

Queuesystem SLURM

Zusammenfassung

Im November 2016 wurde mit Arbeitsgruppen aus der Informatik, Physik, Biologie und Chemie ein HPC-Cluster für Forschung und Lehre in Betrieb genommen. Dieser Cluster wird vom URZ administrativ betreut und bietet den beteiligten Arbeitsgruppen antragslos Rechenkapazität für ihre numerischen Forschungsprojekte zur Verfügung. Der HPC-Cluster wurde in zwei Phasen aufgebaut und wird kontinuierlich erweitert. Derzeit umfasst der Cluster 313 Rechenknoten, von denen einige zusätzliche GPU-Beschleuniger oder große Arbeitsspeicher besitzen.

Diese Anleitung richtet sich besonders an folgende Zielgruppen:

- **nur am Antrag beteiligten Arbeitsgruppen**

SLURM (Simple Linux Utility for Resource Management)

Laden: `module load slurm`

Verfügbare Partitionen:

Partition	Nodes	Rechenknoten	Shared	MaxTime
b_test	node001	Broadwell Standard	no	3:00:00
b_standard	node[002-126,131-136]	Broadwell Standard	yes	8-08:00:00
b_fat	node[137-140]	Broadwell Fat	yes	3:00:00
s_test	node141	Skylake Standard	no	8-08:00:00
s_standard	node[142-268]	Skylake Standard	yes	8-08:00:00
s_fat	node[269-272]	Skylake Fat	yes	8-08:00:00
gpu_p100	node[127-129]	2 NVIDIA Tesla P100	yes	8-08:00:00
gpu_v100	node130	2 NVIDIA Tesla V100	yes	8-08:00:00

Standardmäßig werden die Knoten im Shared- Mode genutzt, d.h. es sind mehrere Jobs und/oder mehrere Nutzer pro Knoten zugelassen. Dies wird über das Resource Allocation Plugin konfiguriert.

SelectTypeParameters=CR_CPU_Memory

D.h. neben den Cores wird auch der Speicher beachtet.

Ohne Angabe des benötigten Speichers geht SLURM davon aus, dass der gesamte Speicher des Knotens benötigt wird. Ein zweiter Job wird dann auf diesem Knoten nicht gestartet.

Mit dem Job- Parameter `--mem-per-cpu` allokiert SLURM entsprechend. Weitere Jobs können auf dem Knoten anlaufen.

SLURM Kommandos:

sinfo	Partition- und Nodestatus auflisten
squeue	Jobs auflisten
sbatch	Job abschicken
sstat	Status laufender Jobs
scancel	Job abbrechen
scontrol	Informationen über Konfiguration und Status
salloc	Ressourcenzuordnung für interaktive Jobs
srun	Jobschritte in Skript oder interaktiven Job starten
sview	graphisches Nutzerinterface

Job- Parameter:

Die Parameter können über die Direktive #SBATCH im Script definiert werden.

Parameter	#SBATCH
Job-Name	---job-name= <i>name</i>
Partition	---partition= <i>partition</i>
Anzahl Nodes	---nodes= <i>n</i>
Prozesse pro Nodes	---tasks-per-node= <i>n</i>
Walltime-Limit	---time= <i>hh:mm:ss</i>
Memory pro CPU	--mem-per-cpu==< <i>MB</i> >
Stdout	---output= <i>filename</i>
E-Mail-Adresse	---mail-user=address
E-Mail-Benachrichtigung	---mail-type=BEGIN ---mail-type=END ---mail-type=FAIL ---mail-type=ALL

Eine vollständige Beschreibung der parameter finden Sie in der [SLURM Dokumentation](#).

Titel: "Queuesystem SLURM"

Stand: 07.04.2020



FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA