自動シグネチャ生成システムにおけるトラヒック情報収集方式に関する検討

東島　慶[†]　　中村　豊[††]　　池永　全志[†††]　　飯田　勝吉[††††]

九州工業大学　工学部電気工学科[†]　　九州工業大学　情報科学センター[††]　　九州工業大学大学院工学研究科ネットワーク工学講座[†††]　　東京工業大学　学術国際情報センター[††††]

1. はじめに

家電製品は携帯電話の IT 化により，これまでコンピュータのみで行われていたインターネット通信が，家電製品や携帯電話のような一般家電で実現できるようになってきている．このような一般家電がインターネット接続可能となると，ノード数が爆発的に増加することが予想できる．IPv6 の普及により，家電製品や携帯電話１つ１つに IP アドレスを割り振ることが可能となる．一方，これまでのサーバ・クライアントモデル通信だけではなく，様々な P2P 型のアプリケーションが登場しており，ネットワークトラヒック複雑化，多様化している．そして，e-コマースと呼ばれる電子商取引が社会基盤化していることで，インターネットの重要性は益々増加している．

このような，ノード数の爆発的増加やアプリケーションの多様化，社会基盤化に伴い，インターネット環境は大規模化・分散化している．このような大規模分散環境において，ネットワークを維持，管理するには，ネットワーク内で自動的に異常トラヒックのシグネチャを生成する必要があると考える．

そこで本研究では，自動的に異常トラヒックのシグネチャを生成するためのトラヒック情報を収集する方式について検討する．

2. 関連研究

トラヒック情報を共有する手法として，SNMP[1]ベースのシステムが数多く存在する．主にルータやスイッチが SNMP をサポートし，L2 情報をクライアントへ送信する．ルータやスイッチの情報を収集したクライアントは視覚化し管理者に情報を提供する．同様に，NetFlow[2]や sFlow[3]を用いたトラヒック情報収集システムが存在する．これらの手法では，トラヒックデータを収集するサーバが一極集中管理することを前提としているため，情報爆発時代の大規模分散ネットワークに適用することは困難である．一方，P2P フレームワークを分散環境におけるトラヒック計測フレームワークとしている研究も存在する[4][5]．iPlane は経路発見方法であり，数箇所の観測点を持って最適経路を推定する情報基盤を提供している．iPlane は信頼性の無いユーザーからの計測結果の寄贈を許すため，情報汚染目的の攻撃に弱くなるという問題点がある

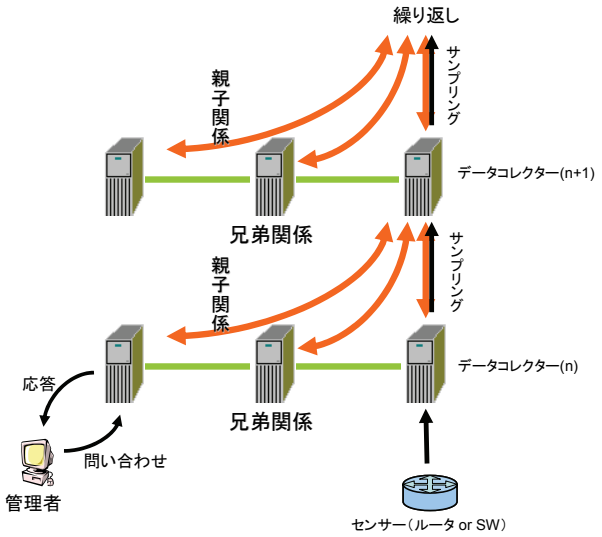


図1.提案アーキテクチャ

3. 提案アーキテクチャ

これまでの研究では，一極集中管理であるためノード数の増加に対応できない場合や，ノード数の増加に伴い，オーバレイネットワーク維持が困難になっている．そこで本研究では，大規模分散環境においても，円滑にトラヒック情報を収集するために，データベースサーバを階層的に配置し，収集したトラヒックデータをサンプリングして縮退させることにより，大規模分散環境に適応できるトラヒック収集方式を提案する．本提案方式の概要を図１に示す．

我々の提案するトラヒック情報収集方式では以下の構成要素が存在する．

・センサー

全ての通信ノードおよび中継ノード．全てのセンサーがデータベースサーバへトラヒック情報を送信する．センサーでは，ミスユース型の異常検知システムが動作しており，明確な異常ト

Traffic Statistics Collecting Method for Automatic Creation of Anomaly Signature

† †† ††† Kyushu Institute of Technology

†††† Tokyo Institute of Technology

ラヒックを判断することができる.

・データベースサーバ
センサーが送信してきたトラヒック情報を蓄積するサーバ．データベースサーバ間で連携してトラヒック情報を管理する.

・管理者
データベースサーバへトラヒック情報の問い合わせを行う.

4. シミュレーション

管理者が必要なネットワーク情報をデータコレクタへ問い合わせ，必要なトラヒック情報を取得するまでの応答時間をシミュレーションにより計測する.

問い合わせを行うデータコレクタの階層を横軸にとり，縦軸を応答時間とした．シミュレーションのシナリオは，以下の通りである．図２にシミュレーションネットワークのトポロジを示す.

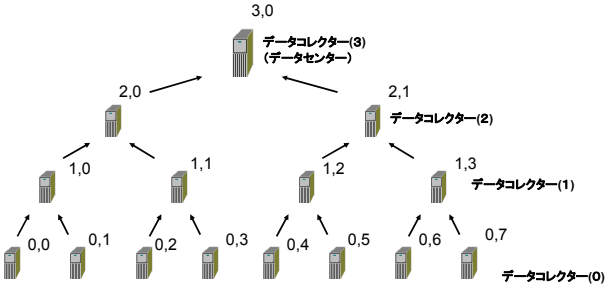


図2. トポロジ図

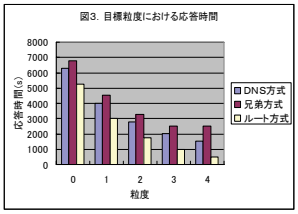
それぞれのデータコレクタには，1TB の容量のディスク容量が存在する．データコレクタ間の帯域は 1Gbps とする.

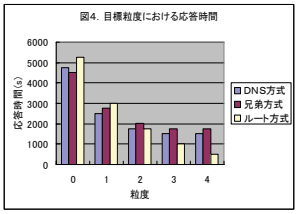
検索の方式として３通りの検索手法について比較検討する.

・兄弟方式
要求する同一ドメインにいるデータコレクタを兄弟と定義する．管理者からの問い合わせは，最初に兄弟へ転送され，発見されない場合は１つ上位のデータコレクタに問い合わせを行う検索方式．上位のデータコレクタにも存在しない場合，上位のデータコレクタの兄弟へ問い合わせを行う．以下，要求するトラヒック情報が発見されるまで繰り返す.

・DNS 方式
要求するトラヒック情報がデータコレクタに存在しない場合，１つ上位のデータコレクタに問い合わせを行う検索方式．この方式の場合，全ての問い合わせがルートに集中するため，大規模分散環境における適応性は低い.

・ルート方式
管理者の問い合わせは，最初にデータコレクタの最上位（ルート）へ行う方式．管理者の要求するデータ階層までデータコレクタを１階層ずつ降りて目的のトラヒック情報を発見する方式.





５．結果

図４に管理者と目的とするトラヒック情報が同一ドメインの場合の結果を示す．図４に示すように，同一ドメインの場合は兄弟方式が最も応答時間が短い.

図３に管理者と目的とするトラヒック情報が異なるドメインの場合の結果を示す．図３に示すように，ドメインが異なる場合はルート方式が最も応答時間が短い.

６．まとめ

シミュレーションの結果より，階層の低い検索で，かつ，同一ドメイン内のトラヒック情報を検索する場合は，兄弟方式が最も優れていることが明らかとなった．また，異なるドメインのトラヒック情報を収集する場合は，ルート方式が最も優れていることが明らかとなった．今後は，兄弟数や兄弟内の検索手法と階層の関係についての調査や，より効率の高い手法の検討を行っていく予定である.

謝辞

本研究は科研費特定領域研究「情報爆発時代に向けた新しい IT 基盤技術の研究」の援助を受けています.

参考文献

[1]“SNMP”, RFC1157, May, 1990

[2]NetFlow,http://www.cisco.com/warp/public/732/Tech/nmp/netflow/index.shtml

[3] sFlow, RFC3176, Sep, 2001

[4] Kenji M. and Youki K. N-TAP: A Platform of Large-Scale Distributed Measurement for Overlay Network Applications. In Proc. of the Second International Workshop on DAS-P2P 2007, January

[5] Harsha V. Madhyastha, Tomas I., Michael P., Colin D., Thomas A., Arvind K., Arun V., “iPlane: An Information Plane for Distributed Services”, 7th USENIX Symposium, pp. 367-380, Seattle