

コンシューマ・サービス論文

アクセス予測に基づいた 広域冗長型安否システムの提案と基礎評価

永田 正樹^{1,2,a)} 阿部 祐輔² 金原 一聖² 福井 美彩都² 峰野 博史¹

受付日 2015年10月1日, 採録日 2016年2月23日

概要: 本論文では, WEB システムを構成するサーバ群を世界規模で冗長化し, アクセス数を事前予測して適切なサーバ数で負荷分散を行う広域冗長型安否システムを提案する. 既存の安否システムには, 災害時の確実な稼働とアクセス状況に応じた適切なサーバ数算出が困難という課題がある. そこで提案システムでは, 日本, シンガポール, カリフォルニア, など多地点および地理的に離れた広域拠点にシステムを分散配置し可用性を向上させる広域冗長化機構と, 災害規模に応じたアクセス数を事前予測し適切なサーバ数で負荷分散を行うアクセス予測に基づくオートスケーリング機構を実装した. 過去の災害データを基にしたシミュレーション評価の結果, 多地点での広域冗長化稼働とアクセス予測に基づくオートスケーリング機構を用いて年間費用約 32%削減の見込みを得ることでき, 提案システムの有効性を確認した.

キーワード: 広域冗長化, 負荷分散, アクセス予測, スケーラビリティ, 安否情報システム

A Proposal and Evaluation of a Global Redundant Safety Information System Based on Access Prediction Model

MASAKI NAGATA^{1,2,a)} YUSUKE ABE² ISSEI KINPARA² MISATO FUKUI² HIROSHI MINENO¹

Received: October 1, 2015, Accepted: February 23, 2016

Abstract: In this paper, we propose global redundant web safety information system which implemented global redundant on the world area and load balancing by the access prediction model. Conventional web safety information system has a problems, which are certain operation at the disaster and adjustment of the number of servers in conformity with access. So we have developed global redundant and auto scaling based on access prediction model to settle a problems. Global redundant arranges a system separately in the distant area and multiple locations (e.g., Japan, California and Singapore). Auto scaling based on access prediction model predicts the access amount according to the disaster scale before access concentration. And load balancing executes. A result of the simulation experiment expected a cost reduction about 32% by global redundant and auto scaling based on access prediction model, and the validity of the proposal system was confirmed.

Keywords: global redundant, load balancing, access prediction, scalability, safety information system

1. はじめに

安否システムとは災害（本論文では地震とする）時に,

システム対象ユーザの安否情報を収集・公開する WEB システムである. 災害時の被災状況や被災者の安否情報を公開する安否システムは, 家族・組織間での安否確認や災害後の復旧活動の迅速化など, 災害情報を多数の関係者に公開する仕組みとして WEB システムでの実装が適している [1], [2]. 安否システムへの要件として災害時の持続稼働があげられる. 東日本大震災のような深刻な被害を出した災害では早期の被災者安否の収集・公開が多数の人命救助につながるため [3], 安否情報を公開する WEB システムの

¹ 静岡大学創造科学技術大学院
Graduate School of Science and Technology, Shizuoka University, Hamamatsu, Shizuoka 432-8011, Japan

² 株式会社アバンセシステム
AvanceSystem Corporation, Hamamatsu, Shizuoka 435-0037, Japan

a) nagata@avancesys.co.jp

停止は避けなければならず災害時に途絶することなくシステムを稼働し続けることが重要である [4]。一般的な安否システムの動作仕様は、災害発生時システムからユーザへ安否報告を促すメールが送信され、ユーザは受信したメールに対しアクションして自己の安否情報を登録・公開し利用者間で安否情報を共有するものである。昨今の安否システムはクラウドコンピューティングを利用した実装が主流であり、スケーラブルな基盤上での冗長化や負荷分散技術を用いて持続可能なサービス提供を実現している。運用面では、サーバ、スイッチなどのハードウェアを資産として保持せずクラウドへアウトソースすることで、経営資源、人件費などの費用削減を可能とし収益性向上に寄与する。しかし安否システムを既存のクラウド環境上で運用する場合、2つの課題がある。

1つ目の課題は、広域拠点での冗長化である。安否システムは災害時においてユーザの安否情報収集・公開を目的とするため、災害時は確実に稼働し続けなければならない。東日本大震災のような大規模災害では災害地域周辺のデータセンタ施設群が倒壊する恐れがあり、災害地域のデータセンタで安否システムを稼働している場合、システム停止が懸念される。つまり一地域・一大陸内での冗長化対策では大規模災害時のシステム停止リスクを回避できない。

2つ目の課題は、アクセス状況に応じた適切なサーバ数算出である。安否システムは災害時にユーザからの安否報告や安否情報公開リクエストが集中するため、平常時と比較するとアクセス処理に必要なサーバ数に開きがある。仮に災害発生に備え災害時のアクセスに対処可能なサーバ数で常時運用する場合、アクセスが少ない平常時では過剰リソースとなり安否システムのように常時稼働を求められるサービスでは費用対効果の面から困難な運用を強いられる。

本研究では上記2つの課題を解決するため、WEBシステムを構成するサーバ群を世界規模で冗長化し、災害時のアクセス数を事前予測し適切なサーバ数で負荷分散を行う広域冗長型安否システムを試作開発し評価結果から有効性を示す。

2章では従来研究と課題、3章では提案システム構成、4章ではアクセス予測モデル、5章では実装と評価、6章では考察、7章でまとめとする。

2. 従来研究と課題

2.1 安否システム

安否システムの関連研究には、図1に示すように、情報収集、WEBシステム、通信など様々な分野の研究が関係するが、本研究では冗長化、負荷分散といったWEBシステム基盤を対象とする。通信関連の研究では、ワンセグや無線など [5], [6] を用いて災害時のネットワーク輻輳に対して輻輳回避や回線確保のためのアルゴリズムを用い確実な通信手段を実現し有効性を示している。情報収集関連の研

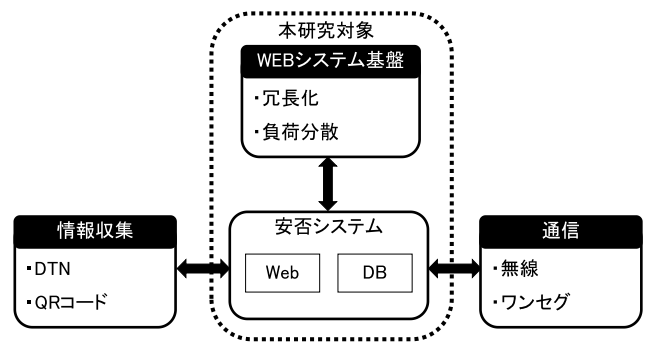


図1 安否システム関連研究

Fig. 1 Related study of safety information system.

究では、DTNを用いたすれ違い通信で安否情報を中継し最終的にSNSへ登録する提案 [7] や、QRコードを用いた安否情報の収集 [8] では、収集後の情報管理までを含んだ情報マネジメントシステムの構築を行い利便性の高い安否情報管理を実現している。これら通信および情報収集関連の研究においても情報管理には一般的にWEBシステムが用いられているため、災害時の総合的な安否情報管理にはWEBシステム基盤の持続稼働が重要となる。

WEBシステム基盤関連の研究では [9], [10]、災害時にシステムのロバストネス向上を目的とし、複数サーバを用いたミラーリングでの冗長化や、DNSラウンドロビンやリダイレクトでのアクセス先振り分けを用いた複数サーバでの負荷分散の提案がある。しかし冗長化は国内での実装評価にとどまっておき東日本大震災規模の災害ではシステム持続稼働が課題となる。負荷分散は複数サーバを用いている点は本研究においても同じ立場をとるが、平常時・災害時を問わず常時複数サーバでの構成のため平常時でのサーバ費用や余剰リソースが課題となる。

2.2 課題1：広域冗長化

災害時の確実な稼働を要求される安否システムは一地域内でなく多地域での冗長化運用が有効である。AWS (Amazon Web Services) [11] や Azure [12] に代表されるクラウドベンダは、世界各地にデータセンタを有しており広域冗長化を実現するシステム基盤として適している。しかし現状の各サービスは主に同一の地域 (以下、「リージョン」) 内での提供が主流であり、システムを複数地域にまたぐ構成にする場合は課題がある。負荷分散を行うロードバランシングサービスではロードバランサが稼働しているリージョン内のサーバに対してのみ通信可能で他リージョンのサーバへ通信できない。冗長化では、リージョン内の同ネットワークセグメントのサーバに対してルーティングテーブルを変更し参照先サーバを切り替えることでフェイルオーバーが可能だが、ルーティングテーブルへは他リージョンのネットワークセグメントを指定できないため、やはりリージョン内での冗長化になる。つまり1リージョン全域が大

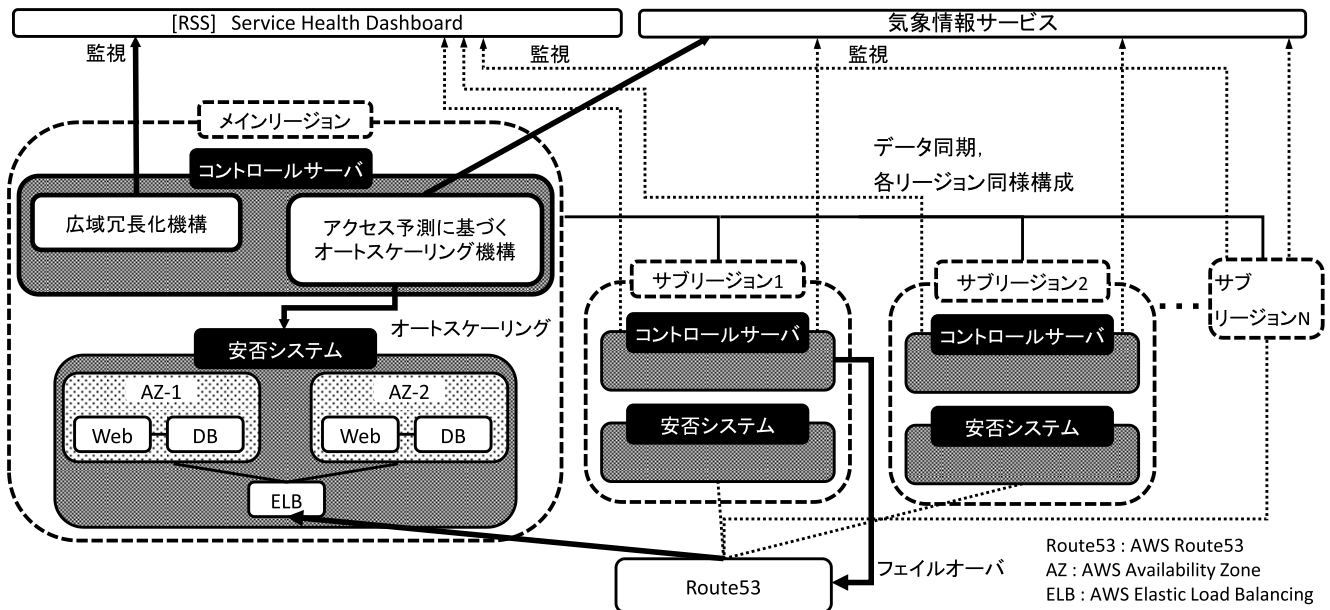


図 2 広域分散 WEB 安否情報システム

Fig. 2 Architecture of global distributed Web safety information system.

災害などで被災しリージョン内のすべてのサービスが停止した場合、複数リージョン間での広域冗長化が必要になるためクラウドベンダの標準サービスに追加した機能実装が必要となる。

2.3 課題 2：状況に応じたサーバ数

平常時と災害時でシステムへのアクセス数に開きがある安否システムは、アクセス状況に応じたサーバ数で運用することで費用削減が可能である。アクセス集中に対するシステムのリソース管理として最も安直な施策は、状況によらずあらかじめ多数のサーバで常時稼働しておくことである。しかし常時多数サーバでの運用は、アクセスが少ない平常時のサーバリソースは余剰リソースとなり、そのまま余分費用に直結する。広域冗長化は災害の発生予測が困難なため常時敷設でなければ意味をなさないが、アクセス状況に応じたサーバ数をそのつど確保する仕組みがあればアクセスの少ない平常時の余剰リソースを削減でき、安否システムのリソース管理としては理想的である。またアクセス集中後のサーバ追加では、システムが高負荷となった状態での事後対応となりレスポンス低下などユーザーの利便性を損ねるため、アクセス集中前に適切なサーバ数算出および追加が望まれる。アクセス状況に応じた適切なサーバ数を算出するためには、災害発生からその後のアクセス分布の予測が必要である。

3. 提案システム

3.1 システム概要

広域冗長型安否システム（以下、「提案システム」）は、システム基盤に AWS を用い、図 2 に示すようにメイン

リージョンとサブリージョンで稼働し、各リージョンにそれぞれ安否システムとコントロールサーバを配置する。各リージョンは、主たるユーザーのアクセス地域（本論文では東京）から、WEB アクセスの応答速度を計測する httping コマンドでの応答速度が速い順に、メインリージョン、サブリージョン 1, 2 とする。AWS がサービス展開している 12 リージョンのうち、メインリージョンを中心として位置的に離れているリージョンをサブリージョンに選択することで、近接地域に偏ることなく広域な冗長化が可能となる。サブリージョンは数が多いほどシステムの可用性が向上するが、本研究では広域冗長化の基礎評価のため 2 つのサブリージョンを用いた。システム正常稼働中のアクセスは、つねにメインリージョンに向けられ、サブリージョンはメインリージョンのバックアップサイトとしてホットスタンバイする。各リージョンには異なる地点に配置されているアベイラビリティゾーン（以下、「AZ」）があり、AZ は一般的なデータセンタと同意である。安否システムはリージョン内の 2 つの AZ に WEB・DB サーバ 1 台ずつ合計 WEB・DB サーバ 2 台を標準構成としてシステムを構成し、リージョン内単体においても閉域冗長構成を敷く。

提案システムは、コントロールサーバと安否システムで構成する。コントロールサーバは、広域冗長化機構とアクセス予測に基づくオートスケーリング機構を有する。広域冗長化機構は、障害発生時バックアップサイトでシステム稼働を継続しシステム全体停止を回避する。アクセス予測に基づくオートスケーリング機構は災害時のアクセス集中に対して適切なサーバ数で負荷分散を行う。安否システムは、図 3 に示すとおり複数顧客がサーバリソースを共有する運用形態であり、あらかじめシステム登録済みのユー

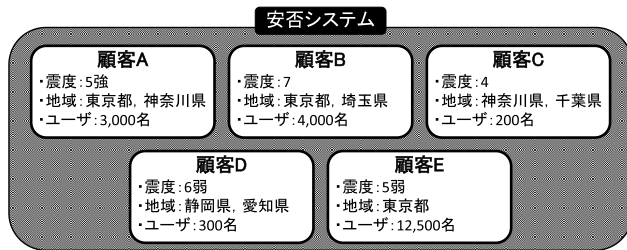


図 3 安否システムの運用形態

Fig. 3 The practical use form of safety information system.

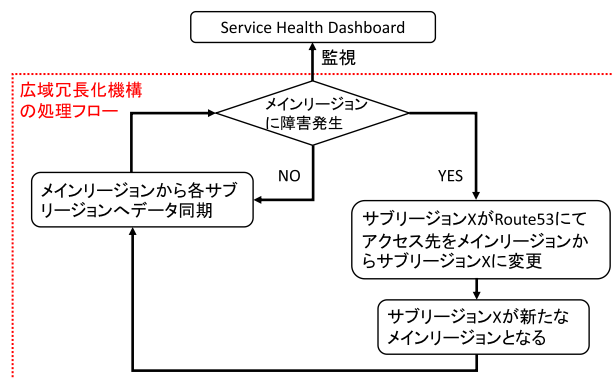


図 4 広域冗長化機構

Fig. 4 Global redundant.

ザが利用可能となる。動作概要は、顧客単位で設定した地域および震度閾値に該当する災害が発生した際、対象ユーザへ安否報告を促すメールを送信し、受信したユーザが安否報告アクセスを行う流れとなる。災害時の対象ユーザ数は、たとえば東京都、神奈川県に震度5強の災害が発生した場合は、図3から顧客A、C、Eとなり15,700名となる。つまり提案システムは、災害規模に応じてシステムに登録済みユーザ数の範囲内で対象ユーザ数が変化し、対象ユーザ数に応じてアクセス予測および負荷分散を行う。

3.2 広域冗長化機構

広域冗長化機構は、各リージョン間のデータ同期や障害発生時のフェイルオーバーを行い、図4の処理フローとなる。障害発生はAWSの障害状況サイト「Service Health Dashboard[13]」から検知する。AWS各サービスはAWSソフトウェア層の不具合だけでなくAWS物理層の障害などにも起因して影響を受けるため、「Service Health Dashboard」の監視でソフトウェア層および物理層両者の障害検知が可能となる。また図2に示すように各リージョン相互に「Service Health Dashboard」を監視し、1リージョン停止によるシステム全体停止を回避し可用性向上を実現する。

フェイルオーバーは図4に示すように障害発生を起点に実行し、アクセス先をメインリージョンからサブリージョンに変更することで実現する(図5)。アクセス先リージョンの変更はAWSのDNSサービスであるRoute53[14]で

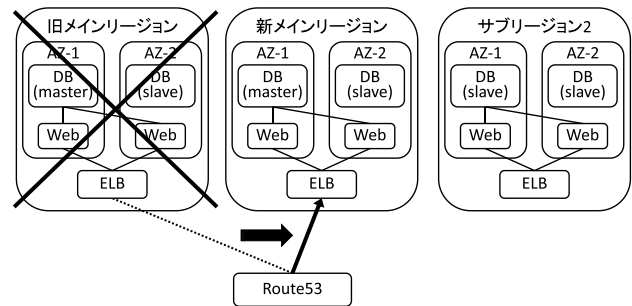


図 5 フェイルオーバー

Fig. 5 Failover.

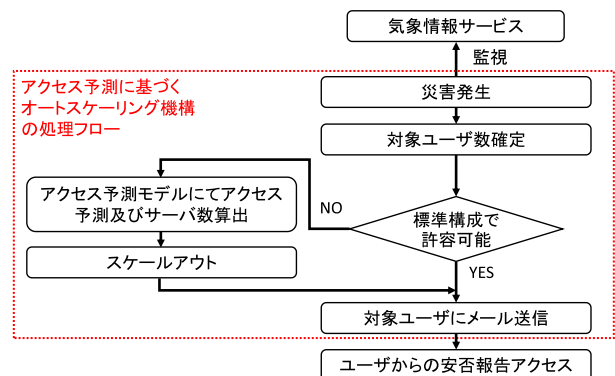


図 6 アクセス予測に基づくオートスケーリング機構

Fig. 6 Auto scaling based on access prediction model.

行う。メインリージョンに障害が発生した場合、サブリージョンはRoute53を用いてアクセス先の重み付けを変更しサブリージョンにアクセスを向ける。新たにアクセスを受けるサブリージョンはメインリージョンに昇格し元のメインリージョンが復旧するまでその役目を果たす。新たなメインリージョンの選択基準は、アクセス地域からhttpingコマンドでの応答速度が旧メインリージョンの次に速いサブリージョン1であるため、メインリージョンはユーザアクセス地域からつねに最も応答速度の速いリージョンとなる。

メインリージョンとサブリージョンのデータ同期は、図4に示すように障害発生時でない平常時にメインリージョンのDBデータをサブリージョンのDBサーバへレプリケーションで行う。平常時はメインリージョンにアクセスが向けられるため、メインリージョンAZ-1のDBサーバがマスターとなりAZ-2がスレーブとなる。同様にサブリージョンのDBサーバはメインリージョンのスレーブとなり広域間でデータを冗長保持する。

3.3 アクセス予測に基づくオートスケーリング機構

アクセス予測に基づくオートスケーリング機構は、災害の対象ユーザ数に応じて適切な負荷分散を行い図6の処理フローとなる。災害発生は気象情報を電文形式で公開するサービス(たとえば文献[15])から検知する。サーバ数の増加をスケールアウト、削減をスケールインと呼ぶ。災害

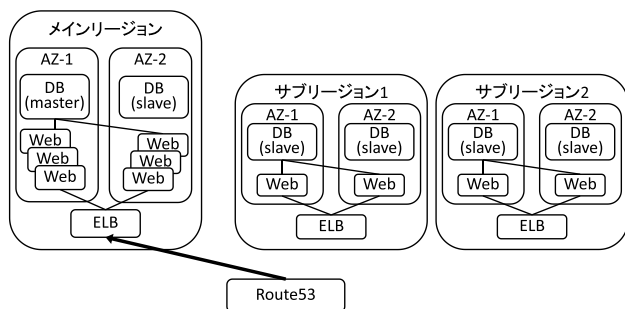


図 7 負荷分散

Fig. 7 Load balancing.

表 1 スケールアウトのタイミング

Table 1 Scale out timing.

	常時 配置	事後 追加	事前追加 (予測なし)	事前追加 (予測あり)
利便性	○	×	△	○
費用	×	○	○	○

の対象ユーザ数に対してシステム標準構成のサーバ数でアクセスを許容可能であればスケールアウトは実行されず、許容不可能な場合にアクセス予測モデルでアクセス予測および適切なサーバ数でスケールアウトし負荷分散を実行する。スケールアウトの規模は対象ユーザ数に応じて変化する。負荷分散は、AWSのロードランシングサービスであるELB[16]配下の各AZ内にWEBサーバを同数配置することで均等に実行される(図7)。アクセス予測モデルは、過去の災害・訓練時のアクセス分布を分析し構築した確率分布モデルでありあらかじめシステムへ組み込まれている。このため対象ユーザ数をモデルに与えるだけでサーバ数算出が可能であり、災害状況下での迅速なスケールアウトを実現する。また、スケールアウト完了後に対象ユーザへメール送信することで、負荷分散環境構築前のアクセス集中を回避する。

WEBサーバをスケールアウトするタイミングは表1に示すように、サーバを平常時からつねに複数台配置する常時配置、アクセス集中後に追加する事後追加、アクセス集中前に追加する事前追加がある。常時配置はアクセスが少ない平常時では余剰リソースとなり高費用となる。アクセス集中後の事後追加は、システムが高負荷となった状態での対応のためレスポンスが低下しユーザ利便性を損ねる。アクセス予測なしの事前追加では、サーバ数算出に明確な根拠がないためリソース余剰もしくは不足の懸念がある。提案システムで用いるアクセス予測ありの事前追加では、アクセス予測を用いて必要サーバ数を算出するため、適切なリソース管理が可能である。また、提案システムで用いるサーバAWS EC2[17]は、起動時間で費用が積算されるため使用時のみのサーバ起動で費用削減が可能となる。つまり提案システムは、アクセス集中前にアクセス予測をす

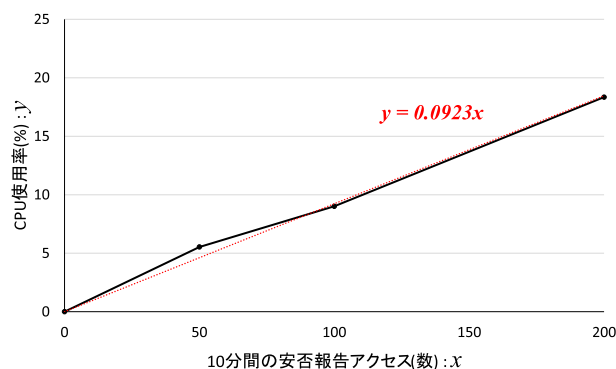


図 8 安全報告アクセス数とCPU使用率

Fig. 8 Safety report accesses and the CPU usage.

ることで適切なサーバ数でのスケールアウトを実現し、利便性だけでなく適切な費用での運用が可能となる。

WEBシステムへのリソース管理において、先行研究[18]はサーバ単体のベンチマーク結果からシステム全体のリソース管理のモデル化を試みている。本研究では先行研究を参考に、予測で得られたアクセス数に対しサーバ1台のアクセス許容量から必要台数を算出する。本研究でのアクセス許容量とは、単位時間あたりに処理可能な安全報告アクセス数とする。提案システムで採用するサーバは、EC2タイプt2シリーズのt2.smallとする。t2シリーズは、CPU使用率のベースラインが定められておりCPU使用率がベースラインを超えた場合バースト状態となる。t2シリーズにおけるバーストとは一時的にCPU性能が向上する状態であり、AWSのCPUクレジットを消費して継続可能となる。バースト状態ではベースラインを超えたCPU使用が可能だが徐々にCPUクレジットを消費していきクレジットが尽きた場合、ベースライン以上のCPU性能が発揮されないという特性がある。本研究ではバースト状態でのアクセス許容量は考慮せず、t2.smallのベースラインとして定められているCPU使用率20%までの安全報告アクセス数を1台のアクセス許容量とする。本研究で対象とする安全システムへの災害時の安全報告アクセスは、主にWEBサーバのCPUリソースを消費する。したがって、サーバ1台のアクセス許容量は安全報告アクセス数に対してCPU使用率の関係から算出する。図8は10分間の安全報告数アクセス数とt2.smallのCPU使用率の関係である。図8からCPU使用率20%時点でのアクセス許容量は約200安全報告アクセス/10分間となり、この値を基に必要台数を算出する。なお、標準構成ではWEBサーバに2台のt2.smallを用いるため、約400安全報告アクセス/10分間が標準構成時のアクセス許容量となる。また、安全システムへの災害発生後のアクセス数は分単位で増減が見られるため、予測単位時間を細かくするほど詳細予測が可能となるが、現状のEC2課金単位が1時間であることと、アクセス予測からWEBサーバのスケールアウト完了までに数分程度要することを考慮して、本研究ではアクセス予測

を10分単位で行うことを想定し、アクセス許容量を10分単位で算出した。

4. アクセス予測モデル

4.1 安否システムのアクセス分布特性

アクセス予測モデルは、過去の災害データの分析からアクセス傾向をモデル化し災害発生後、システムに対してのアクセス分布を予測する。アクセス分布の予測には過去の災害時のアクセス分布特性の把握が必要である。図9は災害発生時の安否報告アクセス分布である。アクセス分布はWEBサーバに用いているApacheのログファイルから抽出した。図9からアクセス分布はシステムからのメール送信を起点(0分時)とし、しばらくしてピークを迎え時間経過とともに減衰する。ある動作を起点としたシステムへのアクセス分布に対し先行研究[19]では正規分布を仮定している。また文献[20]ではネットワークのトラフィック分布に対し対数正規分布を用いて解析を試みている。本研究では図9の形状からメール送信を起点とした安否システムへのアクセス分布は対数正規分布に従うと仮定した。アクセス時間の対数をとった値でQ-Qプロットしたものが図9のQ-Q Plotである。アクセス分布は集計時間単位が短いほど対数正規分布の傾向がより顕著なため、ここでは1秒単位で集計した。Q-Q Plotから経過時間における安否報告アクセス分布の大部分が対数正規分布に従っている。文献[21]では、HTTPにおけるサービス時間分布に対して対数正規分布を用いたモデル化や、文献[22]ではサーバへの実アクセス分布に対数正規分布を用いており、本研究においても安否システムへのアクセス分布予測に対数正規分布を用いた。

4.2 対数正規分布を用いたアクセス予測モデルの構築

安否報告アクセスの分布を予測するため、対数正規分布の確率密度関数に必要なパラメータ決定について明らかに

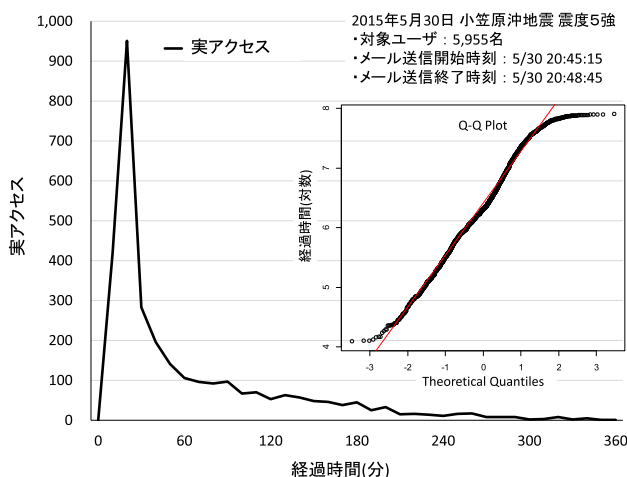


図9 災害時のアクセス分布

Fig. 9 The access distribution of the disaster.

する。対数正規分布の定義式は、確率密度関数 $f(x)$ は式(1)、最頻値 M は式(2)、平均値 E は式(3)となる。 μ は正規分布の平均値、 σ は正規分布の標準偏差である。式(1)の確率密度関数に対してパラメータ μ 、 σ を与え確率変数 x (分) の関数としてアクセス分布の確率を求める。 μ 、 σ は、式(2)、式(3)を連立して解くと式(4)、式(5)となる。式(4)、式(5)へ最頻値 M と平均値 E を与えることで μ 、 σ が決定できる。次に災害発生時その災害に対しての最頻値 M と平均値 E の算出方法を示す。

最頻値 M はアクセスが最も多い時間であり、平均値 E は平均アクセス数の時間となる。最頻値 M と平均値 E は過去の災害の対象ユーザ数 (TU) と災害発生後のアクセス分布を分析し算出する。表2は過去の災害・訓練時のアクセスデータを10分間隔で集計したものであり、各パラメータを求めるための分析対象データである。最頻値 M は表2から災害および訓練の対象ユーザ数 TU によらず20分以下でピークを迎えている。したがって、今回構築するモデルでは最頻値 M を固定値20とする。平均値 E は、アクセスが開始された時間からアクセス数が1桁台になる時間までのアクセス数の平均値をとり、その平均値のアクセスがあった時間とする。表2から平均値 E は、対象ユーザ数 TU に影響を受けない最頻値 M と異なり、対象ユーザ数 TU に応じて値に開きがある。図10は対象ユーザ数 TU と、平均値 E から最頻値 M を引いた差 (D) の関係である。図10の近似式が式(6)となり対象ユーザ数 TU から D を求めることで平均値 E が決まる。これまでで、災害発生時その災害の対象ユーザ数 TU を基に最頻値 M と平均値 E を推定できる。また求めた最頻値 M と平均値 E を式(4)、式(5)に代入すると μ と σ が決定し、同時に式(1)の確率密度関数のパラメータが決まる。

$$f(x) = \frac{1}{x\sqrt{2\pi}\sigma} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right) \quad (1)$$

$$M = \exp(\mu - \sigma^2) \quad (2)$$

$$E = \exp\left(\mu + \frac{\sigma^2}{2}\right) \quad (3)$$

$$\mu = \frac{(\ln(M) + 2 * \ln(E))}{3} \quad (4)$$

$$\sigma^2 = \frac{2 * (\ln(E) - \ln(M))}{3} \quad (5)$$

$$E = M + 4.6039 * TU^{0.2802} \quad (6)$$

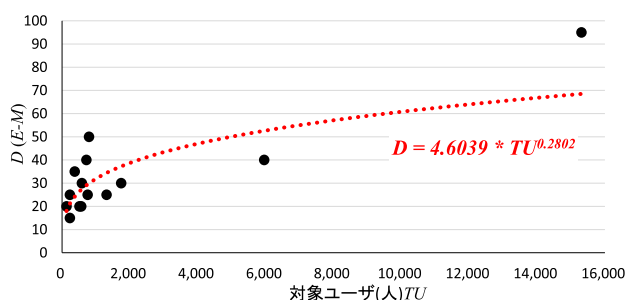
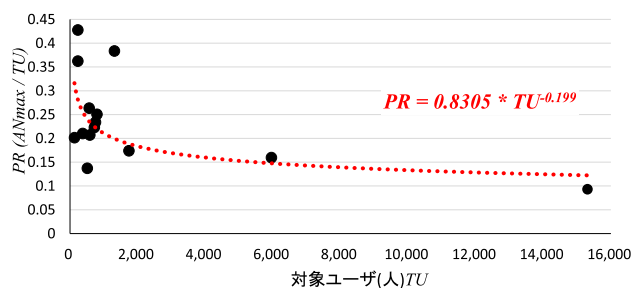
$$AN = A * f(x) \quad (7)$$

次に災害時のアクセス分布曲線を決定するために式(1)に付与する係数 A を算出する。 A は式(1)を対象ユーザ数 TU に応じたピークアクセス数に合わせるための係数であり、式(7)が x 分時のアクセス数 AN を予測するアクセス予測モデルとなる。係数 A は式(7)の x に最頻値 M を代入し、 AN がピークアクセス数 (AN_{max}) となるよう求

表 2 過去の災害・訓練データ

Table 2 Data of the past disaster and training.

区分	対象組織, 災害	TU	M	E	AN_{max}	$D (E-M)$	$PR (AN_{max}/TU)$
訓練	教育機関 A	1,311	20	45	503	25	0.3836
訓練	教育機関 B	371	10	45	78	35	0.2102
訓練	教育機関 C	15,308	20	115	1,425	95	0.0930
訓練	企業 A	229	10	25	98	15	0.4279
訓練	企業 B	794	10	60	199	50	0.2506
訓練	企業 C	229	10	35	83	25	0.3624
訓練	企業 D	129	20	40	26	20	0.2015
災害	宮城県沖 (サーバ 1)	717	10	50	160	40	0.2231
災害	奄美大島近海 (サーバ 1)	565	10	30	149	20	0.2637
災害	奄美大島近海 (サーバ 2)	584	10	40	121	30	0.2071
災害	埼玉県北部 (サーバ 1)	510	20	40	70	20	0.1372
災害	埼玉県北部 (サーバ 2)	752	10	35	176	25	0.2340
災害	小笠原諸島 (サーバ 1)	1,740	20	50	303	30	0.1741
災害	小笠原諸島 (サーバ 2)	5,955	20	60	951	40	0.1596

図 10 TU と $D (E-M)$ の関係Fig. 10 Relationship between TU and $D (E-M)$.図 11 TU と $PR (AN_{max}/TU)$ の関係Fig. 11 Relationship between TU and $PR (AN_{max}/TU)$.

める。表 2 からピークアクセス数 AN_{max} は対象ユーザ数 TU が増加するほど対象ユーザ数 TU に対する割合 (PR) が減少していく傾向にあり図 11 の関係となる。図 11 の近似式で対象ユーザ数 TU からピークアクセス数 AN_{max} を求める。例として、アクセス予測モデル式 (7) を用いて対象ユーザ数 20,000 人でのアクセス分布を予測すると図 12 の各パラメータおよびアクセス分布曲線となる。各時間の予測アクセス分布に対して許容可能なサーバ数を割り当て

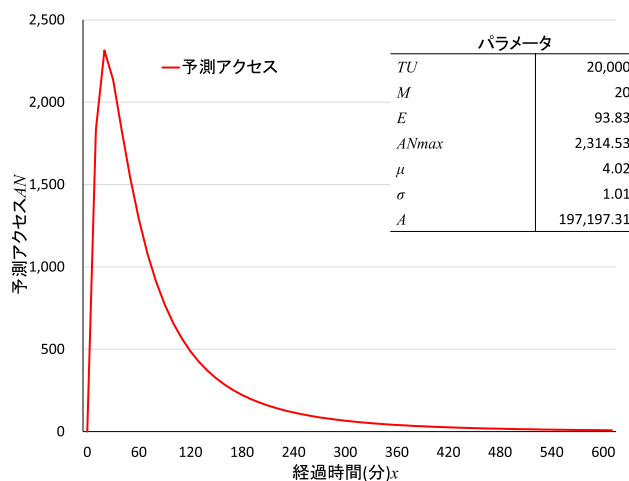


図 12 ユーザ数 20,000 名でのアクセス分布予測

Fig. 12 The access distribution prediction of the 20,000 users.

ることでアクセス数に応じた負荷分散を行う。

5. 実装と評価

5.1 実装

本研究では提案システムに対して疑似ユーザアクセスを行う評価サーバを用いて挙動を評価した。提案システムと評価サーバの構成を図 13、システム環境を表 3 に示す。図 13 から提案システムには 3 つのリージョンを用い、メインリージョンを日本 (httping: 71.5 ms)、サブリージョン 1 をシンガポール (httping: 226.2 ms)、サブリージョン 2 をカリフォルニア (httping: 303.7 ms) とした。フェイルオーバーやオートスケーリングは表 4 の AWS API を用いて実装し、アクセス予測モデルの計算は PHP の Math 関数で実装した。評価サーバには、「Service Health Dashboard」

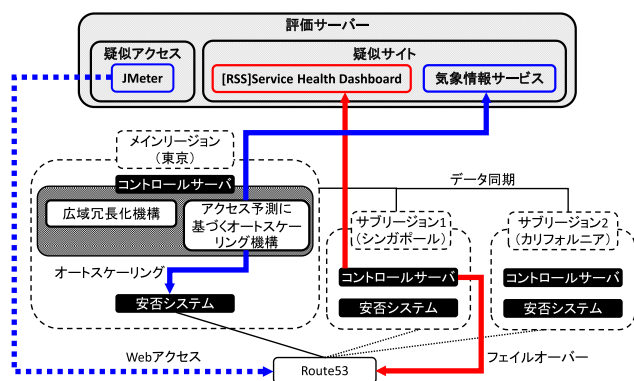


図 13 評価構成

Fig. 13 Evaluation constitution.

表 3 システム環境

Table 3 System environment.

	Hardware	Software
WEB	EC2 t2.small	CentOS 6.4 Apache 2.2.15
DB	EC2 c3.2xlarge	CentOS 6.4 PostgreSQL 8.4.12
コントロールサーバ	EC2 t2.micro	CentOS 6.4 PHP 5.3.3
評価サーバ	EC2 m3.medium	Windows Server 2008 JMeter 2.13

表 4 AWS API

Table 4 AWS API.

EC2 操作	ec2-api-tools 1.6.7
スケールアウト・イン	AutoScaling CLI 1.0.61
DNS 操作	Route 53 Authentication Tool for Curl(dnscurl.pl)

の RSS と気象情報サービスの疑似サイトおよび、負荷分散評価のために JMeter を用いた疑似アクセス環境を実装した。JMeter とは WEB サーバに対し複数のリクエストを送信し負荷をかけることができるツールである。

5.2 広域冗長化機構に関する評価

広域冗長化機構の評価では、評価サイトにメインリージョンである日本に障害が発生した情報を配信し、フェイルオーバーの動作を確認した。本研究ではアクセス予測に基づくオートスケーリング機構のアクセス予測モデルに焦点を絞っているため、広域冗長化機構の実験詳細や考察は省略し結果のみ述べる。フェイルオーバーの動作確認として、評価サイトに障害情報発信後、サブリージョン 1 が Route53 でメインリージョンの重みを 0、サブリージョン 1 の重みを 1 に設定することで、約 70 秒後にアクセス先が新たなメインリージョンであるシンガポールになることを確認した。また、フェイルオーバー時、リージョン間のデータ同期も問題なく実施されていることを確認した。

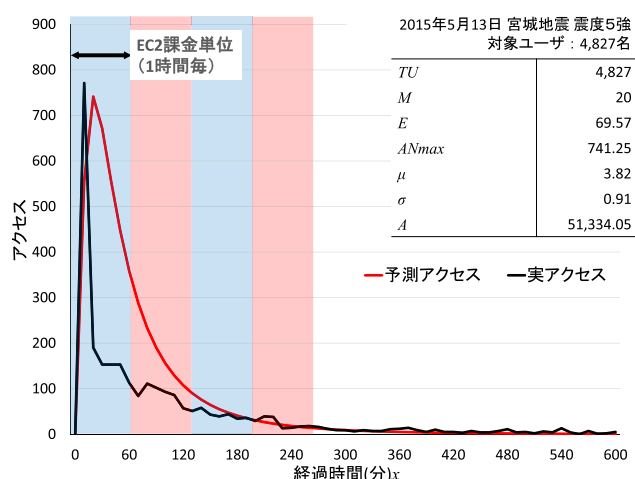


図 14 実災害とアクセス予測モデルの分布

Fig. 14 Distribution of disaster and access prediction model.

5.3 アクセス予測に基づくオートスケーリング機構に関する評価

5.3.1 アクセス予測モデルでのサーバ数算出

実災害のアクセス分布に対してアクセス予測モデルを用いてサーバ数を算出した。図 14 は 2015 年 5 月 13 日に発生した宮城県沖地震の実際のアクセス分布と、対象ユーザ数 4,828 名に対してアクセス予測モデルを用いて算出した 10 分単位の予測アクセス分布である。EC2 は 1 時間単位で課金されるため、1 時間単位の予測アクセス分布に対して 1 台のサーバ許容量を基にサーバ数を算出する。図 14 の 0~60 分の間では最高 741 アクセス/10 分間が予測される。t2.small は 200 アクセス/10 分間の許容能力があるため、t2.small を各 AZ に 2 台ずつ合計 4 台で負荷分散を行う。ここで 0~60 分の間で実アクセスとモデルにわずかに差があるが、t2.small の単体の許容能力内もしくは負荷分散に用いる合計台数の許容能力内でこの差を吸収可能なため、モデルの精度としては妥当だと考える。次の 60~120 分間では最高 358 アクセスが予測されるため、各 AZ に 1 台ずつ合計 2 台で負荷分散を行う。また、アクセス予測モデルは災害対象ユーザに対するシステムへのアクセス分布を予測するが、システムへのアクセスが災害以外の事態で行われた場合は予測の対象外となる。例として、インフルエンザによるパンデミック時の情報共有などがあげられ、この際 CPU 使用率が 20% 超となった場合はアクセス予測モデルとは別にサーバを追加し負荷分散を実施する。

5.3.2 シミュレーション評価

実災害のアクセスに対して負荷分散のシミュレーション評価をした。評価内容は、図 14 の実災害アクセスを基に JMeter でテストシナリオを作成し、評価サーバから提案システムへ実行した。図 14 では発生から 1 時間単位の最高アクセスは、0~60 分間で 771 アクセス/10 分間、60~120 分間で 111 アクセス/10 分間、120~180 分間で 58 アクセス/10 分間であり、テストシナリオはアクセスピーク

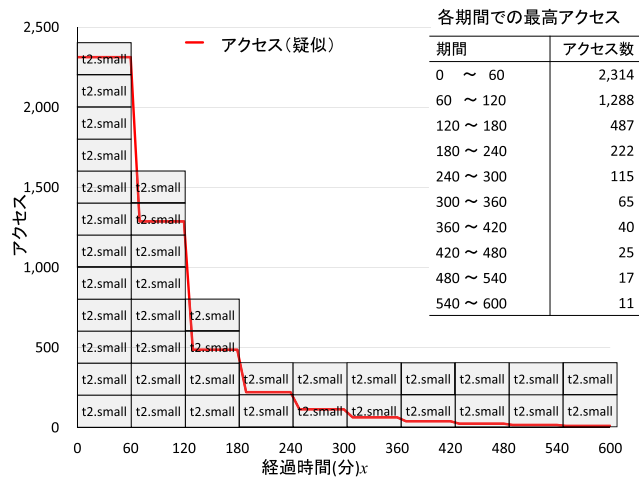


図 15 20,000 名ユーザーでのオートスケーリング

Fig. 15 AutoScaling of the 20,000 users.

を 20 分に設定し、その後の各時間帯の最高アクセスを 1 時間継続するよう作成する。このテストシナリオは各時間帯で実災害の最高アクセスを継続する疑似アクセスとなるため、実災害アクセスのシミュレーション評価が可能となる。図 14 のパラメータをアクセス予測モデルに与えテストシナリオを実行した結果、対象ユーザー 4,827 名の実災害に対して、0~60 分間では t2.small を 2 台スケールアウトし合計 4 台、60~120 分間では 2 台スケールインし合計 2 台、120~180 分間では標準構成の合計 2 台で負荷分散を確認した。負荷分散中の各サーバの CPU 使用率は、t2.small の許容値である 20%以下を推移し、想定内の CPU 使用率で負荷分散実行を確認した。また、アクセス予測に基づくオートスケーリング機構の各処理時間は、気象情報サービスでの災害検知からアクセス予測モデルでアクセス予測およびサーバ数算出に約 10 秒程度、サーバ起動からスケールアウト完了に約 5 分程度、対象ユーザー 4,827 名へのメール送信に約 3 分となり、メール送信 20 分後のアクセスピーク前にスケールアウトが完了し負荷分散が問題なく行えることを確認した。

次に、筆者が所属する企業でサービス提供している安否システムの顧客 20,000 名に対して、シミュレーション評価をした (図 15)。図 15 から対象ユーザー 20,000 名では、0~60 分間では t2.small 12 台、60~120 分間では 8 台、120~180 分間では 4 台と、図 14 の対象ユーザー 4,827 名と比較しより大規模なスケールアウト・インとなる。つまり対象ユーザー数が増加するほどオートスケーリングの規模が大きくなるため、顧客数増加にともない費用対効果の向上が期待できる。

6. 考察

6.1 対数正規分布の仮定

災害発生後の安否システムへのアクセス分布は図 9 から対数正規分布に従うと仮定した。対数正規分布への仮定の

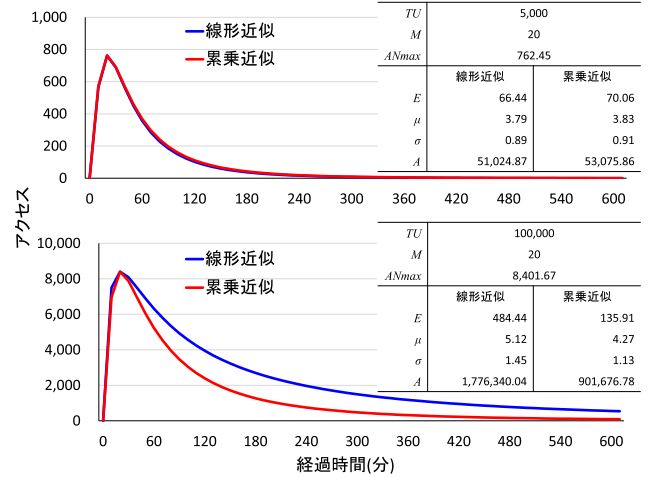


図 16 線形近似と累乗近似

Fig. 16 Linear approximation and power approximation.

背景は、図 9 のアクセス分布が Q-Q プロットにおいてほぼ直線に乗るためであるが、シャピローウィルクでの正規性検定では P 値は 0.05 以下となり正規性があるとはいえない。ここで図 9 の Q-Q プロットが直線から外れるのは 7.6 付近からとなる。7.6 はアクセス時間の対数をとった値なので $Exp(7.6) \approx 2,000$ (秒) となり災害発生後約 30 分後となる。図 9 の災害では発生から 30 分後ではアクセスが少なくなっており、アクセス予測モデルでのオートスケーリングを行うことなく標準構成のサーバ数でアクセスを許容可能である。つまり災害発生後、システムへのアクセス分布は最繁忙期間において対数正規分布に従っており、この期間において適切にアクセス予測ができれば安否システムにおける災害時の負荷分散は問題ないと考える。以上が厳密な正規性ではないものの対数正規分布を基にアクセス予測モデルを構築した理由である。また現状の調査では対象ユーザー数が多いほどアクセス分布が Q-Q プロットにおいて直線になる傾向があり、たとえば数万名規模であれば正規性が認められる可能性がある。対象ユーザー数が多いほど対数正規分布に近づけばアクセス予測モデルの精度を向上させることができ、さらに適切なサーバ数算出が期待できる。

6.2 平均値 E の決定

アクセス予測モデル構築過程において平均値 E は、対象ユーザー数 TU と D (平均値 E と最頻値 M の差) の関係 (図 10) から近似式で決定した。ここで採用した近似式は累乗近似である。一方、対象ユーザー数 TU と D が線形的な関係であれば線形近似も考えられる。図 10 の両者の相関係数は $R = 0.77$ となり高い相関があった。図 16 は対象ユーザー数 5,000 名と 100,000 名で、累乗近似および線形近似を用いて平均値 E を算出した場合の予測アクセス分布である。5,000 名では累乗近似と線形近似に目立った差はないが、100,000 名では線形近似のほうがピークからの

減衰が緩やかであり、対象ユーザ数が多いほどこの傾向は顕著になる。この結果だけ考えれば対象ユーザ数に閾値を設定し累乗近似もしくは線形近似を選択することは可能だが、現状のデータ量では両者の挙動差に対する根拠に乏しいため現時点での断定は避けた。またサーバ数算出であれば線形近似のほうが台数削減は緩やかになりシステムリソースに余裕が生まれ安定性が向上するが、過剰な余裕は費用向上にもつなげるため安定性と費用面を考慮したうえで線形近似もしくは累乗近似を選択しなければならない。両者の挙動を明らかにするために今後のデータ収集および分析が必要である。

6.3 運用費用

現在サービス提供中の 20,000 名ユーザに対して、平常時費用と災害時費用について各システムの費用比較を行った。比較対象は、(1) 現行システム、(2) 平常時は最小構成、災害時は対処可能サーバに変更、(3) 提案システムの 3 システムである。(2) の平常時は (3) と同構成とし、災害時は、過去の災害・訓練データから求めた最大アクセス数を許容するサーバを用いる形態である。使用する EC2 タイプ、費用は表 5 に示し、各システムの平常時のサーバ構成および年間費用は表 6 である。表 6 から (1) は過去の災害のアクセス結果から余裕のある c3.2xlarge を常時用いているため WEB・DB 合計費用は \$17,070 となるが、(2) や (3) は、平常時は最小構成のため \$11,570 となり年間費

用を約 32%削減できる。ここで (2) と (3) は、平常時は最小構成、災害時に負荷分散環境を増強する点において同じ立場をとるが、両者には災害時のサーバ数算出手法に違いがある。

表 7 は対象ユーザ数 20,000 名での災害対処費用であり、(2) は現行システムと同様の c3.2xlarge 2 台を 24 時間起動、(3) は図 15 のアクセス予測モデルでの算出である。表 7 では (2) の対処用サーバを c3.2xlarge としたため、(3) の t2.small と比較し差が生じているが、両者の本質的な違いはアクセス予測の有無である。(2) は、実際に経験した過去の災害・訓練のユーザ数に対してはアクセス結果から必要サーバ数の事前算出が可能だが、未経験のユーザ数に対してはアクセス結果を得ていないため必要サーバ数の事前算出が困難である。一方、(3) 提案システムは、過去の災害・訓練のアクセス分布を統計的に分析し構築したアクセス予測モデルで、経験・未経験によらず対象ユーザ数に応じたアクセス予測が可能のため、経験済み対象ユーザ数のアクセス結果を基にする (2) と比較して優位性がある。また安否システムは新規顧客参入にともないユーザ数が増加していくため、ユーザ数変化のつど、あらかじめ訓練などで必要サーバ数を求めておくことは相応の手間が発生するが、アクセス予測モデルではその手間は不要となる。採用 EC2 タイプは、今回は t2.small としたが、各タイプの性能特性や費用を考慮すれば、さらに精度の高い負荷分散や費用効果が期待できる。

表 5 c3.2xlarge と t2.small の稼働費用

Table 5 c3.2xlarge, t2.small cost.

	1 時間	月間	年間
c3.2xlarge	\$0.478	\$355.6	\$4,267.5
t2.small	\$0.034	\$25.2	\$303.5

表 7 (2)、(3) の災害対処費用

Table 7 Disaster handling cost of (2), (3).

	(2)	(3)
サーバ数	c3.2xlarge : 2 台	t2.small : 18 台
費用	\$22.9	\$1.8

表 6 平常時のサーバ構成および年間費用

Table 6 Server composition of the normal situation and 1 year cost.

	(1) 現行システム	(2) 平常時は最小構成、災害時は対処可能サーバに変更	(3) 提案システム
メインリージョン	WEB : c3.2xlarge : 2 台 DB : c3.2xlarge : 2 台	WEB : t2.small : 2 台 DB : c3.2xlarge : 2 台	WEB : t2.small : 2 台 DB : c3.2xlarge : 2 台
サブリージョン 1	—	WEB : t2.small : 2 台 DB : t2.small : 2 台	WEB : t2.small : 2 台 DB : t2.small : 2 台
サブリージョン 2	—	WEB : t2.small : 2 台 DB : t2.small : 2 台	WEB : t2.small : 2 台 DB : t2.small : 2 台
WEB 費用	\$8,535	\$1,821	\$1,821
DB 費用	\$8,535	\$9,749	\$9,749
WEB・DB 合計費用	\$17,070	\$11,570	\$11,570
(1) からの費用削減率	—	WEB : 78.6%, DB : 114.2% WEB・DB : 32.2%	WEB : 78.6%, DB : 114.2% WEB・DB : 32.2%

7. まとめ

本研究ではWEBシステムを構成するサーバ群を世界規模で冗長化し、災害時のアクセス数を事前予測し適切なサーバ数で負荷分散を行う広域冗長型安否システムを提案し評価結果から有効性を示した。過去の災害のアクセスデータに対して対数正規分布を仮定したアクセス予測モデルでは、過去の災害データに厳密な正規性は見られないもののアクセス減衰期までは対数正規分布に従っており、実運用時のサーバ数算出において期待した効果を得られた。

今後の課題として、アクセス予測モデルの精度向上があげられる。本研究ではアクセス予測モデル構築の概念と基礎評価に対して一定の効果が認められたが、モデル構築の際に分析した実災害データが乏しいためモデル妥当性の根拠も同様に乏しい。今後はより多くの実災害データを収集しモデル精度向上および修正を行う予定である。また今回は対象ユーザ数のみのアクセス予測を行ったが災害時のその他の要素（発生時刻、対象組織の特性など）がアクセス分布に及ぼす影響を調査する予定である。

参考文献

- [1] 長谷川孝博, 井上春樹, 八巻直一: 低コスト運用でユーザフレンドリーな安否情報システムの開発, 学術情報処理研究誌, No.13, pp.91-98 (2009).
- [2] 梶田将司, 太田芳博, 若松 進, 林 能成, 間瀬健二: 高等教育機関のための安否確認システムの段階的構築と運用, 情報処理学会論文誌, Vol.49, No.3, pp.1131-1143 (2008).
- [3] 白井真人, 畑山満則, 福山 薫: 地域コミュニティでの情報システムを用いた安否確認に関する研究, 地域安全学会論文集, No.16, pp.11-20 (2012).
- [4] 白鳥則郎, 稲葉 勉, 中村直毅, 菅沼拓夫: 災害に強いグリーン指向ネバーダイ・ネットワーク, 情報処理学会論文誌, Vol.53, No.7, pp.1821-1831 (2012).
- [5] 西谷 薫, 杉浦彰彦: ワンセグ用データ放送を用いた災害時安否情報配信, 情報処理学会論文誌, Vol.50, No.2, pp.839-845 (2009).
- [6] 大瀧 龍, 重安哲也, 浦上美佐子, 松野浩嗣: 自律的無線ネットワークを用いた被災情報提供システム—被災地域の地形を考慮した無線ノード置局アルゴリズムの提案, 情報処理学会論文誌, Vol.52, No.1, pp.308-318 (2011).
- [7] 小山 由, 水本旭洋, 今津真也, 安本慶一: 大規模災害時の安否確認システムと広域無線網利用可能エリアへのDTNに基づいたメッセージ中継法, 情報処理学会研究報告, 2012-MBL-62, No.29, pp.1-7 (2012).
- [8] 東田光裕, 林 春男, 松下 靖, 三宅康一: 社会サービスとしての被災者対応の質を向上させる情報マネジメントシステムの構築—QRコードを利用した安否情報収集システムの開発, 地域安全学会論文集, No.9, pp.147-156 (2007).
- [9] 越後博之, 湯瀬裕昭, 干川剛史, 沢野伸浩, 高畑一夫, 柴田義孝: 大規模分散環境におけるロバストネスを考慮した広域災害情報共有システム, 情報処理学会論文誌, Vol.48, No.7, pp.2340-2350 (2007).
- [10] 太田芳博, 梶田将司, 林 能成, 若松 進: 名古屋大学安否確認システムの構築と運用, 電子情報通信学会技術研究報告, IA, Vol.108, No.409, pp.77-82 (2009).

- [11] Amazon Web Services (online), available from <http://aws.amazon.com/> (accessed 2015-08-30).
- [12] Microsoft Azure (online), available from <http://azure.microsoft.com/> (accessed 2015-08-30).
- [13] Service Health Dashboard (online), available from <http://status.aws.amazon.com/> (accessed 2015-08-30).
- [14] Route53, Amazon Web Services (online), available from <http://aws.amazon.com/route53/> (accessed 2015-08-30).
- [15] 一般財団法人気象業務支援センター, 入手先 <http://www.jmbc.or.jp/> (参照 2015-08-30).
- [16] Elastic Load Balancing, Amazon Web Services (online), available from <http://aws.amazon.com/elasticloadbalancing/> (accessed 2015-08-30).
- [17] EC2, Amazon Web Services (online), available from <http://aws.amazon.com/ec2/> (accessed 2015-08-30).
- [18] 藤田靖征, 村田正幸, 宮原秀夫: Web サーバシステムのモデル化と性能評価, 電子情報通信学会論文誌 B, Vol.J82-B, No.3, pp.347-357 (1999).
- [19] 石原 進, 岡田 稔, 岩田 晃, 櫻井佳一: イベント駆動方式による LAN 通信量解析モデル, 電子情報通信学会論文誌 A, Vol.J78-1, No.8, pp.961-964 (1995).
- [20] Antoniou, I., Ivanov, V.V., Ivanov, V.V. and Zrelov, P.V.: On the log-normal distribution of network traffic, *Physica D: Nonlinear Phenomena*, Vol.167, No.1, pp.72-85 (2002).
- [21] Murta, C.D. and Dutra, G.N.: Modeling HTTP service times, In *Global Telecommunications Conference, GLOBECOM'04*, Vol.2, pp.972-976, IEEE (2004).
- [22] 稗園泰彦, 上村郷志, 小頭秀行, 中村 元: 一斉報知を用いた遅延発呼制御方式におけるサーバ同時接続数の安定化に関する一考察, 電子情報通信学会論文誌 B, Vol.J95-B, No.3, pp.414-424 (2012).



永田 正樹 (学生会員)

2012 年静岡大学大学院工学研究科修士課程修了。同年静岡大学創造科学技術大学院自然科学系教育部情報科学専攻博士課程進学。株式会社アバンセシステム勤務 (社会人学生)。サーバ負荷分散, アクセス予測, 並列処理等の

研究に従事。



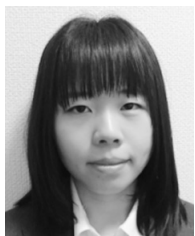
阿部 祐輔

株式会社アバンセシステム勤務。2009 年静岡大学大学院理学研究科修士課程修了。安否情報システム「ANPIC」, 一斉メール配信システム, 教育支援システム等の WEB 開発およびデータ解析に従事。



金原 一聖

株式会社アバンセシステム勤務（2015年まで）。安否情報システム「AN-PIC」、一斉メール配信システム、教育支援システム等のWEB開発に従事。



福井 美彩都

株式会社アバンセシステム勤務。安否情報システム「ANPIC」、一斉メール配信システム、教育支援システム等のWEB開発に従事。



峰野 博史 （正会員）

1999年静岡大学大学院理工学研究科修士課程修了。同年日本電信電話(株)入社。NTT サービスインテグレーション基盤研究所を経て、2002年10月より静岡大学情報学部助手、博士(工学)。2011年4月より静岡大学情報学部准教授。モバイルコンピューティング、センサネットワーク応用システムに関する研究に従事。