

Umfrage zur Bedarfsermittlung von Ressourcen zum wissenschaftlichen Rechnen an der TU Dortmund

Das Wissenschaftliche Rechnen hat sich in Forschung und Lehre in den letzten Jahren zu einem wichtigen Standbein neben Theorie und Experiment entwickelt. Die Situation an der TU Dortmund in diesem Bereich ist unübersichtlich. Dies betrifft sowohl die Software- als auch die Hardwareseite. Viele Arbeitsgruppen bewegen sich in diesem Feld, ohne voneinander zu wissen. Mathematiker sprechen von der "Numerik partieller Differentialgleichungen", die Ingenieure verwenden den Begriff "FEM-Simulation" und die "Datenanalyse" wird in der Physik und Statistik verwendet. Trotz der unterschiedlichen Begriffe und Methoden stehen sie teilweise vor denselben Problemen. DoWiR hat es sich zur Aufgabe gemacht, an dieser Stelle einzugreifen und die Arbeitsgruppen miteinander zu verbinden, um gegenseitig vom Know-How anderer zu profitieren. Dies soll zum einen in der Vermittlung von Methodenwissen liegen, zum anderen auch in der Unterstützung bei konkreten technischen Fragen, z.B. wie kann ich meine Anwendung auf den zentralen Hochleistungsrechner LiDO3 bringen.

Der vorliegende Fragebogen will einerseits die Ist-Situation erfassen, sowohl im Hinblick auf die eingesetzte Software als auch auf die vor Ort bei den Arbeitsgruppen installierte Serverkapazität für wissenschaftliche Anwendungen. Andererseits soll auch der zukünftige Bedarf (im Hinblick auf Kapazität und hardware-technische Möglichkeiten) ermittelt werden, um daraus ein Konzept zu entwickeln, wie zukünftig an der TU Dortmund Ressourcen zum Wissenschaftlichen Rechnen effizienter zur Verfügung gestellt werden können.

Bei Fragen und Anmerkungen zur Umfrage schreiben Sie bitte eine E-Mail an:

service.dowir@lists.tu-dortmund.de

Die Bearbeitung des Fragebogens soll bis zum **1. Juli** erfolgen. Bitte schicken Sie den ausgefüllten Fragebogen per E-Mail an *service.dowir@lists.tu-dortmund.de*.

1. Allgemeines

Name AG-Leiter:

Arbeitsgruppe:

Fakultät/Lehrstuhl:

Größe der Arbeitsgruppe: davon Drittmittel (in %):

Anzahl der Mitarbeiter in der Arbeitsgruppe, die Software zum Wissenschaftlichen Rechnen nutzen:

Spezifikation der Software¹: Numerische Simulation Datenanalyse
 Maschinelles Lernen

1.1 Lehre

Beschreiben Sie bitte Ihr Lehrgebiet:

1.1.1 Welche Bedeutung hat das Wissenschaftliche Rechnen in Ihrer Lehre?

- unverzichtbar sehr wichtig wichtig hilfreich, aber nicht essentiell
 unwichtig

1.1.2 Verwenden Sie entsprechende Programme in der Lehre?

- Ja, bitte geben Sie Details in 2.1 an Nein

1.1.3 Lehren Sie die zugehörigen Grundlagen und Methoden?

- Ja, bitte geben Sie Details in 2.2 an Nein

1.1.4 Vermitteln Sie Methodenwissen zur Softwareerstellung?

- Ja, bitte geben Sie Details in 2.3 an Nein

1 Alle drei Bereiche werden im Weiteren unter dem Begriff „Wissenschaftliches Rechnen“ subsummiert.

1.2 Forschung

Beschreiben Sie bitte Ihr/e Forschungsgebiet/e:

1.2.1 Welche Bedeutung hat das Wissenschaftliche Rechnen in Ihrer Forschung?

- unverzichtbar sehr wichtig wichtig hilfreich, aber nicht essentiell
 unwichtig

1.2.2 Verwenden Sie entsprechende Software in Ihrer Forschung?

- Ja, bitte geben Sie Details in 3.1 an Nein

1.2.3 Erweitern Sie vorhandene Software zum Wissenschaftlichen Rechnen um eigene Bausteine?

- Ja, bitte geben Sie Details in 3.2 an Nein

1.2.4 Entwickeln Sie eigene Software zum Wissenschaftlichen Rechnen?

- Ja, bitte geben Sie Details in 3.3 an Nein

1.3 Hardware-Infrastruktur in der Arbeitsgruppe

1.3.1 Welche Arbeitsplatzrechner verwenden Sie in ihrer Arbeitsgruppe?

Anzahl Windows-Workstations: Anzahl Linux-Workstations:
Anzahl Mac-OS-Workstations: Anzahl Thin-Clients/Virtual Desktops:

1.3.2 Betreiben Sie selbst lokale Server für das Wissenschaftliche Rechnen?

- Ja, bitte geben Sie Details in 4.1 an Nein

1.3.3 Nutzen Sie externe Server für das Wissenschaftliche Rechnen?

- Ja, bitte geben Sie Details in 4.2 an Nein

1.4 Volumen der Nutzung von WiR-Ressourcen

Schätzen Sie bitte das Gesamt-Volumen an Ressourcen, welches Ihre Rechnungen (auf lokalen und externen Servern) benötigen.

1.4.1 typisches, externes Transfervolumen in GB/Monat (von/nach außerhalb des TU-Netzes), z.B.

Download von Mess- und Simulationsergebnissen:

1.4.2 typisches Transfervolumen in GB/Monat innerhalb des TU-Netzes, z.B. Download von Mess- und Simulationsergebnissen:

1.4.3 durchschnittlicher Hauptspeicherbedarf pro Job in GB (falls Sie weder die Techniken MPI und/oder OpenMP benutzen, lassen Sie die entsprechenden Felder frei):

seriell:

MPI:

OpenMP:

1.4.4 durchschnittlicher Plattenspeicherbedarf dauerhaft auf der Hochleistungsmaschine in TB:

1.4.5 durchschnittlicher Plattenspeicherbedarf für dauerhafte Lagerung von Forschungsdaten in TB (außerhalb der Hochleistungsmaschine):

1.4.6 Falls Sie Parallelisierungstechniken benutzen:

1.4.6.1 typische Anzahl an Prozessen in einem Programmlauf:

MPI:

OpenMP:

CUDA:

1.4.6.2 Ist die Verfügbarkeit des Hochleistungsnetzwerks essentiell (niedrige Latenz):

ja nein

1.4.6.3 Ist die Reproduzierbarkeit von Zeitmessungen wichtig:

ja nein

1.4.7 Falls Sie mehr als 256 GB Hauptspeicher pro Rechenknoten benötigen, beschreiben Sie bitte die Anwendung/en, die diese Ressourcen in Anspruch nehmen sollen:

1.4.8 Falls Sie mehr als 64 Prozesse pro Rechenjob benötigen, beschreiben Sie bitte die Anwendung/en, die diese Ressourcen in Anspruch nehmen sollen:

1.4.9 durchschnittlicher Laufzeitbedarf pro Simulation in Stunden und Anzahl benötigter CPU-Cores:

1.4.10 Ist es möglich, Ihre Berechnungen nach einer Unterbrechung fortzusetzen? (Restartmöglichkeit)²

ja nein

1.4.11 Wären kürzere maximale Laufzeiten in den Queues/Partitionen mit bisher 2 bzw. 28 Tagen Maximum für Ihre Berechnungen ein Showstopper?³

ja nein

1.4.12 Falls „ja“ unter 1.4.11: Akzeptabel wären maximale Laufzeiten von

1.4.13 Sind die Simulationscodes für Ihre Berechnungen memory- oder compute bound?

memory bound compute bound trifft beides zu, je nach Code

2 Eine Restartmöglichkeit erlaubt kürzere maximale Laufzeiten. Die derzeit maximal möglichen Laufzeiten von 28 Tagen erschweren "rolling updates" (zeitlich eng geblockte Durchführung von Software-Updates auf allen Servern im laufenden Betrieb ohne Downtime). Bei einem Zeitfenster von 28 Tagen zwischen der Aktualisierung der ersten Rechenknoten bzw. redundant ausgelegten zentralen Server und dem zuletzt leergelaufenen Rechenknoten ist die Vorgabe "alle Rechenknoten eines Typs verhalten sich gleich" ohne Downtime nicht einzuhalten.

3 Kürzere maximale Laufzeiten sorgen für eine höhere Auslastung des Cluster, kürzere Wartezeiten für Gelegenheitsnutzer, haben Vorteile in der Administration ("rolling updates"), aber sorgen typischerweise für Reibungsverluste und höheren logistischen Aufwand beim Fortsetzen von Simulationen.

1.5 Allgemeine Anregungen und Kritik

2. Wissenschaftliches Rechnen in der Lehre

2.1. Verwendung von Software in der Lehre

Welche Software verwenden Sie in der Lehre?

Bieten Sie Lehrveranstaltungen an, bei denen die Studierenden selbst am Rechner sitzen?

Ja Nein

Allgemeine Bemerkungen zur Verwendung von Software in der Lehre:

2.2. Lehre von Methoden zum Wissenschaftlichen Rechnen


Beschreiben Sie bitte, wie Sie Rechenverfahren in der Lehre einsetzen:

2.3. Lehre von Methodenwissen zur Softwareerstellung

Beschreiben Sie bitte, wie Sie Methodenwissen zur Softwareerstellung vermitteln:

2.4 Bedarf an Lehre

Was ist Ihr Bedarf an Lehre in Bezug auf das Wissenschaftliche Rechnen, der zur Zeit nicht abgedeckt ist:

A large, empty rectangular box with a thin black border, intended for the user to write their response to the question above.

3. Wissenschaftliches Rechnen in der Forschung

3.1. Verwendung von Software in der Forschung

Beschreiben Sie bitte, wie Sie Software in der Forschung einsetzen:

3.2. Erweiterung von Software um eigene Programmbausteine

Beschreiben Sie bitte, wie Sie vorhandene Software um eigene Bausteine erweitern:

3.3 Entwicklung eigener Software mit interaktiven Umgebungen

3.3.1 Welche interaktiven Umgebungen verwenden Sie:

- Matlab Maple Mathematica Python-IDE
 R Studio Julia andere, und zwar:

3.3.2 Welche Vor- und Nachteile sehen Sie in der Verwendung dieser Software?

3.4 Entwicklung eigener Software mit höheren Programmiersprachen

3.4.1 Welche Programmiersprache(n) benutzen Sie:

C C++ Fortran Java Python R

andere, und zwar:

3.4.2 Art der Parallelisierung

Welche Art der Parallelisierung nutzen Sie?

CPU: shared memory distributed memory

GPU: shared memory distributed memory

3.4.3 Skalierbarkeit

Maximale Skalierbarkeit der Anzahl Prozesse pro Job (unabhängig von der Art der Parallelisierung (Threads/MPI):

3.4.4 Spezielle Hardware

Planen Sie die Nutzung folgender Hardware oder benutzen Sie diese bereits?

schnelles latenzarmes Hochleistungsnetzwerk (Infiniband)

Einsatz von GPUs

Einsatz von FPGAs

Einsatz von TPUs

Allgemeine Bemerkungen zur Entwicklung eigener Software:

3.5 Lehrangebote für Mitarbeiter

Welche Lehrangebote für Mitarbeiter in Bezug auf das Wissenschaftliche Rechnen benötigen Sie (Programmierkurse etc.):

4. Hardware-Infrastruktur

4.1 Lokale Server

Beschreiben Sie bitte zusammenfassend die Ausstattung Ihrer lokalen Server:

Anzahl der Rechenknoten:

Gesamtzahl der Cores pro Knoten:

Gesamthauptspeicher über alle Knoten in GB:

4.2 Externe Server

Geben Sie bitte an, auf welchen der nachfolgenden externen Rechnersysteme Sie Software zum Wissenschaftlichen Rechnen ausführen.

- lokaler, zentraler Hochleistungsrechner, z.Z. LiDO3
- externer Hoch-/ Höchstleistungsrechner, z.B. HLRS/LRZ/FZJ/Gauß-Allianz/Prace-Calls
- Cloud (Amazon Web Services, Microsoft Azure, Google Cloud Platform, IBM Cloud, Oracle Cloud etc.)

5. Betriebsmodell / Zugriff auf System-Ressourcen

Hinweis: Der Zugriff auf System-Ressourcen (Betriebssystem, Scheduler, besondere Bibliotheken etc.) auf einem gemeinsamen Rechencluster kann nur über Virtualisierung erfolgen.

5.1 Benötigen Sie für Ihre Lehr- und Forschungstätigkeit Containerisierung?

Ja Nein Evtl. geplant

Wenn ja oder eventuell geplant, dann bitte die nachstehenden Fragen beantworten. Falls Ihr Anliegen nicht durch Container abgedeckt werden kann, geben Sie Details in 5.2 an.

5.1.1 Welche Container würden Sie für Ihre Lehr- und Forschungstätigkeit benötigen?

5.1.2 Wie viele Container benötigen Sie für Ihre Lehr- und Forschungstätigkeit?

Anzahl:

5.1.3 Durchschnittliche Anforderungen eines Containers?

A large, empty rectangular box with a thin black border, intended for the user to provide an answer to the question above.

5.2 Container können meine Anforderungen nicht erfüllen, weil:

A large, empty rectangular box with a thin black border, intended for the user to provide an answer to the question above.

6. Sonstiges

Welcher Bereich ist im Fragebogen nicht abgedeckt?

Im Namen des DoWiR-Vorstandes danken wir für Ihre Mithilfe!