

300サーバ超の大規模システムにおけるキャパシティ管理実例

石井直子、名倉利幸

株式会社 アイ・アイ・エム

管理対象サーバ台数が300台を超えるユーザにおける、キャパシティ管理体制構築の事例をもとに、大規模システムのための評価・管理の手法と問題点について論じる。管理においては、管理用データ項目の絞り込みと適切なグラフの選定が重要である。評価においては、全体評価と詳細評価を分離し、全体評価はシステム全体の傾向把握と問題サーバの抽出のみを目的とし、詳細評価は全体評価で抽出されたサーバのみを対象にする。これにより、管理工数の削減と適切なキャパシティ管理の両立を図ることが重要である。

Case study: Capacity Management on the large system with more than 300 servers
Naoko ISHII, Toshiyuki NAGURA
IIM Corporation

There are specific problems involved in capacity management when you have a large number of servers. We will discuss problems and methods on how to solve these capacity management issues based on a specific client case. The system consists of more than 50 different services that contain a total of more than 300 servers.

In a daily management phase, it is important to select a small number of data entries to routinely collect and choose an appropriate graph format. In an evaluation phase, an entire system evaluation must be done for two purposes: to understand the trend in the system, and to select servers to be evaluated in detail.

With this methodology, we can reduce the number of man-hours spent and we can implement efficient capacity management.

1 はじめに

以前は大型のメインフレームで処理されていた重要業務は、複数のサーバにまたがって処理されるようになった。それにともない、各企業で管理するサーバ台数は増加の一途をたどっている。比較的安価なサーバで重要業務を処理できるようになった一方、台数の増加に伴う管理工数の増大はシステム担当者の悩みの種となっている。

効率がよく安定したシステムの構築にはキャパシティ管理、すなわち、システム全体の稼働状況と傾向の把握が必要である。しかし、現実には、サーバで処理される業務が次々と追加され、システム担当者は日々の業務が無事に処理されることのみ注力せざるを得なくなっている。システム担当者が忙しいために全体像が把握できず、とりあえず新規業務のためのサーバを次々に追加してゆき、それがさらに担当者の

管理工数を増やしてしまうという悪循環に陥っているケースも少なくない。その結果、だれもシステムの全体像を把握しないまま日々の業務が稼働しているのである。

本稿では、300台を超えるサーバを管理するユーザでの、キャパシティ管理体制の構築事例をもとに、大規模システムにおけるキャパシティ管理の手法と問題点について論じる。

2 システムの概要

2.1 環境

本稿で事例とした大規模システムは、業務種別で50種類以上、パフォーマンス管理対象（パフォーマンス情報取得対象）は約400サーバある。現在もキャパシティ管理体制の構築を継続中であり現時点で実際に管理対照となっているのは約300サーバである。各サーバで使用されるOSはUnix、Windows、Linuxとさまざまである。また、単

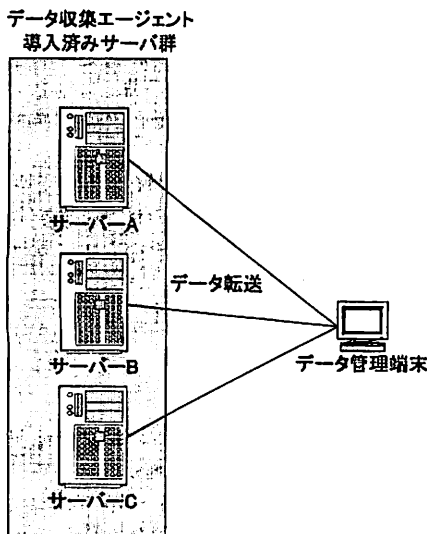
一業務システム内で異なるOSを使用しているケース（UnixとWindowsが混在するなど）もある。各業務における種別（社内／社外向け）、ピーク日、ピーク時間帯、オンライン／バッチ時間帯などの業務特性はすべて異なり、いずれもミッションクリティカルな業務ばかりである。

2.2 キャパシティ管理データ収集体制

キャパシティ管理において、必要なデータが欠落することなく継続的に収集されていることは必須の前提条件である。データの欠落は、問題発生時の原因調査を困難にし、また、長期的な傾向を見る際にも平均や最大の値に影響を与える。そのため、本事例では各サーバにデータ収集エージェントを導入し、収集されたデータをデータ管理端末が定期的に転送・蓄積する仕組みを使用している。（図1）各サーバからデータ管理端末へのデータ転送はネットワーク経由で行われる。実際には、図1におけるサーバは9群に分割され、9台のデータ管理端末がデータを蓄積している。

収集エージェントは、UNIX系OSでは、sarやvmstatなどの標準コマンド、Windows系OSではパフォーマンスカウンタの値を収集する。OS標準のデータを使用することにより、OSベンダを含むキャパシティ管理関係者がすべて同じ基準でシステム評価や稼働管理を行うことができる。

図 1



3 目的

3.1 全体把握（見える化）

キャパシティ管理サービスの導入に当たって、第一に目指したのは、システム全体の把握、いわば「見える化」である。どのサーバに余裕があり、どのサーバが増強を必要としているか、システム全体でバランスがとれているかなど、現状の把握がキャパシティ管理の基本となる。また、各種データを長期的に蓄積することにより、キャパシティ計画に活用することができる。

3.2 キャパシティの最適化

システムの見える化が実現されると、サーバの整理・統合の基礎資料ができる。これにしたがって、サーバの整理・統合、仮想化によるサーバの集中化を行い、システム管理工数・コストの削減が二番目の目的である。

3.3 稼働管理(定常的)

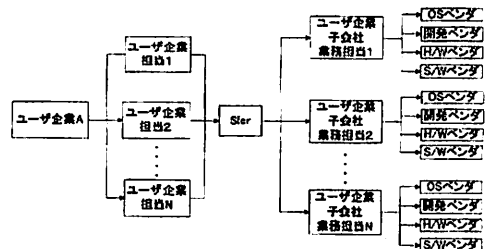
システム全体の見える化とともに、ループなどの異常な動きをしているサーバはないか、バッチは予定終了時刻までに終了しているかなど、各サーバ、業務システムの日次レベルでの定常的な稼働管理も重要である。このとき、各業務システムで共通の管理指標を用い、誰が見ても同じ管理をできるようにすることが三番目の目的である。

4 従来のキャパシティ管理手法と問題点

4.1 問題点

従来、本システムでは業務毎の縦割りによる管理が行われていた。その背景として、大規模システムにおいては体制が複雑化しやすいということがあげられる。本事例における体制は図2に示すとおりである。

図 2



ユーザ企業内の各担当がSIerにシステムの開発・保守運用を依頼し、実際にはユーザ企業のシ

システム会社がその実務を取りまとめている。さらにシステム会社の各担当は、ハードウェアやソフトウェアのベンダと連絡を取り合いながら、日々の運用を行っている。このため、運用体制は縦割りとならざるを得ず、各業務システム間との情報共有を行うことは困難であることがわかる。

その結果、業務システムの評価、チューニングなどはすべて業務システム内で完結しており、システム全体を見渡せる情報を一切持つことができなかつた。そのため、サーバ統合などを実施することが事実上不可能であった。

システム全体を把握できない背景には、他に、複数種類のOSを一元管理するための統一的指標がないということがある。たとえば、プロセッサ使用率ひとつをとってみても、各サーバの処理能力は異なっており、全サーバを比較して、どのサーバで稼動している業務がどの程度プロセッサ能力を必要としているのかを発見することは困難である。そのため、サーバ統合を検討する場合、どのサーバの業務をどのくらいの処理能力を持つサーバに統合したらよいかということを判断できない。

サーバ台数が多いことにより、管理工数はサーバ台数の増加とともに増大する一方である。多くの現場の管理者は、業務を平常稼動させることに精一杯で、パフォーマンス管理・キャパシティ管理にまで手が回らないというのが現状である。また、そのために全体が把握できず、新規業務を既存サーバに追加することはほとんどなく、新規に増設することで対応してきた。そのため、サーバ台数の増加を防ぐ手立てがなかった。IT技術の進歩により、仮想化などのサーバ統合、管理工数の削減のための技術も次々と発表されてはいるが、現場の技術者は、新技術を習得するための時間すら確保できないというのが現状である。

前述の問題点を踏まえ、どのような点が課題として挙げられるだろうか。

4.2 指標の選定

どのような項目を管理指標として使用するかということを決めることがパフォーマンス管理の第一歩となる。本事例のように、対象サーバ数が多い場合、必要十分でありながら少ない管理指標に絞り込むことが重要である。

指標の選定に当たって重要となるのが、何を、どこまで詳細に見るのか、ということである。

本システムにおいては、最重要、重要、その他のサーバの3種類に全サーバを分類し、最重要、重要サーバについてのみ、パフォーマンスデータの収集を行うことにした。このような分類を行うことによって、運用管理者がどのサーバ、業務システムを最初に確認し、対応を行っていく必要があるのかを明確にすることができる。最重要・重要サーバの選定は、各業務担当者にサーバの分類をヒアリングして決定する。この時、あわせて各サーバの重要時間帯の情報も入手する。そこで直面するのが、業務担当者の「重要」とキャパシティ管理上の「重要」が指すサーバが必ずしも一致しないという問題である。業務担当者にとっての「重要」とは、障害などによる停止が許されない、クリティカルなサーバを指すが、一方キャパシティ管理の観点に基づく「重要」は近い将来問題を起こしそうな、注意すべきサーバを指している。

今回の事例においては、業務担当者によって「最重要」と分類されたサーバのプロセッサ使用率が5%前後で推移している、あるいは「ピーク時間帯」は一日24時間継続しているといった回答をもとに稼動状況を監視していると、実際は夜間にほとんど動いていなかったというようなことがしばしば起きている。これには、1)業務担当者がピーク時間帯を把握していない、2)サーバがオーバースペックになっている、などの背景がある。このような「キャパシティ管理上、最重要サーバとは言えない」サーバ群をパフォーマンス管理開始当初から把握することは難しいが、定期的なシステムの評価を行っていくうちに、「重要な(注意すべき)サーバ」を選定することができる。

これら最重要・重要システムの中から、実際の稼動状況を元に問題が発生していると思われるサーバの発見、さらに分析が必要なサーバの選定を行う。

多数のサーバを一元的に管理していくに当たって、重要サーバの選定とともに重要なのが、管理指標を見るとき単位である。本事例の目的を達成するには、詳細なデータは必要ない。データ収集時は15分間隔でデータ収集を行うが、実際にデータを分析する際には収集された詳細データを日単位でサマリー化(1日の平均値・最大値の算出)を行い、さらにそれを1ヶ月間で俯瞰することで、全体の傾向を把握しやすくなる。各サーバ間の比較には、さらにこれをサマリー化し、1ヶ月間の平均値と最大値を使用する。データをサマリー化することで、データ処理も容易になり、

また、時折発生する異常値に左右されない評価を行うことができる。

このように、データのサマリー化を行っていくため、バッチ時間帯とオンライン時間帯、ピーク・非ピーク時間帯の把握は重要である。これらの時間帯を区別せずにサマリー化を行えば、システムの実態に即したキャパシティ管理を行うことはできない。

キャパシティ管理において重要なことは、リアルタイム監視に行うような、閾値に基づいた判断は必ずしも有効ではないということである。現状を把握しないまま、閾値を決めることはできず、「とりあえず」閾値を決めたとしても、その閾値そのものに意味はない。キャパシティ管理における第一の目標はシステムの現状・傾向把握をすることであり、ある一時点での使用率を詳細に見ていくことは必要ない。

5 管理

5.1 管理項目

パフォーマンス管理において、収集可能なデータ項目は、システム資源、OracleやSQLServerなどのDB、ミドルウェア、ネットワークなどさまざまな項目がある。しかし、これら取得可能なデータをすべて扱うことは、管理対象サーバの台数から考えて、現実的ではない。そこで、本システムでは、当面はシステム資源関連情報のみを取得している。(表1)

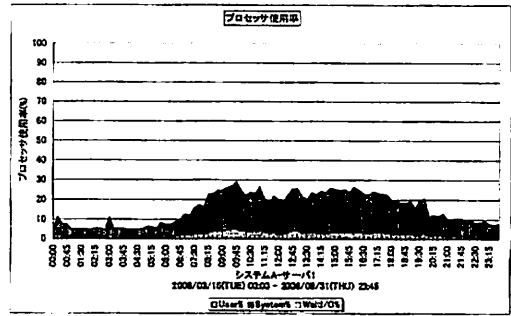
表 1

資源	項目名
プロセッサ	プロセッサ使用率
メモリ	フリーメモリ量
メモリ	ページング回数
ディスク	レスポンス時間
ディスク	ビジー率
ディスク	空き容量

上記6種類のデータ項目を元に、日次、週次、月次の各レベルでグラフを作成し、稼働管理を行っている。これらのグラフは、ウェブで公開され、各業務システム担当者など、システム関係者が必要なときに確認できるようになっている。

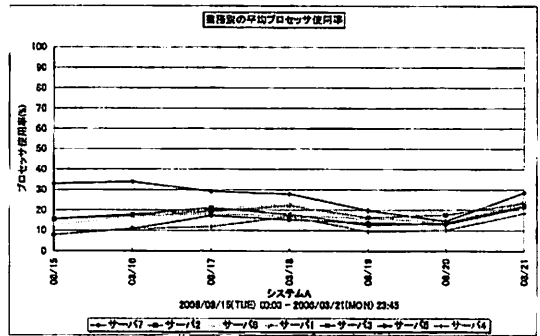
日次レベルではもっとも詳細なデータ(15分インターバル)で各項目について各サーバの一日の動きをグラフで確認できる。(図3)

図 3



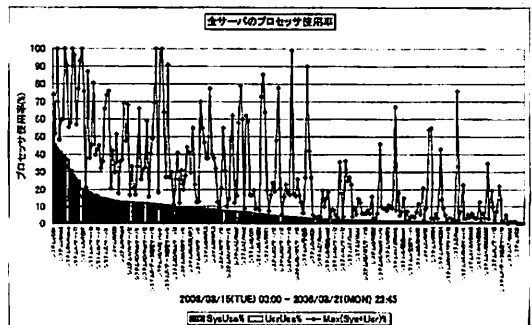
週次レベルでは、業務システム内のサーバ毎・1週間の動きを1日ごとにサマリー化されたグラフで確認できる。(図4)

図 4



また、サーバ単位だけではなく、各業務システムのレイヤー単位で、全業務システムを一覧できるグラフを作成し、業務システム全体の稼働状況を確認している。(図5)

図 5



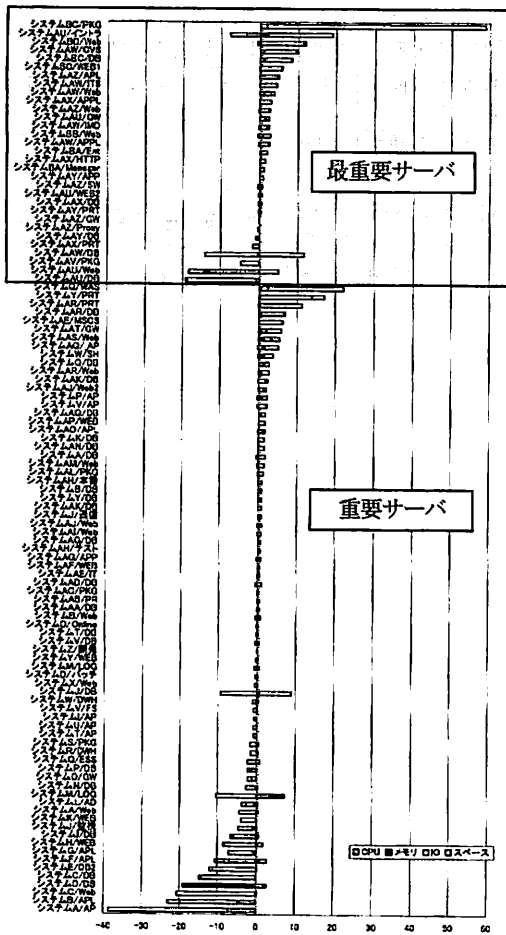
月次レベルでも、週次と同様のグラフを1か月分のデータで作成・確認している。

6 評価

6.1 全体評価

現在、本システムでは毎月1回、月次評価を行っている。この際、最も重要となるのは、300以上あるサーバ群の中から、いかに問題のあるシステムやサーバを発見するかということである。本システムはすでに本番運用で使用されていることから、前提条件として「システムのほとんどのサーバは問題なく運用されている」という想定のもと、「先々月と比べて、先月の運用状況に変化があったかどうか」という観点で評価を行っている。そのときにもっとも有用なのが以下のグラフである。(図6)

図 6



このグラフは、先々月を基準として、先月の各資源の使用状況を示している。棒グラフが中心

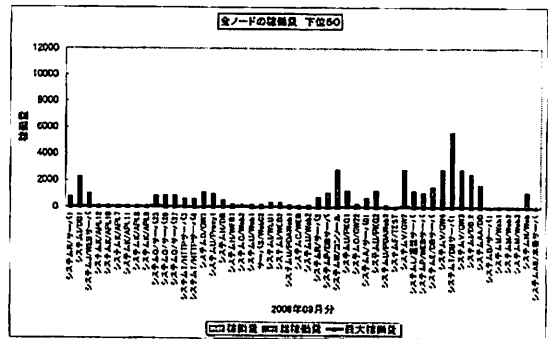
の縦線より右側にあるときは、先々月よりも資源使用率が増加しており、反対に左側にあるときは資源使用率が減少していることを示す。左端には業務システム名を表示する。最重要業務システム群が上に、重要業務システム群が下に表示され、各群で資源使用率の増加率が高いものから表示されるようにソートされている。

このグラフによって、使用率の変動が大きいものを容易に発見することができる。本グラフは、各資源の月間平均使用率を用いて作成され、具体的には変動幅(先々月と先月の差分)が10%を越える業務システム群が、問題がある可能性のある業務システムとして抽出される。

このようにして抽出された各業務システムについてのみ詳細な稼働状況を確認することで、月次レベルでの評価は完了となる。このような基準で抽出を行うと、詳細な稼働状況を確認する業務システムを数システムにまで削減することができるため、大幅な工数削減となる。詳細評価対象システムを絞り込んだ後は、一般的に使用される評価手法で各システムを評価する。

そのほか、別の観点として、プロセッサ使用状況を「稼働量」で見るグラフも有用である。(図7)

図 7



稼働量とは、サーバが持つプロセッサの処理能力に対して、どれだけの処理能力を使用したかを示す。この稼働量が総処理可能量に対して著しく低ければ、そのサーバはオーバースペックであるということができる。プロセッサ使用率を使用した場合、サーバ毎の負荷状態を知ることができるが、異なる処理能力を持つサーバ間の相対的な比較はできない。このグラフを使用して相対的に比較することにより、サーバ統合の対象候補の抽出と、統合時の組み合わせを決める際の参考資料と

して使用することができる。総処理可能量の算出には、SPECINT¹などのベンチマーク値を利用することができる。

プロセッサ使用率5%未満など、低稼働状態で特に問題がないサーバ群については、傾向把握を継続して行いながら、サーバ統合の検討を行うことになる。

6.2 個別業務評価

全体の評価で発見された問題を抱えると思われる業務システムや、レスポンス遅延などの問題が発生したことがわかっている業務システムに対しては、個別で詳細な評価を行う。

この個別評価に関しては、小規模のシステムにおける評価手法との違いは特にない。

7 データ収集上の問題点

すでに述べたように、キャパシティ管理において、データ収集が継続的に間断なく収集されていることは必須の前提条件である。しかし、現実にはさまざまな問題が存在する。

7.1 データ取得状況の把握

キャパシティ管理用データの取得状況の把握は予想以上に困難であった。理由としては、ネットワーク上の構成の複雑さから、データ収集管理サーバと各サーバ間のネットワーク接続の疎通確認には多くの工数がかかっている。また、疎通確認後も切戻・復旧を繰り返すことも多く、対象サーバ台数が多いため、工数が高くなっている。

7.2 OS特有の問題

前述のようなネットワーク環境によるデータ収集の停止のほか、Windows系OSではパフォーマンスカウンタが不定期に停止するといった現象がある。このため、一部のパフォーマンスデータが欠損することが増えてきている。このような事象を完全に回避するための対応方法は公開されておらず、安定的なデータ収集の障壁となっている。

¹ SPECINT:

Standard Performance Evaluation Corporation(SPEC)が提供する標準手順に基づいてハードウェアメーカが算出したプロセッサ処理能力のベンチマーク値。

<http://www.spec.org/>

7.3 ログインパスワードの定期的な変更

近年、セキュリティ意識の高まりから、定期的にパスワード変更を行うシステムも増えてきている。さらには、複数回の正しくないログインがあると、自動的にアカウントをロックするケースもある。当然、サーバ側のパスワード変更にあわせ、データ収集端末側の設定もサーバごとに変更する必要があり、頭の痛い問題である。

8 今後の取り組み

8.1 サーバの統合(仮想化)

本事例のキャパシティ管理体制の構築はまだ始まったばかりであるが、今後管理対象の全サーバのデータを蓄積し、すでに明らかになっているオーバースペックのサーバ群の統合に取り組む予定である。この際には、仮想化技術を使用し、処理能力の高いサーバで複数のOSを搭載して最適なサイジングを行うとともに、サーバ毎の管理工数の削減を目指す。

8.2 サイト/システムごとの閾値設定

今後データが蓄積され、全体の状況・各業務システムの状況が把握されれば、当システム固有の閾値を決定し、より容易な稼働管理、キャパシティ管理を行える体制を作りたい。

9 まとめ

データ収集やグラフ作成が自動化されていたとしても、最終的にこれらを利用して判断を下すのはシステム管理者である。これまで見てきたように、大規模システムにおけるキャパシティ管理において重要なことは、適切な管理項目の選択による管理工数の最小化である。これが実現して初めて、適切なサイジングやサーバの整理・統合、管理工数・コストの削減など、戦略的なキャパシティ管理が実現できる。サーバ台数が多いほど、管理体制の構築そのものに工数がかかり、7. で述べたような問題に対する完全な解決策はまだ見つかっていない。今後管理体制が確立されるまでの間に、現在残っている問題についても検討していきたい。