

九州大学キャンパスクラウドシステムの導入

伊東, 栄典
九州大学情報基盤研究開発センター | 九州大学情報システム部

上田, 将嗣
九州大学情報基盤研究開発センター | 九州大学情報システム部

<https://hdl.handle.net/2324/22547>

出版情報 : 第33回 全国共同利用情報基盤センター 研究開発連合発表講演会 研究開発論文集, pp.109-114, 2012-07-10

バージョン :

権利関係 :



九州大学キャンパスクラウドシステムの導入

上田将嗣† 伊東栄典‡

†九州大学情報システム部, ‡九州大学情報基盤研究開発センター

〒812-8581 福岡市東区箱崎 6-10-1

{ueda, itou}@cc.kyushu-u.ac.jp

概要

高速ネットワークの普及, 計算機性能の向上, 仮想化技術および Web 技術の進展から, クラウド技術が進展している。九州大学情報基盤研究開発センター (以下, センターと略記) では, 主に大学院での情報環境を用いた教育研究に用いるためのプライベートクラウドシステムを導入し, 各大学へ提供することを計画しており, その手始めとして, 情報系の大学院への提供を想定したクラウドシステムを導入中である。本稿では導入中の九州大学キャンパスクラウドシステムについて報告する。

1. はじめに

昨今の計算機ハードウェアおよびネットワークの高速化と, 仮想化技術の発達に伴い, クラウドと総称される計算機システムが実現されている[1,2,3]。大学の情報基盤についてもクラウド化は進んでいる[4,5,6]。九州大学でも, 事務システムに仮想化技術も利用可能な計算機システムが基盤として導入されており, 成績管理システムや人事給与データベースなどを仮想マシン上で動かし計算資源を効率的に使っている。大学の教育用計算機システムについては, 従来のシンクライアントから仮想デスクトップを活用する事例が増えている。

米国 NIST の定義[1]によると, クラウドとは, 簡単にオンデマンドでネットワーク経由から利用可能な, 共有された計算資源であり, 5つの本質的な特徴, 3つのサービスモデル, 4つの展開モデルがあるとしている。5つの本質的な特徴とは, On-demand self service, 広帯域ネットワーク(broad network access), リソースプール(resource pooling), 弾力性(rapid elasticity), そして measured service である。サービスモデルには, SaaS (Software as a Service), PaaS (Platform as a Service), IaaS (Infrastructure as a Service)の3つがある。最後の展開モデルには, プライベートクラウド, コミュニティクラウド, パブリッククラウド, ハイブリッドクラウドの4つを定義している。

本稿ではセンターが導入する九州大学キャンパスクラウドシステム (以下, Qcloud と略記) について説明する。Qcloud は学内で利用するプライベートクラウドとして運用を開始するもので, 九州大学大学院システム情報科学研究院 (以下, システム情報科学研究院と略記) と共同で構築しており, 情報系大学院であるシステム情報科学研究院での教育研究利用を想定している。最初は試験的なサービスとして開始し, 用途や運用方式についての知見を深め, 学内の他学部や, 他大学へ提供していく予定である。

2. 九州大学キャンパスクラウドシステムの利用シナリオ

導入するシステムの構成を決めるには, それがどのように利用されるかを明確化する必要がある。情報系の大学院がどのような用途で計算資源を必要とするか, 具体的な利用シナリオを明確化した。

まずは情報系の大学院である, システム情報科学研究院が現状の計算機活用状況を調査した。次に,

クラウド化した場合の利用形態はどうあるべきかについてセンターとシステム情報科学研究院が共同で検討し、利用形態として以下の5項目を想定することとした。

- (1) 講義および演習での利用
- (2) 常時稼働サーバおよび特定アプリケーションの利用
- (3) 情報サービス開発用
- (4) データ処理用
- (5) 数値計算

上記のうち、(5)の数値計算については、従来からスーパーコンピュータを含む HPC 系のシステムが担ってきた。現在も大規模システムが稼働しており、数値計算のための機能は今回のクラウドシステムには不要であると判断した。以下、上記(1)～(4)について検討した利用形態を述べる。

2.1 講義および演習での利用

大学院では、座学を中心とする講義と、演習を行う講義があり、その中に実際に計算機を用いて行う講義がある。講義でクラウドシステムを利用できるかについて授業担当者への調査を実施した。その結果、表1に示す講義・演習が計算機を利用している、あるいは将来の利用を想定していることが分かった。

表 1 システム情報科学研究院での講義・演習における計算機利用例

講義	LSI シミュレーション	暗号理論	データベース	Web データ処理
要望	Linux OS とシミュレータソフト	Java 言語処理系	MySQL, できれば Windows 系 OS	Perl 言語処理系 HTML 処理ライブラリ

講義および演習でクラウドシステムを用いるには、授業の時間割にあわせて計算資源を用意しておく必要がある。授業での利用を考えたクラウドシステム運用基盤として、ノースカロライナ州立大学が開発した VCL (Virtual Computing Lab.) がある[2]。日本でも明治大学や九州大学が VCL を導入している[3]。

システム情報科学研究院は、2011 年度から VCL の運用を開始し、主に大学院教育で用いている。まだ授業での利用は多くないものの、表1に示す講義での利用が始まっている。センターで導入する Qcloud システムでも、表1に示す講義・演習の要望が実現できる必要がある。他学部での Qcloud 利用を考えるなら、表1には無い授業での利用も考える必要がある。例えば、統計に関する講義・演習のために、統計用ソフトを搭載した VM を準備する必要があるかもしれない。

2.2 常時稼働サーバおよび特定アプリケーションの利用

研究室で、メールサーバや Web サーバなど、常時稼働するサービスを動かしている場合は多い。大学院全体で使う大学院内 SNS や、研究業績蓄積 Web DB システムもある。他にも、ある研究室が特定のアプリケーションを使うための研究室共用サーバもある。これらの常時稼働するサーバは、長期間運用することが多いためクラウドシステム内の仮想環境で動作させる場合でも、長期間の利用が実現できるようにする。

2.3 情報サービス開発用

情報系の大学院では、情報サービスの研究開発を行う研究室も多い。これらの開発環境は従来、実際の計算機上に構築していたが、ハードウェアの用意、OS の導入、アプリケーションの設定、ネットワークの設定など、数時間かかる作業であった。仮想化により、OS およびアプリケーションの準備時間が早くなれば、情報サービスの開発効率が上がる。

新しい OS やアプリケーションを試してみたいという要望もある。自分が使ったことのない OS や、話題になっているアプリケーションを試したい、使用中のソフトの最新版を試してみたい、という要望もある。このような新しい OS やアプリを試す場合、既存の環境を損なわずに、手軽に使える環境があると良い。

また、近年ではソフトウェアシステムを、VM イメージとして配布する場合もある。例えば、Cloudera 社は Hadoop 環境一式を VM イメージとして公開しているし、国立情報学研究所も Shibboleth 認証サーバを VM イメージとして配布している。これらを試す環境も望まれている。

2.4 データ処理用

システム情報科学研究院では、Web 情報検索、データマイニング、バイオインフォマティクス、画像処理など、大規模データ処理を研究しているところも多い。しかしながら、大規模データ処理用の計算機システムで使い勝手の良いものは用意されていなかった。

近年、分散データ処理のためのミドルウェアとして、Apache Hadoop が注目されている[7]。Google 社は Web データのインデクシングに Map-Reduce の仕組みを用いている。Google 社の処理機構をオープンソースとして実現したものが Hadoop である。Hadoop はネットワーク接続された複数台の計算機を、一つの塊として扱う仕組みを備えている。分散ファイルシステム HDFS (Hadoop distributed file system) が用意されており、複数の計算機上にデータが分散配置される。データ処理には Map-Reduce の仕組みが用意されている。データとプログラムの分散配置 (Map 操作) の後、各計算機上でデータ処理を行い、最後に処理結果の集約 (Reduce 操作) を行う (Reduce 操作の前に、Shuffle 操作が入る場合もある)。

Hadoop による大規模データ分散処理基盤を用意すれば、情報検索やバイオインフォマティクス、画像処理などの研究に寄与できる。また、データ処理の講義や演習での利用も可能である。

3. キャンパスクラウドシステムの構成

上記 2 節の検討に基づき、導入する Qcloud システムでは次に示す 4 つの部分システムを構築することとした。

- | | |
|----------------|-------------------------|
| (1) 高年次教育用クラウド | 講義および演習での利用 |
| (2) サーバ用クラウド | 常時稼働サーバおよび特定アプリケーションの利用 |
| (3) 開発用クラウド | 情報サービス開発用 |
| (4) データ処理クラウド | 大規模分散データ処理 |

図 1 に導入する Qcloud システムの構成を示す。

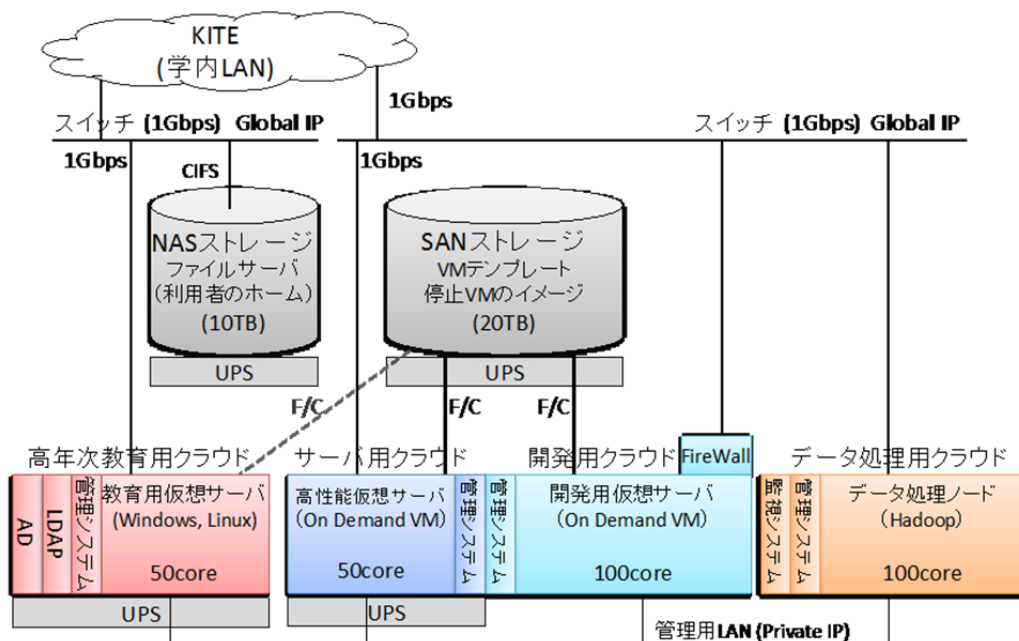


図 1 九州大学キャンパスクラウドシステム (Qcloud) システム構成

以下の表 2 に、各部分システムの特徴を記載する。

表 2 部分システムの特徴

	総コア数	コア当たりの主記憶量	ハイパーバイザ	VM で利用可能な OS	VM の稼働を行う人
(1)高年次教育用	50	4GB	VMware	CentOS Windows7	教員・学生
(2)サーバ用	50	4GB	VMware	CentOS, Windows7	教員のみ
(3)開発用	100	4GB	または	Windows Server	教員・学生
(4)データ処理用	100	2GB	XenServer	CentOS	Qcloud 管理者

3.1 高年次教育用クラウドシステム

高年次教育用クラウドシステムは、大学院の講義での利用を考えているため、VCL と同程度の機能を持つ管理システムを導入する。VM のテンプレートは、講義に使うものを Qcloud 管理者が用意する。時間割に応じた VM の起動設定も Qcloud 管理者が行う。

教員および学生は、Web インターフェイスの利用者認証 (Shibboleth 認証) を経て、VM の稼働操作を行う。

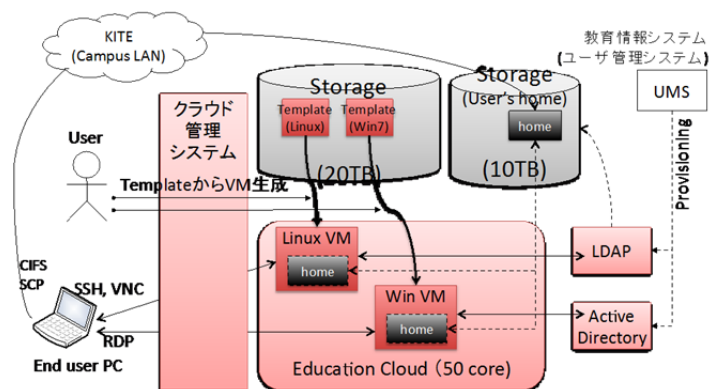


図 2 高年次教育用クラウドシステムイメージ

3.2 サーバ用クラウドシステム

サーバ用クラウドシステムでは、Amazon Elastic Compute Cloud (Amazon EC2) [8]と似た機能を持つ管理システムを提供する。長期的な利用が多いと想定されるため、VM の稼働を行えるのは教員のみに限定する。なお、Web インターフェイスの利用者認証は Shibboleth 認証である。

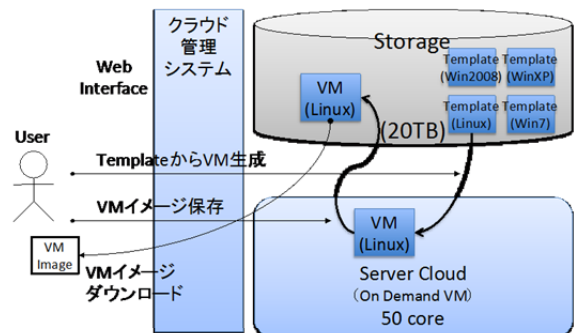


図 3 サーバ用クラウドシステムイメージ

3.3 開発用クラウドシステム

開発用クラウドシステムは、サーバの開発やソフトの試用のための環境である。前述のサーバ用クラウドと同じように Amazon EC2 と同等の管理システムでサービスを行う。開発環境は主に大学院生が使うと想定しているため、学生も VM の稼働を行える。ただし VM の最長稼働期間は制限する。稼働期間を過ぎた VM は、自動的に停止して、VM イメージのファイルとして保持しておく。また、外部から VM のイメージをアップロードして、それを稼働させることも可能である。

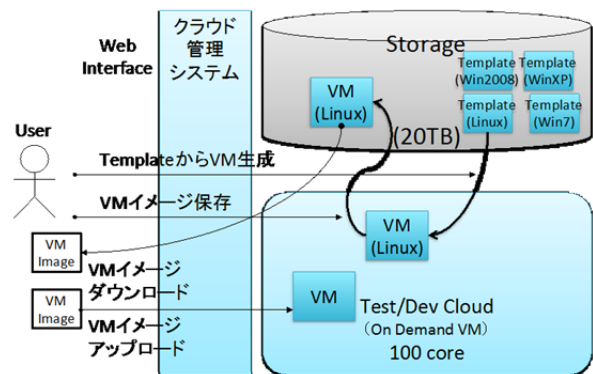


図 4 開発用クラウドシステムイメージ

3.4 データ処理用クラウドシステム

最後のデータ処理用クラウドは、100 コア分 (50 台) の実サーバ機で、Apache Hadoop を搭載した CentOS マシンを多数動かす。利用者は教員と学生で、SSH などで代表サーバ(name node)にログインして使う。

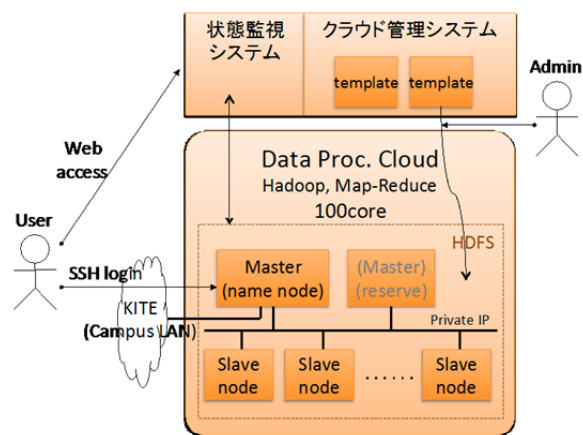


図 5 データ処理用クラウドシステムイメージ

4. 議論

Qcloud システムの運用にあたり、いくつかの問題がある。一つは情報セキュリティの問題である。VM を気軽に動かすことができると、セキュリティ対応が充分でないマシンが稼働する可能性も上がる。特に開発用クラウドのセキュリティ対策が問題になるため、開発用クラウドは通信を学内の計算機との間だけに制限する。

ソフトウェアのライセンスも問題である。従来のように実機であれば、Mathematica や MATLAB のようなソフトウェアを研究室で共用することの問題は少なかった。しかし VM では VM 全体をイメージファイルとして取得可能であるため、きちんとライセンス通りにする制御が

難しい。そもそも、VM 上で個別ソフトウェアを用いる場合のライセンスは、確認が必要である。

5. おわりに

本稿では、九州大学で導入を進めているキャンパスクラウドシステム（Qcloud システム）について説明した。Qcloud は、当初は学内向けのプライベートクラウドとして開始する。将来は国内の大学院における高年次教育用および研究用としてサービス展開を行い、大学のコミュニティクラウドとしていきたい。

参考文献

- [1] P. Mell and T. Grance, The NIST Definition of Cloud Computing, NIST (National Institute of Standards and Technology), Vol. 53(6), No. 50. 2009.
http://csrc.nist.gov/publications/drafts/800-145/Draft-SP-800-145_cloud-definition.pdf
- [2] M. Armbrust, et.al., Above the Clouds: A Berkeley View of Cloud Computing, Technical Report No. UCB/EECS-2009-28, 2008.
- [3] Qi Zhang, Lu Cheng and Raouf Boutaba, Cloud computing: state-of-the-art and research challenges, Journal of Internet services and applications, Vol.1, No.1, pp.7-18, 2010.
- [4] 井上春樹, 長谷川孝博, 八巻直一, 水野信也, 山崎國弘, 吉田仙良, 戸部剛男: クラウドコンピューティング全面適用のインパクト 静岡大学クラウド情報基盤 SUCCES(Shizuoka University Cloud Computing Eco System)の紹介, 静岡学術出版, ISBN-10: 4903859436, 2010.
- [5] VCL, Virtual Computing Lab.: <http://vcl.ncsu.edu/> , accessed at Aug. 2011.
- [6] 明治大学: キャンパスクラウドシステム利用マニュアル,
<http://www.isc.meiji.ac.jp/vpc/VCL-userguide.pdf>, accessed at Aug. 2011.
- [7] Apache Hadoop: <http://hadoop.apache.org/> , accessed at Aug. 2011.
- [8] Amazon EC2: <http://aws.amazon.com/jp/ec2/> , accessed at Aug. 2011.