OSS/IaaS管理基盤技術「OpenStack」の最新動向とNTTデータの取り組み

NTTデータではOpenStackを活用したクラウド基盤技術の研究開発に取り組んでいます．本稿では，OpenStackの仕組み，およびNTTデータによるOpenStack機能開発，OpenFlow連携技術開発，クラウドセキュリティ機能開発，Swiftの活用展開について紹介します．

花舘 蔵之（はなだて まさゆき）
NTTデータ

OpenStackの概要

OpenStack[(1)]は，クラウドを構成する物理サーバや仮想マシンの運用管理を一元的，かつ効率的に行うためのオープンソースソフトウェア（OSS）です．2012年4月26日現在で世界160以上の企業によって開発が進められており，同年9月末にVersion6（Folsom版）がリリース予定です．LinuxディストリビューションのUbuntu 12.04LTSにはOpenStackが同梱されています．

NTTデータはプロジェクト立ち上げメンバとして2010年からOpenStackプロジェクトに参画し，OpenStackを活用したクラウドサービスの研究開発，および普及展開にNTT研究所と連携し取り組んでいます．

OpenStackが提供するサービスは，NIST（National Institute of Standards and Technology）[(2)]が提唱するクラウドサービス種別のIaaS（Infrastructure as a Service）層に相当します．OpenStackの利用者（クラウド利用者）は，KVM（Kernel-based Virtual Machine）やXenServer等の仮想化環境上で動作する仮想マシンにネットワーク経由で外部からアクセスし，計算資源（CPU，メモリ，HDD，IPアドレス等）をいつでも，すぐに利用できます．

次にOpenStackの主な特徴を説明します．

⑴ マルチテナント

さまざまなクラウド利用者の仮想マシンを同一物理マシン上に配備できます．クラウド全体で余剰な計算資源を削減できるため，物理マシンのコスト削減を期待できます．

⑵ オンデマンド・セルフサービス

クラウド利用者自身がWebGUI（Graphical User Interface）やAmazon EC2互換API（Application Programming Interface），OpenStack APIを用いて仮想マシンの運用管理（仮想マシンの起動や停止など）を行えます．クラウド利用者はこれらのインタフェースを用いてクラウドの管理者（クラウド提供者）を介することなく迅速にサービスを開始できます．さらに，クラウド提供者の運用管理業務の一部をクラウド利用者が行うため，クラウド提供者の運用管理コストも削減できます．

⑶ ライブマイグレーション機能

ライブマイグレーション機能とは，ある物理マシン上で動作している仮想マシンのメモリ状態を，無停止でほかの物理マシン上の仮想マシンに載せ替える機能です．物理マシンの入れ替えなど，物理マシンの保守性が向上します．

⑷ セキュリティ機能

OpenStackはクラウド提供者やクラウド利用者の個人認証，Amazon EC2互換APIのHMAC認証，クラウド利用者ごとのファイアウォール設定，VPN（Virtual Private Network）機能，仮想マシン利用時の通信路暗号化（SSH: Secure SHell）等の基本的なセキュリティ機能を備えています．

OpenStackの仕組み

OpenStackは複数のコンポーネントから構成され[(3)，(4)]，これらのコンポーネントが連携することによりIaaS基盤機能を提供するアーキテクチャとなっています（**図1**）．次に，これらのコンポーネント（コードネーム）を説明します．

Compute（Nova）は「計算資源管理」「計算資源割り当て」「メッセージ通信」を行います．計算資源管理では，OpenStackが管理する物理マシンのCPU，メモリサイズ等を管理します．計算資源割り当てでは，計算資源管理で管理されている物理マシンから，

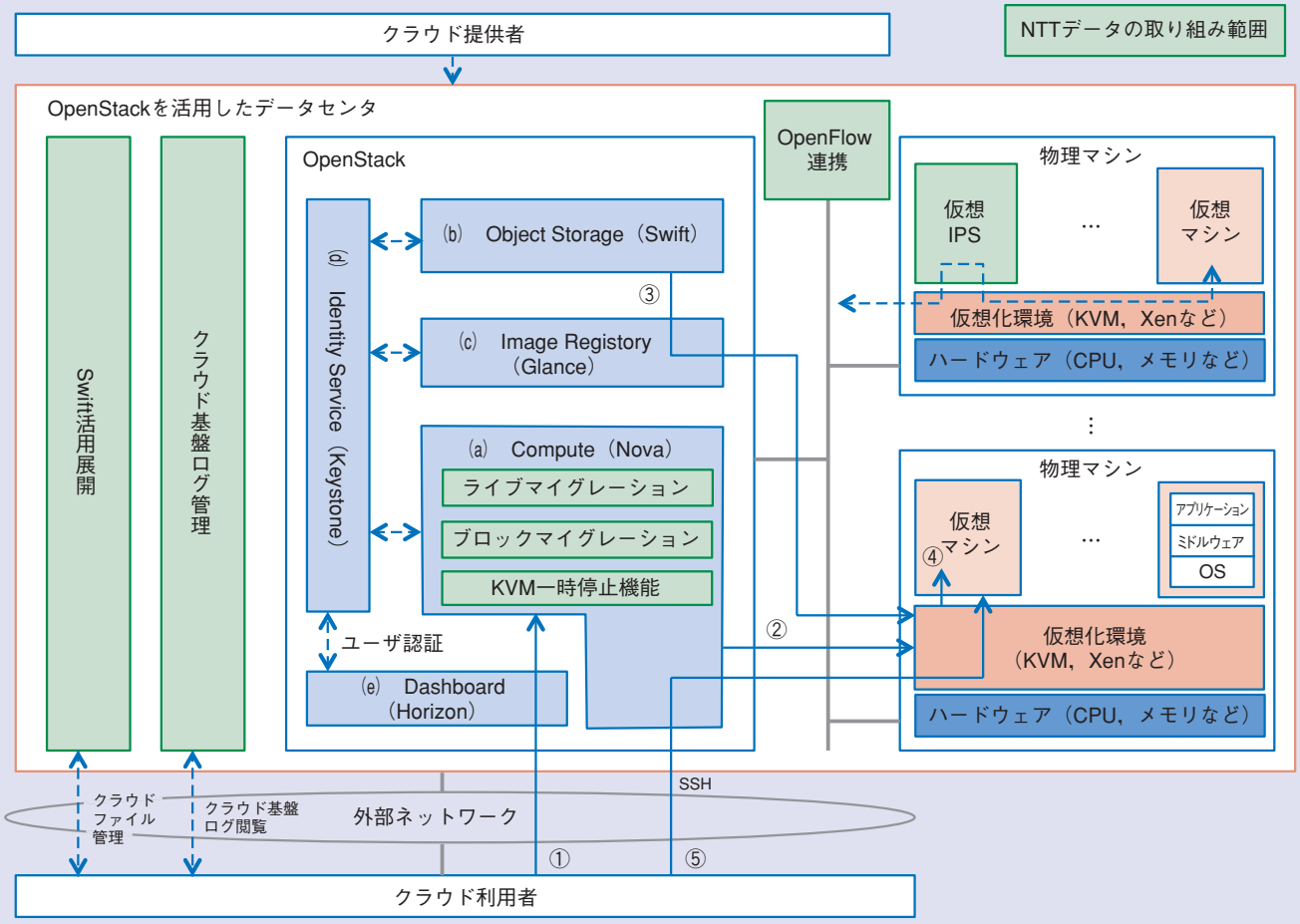


図1　OpenStackの構成

クラウド利用者が利用する計算資源を決定します．メッセージ通信は，仮想マシン起動・停止等，クラウド利用者によるさまざまな制御メッセージを送受信します（図1(a)）．

Object Storage（Swift）は利用可能な仮想マシンのテンプレート情報（VMイメージ）を保管します（図1(b)）．

Image Registry（Glance）はComputeが決定した物理マシンやVMイメージの内容に従い，VMイメージをObjectから読み出し物理マシンに転送します（図1(c)）．

Identity Service（Keystone）はクラウド利用者やクラウド提供者のID・パスワードを一元管理します．また，個人認証処理，および各コンポーネントの認可判定処理を行います（図1(d)）．

Dashboard（Horizon）はクラウド利用者にWebGUIを提供します（図1(e)）．

次に，これらのコンポーネントを活用してクラウド利用者が仮想マシンを起動する流れを説明します．

① メッセージ受信：クラウド利用者は，Computeに仮想マシン起動要求を送信します．その際，クラウド利用者は，希望する計算資源（CPU・メモリ使用量，ディスク容量等），仮想マシンの種類

（OS等）を選択します．

② 計算資源の割り当て：Computeは，希望する計算資源に見合う物理マシンを未使用の物理マシンから決定します．また，使用するIPアドレスも決定します．そして，決定した物理マシンの仮想化環境に仮想マシン起動を通知します．

③ 仮想マシンイメージ・ロード：仮想化環境は，Image Registryに仮想マシンイメージの転送を依頼します．Image Registoryは，Object Storageに格納された仮想マシンイメージから，起動する仮想マシンイメージを特定し仮想化環境に転送します．

④ 仮想マシン起動：仮想化環境は，転送されたVMイメージを起動します．

⑤ 仮想マシン利用：クラウド利用者は起動した仮想マシンにネットワーク経由でアクセスし，利用します．

NTTデータの取り組み

NTTデータは，OpenStackを用いた次の取り組みを行っています（図1）．

① OpenStack機能として，「ライブマイグレーション」「ブロックマイグレーション」「KVM一時停止機能」等を実装し，OpenStackコミュニティに提供しています．

② クラウド導入時の上位阻害要因である，クラウドの安全性に対する不安を解決するために，前述のOpenStackの既存セキュリティ機能をさらに強化する補完的なセキュリティ機能（クラウド基盤の統合ログ管理，仮想不正遮断システム）を開発しています．

③ 「NTTデータ版OpenFlowコントローラ」と「OpenStack」を連携する取り組みを行っています．

④ Object Storage（Swift）を用いてペタバイトクラスの安価で高信頼な大規模分散ストレージシステムの構築技術（Swift参照アーキテクチャ）の確立に取り組んでいます．

ブロックマイグレーション

ブロックマイグレーションはVersion 4（Diablo版）にてリリースしました．KVM上で動作する仮想マシンをメモリだけではなくディスクイメージも含めてほかの物理マシンに移行する機能です（**図2**）．ライブマイグレーションでは，メモリ内容だけが移行されディスク内容は移行されません．移行元と移行先の2つの物理マシンが同じディスクを共有していることが前提条件となるため，ディスク共有が困難な状況（例えばデータセンタが異なる場合など）ではライブマイグレーションも困難となります．これに対し，ブロックマイグレーションは，異なるディスクに接続された物理マシン間でもメモリ内容とディスク内容を移行できるため，異なるデータセンタ間に退避させる場合でも対応でき，ライブマイグレーションを組み合わせることにより保守性が向上します．

KVM一時停止機能

この機能はVersion 4（Diablo版）にてリリースしました．仮想マシンを一時停止する方法は，①仮想マシンの保存，②仮想マシンの停止，③①で保存された仮想マシンの再起動，という手順により実現できます．しかし，Version3（Cuctus版）までは，仮想マシンの状態をディスクやメモリに保存することができなかったため，仮想マシンの一時停止を実現できませんでした．NTTデータは，仮想マシンの状態をディスクに書き出す「Suspend機能」，およびメモリに書き出す「Pause機能」を実装することにより，仮想マシンの一時停止を実現しました．

クラウド基盤の統合ログ管理

クラウド利用者は，自らが管理する仮想マシンに直接アクセスできるため，仮想マシン内のログ（OS，ミドルウェア，アプリケーション等のログ）を自分で確認できます．一方，クラウド利用者がクラウド基盤ログ〔仮想化環境，OpenStack（Compute，Image Registory等）のログ〕を閲覧する場合，すべてのクラウド利用者のログが1つのログファイルに混在しており，またログレコード単位の開示制御が必要であるため，容易に閲覧することはできませんでした．

NTTデータは，1つのログファイル

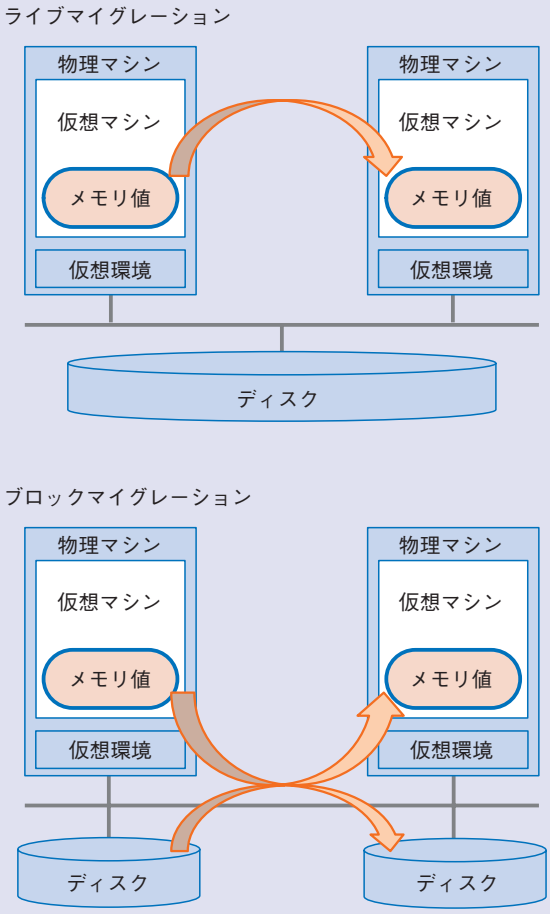


図2　ライブマイグレーションとブロックマイグレーションの仕組み

からリアルタイムにクラウド利用者が関係するクラウド基盤ログだけを抽出するクラウド基盤ログ管理システムを開発しました．これにより，クラウド利用者はクラウド基盤ログを用いた動作状況をリアルタイムかつ安全に閲覧できます．また，見えにくいクラウド基盤に対する安心感の向上も期待できます．

仮想不正遮断システム

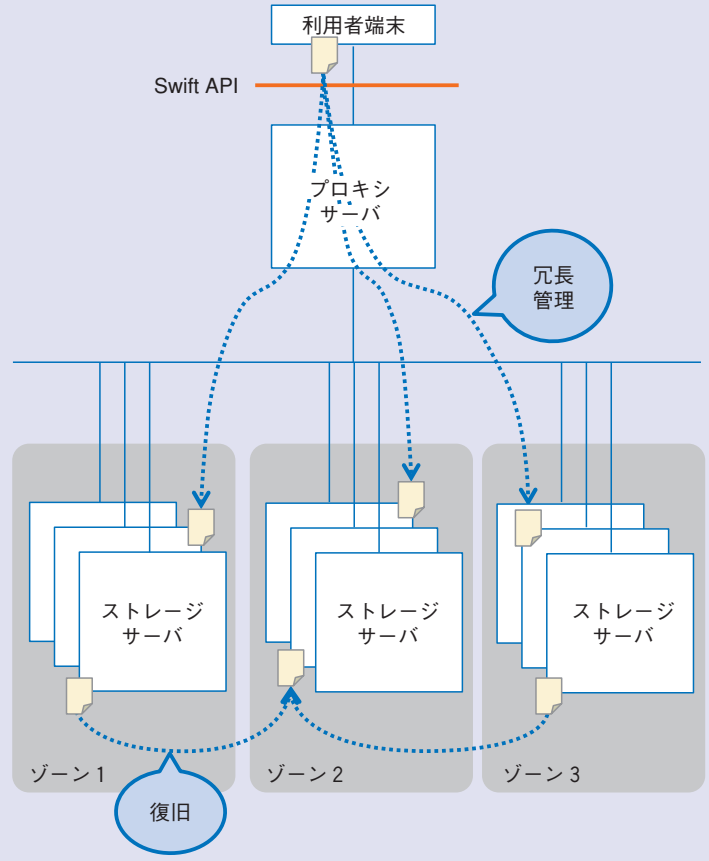
不正遮断システムは，レイヤ3以上の通信パケットを監視し，不正パケットを検知・遮断するシステムです．複数のクラウド利用者が共同利用するシステム上に既存の不正遮断システムを配備する場合，クラウド利用者ごとの通信内容を用途やニーズに合わせて監視することが性能上必要不可欠となります．また，1つの物理サーバ内で仮想マシン間通信が行われる場合，その通信を監視するためには，どうしても一度，物理マシン外のネットワークを経由させ，そこで監視を行う必要があります．

NTTデータは，OSS不正遮断システムであるSuricataを配備した仮想マシンを「仮想IPS」として用意し，仮想IPSを経由した通信を実現することにより，仮想マシン間通信を含むさまざまな通信パケットを仮想マシン単位で監視する技術を開発しています．

Swift参照アーキテクチャ

Swiftは，米国Rackspace社のファイル保管サービスのために開発されました．Swiftのシステム基本構成（**図3**）は，大きくプロキシサーバとストレージサーバに分けることができます．ストレージサーバにはオブジェクトが格納されます．オブジェクトは，テキスト，画像，動画等のファイルや，保管先のディレクトリ等のメタ情報です．プロキシサーバは，Swift APIを用いて，クラウド利用者とストレージサーバ間のオブジェクト送受信を中継します．

ストレージサーバは，複数台で構成され，「ゾーン」と呼ばれる単位でグループ管理されます．オブジェクトは異なるゾーンのストレージサーバに冗長保管されます．オブジェクトは監視プロセスによって監視されており，ファイルやディスクが壊れた場合，冗長化されたオブジェクトを用いて失われたオ



※各ストレージサーバは，メタ情報を管理するアカウントサーバ，コンテナサーバ，ファイルを管理するオブジェクトサーバから構成されます．

図3　Swiftのシステム基本構成

ブジェクトを自動的に複製し，復旧します．

このSwiftを用いてペタバイトクラスの大規模分散ストレージシステムを構築する場合，格納するファイルサイズやネットワーク帯域，ハードウェアの処理性能や故障率，許容される復旧時間等によりシステム構成やパラメータ設定に工夫が必要です．

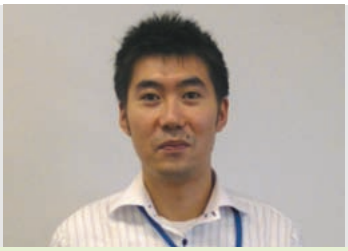
NTTデータは，この設計・構築ノウハウを「Swiftリファレンスアーキテクチャ」としてパターン化することにより，安価で迅速，かつ安定した大規模分散オブジェクトストレージの提供を目指しています．

NII案件への取り組み

クラウド上のマシンのパフォーマンスを重視するケースや，ライセンス上仮想化が使えない用途向けに，クラウドAPIで物理マシンを払い出す仕組みを，国立情報学研究所（NII）を中心としたdodaiプロジェクトで実装しました．そして，2011年度にこのdodaiを使った研究用のクラウドシステムである「アカデミックコミュニティクラウド　プロトタイプシステム」をNII向けに構築しました．今後，運用評価のうえで物理マシンクラウドの仕組みを実用化したいと考えています．

■参考文献

(1) http://openstack.org
(2) P. Mell and T. Grance：“NISTによるクラウドコンピューティングの定義，”NIST SP，800-145（Draft），Sept. 2011.
(3) 野口：“いま注目のOpenStackで学ぶクラウド構築の基本，”日経Linux，第151号，pp.120-124，2012.
(4) http://techtarget.itmedia.co.jp/tt/news/1101/13/news06.html



花舘　蔵之

Internap，KT，Rackspace，HP等，OpenStackを活用したIaaSが世界中で始まっています．このOpenStackを活用したクラウド化の流れに合わせ，NTTデータはOpenStackをコアとしたクラウド基盤技術の展開を推進していきます．

◆問い合わせ先
NTTデータ
技術開発本部
TEL　050-5546-2308
FAX　03-3532-0487
E-mail　rdhkouhou@kits.nttdata.co.jp