

## BoF セッション「クラウドって正直どう？」実施報告

オーガナイザ：渡辺知恵美（お茶の水女子大学 大学院人間文化創成科学研究科）

進行補佐：上田真由美（京都大学 学術情報メディアセンター）

パネリスト（敬称略，順不同）：

宮崎純（奈良先端科学技術大学院大学 情報科学研究科）

小林隆志（名古屋大学 大学院情報科学研究科）

神崎映光（大阪大学 大学院情報科学研究科）

平手勇宇（早稲田大学 メディアネットワークセンター）

### 概要

近年インターネット上にグローバルに拡散したコンピューティングリソースを使って，利用者に情報サービスやアプリケーションサービスを提供する「クラウドコンピューティング」が大変注目を集めている．しかしながらクラウドコンピューティングとは何かという解釈が人や場合によって異なったり，クラウドコンピューティングという概念の登場によってデータ工学として何か新たな研究テーマが生まれたのかというところが明確になっているわけではない．

そこで本セッションではクラウドコンピューティングにおける以下の2点についてパネリストおよび参加者の思うところを議論した．パネリストには上記の4名を迎え，また30名程度の参加者を交え，以下の2点を議題にし，活発な議論が展開された．

#### 議題1：クラウドとは何か

「クラウド」の定義や異義を改めて議論した．各パネリストや各参加者にとって何をクラウドと  
いうのか，もしくはクラウドコンピューティングに係る技術として注目するトピックについて述べた．

#### 議題2：クラウドにおけるデータ工学と情報マネージメント的な研究のしどころ

クラウドコンピューティングに係わる諸技術の中で，データ工学および情報マネージメントからの切り口で興味深いと思われる特徴やこれから取り組むべき課題について議論した．

本稿ではこれらの議題で挙げられた話題について紹介する．

#### 議題1：クラウドとは何か

クラウドコンピューティングの定義はWikipediaによると「インターネットを基本にした新しいコンピュータの利用形態である．利用者はコンピュータ処理を，ネットワーク（通常はインターネット）経由で，サービスとして利用できる．」となっている．この記述自体非常に曖昧であり，その解釈は人によって異なる．クラウドという言葉に関連するキーワードとして以下のものがあげられる．

(1) 広域および大規模分散環境：

Google や Yahoo!などの検索エンジンのように、世界中に分散された広域かつ大規模な分散コンピューティング環境を雲の中の処理にたとえてクラウドと呼ぶことがある。これに関連するキーワードとして、グリッドコンピューティング、P2P ネットワーク（広域に分散するという意味で）、ユビキタスコンピューティングなどがあげられる。

(2) サービス指向アーキテクチャによる多様なレベルのコンピューティング環境の提供

利用形態の視点から、(1)により整備された環境を利用して個人や中小企業などの高性能計算機資源を持たない利用者に対し、ネットワークを介して多様なレベルのコンピューティング環境をサービスとして提供することを「クラウドコンピューティング」と称する場合もある。Amazon Web Service[1]や Google App Engine[2], Windows Azure[3]などが有名なサービスとしてあげられる。関連するキーワードとして、○○ as a Service (Infrastructure as a Service, Hardware as a Service, Database as a Service, Software as a Service), サービスの仮想化, パブリッククラウド/プライベートクラウド, サービス指向アーキテクチャ (SOA), アウトソーシングがあげられる。

まず議題 1 として「クラウド」についての定義や異義, 注目する技術について各パネリストによって以下の意見が得られた。

・このような技術も「クラウド」ではないか, という提言

まず神崎氏よりクラウドコンピューティングというものが「ネットワークの雲の中から何かしらを享受する」のであれば, 神崎氏の専門である P2P ネットワークやモバイルアドホックネットワーク, センサネットワークもクラウドと呼んでよいだろうかという疑問と提言が提示された。ただし, クラウド環境における利用者側を「こちら側」サービス側を「あちら側」と呼ぶ場合, P2P ネットワークやモバイルアドホックネットワークの場合はそれらの区別は明確でなく, ネットワークに参加する各ノードはいわゆる「こちら側」であり同時に「あちら側」でもあるところが異なる点であるとした。



図 1：セッションの様子

#### ・運用を「外注」することの誘惑と問題点

平手氏は高性能な計算環境の提供という面ではスーパーコンピュータやグリッドなど従来から既に存在するものであるが、従来の数値計算やエンタープライズ用途など特化した利用だけでなく、よりサービスが一般化し利用者の幅が広がった点がクラウドと呼ばれるものなのではないかとの見解を述べた。

平手氏が参加している研究プロジェクト「インターネット上の知識集約を可能にするプラットフォーム構築技術」[4]では現在144億のWebページを収集し管理している。そこで必要となる大規模なコンピューティングリソースについて、その運用管理のコストが研究そのものを圧迫するほどかかることをあげ、その運用管理を「外注」する魅力に関して氏の実体験に基づいた率直な意見が述べられた。しかしながら外注をした場合研究の一部を外注先に依存することになる危険性についても言及された。この部分は参加者にとっても非常に興味のある部分であり活発な議論が得られた。たとえばハードウェア面を外注する場合、それらの運用に関するノウハウが蓄積されなくなり、委託した企業にそれらのノウハウが集中することになる。ソフトウェア面でもたとえば検索エンジンAPIを利用する以上その検索エンジンのアルゴリズムに依存することになる。つまり外注することによって相手への依存が断ち切れなくなる（ベンダーロックイン状態になる）ことを十分に注意すべきとの参加者内での共通見解が得られた。

#### ・「クラウド」という言葉でごまかさず本質を見るべき

小林氏および宮崎氏は安易にクラウドコンピューティングという言葉を使うことについての異議を述べた。宮崎氏はクラウドコンピューティングの形態は1990年初頭の分散OS Plan9[5]の目指すものと何ら変わりがなくそれが「クラウド」という新しいキーワードに置き換えられただけで特に新しいものがあるわけではないと主張した。小林氏は「クラウド」という言葉を使うことによってなんとなく現在注目されている新しい技術であると錯覚を覚えてしまうが、冷静に実態を眺めれば従来より続く分散処理や分散オブジェクト技術の延長線上である、と述べた。またInternet.comの記事[6]を引用し、我々が進めている研究の一端を流行のキーワードで飾ってアピールすることは手段の一つであるがその本質を見失ってはいけないと指摘した。

#### ・運用を外注する故の過剰な不安に注目

オーガナイザである渡辺はミドルウェアのアウトソーシングに焦点を当て、特にAmazonやGoogleなどが現在提供しているパブリッククラウド環境ではその利便性に相反し利用者与管理者のお互いの顔が見えない不安から信頼性及びプライバシー保護に関する要求が過剰になる傾向を指摘した。各サービスともSLA (Service Level Agreement, サービス品質保証) による信頼性の保障やプライバシー保護におけるガイドラインの整備[7,8]などを行っているが、利用者側の疑念は尽きないことが[9]にて指摘されている。サービス提供側の責任や義務が増大するなか、サービス側だけでなく利用者側の自衛手段も含めたトータルな信頼性の確保及びセキ

セキュリティ対策が必要であることを主張した。

## 議題 2: 「クラウド」におけるデータ工学と情報マネージメント的な研究のしどころ

### ・ログの多様化と爆発的増加, ローカルとクラウドの同期

平手氏は利用者がクラウドのサービスを利用することで利用にあたっての行動モデルをクラウド側に集約させることができる, これを利用した大規模なログの解析は今後興味深いテーマであるとした。基本的にクラウドはネットワーク環境での利用が前提とされているがオフラインである場合のローカルキャッシュが必要であり, ローカルとクラウドとの同期をどのように図るかも面白いテーマであると指摘した。同期の問題は従来からキャッシュ戦略など多くの手法が考えられているがキャッシュの場合と異なりオフラインの場合はローカルにない場合は取得する手段もない。実際問題オフライン時に限って普段使わないファイルがどうしても必要になる場合もあり, 従来のキャッシュの技術がそのままでは使えず難しい問題であるが挑戦のしがいもあると述べた。

### ・ユビキタスネットワークにおけるクラウド的サービス

神崎氏は, 議題 1 におけるセンサやモバイル機器によるアドホックネットワークもクラウドの研究と言えるのではという主張に基づき, 個人/公的機関/企業など, 様々な人/機関が「好き勝手」に設置したセンサデバイスの統合利用環境を実現するための研究テーマについて述べた。各ピアが持つセンサデバイス, 収集しているデータ, スペック等が全く統一されないネットワークの中で利用者が必要なデータを効率よく取得するためのフレームワークが必要となる。その 1 アイデアとして各センサ共通のコンテンツである時刻に着目した動的ネットワーク構築について紹介した。また「みんなで助け合うクラウド」の一例として, 電力資源の小さいセンサ間の通信処理をモバイル機器が補助する等のアイデアを紹介した。この手法のベースは Data Mule[10]という概念で既出であるが, センサとモバイル, モバイルとモバイルなど, プロトコル的に協調させる点が多く大変興味深いテーマであると述べた。

### ・SaaS におけるフレームワークの提供

小林氏はサービス指向アーキテクチャ (SOA) あるいは Software as a Service (SaaS) の現状と今後の展望について述べた。SOA や SaaS はソフトウェアコンポーネントをサービス化し, 利用者はサービスの組合せで柔軟にアプリケーションが開発できるとしている。しかしながらソフトウェア開発の負担がそれで軽減されているわけではない。SOA を有効に利用するためにはビジネスロジックを内包するフレームワークが必要であるが, 現時点ではそれらのフレームワークは各アプリケーション開発者自身が開発する状況になっており, ソフトウェア開発においてはまだまだ難しいのが現状である。さらにパブリッククラウドまたはエンタープライズクラウドの世界ではどのようになっているかというところ定着しているものはストレージや OS 等のインフラ部分の仮想化部分ではないかという指摘があった。

今後の展望では、参加者からの質問のあった「Amazon や Google の巨大ベンダーに対して小さな新規参入ベンダーが対抗していく手段は」という話と併せて、上記に述べたフレームワーク込みのサービスを提供していくことで差別化を図っていくのだろうと述べた。

#### ・セキュリティや信頼性における責任と義務の共有

時間の超過により当セッション時間中に話すことができなかったが渡辺によるクラウド環境における研究の可能性についてここに記しておく。クラウド環境においてはサービス提供者と利用者との密な関係を期待できないため、十分な信頼関係を築くことができない。その際、厳しい要求をサービス提供者に課すだけでなく、利用者自らも責任をもってデータを扱う必要がある。そのための利用者の自己防衛の技術は新しい視点として興味深い。例えばプライバシー保護検索[11,12] 技術はサービス提供者には暗号化データを預け暗号化したままの処理を行うことにより、サービス提供者におけるデータの閲覧を不可能にする。プライバシー保護だけでなく、トラブル時の原因追跡など利用者側による最小限の労力による自己防衛技術は今後面白いテーマとなると考えられる。

#### まとめ・本セッション内にて議論しきれなかった内容

クラウドコンピューティングというキーワードに関連する様々なトピックを本セッション内で率直に議論できたことは大変有意義であった。時間が足りなかったのも、まだまだ話し足りないところも多く残された。その大きな1点としてクラウド環境、改め、広域かつ大規模分散環境において今後必要となるデータベース技術についての話題をここに記す。2009年3月現在、いわゆるクラウド環境に存在するストレージサービスは従来のリレーショナルDBMSとは異なる。その例として(1)データの一貫性保持の厳密さ、(2)key-value ストアによる単純なデータ格納および問合せ機構、があげられる。たとえばAmazonのDynamoでは結果整合性(eventually-consistent)を採用しており、検索要求の多くがIDによるものという分析の結果key-value storeを採用している。これらは超分散環境においてパフォーマンスを重視するために採用されているが、今後たとえばクラウドによるサービスが普及するにあたってこれらが十分であるのか、もしくは従来のDBMSの持つ諸機能をいかに取舍選択し実現いくべきかという方向で考えるべきか。また、Michael Stonebraker氏が主張するOne size does not fit all[13]に基づいて考えると大規模分散環境において適切なDBMSとは一体どのようなものか、Stonebraker氏らが開発中のH-Store[14]などにも触れるなどして深く議論したかったがオーガナイザ側の十分な準備が足りなかったことも併せて十分な議論ができなかったのが残念であった。

また今後何かの機会があれば、これからも大いに議論を続けていきたいと考えている。

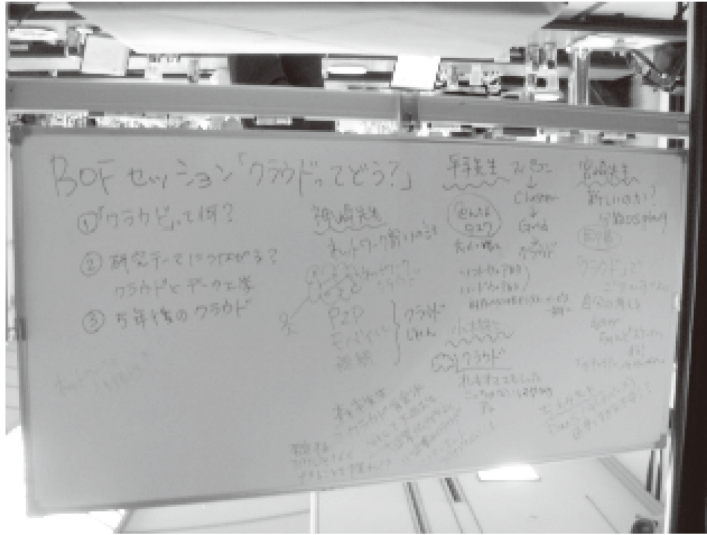


図 2 : 活発な議論を記録したホワイトボード

## 参考文献

- [1] Amazon Web Service, <http://aws.amazon.com/>
- [2] Google App Engine, <http://code.google.com/intl/en/appengine/>
- [3] Windows Azure, <http://www.microsoft.com/azure/default.aspx>
- [4] 文部科学省 e-Society 「基盤ソフトウェアの総合開発プロジェクト」, インターネット上の知識集約を可能にするプラットフォーム構築技術, <http://www.yama.info.waseda.ac.jp/e-society/>
- [5] Plan 9 from Bell Labs, <http://plan9.bell-labs.com/plan9/>
- [6] クラウドコンピューティングがくだらない理由  
<http://japan.internet.com/busnews/20080822/6.html>
- [7] 経済産業省, SaaS 向け SLA ガイドライン,  
<http://www.meti.go.jp/press/20080121004/20080121004.html>
- [8] 総務省, ASP・SaaS における情報セキュリティ対策ガイドライン,  
[http://www.soumu.go.jp/menu\\_news/s-news/2008/pdf/080130\\_3\\_bt3.pdf](http://www.soumu.go.jp/menu_news/s-news/2008/pdf/080130_3_bt3.pdf)
- [9] クラウドに対する“過剰な”セキュリティ不安を払拭しよう,  
<http://www.computerworld.jp/topics/cloud/137069-1.html>
- [10] Arun A Somasundara, Aman Kansal, David Jea, Deborah Estrin, Mani B Srivastava, "Controllably Mobile Infrastructure for Low Energy Embedded Networks," IEEE Transactions on Mobile Computing (TMC), Vol.5, No.8, pp.958—973 (2006)
- [11] Hacigumus H., Iyer B. R., Li C. and Mehrotra S., Executing sql over encrypted data in the database-service-provider model, Proceedings of the 2002 SIGMOD International Conference, pp. 216–227 (2002).
- [12] 渡辺知恵美, 新井裕子, 天笠俊之: ブルームフィルタを用いたプライバシー保護検索における攻撃モデルとデータ攪乱法の一検討, データ工学と情報マネジメントに関するフォーラム (DEIM2009), D4-6 (2009).
- [13] Michael Stonebraker and Uğur Çetintemel: "One Size Fits All": An Idea Whose Time Has Come and Gone, 21st International Conference on Data Engineering (ICDE2005), pp.2-11 (2005)
- [14] Robert Kallman, Hideaki Kimura, Jonathan Natkins, Andrew Pavlo, Alex Rasin, Stan Zdonik, Evan Jones, Yang Zhang, Samuel Madden, Michael Stonebraker, John Hugg, Daniel Abadi. In Proceedings of VLDB, pages 1496 - 1499 (demo), 2008.