

# マルチドメインクラウド環境のための資源管理フレームワーク

竹房 あつ子<sup>†</sup> 中田 秀基<sup>†</sup> 高野 了成<sup>†</sup>  
柳田 誠也<sup>†</sup> 大久保 克彦<sup>†</sup>  
工藤 知宏<sup>†</sup> 田中 良夫<sup>†</sup>

## 1. はじめに

クラウドは、大規模データインテンシブ科学技術計算の計算基盤としても注目されているが、分散する複数組織から提供される資源を統合して計算基盤を構築する技術は成熟していない。このような仮想計算基盤では、以下のような課題がある。

- (1) アプリケーションの性能要求に応じて、計算機やストレージだけでなくクラウド間のネットワーク帯域を含めた資源群を適切に確保する。
- (2) 資源群の所在やドメインを隠蔽し、1つの仮想計算基盤として提供する。
- (3) 仮想計算基盤のモニタリング情報を提供し、アプリケーションの性能品質維持をサポートする。

我々は、マルチドメインクラウド環境で仮想計算基盤を提供する資源管理フレームワーク GridARS を提案している<sup>1)</sup>。GridARS は、資源管理サービス<sup>2)</sup>、プロビジョニングサービス<sup>3)</sup>、モニタリングサービス<sup>4)</sup>を提供し、予約ベースでマルチドメインクラウド上の資源の選択と確保、仮想計算基盤の構築とユーザアプリケーションの実行、モニタリング情報の提供を自動的に行う。デモンストレーションでは、クラスタ上に構築した仮想マルチドメインクラウド環境を用いて、本フレームワークが実際に動作する様子を示す。

## 2. GridARS 資源管理フレームワーク

図1に提案する GridARS フレームワークの概要を示す。橙色が資源管理、水色がプロビジョニング、紫色がモニタリングのサービスモジュールを表す。

資源管理サービスは、資源コーディネータと各資源を管理する資源マネージャからなる。資源マネージャは予約機能を持つ既存のローカルスケジューラと連携することができる。資源コーディネータは各資源マネージャから利用可能な資源の情報を取得し、0-1 整数計画問題にモデル化した手法<sup>2)</sup>を用いてアプリケーションの要求を満たす資源群の予約プランを作成する。

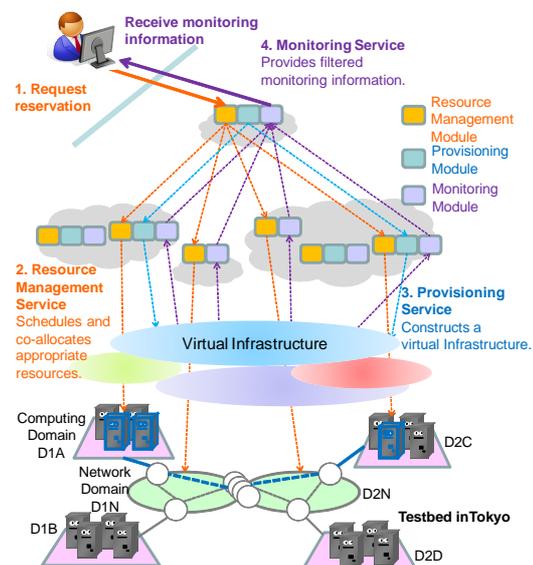


図1 GridARS 資源管理フレームワークの概要

ここで、資源コーディネータは計算機やネットワークだけでなく、仮想計算基盤で用いる VLAN タグ ID や IP アドレスレンジも同時に割り当てる。予約プランのもとに、資源コーディネータは各資源マネージャが連携して実際に資源群を予約する。

プロビジョニングサービスは、ユーザのジョブの実行を司る。実行時間になった時点で、割り当てられた資源上に仮想計算基盤を構築し、ジョブを実行する。この際、割り当てられた計算機に対して動的に決定される IP アドレスを用いて、マルチドメインにまたがった仮想的なネットワークを構成する。構築した仮想計算基盤では、一般のクラスタ向けに書かれたプログラムが変更なしに動作する環境を提供する。計算機の仮想化環境としては Linux コンテナを用いる。

モニタリングサービスは、予約 ID をキーとして関連する資源マネージャと連携し、仮想計算基盤の利用状況をモニタリングする。モニタリングでは、Ganglia<sup>5)</sup> や perfSONAR<sup>6)</sup> のように資源全体を監視する

<sup>†</sup> 産業技術総合研究所 National Institute of Advanced Industrial Science and Technology (AIST)

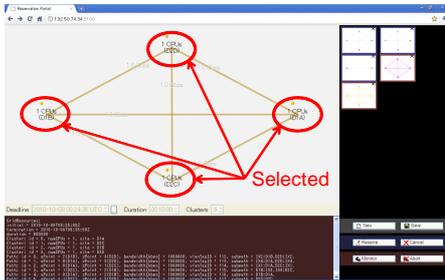


図 2 ポータル画面（予約結果）

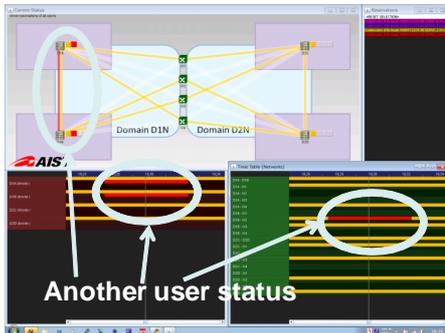


図 3 管理者用モニタリング結果

のではなく、構築した仮想計算基盤の利用状況のみを監視する．具体的には、計算機では Linux コンテナの cgroup インタフェースから情報を取得し、ネットワークでは指定された VLAN タグ ID の情報をネットワーク機器から取得する．各モニタリング情報は各ドメインで管理され、認可された情報のみをモニタリングサービスが収集してユーザーに提供する．これにより、過負荷なときは仮想計算基盤の資源を増やし、不用な資源は解放するなど、仮想計算基盤上で実行されるアプリケーションの品質を保つことができる．

### 3. デモンストレーション

デモンストレーションでは、GridARS を用いてユーザーの要求に応じてマルチドメインクラウドが提供する資源が自動的に確保され、予約時刻に仮想計算基盤が構築され、仮想計算基盤のモニタリング情報が提供されることを示す．マルチドメインクラウド環境として、クラスタ上に 2 つのネットワークドメインと 4 つのクラスタを構築し、ネットワークエミュレータ GNET1 で WAN 環境を再現する．

図 2、図 3、図 4 にポータル画面と管理者およびユーザー用のモニタリング結果を示す．図 2 では、資源と時間に関する要求を入力しただけでクラスタとその間のネットワークが自動的に割り当てられたことを示している．また、図 3 では他のユーザーの予約状況が出力されているのに対し、図 4 では予約したユーザーの情

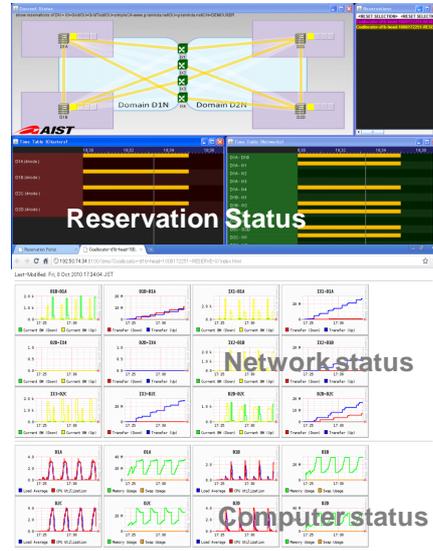


図 4 ユーザー用モニタリング結果

報のみが出力されているのが分かる．さらに、図 4 の下部分には構築された仮想計算基盤を構成する計算機の負荷やメモリ使用量、ネットワークの使用帯域と累積転送量が表示される．

謝辞 本研究の一部は、情報通信研究機構（NICT）の委託研究「ダイナミックネットワーク技術の研究開発」、「新世代ネットワークサービス基盤構築技術に関する研究開発」、および科研費（21700047）の助成を受けたものである．

### 参考文献

- 1) 竹房, 中田, 工藤, 田中, 関口: 多様な資源を事前予約で同時確保するためのグリッドコアロケーションシステムフレームワーク GridARS, 情報処理学会論文誌コンピューティングシステム (ACS20), Vol. 48, No. SIG18, pp. 32–42 (2007).
- 2) 竹房, 中田, 工藤, 田中: 性能を保証する分散実行環境のためのオンラインコアロケーション手法, 情報処理学会論文誌コンピューティングシステム, Vol. 3, No. 3, pp. 126–137 (2010).
- 3) 中田, 高野, 竹房, 工藤: ネットワーク帯域予約と OS 仮想化機構を用いた分散アプリケーション実行環境, コンピュータシステム・シンポジウム (ComSys 2009) 予稿集, pp. 51–58 (2009).
- 4) Takefusa, Nakada, Yanagita, Okazaki, Kudoh, and Tanaka.: Design of a Domain Authorization-based Hierarchical Distributed Resource Monitoring System in cooperation with Resource Reservation, *Proc. HPC Asia 2009*, pp. 77–84 (2009).
- 5) Ganglia Monitoring System: <http://ganglia.info/>.
- 6) perfSONAR: <http://www.perfsonar.net/>.