

クラウドで実現する 大規模Webプラットフォーム

佐藤慎太郎* 吉川晃平*
渡辺善三*
青木勇太*

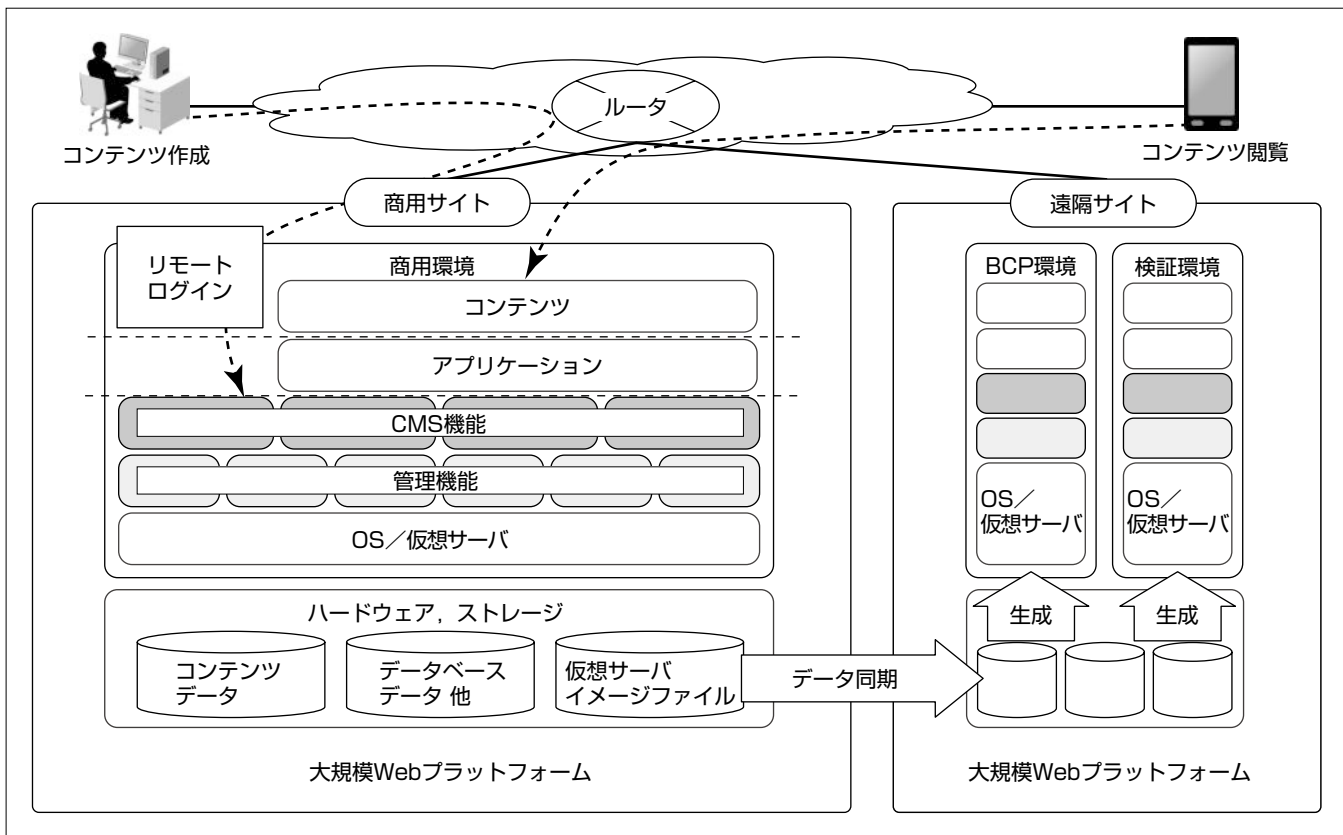
Large-scale Web Platform Based on Cloud Computing

Shintaro Sato, Zenzo Watanabe, Yuta Aoki, Kohei Yoshikawa

要 旨

今日の企業ポータルサイトは様々なサービスを提供するとともに、サービス障害情報や災害情報などの緊急性の高い情報を公開しており、社会的に重要性が高いシステムとなっている。さらに、近年スマートフォンの普及やモバイル回線の高速化による携帯端末向けコンテンツの大容量化が進み、またソーシャルネットワークサービスの普及によって情報がリアルタイムに共有されることで、短時間にアクセスが集中し、巨大トラフィックが頻発する傾向にあるが、そのような場合でも企業ポータルサイトには、安定して高速に応答できることが求められている。このような背景のなか、三菱電機インフォメーションシステムズ株(MDIS)は、プライベートクラウド技術とSaaS(Software as a Service)型クラウドサービスを利用し、数十ドメイン

のWebシステムからなる企業ポータルサイトの集約と、それによる企業ポータルサイトシステムの共通基盤化を行った。Webプラットフォームは多層化したキャッシュ機構を駆使して、巨大なリクエストに対して安定かつ高速な応答を返す仕組みとしている。また遠隔サイトにBCP(Business Continuity Planning)環境を構築し、大規模災害対策を実現した。企業ポータルサイトシステムの開発では、このシステムを利用する多数のステークホルダと密に連携し、システム運用フローの再定義や、性能対策の策定についてPDCA(Plan-Do-Check-Act)サイクルを実施することで、個々の課題を解決し、2013年3月からサービスを開始している。



大規模Webプラットフォームの構築事例

企業Webシステムプラットフォームを仮想サーバ上に構築した。このシステムは数十ドメインを収容し、多様なコンテンツベンダーが複数のCMS(Content Management System)を使用してコンテンツの開発と公開を行っている。また商用サイトと遠隔サイトがあり各種データを遠隔サイトに同期する構成となっている。遠隔サイトでは同期したデータを利用してBCP環境と検証環境を実現している。

1. ま え が き

企業のWebポータルシステムでは、アクセス集中に対して、安定かつ高速に応答することが求められる。また近年サービスの重要性が増しており、大規模災害時の事業継続性が求められる。これらの課題を解決するため、プライベートクラウド技術とSaaS型クラウドサービスを活用し、Webプラットフォームを構築した。

本稿ではWebプラットフォーム構築でのクラウド技術利用上のポイントを述べる。

2. 大規模Webプラットフォームの要件

システム構築に当たり、大規模Webプラットフォームとして必要な要件を表1のとおり定義した。

3. 大規模Webプラットフォームのアーキテクチャ

強固なインフラ基盤構築に向けて、大規模Webプラットフォームの要件を満足するために、このシステムで採用した対応策、アーキテクチャについて述べる。

3.1 応答性能対策

3.1.1 キャッシュを利用した性能対策

常時多数のリクエストに対応するため、次に示すように多段化したコンテンツキャッシュの仕組みを作った。アクセスするデバイスごと、コンテンツの内容ごとに最適なキャッシュを利用することで高速応答を実現している。

(1) SaaS型クラウドサービスの利用

画像や映像等、サイズが大きいコンテンツは帯域を圧迫する。対策としてSaaS型クラウドサービスである外部CDN(Contents Delivery Network)サービスを利用した。CDNサービスを利用することによって、ウェブサイトのコンテンツを世界中のASP(Application Service Provider)型Webサーバにキャッシュし、クライアントからのアクセスを分散させることで、帯域を消費するコンテンツにアクセスが集中した場合でも安定して高速に応答できる。ただし、CDNはコンテンツキャッシュの保持時間が非常に長い場合、コンテンツ変更後にキャッシュが更新されず古いデータが表示されてしまうことがある。そこでこのシステムではコンテンツ更新時に自動でCDNへキャッシュクリアを指示する機能を独自に開発し、常に最新のデータが

表示される仕組みを構築した。

(2) ロードバランサのコンテンツキャッシュ機能の利用

CDNを利用しないドメインでは、ロードバランサのコンテンツキャッシュ応答機能を利用することとした。静的コンテンツへのリクエストの大部分をロードバランサで応答させることで、Webサーバの負荷を大幅に軽減できる。デバイスごとに異なるデザインのページは、ユーザーエージェントごとに異なるキャッシュを保持・応答することで、デバイスごとに最適なコンテンツの高速応答を実現している。昨今増えているユーザーごとに異なるパーソナライズされたページに対して、ページ内の静的コンテンツはキャッシュで応答し、ユーザー固有に生成する動的コンテンツはAjax(Asynchronous JavaScript^(注1)+XML(eXtensible Markup Language))技術を応用し、クライアント(ブラウザ)側で非同期にアクセスさせることで、高速に応答している。

(3) オンメモリKVSの利用

データベースデータの参照は複数の参照用データベースサーバで処理を分散している。しかし、データベースデータの更新は1台のデータベースサーバで行う必要があるため、データベースデータの更新が発生すると応答性能が大きく低下する。そこでセッション管理を必要とするWebサイトでは、セッション管理のトランザクション処理にデータベースを使用せず、高速なオンメモリKVS(Key-Value Store)を利用することで、データベースデータの更新処理を回避して高速化している。

(4) ストレージキャッシュの利用

仮想化基盤システムでは一般的にコンテンツデータ、データベースデータ、仮想サーバのディスクイメージ等を共有ストレージ上に配置するため、共有ストレージのI/Oがボトルネックとなる。このシステムでは共有ストレージに高速なフラッシュデバイスを搭載して二次キャッシュとして利用することで、ディスクアクセス性能を向上させ、仮想化基盤全体のパフォーマンスの改善を図っている。

(注1) JavaScriptは、Oracle Corp.の登録商標である。

3.1.2 性能対策でのPDCAの取組み

画像や映像等のリッチコンテンツは転送時間が長く、インタラクティブなコンテンツを多用したWebページはシステムへの負荷が高いという課題がある。これらの課題はコンテンツの実装による影響が大きいため、コンテンツベンダーと連

表1. 大規模Webプラットフォームの要件

項目	説明	要件
応答性能	トラフィックが非常に多いが、アクセス集中時も安定かつ高速に応答できる必要がある	アクセス集中時(条件: 2,000PV/s, 5Gbps)に次を満たすこと ・全リクエスト正常応答 ・応答時間: 数十ms
システム可用性	社会的に重要性が高く(災害情報の提供など)、災害時にサービス継続が必要である	商用サイト災害時に次を満たすこと ・コンテンツデータを消失しないこと ・商用サイトでサービス停止した時にも、サービスを継続できること
システム利用ユーザー管理	サイト運用者、コンテンツベンダー等、多岐に渡る関係者が利用する	・利用者がアクセス可能なコンテンツデータ、機能を制限できること ・利用者の変更、権限設定を管理できること

PV : Page View

携して次のような性能対策のPDCAサイクル(図1)を実施した。

- (1) Webサイトの要件定義時から、クライアント種別(パソコン, スマートフォン等)とその比率, 及びコンテンツ種別(静的, 動的等)からトラフィックのモデル化を行い, それを基にリソースのサイジングを行う(Plan)。
- (2) システム構築後, モデルトラフィックで負荷をかけ性能試験を実施する(Do)。
- (3) 試験結果を評価し, ボトルネックを調査する(Check)。
- (4) 調査結果を基にインフラ及びコンテンツの性能対策を実施する(Act)。

このシステムではプロトタイプ作成時と完成時にPDCAを回し, プラットフォームとコンテンツでそれぞれ最適な性能対策を実施することで, ユーザーエクスペリエンスを低下させず高速応答を実現した。PDCAサイクルで得られた性能対策のノウハウはコンテンツ作成ガイドラインとして整理し, ステークホルダ間で共有して, システムの安定稼働のための対策を講じている。

3.2 システム可用性対策

このシステムでは, 商用サイトで大規模災害時のサービス障害対策として, サーバ仮想化技術を利用し, 遠隔サイトにはBCP環境を構築した。また遠隔サイトへ同期した各種データを利用し, コンテンツデータ消失対策として遠隔バックアップを実現し, また検証環境を構築した(表2)。次に, 遠隔サイトの各環境の構築, 機能等について述べる。

3.2.1 システム全体でのデータ同期の仕組み

遠隔サイトでのコンテンツデータの保存, 及びBCP環境, 検証環境の実現のため, 商用サイトと遠隔サイト間でシステムが利用する各種データをリアルタイムに同期させる設計とした。データの同期は共有ストレージの遠隔同期機能と, 各種ミドルウェアのレプリケーション機能を利用して実現した(図2)。

3.2.2 遠隔バックアップの実現

大規模Webプラットフォームでは, 様々な障害やオペレーションミスによるデータ消失に備えて, 商用サイト内だけでなくシステム外部にバックアップデータを保管して

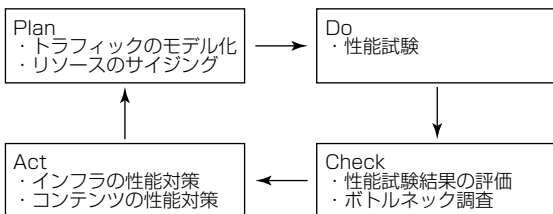


図1. 性能対策でのPDCA

表2. システムが備える環境

名称	説明	ロケーション
商用環境	通常時にサービスを提供する環境	商用サイト
BCP環境	災害発生時に切り替えてサービスを継続提供する環境	遠隔サイト
検証環境	コンテンツやインフラの動作検証を行う環境	遠隔サイト

おくことが重要である。このシステムでは商用サイトから遠隔サイトに同期したコンテンツデータ, 仮想サーバーイメージデータをバックアップデータとして利用することで, 商用サイトでデータを消失した場合に, 遠隔サイトのバックアップからのデータの復旧を可能とした。

3.2.3 BCP環境の実現

(1) BCP環境の要件

BCP環境の構築に当たり, WebプラットフォームのBCP環境として必要な要件を表3に示すとおり定義した。

次にBCP環境の要件を満たすために必要な各種データの最新化タイミングの要件を表4に示すとおり定義した。

この要件を満たすBCP環境と, 災害発生時の切替え運用については, サイト運用者と十分に検討を行い決定した。

(2) BCP環境の構築

BCP環境の構築に当たり, パッケージ製品の採用を検討

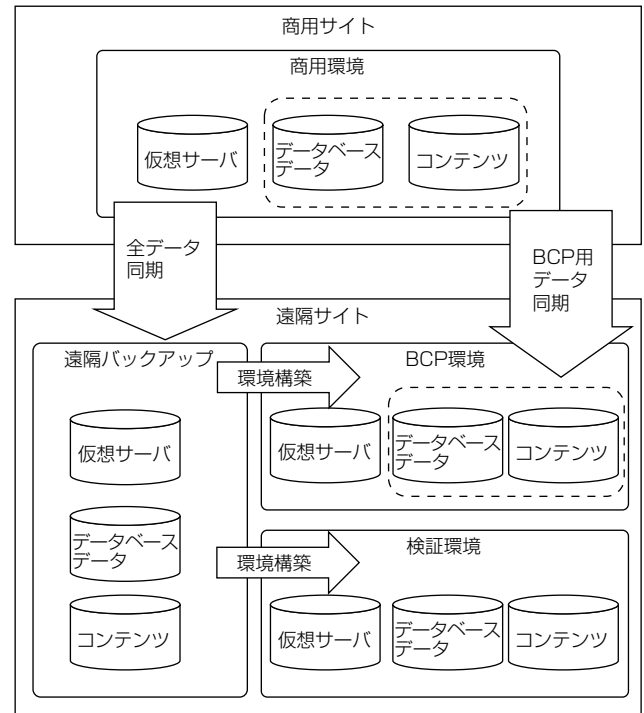


図2. サイト間データ同期とデータの利用方法

表3. BCP環境の要件

要件	目標
データ鮮度	公開中, 編集中のコンテンツデータがリアルタイムに同期されること
機能同一性	商用環境と同じ機能(コンテンツ公開, 編集, ユーザー管理)を提供できること
運用操作の容易性	被災時の切替えを1オペレーションで実施できること

表4. データの最新化タイミングの要件

データ種類	最新化タイミングの要件
コンテンツデータ, データベースデータ	コンテンツデータ, データベースデータは商用環境と同期されること
仮想サーバーイメージデータ	商用環境の変更の都度, 最新化すること(更新頻度: 多い)
ネットワーク設定, ハードウェア設定	商用環境の変更の都度, 最新化すること(更新頻度: 少ない)

したが、先に述べたWebプラットフォームに求められる様々な要件を満たすことができなかつたため、独自にBCP環境を構築することとした。

①BCP環境用コンテンツ設計

BCP環境で使用するコンテンツデータは、商用環境と同じ最新のものである必要がある。先に述べた、システム全体のデータ同期の仕組みによって、常に商用環境と同じデータが利用できる設計とした。

②BCP環境用仮想サーバ設計

BCP環境用の仮想サーバは、商用の仮想サーバを基にホスト名、IPアドレス等のネットワーク設定を拠点に合わせて変更して生成する設計とした。サービスに利用するネットワークのIPアドレスは、コンテンツやプログラム内に埋め込まれており、変更すると動作に影響するため、保守経路のIPアドレス(監視用)だけ変更し、それ以外のネットワークセグメントは商用環境と同じIPアドレスを使用する設計とした。

③短時間でのBCP環境用仮想サーバ構築の実現

BCP環境用仮想サーバの構築は、遠隔サイトに同期した商用環境の仮想サーバイメージを1台ずつBCP環境用に設定変更を行う。しかし仮想サーバの台数が多いため、手動で実施した場合に時間がかかる、また設定ミスが発生するリスクがあった。そこで仮想サーバの設定変更をバッチ処理で行う機能を開発し、短時間での多数のBCP環境用仮想サーバの構築を可能とした。

④BCPコンテンツ確認とリハーサル訓練

BCP環境の上位ネットワーク機器には、サービス用のグローバルIPアドレスとは別に、コンテンツ動作確認用のIPアドレスの設定を行った。この設定によって、平常時にBCP切替えを行うことなくBCP環境のコンテンツ動作確認や、BCPを想定したリハーサル訓練を可能とした。

(3) BCP切替えの運用

商用環境で災害が発生してサイト運営が困難と判断された場合、BCP環境への切り替えを行うことによってサイト運営を継続する。

平常時はBCP環境は次の①、②の状態でスタンバイ状態を保つ。

①ルータの設定

平常時は、システム上位のルータに商用環境に向けてルーティングが設定されている。切替え時には、上位ルータのルーティングテーブルをBCP環境に振り向けることでIPレベルで拠点間切替えを実現する。

②データベースサーバの構成

平常時はデータベースサーバは商用環境がマスター、BCP環境がスレーブとして同期しているため、BCP環境のデータベースサーバは書き込み不可状態になっている。切替え時にはデータベースサーバを書き込み可能状

態に変更し、BCP環境でデータが更新できるようにする。BCP切替えが必要となった際には①' ②' のように設定を変更する運用としている。

①' ルーティング切替え操作

ルーティング変更(①)はBGP-4(Border Gateway Protocol version 4)プロトコル^(注2)によってシステム上位のルータのルーティングテーブルを商用環境からBCP環境に瞬時に切り替える方式とした。これによって、短時間でのBCP切替えを実現した。このシステムでは、あらかじめ商用環境、BCP環境の両方の上位ルータに設定を組み込んでおくことで、災害発生時に簡単な操作で切替えを可能とした。

②' データベースサーバの構成変更

データベースサーバの構成変更(②)は、切替え操作ミスの防止及び作業時間の短縮のため、BCP切替えプログラムを開発し、1オペレーションでの切替えを実現した。

(注2) BGP-4は、インターネット・バックボーン上でルーティングテーブル情報を交換する際に用いる経路制御プロトコル。

3.2.4 検証環境の構築

BCP環境は平常時は利用されないため、遠隔サイトのハードウェアリソースは遊休状態となっている。遠隔サイトのハードウェアリソースを有効活用するため、遠隔サイト上に検証環境を構築し、平常時は検証環境として活用する設計とした。検証環境の構築は、先に述べたBCP環境構築機能と同様に、商用環境の仮想サーバイメージを検証環境用にバッチ処理で変換する方法で行い、検証環境構築作業を効率化した。

3.3 システム利用ユーザーの管理

このシステムでは多数のコンテンツベンダーなどがCMSを利用してコンテンツを作成しているが、セキュリティ確保と情報漏えい防止の観点から、利用者ごとにアクセス可能なコンテンツデータを厳密に分離する必要がある。そのために、このシステムではLDAP(Lightweight Directory Access Protocol)をベースにアカウント管理機能を開発した。この機能によって、利用者ごとのロール設定(権限管理)を実現するとともに、コンテンツ作成機能に関する各認証(リモートログイン認証、CMSログイン認証等)の一元管理を実現している。

4. む す び

このシステムは2013年3月からサービスを開始し、その後安定してサービスを提供している。Webポータルサイトでは、常に新しいコンテンツやサービスの追加が行われており、応答性能の安定化と迅速なサービス基盤の拡張が求められている。引き続きクラウド技術に積極的に取り組み、Webプラットフォームの高機能化を図り、また他システムへの展開を進めることで、更なるインフラ基盤共通化を目指していく。