

MCJ-CloudHub における計算資源量推定機能、及び進捗状況可視化機能の設計

大江 和一¹⁾, 齊藤 智也²⁾, 筒井 優子¹⁾, 丹生 智也³⁾, 西井 淳²⁾,
岡田 耕一²⁾, 爲末 隆弘²⁾, 王 躍²⁾, 竹房 あつ子^{1),4)}

1) 国立情報学研究所

2) 山口大学

3) 国立遺伝学研究所

4) 総合研究大学院大学

koe@nii.ac.jp

Design of Resource Estimation and Progress Visualization Functions for MCJ-CloudHub

Kazuichi Oe¹⁾, Tomoya Saito²⁾, Yuko Tsutsui¹⁾, Tomoya Tanjo³⁾, Jun Nishii²⁾,
Koichi Okada²⁾, Takahiro Tamesue²⁾, Yue Wang²⁾, Atsuko Takefusa^{1),4)}

1) National Institute of Informatics

2) Yamaguchi University

3) National Institute of Genetics

4) The Graduate University for Advanced Studies, SOKENDAI

概要

本研究は、大学ごとに運用形態が異なるコンピュータシステム環境上において容易に構築・運用が可能なオンライン講義演習システムの構築・運用技術の確立を目指している。この技術により、情報システム部門が手薄な大学においても、オンライン講義演習システムの構築や運用が可能になる。その実現のために、山口大学のオンプレミス環境で運用中の講義演習システム（Moodle+JupyterHub+Nbgrader+山口大学独自機能）の汎用化を行い、さらに国立情報学研究所が提供している学認クラウドオンデマンド構築サービス（OCS）からの構築・運用が可能となる MCJ-CloudHub (Multiple Course Jupyter-based Cloud Hub) の開発を進めている。本報告では、MCJ-CloudHub の新規機能として開発を進めている計算資源量推定機能、及び進捗状況可視化機能について説明する。

1 はじめに

近年、世界各国で AI やデータサイエンス (DS) に関する人材育成が重要視されており [1, 2]、プログラミング演習及びデータ処理演習を含む実践的教育を推進するため、Web 型プログラミング教育支援システムの需要が高まっている。プログラミング演習を伴うオンライン講義を効率的・効果的に進めるためには、受講生間でプログラム開発・実行環境の設定やライブラリのバージョン等が統一されていること、課題の配布・回収が容易であること、受講生の進捗状況を把握可能であること、並びに時間帯や場所を問わず課題への取り組みが可能であることが求められる。

山口大学では、文系学部を含む全学部におけるデー

タサイエンス関連科目の標準的な演習環境とするため、Moodle[3] と JupyterHub[4]+nbgrader[5] を組み合わせた Web 型プログラミング教育支援システムを構築・運用している [9]。このシステムは、受講生が選択した科目ごとに統一した実行環境が提供され、nbgrader から課題の配布・回収が可能であり、Web 型であるため授業時間以外でも課題への取り組みが可能である。さらに、各学部の授業担当者による JupyterHub の利用開始を容易にし、且つ、システム管理者の管理業務の負担を大幅に削減する目的で、単一の JupyterHub システムにおいて、授業科目ごとに受講生一覧や個別の共有フォルダ等を提供しつつ、多数の授業科目が自由にかつ同時にシステムを共同利用することが可能となるシステム構成を採用した。授

業科目ごとの利用環境は自動的に生成され、利用者の Jupyter 環境に適用される。授業科目名や受講生一覧の情報も自動的にシステムに反映されるため、授業科目ごとの設定作業等を必要としない特徴も有する。しかしながらこの山口大学のシステムは、山口大学のオンプレミス環境を前提にした設計であり、クラウド環境への拡張や他機関環境での運用は困難である。

そこで我々は、山口大学の Web 型プログラミング教育支援システムを国立情報学研究所 (NII) が提供している学認クラウドオンデマンド構築サービス (OCS) [10] から構築・運用が可能となる MCJ-CloudHub (Multiple Course Jupyter-based Cloud Hub) を開発した。MCJ-CloudHub は他機関での運用を目指しており、1) 山口大学環境依存の設定、実装の分離、2) ユーザ認証や属性情報取得方法の汎用化、3) 学習管理システムからの履歴者一覧取得方法の汎用化、4) 最新の OSS を前提にしたシステム構築、を行ったうえで、OCS のアプリケーションテンプレート [11] として実装した。MCJ-CloudHub は、昨年度開発と山口大学での試運用まで完了した。今年度は、山口大学での試運用を継続して進めており、他機関での利用を促す目的で GitHub からソース公開も実施した [12]。さらに、より使い勝手を向上させる目的で計算資源推定機能と進捗状況可視化機能の開発を進めている。計算資源推定機能は、演習で使用する Jupyter Notebook を事前実行してそのメトリクス情報を収集しておき、その情報を分析することで演習に必要な計算資源を推定する機能である。進捗状況可視化機能は、学生単位で演習実施時と実施後の時間経過ごとの進捗状況を可視化する機能である。本報告では、この新規開発機能を中心に MCJ-CloudHub 普及のために取り組んでいる内容について説明する。

2 基盤ソフトウェア VCP の概要

国立情報学研究所 (NII) クラウド基盤研究開発センターは、オンプレミスやクラウドのアクセス方法の違いをユーザが意識することなくアプリケーション環境の構築・運用が可能な学認クラウドオンデマンド構築サービス (OCS) [10] の運用を行っている。OCS は、仮想プライベートネットワーク (VPN) 内に利用する資源を囲い込み、仮想クラウドコントローラ (VC コントローラ) から操作することで、全ての資源を統一的に操作できる。また、VC コントローラは VCP SDK と呼ばれる Python ライブラリを用いて制御可能であり、典型的な操作は可読性が高い Jupyter Notebook

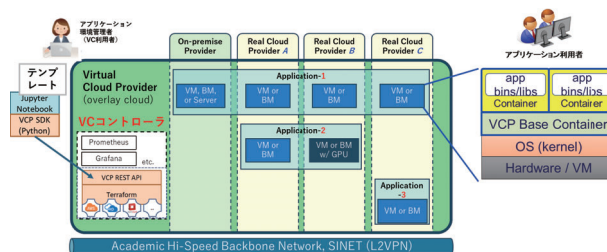


図1 VCPの概要

形式で記述して OCS テンプレートとして提供している。これにより、利用者は OCS テンプレートから容易に操作ができる。

この OCS を実現する基盤ソフトウェアは VCP (Virtual Cloud Provider) [13] である。図1は、VCP の概要図である。VCP は、SINET に接続された1つ以上のクラウドやオンプレミスの物理計算機や仮想計算機を仮想プライベートネットワークで分離し (L2VPN)、1つの仮想クラウドとして利用できる機能を提供する。

アプリケーションの配備では、まず、オンプレミス若しくはクラウド環境の仮想/物理計算機 (VM/BM) を確保する。この確保した計算資源を VC ノードと呼ぶ。次にその上に VCP Base Container が構築され、さらにその上に Application Container が構築される。このような利用形態は、Docker-in-Docker と呼ばれており、多少の性能劣化が懸念されるものの、柔軟なアプリケーション・ライブラリの配備を可能にする。VCP Base Container は、ネットワーク設定や起動した VC ノードの死活監視を行う。Application Container は、目的のアプリケーションで必要とされるコンテナを配備する。VCP Base Container は各 VM 又は BM に1つ配備されるものであり、Application Container は VCP Base Container 上に複数配備することができる。

OCS では、MCJ-CloudHub 以外にも Moodle を用いた学習管理システム (LMS) 構築、HPC クラスタ環境構築、Open OnDemand 環境構築テンプレートをすでに公開している [11]。今後もユーザからの利用が見込まれるテンプレート開発を進める予定である。

3 MCJ-CloudHub の概要と新規開発機能

3.1 MCJ-CloudHub の概要

MCJ-CloudHub は、山口大学のオンプレミス環境上で運用が行われている Web 型プログラミング教育支援システム [9] をベースに他機関でも運用可能とな

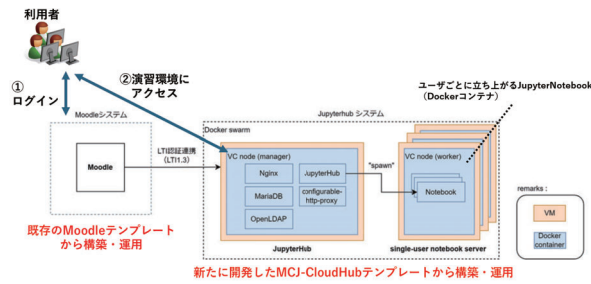


図 2 MCJ-CloudHub の概要

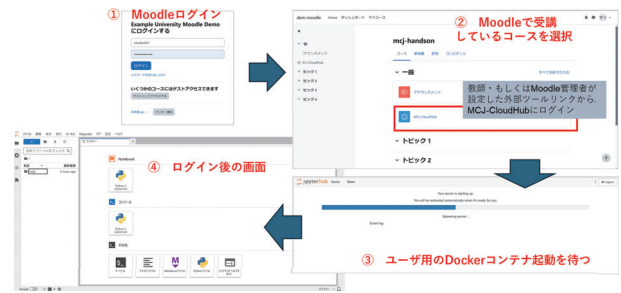


図 3 MCJ-CloudHub への login 方法

るように拡張したシステムである。MCJ-CloudHub は他機関での運用を目指しており、1) 山口大学環境依存の設定、実装の分離、2) ユーザ認証や属性情報取得方法の汎用化、3) 学習管理システムからの受講生一覧取得方法の汎用化、4) 最新の OSS を前提にしたシステム構築、が山口大学で運用中のシステムとの違いである [14, 15]。

図 2 は MCJ-CloudHub の概要である。MCJ-CloudHub は、教員や受講生が選択した科目ごとに統一した実行環境が提供され、nbgrader から課題の配布・回収が可能であり、Web 型であるため授業時間以外でも課題への取り組みが可能である。全ての Web からの GUI 操作で完結する設計である。

システムアーキテクチャの観点では、単一の JupyterHub システムにおいて、授業科目ごとに受講生一覧や個別の共有フォルダ等を提供しつつ、多数の授業科目が自由にかつ同時にシステムを共同利用することが可能である。また、授業科目ごとの利用環境は自動的に生成され、利用者の Jupyter 環境に適用される。授業科目名や受講生一覧の情報も自動的にシステムに反映されるため、授業科目ごとの設定作業等を必要としない。これら機能により、システム管理者の管理業務を大幅に減らすことも可能である。

MCJ-CloudHub の具体的なアクセス方法について説明する。利用者（教員・学生）は、まず、MCJ-CloudHub と連携する Moodle 側にログインし、MCJ-CloudHub を使用するコース選択を行う。すると、Moodle から JupyterHub に対して利用者の Jupyter Notebook 環境立ち上げ要求が出される。この立ち上げ要求の中に利用者が Moodle 選択したコース情報が添付される。JupyterHub は、利用者のコース情報に対応する教材などを展開した Jupyter Notebook を立ち上げ、利用者の演習環境が立ち上がる（図 3 参照）。

MCJ-CloudHub の構築は OCS から行う。最初に既存の Moodle テンプレートを用いて Moodle 環境の立ち上げを行い、その後、新規開発した MCJ-

CloudHub テンプレートを用いて JupyterHub システムの立ち上げと Moodle 側との連携設定を行うことで利用可能となる。

3.2 新規開発機能

3.2.1 計算資源推定機能

MCJ-CloudHub の運用を円滑に進めるためには、教材の内容やコースに参加する利用者に見合った計算資源を事前に割り当てておく必要がある。MCJ-CloudHub の構築・運用を行う VCP は、Prometheus[7] と Grafana[8] がインストールされており、MCJ-CloudHub 運用時のメトリクス情報収集を行うことが出来る。そこで、コースで用いる Jupyter Notebook を事前実行してその時のメトリクス情報を VCP の Prometheus に収集し、その情報を VCP の Grafana 用いて可視化する機能を開発した。この機能により、コースで用いる Jupyter Notebook を単体で動かした時に必要な計算資源量の把握は可能になる。

さらに、指定した利用者数で Jupyter Notebook の起動確認を行う機能開発も行った。この機能は、コースに参加する利用者数で問題なく Jupyter Notebook が実行出来るのかを確認する機能である。この機能も VCP の Prometheus と Grafana を用いて計算資源が足りているのか等を確認する。

進捗状況と今後の計画について説明する。両機能とも、2024 年 7 月に実施した山口大学の演習授業（90 人規模）向けトライアルで試用し、正常に動作することを確認した。現在の実装では、各 Jupyter Notebook で利用可能な計算資源量の上限値設定や MCJ-CloudHub 全体で準備するサーバ数などは、収集したメトリクス情報を用いて人手で行う前提である。今後は、現在人手で行う操作の自動化を検討していきたい。

3.2.2 進捗可視化機能

次に進捗可視化機能に関して説明する。MCJ-CloudHub は、利用者（学生）がリアルタイムに参

セル番号	項目番号	511	512	513	521	522	523	524	525	526	527	528	529	540
cell_1	2.1	o			o	o	o	o	o	o	o	o		
cell_2	2.2	o			o	o	o	o	o	o	o	o		
cell_3	2.2	o			o	o	o	o	o	o	o	o		
cell_4	2.2	o			o	o	o	o	o	o	o	o		
cell_5	2.3	o			o	o	o	o	o	o	o	x		
cell_6	2.3	o			o	o	o	o	o	o	o	x		
cell_7	2.3	o			o	o	o	o	o	o	o	x		
cell_8	2.4	o			o	o	o	o	o	o	o	x		

図 4 演習実行時の可視化例

加する演習での利用を想定している。演習を円滑に進めるには、演習実行時の各学生の状況把握が必要である。図 4 は、開発した演習実行時可視化機能である。縦軸は各演習課題（CELL 番号）であり、横軸は各利用者を表す。利用者 528 は CELL5 がエラーとなっており、利用者 512,513,529,540 は全く演習課題が進んでいないことが分かる。リアルタイムに更新されるため、各学生の現在の進捗状況の把握が可能となる。

また、演習終了後の分析機能（例、演習課題のどの部分につまずく学生が多いのか）も必要である。そこで群馬大学における可視化事例 [16] などを参考に、時間経過と共に各学生の進捗推移を可視化する機能の開発を行い、動作確認済である。

今後は、MCJ-CloudHub を用いて演習を行う先生方からご要望が出た機能の追加を検討していきたい。

4 まとめ

山口大学と国立情報学研究所（NII）は、MCJ-CloudHub の開発を共同で進めている。MCJ-CloudHub は、複数科目で共同・同時利用が可能な Web 型プログラミング教育支援システムである。他機関においても容易に導入出来るようする目的で、MCJ-CloudHub は学認クラウドオンデマンド構築サービス（OCS）から構築・運用出来るように開発し、GitHub から構築に必要な OCS テンプレート公開済みである。

本報告では、MCJ-CloudHub の使い勝手向上を向上させる目的で開発を進めている計算資源推定機能と進捗可視化機能について説明した。計算資源推定機能は、演習時に計算資源不足が起きないようにする目的で、演習で使用する JupyterNotebook を事前実行してそのメトリクス情報を収集し、その情報を分析することで演習に必要な計算資源を推定する機能であり、動作確認まで完了した。現在の実装は、メトリクス情報から計算資源を確保するところは人手となっている。この人手の部分を自動化するのが今後の課題である。

進捗状況可視化機能は、学生単位で演習実施時と実施後の時間経過ごとの進捗状況を可視化する機能である。現在、基本的な機能の実装が終わった段階である。今後は、演習等での試用を進め、教員のからの意見を踏まえて機能拡充を行う予定である。

謝辞

本研究は 2024 年度国立情報学研究所公募型共同研究（24S0802）と情報・システム研究機構“戦略的研究プロジェクト”（2024-SRP-06）の支援を受けたものである。

本研究にご協力いただいた株式会社アスケイドの那須野淳様、羽鳥文子様、増山隆様、神田千咲様、鳥井貴史様、寺山流晟様に深く感謝いたします。

参考文献

- [1] Jeremiah W. Johnson, Karen H. Jin; Jupyter Notebooks in education; Journal of Computing Sciences in Collages, Volume 35, Issue 8, pp.268-269, 2020, April
- [2] Saman Rizvi, Jane Waite, Sue Sentance; Artificial Intelligence teaching and learning in K 12 from 2019 to 2022: A systematic literature review; Computers and Education: Artificial Intelligence vol.4 100145 p.1-15;
- [3] MOODLE: <https://moodle.org/>
- [4] JupyterHub: <https://jupyter.org/hub>
- [5] Nbgrader: <https://github.com/jupyter/nbgrader>
- [6] JupyterHub: <https://jupyter.org/hub>
- [7] Prometheus: <https://prometheus.io/>
- [8] Grafana Labs: <https://grafana.com/ja/>
- [9] 齊藤智也, 王躍, 西井淳, 河野倫華, レールマルク; Moodle と JupyterHub を用いた Web 型プログラミング教育環境の改善および運用; IPSJ CLE39, Fukuoka, Japan; 2023, March
- [10] OCS: <https://cloud.gakunin.jp/ocs/>
- [11] 学認クラウドオンデマンド構築サービスのアプリケーションテンプレート: <https://github.com/nii-gakunin-cloud/ocs-templates>
- [12] MCJ-CloudHub template: <https://github.com/nii-gakunin-cloud/mcj-cloudhub>
- [13] Takefusa, A., Yokoyama, S., Masatani, Y., Tanjo, T., Saga, K., Nagaku, M. and Aida, K.:

Virtual Cloud Service System for Building Effective Inter-Cloud Applications, Proc. IEEE CloudCom2017, pp. 296-303 (2017).

- [14] 齊藤智也, 大江和一, 西井淳, 岡田耕一, 爲末隆弘, 王躍, 筒井優子, 丹生智也, 竹房あつ子; 複数科目で共同・同時利用可能な Web 型プログラミング教育支援システムのアプリケーションテンプレート開発; IPSJ 教育学習支援情報システム (CLE) 研究会 第 42 回研究発表会
- [15] 大江和一, 齊藤智也, 筒井優子, 丹生智也, 西井 淳, 岡田耕一, 爲末隆弘, 王 躍, 竹房あつ子; 複数科目で共同・同時利用可能な Web 型プログラミング教育支援システムの開発; IEICE CPSY 研究会 (SWoPP 2024)
- [16] 青木悠樹, 浜元信州, 大江和一, 横山重俊, 井上仁, 竹房あつ子; 学習者の理解度を時系列把握した Python 教育; 教育システム情報学会 2024