

Dortmunder Zentrum für Wissenschaftliches Rechnen (DoWiR)

3. Mitgliederversammlung

Mittwoch, 28. November 2012, 17.00 – 18.30 Uhr, Mathematik-Gebäude, M 614/616

Teilnehmerinnen und Teilnehmer:

- Frithjof Anders, Physik
 - Franz-Joseph Barthold, Bauwesen
 - Christian Becker, Mathematik / ITMC
 - Bernd Bischl, Statistik
 - Peter Buchholz, Informatik
 - Peter Ehrhard, BCI
 - Roland Frach, Chemie
 - Jörg Gehrke, ITMC
 - Jochen Heil, Chemie
 - Leonhard Henkes, Chemie
 - Franziska Hoffgaard, Chemie
 - Stefan M. Kast, Chemie
 - Jan Kierfeld, Physik
 - Martin Kötterheinrich, ITMC
 - Andreas Menzel, Maschinenbau
 - Christian Meyer, Mathematik
 - Katharina Morik, Informatik
 - Florian Mrugalla, Chemie
 - Sven Odendahl, Maschinenbau
 - Carsten Raas, Physik
 - Andreas Rademacher, Mathematik
 - Jakob Rehof, Informatik
 - Peter Resch, ET/IT (in Vertretung für Uwe Schwiesselshohn, ET/IT)
 - Wolfgang Rhode, Physik
 - Stefan Turek, Mathematik
 - Jan Ziemann, ITMC / Physik
- Gäste:
- --

Protokoll:

- Magdalena Thöne, Mathematik

Tagesordnung:

1. Begrüßung und Festlegung der endgültigen Tagesordnung
2. Genehmigung des Protokolls der 2. Mitgliederversammlung vom 10.10.2012
3. Bericht des Vorstandes
4. Rechnerkonzepte - Aktueller Zwischenbericht zur Umfrage / Bedarfsanalyse
5. Aktivitäten von DoWiR: Aktueller Stand (Internet, Veranstaltungen, Mitgliederbefragung, ...)
6. Sonstiges (u.a. Termin Mitgliederversammlung - Vorschlag erste Februarwoche)

1. Begrüßung und Festlegung der endgültigen Tagesordnung

Herr Turek begrüßt die Anwesenden (26 Mitglieder, davon 14 ordentliche=stimm-berechtigte Mitglieder; Protokollantin) zur dritten Mitgliederversammlung von DoWiR (Dortmunder Zentrum für Wissenschaftliches Rechnen). Mehrere Mitglieder haben sich für die Versammlung entschuldigt.

Die Beschlussfähigkeit wird festgestellt, da mehr als ein Viertel der ordentlichen Mitglieder (i.W. Hochschullehrer/innen und Projektleitungen) anwesend ist. (Erforderlich sind hierzu aktuell 10 ordentliche Mitglieder.)

Aktuell hat DoWiR 65 Mitglieder, davon 37 ordentliche Mitglieder.

Herr Turek stellt einen Vorschlag für die Tagesordnung vor. Es bestehen keine Änderungswünsche bezüglich der Tagesordnung. Die Sitzungsleitung wird von Herrn Turek übernommen.

2. Genehmigung des Protokolls der 2. Mitgliederversammlung vom 10.10.2012

Der Protokollentwurf war mit der Einladung zur Mitgliederversammlung verschickt worden. Es liegen keine Änderungsvorschläge vor. Das Protokoll wird per Akklamation **genehmigt**.

3. Bericht des Vorstandes

Die Satzung und die Satzungsänderung aus der letzten Mitgliederversammlung liegen noch in der Verwaltung zur Prüfung. Bisher gab es noch keine Rückmeldung, nur einige Nachfragen.

Der Vorstand hat einige neue Mitglieder aufgenommen.

Herr Turek verweist kurz darauf, dass das aktuelle Procedere zur Aufnahme von neuen Mitgliedern derart geregelt ist, dass Aufnahmewünsche formlos per Mail an eines der Vorstandsmitglieder gerichtet werden können. Die Anträge werden dann vom Vorstand gemeinsam behandelt.

Die **Aktivitäten** von DoWiR sind bislang im Wesentlichen **intern**.

Es gibt **Kontakte mit** Kolleg/inn/en der **Universität Duisburg-Essen**; ein Treffen (zum gegenseitigen Kennenlernen und "Beschnuppern") ist geplant, um sich über Konzepte und Ideen auszutauschen. Als langfristige Idee steht immer noch eine Beteiligung an der Gauß-Allianz im Raum. Es soll noch abgefragt werden, wer aus den DoWiR-Reihen Interesse hat, an einem gemeinsamen Treffen mit den "Nachbarinnen und Nachbarn" teilzunehmen; ein solches Treffen könnte etwa im Januar stattfinden.

Ganz frisch eingetroffen ist eine **Antwort von IBM** bzgl. der Möglichkeiten im Bereich Cloud Computing etc. Die Antwort kam allerdings so kurzfristig an, dass sie noch nicht ausgewertet werden konnte und daher hier noch nicht vorgestellt werden kann. Die Anfrage war wohl bei IBM zeitweise "untergegangen".

4. Rechnerkonzepte - Aktueller Zwischenbericht zur Umfrage / Bedarfsanalyse

Ein wesentlicher Punkt der heutigen Mitgliederversammlung ist die Vorstellung und Diskussion des **Zwischenberichts zur DoWiR-Umfrage**, mit der versucht werden soll, einen Überblick über Bestand und Bedarf an Wissenschaftlichem Rechnen innerhalb der TU Dortmund zu bekommen.

Im Vorfeld war ein erster Zwischenstand zur Auswertung der Umfrage verschickt worden; bis zur Sitzung liegen weitere Rückmeldungen vor, die im Folgenden vorgestellt werden. Christian Becker stellt die Resultate vor; diese Präsentation ist im *Anhang* beigelegt.

Auch diese Präsentation stellt einen Zwischenstand dar, da noch weitere Rückmeldungen erwartet werden. Sie stellt allerdings eine gute erste Übersicht dar und ist Grundlage für die anschließenden Diskussionen. Bis Ende Dezember sollen noch Rückmeldungen gesammelt werden.

Nach dem heutigen Kurzbericht soll bis zur nächsten Sitzung eine detailliertere Auswertung mit einem Bericht erstellt werden; in einem dritten Schritt sollte dann ein DoWiR-Team auf dieser Grundlage an der Konzeptentwicklung arbeiten.

Kurzzusammenfassung der Präsentation und **Diskussion** (mehr Details vgl. *Anhang*):

Bislang haben **23 Gruppen** Fragebögen ausgefüllt und zurückgeschickt; aus fast allen MINT-Fakultäten liegen bisher Antworten vor. Die Angaben sind hier zunächst ungewichtet, unabhängig von der Größe der jeweiligen Gruppe, dargestellt, wobei die Gruppen zwischen 3 und 30 Mitarbeiter/inne/n umfassen.

Auch in der **Lehre** spielt Wissenschaftliches Rechnen eine wichtige Rolle. Die **verteilte Ausstattung** ist von der Größenordnung mit LiDo vergleichbar. Diskutierte Aspekte sind die Nutzung externer Rechner (incl. LiDo), Hauptspeicher, Parallelisierung, Hochleistungsnetzwerke, Zeitmessung und GPU-Nutzung.

In der Diskussion wird nachgefragt, ob etwa die Nutzung von Hadoop mit LiDo oder anderen Großrechnern möglich ist. Hierzu wird ausgeführt, dass ein "Verheiraten" der Konzepte möglich ist, dass aber die jeweiligen Vorteile dabei evtl. verloren gehen können. Daher wird eher separiert vorgegangen, Testversionen sind aber machbar.

Herr Kast ergänzt noch zu den **Kontakten mit Duisburg-Essen**, dass über einen gemeinsamen Antrag bei Mercator (Strukturförderung MERCUR) nachgedacht wird. In

diesem Fall ist der Zeithorizont bis Mitte Mai, um einen solchen Antrag zu formulieren (mit entsprechenden Gremienwegen). Hierzu ist es vielleicht sinnvoll, den Fragebogen auch an Duisburg-Essen zu geben, so dass dort eine ähnliche Analyse vorgenommen werden kann.

Zur **Software-Entwicklung** wird die Bedeutung in der Lehre und der Bedarf entsprechender Lehrangebote thematisiert, um Konzepte erlernen zu können (Compiler, Debugger, ...).

Zur Frage der **Betriebsmodelle** wird diskutiert, dass man mehr selbst bestimmen können muss, wenn Algorithmen selbst der Gegenstand der Forschung sind. Auch bei Software als Forschungsgegenstand sind eigene Cluster o.ä. sehr wichtig.

Es wird noch einmal betont, dass **LiDo** mehr benutzt würde, wenn mehr Personen damit umgehen können; hier ist die Hemmschwelle noch relativ groß und es besteht Bedarf an entsprechenden Grundschulungen. Aktuell wird vielfach lieber auf eigenen Clustern gerechnet. Es sollte noch mehr "Übertragung" stattfinden, indem man lokal testet und dann größere Programme auf LiDo schickt.

Insgesamt zeigt sich, dass mit mehr entsprechender **Ausbildung** auch mehr Nachfrage auf allen Bereichen entsteht; die Ausbildung sollte auch für Nicht-Fachleute verbessert bzw. überhaupt erst ermöglicht werden.

Zur Abschätzung von **Energiekosten** sind sowohl Server als auch Workstations zu betrachten; hier ist auch die Anzahl der Arbeitsplatzrechner eine relevante Größe.

Die Ergebnisse müssen noch weiter aufbereitet werden und ggf. mit entsprechenden Bewertungen ergänzt werden.

Die Datenbasis kann als Basis für ein erstes **Betriebskonzept** (max. zwei Seiten, Grobskizze, "Managementvorlage") verwendet werden. Hierzu sollte sich eine kleine Arbeitsgruppe zusammenfinden. Wenige Kenngrößen sollten hierfür zusammengestellt werden. Die Auswertung ist hierzu als Hintergrundinformation sehr wichtig. Hybrid- und Cloudansätze können in verschiedenen Szenarien eine Rolle spielen (Beteiligung in der Gauß-Allianz als Idee, Cloud als anderes Konzept, ...).

Ergänzend wird noch darauf hingewiesen, dass LiDo noch bis Ende 2013 läuft.

Die Nutzung **externer Rechenkapazitäten** (z.B. im Forschungszentrum Jülich) könnte noch weiter gesteigert werden. Hier ist es aber wichtig, die Skalierbarkeit nachzuweisen. Templates, um Rechenkapazitäten in Jülich zu beantragen, sind online verfügbar; aktuell sind die Chancen etwa 1 zu 7, dort mit einem Antrag Erfolg zu haben.

Gerade **Nachwuchswissenschaftler/innen** müssen **stufenweise** an das wissenschaftliche Rechnen herangeführt werden. Dies fängt bereits mit FORTRAN an; es sollte mehr Aufmerksamkeit in der Lehre auf diese Aspekte gelegt werden, da ansonsten ein beträchtlicher Anteil teuer bezahlter Arbeitszeiten (z.B. im TV-L) noch für Grundlagenausbildung aufgewendet wird.

Die Bündelung verschiedener Kursangebote kann weitere Synergie-Effekte mit sich bringen.

Bei einigen Gruppen geschieht die Nichtnutzung von großen Rechnern allerdings bewusst, das sie z.T. eine andere Architektur benötigen.

Es sind daher insgesamt verschiedene Arbeitsebenen (Ranges) für Wissenschaftliches Rechnen erforderlich:

Arbeitsplatz – kleiner Cluster – LiDo o.ä. – Hochleistungsrechnen

Es wird noch einmal betont, dass eine **Arbeitsgruppe** gebildet werden sollte, um ein Fazit aus den vorliegenden Ergebnissen und Diskussionen zu ziehen. Wer dabei mitwirken möchte, sollte sich bei Herrn Turek melden.

Herr Kast betont noch einmal die Bedeutung der Betriebskosten, die auch innerhalb der UAMR-Diskussion wichtig sind. Hier deutet sich eher eine Zusammenarbeit von Dortmund und Duisburg-Essen an, da Bochum sehr dezentral organisiert ist. Eine Kopplung der Betriebskosten mit den Gesamtkosten ist daher ebenfalls wichtig. (Innerhalb der Gauß-Allianz ist eine solche Kopplung möglich, etwa wenn für die Energiekostenfrage innovative Ansätze verfolgt werden.)

Frau Morik erwähnt, dass in diesem Zusammenhang eine interessante Neuberufung in der Informatik erfolgen könnte.

Alle Beteiligten werden aufgerufen, in ihren Arbeitsgruppen und im weiteren Umfeld noch an die Befragung zu erinnern, um den Rücklauf weiter zu erhöhen und die Datenbasis weiter zu untermauern.

5. Aktivitäten von DoWiR: Aktueller Stand (Internet, Veranstaltungen, Mitgliederbefragung, ...)

Verschiedene Aspekte werden angesprochen und diskutiert.

Die **Internetpräsenz** von DoWiR (www.dowir.de) ist inzwischen freigeschaltet, benötigt aber weiter Inhalte und Beiträge aus den Reihen der Mitglieder sowie Kurzlebensläufe o.ä. Auch weitere Kurse und Lehrveranstaltungen sollten hier noch gemeldet werden (z.B. MPI, C++, Fortran, ...).

6. Sonstiges (u.a. Termin Mitgliederversammlung - Vorschlag erste Februarwoche)

Bis Weihnachten sollen noch Rückmeldungen zur Umfrage erfasst werden. Wer in der Arbeitsgruppe zur Auswertung der Umfrage mitwirken möchte, sollte sich ebenfalls bis zum Jahresende melden. Die Auswertung sollte mit einem ersten Fazit versehen werden und dann Anfang des Jahres weiter diskutiert werden.

Als Termin für die nächste **Mitgliederversammlung** wird ein Termin in der ersten Februarwoche (erste Woche der vorlesungsfreien Zeit) vorgeschlagen.

Es wird vereinbart, die nächste Vorstandssitzung spätestens am Dienstag, 05. Februar 2013, durchzuführen; die Mitgliederversammlung soll am **Mittwoch, 06. Februar 2013, um 15 Uhr (Mathematikgebäude, Raum E19)** stattfinden.

Die nachfolgende Mitgliederversammlung sollte dann Anfang März terminiert werden.

Weitere Wortmeldungen liegen nicht vor.

Herr Turek dankt allen Beteiligten für ihre engagierte Mitwirkung, schließt die Sitzung und dankt allen Beteiligten für ihr Kommen.

Ende der Sitzung: 28.11.2012, 18.30 Uhr

Protokoll: Magdalena Thöne

Sitzungsleitung:

Stefan Turek, Mathematik

Stefan M. Kast, Chemie

Zwischenergebnis DoWiR-Umfrage WiR

DoWiR-Mitgliederversammlung

28.11.2012

Umfrage WiR

Anzahl Gruppen: 23
Anzahl WiR-Benutzer: 267

Fakultäten:

Architektur und Bauingenieurwesen:	4%
BCI:	4%
Chemie:	13%
Informatik:	17%
MB:	13%
Mathematik:	13%
Physik:	17%
Statistik:	17%

Spezifikation Software: Numerische Simulation:	73%
Datenanalyse :	47%

Umfrage WiR: Lehre

Welche Bedeutung hat das WiR in Ihrer Lehre?

Teilnahme insgesamt: 23 100%

unverzichtbar:	43%
sehr wichtig:	39%
wichtig:	13%
weniger wichtig:	4%
unwichtig:	0%

Verwenden Sie entsprechende Programme in der Lehre?

ja: 96% nein: 4%

Lehren Sie die zugehörigen Grundlagen und Methoden?

ja: 91% nein: 9%

Vermitteln Sie Methodenwissen zur Softwareerstellung?

ja: 70% nein: 30%

Umfrage WiR: Forschung

Welche Bedeutung hat das WiR in Ihrer Forschung?

Teilnahme insgesamt: 23 100%

unverzichtbar:	74%
sehr wichtig:	22%
wichtig:	4%
weniger wichtig:	0%
unwichtig:	0%

Verwenden Sie entsprechende Software in Ihrer Forschung?

ja: 100% nein: 0%

Erweitern Sie vorhandene Software zum WiR um eigene Bausteine?

ja: 87% nein: 13%

Entwickeln Sie eigene Software zum WiR?

ja: 87% nein: 13%

Lokale und externe Server

Ausstattung lokaler Server:

Anzahl lokaler Server: 305
Anzahl lokaler CPUs: 674
Anzahl lokaler Cores: 2433
Größe Hauptspeicher: 5397

Zum Vergleich LiDOng:

Anzahl Server: 440
Anzahl CPUs: 912
Anzahl Cores: 3648
Größe Hauptspeicher: 8128

Nutzung externer Rechner

LiDOng:	ja: 48%	nein: 52%
HLRS/LRZ/FZJ:	ja: 9%	nein: 91%
Grid:	ja: 4%	nein: 96%

Transfervolumen nach außen: 32288 GB/m

Hauptspeicheranforderung

Hauptspeicher seriell (in GB):

keine Angabe:	26%
1-31:	34%
32-64:	35%
>64:	5%

Hauptspeicher MPI (in GB):

keine Angabe:	65%
1-256:	26%
>256:	9%

Hauptspeicher OpenMP (in GB):

keine Angabe:	56%
1-31:	13%
32-64:	13%
>64:	18%

Permanenter Plattenspeicher (in TB): 945.7

Parallelisierungstechniken:

rein seriell	ja: 83%	nein: 17%
Shared Memory	ja: 43%	nein: 57%
Distributed Memory	ja: 39%	nein: 61%

Anzahl Prozesse MPI:

keine Angabe :	56%
<64 :	18%
64-256 :	21%
>256 :	5%

Anzahl Prozesse OpenMP:

keine Angabe :	52%
1-8 :	26%
9-32:	17%
>32:	5%

Spezielle Hardware

Verfügbarkeit eines Hochleistungsnetzwerk wichtig

ja: 22% nein: 78%

Reproduzierbarkeit von Zeitmessungen wichtig

ja: 30% nein: 70%

Einsatz von GPUs (in Nutzung/geplant)

ja: 35% nein: 65%

- homogene Entwicklungsplattform und homogenes Runtime Environment auf Workstations und lokalen Clustern
- großer Bedarf an schnellem lokalem HD-Scratch (ab 1 TB) auf allen Compute-Nodes
- hinreichend schnelle Vernetzung der Compute-Nodes untereinander
- viele schnelle Floating-Point-Einheiten auf jedem Compute-Node, min. eine FP-Einheit/Core
- sehr hoher Bedarf an RAM pro Compute-Node (ab 32 GB, nach oben offen)

- flexibles Handling von Queues, Scheduler-Parametern und Job-Prioritäten ermöglicht die einfache Anpassung an Deadlines (Abgabefristen, Konferenzen)
- Statistiker/Datenanalysten haben selten parallele Jobs im Sinne von MPI und OpenMP. Vielmehr haben diese i.d.R. viele tausend bis hunderttausend weitestgehend unabhängige Jobs, die ausgeführt werden. Viele Aufgaben lassen sich gut als Map-Reduce-Job modellieren. Diese Art Aufgaben mit einem regulären Job-Scheduler abzuarbeiten, ist bisweilen sehr umständlich.
- wünschenswert wäre der Support von Java inkl. RMI und Hadoop

Nutzung von interaktiven Entwicklungsumgebungen:

Matlab	ja: 70%	nein: 30%
Maple	ja: 43%	nein: 57%
Mathematica	ja: 39%	nein: 61%

Nutzung von höheren Programmiersprachen:

C	ja: 57%	nein: 43%
C++	ja: 74%	nein: 26%
Fortran	ja: 48%	nein: 52%
Java	ja: 30%	nein: 70%

- Grundlagenkurs C++, Fortran
- Crashkurs Benutzung von Cluster-Systemen mit Batch-Betrieb
- Crashkurs MPI, pthread, OpenMP
- Crashkurs GPU/CUDA
- Linuxcluster-Administrationskurs

Wenig Bedarf an direktem Systemzugriff/Virtualisierungslösungen

MapReduce automatisch

Teilnahme insgesamt: 18 78%

unverzichtbar:	11%
sehr wichtig:	17%
wichtig:	17%
weniger wichtig:	28%
unwichtig:	28%

MapReduce parametrierbar

Teilnahme insgesamt: 18 78%

unverzichtbar:	0%
sehr wichtig:	11%
wichtig:	17%
weniger wichtig:	39%
unwichtig:	33%

dienstorientierte Verfahren

Teilnahme insgesamt: 18 78%

unverzichtbar :	0%
sehr wichtig:	17%
wichtig:	11%
weniger wichtig:	28%
unwichtig:	44%

Anwenden von dienstorientierten Verfahren

Teilnahme insgesamt: 16 70%

unverzichtbar:	0%
sehr wichtig:	6%
wichtig:	19%
weniger wichtig:	38%
unwichtig:	38%