

Web アクセス集中対応 3 層データセンタ制御方式

増田 峰義[†], 垂井 俊明[†], 吉村 裕^{††}, 庄内 亨[†], 杉江 衛^{†††}

[†]日立製作所中央研究所 E-mail:{mine-m,tarui,shonai}@crl.hitachi.co.jp

^{††}日立製作所エンタープライズサーバ事業部 E-mail:yutaka.yoshimura@itg.hitachi.co.jp

^{†††}日立製作所システム開発研究所 E-mail:sugie@sdl.hitachi.co.jp

内容梗概

高品質な Web サービス提供には、アクセス集中時のシステム応答確保が必須である。VPDC (Virtual Private Data Center) では、複数顧客企業間で予備サーバ共有し、負荷に応じて予備サーバを動的に割当ることにより、低コストでアクセス集中に対応する。大規模 Web システムの標準的な構成である、3 層アーキテクチャにおける VPDC 詳細制御方式を提案した。プロトタイプにより、3 層アーキテクチャにおける負荷に応じたサーバ自動割当を実現し、アクセス集中が発生してから、Web 層/アプリケーション層は約 20 秒、データベース層は約 2 分の遅延で、システム応答時間を改善できることを確認した。

Dynamic Server Allocation in 3-tier Datacenter

Mineyoshi Masuda, Toshiaki Tarui, Yutaka Yoshimura, Toru Shonai, Mamoru Sugie

[†]Hitachi, Ltd., Central Research Laboratory E-mail:{mine-m,tarui,shonai}@crl.hitachi.co.jp

^{††} Hitachi, Ltd., Enterprise Server Division E-mail:yutaka.yoshimura@itg.hitachi.co.jp

^{†††}Hitachi, Ltd., Systems Development Laboratory E-mail:sugie@sdl.hitachi.co.jp

Abstract

This paper proposes VPDC (Virtual Private Data Center) mechanism in large scale 3-tier datacenter, a data center solution that responds to unexpected peaks in internet accesses, thus providing high-quality web services. VPDC has pools of servers shared among users, and dynamically changes the allocation of servers to users according to the access load, thus avoiding performance degradation and reduces the system cost. Evaluation of the VPDC prototype shows that servers in a 3-tier datacenter can be allocated automatically according to the access load. It is also shown that system response time can be improved approximately within 20 second after access peak in the case of Web/Application server bottleneck, and approximately within 2 minutes in the case of DB server bottleneck.

1. 序論

近年、インターネットの急激な普及に伴い、電子商取引等をサポートする Web システムの重要性が増大している。Web システムでは、スケーラブルなシステムを構成するために、Web サーバ、アプリケーション(AP)サーバ、データベース(DB)サーバの 3 種類のサーバを階層的に接続する 3 層アーキテクチャが広く用いられている²⁾。

Web アプリケーションを実行する企業情報システム、データセンタにおいては、24 時間 365 日連続した高品質なサービスを提供することが求められる。特に、急激なアクセス集中時にも、システムの応答を確保することは重要な課題である³⁾。不特定多数のユーザにサービスを提供するデータセンタでは、「サービス開始日」、「キャンペーン」、「TV 番組での URL 紹介」等により、通常の数倍から数百倍の負荷が集中する場合がある。アクセス集中に対して十分なサーバが提供できない場合、応答時間遅延によるサービスの悪化、さらに

は、システムダウンによるサービス停止を招く。このようなシステムの停滞・停止は、ビジネス機会の損失につながり、企業の社会的信頼を損なう。ここで、アクセス集中の発生頻度は低いため、アクセス集中に対して十分な数の予備サーバを用意しておくことは現実的でない。現用系の数倍規模の予備系を持たなければならず、コスト的に大幅に不利である。そこで、低コストでアクセス集中に耐えることのできるシステムが求められる。

複数の顧客企業を持つデータセンタで、低コストでアクセス集中に対応するために、負荷に応じてサーバを動的に割当てて方式が提案されている。データセンタの複数顧客企業間で予備サーバを共有し、アクセス集中時に相互に融通することを可能にする。それにより、顧客毎に予備サーバを持った場合に比べて、予備サーバのコストを大幅に削減する事が可能である。

このようなシステムとして、本報告では VPDC (Virtual Private Data Center)を提案する。VPDCは、Web/AP/DB 各層に複数企業で共有の予備サーバを置き、負荷に応じてサーバを追加/削減する。類似の研究として IBM 社の Oceano⁴⁾があるが、Oceanoでは、Web/APの二層のみをサポートするのに対し、VPDCでは、DB層も含めた3層をサポートする。コストの高いDBサーバを動的に追加・削減する機能を持たせることにより、より柔軟にアクセス集中に耐えるシステムを実現する。これにより、顧客企業はシステムの性能上の制限を気にすることなく、エンドユーザに対して、DB層への負荷が大きい1 to 1マーケティングのような魅力あるサービスを提供できる。

我々は、現在までに、Web層部分の先行プロトタイプ開発、評価を行なったが¹⁾。今回は、プロトタイプをWeb/AP/DBの3層まで拡張した。本報告では、この3層プロトタイプ上で行った、VPDCのアクセス集中対策効果を検証する実験結果について報告する。

2. データセンタシステムにおける課題

2.1 データセンタシステム構成

図1に3層アーキテクチャを持つデータセンタの構成を示す。大規模なWebシステムをサポートするデータセンタは、機能毎にサーバを分割し、スケーラブルなシステムを構成するために、HTMLを生成するWebサーバ、ビジネスロジックを実行するAPサーバ、データを保持するDBサーバの3層により構成される。各層は、複数のサーバによるクラスタ構成を取る。クラス

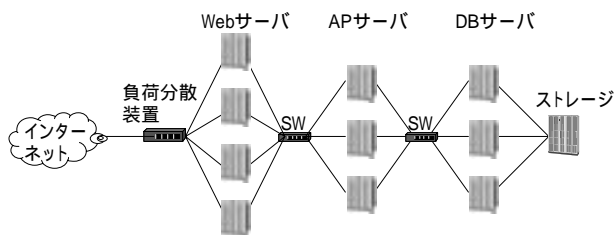


図1 3層データセンタ構成

Fig. 1 Structure of 3-tier Datacenter

タ構成により、負荷に応じたシステム拡張が容易になるとともに、サーバに不具合が発生した場合に他のサーバにより処理を継続するフェイルオーバー処理が可能になる。Web層のクラスタ化は、負荷分散装置により行なわれる。AP、DB層のクラスタ化は、AP、DBソフトウェアの持つクラスタ機能により実現する。

2.2 アクセス集中

Webサービスを提供するデータセンタでは、アクセス集中への対応が大きな課題である。Webアクセス集中の特徴は下記である。

(1) 予測困難

新規サイト開設時、キャンペーンを行なった場合等にどの程度の負荷が発生するかは予測困難である。

(2) ピーク負荷

携帯電話や常時接続環境の普及により、TV等によるURL放映時には、一瞬にして通常の数倍～数百倍の負荷が課せられる場合がある。

アクセス集中への対応を誤ると、応答時間の悪化によるサービス低下を招くだけでなく、過負荷によるシステムダウンにつながる場合があるため、アクセス集中への対応は避けて通れない課題である。

ここで、アクセス集中の解決のためにアクセスピークに十分なサーバを持つ方法が考えられる。しかし、アクセスピークはごく例外的に短期間だけ生じる場合がほとんどであるため、アクセスピークに合わせた設備増強を行なうと、稼働率の低い予備機器を多数抱えることになり、コストを考えると現実的な解ではない。そのため、低コストでアクセス集中に対応することのできるシステムが求められる。特に、Web/AP/DBの3層構成を持つデータセンタの場合、DBサーバの負荷は重く、高価なサーバが使用されるため、DBサーバのアクセス集中対策は重要である。

3. VPDC

3.1 基本概念

VPDCでは、データセンタがサポートする複数の顧客企業間で共有の予備サーバを持つ。各顧客企業に割当てられたサーバを、顧客企業の負荷に応じて割当て・削減することにより、アクセス集中に対応する。VPDC制御システムは、顧客企業に割当てられたサーバの稼働状

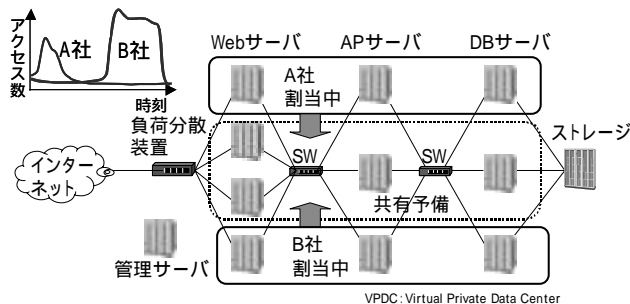


図2 VPDC構成

Fig.2 VPDC Architecture

況、ネットワーク等の資源使用状況を監視し、サーバの負荷が過大・過小と判定された場合には、該顧客企業の処理を行なうクラスタに割当てられたサーバの台数を追加・削減する。

図2にVPDCの構成を示す。Web/AP/DBの各層で負荷に応じたサーバの割り当て・削減を実現する。それにより、Web/AP/DBのどの層がネックになった場合にも柔軟に対応できるため、種々のアクセスパターンにおけるアクセス集中へ対応することができる。図2においてデータセンタ内のWeb/AP/DBサーバは、A社・B社の2つの顧客企業に割り当てられているサーバ、共有予備サーバに分類される。

VPDCでは、OceanicのAP/DB層サーバ動的割当に加え、DB層のサーバ動的割当を実現する。下記に示すDBサーバの性質より、DB層VPDCによりアクセス集中に低コストで効果的に対応するシステム実現することができる。

- Web/APサーバは比較的安価で多数台の予備を比較的容易に持つ事が出来るのに対し、DBサーバは高価で予備のコストが高い。
- 一般にDBサーバの負荷は重く、VPDCによるサーバ融通の効果が大きい。

以下でVPDCの全体制御方式について述べ、続く節でWeb/AP/DB各層での制御方式を個別に説明する。

3.2 全体制御方式

VPDCシステムの全体制御は、システムに1台の管理サーバと、各サーバ上で動作するVPDCエージェントで行う。管理サーバとVPDCエージェントは、図3に示すように連携して動作し、以下のステップでシステム全体を制御する。各ステップの詳細を表に示す。

- (1) 検出： 負荷をモニタリングし、高負荷を検出
- (2) 判断： サーバ割り当てを判断
- (3) 指示： 割り当て処理の指示
- (4) 実施： 割り当て処理の実行

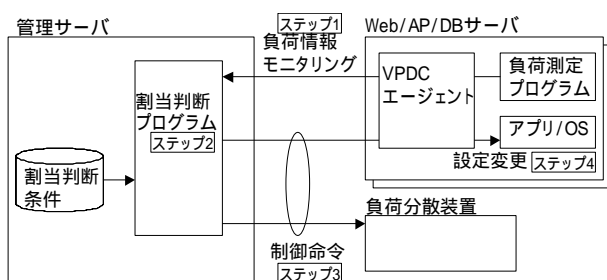


図3 VPDC制御方式
Fig.3 VPDC Control Mechanism

次節以降では、上記(4)の項目、つまりサーバ割り当て処理の詳細について述べる。まず、Web/AP/DB各層で共通な方式である、“安全にサーバを削減する方式”について述べる。その後、Web/AP/DB各層固有な割り当て処理の詳細について述べる。

ステップ	処理内容
1.検出	管理サーバは、各顧客企業に割当てられたサーバの負荷を表わす CPU 稼働率・ネットワーク資源使用状況等のメトリックを各サーバに置かれたエージェントを通じて定期的にモニタリングする。
2.判断	管理サーバは、各サーバの負荷とあらかじめ定められたサーバ判断条件を比較し、サーバの割当判断を行なう。具体的には、負荷が高い場合には該当する顧客企業に予備サーバからサーバを追加し、負荷が低い場合には該当する顧客企業の処理を行なっているサーバを1台削減し予備サーバとする。
3.指示	管理サーバは、割当判断結果を各サーバに置かれたエージェントに通知すると同時に、負荷分散装置に対し、負荷分散の設定を変更するための制御命令を送出する。
4.実施	各サーバのエージェントは、管理サーバからの指示に従い、Web/AP/DB アプリケーションソフトウェア・OS の設定を変更し、該当する顧客企業の処理を行なうサーバ台数を変更する。

3.3 安全なサーバ削減方式

各層VPDC制御での課題は、システムを停止させることなくサーバを追加/削減する技術の開発である。ここでポイントとなるのは、“安全に(Safe)サーバを削減する方式”である。

Web層を例にとり、安全なサーバ削減方式について説明する。稼働中のWebサーバには、ユーザ・リクエストを処理中のプロセスが存在する。このサーバを単純に削減すると、処理の途中でプロセスが終了してしまい、ユーザにはエラーが返る。これは安全(Safe)ではない。

安全にWebサーバを削減するためには、削減によってデータが喪失しない状態、つまりWeb層ではサーバ内に処理途中のプロセスが存在しない状態で行う必要がある。

VPDCでは、以下の手順で安全なサーバ削減を実現する。この手順は、Web/AP/DB各層で共通である。

- 1) 前層の負荷分散機能に対し指示を出し、削減対象サーバへのリクエスト要求を停止させる。
- 2) 当該サーバを削減してもデータが失われない状態にする。Web層の例では、タイムアウトを待つ、プロセスを別のサーバへ移送するなどの実装が考えられる。
- 3) 当該サーバを削減する。

3.4 Web層制御方式

図4にWeb層VPDC制御方式の概要を示す。Webサーバへの負荷分散には負荷分散装置を用いる。負荷分散装置は、各顧客企業毎に一つの代表IPアドレスを持つ。クライアントより、代表IPアドレスに送られてきたWebアクセスは、負荷分散装置により各顧客企業に割当てられた複数のサーバにラウンドロビン等により分配される。それにより、クライアントからは複数のWebサーバを意識せずにアクセスを行なう事ができる。

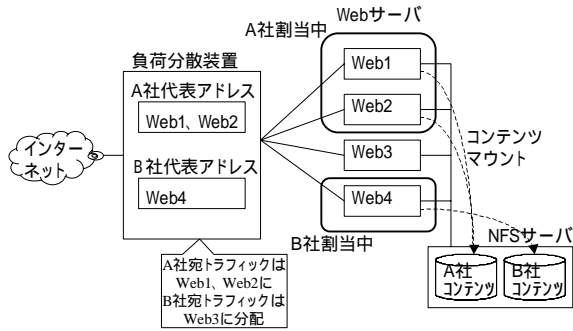


図4 Web層VPDC制御方式

Fig.4 VPDC Mechanism (Web Server)

Webサーバの追加/削減の処理は、以下の手順で行われる。

ステップ	処理内容(追加)	実行者
1:	顧客企業に新しく割当てるサーバを、予備サーバから一台を選択する。	管理サーバ
2:	上記サーバへ顧客企業のコンテンツを配備する。コンテンツの配備方法には、NFS等によるマウント方式、コンテンツをWebサーバへコピーする方式などが考えられる。	Webエージェント
3:	Webサーバプログラムを起動する。	Webエージェント
4:	管理サーバが負荷分散装置へSNMP命令を送り、該当するWebサーバを負荷分散対象に加える。	管理サーバ
ステップ	処理内容(削減)	実行者
1:	削減対象のWebサーバを選択する。	管理サーバ
2:	負荷分散装置の負荷分散対象から除外する。	負荷分散装置
3:	Webサーバプログラムを停止する。	Webエージェント
4:	コンテンツの割当を外す。	Webエージェント
5:	該当するサーバを予備サーバに変更する。Web層のより詳細な処理方式は参考文献 ¹⁾ に記す。	Webエージェント

3.5 AP層制御方式

3.5.1 セッション引き継ぎ機能

AP層の基本的な制御方式はWeb層と同じであるが、AP層で安全にサーバを削減するためには、“セッション”を考慮に入れる必要がある。一般に、APサーバはセッション維持機能を持つ。セッション維持機能とは、同一エンドユーザのリクエストを特定のAPサーバに割当てる機能である。セッション維持機能は、電子商取引の「バスケット」等の、エンドユーザ毎の情報保持に使われる。エンドユーザが買い物中にバスケットに入れた商品、等のエンドユーザの状態は、特定のAPサーバ上に、セッション情報として記憶される。

APサーバ削減時に、セッションを考慮せず単純にAPサーバを削除すると、当該APサーバが保持していたセ

ッション情報が失われる。セッション情報の喪失は、買い物中にバスケットに入れた商品が失われる、といった不具合を招く。

したがって、VPDC制御によりAPサーバが削減された場合にセッション情報が失われることを避けるために、当該APサーバのセッション情報を、別のAPサーバが引き継ぐ機能が必要となる。

セッション引き継ぎ機能の実装例として、APサーバ間でセッション情報を共有することが挙げられる。各APサーバは、セッション情報を自サーバ内だけでなく共有機能にも書込む。通常APサーバはセッション情報を自サーバ内で検索する。VPDC制御によりAPサーバが削減され、今までとは異なるAPサーバにセッションが割当てられたために、自サーバ内にセッション情報がみつからない場合には、セッション情報は共有機能から読み出される。これにより、顧客企業に割当てられたサーバが削減された場合にも、別のAPサーバが処理を引き継ぐことができる。

3.5.2 制御手順

図5にAP層のVPDC制御方式を示す。ここでは、APサーバプログラムとしてTomcat⁶⁾**を採用しているAPサーバの追加/削減をWeb層の負荷分散機能に反映させるため、Webサーバに置かれたAPサーバへの負荷分散プラグインプログラムの負荷分散設定ファイルを動的に変更する機能を追加する。

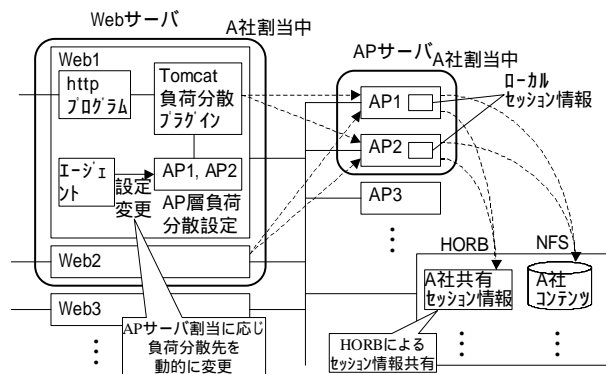


図5 AP層VPDC制御方式

Fig.5 VPDC Mechanism (Application Server)

セッション引き継ぎ機能の実現のために、セッション情報の共有機能を置くサーバを設け、HORBB⁷⁾によりセッション情報を共有する。各APサーバは、自サーバ上にセッション情報が記憶されていない場合には、HORBB経由で共有機能上のセッション情報を読み出す。本プロトタイプでは、上記のセッション管理機能をJavaクラスとして提供する。

APサーバの追加/削減は下記の手順で行なわれる。

* NFS,Java は、米国 Sun Microsystems の登録商標である。

** tomcat は、Apache Software Foundation の登録商標である。

ステップ	処理内容(追加)	実行者
1:	顧客企業に新しく割当てるサーバを、予備サーバから一台を選択する。	管理サーバ
2:	上記サーバに、追加する顧客企業のコンテンツを NFS 等によりマウントし、コンテンツを配備する。	AP エージェント
3:	AP サーバプログラムを起動する。	AP エージェント
4:	顧客企業に割当てられている全 Web サーバの負荷分散プラグインの設定を変更し、該当する AP サーバを、負荷分散対象に加える。	Web エージェント
ステップ	処理内容(削減)	実行者
1:	削減対象の AP サーバを選択する。	管理サーバ
2:	Web サーバ負荷分散プラグインの負荷分散対象から除外する。	Web エージェント
3:	AP サーバプログラムを停止する。	AP エージェント
4:	コンテンツの割当を外す。	AP エージェント
5:	該当するサーバを予備サーバに変更する。	AP エージェント

3.6 DB 層制御方式

3.6.1 Shared 型と Shared Nothing 型

DB サーバをクラスタ化する構築方式は、大きく二つの方式に分かれる。一つは、複数の DB サーバで一つのストレージを共有し、複数の DB サーバが並列に処理を行う Shared Disk 型、もう一つは、各 DB サーバがデータの一部分を持ち、分担して処理を行なう Shared Nothing 型である。それぞれの方式での VPDC 制御を以下に示す。

(A) Shared Disk 型

Shared Disk 型 DB でのサーバの追加/削減には、DB サーバのフェイルオーバー機能を利用する。ストレージを共有しているため、DB サーバの追加/削減は Shared Nothing 型に比べ比較的容易である。ただし、ストレージがボトルネックになっている場合、Shared Disk 型では対応ができない。

(B) Shared Nothing 型

Shared Nothing 型の DB では、サーバ毎に異なる範囲のデータを処理する。そのため、サーバを追加/削減すると同時に、データの再配置を行う必要がある。この方式は、ストレージネックを解消することができるが、データの配置を適切に行わないと、DB サーバの負荷に偏りが発生するという欠点がある。

VPDC プロトタイプでは、DB 層での負荷に応じたサーバ動的割当を実現し、アクセス集中対策効果を検証することを目的としている。そこで、本プロトタイプでは構築が容易な(A)の Shared Disk 型 DB である Oracle Parallel Server(OPS) * を使用して実装した。

* Oracle Parallel Server(OPS)は、米国 Oracle 社の商標登録である。

3.5.2 制御手順

各 DB サーバには VPDC 層制御のためのエージェントが置かれる。エージェントは、OPS に対して、クラスタ構成変更コマンドを送出する。OPS では、AP 層制御で行われていたような、前段におかれたサーバの負荷分散プログラムへの構成変更指示は不要である。その理由は、AP サーバから DB サーバへの負荷分散が、物理サーバ名ではなく、各顧客企業の DB サービス名で行われるためである⁹⁾。それにより、DB サーバ割当が変更になった場合でも、AP サーバ側の構成を変更する必要は無い。

DB サーバの追加は下記の手順で行なわれる。

ステップ	処理内容(追加)	実行者
1:	顧客企業に新しく割当てるサーバを、予備サーバから一台選択する。	管理サーバ
2:	上記サーバ上で、顧客企業の DB サービスを開始させる。	DB エージェント
ステップ	処理内容(削減)	実行者
1:	削減対象の DB サーバを選択する。	管理サーバ
2:	上記サーバ上で、顧客企業の DB サービスを停止させる。	DB エージェント
3:	該当するサーバを予備サーバに変更する。	DB エージェント

4. システム評価

4.1 目的

3 層アーキテクチャを持つデータセンタにおける DB 層を含めたアクセス集中対策効果を検証するために、Web/AP/DB 各層における負荷に応じたサーバ動的割当を実現する VPDC プロトタイプを試作・評価した。

評価では、システムの負荷を急激に増加・減少させ、下記を確認する。

- Web/AP/DB 層の各層のサーバが負荷に応じて割当/削減されるか。
- アクセス集中時のサーバ動的割当により、エンドユーザから見た応答時間が改善されるか。

以下では、実験システム、評価に使用した負荷について述べた後、評価結果について述べる。

4.2 実験システム

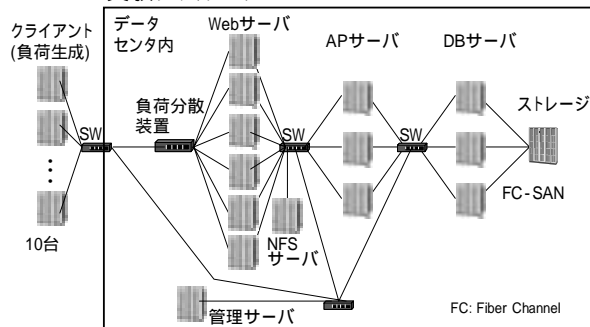


図6 実験システム構成
Fig.6 VPDC Prototype

図6に実験システムの構成を示す。表1に各装置の諸元を示す。なお、ここで使用した製品、装置名は、各社の登録商標である。

負荷分散装置、Webサーバ6台、APサーバ3台、DBサーバ3台、RAIDがネットワークで接続されている。WebコンテンツはNFSサーバに置かれる。AP層のセッション情報は、APサーバの1台におかれたHORBサーバプログラムにより共有される。サーバ割当て判断・割当て指示等のVPDC制御は、管理サーバで行われる。Web、AP、DBサーバには、CPU使用率等のリソース使用量モニタリング、サーバ割当て・削減処理を行なうためのVPDCエージェントが置かれる。エージェントは管理サーバ上のマネージャとSNMP⁽¹⁰⁾により通信する。負荷分散装置の設定変更もSNMPで行われる。

表1 実験システム諸元
Table 1 Prototype Specification

種別	ハードウェア	ソフトウェア
Webサーバ	PC/AT, Pentium III 750 ~ 800MHz	RedHat 6.2 (Kernel 2.2.14), Apache 1.3.12
NFSサーバ	PC/AT, Pentium III 750 ~ 800MHz	RedHat 6.2 (Kernel 2.2.14)
APサーバ	PC/AT, Xeon 550 ~ 700MHz	RedHat 6.2 (Kernel 2.2.14), Tomcat 3.2.1
DBサーバ	IBM eServer pseries 640 B80, Power3-II 375MHz	AIX 4.3, Oracle 8i enterprise edition 8.1.6, OPS
管理サーバ	PC/AT Pentium III 930MHz x 2	RedHat 6.2 (Kernel 2.2.14)
ストレージ	日立 RAID A-6573/A-6576	-
負荷分散装置	Alteon 180E	-
クライアント	PC/AT, Pentium III 1GHz	RedHat 6.2 (Kernel 2.2.14)

4.3 サーバ割当て・削減条件

VPDCシステム実用化のためには、顧客企業との間で取り交わしたSLA(Service Level Agreement)を実現するために、

- ・各層のサーバ割当て・削減の条件
- ・複数の顧客企業間でのサーバ割当て優先順位等の課題を解決し、サイトの性質に応じた最適なサーバ追加・削減ポリシーを決定しなければならない。しかし、本報告においては、VPDC基本機能であるWeb/AP/DBの3階層システムにおけるサーバ動的割当て・削減の検証を目的とする。詳細なサーバ割当てポリシーは将来課題とし、実験では下記に示す基本的なサーバ割当てアルゴリズムのみを使用する。

システムの応答時間の低減に向け、管理サーバは、下記の2条件よりサーバ割当て・削減を決定する。

(1) システム応答時間

管理サーバより、各顧客企業に割当てられたWebサーバ・APサーバを通りDBサーバをアクセスするダミートランザクションを定期的に出す事により測定する。

(2) サーバ平均CPU稼働率

各顧客企業に割当てられたWeb/AP/DBサーバのCPU稼働率をSNMPエージェントにより測定し、顧客企業毎に平均する。

(1)の値が応答時間の限界値より大きい場合には、システムが過負荷と判断する。その場合、Web/AP/DB各層毎に、CPU稼働率の平均値が定められた限界値を超えていないか調べられる。限界値を超えている層のサーバが性能ネックになっていると判断し、該当する層にサーバを追加する。その後負荷が減少し、サーバの平

均稼働率が定められた値を下回ると、該当する層のサーバを1台開放する。表2に各層のサーバ割当て・削減条件をまとめる。管理サーバは5秒毎にサーバ割当て判断を行なう。

表2 サーバ割当て・削減条件
Table 2 Server Allocation Policy

種別	サーバ追加	サーバ削減
Web	応答時間1.5秒, CPU使用率30%	CPU使用率 30%
AP	応答時間1.5秒, CPU使用率85%	CPU使用率 85%
DB	応答時間1.5秒, CPU使用率60%	CPU使用率 60%

4.4 テストプログラム

テストでは書籍の電子商取引サイトを模擬するTPC-W⁽¹¹⁾をベースとしたベンチマークプログラムを使用した。TPC-Wの各ページ(管理ページを除く)を独立したJSPプログラムでコーディングした。

テストでは図7で示すTPC-Wページのシーケンスを各クライアントプロセスから繰り返しアクセスさせる。アクセス集中時にサーバの動的割当てが正しく行われることを確認するために、低負荷・高負荷・低負荷のシーケンスで負荷を課した。クライアント側のプロセス数を変化させる事により負荷量を調節する。低負荷時には25クライアントプロセス、高負荷時には205クライアントプロセスからアクセスを発行させる。

本報告では、DB層サーバ動的割当て・削減の効果を検証することに主眼を置いた。そのため、DBネックとなる状況でアクセスを行なった。

Home	ShoppingCart	SearchResults	SearchResults
ShoppingCart	CustomerRegistration	ProductDetail	ProductDetail
ShoppingCart	BuyRequest	ShoppingCart	Home
Home	CustomerRegistration	CustomerRegistration	ProductDetail
NewProducts	SearchRequest	Home	ShoppingCart
ProductDetail	SearchResults	Home	CustomerRegistration
SearchRequest	SearchRequest	NewProducts	BuyRequest
Home	SearchResultsc	Home	BuyConfirm
ProductDetail	Home	Home	
SearchRequest	SearchRequest	SearchRequest	

図7 クライアントアクセスパターン

Fig.7 Client Access Pattern

4.5 テスト結果

図8にクライアントから課した負荷量(相対値)、クライアントから測定されたシステム応答時間、Web/AP/DB各層に割当てられたサーバの稼働率の時間変化を示す。図8では、実際に割当てられたサーバのみ稼働率のグラフを示す。アクセス集中によりシステム負荷が低負荷から高負荷になると、システム応答時間が増大し、管理サーバがサーバ追加を判断する。その結果、Web/AP/DB各層のサーバが3台まで順次追加され、応答時間が改善される。

システムのサーバ追加・応答時間は2段階で改善される。本実験で課した負荷はDBネックであったため、DBサーバ動的割当てのシステムの応答時間短縮への寄与が大きい。

クライアントからの負荷が増大すると、まず、DBサーバのCPU稼働率が限界値を超えるため、DBサーバが追加される。DBサーバが追加される前の段階では、DBネックのため、Web/AP層のサーバの負荷は低く、サーバ追加条件に達しない。DBサーバが追加された結果、前段のWeb/APサーバにも高負荷が課せられるように

なるため、Web/AP サーバも追加される。
DB サーバが追加されてから、DB サーバ追加の効果が

あらわれ、システムの応答時間が改善されるまでには
2分程度の待ち時間がある。その理由は、DB サーバの

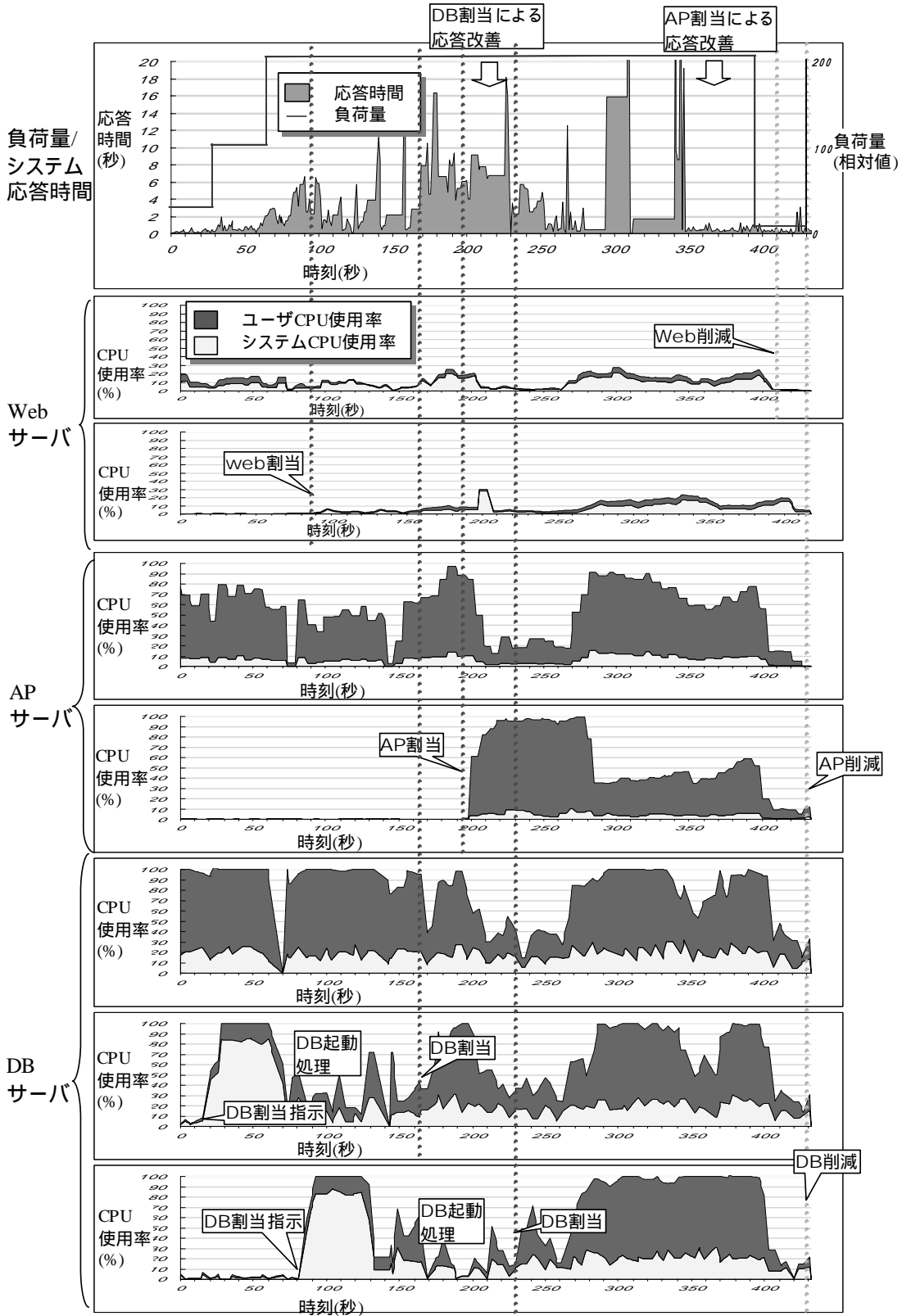


図 8 負荷量、応答時間、各サーバ稼働率の関係

Fig.8 Relationship between Workload, Response Time and Server Utilization

起動時間が2分程度必要であり、追加されたサーバ上でDBサーバのサービスを起動してから実際にアクセスを受け付けられるようになるまでには遅延があるためである。それに対して、Web/APサーバでは、プログラムを起動すると即座に処理を受け付けることができるため、サーバ追加直後に応答時間が改善される。

各層のサーバの負荷が増大してから、管理サーバで判断を行い、サーバが追加されるまでの待ち時間は20秒である。プログラムの起動時間の大きいDB層以外では、負荷増大から20秒程度で処理能力を増強することができる。

DB層ではDBサーバの起動時間のために、応答時間が改善されるまでには、2分程度待たなければならない。そのためDB層は、1分以内に負荷が急増するような急激なアクセス集中に対応することができない。しかし、上記のような急激なアクセス集中が生じた場合にも、下記の理由により、DB層の立ち上がりの遅延をカバーすることができる。

(1) アクセス集中の影響を直接受けるフロント層(Web/AP層)では、VPDC制御の効果は20秒程度で現れる。

(2) DB層のみがbusyの場合、Web/AP層で「ただいま混みっております」等のメッセージを出す等の手段により、エンドユーザのアクセスをある程度待たせることができる。

さらに、必ずしも全てのアクセス集中において、2分以内の高速な応答が必要になる事は無い。したがって、本報告で提案したVPDC制御により、DB層まで含めたアクセス集中対策が実現できる。

以上のように、3層データセンタにおけるVPDC制御のためには、

- ・各階層のプログラムの起動時間
- ・Web/AP/DB各層の応答時間に対する寄与

を考慮したVPDC制御ポリシーが必須であり、今後の検討課題である。

5. 結論

Web/AP/DBの3層アーキテクチャを持つデータセンタにおいてアクセス集中を実現するために、負荷に応じて自動的にサーバを割当・削減するVPDCを提案、プロトタイプを試作し、下記の結果を得た。

- (1) Web/AP/DBの各層において、処理を行なうサーバ数の動的な割当・削減を実現した。
- (2) システム負荷に応じて、Web/AP/DB各層のサーバ数を自動的に割当・削減できることを確認した。
- (3) アクセス集中時のWeb/AP/DBサーバの動的割当により、システムの応答時間を改善出来る事を確認した。
- (4) Web層/AP層では、アクセス集中が発生してから約20秒で、サーバを追加し、システム応答時間を改善することができる。
- (5) DB層においては、アクセス集中が発生してから2分程度で、サーバ割当・データベースプログラム起

動を行い、システム応答時間を改善することができる。急激なアクセス集中が生じた場合にも、システムの入口のWeb/AP層で高速なVPDC制御が実現できているため、DB層VPDC制御の遅延をカバーすることが可能である。

謝辞

本研究を進めるにあたり、当所エンタープライズシステム研究部の河辺峻(現在日立電子エンジニアリング)、マシエル・フレデリコ、亀山伸、川本真一の4氏には、本研究に関して貴重なアドバイス及び議論をさせて頂いた。深く感謝致します。

本研究の機会を与えて下さいました、当所情報システム研究センタセンタ長(現在ソフトウェア事業部)辻畑好秀氏、当所エンタープライズシステム研究部部長(現在エンタープライズサーバ事業部)稲上泰弘氏に感謝いたします。

(株)日立超LSIシステムズの小豆畑正一、蛇石武、三浦政宏、橋本孝子、真鍋和英、秋元貴樹の6氏、及び(株)日立ソフトウェアエンジニアリングの朝日通晴氏には、VPDCシステムに関する助言を頂くとともにシステム開発において多大な御支援を頂いた。深く感謝致します。

また、(株)日立情報システムズの富名則之、内山幸治、竹ノ内丈児の3氏にはシステム構築時に多くの御支援を頂いた。深く感謝致します。

本報告の研究を遂行するに当たり、ここで御紹介した以外の方々からも御指導と御支援を頂いた。これらの方々を含め、改めて感謝致します。

参考文献

- 1) 吉村 裕, 他: Webアクセス集中対応データセンタプラットフォームの研究, 信学技報 2001-29, pp.33-40 (2001)
- 2) 藤沼 彰久, 他: C/Sシステムの設計・構築 ~3階層型, Web系にいたるC/S開発のすべて~, 日経BP社 (1999)
- 3) インターネットトラブル続出, 日経コンピュータ 1999年8月16日号, pp.98-118 (1999)
- 4) K. Appleby, 他: Oceano SLA Based Management of a Computing Utility, IEEE International Symposium on Integrated Network Management 2001, (2001)
- 5) <http://jakarta.apache.org/tomcat/jakarta-tomcat/src/doc/Tomcat-Workers-HowTo.html>
- 6) <http://jakarta.apache.org/tomcat/>
- 7) <http://www.horb.org>
- 8) Oracle社, Oracle 8i Parallel Server マニュアル
- 9) Oracle社, Oracle 8i マニュアル
- 10) M. Miller: SNMP インターネットワーク管理, 翔泳社 (1998)
- 11) <http://www.tpc.org/tpcw/>