

Notizen aus dem 2. RWTHjupyter Community-Treffen

Termin: 28.01.2022 11.00 – 12.00 Uhr

Online Veranstaltung (Zoom)

Vorstellung aktueller Stand RWTHjupyter

Frage: Gibt es jetzt eine Regellinuxversion im IT Center?

Antwort: ja und nein, es gibt eine Zwischenregelung für dieses Jahr, weitere Wechsel sind aber nicht ausgeschlossen

GPU Nutzung

- Kubernetes kann GPUs nur einem User zuweisen
- Mit einem neuen Scheduler sind mehrere Zuweisungen möglich, allerdings mit einer möglichen Überschneidung der VRAM-Nutzung
 - eine mögliche Erweiterung wäre das Terminieren von Jobs, die die VRAM-Grenzen überschreiten
 - Wir sind offen für weitere Vorschläge, wie wir hier vorgehen
- Ergänzung zu GPU:
 - Herausforderungen: Einhaltung der Limits, Zugriffsregelung
 - Schön wäre eine Zuordnung der Ressourcen in bestimmten Zeitslots an bestimmte Kurse
- Warum braucht es physikalische GPUs?
 - Weil einige Machine-Learning-Use-Cases die Power brauchen
 - Alternativ könnten auch virtuelle GPUs genutzt werden, falls der Use Case dies erlaubt

Projekte

1) Second Life for Energiewende

- Unterscheidet sich von den bisherigen Projekten deutlich
- Studierende sollen verschiedene Konfigurationen von Energienetzen durchführen
- Ähnlich dem SecondLife-Ansatz sollen die Teilnehmenden sich in einer virtuellen Simulationsumgebung gemeinsam bewegen
- Die Simulation wird dabei in Echtzeit durchgeführt
- Unterschieden wird das Projekt von "normalen" Jupyter Notebooks dadurch, dass ein eigener Kernel zur Simulation genutzt wird,
- das Jupyter-Notebook wird für das Angebot der Dokumentation, den Einstieg und den Blick auf die Simulation genutzt
- Alle Komponenten laufen im RWTHjupyter-Cluster als Kubernetesprozesse, wodurch die Last vieler teilnehmender Studierender aufgenommen und verteilt werden kann
- Es wird ein Video gezeigt, die gezeigten Handlungen in der Simulation wären von allen Nutzenden einzusehen

- Es werden die Authentifizierungsmechanismen dargestellt
- Fragen:
 - (Uni Bonn) würde das auch gerne so implementieren
 - Realisation über das Deligieren der Simulation in einzelne Container in der Kubernetes-Umgebung, Inhalte sind Open Source im GitLab verfügbar
 - Standard PAM-Authentication-Modul genutzt?
 - Nein, ist wahrscheinlich nur für Linux-Systeme. Im Jupyter werden Docker-Container genutzt, die eine andere Authentifizierung per Shibboleth nutzen. Die Strukturen unterscheiden sich davon maßgeblich.
 - Frage zur Authentifizierung: (Uni Bonn) wollte externe Nutzende per DFN-AAI zulassen, der Datenschutzbeauftragte hatte hier die Befürchtung, dass zu viele Nutzende Zugriff hätten.
 - DFN-AAI wird per White-List freigeschaltet. Geprüft wird auch noch die weitere Einschränkung des Nutzendenkreises per Shibb. In der RWTH wird hier noch die Moodle-Information eingesetzt.

2. Jupyter4Power

- Ausblick auf das Projekt mit einer italienischen und einer spanischen Universität
- Erweiterung des Login-Kreises per eduGAIN für ausgewählte Partner möglich

Vorstellung HPCJupyterHUB für HPC

- Mündliche Vorstellung des Standes
- Zielsetzungen:
 - Ansatz ist die Bereitstellung eines Zugangs per UI zum HPC-Cluster, statt eines Konsolen-basierten Zugangs
 - Multicore und Multinode sollen über diesen Zugang auch bereitgestellt werden, um auch existierende Herangehensweisen nutzen zu können
- Aktuell noch einige Herausforderungen, die im Vergleich der Systeme aufgefallen sind
- Im HPC-Umfeld soll Singularity statt Kubernetes genutzt werden und als Scheduler SLURM
- Es wird die HPC-Umgebung und damit auch die Software und der etablierte Konsolenzugang genutzt werden können
- Views zeigen die Nutzung der Ressourcen
 - LiveDemo per ScreenShare zeigt Simple und Advanced View, über die Ressourcen angefordert werden können
- Die Bereitstellung von Ressourcen aktuell noch langsam, Überprüfung der Bereitstellung von dedizierten Ressourcen zum Beschleunigen
- Darstellung der Nutzung der Cluster-Ressourcen mittels Jupyter-Notebook
- Aktuell noch in Klärung:
 - Baremetal, Umstellung auf Singularity in Vorbereitung
- Frage: Wie können mehrere Knoten genutzt werden?
 - Alvaro: Für den Nutzenden sind alle Ressourcen z.B. CPUs in der Umgebung nutzbar

RWTHjupyter Community

- Hinweis auf die zentrale Seite für das Jupyter Board <https://www.itc.rwth-aachen.de/go/id/rfzei>
- Einstiegsoption wurde schon per Mailingliste bekannt gegeben

Vorstellung der Beteiligungsprozess

- Zusammenspiel der verschiedenen Elemente bezüglich der Entwicklung des Services
 - Anfragen werden daraufhin überprüft, ob es sich um Anfragen zur Entwicklung handelt und in diesem Falle in das Issue-System eingetragen
 - Das Jupyter Board wird diese Issues beobachten, um ggf. Interessenten zusammenzubringen oder diese nach Möglichkeit umzusetzen oder zu priorisieren
 - Als Beispiel wird z.B. die automatisierte Abgabe und Auswertung von Übungsblättern zu ermöglichen genannt
 - Herangehensweise wäre eine Zusammenführung der Interessenten zu einer Gruppe
 - Hier dann ggf. gemeinsame Umsetzung oder Betrachtung der Möglichkeiten aus der weltweiten Jupyter-Community
 - Das IT Center unterstützt die Umsetzung mit eventuell notwendigen Anpassungen
 - In einer Staging-Umgebung kann die Umsetzung von Entwicklungen und Neuerungen getestet werden
 - Die eigentliche Entwicklung neuer Punkte wird nicht durch das IT Center erfolgen, sondern hier unterstützt
-
- Frage: Wie kann ich Probleme ans IT Center bringen? Offensichtlich nicht per Issue-Tracker. Das Anlegen eines neuen Profils wird über ein Template-Projekt durchgeführt
 - Antwort: Es geht bei dem gezeigten Prozess nicht um das Erstellen neuer Profile oder Supportanfragen, sondern um Bug Fixes und Feature Requests. Das Anlegen von Profilen erfolgt nicht in diesem Prozess
 - Frage: HPC wird immer interessanter für den akademischen Betrieb. Können wir das JupyterHUB dafür nutzen?
 - Antwort: Förderung für RWTHjupyter/JupyterHUB ist für die Lehre ausgerichtet, HPC entsprechend dem Kontext. Erweiterungen könnten angefragt werden.
-
- Anfrage für ein weiteres Nutzendentreffen zum Austausch: soll per Abstimmung über die Mailingliste erfolgen