



Dr. Andreas Knüpfer ZIH / TU Dresden

Data-intensive High Performance Computing at TU Dresden: Combining Computer Science Research with Support for Computational Sciences

11th Dresden Probabilistic Workshop 2019-10-10



Overview

- Overview about ZIH
- HPC Hardware and Infrastucture
- Selected Computer Science Challenges
- Selected Computational Science Challenges









Overview about ZIH and HPC



ZIH Structure

- Director Prof. Dr. Wolfgang E. Nagel, deputy director Dr. Andreas Knüpfer
- 7 departments, >160 staff including apprentices, at 6 locations on the campus







ZIH Topics

IT Services for TU Dresden	 Campus network, internet uplink, phone network E-Mail, groupware, data exchange, backup etc. Software procurement 	
Computational Science Services	 Virtual machines, hosting Supercomputing Big Data methods 	
Research and Development	 Parallel programming and algorithms Performance optimization and scaling, energy efficiency, Data Analytics applications, machine learning, 	





ZIH Topics

IT Services for TU Dresden	 Campus network, internet uplink, phone network E-Mail, groupware, data exchange, backup etc. Software procurement 	
Computational Science Services	 Virtual machines, hosting Supercomputing Big Data methods 	
Research and Development	 Parallel programming and algorithms Performance optimization and scaling, energy efficiency, Data Analytics applications, machine learning, 	







IT Services and the Service Catalog

- <u>https://tu-dresden.de/zih/dienste/service-katalog</u>
- Topic lists and service descriptions

TECHNISCHE UNIVERSITÄT DRESDEN	TU Dresden	7	Sprache 🎸	Suche 🔎	Intern 🥑
ZENTRUM FÜF DIENSTE UND RECHNEN (ZIH	R INFORMATIONS- HOCHLEISTUNGS I)	-			≡
Service k	Katalog				
ANGEBOT	ENE DIENSTE	FÜR E	INZELNE NU	TZER	
Dienst			Studierende	Mitarbeiter	Gast
Zugangsvora	ussetzung				
>ZIH Ben	utzer-Login A	ndreas Kn	+ üpter – Dresden Pro	+ babilistic Worksho	+

		Studierende	Mitarbeiter	Gast
Zugængsvoraussetzung				
321H Benutzer-Lo	gin			
Arbeitsumgebung				
>PC-Arbeitsplatz		(PC-Pools)	(PC-Service)	[PC-Service]
Software		[MS Office]		Igegen Gebühr!
NZutang Campus	netz via WLARE& VPN			
Talatas				
Dateromaicher				
3				-
s Backup & Archiv		Datemspeicher		
		Projekt]		
» E-Mail				•
3 Zeitdionate				•
3 Drucken / Scann	es / Piotien			
3 Handbücher				
Zusammenarbeiten & Fors	ichen			
5 Daterwastausch		(Cloudstore)		
Weter-Claimer	vilanenten	(nur als		
January millions		Teilnehmer)		
yE-Mail-Listen				•
> Software-Entwick	Kungsumgebung	-		•
3 Broupware (Kale	nder/Aulgabern]			•
Simulation / Hochleistungs	rechtern			
	20 Berry 10	(nur über	(nur über	[nur über
100,000		Projekte)	Projekta)	Projekte)
GEBDTENE DIENSTE I	FÜR EINRICHTUNGEN DER TU DRESDEN			
GEBOTENE DIENSTE I Dierat	FÜR EINRICHTUNGEN DER TU DRESDEN			
GEBOTENE DIENSTE I Dierat Arbeitsumgebung	FÜR EINRICHTUNGEN DER TU DRESDEN			
IGEBDTENE DIENSTE I Dianat Acheituumgebung SZentrala Verw	PÜR EINRICHTUNGEN DER TU DRESDEN wiltung von PC-Actaringtotown / PC-Pasta			
IGEBDTENE DIENSTE I Disrat Arbeitsumgebung >Zentrala Vers >Verzeichnickli	FOR EININGHTUNGEN DER TU DRESDEN wittung von PC-Arbeitigstationn / PC-Procise entel (Active Directory/DMP)			
INGEBOTENE DIENSTE I Diarat Arbeitsumgebung > Zentrale Vors > Verzeichnické > Authentfüher	FOR EXHIBITION DER TU DRESDEN anberg ein PC-Artschiptibum / PC-Parls and Antra Constry/LDAM ungsderst Edenty Provided			
GGEDTENE DIENSTE I Dierat Sebelsungsbung > Zentrals Vern > Verzeichnis Ge > Authentifier > Gereinamu	FOR EINRICHTUNGEN DER TU DRESDEN anhung von PC-Arissingsbisson / PC-Pauls and Retries Drestrag/Dably ungedienes Telensty Proceder Naturg on Schware Europarented			
Activities of the second secon	POR EININGHTUNDERN DER TU DRESDEN sohnerg von FC-Arkengelässen / PC-Puela and Marka Onempry/LBM9 urguskenet Marka / Prevaled Markeng inn-Schlmere Kuensen-sold E Ernsthangen von antmäs Datensetz			
Activities of the second secon	TOR ENRICHTUNEEN DER TU DRESDEN antrag von FC-Astenigizier (FC-Paris erst Erkeit Dimenschäßt) orgadient Barbeit Produkt orgadient Barbeit Produkt forschäftiger ein streich Dimensi d dresserschäftig dresserschäftig			
AGEBOTENE DIENSTE I Dierat Schnitzungsburg Schnitzungsburg Schnitzung Schnitzung Schnitzung Schnitzung Schnitzung Schitzung Sc	CIR EINECHTUNEEN DER TU DRESDEN auflang an PC-Assinghtein (PC-Parls mer Hons Dimetry/DM) ungelande Barris (Parland Franzigues Catalante Barris			
Actional and Comparison of Com	THE ENRICHTUNEEN DER TU DRESDEN anbeg an 19 Anterphone (19 Den) and hars pansyllable seutragenen Denis (19 Anterphone) Telephone Statisticken Lassenanisel Er erstrangen aus versten Bassena die auszuhlung die auszuhlung seutragen berging und Gengenann			
AGEBOTENE DIENSTE I Dianat Schattsungsbung Schattsungsbung Schattenfricklic Schattenfricklic Schattenfrick Schatte	TOR ENRICHTUNEEN DER TU DRESDEN wihrig von PC Adventigitism (PC Polis) wird bein Densing/SDB) ungekenst besinsensing/SDB) ungekenst besinsensing im Erstriktungen von Jeherse Besamstend im Erstriktungen von Jehersen Beging von Gengensten wird			
KGEBOTENNE DIENSTE I Dianat // Statisampilusy //	Cirk Ennici-Hundelin CER TU DRESORN anharg on PC-Mainglition (PC-Puin anharg on PC-Mainglition (PC-Puin anharg on PC-Mainglition (PC-Puin anharg on PC-Mainglition PC-Mainglition (PC-Puin Anhard (PC-Puin) Anhard			
2 Sender Since DUPSETE ID Dente 4 Enterruption 3 Sender Since S	The ENRICHTUNEEN DER TU DRESDEN The ENRICHTUNEEN DER TU DRESDEN andhag am 70 Andrighten / 70 Parla andhas par 90 Andrighten / 70 Parla andhas par 90 Andrighten / 70 Parla Andreg and Schwarter / 70 Parla Andreg andreg andreg andreg andreg andreg andreg andreg			
Parameter Sector	TOR EINRICHTUNDEDN DER TU DIESDEN TOR EINRICHTUNDEDN DER TU DIESDEN anbeig von PC Arbeigkeisen /PC Prote ant Unsch Diesensyn (2005) ungekonnte Unschweisen / 2005) andersongen Schweise Kassenstein anternahmen Keitersonahmen senternahmen senternahmen senternahmen anternahmen senternahm			
Reserve Sectors (Constraint) Sectors	Cirk ENRICHTUNEEN DER TU DRESDEN anting von FC-Aktingitisten (FC-Parls anting von FC-Aktingitisten (FC-Parls angeland Enterpressen angeland Enterpressen angeland Enterpressen anderen einer Anterpressen Anterpressen an			
Electronic Converter La Co	Concentration of the Concentra			
Electropic Contest of	The ENRICHTUNEEN DER TU DRESDEN The ENRICHTUNEEN DER TU DRESDEN erstlichter zum Für Anteruptionen / Für Paule erstlichter Ummer Fahrenden Natureg von Schwaren Lanzenstreich Er erschlagenes Anterny Paule Restrengenes Schware Lanzenstreich erstlichter Schware Lanzenstreich erstlichter Schware Lanzenstreich Restrengenes Schware Lanzenstreich Res	Liña		
december 2004 Sector 100	TOR ENRICHTUNEDN DER TU DRESDEN anlang war FC Advelaption (FC P Ani anlang war FC Advelaption (FC P Ani and Advelaption (FC P Ani and Advelaption (FC P Ani and Advelaption (FC P Ani Advelaption (FC P	Links	1.5 bibernateraate	
delastrituma dana dana dana dana dana dana dana da	TOR EINRICHTUNEDN DER TU DIESDEN TOR EINRICHTUNEDN DER TU DIESDEN weberg von PC Advangeteinen /PC Parte weberg von PC Advangeteinen /PC Parte webergen Schweisenschaft webergen Schweisenschaft webergen Schweisenschaft webergen Schweisenschaft webergen Schweisenschaft of Schweisenschaft Schweisenschaft um Schweisenschaft BEREIDER DER TU DIESDEN BEREIDER DER UM Besternt um Schweisenschaft um Schwe	Lana survey and survey	1.5.5Memoretergat,tw	
Idea STAN DURNET (For EININGCHUNGEN DER 11 DEESDEN andrag sein FC-Makeyalanis (FC-Paus eral Natura Smartery (LBN) andrag sein FC-Makeyalanis Ratery aus Schwart Lansenstein Ersteinung aus anterkte Talenenst andragen and Response andragen Aller State State (LENEN) Bereitung auf das Angelen andragen andragen BEREINEL DER TU DEESDEN ERST Ersteinung aus das Angelen andragen Angelen aus andragen andragen Bereitung auf das Angelen andragen Angelen andragen aus andragen andragen Bereitung aus das Angelen andragen Bereitung aus das Angelen andragen Bereitung aus das Angelen andragen Bereitung aus das Angelen andragen Angelen andragen aus andragen angelen andragen Angelen andragen aus andragen angelen andragen Angelen andragen aus andragen angelen angelen angelen Angelen angelen angelen angelen angelen Angelen angelen angelen Angelen angelen angelen Angelen angelen angelen Angelen angelen angelen Angelen angelen Angelen angelen Ange	Laks Scottgare Maintensor Second	2.5 bitamatematematematematematematematematema	

Topics overview

IT Services for TU Dresden	 Campus network, internet uplink, phone network E-Mail, groupware, data exchange, backup etc. Software procurement 	
Computational Science Services	 Virtual machines, hosting Supercomputing Big Data methods 	
Research and Development	 Parallel programming and algorithms Performance optimization and scaling, energy efficiency, Data Analytics applications, machine learning, 	





HPC Resources at ZIH

HRSK II 2015

- ~ 44,000 cores Intel (mostly Haswell)
- 256 GPUs Nvidia Tesla K80 +
- 44 GPUs: Nvidia Tesla K20
- 136 TB RAM, >5 PB scratch file system

HPC-DA extension 2018

- 22 Machine Learning nodes IBM AC922
- 2 PB NVME storage (90 nodes, NVMEoF) with 2 TB/s bandwidth in total
- 10 PB Object Storage

Follow-up extension in 2019 (approx. 4 M€)







HPC-DA Extension 2018

HPC-DA extension towards extremely fast I/O

- Redesigned one compute island of HRSK II
- Strong focus on highest bandwith and low latency
- 612 existing CPU compute nodes
- 32 Machine Learning Knoten IBM AC922
 - 2x Power-9 CPUs, 6x NVIDIA V100 GPUs, NVLink
- 90 NVME storage nodes
 - Each node with 8 3,2 TB PCIe x4 NVME cards
 - Dual-link EDR IB, NVME over fabric
- 10 TB Object Storage with 50 GB/s





HPC-DA ML Nodes

Hardware

- 32 IBM AC922 nodes
- 2x POWER9 CPU, 22 core, 4-way HT (176 threads per node in total)
- 2.80 GHz, 3.10 GHz boost
- 256 GB RAM DDR4 2666MHz
- 6x NVIDIA VOLTA V100 with 32GB HBM2
- NVLINK with 150 GB/s between GPUs and between host and GPUs
- CPUs and GPUs direct water cooled
- 0.4 TB/s aggregated bandwidth to NVME nodes



Image: https://www.ibm.com/it-infrastructure/power/accelerated-computing





HPC-DA NVME Storage Nodes

Hardware

- 90 NVME storage nodes
- 2 sockets Intel Xeon E5-2620 v4 (16 cores each, 2.10GHz)
- 64 GB RAM
- 8x Intel NVMe Datacenter SSD P4610,
 3.2 TB, PCIe 3.1 x4 3DNAND ME 2.5" U.2,
 3,2 GB/s (8x 3.2 ==25.6 GB/s)
- 2 Infiniband EDR links, Mellanox MT27800, ConnectX-5, PCIe x16, 100 Gbit/s each





in total 2 TB/s peak bandwidth





HPC-DA NVME Usage Models

Allocation strategy

- NVME shares allocated as long-term "NVME leases" (weeks to months)
- Granularity of 1/8th node (1 NVME card) or full NVME nodes



Separate BeeGFSes in own NVME lease

- Instantiate separate BeeGFS
- Granularity of ½ or 1 NVME nodes, including the EDR links and the CPU cores
- Separate MDSs an OSTs for this FS
- No meta-data interference with everyone else
- Full nominal bandwidth per NVME node
- Have it mounted automatically to compute jobs of your HPC project





Software for HPC-DA: Modules vs. Containers

- A "basic" software installation
- Classical HPC software packages and libraries
- Big Data frameworks
- AI frameworks

Software modules

- Long list of software packages, multiple versions each, dependencies
- Open Source SW, scientific community software packages, commercial SW
- Application software, libraries, software tools
- Compiled vs. interpreted (Python etc.)
- Central and individually per project/user

 \sim \wedge \times () taurus — Konsole ATK/2.28.1-foss-2018a basemap/1.0.7-intel-2018a-Python-3.6.4 cairo/1.14.12-GCCcore-6.4.0 GL2PS/1.4.0-foss-2018a GL2PS/1.4.0-foss-2018b GL2PS/1.4.0-intel-2018a libGLU/9.0.0-fosscuda-2018 libGLU/9.0.0-GCCcore-8.2.0 Pillow-SIMD/5.0.0-foss-2018a-Python-3.6.4 Pillow-SIMD/5.0.0-intel-2018a-Python-3.6.4 Pillow/5.0.0-intel-2018a-Python-3.6.4 (D) (D) libGLU/9.0.0-intel-2018a (D) lib61U/9.0.0-intel-2018a Little(ENS)2.9-6CCcore-6.4.0 matplotlib/2.1.2-fors-2018a-Python-2.7.14 matplotlib/2.1.2-fors-2018a-Python-3.6.4 matplotlib/2.1.2-intel-2018a-Python-3.6.4 matplotlib/2.1.2-intel-2018a-Python-3.6.4 cairo/1.14.12-GCCcore-7.3.0 (D) GLib/2.54.3-GCCcore-6.4.0 pixman/0.34.0-GCCcore-6.4.0 pixman/0.34.0-GCCcore-7.3.0 PyQt5/5.10.1-intel-2018a-Python-3.6.4 GLib/2.54.3-GCCcore-6.4.0 GLib/2.54.3-GCCcore-7.3.0 GLib/2.60.1-GCCcore-8.2.0 gnuplot/5.2.2-foss-2018a gnuplot/5.2.4-foss-2018a gnuplot/5.2.4-foss-2018a FFmpeg/3.4.2-foss-2018a (D) (D) (D) FEmpeg/3.4.2-intel-2018a flair-geoviewer/2.3-0-foss-2018a flair-geoviewer/2.3-0-foss-2018b flair/2.3-0-foss-2018a flair/2.3-0-foss-2018b Tk/8.6.8-6CCcore-6.4.0-test Tk/8.6.8-6CCcore-6.4.0 Tk/8.6.8-6CCcore-7.3.0 torchvision/0.2.1-foss-2018a-Python-3.6.4-CUDA-9.1.85 (D) (D)(D) onuplot/5.2.4-intel-2018a matplotlib/3.0.3-fosscuda-2018b-Python-3.6.6 (D) gnuplot/5.2.5-foss-2018b GraphicsMagick/1.3.28-foss-2018a GraphicsMagick/1.3.28-intel-2018a Mesa/17.3.6-foss-2018a Mesa/17.3.6-intel-2018a VMD/1.9.3-foss-2018a-Python-2.7.14 VMD/1.9.3-intel-2018a-Python-2.7.14 FLTK/1 3 4-foss-2018; (D) FLTK/1.3.4-foss-2018b (D) FLTK/1.3.4-fosscuda-2018 Mesa/18.1.1-foss-2018b VTK/5.10.1 FLTK/1.3.4-intel-2018a (D) GraphicsMagick/1.3.31-foss-2018b GTK+/2.24.32-foss-2018a HarfBuzz/1.7.5-foss-2018a HarfBuzz/1.7.5-intel-2018a (D) Mesa/18.1.1-fosscuda-2018 VTK/8.1.0-foss-2018a-Python-3.6.4 Mesa/18.1.1-fosscuda-2018b Mesa/19.0.1-GCCcore.8.2.0 motif/2.3.8-intel-2018a OpenCV/3.4.1-foss-2018a-Python-2.7.14 OpenCV/3.4.1-foss-2018a-Python-3.6.4 Pango/1.41.1-foss-2018afontconfig/2.12.6-GCCore-6.4.0 fontconfig/2.12.6-GCCore-7.3.0 fontconfig/2.13.0-GCCore-7.3.0 (D) VTK/8.1.0 intel-2018a-Python-3.6.4 VTK/8.1.1-foss-2018b-Python-2.7.15 VTK/8.1.1-foss-2018b-Python-3.6.6 (D) fontconfig/2.13.1-GCCcore-7.3.0 fontconfig/2.13.1-GCCcore-8.2.0 HarfBuzz/2.2.0-foss-2018b HDFV1ew/2.14-Java-1.8.0_172-centos7 ImageMagick/7.0.7-39-GCCcore-6.4.0 X11/20180131-GCCcore-6.4.0 X11/20180604-GCCcore-7.3.0 (D) (D) freetype/2.9-GCCcore-6.4.0 Pango/1.41.1-intel-2018a X11/20190311-GCCcore-8.2.0 (D) freetype/2.9.1-GCCcore-7.3.0 freetype/2.9.1-GCCcore-8.2.0 FriBidi/1.0.1-foss-2018a FriBidi/1.0.1-intel-2018a JasPer/1.900.1-GCCcore-6.4.0 Pango/1 42 4-fors-2018b x264/20190129-6(Ccore-6.4.0 JasPer/2.0.14-GCCcore-6.4.0 JasPer/2.0.14-GCCcore-6.4.0 JasPer/2.0.14-GCCcore-7.3.0 libGLU/9.0.0-foss-2018a PangO/1.42.4-Toss-2018b ParaView/5.4.1-foss-2018b-mpi ParaView/5.4.1-intel-2018a-mpi ParaView/5.5.2-intel-2018a-mpi PCL/1.9.1-intel-2018a-Python-3.6.4 x265/2.6-GCCcore-6.4.0 xprop/1.2.2-GCCcore-6.4.0 xprop/1.2.3-GCCcore-6.4.0-X11-20180131 xprop/1.2.3-GCCcore-7.3.0 (D) (D) (D) Gdk_Pixbuf/2.36.11_foss_2018: 1 ibGLU/9.0.0-foss-2018 /sw/modules/scs5/tools ---- /sw/modules/scs3/tools ----help2man/1.47.4-GCCcore-6.4.0 help2man/1.47.4-GCCcore-7.3.0 help2man/1.47.4 help2man/1.47.6-GCCcore-8.2.0 help2man/1.47.10-GCCcore-9.1.0 Inspector/2010_update3 ANSYS/19 2 EncyRuild/2.8.0 Subversion/1.9.7-600core-6.4.8 APR-util/1.6.1-GCCcore-6.4.0 APR/1.6.3-GCCcore-6.4.0 EasyBuild/3.8.1 EasyBuild/3.9.1 EasyBuild/3.9.1 Szip/2.1.1-GCCcore-6.4.0 Szip/2.1.1-GCCcore-7.3.0 Szip/2.1.1-GCCcore-8.2.0 binutils/2.26 binutils/2.28-GCCcore-6.4.0 EasyBuild/3.9.2 EasyBuild/3.9.3 EasyBuild/3.9.4 (D) (D) 521p/2.1.1-GCLCore-8.2.0 tcsh/6.20.00-GCCcore-6.4.0 tcsh/6.20.00-GCCcore-7.3.0 (0) (D) binutils/2.28 binutils/2.30-GCCcore-7.3.0 (D) erCodes/2_8_2_intel_2018a
 Inspector/2019_update1
 (D)

 IPython/6.3.1-intel-2018a-Python-3.6.4
 (D)

 IPython/6.4.0-foss-2018a-Python-3.6.4
 (D)

 itar/2018.3.022
 (D)

 Ubssh/1.8.0-6Cccre.6.4.0
 (D)

 Mercurial/4.6.1-foss-2018a-Python-2.7.14
 (M)

 Mercurial/4.6.1-foss-2018a-Python-2.7.14
 (D)

 Mesand/0.30_6Ccccre.8.2.0-Python-3.7.2
 (D)
 Inspector/2019 update1 tmux/2.3 binutils/2.30-GCCcore-8.2.0 binutils/2.30 binutils/2.31.1-GCCcore-8.2.0 expat/2.2.5-GCCcore-6.4.0 expat/2.2.5-GCCcore-7.3.0 till/inux/2.31.1-GCCcore-6.4.0
util-linux/2.32-GCCcore-7.3.0
util-linux/2.33-GCCcore-8.2.0 expat/2.2.6-GCCcore-8.2.0 (D) (D) binutils/2.31.1 binutils/2.32-GCCcore-9.1.0 binutils/2.32 bzip2/1.0.6-GCCcore-6.4.0 gettext/0.19.8 Vampir/unstable gettext/0.19.8 gettext/0.19.8.1-GCCcore-6.4.0-libxml2-2.9.7 gettext/0.19.8.1-GCCcore-6.4.0 gettext/0.19.8.1-GCCcore-7.3.0 Vampir/9.5.0 (D) Vampir/9.6.1 Vampir/9.7.0 bzip2/1.0.6-GCCcore-7.3.0 bzip2/1.0.6-GCCcore-8.2.0 gettext/0.19.8.1-GCCcore-8.2.0 gettext/0.19.8.1 Ghostscript/9.22-GCCcore-6.4.0-cairo-1.14.12 Ninja/1.9.0-GCCcore-8.2.0 numactl/2.0.11-GCCcore-6.4.0 numactl/2.0.11-GCCcore-7.3.0 Vampir/9.7.1 Vampir/9.8-dev (D) (D) bzip2/1.0.6-GCCcore-9.1.0 VTune/2018 update3 VIUne/2018_upuate3 VIune/2019_update2 wheel/0.31.0=foss-2018a=Python-3.6.4 wheel/0.31.1=fosscuda=2018b=Python-2.7.15 wheel/0.31.1=fosscuda=2018b=Python-3.6.6 (D) chroath/0 14 Ghostscript/9.23-GCCcore-6.4.0 chrpath/0.16 cURL/7.58.0-GCCcore-6.4.0 cURL/7.60.0-GCCcore-7.3.0 cURL/7.63.0-GCCcore-8.2.0 cURL/7.63.0-GCCcore-8.2.0 dtt/18.7.2-Redhat-7.0 Ghostscript/9.23-GCCcore-6.4.0 Ghostscript/9.23-GCCcore-7.3.0 git-cola/3.2-intel-2018a git-1fs/2.7.2 git/2.16.1-GCCcore-6.4.0 git/2.18.0-GCCcore-6.4.0 numact1/2.0.12-6CCcore-8.2.0 numact1/2.0.12-6CCcore-9.1.0 p7zip/9.38.1 paralle1/20108822-6CCcore-7.3.0 patche1/0.9-6CCcore-6.4.0 patche1/0.9-6CCcore-7.3.0 (D) (D) (D) XZ/5.2.2-foss-2018b XZ/5.2.2-intel-2018a (D) DMTCP/2.5.1 it/2.19.1-GCCcore-7.3.0 (D) PDFCrop/0.4b XZ/5.2.3-GCCcore-6.4.0 git/2.19.1-GCCcore-7.3.0 GLPK/4.65-GCCcore-6.4.0 GLPK/4.65-GCCcore-7.3.0 gperftools/2.7-GCCcore-6.4.0 gzip/1.8-GCCcore-6.4.0 gzip/1.9-GCCcore-7.3.0 gzip/1.9-GCCcore-8.2.0 petsc4py/3.9.1-foss-2018a-Python-3.6.4 DMTCP/2 5 2 (D) XZ/5.2.4-GCCcore-7.3.0 DMTCP/2.5.2 Dyninst/9.3.2-GCCcore-6.4.0 EasyBuild/3.6.1 EasyBuild/3.6.2 EasyBuild/3.7.0 EasyBuild/3.7.1 petscapy/3.9.1-toss-2018a-Python-3.6 pigz/2.4-GCCcore-7.3.0 pytest/3.8.0-foss-2018a-Python-3.6.4 XZ/5.2.4-GCCcore-8.2.0 XZ/5.2.4-GCCcore-8.2.0 XZ/5.2.4-GCCcore-9.1.0 (D) (D) Serf/1.3.9-GCCore-6.4.0 slepc4py/3.9.0-foss-2018a-Python-3.6.4 source-highlight/3.1.8-foss-2018a YAXT/0.6.0-intel-2018a (D) sw/modul as /scs5/tool chair gompi/2019a iccifort/2018.3.222-GCC-7.3.0-2.30 iccifort/2019.0.117-GCC-7.3.0-2.30 foss/2018; fosscuda/2018b gomkl/2019a (D) iimpi/2018a intel/2018; ionk1/2018 gcccuda/2018b gompic/2018b iccifort/2018.1.163-GCC-6.4.0-2.28 foss/2018 compi/2018a iimpi/2018b intel/2018 iompi/2018a foss/2019a (D) golf/2018a gomp1/2018b iccifort/2019.1.144-GCC-8.2.0-2.31.1 (D) iimpi/2019a (D) intel/2019a (D) pompi/2018.04 CUDA/9.1.85 CUDA/18.0.138 (D) hwloc/1.11.10-GCCcore-8.2.0 hwloc/1.11.11-GCCcore-8.2.0 hwloc/2.0.3-GCCcore-8.2.0 hwloc/2.0.3-GCCcore-9.1.0-metloc hwloc/2.0.3-GCCcore-9.1.0 libpciaccess/0.14-GCCcore-8.2.0 liberiaccess/0.14-6(Ccore-9.1.0 (0) CUDA/9.2.88-GCC-7.3.0-2.30 CUDA/9.2.88 CUDA/10.0.130 hwloc/1.11.8-GCCcore-6.4.0 hwloc/1.11.10-GCCcore-7.3.0 (D) OpenPGM/5.2.122-GCCcore-6.4.0 OpenSSL/1.0.21-GCCcore-6.4.0 -- /sw/modules/scs5/pnys ----Siesta/4.1-b3-intel-2018a SWASH/5.01-intel-2018a UDUNITS/2.2.26-foss-2018b AnsysEM/19.0 COMSOL/5.4 DFTB+/18.1-intel-2018a-Python-2.7.14 DFTB+/18.2-intel-2018a-Python-2.7.14-patched DFTB+/18.2-intel-2018a-Python-2.7.14-mpi DFTB+/18.2-intel-2018a-Python-2.7.14 UDUNITS/2.2.26-fosscuda-2018b (D) VASP/5.4.4-intel-2018a-optics VASP/5.4.4-intel-2018a (D) (D) /sw/modules/srs%/nert Advisor/2018_update3 CubeW/4,4-foss-2018b lo2s/1.1.1-GCCcore-6.4.0 OTF2/2.1.1-intel-2018a (D) Score-P/4 8-foss-2018 (D) ANDI 168/2 4 8 CubeW/4_4-intel-2018a (D) PAPT/5 6 8-form-2019h Score-P/4 8-intel-2018 AMDuProf/3.1.35 CubeGUI/4.4-foss-2018b CUDEW/4.4-1ncc-celoa Devel-NYTProf/6.06-6CCcore-6.4.0-Perl-5.26.1 Devel-NYTProf/6.06-6CCcore-7.3.0-Perl-5.28.0 Devel-NYTProf/6.06-6CCcore-8.2.0-Perl-5.28.1 (D) PAPI/5.6.0-T055-2018D PAPI/5.6.0-GCCcore-6.4.0 PAPI/5.6.0-intel-2018a (D) scorep_plugin_fileparser/1.3.1-GCCcore-6.4.0 MUST/1.6.0-rc3-intel-2018a 0PARI2/2.0.3-foss-2018b 0PARI2/2.0.3-intel-2018a (D) 0TF2/2.1.1-foss-2018a 0TF2/2.1.1-foss-2018b CubeGUI/4.4-intel-2018a (D) CubeLib/4.4-foss-2018b CubeLib/4.4-intel-2018a (D) PDT/3.25-foss-2018a PDT/3.25-foss-2018b PDT/3.25-intel-2018a lo2s/1.0.1-GCCcore-6.4.0 lo2s/1.0.2-GCCcore-6.4.0 (0) /sw/modules/scs5/numlib NLopt/2.4.2-GCCcore-7.3.0 NLopt/2.4.2-intel-2018a OpenBLA5/0.2.20-GCC-6.4.0-2.28 OpenBLA5/0.3.5-GCC-3.20-2.30 OpenBLA5/0.3.5-GCC-0.2.0-2.31.1 ScaLAPACK/2.0.2-qompi-2018b-0penBLAS-0.3.1 ScaLAPACK/2.0.2-qompi-2019a-0penBLAS-0.3.5 ScaLAPACK/2.0.2-qompi-2019a-0penBLAS-0.3.1 (D) SLEPC/3.9.2-foss-2018a SuiteSparse/5.1.2-foss-2018a-METIS-5.1.0 SuiteSparse/5.1.2-foss-2018b-METIS-5.1.0 arpack-ng/3.5.0-foss-2018a FFTW/3.3.8-gompi-2018b FFTW/3.3.8-gompi-2019a FFTW/3.3.8-gompi-2019a GSL/2.4-GCCcore-6.4.0 arpack-ng/3.5.0+fos5-2018a arpack-ng/3.5.0+fos2-2018b arpack-ng/3.5.0+intel-2018a arpack-ng/3.6.1-intel-2018a (D) CGAL/4.11.1-fos5-2018a-Python-3.6.4 (CGAL/4.11.1-intel-2018a-Python-2.7.14 (D) (D) (D) GSL/2.4-GCCcore-6.4.0 GSL/2.5-GCC-7.3.0-2.30 GSL/2.5-GCC-8.2.0-2.31.1 (D) (D) PETSc / 3, 8, 3, foss, 2018a, downloaded, dens CUDNN/7 @ 5-CUDA-9 1 85 Hypre/2.14.0-foss-2018a imkl/2018.1.163-gompi-2018a PETSc/3.8.3-foss-2018a-64bit-downloaded-deps PETSc/3.8.3-foss-2018a-64bit-downloaded-deps SuiteSparse/5.1.2-intel-2018a-METIS-5.1.0 (D) SuperLU/5.2.1-intel-2018b cuDNN/7.1.4-CUDA-9.2.88 SuperLU/5.2.1-intet-2018b SuperLU DIST/6.1.0-intel-2018b-CUDA cuDNN/7.1.4.18-fosscuda-2018 inkl/2018.1.163-iimpi-2018a PETSc/3.9.3-foss-2018a UDAN/7 4 2 24-CUDA-18 8 128 (D) imkl/2018.1.163-icmpi-2018a imkl/2018.3.222-impi-2018b imkl/2019.1.144-gompi-2019a imkl/2019.1.144-impi-2019a (D) PETSc/2 9 A-forr-20185 (D) SuperLU_DIST/6.1.0 intel-2018b SuperLU_MT/3.1-intel-2018b Trilinos/12.12.1-foss-2018a-Python-3.6.4 (0) 4.2.24-CUDA-10.0.130 4-intel-2018a-PFFT-20150905 .7-GCC-6.4.0-2.28-serial PET/1.0.8-alpha-intel-2018a qrupdate/1.1.2-foss-2018b qrupdate/1.1.2-foss-2018b FFTW/3.3.7-gompi-2018a FFTW/3.3.7-intel-2018a NLopt/2.4.2-foss-2018a ScaLAPACK/2.0.2-gompi-2018a-OpenBLAS-0.2.20 (sw/modul as /srs5/moj 5/mpi OpenMPI/2.1.5-GCC-6.4.0-2.28 OpenMPI/3.1.1-GCC-7.3.0-2.30 (D) OpenMPI/3.1.1-gcccuda-2018b OpenMPI/3.1.2-GCC-8.2.0-2.30 OpenMPI/3.1.2-PGI-18.4 impi/2018.1.163-iccifort-2018.1.163-GCC-4.4.8-2.28 OpenMPT/1 10 7-PGT-18 4-GCC-6 4 0-7 28 OpenMPT/2 1 2-600-8 2 8-2 21 1 imp1/2018.1.863-6CC107T-2018.1.163-6C-6-4.0-2.28 imp1/2018.3.222-icc170-72018.3.222-6CC7.3.0.2.30 imp1/2018.4.274-1cc1fort-2019.1.144-6CC-8.2.0-2.31.1 (D) OpenMP1/1.10.7-6CC-6.4.0-2.28 OpenMP1/1.10.7-6CC-6.4.0-2.28 OpenMPI/1.10.7-901-18.4 OpenMPI/1.10.7-901-18.4 OpenMPI/2.1.2-6CC-6.4.0-2.28 OpenMPI/2.1.2-FGI-18.4-6CC-6.4.0-2.28 OpenMPI/2.1.2-iccifort-2018.1.163-6CC-6.4.0-2.28 OpenMPI/3.1.3-GCC-8.2.0-2.31.1 OpenMPI/4.0.1-GCC-9.1.0-2.32 OpenMPI/4.0.1-PGI-19.4-GCC-8.2.0-2.31.1 /sw/modules/srs5/math DOLETN/2018 1 0 post1_foss_2018a_Python_3 6 4 Gurobi/8 0 1 MATLAR/20185 (D) ParMETTS/A @ 3_intel_2018 (D) Eigen/3.3.4-GCCcore-6.4.0 Keras/2.2.0-foss-2018a-Python-3.6.4 Keras/2.2.4-foss-2018a-Python-3.6.4 Maxima/5.42.1 METIS/5.1.0-GCCcore-6.4.0 Qhull/2015.2-GCCcore-6.4.0 Ohull/2015.2-GCCcore-7.3.0 (D) (D) igen/3.3.4 1gen/3.3.4 LPA/2016.11.001.pre-iomk1-2018a keras/2.2.4-fosscuda-2018b-Fython-3.6.6 (D)
libcerf/1.5-GCCcore-6.4.0
libcerf/1.5-GCCcore-7.3.0
libcerf/1.7-GCCcore-7.3.0
(D) METIS/5.1.0-GCCcore-7.3.0 METIS/5.1.0-GCCcore-8.2.0 MPFR/4.0.1-GCCcore-6.4.0 SCOTCH/5.1.12b_esmumps-foss-2018a SCOTCH/6.0.4-intel-2018a SCOTCH/6.0.5-foss-2018a ELPA/2018.11.001-intel-2019a FFC/2018.1.0-foss-2018a-Python-3.6.4 FIAT/2018.1.0-foss-2018a-Python-3.6.4 (D) (D) (D) MPER/4 0.1-GCCcore-7.3 0 (D) SCOTCH/6 @ 6-foss-2018; Scauspill Andreas Knüpfer – Dresden Probabilistic Workshop – 2019-10-10 DRESDEN LEIPZIG



Software for HPC-DA: Modules and Containers

Singularity containers

- Tailored software environments that you can take with you or share with others
- Can be defined/built on top of each other
- Independent (mostly) from central software installation, updates, version changes, ...
- Cannot combine two existing containers, though
- Automatic building from definition files
- Combine with version control for definition files
- Some challenge to build containers for Power9
 ... because few have Power9 laptops yet;)
- Automatic solution with Singularity definition files

- Designed for HPC
- Works well with MPI, Infiniband, GPUs, ...
- Thin run-time layer
- No danger of privilege escalations



- Huge supply of prebuilt containers available
- Also compatible with Docker images

See https://sylabs.io/docs/





Topics overview

IT Services for TU Dresden	 Campus network, internet uplink, phone network E-Mail, groupware, data exchange, backup etc. Software procurement
Computational Science Services	 Virtual machines, hosting Supercomputing Big Data methods
Research and Development	 Parallel programming and algorithms Performance optimization and scaling, energy efficiency, Data Analytics applications, machine learning,









Selected Computer Science Challenges



ZIH Computer Science Research Topics

- Scalable software tools to support the optimization of applications for HPC systems
- Data Intensive Computing and Data Life Cycle
- Performance and energy efficiency analysis for innovative computer architectures
- Distributed Computing and Cloud Computing
- Data analysis, methods, and modeling in life sciences
- Parallel programming, algorithms and methods















Andreas Knüpfer – Dresden Probabilistic Workshop – 2019-10-10





Parallel Performance Analysis Tools



High Performance Computing

Parallel Performance Analysis

Enable or improve computational sciences

- Throughput and Response time
- Strong and weak Scaling
- Capability for larger / more complex scenarios
- Additional functionality / novel aspects
- Reliability
- Development Cycle











Performance Consulting is a Cyclic Process





DRESDEN LEIPZIG

Analyzing Unexpected Memory Demand







Analyzing Bottleneck in Multi-Stage-I/O







Load Balance Analysis of Weather Forecast Model







Load Balance Analysis of Weather Forecast Model After Tuning











Parallel Programming Abstractions with C++



High Performance Computing

Spectrum of Computing Architectures







Shared Memory vs. Distributed Memory Programming







Node

Message Passing



Performance, runs everywhere

Productivity





PGAS – Combining the Advantages of both Approaches PGAS: Partitioned Global Address Space



Mem. Sys. + PGAS Runtime Layer

PGAS Languages

Chapel, CoArray Fortran, UPC, ...

PGAS *Libraries*

Global Arrays (GA), GASPI, OpenShmem, MPI3.0 RMA

Locality control, runs everywhere, performance and productivity





DASH C++ Template Library for Parallel Programming

- C++ template library for application programmers
- Distributed data container classes
- Similar to the C++ STL container classes, compatible
- Built-in knowledge about distribution
- Algorithms similar to STL on distributed containers









DASH Array

DASH n-dimensional array

- Global random access with begin(), end() and [] ... via slow element-wise get
- Dedicated local access with
 myarray.local.begin() / .end()
 and .local[] ... direct and fast
- Configurable data distribution patterns in n dimensions
- STL-like algorithms considering actual data distribution patterns









Selected Results from Computational Science Projects at ZIH



Prestige Project

Ongoing LuFo V project PRESTIGE "Interdisciplinary probabilistic analysis of compressor blades accounting for real geometry effects"

TU Dresden partners:

- Chair of Turbomachinery and Flight Propulsion (TFA)
- Institute of Lightweight Engineering and Polymer Technology (ILK)
- ZIH
- Coordinator Rolls-Royce Deutschland Ltd & Co KG
- DLR Braunschweig and Köln
- TU Darmstadt
- University of Surrey (UK)
- Kitware
- Intelligent Light
- T-Systems Solutions for Research GmbH

Gefördert durch:



aufgrund eines Beschlusses des Deutschen Bundestages

... not the focus today









Which one next?

Laser Particle Acceleration

or

Cartography of the Milky Way



High Performance Computing





Courtesy of Dr. Michael Bussmann et.al.

Novel Particle Accelerators and Highly Scalable GPU-based Simulation



Center for Information Services & High Performance Computing

— Advanced method of

electron acceleration

- Based on highly non-linear laser plasma interaction
- Requires large scale particle-in-cell simulation for modeling
- Hundreds of simulations on up to 146 K80 GPUs performed with PIConGPU at Taurus/ZIH

PICon CPU



<u>Video</u>



Laser Wakefield Acceleration Experiments at HZDR

J.P. Couperus et al.: Demonstration of a beam loaded nanocoulomb-class laser wakefield accelerator. **Nature Comunications** 8.1 (2017) A. Irman et al.: Improved performance of laser wakefield acceleration by tailored self-truncated ionization injection. **LPAW proceeding** pp.1-13 (2017)



Simulations accompanying experiment at HZDR



Studying the influence of higher order laser modes





Laser-Ion Acceleration with Mass-Limited Targets

P. Hilz, T.M. Ostermayr, A. Huebl et al.: Isolated proton bunch acceleration by a petawatt laser pulse. Nature Comunications 9.423 (2018) A. Huebl et al.: On the Scalability of Data Reduction Techniques in Current and Upcoming HPC Systems from an Application Perspective, ISC'17, LNCS 10524

- 3D simulation of novel, fully isolated target for laser-ion acceleration
- 15 M CPUhrs (½ MGPUhrs), INCITE Award Highlight
 PByte-Scale I/O through ADIOS at Titan/OLCF









Courtesy of Prof. Sergei A. Klioner et.al.

Astrophysics: The Gaia Project



Center for Information Services & High Performance Computing







One of the main problems of astronomy: distance

Without knowing how far the object is, physical understanding of that object is

impossible...





A comet: far away and very big or inside the Earth atmosphere and rather small? Tycho Brahe, 1577







Astrometry: the art of measuring stellar positions

Astronomy cannot touch its objects!

Astronomy cannot make experiments!

Astronomy analyses stellar light:

Astrometry Photometry Spectroscopy Polarimetry

- direction
- quantity
- colour and more
- polarization

+ cosmic particles ++ gravitational waves









FSDF

Why to bother?

- We need to understand stars.
 - (our Sun is a star!)

Without knowing the distance it is not possible to judge if a star is big or small, etc.

– We live in a galaxy.

We need to understand how our Galaxy was formed.



Accuracy of astrometric observations



TECHNISCHE UNIVERSITAT 1 µas is the thickness of a sheet of paper seen from the other side of the states and reas Knupper bresden Probabilistic Workshop - 2019-10-10 other side of the states and reas Knupper - Dresden Probabilistic Workshop - 2019-10-10 other side of the states and reas the states and rea

What do we know about our Galaxy?

The Sun should be here:











What do we know about our Galaxy?

The stars with distances known till 2016 are all in the small red spot:









Andreas Knüpfer – Dresden Probabilistic Workshop – 2019-10-10

What do we know about our Galaxy?

With Gaia we can explore a significant part of our Galaxy:









Andreas Knüpfer – Dresden Probabilistic Workshop – 2019-10-10

The challenge of data processing

- Parameters
 - At least 5 parameters for each star: 5 x 1.7 10⁹
 - 4 parameters of orientation each 15 seconds: 10⁸
 - 2000 calibration parameters per day: 4 10⁶
 - global parameters: <10⁴
- Observations

about 1000 raw images for each star: 10¹²

- Data volume: 1 PB (iteratively!)
- Computational efforts: ~10²² flops
- Direct least squares solution is impossible









Gaia in Dresden

- 1. The model of observations in Einstein's General Relativity
- 2. Tests of fundamental physical laws with Gaia data
- 3. Analysis of the Gaia reference frame: quasars
- 4. Synchronization and monitoring of Gaia's atomic clock
- 5. Special astrometric solutions: stability and quality verification, special calibration of the instrument, relativistic tests

Special thanks to ZIH for about 3 Million CPU-hours by now!

<u>Video</u>









Summary



Center for Information Services & High Performance Computing

Computer Science and Computational Science at ZIH

- Exciting computer science research
- Broad spectrum of computational science topics together with application science projects

Contact:

Dr. Andreas Knüpfer Deputy Director / CTO of ZIH andreas.knuepfer@tu-dresden.de Tel. +49 351 463-38323 Willersbau A113





