

## DynamicSVD によるサーバ負荷予測手法について

## On Server Load Prediction Method Using DynamicSVD

○平田 飛仙<sup>†‡</sup>、鹿島 理華<sup>†</sup>、菅野 幹人<sup>†</sup>○Takahisa HIRATA<sup>†‡</sup>, Rika KASHIMA<sup>†</sup>, Mikihiro KANNO<sup>†</sup>三菱電機株式会社 情報技術総合研究所<sup>†</sup>

## 1 背景

FTTH (Fiber To The Home) の普及、大手通信会社による NGN (Next Generation Network) サービスの開始など、ブロードバンド通信インフラの発達に伴い、ネットワーク経由で機能を提供するソフトウェア形態である SaaS (Software as a Service) が注目されている[1]。またSaaSをはじめとした、ネットワークサービスを支えるITインフラとしてVM (Virtual Machine) 技術への関心も高まっている。

本研究の目的は、ネットワークサービスを効率的に運営するための基礎技術として、サーバの負荷分散を最適化する手法を開発することである。サーバの負荷を適切に分散することで、スループットの向上によるサービス品質の改善や、遊休設備の削減によるコスト低減などの効果が期待される。また、将来的には本手法をネットワークの負荷分散などに応用することも検討している。

## 2 問題定義

## 2.1 サービス提供モデル

現在我々が想定しているネットワークサービスは、主として業務システムを SaaS 化したものであり、そのサービス提供モデルを図 1に示す。このモデルでは1つのホストサーバ上で複数のVMが動作しており、それぞれのVMは1つのサービスを提供するパッケージに対応しているものとする。またサービスの提供者は、このようなホストサーバを複数運営することで、多数のユーザにサービスの提供を行うものとする。

## 2.2 問題定義

我々の目的は、上記のような環境下で、VM 毎の負荷情報に基づき、VM への動的な資源割当て

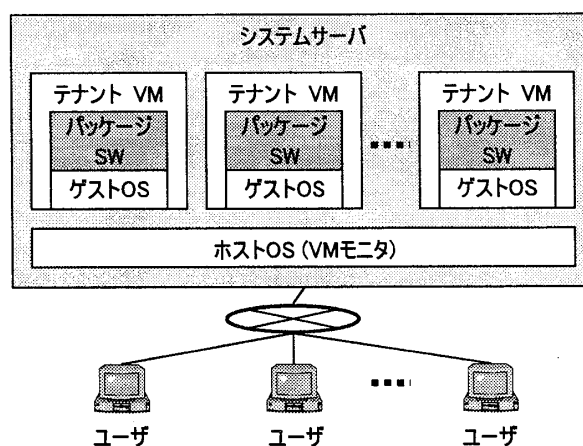


図 1 VM によるサービス提供モデル

や、ホストサーバ間での VM の再配置などにより、ホストサーバの負荷を均等にすることである。前者の例としては一時的なアクセス集中や不定期な高負荷処理への対応が挙げられる。このような目的に対してはオンライン計算による短期的な負荷予測が求められる。一方、後者の例としてはサービスの定常的な負荷パターンによる分類により、夜間に負荷の高いサービスと日中に負荷の高いサービスを組み合わせて負荷を平均化することなどが挙げられる。このような目的に対してはオフライン計算による長期的な負荷予測が必要である。

また実際にネットワークサービスを運用する上では、多数の VM に対して上記予測を行う必要があり、きめ細かな制御を行うには短い間隔でのサンプリング、すなわち大量の標本の取得が必要である。そのため、これらの膨大なデータを有意な時間内に処理するための高速な処理系が必要となる。

以下では、まずはじめに我々の採用した負荷予測手法について述べ、つぎにDynamicSVD[2][3]を利用することで、これらの手法が実システムに適用可能であることを示す。

† 三菱電機株式会社 情報技術総合研究所

247-8501 神奈川県鎌倉市大船 5-1-1

Tel: 0467-41-2525

Information Technology R&amp;D Center, Mitsubishi Electric Co.

5-1-1, Ofuna, Kamakura, Kanagawa, 247-8501, JAPAN

Tel: +81-467-41-2525

‡ E-mail: Hirata.Takahisa@bp.MitsubishiElectric.co.jp

### 3 提案システム

#### 3.1 負荷予測手法

はじめに短期的な負荷予測手法について述べる。ここで我々が採用するのは、SVD (Singular Value Decomposition) を使用した時系列解析手法である[3]。この手法は短時間に区切った時系列データ群から、SVDによって主成分を抽出し、図 2に示されるように主成分空間上で定常状態と非常状態とを区別するものである。この手法の利点は、優れた異常識別能力を持つことに加え、運用に必要なパラメータが、定常状態で取得されたデータと定常状態とみなす範囲を定めるための係数との2点だけであるという簡便さである。

我々は、この手法を負荷予測に応用することで、定常状態からの逸脱、すなわち低負荷状態から高負荷状態への遷移を素早く検出することができることを実験により示した。

つぎに長期的な負荷予測手法について述べる。これについても、短期的な予測と同様なSVDによる解析手法を検討中である。短期的な予測においては、一つの対象の時系列データを短時間に区切って比較することで、その対象の定常状態と非常状態とを区別したが、長期的な予測においては1日あるいは1週間といった単位で対象毎に典型的なパターンを抽出した後、複数の対象間でパターンを比較し分離を行う。これは一般にはクラスタリング問題として定義され、短期的な予測における定常状態に相当するパターンを抽出することは容易ではない。しかし、ネットワークサービスにおいては、同一のサービスを大量に提供することが多く、また類似のサービスは類似の負荷パターンを示すと予想されるので、サービスの種類毎に定常的な負荷パターンを抽出し分離することは比較的容易なものと考えられる。

我々は実験環境において、上記仮定の妥当性を検証するとともに、本手法で負荷パターンの分離が可能であることを示すプロトタイプを設計中である。

#### 3.2 DynamicSVD

前述の二つの負荷予測手法に共通するのは、SVDを用いた解析という点である。さらに、いずれの手法も日々刻々と取得される観測値を追加しながら解析を行うという特徴が共通している。特に長期的な予測に関しては、長期にわたる観測値の量が大规模であるのに対して、日々追加される観測値の量は非常に少ないという特徴が強い。一般に、このような状況で解析結果を最新に保つためには、観測値全体に対して再計算を行う必要があるため、計算にかかるコストが高くなる。

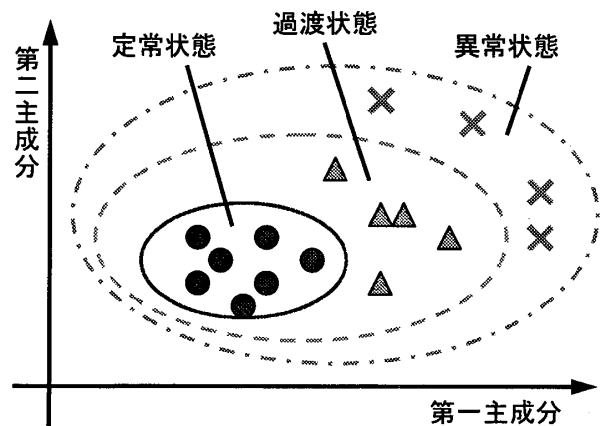


図 2 SVD による特徴抽出

このような逐次更新される行列に対して、効率的にSVDを計算するアルゴリズムに、我々が開発したDynamicSVDがある[2][3]。我々の過去の研究によれば、100行×10,000列の行列に対して、新たなデータを1列追加してSVDを再計算するのに必要な時間は、市販のPCを用いても1秒未満であった。ここで、行の数は1サンプルのデータ長、列の数は取得したサンプル数に相当する。例えば、1分毎に取得したデータを100分間隔で区切ったデータの過去10,000件分(1週間相当)の解析が1秒未満で終わることを意味する。これは、提案システムを実システムへ適用するのに十分な性能といえる。

#### 4 まとめ

本稿では、ネットワークサービスにおける負荷分散を目的とした負荷予測手法について述べた。予測すべき負荷変動として短期と長期との2種を想定し、両者に対してSVDを用いた予測手法を提案した。また、SVDの計算にDynamicSVDを用いることで、これらの手法を実システムに適用するのに十分な計算速度が得られることを示した。今後は引き続き提案システムの実装および性能評価を行う予定である。

#### 参考文献

- [1] "Gartner Says Service Providers Must Prepare Now for the Software as a Service Wave," Gartner Inc., March 2007.
- [2] M. Brand, "Incremental Singular Value Decomposition of Uncertain Data with Missing Values," European Conference on Computer Vision, May 2002.
- [3] R. Kashima, S. Fujii, et al., "An Intrusion Detection System Based on Network Stationary Monitoring," Proc. of the 69th National Convention of IPSJ, March 2007.