情報爆発に対応する高度にスケーラブルなモニタリングアーキテクチャ

研究代表者	中島　達夫	早稲田大学・理工学術院・教授
研究連携者	村岡　洋一	早稲田大学・理工学術院・教授
	後藤　滋樹	早稲田大学・理工学術院・教授
	山名　早人	早稲田大学・理工学術院・教授
	甲藤　二郎	早稲田大学・理工学術院・教授
	秋岡 明香	電気通信大学大学院・システム情報学研究科・助教

１．研究概要

大規模な分散システムを安定して動作させるためにはシステムが置かれた状況を理解することを可能とする必要がある．本研究では，そのためのインフラストラクチャとして様々な実時間に生成された情報を収集，分析することを可能とするためのモニタリングアーキテクチャに関する研究をおこなう．

２．モニタリングアーキテクチャ

今日、top, nmon, sar, iostat 等の性能モニタリングツールは Web アプリケーションの性能を解析するための十分なツールではない．それらのツールを用いてアプリケーションの大域的なリソースの使用量を知ることは可能だが，個々の HTTP リクエストの詳細なリソース使用量を調べ，その情報に基づいてシステムのハードウエアリソースを割当てたり，性能に関するバグを調査することは困難である．例えば，１つの CGI アプリケーションが多くの異なる HTTP リクエストを処理する場合，現状のモニタリングツールは ″login″ と ″addtocart″等の異なるリクエストを区別することが出来ないので，詳細な性能分析が出来ない．また，まれなケースであるが，各リクエストが異なる CGI アプリケーションにより処理される場合は，これらのリクエストを同様の物として扱うことが不可能となってしまう．さらに，最近の Web サーバは PHP や Perl 等のインタープリターを内蔵し，それらを独立したスレッドにより実行するため，現状の荒い粒度のリソースしかモニタリングできないツールを用いて性能解析することは困難である．同様に，それぞれのクライアントのコネクションを異なるスレッドで処理する SQL サーバの性能解析も極めて困難である．

本研究では，以上の問題点を解決するため，オペレーティングシステムとツール間のギャップを検討し，HTTP リクエスト毎の詳細なリソースの管理を可能とする．本研究により開発するツールは以下の２つの部分から構成されている．

１）スレッドベースの詳細なリソースモニタリングを可能とする perfmon2 と libprf ライブラリを使用して各 CPU が提供するパフォーマンスカウンタをアクセスすることを可能とする．これらのツールを利用することにより，Web サーバに埋め込まれたインタープリタやデータベースサーバのコネクションハンドラの詳細なリソース使用量の分析が可能となる．また，Apache 内の HTTP リクエストが CGI を実行するためにプロセスを fork する場合は，ptrace システムコールを用いて子プロセスの終了時に必要な情報を抽出する．

２）既存のミドルウエアの拡張をおこなった．本研究では，Apache2 と MySQL を使用することを前提としている．これらのソフトウエアはオープンソースとして開発され，Web アプリケーションを実行するために広

く利用されている．本研究により開発するツールを利用可能とするため，個別の HTTP リクエストのリソース使用量をモニタリングしてログとして書き出せるように Apache2 を改良した．また，MySQL に関しては，各 HTTP リクエスト毎のリソース使用量を取得可能とするため，各 SQL キュエリーをモニタリング可能とした．

本年度は，提案するツールのプロトタイプの実装と性能解析のためのビジュアルインタフェースを構築した．今後は，実用的な Web アプリケーションを利用して開発したツールの有効性を検証する予定である．

３．分散システム情報収集分析・活用ミドルウェアの研究

本研究の主目的は、収集したログ情報の分析手法と、ログ分析結果の活用方法を検討することにある。ログ分析手法としては、データマイニングの一手法であるシーケンシャルパターンマイニングを用いる。本年度はまず、シーケンシャルパターンマイニングのログ情報に対する適用が有効であることを示すために、単一マシン上においてファイルアクセスログからファイルアクセスパターンを抽出し、システム性能向上のために活用できることを確かめた。そして、既存のシーケンシャルパターンマイニング手法を分散環境におけるログに適用できるよう拡張を行った。以下、順に説明する。

シーケンシャルパターンマイニングのファイルアクセスログへの適用

情報爆発に伴い、ストレージのボトルネックがますます問題となってきている。そこで、OS が空きメモリ上に確保するファイルキャッシュの効率を上げるために、ファイルアクセスログを活用することを目指した。OS 内からは、各アプリケーションが、どのファイルのどの領域にアクセスしたかというファイルアクセスログを取得することができる。ファイルに対するアクセスには、たとえば、「ある HTML ファイル A がアクセスされると、A からリンクされている画像 B がアクセスされる」といった、一定のアクセスパターンがあることが普通である。このようなアクセスパターンを抽出するため、アクセスログに対してシーケンシャルパターンマイニングを適用した。

本年度は単一マシン上で研究を行った。提案手法を Linux に実装し、実機上で評価を行った。Linux カーネル内の専用スレッドがアクセスログからアクセスパターンを抽出し、抽出したアクセスパターンに基づいてキャッシュアルゴリズムを最適化する。実験の結果、データベースのベンチマークである DBT-2 の性能において 5%の性能向上を確認した。数%のアクセス性能向上パッチでも Linux カーネルに採用されることを考えると、これは決して小さな性能向上ではない。

シーケンシャルパターンマイニングの拡張

分散環境で収集されるログには、CPU やメモリの使用情報、前述のアクセスログなど多数の情報が含まれる。また、イベント発生源のノードも区別しなければならない。さらに、イベントが発生した時間も重要である。つまり分散環境でのログは、イベント内容・イベント発生源・イベント発生時間の 3 つの属性からなる。ここでイベント発生時間は特に重要である。例えばあるイベント A と B が狭い時間内に同時に発生する場合、A と B は何らかの関係があるといえる。しかし、A と B の時間間隔が離れていると、A と B の関係は弱いと考えられる。通常のシーケンシャルパターンマイニングアルゴリズムは、イベントの発生間隔を考慮

することができず、分散環境のログに適用するには不適切である。

そこで、我々が以前に開発した、時間間隔を考慮可能なシーケンシャルパターンマイニングアルゴリズムを拡張し、3 つの属性からなる分散環境のログのシーケンシャルパターンマイニングを可能にした。時間間隔を考慮したシーケンシャルパターンを効率よく抽出するために新しいアルゴリズムを開発した。

次年度への展望

本年度は、単一ノード上でのファイルアクセスログ分析・活用と、分散環境でのログ解析手法の開発を行った。次年度は分散環境において実機による評価を行うことが第一目標である。分散環境で得られたログから有益な情報を抽出し、各ノードにフィードバックすることで各ノードの性能を向上させることを目指す。具体的には、

(1) アクセスパターンを利用した分散ファイルシステムの性能向上

(2) アイドルメモリやアイドル CPU を活用した性能向上

といったことを検討する。(1) は、本年度のアクセスパターンの研究を拡張することで達成することを目指す。(2) は、低負荷のノードを高負荷のノードの補助リソースとして活用することを目指すものである。各ノードの負荷の変化を予測するために、ログ情報が使えると考えている。

４．大規模分散セキュリティアナリストシステムの構築

研究概要

コンピュータやネットワークが人々の生活に欠かす事ができないインフラストラクチャとなった現在，コンピュータシステムを常時安全に安定稼働させることが必須である．しかし，ネットワークやサーバへの攻撃は日々進化を続け，サーバ管理者は自らが管理するシステム群のログ管理や新たな攻撃への対応に追われている．

本研究では，このようなサーバ管理者の負担を軽減するために，サーバのログを収集・解析することで攻撃の検知を自動で行ない，さらにこうした攻撃情報を各地から集約することで，攻撃のトレンドや大規模攻撃の予測，新たな攻撃手法の発見といった知見を洗い出して共有する，大規模分散セキュリティアナリストシステムを構築する．

セキュリティアナリストシステムの概要

本年度は，大規模分散セキュリティアナリストシステム実現へ向けた一歩目として，プロトタイプシステムの構築を行ない，実証実験として SSH への辞書攻撃の検知と攻撃ルートの検出を行なった．プロトタイプシステムは，以下の 3 つのモジュールで構成される．

- フィルタリングモジュール

  システムログから，通常のユーザログインや定常状態のログを排除し，SSH への辞書攻撃に該当すると見られるログのみを取り出し，知識集約モジュールへ送る．
- 知識集約モジュール

各サーバから収集された SSH への辞書攻撃に関連すると見られるログを集約し，攻撃ルートの特定を行なう．

- 可視化モジュール
  知識集約モジュールが特定した攻撃ルートを可視化し，利用者にわかりやすく提示する．

プロトタイプの実証実験環境として仮想ネットワークを用意し，このネットワーク内の 8 台の linux 仮想マシンにおいて，フィルタリングモジュールを稼働させた．この仮想ネットワーク内で SSH の辞書攻撃を次々と行なったところ，各仮想マシンで SSH の辞書攻撃を受けたことを検知し，知識集約モジュールで攻撃者が辿ったルートを特定することに成功した．また，利用者が通常のログインを行なう際にパスワードを誤入力したケースを，攻撃と誤判定しないことも確認した．

まとめと次年度への展望

本年度は，大規模なセキュリティアナリストシステムの構築へ向けてプロトタイプの実装を行ない，SSH への辞書攻撃の検知と攻撃ルートの検出を行なう検証実験を行なった．次年度の課題として，以下の 2 つを予定している．

1）より汎用性が高い手法による攻撃検知アルゴリズムの研究開発

さらに多くの既知の攻撃に対応し，今後新たに生じる未知の攻撃の検出も行なうためには，攻撃手法とログの一般的な関係性の洗い出しを行ない，特定のログメッセージに依存しない検出アルゴリズムの開発が重要である．

2）大規模な環境での実証実験

プロトタイプの検証実験は，仮想ネットワーク内での小規模なノード群で行なった．しかし，本研究の真価は，大規模にシステムログを収集し，収集したログを解析することで新たな知見を生み出し，さらにその知見を共有する形でフィードバックする点である．したがって，スケーラビリティの確保は不可欠であるため，大規模分散環境での実証実験を行ない，その結果を踏まえた改良を行なう予定である．

５．自己組織型セキュリティミドルウェア

本年度は、継続している自己組織型セキュリティミドルウェアに関する検討をさらに進め、ネットワーク構成技術の面からは、まずメッシュ型配信とツリー型配信を組み合わせたP2P型データ配信方式に関する検討を行ない、従来方式よりもオーバーヘッドの少ない配信性能を実現した。また、モバイルCDN環境における動的コンテンツの一貫性制御と更新方式に関する検討を行い、同じく従来方式に対して総トラヒック量を抑える方式を実現した。これらは共に自律的にオーバーレイネットワークを構成し、障害に対するロバスト性も兼ね備えている。また、無線通信におけるRSSIの変化に着目した居室内の混雑度推定システムに関する検討も行った。

またネットワーク測定技術の面から、P2P型のデータ通信を可能とするためのNAT越え(traversal)の研究を進めた。これは従来の方法では困難とされていたsymmetric型のNATを対象とするものである。この方法を適

用するとUDPだけではなくTCPにおいてもNAT越えが可能となる。さらに通信に参加するノードが異種のネットワークをスムーズに切り替えることを可能とするために、マルチキャストを用いる方法を提案した。通信の品質に関して研究を進めて、高品質の画像を簡単に通信できる方法を提案した。またサーバの負荷を分散する従来の方式を改良して、実地のネットワークに適用して実証実験を行った。セキュリティの面からの研究を進めて、昨年度の研究成果を拡張し、Dark IPを用いてネットワークの挙動を監視する方法を広い範囲に適用した。

メッシュ型配信とツリー型配信を組み合わせたP2P型データ配信方式においては、オーバーヘッドの少ないツリー型配信の利点と、障害に強いメッシュ型配信の利点を組合せ、両者のハイブリッド構成から成るデータ配信方式の提案を行った。その結果、従来方式に対するスループット性能の向上とオーバーヘッドの削減効果を同時に実現し、かつ、ノードの離脱や回線障害にも耐性を有するデータ配信を実現した。本提案方式はいかなるP2P型配信にも適用可能であり、元々はマルチメディア情報の配信手段としての検討を進めていたが、ネットワーク上のコラボレーションツールや、基幹サーバシステムなどへの応用も検討を進めている。

モバイルCDN環境における動的コンテンツの一貫性制御と更新方式に関しては、時間的に更新される動的コンテンツの複数サーバ上のキャッシュ管理に関して、コンテンツへのアクセス頻度を考慮した問題の定式化を行い、アクセス頻度に応じて適応的にコンテンツの複製個数を制御したり、コンテンツの配信元のサーバを適応的に切り替えたりする方式の提案を行った。ここではまた、モバイル環境を示すパラメータとして、アルゴリズムに信号の到達範囲と信号強度とを組込み、モバイルCDNとしての実現性を探った。その結果、従来方式に対して、コンテンツのヒット率の改善や総トラヒック量の削減などの利点を実現した。また本提案は、Chinacom 2008 の Best Paper Awardを受賞した。

無線通信の受信電波強度RSSIの変化を用いた混雑度推定システムに関しては、無線通信方式として802.11bとZigBeeを取り上げ、実環境における無線通信特性の取得から、実際のアプリケーション試作までを行った。無線通信では、送受信端末間に人間が立っていたり障害物が存在すると、電波の吸収などによってRSSIが減衰する。そこで、事前に人間の人数を変化させながらRSSIの変化を計測し、RSSIの平均と分散に関する近似曲線を求めた。以後は、RSSIを計測しては近似曲線から逆算することで人間の人数（混雑度）の推定が可能になるが、これをさらに遠隔地からアクセスできるようにしたプロトタイプの試作を行った。本検討は、例えば無線センサーネットワークを本来の目的に使用しながら、構成を変更することなく、無線通信の特性を活かした別の応用も探ろうとするものである。

従来から幾つかのNAT 越えの方法が提案されている。本研究では、従来の方法で最も難しいとされているSymmetric NAT を越える方法を提案した。この新しい方法の特徴は、測定技術に基づきNATのポートマッピングを予測して、通信に使用できるポートを多数オープンすることである。このようにして、NAT を経由する通信が成功する確率を高める。この方法をUDP Multi Hole Punching と呼ぶ。本研究では新しい方法を実装してその特性を評価・考察した。あわせてTCP Hole Punching の新しい手法も提案した。

携帯端末を使うときには、移動中にも通信を継続したい。ネットワークを切替え
る時の通信の品質を保つために、従来から1台の端末に2つのIPアドレスを付与する方法が知られている。ただし単に2つのIPアドレスを使用するのでは、切替え時にネットワークの帯域を2倍使うことになる。本研究では、マルチキャストネットワークを活用することにより、ネットワークの切替えを円滑に行ない、帯域を能率よく使用する方法を提案した。さらに、新しい方法を実装して実証した。

ブロードバンド接続の普及にともなって、家庭でもインターネット上の映像配信が広く利用されている。ただし、HD画質の映像配信は配信設備が高価なことから、まだ普及するには至っていない。本研究ではオンデマンドコンテンツの配信において、FLASHを用いたフルHDの映像配信システムを構築する方法を提案する。ここで提案する方法の特徴は、市販のAVCHDカメラで撮影した映像をH.264(.mp4) にエンコードして、良く使われているFLASH配信ソフトをWEBサーバに組み込むことによって、プログレッシブなダウンロードを可能にしたことである。これは安価にHD配信システムを実現したことになる。ところで本システムを使用する際には、H.264の映像を再生するために視聴者側のPCに負荷が生じる。そこで本研究では視聴ページを開くときのパラメータを調整して、最適な視聴環境を整えると共に、各設定ごとのCPU使用率を測定した。

大容量のコンテンツを扱うサービスを、単体のサーバで実現するのは困難になる。従来から、サーバの負荷分散を実現するための幾つもの技術が提案されている。既存技術の中でも導入が容易なDNSラウンドロビンでは、サービスの種類によってサーバごとの接続数に偏りが出やすいという問題が知られている。本研究では、低コストで精度の高い負荷分散を実現する。具体的には、DNSラウンドロビンの対象となるDNSレコードを動的に更新する方法である。評価のために、実際のサーバのデータを用いて評価を行った。

セキュリティの観点からは、昨年度の研究成果であるDark IPによるネットワーク上の脅威検出の適用範囲を拡大して、局所的なネットワークにおいて、パケットの全数検査に基づく方法を考案した。昨年度までの研究は、広域のネットワークにおいて、パケットのサンプルに基づくフローデータを分析の対象としていた。このように、大学のような局所的なネットワークと、大学間ネットワークのような広域のネットワークの両方を適用範囲とすることにより、局所的なネットワークにおける精度評価と、広域的なネットワークにおける脅威検出の両方を実現することができた。

今後の展望は下記のとおりである。

P2P型データ配信に関しては、障害の発生を、ノードの離脱や回線障害によるものだけではなく、外部からの攻撃等、よりセキュリティ対策としての具体性の高い方式に拡張する。モバイルCDNも同様であり、攻撃や異常トラヒック存在時の有効性の検証を進める。無線通信システムに関しては、現状でも侵入検知などへの応用が可能だが、さらに映像情報を加えるなど、センサフュージョン的な拡張を進めていく。

ネットワーク測定に関しては、NATが多段に重なった場合のtraversal技術の確立に向けて研究を進める。さらにP2Pのトラフィックを有効に把握するための技術を検討する。ルータ等のネットワーク機器の測定を行い、その結果を用いてP2Pトラフィックを制御するのが目標である。セキュリティに関しては本研究において提案した方式を、従来方式と精緻に比較して利点を明らかにする予定である。

６．センサーを利用した応用サービス

センサー情報を利用するサービスとして Ambient Lifestyle Feedback System を提案した. 本システムは, ユーザの行動をユーザに意識させずにモニタリングして, その行動を改変するためにフィードバック情報をユーザに心理的負荷を課さないように提示する.

本年度は, 今まで開発した４つのケーススタディの検討をおこない、開発／使用経験をもとまとめた. その結果、ポジティブなインセンティブとネガティブなインセンティブを組み合わせて使用することが重要であることが明らかになった. また、短期的なゴールを明示することもユーザのモティベーションを高めるために有効であることが明らかになった. 今後は, 社会的なインセンティブや経済的なインセンティブを利用することでどのようにユーザの行動を改変できるかを検討する予定である.

７．研究成果リスト

著書、論文、国際会議

- Fahim Kawsar, Kaori Fujinami, and Tatsuo Nakajima; "Deploy Spontaneously: Supporting End-Users in Building and Enhancing a Smart Home"; The Tenth International Conference on Ubiquitous Computing (Ubicomp 2008), Seoul, South Korea, September 21 - 24, 2008
- Tatsuo Nakajima, Hiroaki Kimura, Tetsuo Yamabe, Vili Lehdonvirta, Chihiro Takayama, Miyuki Shiraishi, Yasuyuki Washio: Using Aesthetic and Empathetic Expressions to Motivate Desirable Lifestyle. Smart Sensing and Context, Third European Conference(EuroSSC 2008): 220-234, Zurich, Switzerland, October 29-31, 2008
- Yu Hirate, Hayato Yamana, “Profiling Node Conditions of Distributed System with Sequential Pattern Mining,” In *Proc. of the 1st Int’l Workshop on Software Technologies for Future Dependable Distributed Systems* (STFSSD 2009), Tokyo, Japan, Mar. 2009 (To Appear).
- Takanori Ueda, Yu Hirate, Hayato Yamana, “Improving Storage Performance by using Idle Memory and CPU on Distributed Systems,” In *Proc. of the 1st Int’l Workshop on Software Technologies for Future Dependable Distributed Systems* (STFSSD 2009), Tokyo, Japan, Mar. 2009 (To Appear).
- 上田高徳、平手勇宇、山名早人、“システムコールレベルのアクセスログを用いたディスクアクセスパターンマイニング、”日本データベース学会論文誌、Vol.7, No.1, pp.145-150 (2008.6).

- Sayaka Akioka, Hideo Fukumori, and Yoichi Muraoka, “Search in the Mood: the Information Filter based on Ambiguous Queries”, International Journal of Computer Applications in Technology (to appear).
- Sayaka Akioka, Hideo Fukumori, and Yoichi Muraoka, “Search in the Mood: the Information Filter based on Ambiguous Queries”, The 5th IEEE/ACM International Conference on Information Technology and Applications (ICITA2008), June 23-26, 2008.
- Junichi Ikeda, Sayaka Akioka, and Yoichi Muraoka, “An Alarm Announciator for Domain Administrators”, In Proc. of the 1st Int’l Workshop on Software Technologies for Future Dependable Distributed Systems (STFSSD 2009), Tokyo, Japan, Mar. 2009 (To Appear).
- 下田晃弘, 後藤滋樹, “フローデータからのDark IP抽出による脅威観測法”, 電子情報通信学会論文誌, Vol.J92-B, No.1, 採録決定, 印刷中, January, 2009
- M.Nakatsuka, S.Iwatani and J.Katto, “A Study on Passive Crowd Density Estimation using Wireless Sensors”, ICMU 2008, Jun.2008.
- S.Zhou, J.Katto, Y.Yasuda and Y.Chen, “Consistency and Update in Mobile Overlay Networks”, Chinacom 2008, Aug.2008.
- S.Awinphan, S.Zhou and J.Katto, “Mesh-based Data Delivery over Multiple Tree-Shaped Routes in P2P Overlay Network”, ICOIN 2009, Jan.2009.

招待講演

- Tatsuo Nakajima, Reflecting Human Behavior to Motivate Desirable Lifestyle, The 3rd Rutgers-Helsinki Workshop on Spontaneous Networking, 2008.5
- 上田高徳 : “メニーコア時代における OS レベルでの I/O 最適化、” 情報研報 (DBS) / jDB ワークショップ、Vol. 2008, No.56, pp.133 (2008.6).

受賞

[1] 上田高徳 : 日本データベース学会 / 電子情報通信学会データ工学研究会 / 情報処理学会データベースシステム研究会 最優秀若手研究者賞

[2] 上田高徳 : 学生発表奨励賞、iDB フォーラム 2008 (2008.9)