

インターネットによるバイオテロ情報の
公開に関する研究

2006年度報告書

慶應義塾大学
環境情報学部
中村 修

1 インターネットによるバイオテロ情報の公開における留意点

1.1 大量なアクセスに対応したサーバとネットワーク

バイオテロに関する情報へのアクセスは、関連する事象が発生したときには、通常のアクセス数からは予想できない程の数になることが予想され、このことを想定したシステム構築をおこなわなければ、システムが機能しなくなることが予想される。事実、米国で炭疽菌事件が発生した際、米国 CDC の Web サーバは、炭疽菌に関する情報を求める利用者のアクセスにより、システムに障害が発生した。現在 CDC は、システムの分散化および商用コンテンツ配信サービス (Akamai) を利用し、利用者からのアクセス数のスケールビリティを確保している。

参考資料 (GCN: Government Computer News, March 4, 2002 Load Balancing) の記事によれば、CDC の Web には、月平均で 400 万利用者 (2001-2002 年のユニークな利用者数の平均) がアクセスをしていて、炭疽菌事件の後、10 月には 910 万人、11 月は 520 万人、12 月は 700 万人の利用者からアクセスがあった。また、Akamai を利用することにより、0.44 秒のレスポンスタイムを維持している。

日本のインターネット利用者数は、総務省通信白書によりば 2005 年で 8,529 万人 (前年比 581 万人増) である。この値から推測すれば、ここ数年で 1 億人の利用者数を視野に入れたシステムを検討すべきである。

米国の炭疽菌事件と同様なことが発生したときの Web サーバへのアクセス数を予測することは、なかなか難しいが、米国炭疽菌事件が起きた頃のアメリカのインターネット利用者数が、約 1 億 6000 万だとすると、約 5% の利用者がアクセスしたことになる。日本の利用者数が約 1 億人であると想定すると、その 5% すなわち 500 万人の利用者が短期間に Web サーバへのアクセスをおこなうことになる。

現在は、5 年前よりもインターネットで情報検索をする習慣が広がっていることなどを考えると、アクセス数集中の度合いは上がる可能性が高いため、アクセスの集中を考慮した情報公開システムを開発する必要性が高まっている。具体的には、1 台のコンピュータでの情報公開は、演算装置・記憶装置・ネットワークインターフェースの性能に限界があるため現実的ではなく、複数のコンピュータによる情報公開システムを構築する必要がある。また、設置拠点についても、その拠点の堅牢性や拠点へのネットワーク接続性を考慮して、複数拠点への分散設置が望ましい。

1.2 コンテンツの更新と整合性

バイオテロ関連の事象が発生した際に、情報公開において最も重要になるのがコンテンツを適切に更新していくことである。コンテンツ更新の技術としては、適宜コンテンツを更新するための仕組みとしてコンテンツマネジメントシステム (CMS) と分散拠点でのコンテンツの整合性を保つ仕組みが必要となる。

コンテンツマネジメントシステムは、情報公開サーバ上での情報編集を統合的におこなうことのできる仕組みであり、大規模な Web サービスのコンテンツ管理に一般的に用いられている。特にバイオテロ関連の情報について考えた場合、事象の検知や情報の利用は主に医療関係者となるため、コンピュータが専門でない者が不自由なく扱えるユーザインターフェースを持つコンテンツマネジメントシステムの開発が重要となる。同時に、複数の作業者が同時にコンテンツの更新をおこなえることが望ましい。バイオテロとの関連が既に想定されている情報については、これらを利用する際に必要となる項目を検討し、情報の正規化をおこなうことによって、コンテンツマネジメントシステムでの管理を実現することができる。しかしながら、事前に想定されていない情報や、動的な統計収集のためのコンテンツについては、今後も検討を続ける必要がある。

一方で、複数のネットワーク拠点に複数の情報公開サーバを設置する際には、それぞれのサーバが公開している情報の整合性を保つ必要がある。分散拠点から情報公開をする際に、それぞれのサーバが公開している情報の同一性を保証することは重要である。具体的には、データベースのクラスタリングの機能を使う方法や、分散ファイルシステムを用いる方法、更新があったタイミングにネットワークでのファイルコピーをおこなう方法等があり、コンテンツマネジメントシステムやサーバの設置拠点のネットワーク状況に応じて適切な手法を検討する必要がある。

1.3 セキュリティ

特にバイオテロ関連に特化したことではないが、情報公開サーバへのアクセスについて、暗号化を用いないアクセス手法を用いるべきではない。具体的には、リモートログインには SSH、Web インターフェースでの情報編集用アクセスには SSL を用いることは必須である。最近では、暗号化通信路を利用していても認証方法がパスワードである場合には、総当たり攻撃 (Brute Force Attack) によって危険にさらされるので、SSH では公開鍵認証方式、Web インターフェースでは個人証明書による認証方式にすることが望ましい。

また、発信する情報が静的なものではなく、逐次更新されることが予想されるため、公開する情報には、情報公開サーバでの証明書提示や情報の作成・更新日時の挿入、デジタル署名の手法について検討をおこなう必要がある。

さらに、ネットワークでのサービス不能攻撃 (DoS) や脆弱性を突いた攻撃に対する対策を検討することも必要であり、システム設計としてこれらへの対策を内包するのか、他のサービスとの組み合わせて考えるのかということが大切なポイントである。

2 既存のシステムに関する調査

2.1 Akamai

米国 Akamai 社が世界規模での情報公開をおこなう際のデータ転送の効率を上げるためにおこなっているサービスである。Akamai 社では、インターネットでの情報転送をファーストマイル・ラストマイル・ピアリングポイント・バックボーンに分類したモデル化をおこなっている。ファーストマイルとは、Web サイトのインターネット接続状況のことであり、ラストマイルとはユーザのインターネット接続状況のことである。また、ピアリングポイントとは、ISP 同士の接続をしている拠点のことであり、バックボーンとは ISP 内のバックボーンのことである。

Akamai 社では、インターネットのどこへでも効率良くデータ転送をおこなうための基盤をコンテンツ公開者がおこなうことは、コスト高であるとの観点から、世界中どこにでも高速な接続性が確保できるコンテンツディストリビューションのプラットフォームを提供している。Akamai 社の情報提供は、オリジンサーバとエッジサーバにより実現されており、オリジンサーバに置かれたコンテンツをユーザの好きなタイミングでスナップショットとしてエッジサーバへコピーすることで実現されている。ユーザはネットワーク上で理論上、最も近いエッジサーバへ誘導されるため、常に最善の性能を得ることができる。

利用者が多くの拠点に分散している場合、情報公開をおこなうために世界中に拠点を設けることはコストの観点から現実的でないため、米国 CDC や米国海軍などもこのサービスを利用した情報公開をおこなっている。

2.2 大規模分散サーバクラスタ技術

近年、インターネットでの情報システムは突発的な大規模アクセスを想定した設計をおこなうことが必須であり、このため特定のサービスへのアクセスを複数のサーバへ分散する負荷分散装置 (レイヤ 4 スイッチやレイヤ 7 スイッチ) を用いることが多い。これらは、1 台のコンピュータの性能では多くても数千同時アクセスを処理することが限界であることに対する技術であり、特に Web やストリーミングの分野では多く用いられている。例えば、レイヤ 4 スイッチでは、Web サーバが主に TCP の 80 番ポートでサービスをおこ

なっていることを利用し、宛先が 80 番ポートのパケットを複数の Web サーバへ分散して転送することにより、同時アクセス数を増やすことを実現している。この手法では、サーバの台数を増やし、クラスタ化することでそれに応じた性能の向上を得ることができると、一般的な手法として広く用いられている。

しかし、大規模サーバクラスタを 1 拠点に設置した場合にも、その拠点のネットワーク接続性がボトルネックとなり、性能向上の限界となる場合がある。そのような場合には、サーバクラスタを分散配置することにより、サーバへの負荷分散と同時にクラスタへの負荷分散をおこなう。多くの検索エンジンやショッピングサイトはこの手法により、同時アクセス数への対策と堅牢性の確保を実現している。

2.3 IBM がおこなったオリンピックでの事例

米 IBM 社が、1996 年アトランタ夏季オリンピック、1998 年長野冬季オリンピックにおいて、インターネットへの情報公開をおこなった手法は、先駆的なものであった。当時、Web サーバへのアクセス数を 100,000,000 ヒット/日 (1157 ヒット/秒) と予測し、システム設計をおこなった。1994 年リレハンメル冬季オリンピックの際に同社がおこなった情報公開では、当時 IP 接続されていた 64 カ国のうち 42 カ国から 1,300,000 ヒット/日のアクセスがあり、1998 年は 171 カ国が IP 接続されていることからの予測である。また、当時の Yahoo! のアクセス数は 5,000,000 ヒット/日であるからオリンピックでの情報公開は非常に大きな規模のものであったといえる。

構築された情報システムでは、主に Web を使った競技結果を含むさまざまなコンテンツの配信やラジオ放送のインターネットへの生中継など、オリンピック会場で起こった事柄をより早く世界へ向けて公開するものであった。ネットワーク構成としては、世界規模のバックボーンを持つ IBM 社の OpenNet を東京の NSPIX-2(現 DIX-IE) をはじめとする世界中の IX(Internet eXchange Point) でインターネットへ接続し、世界 5 拠点に設置されたサーバクラスタへ単一 IP アドレスアーキテクチャを用いて負荷分散を実現した。単一 IP アドレスアーキテクチャでは、世界中に設置されたサーバに全て同じ IP アドレスを割り振りクライアントは、ネットワークの経路制御に基づき最も近いサーバへ誘導される。この技術は IPv6 により、Anycast として定義されており、負荷分散と同時に障害時のバックアップサーバへの動的な切り替えも可能とするものであった。

分散拠点でのコンテンツの整合性は、DFS(Distributed File System) によって実現された。DFS は、米国カーネギーメロン大学の研究成果である AFS の流れを汲む分散型ファイルシステムである。世界 5 拠点に設置された情報公開サーバがそれぞれ DFS によりコンテンツを共有することで、DFS のキャッ

シュ機能が有効に機能し、大元の DFS サーバへのアクセスを全体の約 5% に抑えることができたため、負荷分散の目的を十分に果たしたと言える。

3 動的コンテンツの提供手法に関する評価システム

バイオテロ関連の情報公開を検討するために、動的コンテンツの提供手法を考慮した評価システムの開発をおこなった。本システムは、バイオテロ関連の情報のうち、事前に想定できる感染症についての情報を医療関係者との議論により分類をおこない、コンテンツマネジメントシステムによる情報管理をおこなえる形で開発した。

3.1 症例データの正規化

想定される感染症について、システム化するために各症例について情報の分類を検討した。検討をおこなった症例を以下に挙げる。

- 炭疽
- 天然痘
- 野兔病
- ウイルス性出血熱
- ボツリヌス症
- ペスト
- 鼻疽・類鼻疽
- ブルセラ症
- 消化管感染症
- 重症急性呼吸器症候群 (SARS)
- ウエストナイル熱／脳炎
- 狂犬病
- コクシジオイデス症
- 多剤耐性結核菌
- Q 熱

これらの疾病について、どのような項目の情報が公開されるべきかについて検討し、以下の 4 つに分類した。

バイオテロが疑われる状況と対応

フローチャートによる対応マニュアル

サマリー

比較的短い時間で概観するための情報

詳細

より詳しい事項が確認できる情報

画像の一覧
症例画像

これらをシステム上に実現するにあたり、各分類について、項目の分類をおこなった。

「バイオテロが疑われる状況と対応」については、フローチャートによる対応マニュアルを作成し、図として保存することとした。

「サマリー」に含まれる項目を表1に示す。

表 1: サマリーに含まれる項目

項目	情報形式
病原体の特徴	テキスト
分類と潜伏期間	テキスト
感染経路	テキスト
臨床症状	テキスト
検査所見 (Q 熱のみ)	テキスト
検体の種類および採取法	テキスト
検体の輸送法	テキスト
微生物学的検査法	テキスト
感染症における取り扱い	テキスト
患者の隔離や汚染器材等の管理	テキスト
治療の要点	テキスト
抗菌薬の予防投与	テキスト

「詳細」に含まれる項目を表2に示す。

表 2: 詳細に含まれる項目

項目	情報形式
はじめに (狂犬病のみ)	テキスト
病原体の特徴	テキスト
主な臨床像	テキスト
臨床検査所見	
血液生化学検査	テキスト
画像検査その他	テキスト
確定診断	
検体の採取、輸送、保存など	テキスト
微生物学的検査法	テキスト
微生物学的検査結果に基づく症例定義	テキスト
治療	
薬物療法 (抗菌薬療法)	テキスト
その他治療上の留意点	テキスト
予防 (ワクチン)	テキスト
バイオハザード対策	
患者の隔離	テキスト
検体、菌、汚染器材等の取り扱い	テキスト
感染症法における取り扱い	テキスト
参考文献	テキスト

「写真の一覧」については、症例の資料画像とテキストによるキャプションとして保存することとした。

これらの項目の分類をもとに、症例ごとの情報公開をおこなう評価システムの開発をおこなった。一方で、症例ごとに項目分類に多少の差が残っており、引き続き検討をおこなう必要性がある。

3.2 コンテンツマネジメントシステム

検討した項目分類を Web インターフェースで編集するために、コンテンツマネジメントシステムの開発をおこなった。本システムを用いて分類した項目毎にデータを入力または編集することにより、コンテンツの更新をおこなうことが可能である。

図 1 に、コンテンツマネジメントシステムでのテキスト情報部分の編集画面を示す。「疾患」・「分類」・「タイプ」・「セクション」をメニューから選択することにより、任意の項目の情報を編集することが可能である。



図 1: テキスト情報部分の編集画面

図 2 に、コンテンツマネジメントシステムでの画像情報部分の編集画面を示す。「疾患」・「画像番号」をメニューから選択することにより、任意の項目

の画像とキャプションを編集することが可能である。

バイオテロ対策マニュアル

TOP はじめに 総論 疾患の鑑別 疾患の一覧 画像の一覧 資料リンク

画像の編集

00003: 炭疽

00005: NO TITLE

Caption:
NO TITLE

Source:
CDC
(<http://pathmicro.med.sc.edu/ghaffar/anthraxd4.jpg>)

Comment:
53 year old female, employed 10 years in the spinning department of a goat-hair processing mill. Cutaneous anthrax lesion on right cheek; lesion as seen on 4th day

参照...

JPEG

読み込み 更新 削除

クエリ送信



4日目 8日目 13日目

図5 NO TITLE
53 year old female, employed 10 years in the spinning department of a goat-hair processing mill. Cutaneous anthrax lesion on right cheek; lesion as seen on 4th day

CDC
(<http://pathmicro.med.sc.edu/ghaffar/anthraxd4.jpg>)

図 2: 画像情報部分の編集画面

本評価システムの開発により、項目の分類によってコンテンツマネジメントシステムによるバイオテロ関連の症例情報管理を実現したが、想定される情報の公開について、情報の更新があった場合の更新履歴の表示方法や、想定されていないコンテンツの公開について継続して検討する必要がある。これらの検討を終え、システム全体の設計をおこなうことのできる段階へ進むことが急務である。