

# 実データの分散配置場所を柔軟に制御できる 仮想計算機用ストレージシステムの提案

島 慶一

IIJ イノベーションインスティテュート

分散型データセンターでの仮想計算機基盤の運用は、データセンター間のネットワーク帯域や遅延に柔軟に対応可能でなければならない。本書では地理位置条件に応じて仮想計算機用ディスクイメージを柔軟に分散配置運用できるストレージシステムを提案する。

## A Storage System for Virtual Machines with a Flexible Data Location Management Method

Keiichi Shima

IIJ Innovation Institute Inc.

When operating a virtual machine infrastructure in a distributed datacenter, flexibly adapting to different network latency and bandwidth between datacenters is important. In this document, we propose a virtual image storage system that enables a flexible operation in locating storage data based on geographical conditions.

### 仮想計算機用ストレージの必要性

仮想化技術の著しい発展により、実用に耐える仮想計算機の運用が可能になった。ネットワーク上で提供されるサービスにおいても仮想機材を用いて構成する場合が増えている。仮想化により、サービスの迅速な展開や高負荷時の緊急設備拡張、また利用減に伴う設備縮退など、物理計算機では手間がかかる運用が比較的容易に実現可能になった。今後は、仮想資源の配置や再配置を柔軟に実現する仕組みが重要となる。

CPU やメモリ資源の再配置技術はすでに実用化され、オープンソース、商用を問わず多くのハイパーバイザーに実装されている<sup>1</sup>。また、SDN 技術を用いたネットワーク資源の再配置技術の実用化も近い。残る問題はストレージ資源の再配置である。

データセンターは効率を求めて大規模化している。しかし、我が国のような国土が狭い環境で諸外国にみられる規模のデータセンターを建設することは困難である。そのため、地理的に独立した中規模のデータセンターを仮想的に大規模データセンターとみなす分散型データセンター

運用が検討されている。分散したデータセンターを横断して基盤サービスを構成する場合、地理条件や遅延を考慮して仮想資源を効率的に配置運用する技術が重要となる。

本書では、仮想計算機基盤運用者の視点に立ち、分散型データセンターにおいて仮想ディスクの実データの配置を柔軟に制御できるストレージシステム「UKAI」[1] を解説する。UKAI システムを用いることで、仮想ディスクイメージを構成する実データの冗長性や配置場所を、運用者が自在に制御できるようになる。

### UKAI の設計と実装

UKAI の仮想ディスクは、実データへのポインタの集合として構成される。仮想ディスクは複数の論理ブロックに分割され、それぞれのブロックに対応する実際のデータは地理的に分散された外部のストレージノードに保持される。ストレージノードは複数運用することができ、必要とされる冗長度に応じて複数のブロックの複製を保持できる (図 1)。仮想ディスクの構成操作は仮想計算機の運用とは独立に実行できるようになっており、運用者は仮想計算機を動作

<sup>1</sup>KVM: <http://www.linux-kvm.org/>, VMware: <http://www.vmware.org/>

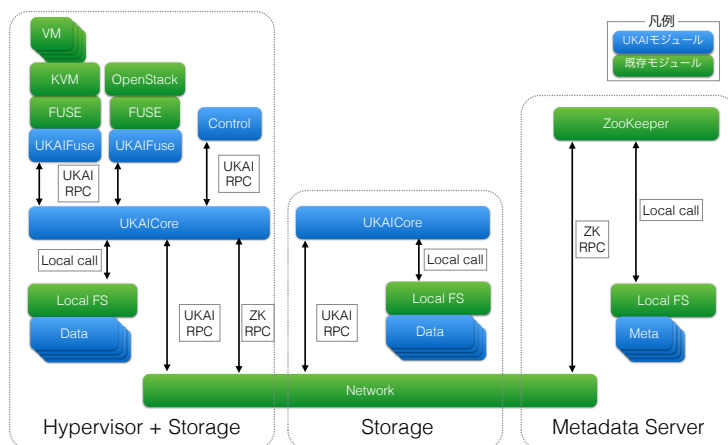


図 2: UKAI モジュール関連図

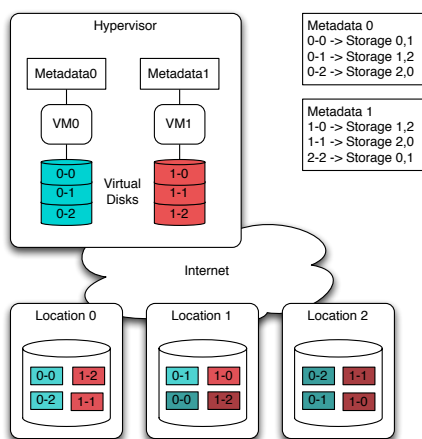


図 1: UKAI 仮想ディスクの構成図

させたまま、仮想ディスクの実データをより適切な場所、例えば遅延の短い場所に移動させたり、地理的に離れた場所に複製を作ったりできる。既存の分散ストレージ [2] ではブロックの自律配置に主眼をおいているが、UKAI では運用者に配置決定の権限と責任が委ねられる。

図 2 に FUSE を用いた参照実装<sup>2</sup> のモジュール関連図を示す。仮想ディスクの構成を記録したメタデータは ZooKeeper クラスタで保持され、仮想ディスクが利用される際にハイパーバイザーから動的に参照される。KVM などに代表される仮想計算機の実行プログラム、および OpenStack などのクラウドコントローラは、FUSE 経由で提供されるファイルを仮想ディスクの単位とし

て操作する。

## まとめ

仮想基盤の活用が進むにつれ、その効率的な運用のために仮想資源の柔軟な配置制御が重要な要素となる。本書では、仮想計算機基盤運用者の視点に立ち、分散型データセンターでの基盤運用に必要な地理位置情報を考慮できるストレージシステムの提案を行った。

## 参考文献

- [1] K. Shima. UKAI: Centrally Controllable Distributed Local Storage for Virtual Machine Disk Images. In *Proceedings of Globecom 2013 Workshop - Cloud Computing Systems, Networks, and Applications (CC-SNA)*, pages 432–438, December 2013.
- [2] S. A. Weil, A. W. Leung, S. A. Brandt, and C. Maltzahn. RADOS: A Scalable, Reliable Storage Service for Petabyte-scale Storage Clusters. In *Proceedings of the 2th international Petascale Data Storage Workshop (PDSW'07)*, pages 35–44, November 2007.

<sup>2</sup><https://github.com/keiichishima/ukai/>