様式 C-19

科学研究費補助金研究成果報告書

平成 23 年 1 月 22 日現在

機関番号：32689
研究種目：特定領域研究
研究期間：2006-2010
課題番号：18049068
研究課題名（和文）
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
研究課題名（英文）
Highly Scalable Monitoring Architecture for Information Explosion Environments
研究代表者
中島　達夫(NAKAJIMA TATSUO)
早稲田大学・理工学術院・教授
研究者番号： 10251977

研究成果の概要（和文）：本研究において目的とするモニタリングアーキテクチャは， 情報基盤， 社会基盤，人々の日常生活を守るためのソフトウエアの一群である．通常は独立に研究されていた研究分野を統合することにより，従来解決が困難であった問題を解決していくことを可能とする．本研究では，計算機システム，ネットワークシステム，実世界に関する統合的なスケーラブルモニタリングに関するシステム構築をおこなうことで将来の計算機システムのあるべき姿を検討した．

研究成果の概要（英文）：In this project, a monitoring system architecture consists of a set of software to protect information infrastructures, social infrastructures and human everyday life. The goal of the project is to integrate research areas that are independently discussed before. The project developed several monitoring systems for computer systems, network systems and the real world to investigate the future information infrastructure.

交付決定額

（金額単位：円）

	直接経費	間接経費	合　計
2006 年度	12,100,000	0	12,100,000
2007 年度	13,700,000	0	13,700,000
2008 年度	12,300,000	0	12,300,000
2009 年度	14,000,000	0	14,000,000
2010 年度	12,200,000	0	12,200,000
総　計	64,300,000	0	64,300,000

研究分野：情報科学
科研費の分科・細目：情報学・ソフトウエア
キーワード：モニタリング，性能解析，実世界センシング，分散システム

1.研究開始当初の背景：

情報システムの安全性は現在の社会の安全性と密接に関連している．心理学者のアブラハム・マズローが示したように人間の安全の欲求を満足できない場合，人間がより快適な社会を築き，付加価値の高い社会を実現することは困難となる．現代社会はグローバル化により複雑化し，様々な危機が人間の安全の欲求の満足度を低下させている．情報環境インフラは社会インフラとして益々重要となり，我々の日常生活における高度な付加価値は情報環境インフラの進歩に大いに依存している．

2.研究の目的: 本研究課題では，システムモニタリングに適した分散システム／サービスアーキテクチャの開発，収集した情報を分析可能とするミドルウエア/サービスの開発をおこなう． また,開発したモニタリングアーキテクチャの有効性を示すため，システム自体の安全性を向上することを可能とする手法やユーザのミスや悪意のある攻撃を未然に防ぐための手法の開発も目指す

3.研究の方法:

(1) システムモニタリングのための基盤ソフトウエアに関する研究： アプリケーション，OS，ネットワークをモニタリングするための基礎技術の開発

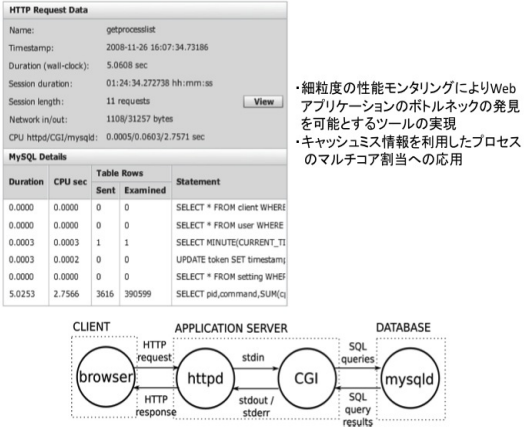
(2) センシングを利用した実世界モニタリングに関する研究： センシングを利用したミドルウエアやサービス構築をおこなった

4.研究成果

(1) Web サービス性能解析ツール mBrace

Web サービスの振舞いに関する異常を発見するためには,HTTP リクエスト毎に使用するリソース量の分析が必要となる．従来のモニタリングツールでは，複数の HTTP リクエストが使用するリソース量の全体の総量しか分析出来ないため，性能解析をピンポイントに分析することが困難であった．mBrace は Linux カーネルが提供する細粒度タイマー機能を利用することにより Web サービスが受け付ける各 HTTP リクエストが Web サーバ，CGI スクリプト，SQL サーバ等の各コンポーネントにおいて消費するリソース量を詳細に分析可能とする．これにより，Web アプリケーションのバグの発見やサーバのキャパシティプランニング等のために必要となる有用な情報を提供することが可能となる．

mBrace の実装を行うに当たり，OS には Linux を，また,Web サーバ・DB サーバには,



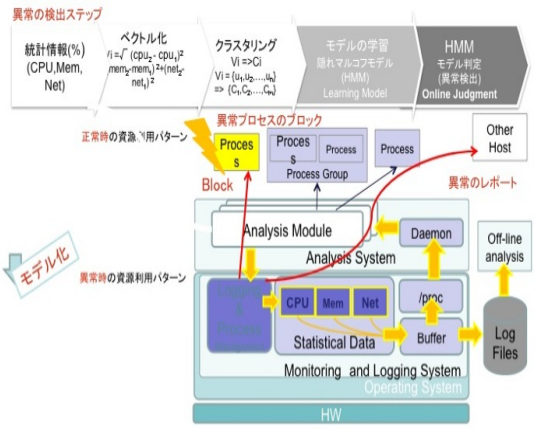
Apache と MySQL を用いて実装を行った.Linux には CPU の Performance Monitoring Unit (PMU)を有効にするためのカーネルパッチ(perfmon2)を適応した．そして，そのパッチと対応するライブラリ(libpfm)を用いてモニタリングシステムの実装を行った．PMU とは，プロセッサに組み込まれているモニタリング機能である.mBrace は PMU を使うことで，低オーバーヘッドなモニタリングを可能にしている．Apache サーバは，Apache サーバ自身と，フォークされた CGI プロセスのリソース使用量をモニタリングする．サーバ自身と CGI プロセスのモニタリングは，モジュールによって実装されている．また，クライアントからの HTTP リクエストを識別するために，ユニークな ID(mBrace アクション ID)を生成する．MySQL は，SQL クエリとその原因となった mBrace アクション ID とを組み合わせたログを出力する．このためには，Apache から MySQL へ，mBrace アクション ID を通知する必要がある．そのため，MySQL コマンドのプロトコルを拡張し,COM_MBRACE コマンドを作成した．COM_MBRACE コマンドは，元となる MySQL コマンドに mBrace アクション ID を付加したものである．そして，MySQL サーバが COM_MBRACE コマンドを受信したとき，自身のモニタリングを開始するように変更を加えた．現在，Linux への修正は，カーネルパッチを適応するだけで可能である．また，Apache・MySQL 本体へは，それぞれ 13 行・846 行の変更を加えている．本研究では，mBrace は極めて低いオーバヘッドで Web アプリケーションの性能異常の解析をおこなうために必要となる情報が抽出可能であることを示した.

本研究の結果は，Journal of Web Services Practices，IEEE RTCSA 2010 等において発表をおこなった.

(2) カーネルログを機械学習を用いて異常検出を可能とする Ayaka

Ayaka は Linux が提供する LTTng というカーネルイベントのロギングのためのツールを利用して，各アプリケーションのカーネル内のアクティビティの振舞いをログとして抽出する．そして，それらを解析することにより，アプリケーションの振舞いの分析を可能とする．Ayaka は LTTng により抽出したログの内、CPU やメモリ，ネットワークリソースの使用量に関する情報だけをフィルタリングし，機械学習アルゴリズムを利用して通常の正常状態の振舞いのパターンを学習させる．その後に，学習したパターンと異なるパターンの発生を定期的にチェックする．異なるパターンの発生は異常状態の発生を意味するので，異常状態の検出が可能となる．

実際に，Ayaka を利用してサイトへの侵入検知が可能となることを示した．Ayaka のカーネル内の振舞いの学習は計算機毎におこなう必要がある．そのため，大量の台数の計算機が存在するクラスタ環境やユビキタスコンピューティング環境には適していないことが明らかになった．また，Ayaka が提供する振舞いを解析する手法は，どこに原因があるかを特定することが容易でないことが明らかになり，それが３つ目のモニタリングシステムを構築する動機となった．

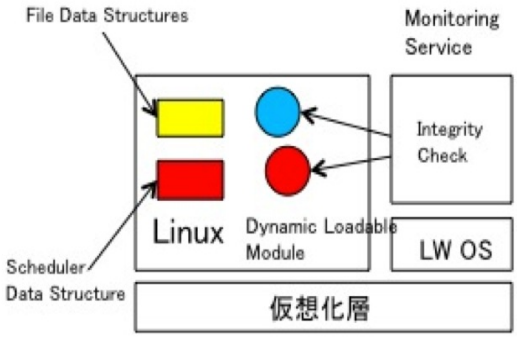


Ayaka: リソース情報を用いた学習モデルによる異常検出・管理システム

本研究の成果は，IEEE ISORC 2009 において発表をおこなった．

(3) スケーラブルモニタリングを可能とする IntegrityMonitor

IntegrityMonitor はLinux カーネル内のデータ構造の状態遷移をモニタリングする．異常を検知するためには，予めどのような異常が発生するかを想定し，異常な状態に陥ったときに，正常な状態との違いを明らかにしておく必要がある．IntegrityMonitor はこの違いを用いることで監視対象の異常を検出することが出来る．しかし，この違いを手作業で定義していった場合でもその異常を検知することが出来る．しかし，この手法では想定外の異常に関しては検知することが出来ず，また，多くの異常を検知するためには，非常に大きなエンジニアリングコストが必要となる．従って，自動的に異常な状態と正常な状態を区別する定義を生成することが求められている．そこで，IntegrityMonitor は Daikon というツールを使って，この違いを定義する．Daikon ツールを使うためには，監視対象の OS が用いるデータ構造の定義を取得し，それらのデータが取る値を集めなければならない．我々の監視対象である Linux はオープンソースであるので，ソースコードを入手することが可能である．従って，ソースコードを解析することで，データ構造の定義を取得することが出来る．また，データを集めるプログラムも同時に自動的に生成することが可能である．この Daikon ツールを用いたモニタリングサービスの自動生成手法の利点は生成した普遍量を再利用出来るということである．データ構造の定義が同じであるため，その振舞いは大きく変わることがない．また，解析にかけるデータを多くすればするほど，異常検知の精度が上がる．従って，個々の計算機毎に学習させることが困難なコンシューマエレクトロニクスやクラスタサーバにおいても，この手法を用いることで不変量を作り出し，異常を検知することが可能である．実際に，IntegrityMonitor は既存のカーネル Rootkit の多くを発見することが可能であり，セキュリティの向上の面で極めて有効な手法であることが示された．



IntegrityMonitor: Linuxカーネルセキュリティモニタリング

本研究の成果は，IEEE ISORC 2010 において発表をおこなった．

(4) ネットワークのモニタに基づくセキュリティ対策

モニタリング（測定）技術を活用してネットワークのセキュリティ対策に有効な方法を確立することができる。基本的なアイデアは次の通り。インターネットのホストには一意の IP アドレスが割当られている。実際には、割当をしていない未使用の IP アドレスが存在する。その IP アドレスに流入するパケットをモニタしてみると、通常の利用ではないパケット、多くは悪意のあるスキャン、ウィルスを拡散する行動、ボットネットなどのマルウェアの活動を観測することができる。このような測定法は既存技術として存在しており、Dark IP などと呼ばれている。本研究は Dark IP を発展させて仮想的に実現す

る。すなわち上記のフローのモニタ技術を活用して、未使用の IP アドレスを宛先としているパケットを収集する。この方法を活用すると、数千台のモニタを設置したのと同等の情報を１台のルータから収集することができるため甚だ効率が良い。本研究では、さらに仮想 Dark IP を発展させた。単に観測するのではなく悪意のあるパケットに対して応答を返す。すると悪意のある送信者が次のパケットを送ってくる。これを観測すると、セキュリティにかんする情報を多く集めることができる。本研究では、この技術を仮想ハニーポットと名付けた.

本研究は，電子情報通信学会 論文賞および電気通信普及財団 テレコムシステム技術賞受賞し高く評価されている.

(5) モニタリングデータからのデータマイニングとシステム高速化

アプリケーションには一般に，どのファイルへいつアクセスするかといったアクセスパターンがあるため，アプリケーションのファイルアクセス時系列をシーケンシャルパターンマイニングすることで，頻出アクセスパターンを抽出することができる．そこで，OS 層に追加した機能を用いて頻出ファイルアクセスパターンを抽出し，OS が持つファイルキャッシュ管理機構へ頻出パターンをヒント情報として与え，ランダムアクセスされるデータを優先的にキャッシュすることでシステム性能の向上を図った．ストレージへのランダムアクセスは低速なため，ランダムアクセスされるデータを優先的にキャッシュすることでシステム性能を向上することができる．我々は，上記の処理を全て OS 内でオンラインに実行する手法を開発した．ファイルアクセスのモニタリングデータは膨大になるため，性能向上に有効でないモニタリングデータをフィルタすると共に，CPU のアイドルタイムを活用してパターン抽出を行うことでオンラインでのシステム高速化を実現した．提案手法は Linux Kernel に実装しており，データベースベンチマークである TPC-H において性能向上を確認できた．本研究は，膨大なモニタリングデータからデータマイニング技術を用いて情報を抽出し，実システムの高速化に利用できることを示す具体例である.

(6) モニタリングデータを利用した各種サービス

スケーラブルなモニタリングシステムの実現により，ひとつの事象について，多次元の情報が詳細に取得できるようになる．その結果，情報量は爆発的に増加し，利用者が優先的に参照したい情報以外の情報も過剰に提供してしまう事態が生じる．また，情報を詳細に取得できるようになったことで，取得した情報の全体的な傾向が明確でなくなったり，多くのノイズが乗った情報となってしまったりといった副作用もある．こうした問題を解決するために，取得したモニタリングデータの利用目的に応じて不要な情報を除去し，利用者が望む情報を明確に提示するサービスが必要となる．そこで本研究では，OS，センサ，モニタリングツール等からのモニタリングデータを解析・加工し，新たな知見を引き出したり，獲得した情報を利用者に分かりやすく提示したりするサービスの提案をおこなった.

セキュリティサービスは実際に研究室したシステムの１つである．システムログの監視は，セキュリティ対策の基本的かつ重要なタスクのひとつである．しかし，膨大なシステムログから問題のある箇所のみを抽出するためには，一定の経験を必要とする．また，マシン単体のシステムログを参照しても問題を特定することが困難な場合も多く，複数のマシンで取得したシステムログを総合的に判断することが必要となる．そこで，管理下にある複数のマシンで取得したシステムログを統合し機械学習させることで，システムの異常や外部からの攻撃を発見し，システム管理者に警告を発するサービスを提案した.

(7) 小型ノード向け仮想化レイヤ:

オペレーティングシステムを外部からモニタリングするためには、仮想化レイヤを導入し、その上でオペレーティングシステムを動作させる必要がある。仮想化レイヤは、これまで主にサーバをターゲットとして開発されてきており、そのため x86 アーキテクチャではいくつかの仮想化レイヤが使用可能である。また、ハードウェアで仮想化をアシストする機能も、プロセッサに搭載されるようになってきている。しかしながら、小型ノードで用いられるような組込み用途のプロセッサでは、仮想化レイヤの利用はほとんど例がなく、そのため仮想化レイヤの開発事例もあまり見られない。そこで、組込み用途で幅広く用いられている ARM プロセッサをターゲットとする仮想化レイヤの研究を行った。ARM プロセッサは、ユーザとカーネルの 2 種類の保護レベル、実行モードによって使用するレジスタが切り替わるバンクレジスタという、組込み用途のプロセッサとしては典型的な特徴を持つ。仮想化レイヤは、ユーザプロセスとカーネルを異なる保護ドメインで、ユーザレベルでカーネルを実行することで、カーネルと仮想化レイヤを保護する。ユーザ

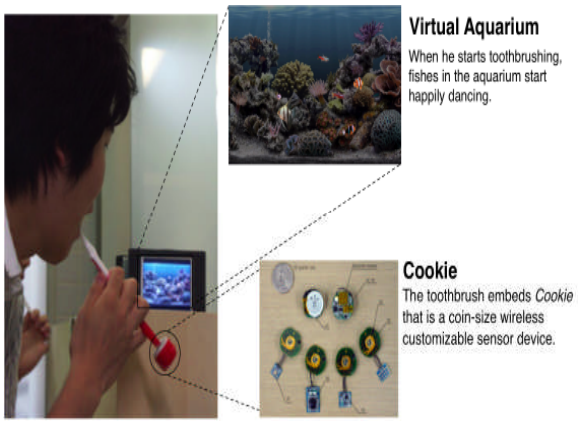
レベルでカーネルを実行するために、仮想化レイヤは、保護レベル、実行モード、バンクレジスタを仮想化する。また、仮想化レイヤでエミュレーションする必要がある ARM プロセッサの特権命令について明確にした.

(8) ハイブリッド型 P2P ライブ配信アーキテクチャ ToMo

ネットワークワイドの研究例として、P2P 技術を用いたライブ配信において、制御オーバーヘッドの少ないツリー型配信方式と障害に強いメッシュ型配信方式を組合せることで、高信頼かつスケーラブルなデータ配信を実現するハイブリッド型 P2P ライブ配信アーキテクチャ ToMo (Two-layer Mesh/tree overlay) の提案を行った。本方式では、まず、上位層に位置する各ノードはメッシュ網を構成し、複数の隣接ノードを互いにモニタリングしながらデータ配信を実行する。障害を検知した場合には、迅速に異なるノードに配信要求を投げてネットワークを再構成し、セッションの継続を試みる。また、下位層に位置する各ノードはツリー網を構成し、自身の親ノードのみをモニタリングしながらデータ配信を実行する。障害対応はメッシュ網ほど迅速には対処できないが、逆にオーバーヘッドは減少し、スケーラビリティの改善に寄与する。また、提案方式のデータセンターや CDN (Content Delivery Networks) への応用についても検討を行った。

本年度はさらに上記のネットワーク構成において、各ノードを貢献度 (インセンティブ) に応じて上位層と下位層に振り分ける仕組みを導入し、システムの実現性を高めた。

(9) センシング応用サービス



センシングにより実世界の様々な情報を取得することが可能になった. 特に, 人間の一般的な活動の状況を取得し, その情報を利用して, 人間の意思決定を支援するために有益な情報を提供することが可能となった. 本研究では, 実世界のセンシングを特別な機器を日常生活内に配置しなくても可能とするための基盤ソフトウエアとその基盤ソフトウエアを利用したサービスの構築をおこなった. はじめに, センサ機器を容易に実世界環境に配置することを可能とするミドルウエアの開発をおこなった. 開発したミドルウエアは, センサーを組込んだ家電機器をドキュメントを記述して組み合わせることだけで複雑なコンテキストアウェアサービスを構築することを可能とする. 実際に, 過去に構築されたサービスを容易に構築することが可能となったことが示された. 次に, 基盤ソフトウエア上のサービスの１つとして, ライフスタイルアンビエントフィードバックシステムを構築した. このサービスは, ユーザの行動を反映するメタフォアを日常の環境内に表示する. Virtual Aquarium は歯磨きという健康の為に必要な行動を洗面所内に置かれた仮想水槽におけるフィードバックとして返す. また, EcoIsland では, ユーザのエコフレンドリな行動を家族が生活する島の状況としてフィードバックをおこなう. 本研究は, 世界的に高く評価され, 様々な研究機関において講演をおこなった.

最後に, UbiPay と呼ぶ経済的インセンティブを様々なサービスにおいて利用するための基盤システムの構築をおこなった. UbiPay はユーザのアクティビティをモニタリングすることにより, 必要に応じて自動的に支払いをおこなうことを可能とする. これにより, 非常に少額のトランザクションを認知負荷なくおこなうことが可能となり, 新たなサービスの可能性を提供することが可能となった.

本研究の成果は, ACM DIS2008, Ubicomp2008, Ubicomp2009 等のトップクラスの国際会議に採録され, 国際的に高い評価を得ている.

5. 主な発表論文等 (研究代表者、研究分担者及び連携研究者に は下線)

〔雑誌論文〕(計11件)

1 Andrej Van Der Zee, Alexandre Courbot, Tatsuo Nakajima, mBrace: Fine-Grained Profiling of Multi-Tier Web Applications, International Journal of Web Services Practices, Vol. 5, No. 1, pp. 10-21, 2010. (査読あり)

2 Midori Sugaya, Yuki Ohno, Andrej van der Zee and Tatsuo Nakajima, "A Lightweight Anomaly Detection System for Information Appliances", IEEE International Symposium on Object-Oriented Real-Time Distributed Computing, pp. 257-266, 2009. (査読あり)

3 Yuki Ohno, Sayaka Akioka, Midori Sugaya, Tatsuo Nakajima, A Library-Based Tool to Improve CPU Assignment for Multicore Processor-Based Pervasive Servers, IEEE 16th International Conference on Embedded and Real-Time Computing Systems and Applications, pp.114-123, 2010. （査読あり）

4 Fahim Kawsar, Kaori Fujinami, Tatsuo Nakajima, Deploy Spontaneously: Supporting End-Users in Building and Enhancing a Smart Home, The Tenth International Conference on Ubiquitous Computing, pp.282-291, 2008. （査読あり）

5 Tetsuo Yamabe, Vili Lehdonvirta, Hitoshi Ito, Hayuru Soma, Hiroaki Kimura, Tatsuo Nakajima, Applying Pervasive Technologies to Create Economic Incentives that Alter Consumer Behavior, The 11th International Conference on Ubiquitous Computing, pp.175-184, 2009. （査読あり）

6 近藤 秀和,村岡 洋一, ウェブブラウザ「Lunascape」, コンピュータソフトウェア, Vol. 24, No. 4, 139-152, 2007. （査読あり）

7 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. （査読あり）

8 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 pp.163-173, 2009. （査読あり）

9 Tetsuya Kusumoto, Jiro Katto, Sakae Okubo, Proactive Route Maintenance for Tree-Based Application Layer Multicast and Its Implementations, IEICE transactions on information and systems E89-D(12), pp.2856-2866, 2006. （査読あり）

10 Shuichi Oikawa, Megumi Ito, Tatsuo Nakajima, Linux/RTOS Hybrid Operating Environment on Gandalf VMM, The 2006 IFIP International Conference on Embedded and Ubiquitous Computing, pp.287-196, 2006. （査読あり）

11 Hiromasa Shimada, Alexandre Courbot, Yuki Kinebuchi, and Tatsuo Nakajima, A Lightweight Monitoring Service for Multi-core Embedded Systems, 13th IEEE International Symposium on Object Oriented Real-Time Distributed Computing, pp. 202 - 209, 2010. （査読あり）

〔図書〕（計1件）
1 後藤滋樹,外山勝保, インターネット工学, コロナ社, 2007.

〔その他〕 ホームページ等
http://www.dcl.info.waseda.ac.jp/

6.研究組織
(1)研究代表者
中島 達夫(NAKAJIMA TATSUO)
早稲田大学・理工学術院・教授
研究者番号: 10251977

(2)研究分担者
村岡 洋一(MURAOKA YOICHI)
早稲田大学・理工学術院・教授
研究者番号: 50182085
(H21→H22:連携研究者)

後藤 滋樹(GOTO SHIGEKI)
早稲田大学・理工学術院・教授
研究者番号: 30287966
(H21→H22:連携研究者)

山名 早人(YAMANA HAYATO)
早稲田大学・理工学術院・教授
研究者番号: 40230502
(H21→H22:連携研究者)

甲藤 二郎(KATTO JIRO)
早稲田大学・理工学術院・教授
研究者番号: 70318765
(H21→H22:連携研究者)

追川 修一(OIKAWA SHUICHI)
筑波大学・システム情報工学研究科・准教授
研究者番号: 00271271
(H21→H22:連携研究者)

(3)連携研究者
秋岡 明香(AKIOKA SAYAKA)
早稲田大学・IT 研究機構・准教授
研究者番号: 9033533