

46. Sitzung des Wissenschaftlichen Lenkungsausschusses der Deutsche Klimarechenzentrum GmbH

Beginn der Sitzung: 1. Dezember 2023 um 10:00 Uhr

Teilnehmende

Prof. Dr. Arne Biastoch, GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel
Dr. Hendryk Bockelmann, DKRZ
Dr. Frauke Feser, Helmholtz-Zentrum Hereon (Vorsitz)
Dr. Bernadette Fritzsich, AWI Bremerhaven (Vorsitzende des DKRZ-Usergroup-Komitees)
Dr. Helge Goessling, AWI Bremerhaven
Dr. Christopher Kadow, DKRZ
Dr. Constanze Leemhuis, DLR-PT (BMBF)
Dr. Patrick Ludwig, Institut für Meteorologie und Klimaforschung, Karlsruher Institut für Technologie
Prof. Dr. Thomas Ludwig, DKRZ
Dr. Armin Mathes, DLR PT (BMBF)
Prof. Dr. Juan Pedro Mellado González, Meteorologisches Institut, Universität Hamburg
Dr. Karsten Peters-von Gehlen, DKRZ
Prof. Dr. Johannes Quaas, Institut für Meteorologie, Universität Leipzig
Dr. Mathis Rosenhauer, DKRZ (Protokoll)
Prof. Dr. Anja Schmidt, Deutsches Zentrum für Luft und Raumfahrt
Hannes Thiemann, DKRZ
Prof. Dr. Uwe Ulbrich, Institut für Meteorologie, Freie Universität Berlin
Dr. Sebastian Wagner, Helmholtz-Zentrum Hereon
Prof. Dr. Sönke Zaehle, MPI für Biogeochemie

1. Annahme der Tagesordnung

Die Tagesordnung wird angenommen.

2. Organisatorisches

a) Annahme des Protokolls der 45. Sitzung

Das Protokoll wird angenommen.

b) Ort und Termin der nächsten Sitzung

Die nächste Sitzung wird am 07.06.2024 in Hamburg stattfinden.

3. Bericht DKRZ

a) Nutzung HLRE-4 (H. Bockelmann)

Aufgrund von Kompensationsleistungen konnte die Zahl der CPU-Knoten im August um 153 gesteigert werden. Der Endausbau von HLRE-4 ist damit erreicht.

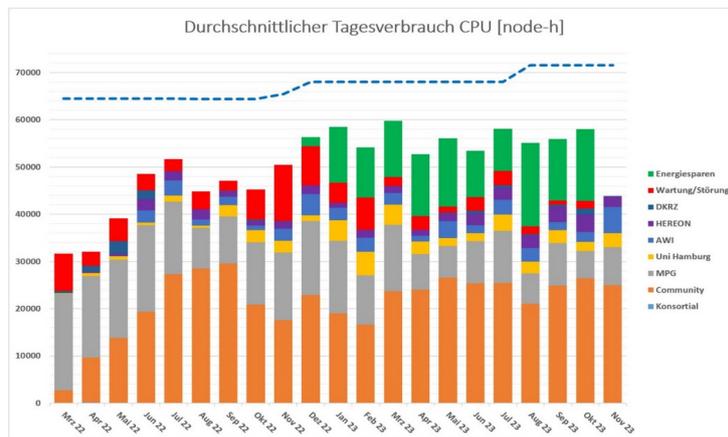


Abbildung 1: Durchschnittliche Auslastung der CPU-Rechenknoten von Levante

Die Auslastung des Systems ist weiterhin nicht sehr hoch. Insgesamt wird der Benutzerbetrieb durch Energiesparmaßnahmen nicht beeinträchtigt. Die Ausfallzeit von CPU-Knoten konnte weiter gesenkt werden und liegt nun nahe der Herstellerzusagen.

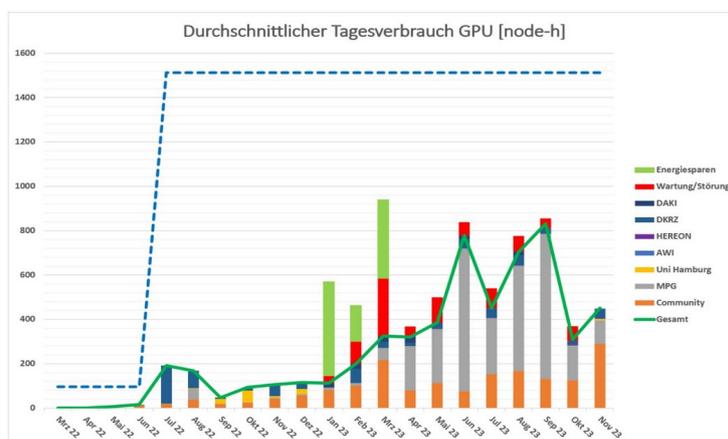


Abbildung 2: Durchschnittliche Auslastung der GPU-Knoten von Levante

Die Auslastung der GPU-Knoten ist noch gering. Aufgrund technischer Gegebenheiten können die Energiesparmaßnahmen, die bei CPU-Knoten vorgenommen werden, nicht auf GPU-Knoten übertragen werden. Im Sommer 2023 wurden große ICON-Läufe durchgeführt, die zeitweilig zu einer stärkeren Auslastung führten.

Ab Anfang 2024 wird der „fair share“ Mechanismus des Batch-Systems für die GPU-Partition aktiviert werden.

Die CPU-Knoten werden etwa zur Hälfte von Community-Projekten genutzt. Die Anteile der Gesellschafterprojekte bewegen sich im vorgegebenen Rahmen. Auf der GPU-Partition ist nur im erste Quartal ein Community-Anteil von etwa 55% zu verzeichnen, danach ist dann ein MPG-Projekt dominierend.

Die Hardware-Probleme im Infiniband sind grundlegender Natur und lassen sich absehbar nicht beheben. Es existieren jedoch „Workarounds“ sodass ein stabiler Betrieb gewährleistet ist.

Nach der Installation einer neuen Version des Batch-Systems SLURM können nun Jobs, die gleichzeitig GPU- und CPU-Knoten nutzen, zuverlässig durchgeführt werden.

Ein verbesserter Update-Prozess erlaubt kleinere Wartungsfenster.

Für Testzwecke sollen die GPU-Knoten auch weiterhin von Projekten ohne bewilligten GPU-Ressourcen genutzt werden können.

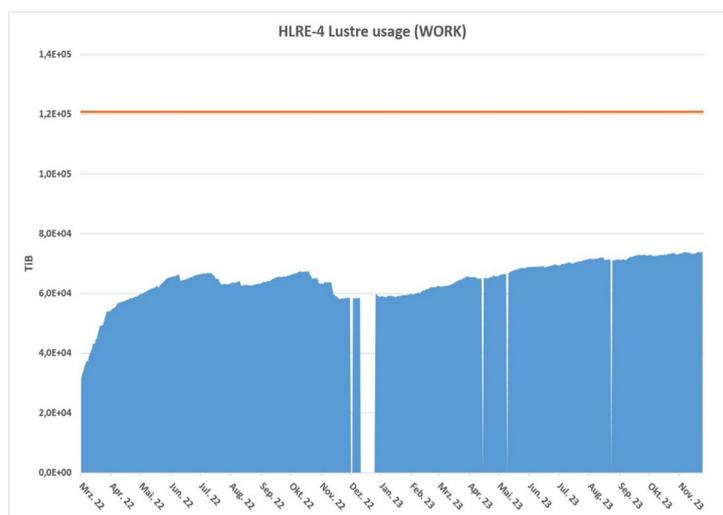


Abbildung 3: Belegung des Lustre Dateisystems auf Levante [TiB]

Die Auslastung des Disk-Systems ist noch unproblematisch und es erweist sich im Betrieb als sehr stabil. Die Nutzung des Scratch-Bereichs bleibt konstant unter 1,5PiB.

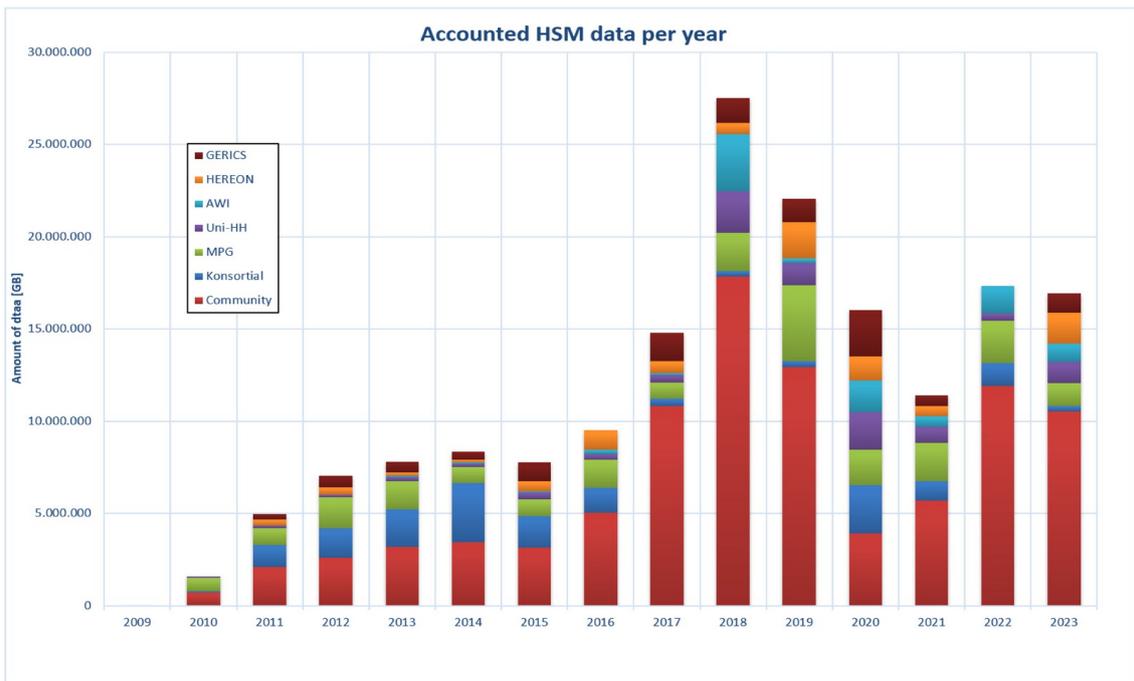


Abbildung 4: Belegung des Bandarchivs pro Jahr [GiB]

Nach der Installation neuer Tape-Libraries erweist sich der Betrieb des gesamten Archivs als deutlich stabiler. Das betrifft insbesondere das Schreiben neuer Daten ins Archiv. Das Lesen archivierter Daten (retrieval) ist noch teilweise schwierig.

Benutzer werden bei unvollständig archivierten Dateien nun per E-Mail auf das Problem hingewiesen.

Vom bewilligten Archivplatz wird nach wie vor nur ein kleiner Teil genutzt. Trotzdem wurde deutlich mehr beantragt als auf der zur Verfügung stehenden Bandkapazität gespeichert werden kann. Da von einer vollständigen Nutzung des bewilligten Archivplatzes ausgegangen werden muss, ist eine starke Kürzung notwendig.

b) /pool-data (H. Thiemann/ K. Peters)

H. Thiemann und K. Peters präsentierten den aktuellen Stand von /pool-data (s. Vortragsfolien im Anhang). Im Anschluss gingen sie auf einige Fragen des WLA ein.

Einige Projektanträge sind reine „Datenaustauschprojekte“, die den HPC überhaupt nicht nutzen. Diese Nutzung könnte auf einen weiteren Server ausgelagert werden.

Die tatsächliche Nutzung der Daten im HPC-Dateisystem ist schwierig zu erfassen. Anhand der Suchanfragen an das GEMS-System (<https://gems.dkrz.de/>) kann man jedoch Rückschlüsse ziehen. GEMS könnte in Zukunft auch Vorschläge für ähnliche Daten machen.

ERA6 in höherer Auflösung wird demnächst starten. Auch Reanalyse-Daten aus Japan könnten am DKRZ vorgehalten werden. Diese höher aufgelösten Daten hätten jedoch deutlich höhere Speicheranforderungen.

c) **Perspektiven (T. Ludwig)**

Das Festplattensystem von HLRE-4 wird möglicherweise auch mit HLRE-5 genutzt und ggf. nur aufgerüstet werden. Die Fragen zu Wartungsverträgen und Ersatzteilversorgung müssten allerdings noch geklärt werden. Vergleichbares könnte auch für Rechenknoten gelten.

Bezüglich Preisen und Verfügbarkeit von GPUs ist die Situation derzeit sehr ungünstig. Es besteht aber die Hoffnung, dass die Bedingungen bis zum Zeitpunkt der Beschaffung von HLRE-5 besser werden.

Die Nutzung der GPU-Knoten konzentriert sich derzeit nur auf ICON mit einem kleineren Anteil an ML-Aktivitäten. Der GPU-Anteil von HLRE-5 wird zentraler Gegenstand von Sitzungen mit Vertretern der Community im kommenden Jahr sein. Ergebnisse dieses Austauschs werden die Grundlage für die Ausschreibung des nächsten Systems bilden.

Im Rahmen von NatESM können auch andere Modelle von der Portierung auf GPUs profitieren. NatESM unterstützt allerdings nur deutsche Modelle, was die Auswahl der von der Community genutzten Modelle einschränkt.

Mit einem Teil der Abwärme von HLRE-4 werden die Labore der benachbarten Chemie beheizt. Das betrifft ca. ein Viertel der Energie. Weitere Abnehmer sind derzeit nicht in Sicht und es gibt auch keine Netze zur Einspeisung.

4. **Bericht aus der DKRZ-User-Group (B. Fritzsich)**

Im öffentlichen Teil der letzten Sitzung wurde der stabile Betrieb der Levante gelobt. Insbesondere der Einsatz der Mitarbeitenden des DKRZ wurde positiv erwähnt.

Es gibt zwei Stellenausschreibungen am DKRZ, deren Befristung auf 5 Jahre befürwortet wird.

Allgemein wird die Besetzung von Stellen mit längerer Befristung als erfolversprechender eingeschätzt.

5. **Neue Geschäftsordnung WLA**

Im weiteren Abstimmungsprozess zur Geschäftsordnung gab es keine größeren Änderungen an der ursprünglichen Version. Die angepasste Satzung des DKRZ wird am 4.12.2023 notariell beglaubigt. Im Anschluss wird auch die Geschäftsordnung des WLA beschlossen.

6. **Konsortialprojekte**

Die Kriterien für Konsortialprojekte sollen überarbeitet werden. Als weiteren Punkt beschließt der WLA eine Laufzeitbegrenzung auf fünf Jahre. Der Text wird im Umlaufverfahren finalisiert.

7. Sonstiges

Frau Dr. Constanze Leemhuis wird als weitere Vertreterin des BMBF in Zukunft an Sitzungen des WLA teilnehmen.

Für Mitte Januar 2024 ist eine Umfrage des DKRZ zu Nutzerzufriedenheit geplant. Die Fragen orientieren sich an einer Umfrage, die das HLRS regelmäßig durchführt. Sie wird auch Punkte zur GPU-Nutzung enthalten. Für die Zukunft ist eine jährliche Durchführung geplant.

8. Rechenzeitanträge (intern)

Im nichtöffentlichen Teil der Sitzung wurde unter anderem über die Rechenzeitanträge für Community- und Konsortial-Projekte beraten.

Es wurden Ressourcen für Neu- und Folgeprojekte über den Zeitraum vom 01.01.2024 bis 31.12.2024, sowie zusätzliche Ressourcen über den Zeitraum vom 01.01.2024 bis 30.06.2024 bewilligt. Im einzelnen sind dies:

	Beantragt	Bewilligt
Levante CPU [Node hours]	22.400.107	10.085.501
Levante GPU [Node hours]	150.199	145.520
Levante storage [TiB]	45.137	35.788
Archive project [TiB]	49.923	20.170
Archive long term [TiB]	6.541	4.453

Ende der Sitzung: 16:23

DKRZ Data Pool

@Levante: /pool/data



/pool/data Konzept

Zentraler Zugriffspunkt auf **relevante Datensammlungen** am DKRZ, lesbar für alle DKRZ Nutzer, u.a.

- Vom DKRZ bereitgestellte Sammlungen (CMIP, ERA) → Konsortialprojekte
- Initialisierungs- und Randbedingungen für am DKRZ betriebene Modelle (ICON, ECHAM, MPI-ESM, etc)
- Projektdatensammlungen mit potentiell Interesse zur Nachnutzung (EUREC4A, ReKlies, fesstval, HAPPI, ...)

Management der DKRZ Data Pool Projekte

Beantragte Datenvolumina

> 80 TiB pro Projekt

“Biggies”

- WLA Konsortialprojekte (ERA, CMIP)
Bedingungen:
 - Rechenzeitantrag
 - WLA genehmigt
- Gesellschafterprojekte (ICDC, EUREC4A)

<i>Anzahl Konsortialprojekte:</i>	2
<i>Levante Speicher:</i>	6,6 PiB bewilligt 6,3 PiB

< 80 TiB pro Projekt

“Smallies” (aka “pd”-Projekte)

- DKRZ Projektdaten, modellspezifische Daten, Feldkampagnen, ...
Bedingungen:
 - Antrag via support@dkrz.de
 - DM genehmigt (kleiner Dienstweg)
 - Speicherbedarf aller “smallies” < 1% des Levante-Speichers, d.h. <1,3 PiB
 - Dokumentation (standardisierte README) erstellt vom Dateneigner, von DM geprüft

<i>Anzahl Projekte:</i>	16
<i>Levante Speicher:</i>	810 TiB bewilligt 625 TiB

Nutzungsfrequenzen (Nachnutzung)

- Keine Statistik über Datenzugriffe vorhanden (technisch limitiert)
 - DKRZ Nutzer sollen die Verwendung der DKRZ Ressourcen im Generellen würdigen
- ⇒ Vorschlag: Nachnutzende von /pool/data Datensammlungen sollen die Zitiervorgabe im README berücksichtigen

Beispiel Zitiervorgabe README für /pool/data/ICON

METHODS USED FOR DATA CREATION

Details of the ICON model are described in

Giorgetta, M. A. et al. (2018). ICON-A, the atmosphere component of the ICON Earth system model: I. Model description. *J. Adv. Model. Earth Syst.*, 10, 1613–1637. doi:10.1029/2017MS001242

Korn, P. (2017). Formulation of an unstructured grid model for global ocean dynamics. *Journal of Computational Physics*, 339, 525–552. doi:10.1016/j.jcp.20171246.03.009

Datenverfügbarkeit

In /pool/data verlinkte/verfügbare Datensammlungen sind generell für alle DKRZ-Nutzer lesbar.

Sichtbarmachung und Nutzbarkeit

Sichtbarmachung

- Daten sind unmittelbar sichtbar über das Filesystem und die web-Doku, vermehrt über gems.dkrz.de
- In Zukunft: Einbindung in einen *zentralen DKRZ-Katalog*
- Kuration der Dokumentation aller /pool/data Daten in Arbeit

Nutzbarkeit

- Datencharakteristika inkl. Nutzungshinweise und Lizenz in READMEs beschrieben
- Kontaktdaten zu Datenverantwortlichen in READMEs und in web-Doku enthalten