Vol.2014-IOT-27 No.14
Vol.2014-SPT-11 No.14
2014/10/10

監視項目の時系列類似度に基づく障害把握を支援するシステムの検討

宮座 菜緒[†1,a)]　桝田 秀夫[†2]

概要：ネットワーク技術の発展により，組織のネットワークは仮想化などといった物理構成に縛られにくい柔軟性のあるサービス基盤を用い，様々なコンポーネントを連携させながら我々の生活に必要不可欠なサービスを提供している．このようなサービスが広まるとともに，利用者にサービスを継続して提供する重要性も高まり，常に安定した運用管理が求められているが，サービス基盤内の構成が複雑であったり，監視対象や監視機器の種類の増加に伴い，互いのコンポーネントの依存関係や関連性の把握が難しく，管理者が障害の原因を突き止めることが困難となっている．そこで本稿では，複数の機器の監視項目に対して，グラフの表示方法や時系列の類似性に着目し，障害 の引き金を探し当てることを支援するシステムの検討を行ったので，現状を報告し考察する．

キーワード：時系列類似度，zabbix，グラフの重ね合わせ

Consideration of a system to support understanding of fault occurrences based on the similarity of the time series

Miyaza Nao[†1,a)]　Masuda Hideo[†2]

Abstract: With the development of network technology, the organizations use the network which is based on flexible service infrastructure such as virtualization ， and provide us with essential services which are linked up with various components. With the spread of such services，organizations should be able to continuously provide the services to the users. However, because of the complicated constitution of the service infrastructure and the increase of the types of surveillance devices，the organization's network managers can't understand the relevance of various monitoring components. Therefore it is difficult for the organization's network managers to find the cause of the fault occurrences. This paper explain the consideration of a prototype to support them to find the cause of the fault occurrences by examining the expression method of graph and the similarity of time series regarding the various components monitored by the systems.

Keywords: similarity of time series , zabbix ， combined graph

1. はじめに

近年，ネットワーク技術の発展により，教育機関や企業などの組織のネットワークは複数のサーバやネットワーク機器で構成され，大規模かつ複雑になってきている．VMware や Xen などの仮想化技術の進歩に加えて，10GbE や Infiniband などの高速なインターコネクト技術により，物理構成に縛られにくい非常に柔軟性のあるサービス基盤が容易に手に入るようになってきた．また，サーバは様々なコンポーネントが連携して動作することで電子メールや WWW など，我々の生活に必要不可欠なサービスを提供している．

このようなサービスが広まるとともに，利用者にサービ

†1 現在，京都工芸繊維大学大学院工芸科学研究科情報工学専攻
Presently with Graduate School of Information Science, Kyoto Institute of Technology
†2 現在，京都工芸繊維大学情報科学センター
Presently with Center for Information Science, Kyoto Institute of Technology
a) n-miyz13@dsm.cis.kit.ac.jp

スを継続して提供する重要性も高まり，サーバやネットワーク機器などといったこれらを支えるシステムに対して，常に安定した運用管理が求められている．サーバやネットワーク機器の障害により，こうしたシステムが停止してしまうと，組織自体の活動も停止してしまう恐れがある．よって，システムを安定に稼働させ，サービスをできるだけ停止させることなく提供し，たとえ障害が発生した場合でもすぐに検知し，復旧できる体制を整えておくことが大切である．システムを効率的に安定して運用させるためには，サーバのハードウェアに故障が発生していないか，システムのリソースは足りているか，アプリケーションやプロセスは正常に稼働しているかなど，システム全体の稼働状況をリアルタイムに把握できるように監視することが重要である [1]．しかし，最近では，サーバやストレージの仮想化によってサービス基盤内の構成が複雑になってきている．また，仮想マシン上では様々な種類の OS が扱えたり，数を増やすことも容易にできたりするため，監視対象や監視機器の種類も増えてきている．さらに，大規模なシステムを運用する組織のサーバ群は，ストレージサーバ，データベースサーバ，web サーバ，各種アプリケーションサーバなどが互いに依存関係を持つことが多い．よって，機器の接続関係といった物理的なコンポーネント間の関係だけを把握していても，本当にそれらが連携して実際のサービスを提供しているのかどうか，および，そのサービスはどの機器に依存して動作しているのかといった関連性までは把握することが難しい．また，連携して動作しているシステムの一部の機器に障害が発生した場合，関連している機器が個別に障害情報を出力するため，管理者が障害の原因を突き止めることも困難となっている．

そこで本稿では，複数の機器の監視項目に対して，グラフの表示方法や時系列の類似性に着目し，障害 の引き金を探し当てることを支援するシステムの検討を行ったので，現状を報告する．

2. システム監視

システム監視とは，システム内で動作しているサーバ、アプリケーション，ネットワークなどが正常に稼働しているかどうかを監視することで，システムで発生した障害やリソース不足を検知し，管理者に通知を行うための作業や仕組みのことである [2]．

2.1 統合監視システム

システム監視には，稼働監視，リソース監視，アプリケーション監視などがあるが，これらの監視を全て手作業や自作のスクリプトで実施したり，それぞれ異なるソフトウェアを組み合わせて利用する場合，ス クリプトの開発や設定管理のメンテナンス作業の増加により監視システム自 体の維持管理にコストがかかるようになる．また，スクリプトや設定が適切に管理され，確実に実行されているかどうかを把握すること自体が難しくなってくる．このような課題を解決する方法として，統合監視システムの利用が挙げられる．統合監視システムは，主に次の 3 つの機能を有している．

(1) データ収集機能
ネットワークを介して複数のサーバやネットワーク機器，リソース，アプリケーションを定期的に監視しつつ，収集したデータを一元管理する機能

(2) 障害検知，通知機能
収集したデータが正常範囲であるかどうかを判断し，障害と思われる場合は管理者に通知する機能

(3) 表示機能
専用の管理インタフェースから，収集した情報の履歴やグラフの表示，監視設定を行う機能

近年では，こうした統合監視システムの開発も進み，Nagios[3] や Zabbix[4]，Hinemos[5]，HP OpenView[6] などといったシステムが数多く存在している．

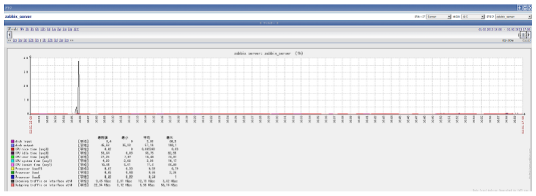


図 1 スケールの大きいものを含む場合
Fig. 1 Example of including a large scale graph

図 2 スケールの大きいものを除いた場合
Fig. 2 Example of excepting a large scale graph

3. 解決すべき問題点

システムを監視する際に解決すべき問題点について述べる．

3.1 コンポーネント間の関連性把握

仮想化の普及により，サービス基盤内の構成が複雑となり，各機器のコンポーネント間の関連性がわかりにくくなってきていたり，複数の機器同士が依存関係を持っていたりするために，どの機器のどの部分が原因となり障害が発生したものなのかを把握しにくいという問題がある．

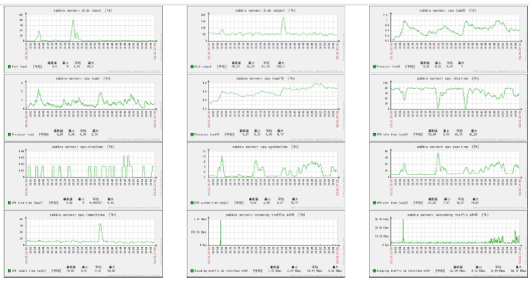


図 3　グラフを並べた場合
Fig. 3　Example of lining up graphs

3.2　グラフエリアの複雑化

統合監視システムを用いて，管理者 が自身のサービス基盤内の機器を管理する場合，項目ごとにグラフを生成することが多く，複数の監視項目をグラフにする場合でも，グラフエリアが複雑になり，どれがどの項目を示すグラフかわかりにくくなってしまう．また，項目同士の関連性もグラフのスケールの違いなどによって，他のグラフの線にかき消されてしまい，わかりにくくなってしまうという問題がある．

図 1 はスケールの違いによって他のグラフの線がわかりにくくなってしまっている例である．図 2 は図 1 でスケールの大きかったものを除外し再度グラフを作成したものである．なお，これらのグラフは統合監視システム Zabbix で監視を行ったものを，Zabbix のグラフを表示する機能を用いて作成したものである．このように図 1 では図 2 のような変動を見ることはできない．また，スケールの大きかった項目は他の項目と比較できないこともわかる．

3.3　画面領域の制限

前節でグラフエリアの複雑化を挙げたが，これに対処するために，監視項目ごとに個別にグラフを作成し，並べて比較してみるという方法が考えられる．図 3 がその例であり zabbix の機能を用いて作成したものである．しかし，この方法では，1 つの画面領域に並べて表示できるグラフの数が限られてしまうため，項目数が多くなると比較できないという問題がある．また，並べ て表示できたとしても，一目ではそれぞれの特徴を把握することが困難となり管理者の負担が大きくなってしまうという問題もある．

4.　システムに求められる要求

システム監視の際に求められる要求について述べる．

要求 1：システム管理者の負担が少ないこと

仮想化などの技術により，監視台数や種類が増加し，組織のネットワークも大規模化している．そのような状況の中で，管理者が監視項目ごとにそれぞれグラフを作成していると，管理者の負担は非常に大きくなってしまう．また，手作業による人為的なミスなども生じる恐れがあるため，管理者の負担が少ないシステムが求められる．

要求 2：コンポー ネント間の関連性把握ができること

障害発生時などに迅速に対応できるように，単純にグラフ化してもよくわからなかったコンポーネン ト間の関連性や依存関係が把握できる仕組みが求められる．

要求 3：限られた画面領域で必要な情報が提示されること

監視項目が増えたとしても，限られた画面領域の中で必要な情報を提示できる仕組みが必要である．そのためにも，スケールの大きい項目に他の項目がかき消されてしまったり，項目数が多くなって，グラフエリアが複雑になり，互いに情報をかき消しあったりしてしまうなどといった問題が起こらないようにしなければならない．

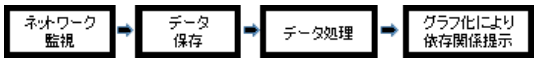


図 4　構成図
Fig. 4　Composition of system

5.　提案手法

今回，上記の要求を満たすためのシステムとして
(1) 重ね合わせグラフ表示法
(2) 時系列類似グラフ抽出法
の 2 手法を提案する．
なお，両手法ともシステムの構成は図 4 のようになる．

手順 1：システム監視

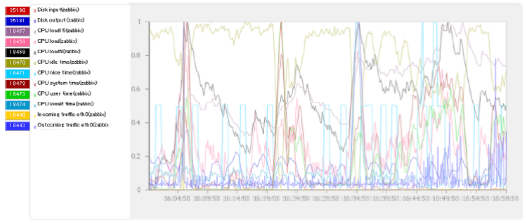
本研究では，システムの障害時の引き金を探し当てることについて重点を置いているため，ネットワークを監視する方法についてはあまり重要視していない．ここではオープンソフトウェアである Zabbix を用いてネットワークを監視する．Zabbix は監視データを扱うサーバに Zabbix サーバをインストールし，監視対象とするサーバには Zabbix エージェントをあらかじめインストールしておく．また，スイッチなど，Zabbix エージェントをインストールできない機器で SNMP エージェント機能を有しているものは SNMP コマンドを用いて監視情報を収集する．

手順 2：データ保存

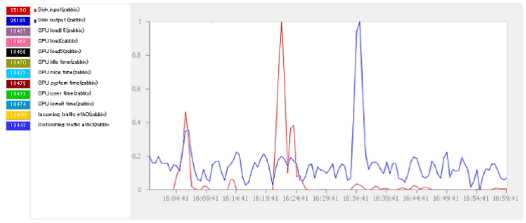
収集したデータはデータベースに保存する．なお，Zabbix は MySQL サーバを基本として開発されていることや，動作の高速さや軽量さの観点から，今回はデータベースとして MySQL を使用する．また，収集する監視項目やデータベースへのデータ保存期間の設定などは Zabbix の web インタフェースで行った．

手順 3：データ処理方法

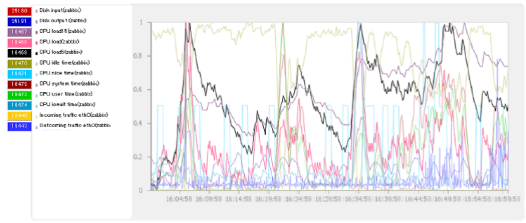
収集したデータは監視項目ごとにスケールが違うため



(a) 表示
(a) Display



(b) 非表示
(b) Hide



(c) 強調表示
(c) Emphasis

図 5 重ね合わせグラフに実装した機能
Fig. 5 Functions of overlay graph

正規化を行う．正規化の方法としては，各監視項目において比較する日のデータの最大値と最小値がそれぞれ 1 と 0 になるように計算を行う．

手順 4：グラフ化により提示

Web インタフェースで生成したグラフを表示させる．

6. グラフ生成法

6.1 重ね合わせグラフ表示法

グラフの表示方法に着目する方法として，1 つのグラフエリアに複数のグラフを重ね合わせて表示する方法を提案する．

方法としては，正規化を行ったデータをグラフ化し重ね合わせて表示する．なお，グラフ化したデータをユーザが操作できるように図 5 のように重ね合わせたグラフをユーザが項目を選択することにより表示，非表示，強調表示できる機能を実装した．

6.2 時系列類似グラフ抽出法

データ自体に着目する方法として，時系列データの類似性に着目する方法を提案する．

6.2.1 グラフの類似度

今回，類似度という監視したデータをグラフにした際の

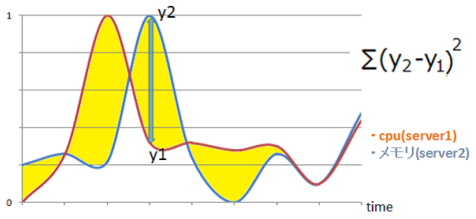


図 6 類似度算出方法
Fig. 6 Calculation method of similarity

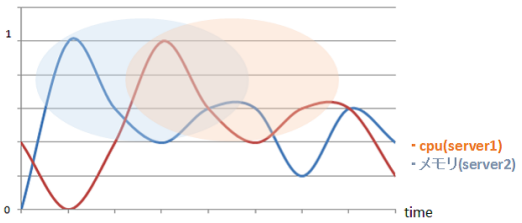
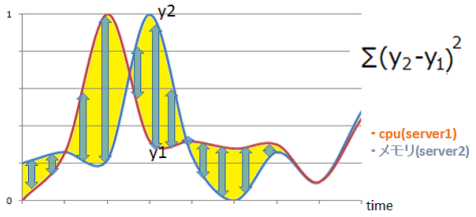
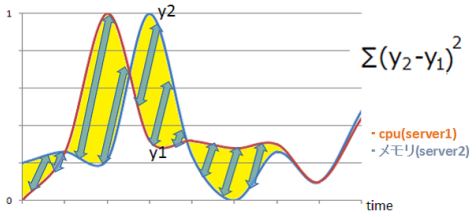


図 7 タイムラグの関係
Fig. 7 Relationship of the time lag



(a) 通常の算出方法
(a) The method of normal calculation



(b) 時間軸をずらした算出方法
(b)The method of calculation using shifted time axis

図 8 時間を考慮した算出方法
Fig. 8 The method of calculation with consideration of time

形状が似ている度合を用いてデータを比較した．類似度算出方法としては図 6 のように 1 日のデータの差の二乗和で算出し，算出された値が小さいほど監視項目同士が似ていると判断し，類似度が高いとする．

また，図 7 のように時間軸がずれて同じような動きをしている項目同士も関係があるのではないかと考え，それらを抽出できるように時間軸を考慮した類似度算出を行う．時間軸を考慮した類似度算出の方法としてデータの時間軸を 1 つ 2 つ…とずらしてそれぞれ類似度比較を行い，一番値が小さかった組み合わせのものを用いる．図 8 はその様子を表す．

図 9　過去データとの比較
Fig. 9　Comparison with the past data

6.2.2　同一項目上の過去データとの比較による抽出

比較したい日のデータを各監視項目ごとに過去 4 週分のデータと類似度比較し，過去のデータと違う変化をしていないか確認する．なお，Zabbix のデフォルトの機能で各監視項目のデータを 30 分ごとに平均してデータベースに保存する機能があるので，今回は 30 分ごとの平均値のデータをサンプル値として比較する．比較方法としては，まず各監視項目ごとに指定した日の 1，2，3，4 週間前のそれぞれの日のサンプル値を抽出しそれらの平均値を求める．求めた平均値のデータと指定した日のサンプル値を類似度比較する．類似度比較した結果を類似度の 低い順にソートし，普段と違う動きをしている順に並べる．図 9 は過去のデータの平均と比較したいデータを描画したものをソートして並べたものの一部である．

6.2.3　指定期間内の複数項目間との比較による抽出

ソートしたときに一番上に上がってきた最も普段と違う動きをしている項目をベースに，他の監視項目と類似度比較し類似度の高い順にソートすることで，ベースに似ている監視項目順にソートする．このとき類似度比較は正規化を行った後のデータで行う．また，引き金関係にある項目同士を抽出するために，時間軸を考慮した類似度比較を行う．図 10 は指定期間内の複数の項目のデータを描画したものをソートして並べたものの一部である．このように並べて表示することにより注目すべき関連性のありそうな監視項目を絞って提示する．

7.　考察

7.1　重ね合わせグラフ表示法

正規化したデータをグラフ化することにより，スケールの大きい項目に他の項目がかき消されてしまう問題を解決した．しかし，一見するとグラフの線により，グラフエリアが複雑になってしまっているようにも思える．これは，本システムのボタン機能を用いることで，注目したいグラフのみを表示させたり強調することによってある程度は解決できると考えられる．しかし，項目数がこれ以上増えてくるとグラフエリアはさらに複雑になり，ボタン機能では解決できないと考えられる．よって，新たなグラフの表示方法を考える必要があると考えられる．

7.2　時系列類似グラフ抽出法

監視項目が増えたとしても類似度比較を行い注目すべき関連性の高い監視項目を抽出し提示しているため，管理者の負担は少なくなると考えられる．また，監視項目を抽出し提示しているため，限られた画面領域の中に必要な情報を提示でき，グラフエリアが複雑になってしまう問題も解決している．しかし，今回の類似度比較ではグラフの形状に注目した計算方法になっているため，それ以外の類似性などの抽出はできていないと考えられる．よって，注目すべき情報が抽出しきれていない場合などがあるため，今後さらなる比較方法の検討が必要であると考えられる．

8.　まとめ

本研究ではシステム監視について複数の機器の監視項目に対して，グラフの表示方法や時系列の類似性に着目し，障害 の引き金を探し当てることを支援するシステムの検討を行った．重ね合わせグラフ表示法では，ある程度の項目数までは実装したグラフの機能で対応することができたが，さらに項目数が増えると対応しきれないといった問題があるため，これらを解決するためにはグラフの表示方法をさらに工夫したり，事前にグラフのデータを絞りこむような処理を施したりなどといったことが必要になってくると考えられる．また，時系列類似グラフ抽出法では，注目すべき関連性の高い監視項目を抽出し提示することにより管理者の負担を少なくすることができた．今後は，必要な情報が確実に抽出できているかどうかや，さらに正確性の高い抽出方法があるかどうかなどといった検討を行ってい

図 10　複数項目間との比較
Fig. 10　Comparison with multiple components

く必要があると考えられる．

参考文献

[1] 片岡巖：サーバ/インフラエンジニア養成読本管理/監視偏，株式会社技術評論社 (2012).
[2] 寺島広大：，Zabbix 統合監視 [実践入門]-障害検知，傾向分析，可視化による省力運用，株式会社技術評論社 (2010).
[3] Nagios Enterprises：Nagios-The Industry Standard in IT Infrastructure Monitoring(online)，入手先 ⟨http://www.nagios.org/⟩ (2014.09.10).
[4] Nagios Enterprises：Nagios(online)，入手先 ⟨http://www.nagios.org/⟩ (2014.09.10).
[5] NTT DATA CORPORATION：Hinemos(online)，入手先 ⟨http://www.hinemos.info/⟩ (2014.09.10).
[6] Hewlett-Packard Development Company：HP OneView(online)，入手先 ⟨http://h50146.www5.hp.com/products/servers/proliant/management/ov/index.html⟩ (2014.09.10).
[7] 宮座菜緒　桝田秀夫：重ね合わせグラフによる項目間の関連性把握を支援するシステムの試作，京都工芸繊維大学工芸科学部情報工学課程卒業論文 (2013).