解析対象のフローを適切に複数の検知モジュールに割り振る機能（同モジュール 103）が含まれる．解析対象管理機能（モジュール 103）では，各検知モジュール，観測されたフローに優先度を設定し，解析の終わっていないフローについて，優先度の高いフローから順に優先度の高い検知モジュールに割り当てる．割り当ての内容はコマンドデータとして選択的ミラーリングのモジュールに対して出力される．

フローの割り当ては，全てのフローが可能な限り全ての検知モジュールに割り当てられるように行われる．各フローの検知モジュールごとの割り当て状況（未割り当て，現在解析中，解析完了などの状況）は，マスタフローテーブルのデータベースに蓄積される．

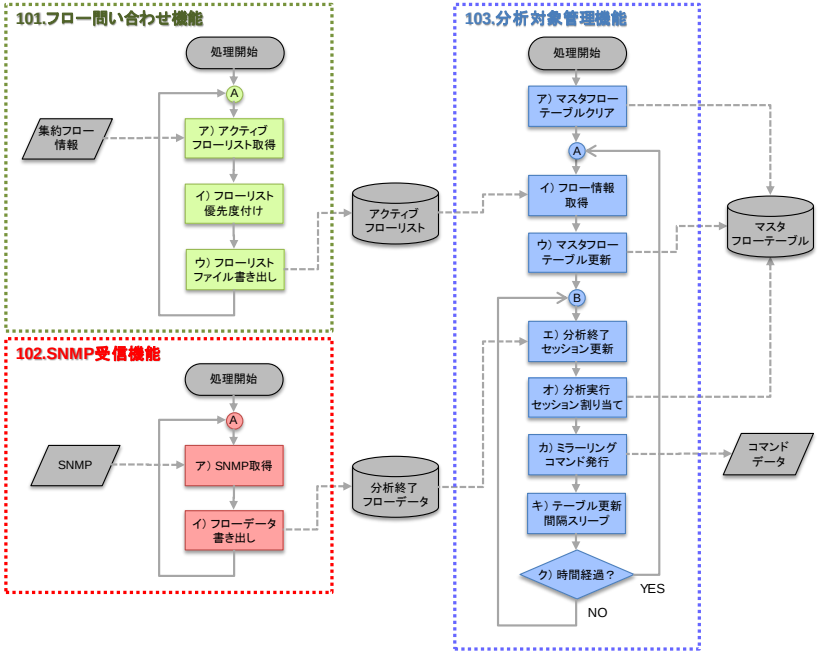

図 3：第 1 ステージマネージャの主要機能のフローチャート

(2) 第 2 ステージマネージャ

第 2 ステージマネージャの概要を図 4 に示す．第 2 ステージでは各フローについて，各検知モジュールが解析を行った検知結果を統合して，制御の判定を下す機能を備える．具体的には，1 つのフローについて検知結果が入力されるたびにそのフローの P2P ソフトウェアらしさを示すスコアを更新し（スコアの更新には，各検知モジュールのスコアの重み付き平均などの評価関数を用いる），統合されたスコアが一定値を超えた場合に制御のコマンドを発行する．

第 2 ステージマネージャの主要な処理を図 5 のフローチャートに示す．フローチャートの右部では，出力されるスコアを一定時間間隔ごとにポーリングして，最新の統合スコアを算出し，発行すべき制御コマンドを決定している．

制御コマンドは制御管理部を介してトラヒック制御モジュールに含まれる階層化シェイパに送信され，帯域制御や遮断，経路変更などの制御が実行される．

ⓒ2009 Information Processing Society of Japan

Vol.2009-CSEC-46 No.34
2009/7/3

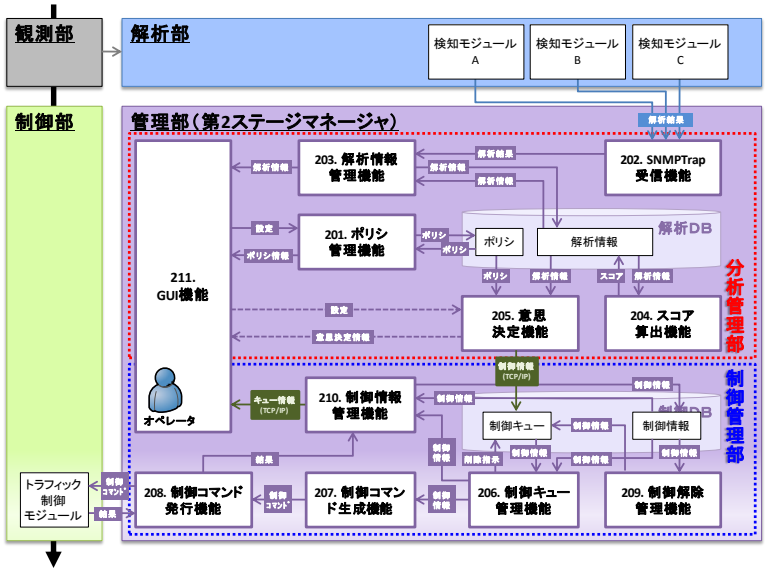

図 4：第 2 ステージマネージャの機能構成図

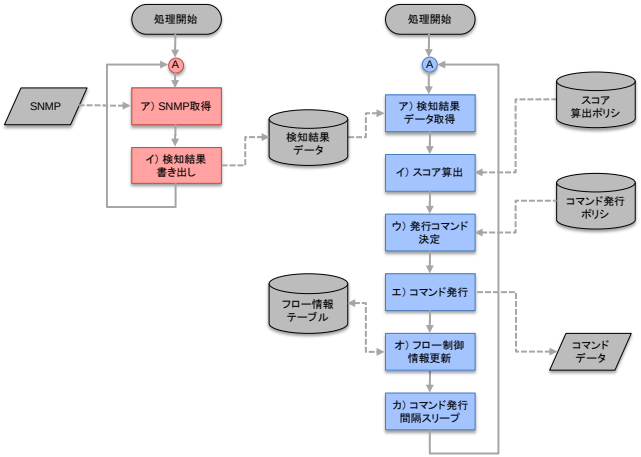

図 5：第 2 ステージマネージャの主要機能のフローチャート

3.2 観測部

観測部はネットワークの基幹トラヒックに影響を及ぼさないようにトラヒックを監視し，必要なトラヒックを解析部にミラーリングするための機能部である．このミラーリングの指示は管理部によってなされる．

観測部は 10 ギガビットイーサネットクラスのトラヒックに対応し，パケットのヘッダ情報をチェックして IP アドレス（フロー）ごとに異なった物理ポート（検知モジュール）へミラーリングを行うため，選択的ミラーリングモジュールの開発を独自に行った．

3.3 解析部

解析部は，異なる P2P ソフトウェアに対応する複数の検知モジュールが含まれる（検知手法自体については本報告の範囲外であるため割愛する．例えば[5]などを参照のこと）．検知モジュールは，検知のために用いる解析の手法によって大きく 2 つの種類に分けられる．1 つは，トラヒックのペイロードの内容を解析し，含まれる文字列などより P2P ソフトウェアを特定するプロトコル解析，もう一つは，主にトラヒックのヘッダ情報から得られる挙動情報を用いて解析を行うコネクション解析である．

プロトコル解析は P2P ファイル交換ソフトに特融の特徴をとらえることができれば高精度の検出が可能であるが，既知のソフトウェアにしか対応できないなどの欠点がある．

コネクション解析は，解析手法がサポートする特徴を持つソフトウェアを，未知の P2P ソフトウェアを含めて汎用的に検知できるという特長がある一方で，誤検知が起こる可能性があるという欠点がある．

それぞれの解析手法は統一された入出力インタフェースを持つ．具体的には，ミラーリングされた生のパケットを入力とし，結果を SNMP として出力する．このため，モジュールを追加・削除することが容易であり，観測対象のトラヒックの状況や新種の P2P ソフトウェアに対して柔軟に対応することが可能である．

これらの解析モジュールを組み合わせることにより，P2P ソフトウェアを幅広く網羅的にカバーし，かつ高精度の検知を行うシステムを構築する．

3.4 制御部

制御部は，トラヒックに対して実際に制御を行う機能部である．ここで行われる制御の種別は帯域制限，経路変更，遮断などであり，管理部によって指示がなされる．この制御は，制御モジュールにはフローを割り当てるためのバーチャルなキューを用意し，管理部か制御対象となるフローとそのフローを割り当てるキューの組み合わせを与えることで実現する．制御モジュールは，各キューに割り当てられたフローに対してあらかじめ指定された種別の制御を行う．

制御部は，上記のポリシに応じたトラヒックの制御を行うハードウェアとして，階層化シェイパの開発を行った．

ⓒ2009 Information Processing Society of Japan

4. 評価実験

本章では，制御マネージャ，選択的ミラーリングモジュール，階層化シェイパ，および検知モジュールを相互に接続し，P2P ソフトウェアのトラヒックに対して有効な制御されていることを確認する実験を行う．

4.1 実験手順

実験にあたっては，50 台規模の実験環境を用意し， P2P ソフトウェアの通信によってそれ以外のトラヒックが逼迫されている状況を人為的に作り出し，通常トラヒック（ここでは，P2P ソフトウェア以外のトラヒックを通常トラヒックと呼ぶこととする）が本システムによる制御（ここでは帯域制限を行う）を行った場合に回復することを確認する．実験手順を以下に示す．

1. 45 台の P2P ソフトウェアがインストールされた端末，5 台の P2P ソフトウェア以外のネットワークアプリケーションがインストールされた端末を用意する．ここでは FTP 通信を行うためのサーバとクライアントを用意した．
2. これら計 50 台の端末，および制御マネージャを含む実験用のネットワークを構成し，それぞれ P2P ソフトウェアを用いたファイルのダウンロード，FTP を用いたファイルのダウンロードをそれぞれ行う．
3. それぞれソフトウェアの動作が安定したら，FTP の通信が P2P ソフトウェアによって逼迫していることを確認し，その後制御マネージャを実行して，ミラーリングモジュールを用いた検知モジュールへのトラヒックのコピー，および検知モジュールを用いた P2P ソフトウェアの検知，検知結果を受けたトラヒックの制御（ここでは帯域制限）をそれぞれ行う．
4. 上記の処理によって，P2P ソフトウェアの帯域が制限され，逼迫していた FTP の通信が回復することを確認する．

4.2 実験結果

実験結果を以下に示す．なお，帯域制限においては，検知された P2P ソフトウェアの通信を，上限 256Kbps（約 30Kbyte/s）に制限した．

評価の期間を通じた全体のトラヒック量の推移を図 6 のグラフに示す．この評価において，グラフの X 軸は経過時間（単位は分）を，Y 軸はトラヒック量（単位は Kbytes/s）を表している．トラヒック量はやや上下するものの，概ね 3Mbyte/s（約 25Mbps）程度を推移している．

次に，FTP の通信の制御の前後における変化を図 7 に示す．なお，経過時間 20 分過ぎに制御モジュールを実行し，トラヒックの解析～制御を行った．また経過時間 40 分の手前でトラヒックの制御を終了した．

FTP クライアントの動作として，ファイルのダウンロードを開始した 8 分程度の時間からは概ね 50Kbyte/s 程度の速度でダウンロードを行っている．

ここで，制御モジュールを実行し，P2P ソフトウェアの検知～制御が開始されると，P2P ソフトウェアが検知されるごとにそのトラヒックが抑制され，代わりに FTP の通信量が増加していることがわかる．P2P ソフトウェアの検知にはある程度時間がかかるため，FTP のトラヒックは緩やかに増加し，最大では 400Kbyte/s 程度のトラヒック量を記録している．つまり，制御開始前の 10 倍程度にまでトラヒック量が向上したことを示す．

またトラヒックの制御を終了すると，P2P ソフトウェアの通信が再び活発になるため，FTP のトラヒック量は急速に低下し，再び 50Kbyte/s 以下のトラヒック量に下落している．

P2P ソフトウェアの通信の通信量の変化を図 8 のグラフに示す．図は 45 台のネットワークに参加している LimeWire のトラヒック量を示すものである．25 分程度に検知が行われてからは，最大 256Kbps（約 30Kbyte/s）を超えない速度で通信を行っており，制御部におけるシェイパが有効に機能していることがわかる．なお，検知の直前にトラヒック量が若干落ちてきているのは，P2P ソフトウェアの通信の不安定さによるものであり，制御が働いているわけではないことに注意する．

また，制御において，適切な検知対象トラヒックの巡回，制御コマンドの発行が行われていることを，制御マネージャのログから確認した．

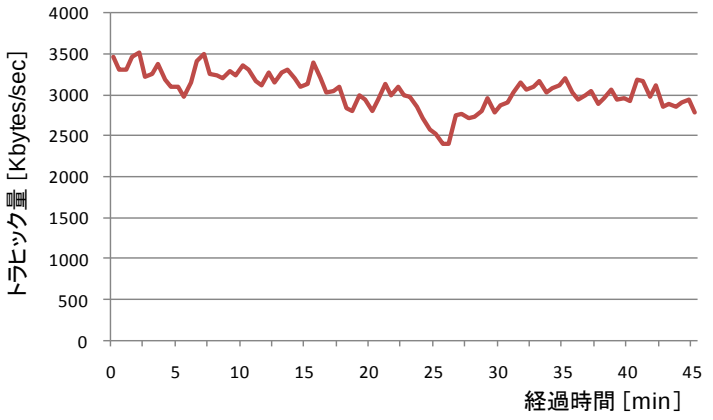

図 6：実験トラヒックの全体量の推移

ⓒ2009 Information Processing Society of Japan

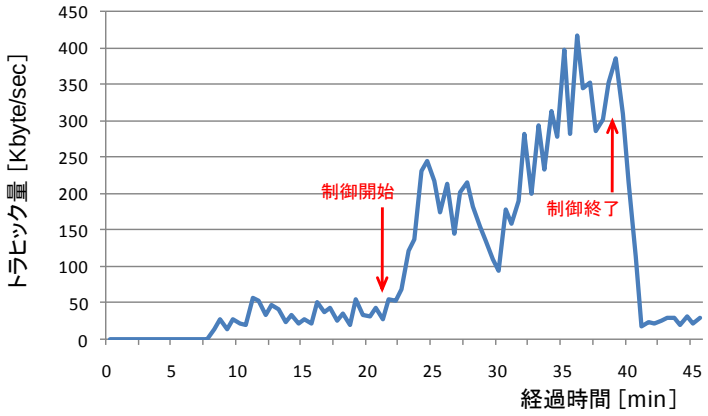

図 7：FTP トラヒックの推移

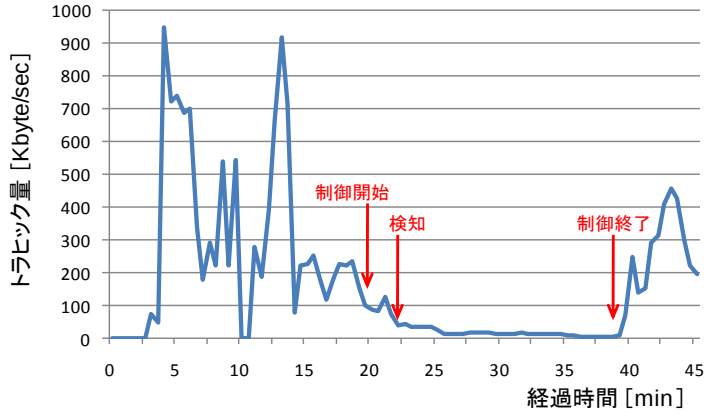

図 8：P2P トラヒックの推移

5. おわりに

本稿では，ネットワークにおいて P2P ソフトウェアによるトラヒックを検出し，制御するための基本的なアーキテクチャを示した．具体的には，大容量トラヒックに対応するためのトラヒック割り当て方式，制御方式などを検討し，これに対応するミラーリングおよびシェイピングのハードウェアを開発した．また，これらのハードウェアを管理するマネージャソフトウェアモジュールを開発した．また以上の開発モジュールについて，実験によってその有効性を確認した．

今後の課題としては，より大規模なトラヒックに対応するための各機能部の機能拡張などが挙げられる．選択的ミラーリング，階層化シェイパのハードウェアにおける制御方式の改善，またそれに対応した高機能の管理部ソフトウェアの開発などが課題である．

謝辞 本研究は総務省から委託を受けた「ネットワークを通じた情報流出の検知及び漏出情報の自動流通停止のための技術開発」の一部として実施しているものです．本研究を進めるにあたって有益な助言と協力を頂いた関係者各位に深く感謝致します．

参考文献

[1] 社団法人日本インターネットプロバイダー協会「『帯域制御の運用基準に関するガイドライン（案）』に係る意見募集について」：
http://www.jaipa.or.jp/other/bandwidth/info_080317.html

[2] 官房長官記者発表「Winny を介した情報漏えいについて」：
http://www.kantei.go.jp/jp/tyoukanpress/rireki/2006/03/15_a.html

[3] One Point Wall：http://www.onepointwall.jp/

[4] NetScreen：http://juniper.co.jp/products_and_services/firewall_slash_ipsec_vpn/

[5] 重本倫宏・大河内一弥・寺田真敏：コネクション解析による P2P 通信端末検知手法，情報処理学会第 42 回 CSEC 研究会，2008

ⓒ2009 Information Processing Society of Japan