

Einführung der Projektbewirtschaftung des RWTH Compute Cluster

26.8.2014

Paul Kapinos

Dieter an Mey

Marcus Wagner

- **Leitfaden zur Begutachtung von Forschungsbauten
Drs 3782-14 des WR vom 11.4.2014**

(Seite 15, sowie in Ergänzende Kriterien für die programmatisch-strukturelle Linie „Hochleistungsrechner“, Seite 42)

„Nachweis der Antragsteller, dass ein wissenschaftsgeleitetes Verfahren der Nutzung etabliert wird, welches sicherstellt, dass der Rechner Voraussetzung für die Durchführung von Forschungsprogrammen von hoher Qualität ist.“

- **Erfüllung der Kriterien des WR**
- **Weiterführung der Nutzungsmöglichkeit für Aus- und Weiterbildung**
- **Geringe Zeitverzögerung beim Starten eines Projektes**
- **Faire Behandlung der Nutzer bei knappen Ressourcen**
- **Strukturierter Einblick des RZ in die Aktivitäten der Nutzer um bessere Unterstützung leisten zu können**

■ Aufwand für

- Wissenschaftler (Anträge, Berichte)
- Aufwand für die Gutachter
- Aufwand für die Betreiber (Implementierung, Scheduling, Abrechnung)
- Verwaltung (Antragswesen, Begutachtung, Berichtswesen)

■ Beschränkung der Ressourcen

Wir bemühen uns diese Nachteile soweit wie möglich zu begrenzen.

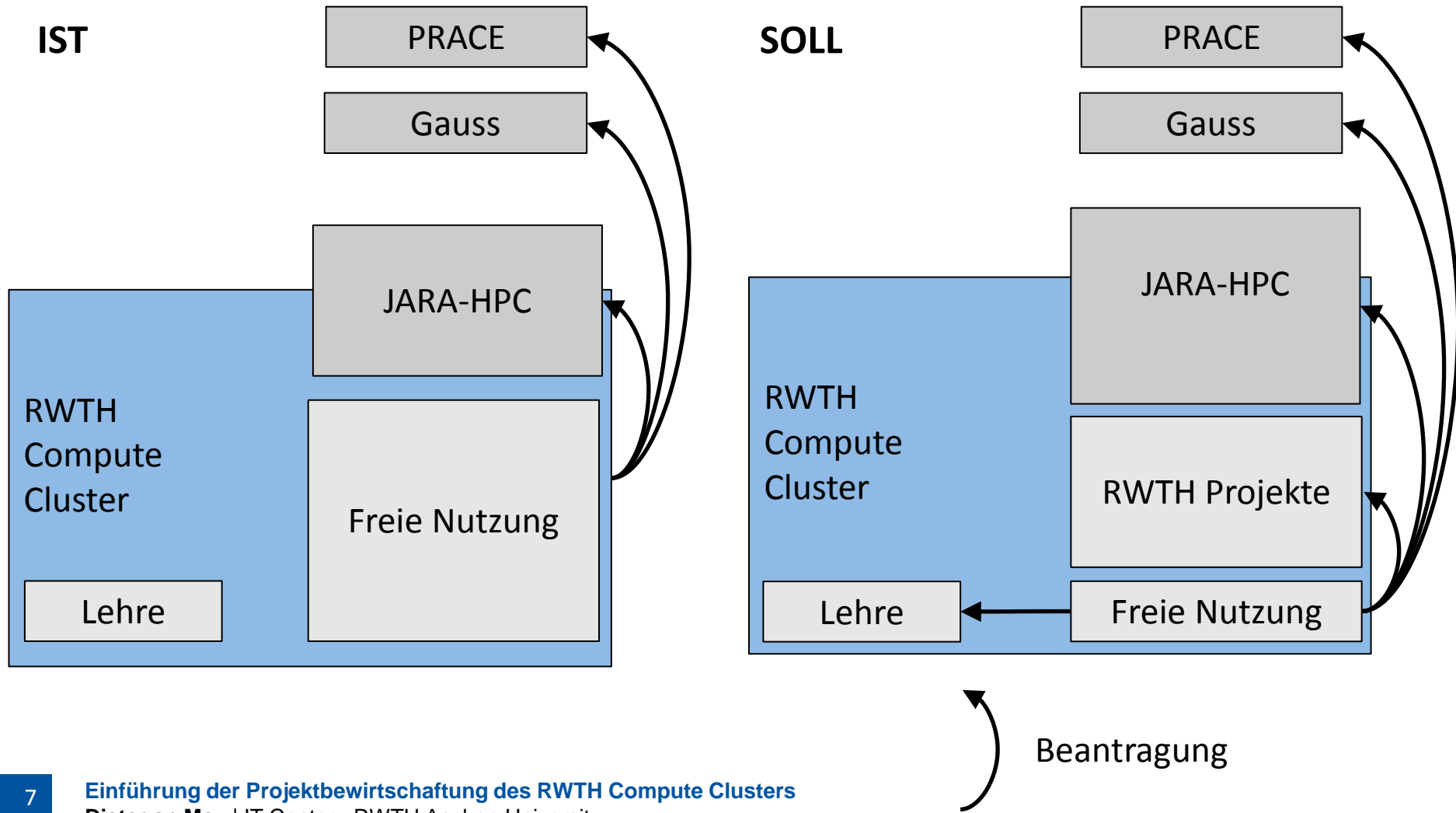
- **Bewirtschaftung von HPC Ressourcen ist üblich in den**
 - Tier-0: PRACE-Zentren der EU (GENCI, CINECA, HLRS, JSC, BSC, LRZ)
<http://www.prace-ri.eu/HPC-access>
 - Tier-1: Gauß-Zentren in D (JSC, HLRS, LRZ)
<http://www.gauss-centre.eu>
 - Tier-1.5: JARA-HPC (JSC, RWTH)
<http://www.jara.org/de/research/jara-hpc/partition/>
 - Tier 2: (einige) Zentren der Gauß-Allianz (RWTH,TUD,KIT,HLRN,ZIB,...)
- **Projektbezogenes Antragsverfahren**
- **Technische und wissenschaftliche Begutachtung**
- **Feste Rechenzeitperioden (meist zweimal pro Jahr, Dauer: 1 Jahr)**
- **Schnupperkontingente**
- **Berichtswesen**

■ Die Zentren der Ebenen Tier 0 – 1.5 (und einige der Tier 2 Zentren)

- bedienen ausschließlich die reine HPC-Nutzung
100% Projektierung – Fokus auf skalierbare Parallelisierung
- LPC-(Low Performance Computing)-Nutzer werden auf andere Ressourcen verwiesen
- Ausbildung und Lehre vorrangig für HPC-Kunden mit Vorwissen
- (Über-) regionale Nutzerschaft und auch Gutachterreservoir

■ Im RWTH Compute Cluster

- möglichst nahtloser Übergang vom LPC zum HPC erwünscht,
- Heranführen von vielen Studenten der Computational Sciences
- HPC- und LPC-Nutzung in Ausbildung und Lehre integriert.



- **Technische und wissenschaftliche Begutachtung für “große und mittlere Rechenprojekte” in Anlehnung an die bewährten Verfahren des JSC wie für die JARA-HPC Partition**

- **Reservierung von kleinen Kontingenten nach Projektbeschreibung ohne wissenschaftliche Begutachtung zeitlich flexibel im Rahmen von**
 - Kleinen Forschungsarbeiten
 - Schnupperkontingenten, vor der endgültigen Begutachtung
 - Lehrveranstaltungen
 - Abschlussarbeiten

- **Definierte Freikontingente für F&L**

Interne Strukturierung in Projektklassen und Aufteilung der Ressourcen



Projektkategorien	Anteil	Begutachtung
JARA-HPC	30%	Technische und externe wissenschaftliche Begutachtung
Große RWTH-Projekte	25%	
RWTH Standard-Projekte	25%	Technische und wissenschaftliche Begutachtung primär RWTH-intern und bei Bedarf durch externe Wissenschaftler
Kleine RWTH Projekte	10%	kurze Projektbeschreibung bei Bedarf technische Klärung
Lehre	10%	Keine Begutachtung
Freie Nutzung		

Welcher Kuchen ist zu verteilen? Ziel



- JARA-HPC ist derzeit überbucht, Kombination mit RWTH Big sinnvoll
- Hier: Anzahl Nutzer, Aber: Bewirtschaftung durch Projekte

Job Category	resources	Mio Core-h	#users 2013	avg resources per user	#users target 2015	avg resources per user 2015
Total	100%	200	>1000			
JARA-HPC	30%	60	100	0,60	40	1,50
RWTH Big	25%	50	10	5,00	50	1,00
RWTH Standard	25%	50	30	1,67	100	0,50
RWTH Small	10%	20	35	0,57	200	0,10
RWTH open	10%	20	925	0,02	1000	0,02
"Normale" Nutzer: (nicht JARA oder IH)			500 user > 50 Jobs			
			500 user > 2000 Core-h			
			140 user > 100.000 Core-h = 1 Westmere-Knoten-Jahr			

- **sukzessive Einführung der Projektierung bis zum Jahresende**
- **Definition der Projektklassen anhand der Nutzungsstatistik 2013 und sukzessive Anpassung**
- **Abrechnungsverfahren in Anlehnung an das JARA-HPC-Verfahren**
- **3-Monatsfenster**
 - in einem Monat können maximal 3 Monatskontingente verbraucht werden
 - Danach kann nur noch gerechnet werden, wenn kein anderer Job wartet.
- **Studentische Mitarbeiter können in Projekte aufgenommen werden**

- **Übernahme und Anpassung der Verfahrensweise und der Richtlinien des JSC**
- **Hinzunahme von externen Fachgutachter für große Rechenprojekte im JARA-HPC Verfahren**
- **Wissenschaftler, die selbst Rechenzeitanträge stellen, werden gebeten, sich als Gutachter zur Verfügung zu stellen**
- **RWTH-interne Begutachtung für Standardprojekte:**
 - Begutachtung durch Fachwissenschaftler und
 - durch Wissenschaftler anderer Fachbereiche
- **Die wissenschaftliche Begutachtung großer Projekte wird nachgelagert durchgeführt ab Q4 2014**
 - Vorab werden Schnupperkontingente eingerichtet

■ **Lehre soll nicht zu kurz kommen**

→ Implementierung mit Hilfe von Rechenprojekten

■ **Lehrveranstaltungen**

→ Blockveranstaltungen

→ regelmäßige Veranstaltungen

→ Projektleiter: Dozent

→ Technischer Kontakt: techn. oder wiss. Mitarbeiter

■ **Abschlussarbeiten**

→ Zusammenfassen von mehreren Abschlussarbeiten zu einem Projekt wird ermöglicht werden

→ Projektleiter: Erstgutachter

→ Technischer Kontakt: wiss. Mitarbeiter / Betreuer

■ Freie unkontingentierte Nutzung:

- Mitarbeiter 2000 Core/h pro Monat
- Studenten 500 Core/h pro Monat
- danach landen die Jobs in “low-prio” Queue, *können aber laufen*

■ Nutzer mit erhöhtem Ressourcenbedarf müssen zukünftig Projektanträge stellen

- Kleine Projekte bis 10000 Core-h pro Monat (unbürokratisch)
- mittlere Projekte (Begründung, technische und wiss. Begutachtung)
- große Projekte (ab 0.6M Core-h pro Jahr) => JARA (wie schon laufend)
- Antragsformulare im DokuWeb verfügbar

Project Category		Reviewing procedures	cpu time quota Mio Core-h per month		cpu time quota Mio Core-h per year	
			min	max	min	max
JARA-HPC / RWTH-Big (research projects)	XL / L	external scientific reviews	0,05	2,5 (current total JARA-HPC budget per call)	0,6	30 (current total JARA-HPC budget per call)
RWTH Standard (research projects)	M	internal scientific reviews	0,01	0,05	0,12	0,6
RWTH Small (research projects)	S	technical reviews only	0,002	0,01	0,024	0,12
RWTH open (scientists) ----- RWTH lecture (students) RWTH thesis (students)	XS	no reviewing ----- technical reviews only	0,0005	0,002	0,006	0,024
RWTH open(students)	XXS	no reviewing		0,0005		0,006

■ Nutzung des Batchsystems:

→ r_batch_usage

→ r_batch_usage -q

■ Nutzung des Batchsystems mit Projekten:

→ r_batch_usage -p <projectname>

→ r_batch_usage -p <projectname> -q

■ Drei-Monats-Fenster:

→ Restquota kann zwischen -101% und 200% schwanken

→ Nutzung des letzten Monats + Nutzung des aktuellen Monats muss kleiner sein als $3 * \text{Monatskontingent}$

→ Ansonsten in die low-queue, auch bereits submittierte, aber noch wartende Jobs, werden verschoben

- **lecture, thesis, jara, rwth, normal**

- Und zwar in der Reihenfolge von hoch priorisiert zu niedrig

- **Am Ende der Fahnenstange sitzt die low-Queue**

- Alle Jobs, die kein Quota mehr besitzen, landen in der low-Queue

- Jobs laufen dort nur an, wenn freie Ressourcen vorhanden sind,
d.h. keine Jobs mit höherer Priorität warten

- **bis zum 31.8.2014**

- jara-low Queue wird *noch* wie bisher gehandhabt

- low Queue hat *noch* die gleiche Priorisierung wie die normal Queue

- **Aktivierung der Prioritätenregelungen zum 1.9.2014**

■ Query your past compute resource consumption

```
da026566@cluster-linux:~[1180]$ r_batch_usage -p jara0001
batchusage from Sat 01.02,2014 00:00:00 to Mon 21.07,2014 23:59:59
of project jara0001 in corehours:
```

#	Queue +	User +	actual +	Jun 2014 +	May 2014 +	Apr 2014 +	Mar 2014 +	Feb 2014 +	total #
#	jara	ad846581				31,15			31,15 #
#		ct747764				7,05K	14,16K	41,19	21,25K #
#		dd794493		47,50					47,50 #
#		ds534486		385,03	515,06	1,07K	689,10	19,49K	22,15K #
#		fg716463	16,91K	1,28K					18,19K #
#		fr356676			1,83	5,91	3,08K	21,20K	24,30K #
#		fr35667a			0,28		6,73K		6,73K #
#		hh480926				100,89			100,89 #
#		hi853255			0,09		1,43K	1,09K	2,51K #
#		mb066302	524,37						524,37 #
#		mh391413	11,36		1,18	0,14	2,96	2,57	18,21 #
#		pk224850	54,92K	41,18K				45,94	96,15K #
#		tc530841			66,84	16,87K		8,11K	25,05K #
#	normal	fr356676					0,04		0,04 #
#	total		72,37K	42,89K	585,28	25,13K	26,09K	49,98K	217,04K #

`r_batch_usage -p <project> -q`



■ Query the status of your quota

```
da026566@cluster-linux:~[1179]$ r_batch_usage -p jara0001 -q
Group:                               jara0001
Start of Accounting Period:          01.05.2014
End of Accounting Period:            01.05.2015
State of project:                    active
Quota monthly (core-h):              100000
Remaining core-h of prev. month:     57110
Consumed core-h act. month:          72369
Consumable core-h (%):               85
Consumable core-h:                   184741
```

Rest vom Vormonat
aktueller Verbrauch
% verf. vom Monatsquota
incl quota vom nächsten
Monat

- Projektbewirtschaftung im Doku Web

→ <https://doc.itc.rwth-aachen.de/display/CC/Projektbewirtschaftung>

- Project-Based Management of Cluster Resource in the DokuWeb

→ <https://doc.itc.rwth-aachen.de/display/CC/Project-based+management+of+the+cluster+ressources>