Scientific Computing @ MPP

Stefan Kluth

MPP Project Review

19.12.2017

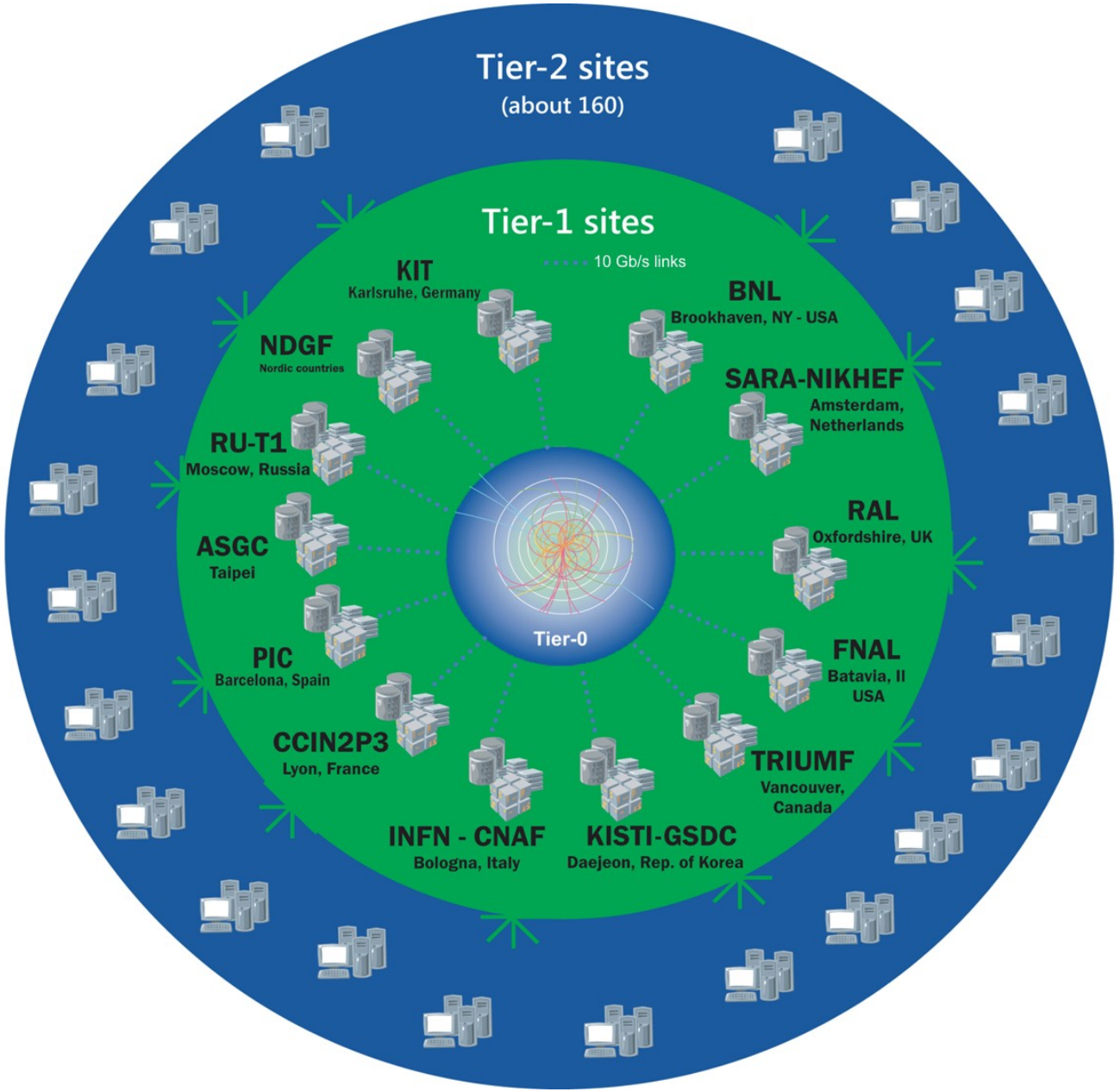
Science with computers

- The scientific method (simplified)
  - Experiment: design a setup and collect data, infer from data underlying principles; test theories
  - Theory: build up from fundamentals a mathematical framework to describe nature and make predictions; learn from experiment data
- With computers
  - Numerical simulation: translate abstract / unsolvable models into practical predictions, discover behavior
  - Find structures in (unstructured) data

Overview

- Some applications
  - ATLAS
  - Theory: see Stephen Jones talk
- Data Preservation
- Software development example
  - BAT
- Resources
  - MPP, MPCDF, LRZ, Excellence Cluster (C2PAP)

ATLAS WLCG



Tier-0: CERN

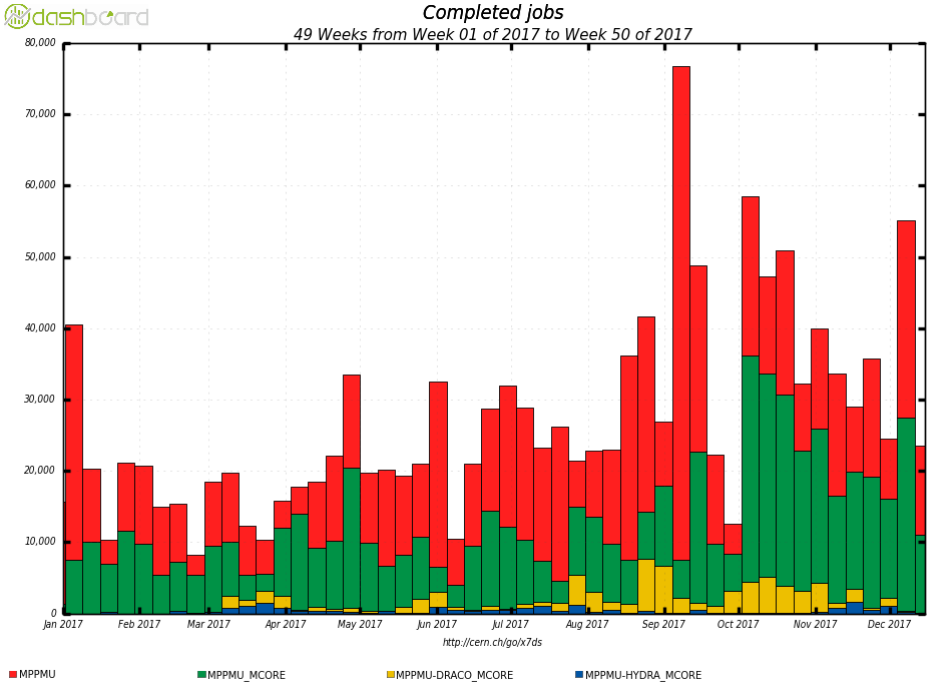
Tier-1: GridKa

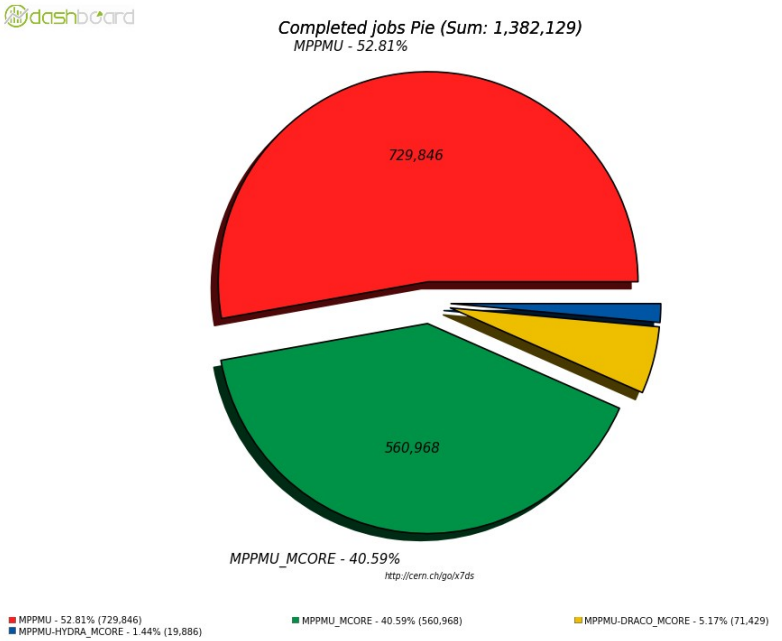
Tier-2: MPPMU

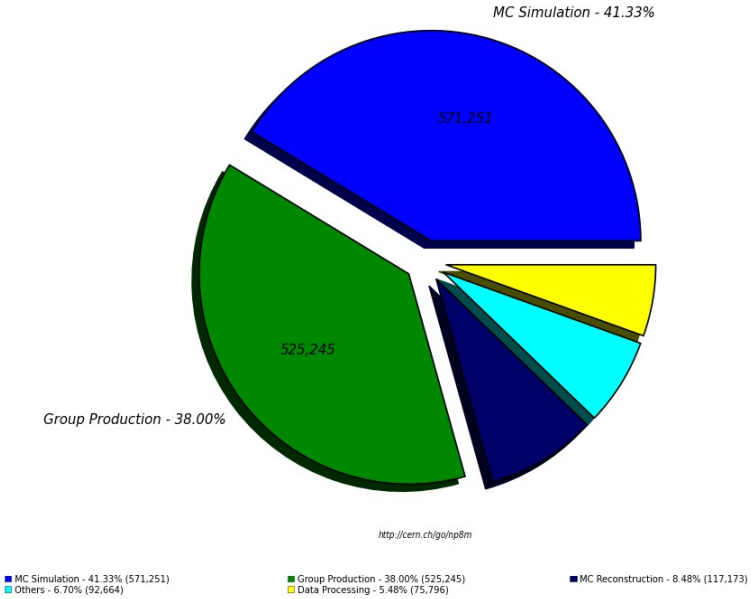
Originally hierarchical, moving to network of sites

MAGIC, CTA, Belle 2 following this model, our Tier-2 supports this

ATLAS MPP Tier-2 & Co







50% nominal Tier-2
1/60 of total ATLAS Tier-2
Incl. “above pledge” contributions

DRACO is HPC at MPCDF
“opportunistic”

DPHEP

Andrii Verbytskyi

- MPP has several experiments with valuable data and ongoing analysis activity
- H1 and ZEUS @ HERA
- OPAL @ LEP and JADE @ PETRA
- See Andrii Verbytskyi talk
  - and previous project reviews since 2000

DPHEP

- Save bits: copy data to MPCDF
  - Provide access via open protocols (http, dcap)
  - Use grid authentication (X509)
  - About 1 PB (H1, ZEUS, OPAL, JADE), goes to tape library
- Save software: installation in virtual machine
  - Provide validated environment (SL5, SL6, ...)
- Save documentation: labs, inspire, …
  - Older experiments: scan paper-based documents

Current status of H1&ZEUS DP

Data/MC	ZEUS	H1
DESY archive	Processed data/MC ntuples	Raw data/MC, processed data/MC
DESY available online	Everything	Everything
DESY access	NFS, +2 machines in DESY	NFS, +2 machines in DESY
MPCDF/MPP online+archive	As in DESY+raw data	As in DESY (online)
MPCDF/MPP access	Multiprotocol, worldwide with ZEUS VO cert.	Multiprotocol, worldwide with H1 VO cert.
Software		
DESY reconstruction	No	Yes
DESY MC generation	No	Yes
DESY analysis	Yes	Yes(up to 5y)
DESY user storage	Yes, limited, on 2 machines in DESY	Yes, limited, on 2 machines in DESY
DESY environment	2 machines in DESY+BIRD(up to 5y)	2 machines in DESY+BIRD(up to 5y)
MPCDF/MPP reconstruction	Yes	No
MPCDF/MPP MC generation	Yes	No
MPCDF/MPP analysis capability	Yes	Generic
MPCDF/MPP user storage	Yes	Yes
MPCDF/MPP environment	CentOS7 virtual machine available	Generic CentOS7 virtual machine
Documentation		
DESY analysis primer/manual	Archived web-server	Archived web-server
DESY legacy notes, drafts etc.	InSpire+DESY library	InSpire+DESY library
DESY MC primer/manual	No	No
DESY preservation paper/note	No	No
MPCDF/MPP analysis primer/manual	Relies on DESY	Relies on DESY
MPCDF/MPP MC primer/manual	Yes	No
MPCDF/MPP legacy notes, drafts etc.	InSpire+DESY library	InSpire+DESY library
MPCDF/MPP preservation paper/note	Non-public draft	No

DESY	Finished	Finished, but not optimal	Significant advance	Moderate advance	Will not be done
MPCDF/MPP	Finished	Finished, but not optimal	Significant advance	Moderate advance	Will not be done

2/3

Current status of OPAL&JADE DP

Data/MC	OPAL, Host=CERN	JADE, Host=DESY
Host data	Raw/processed, data/MC	No
Host access	Multiprotocol, CERN	No
MPCDF/MPP available online	Raw/processed, data/MC	Raw/Processed Data/MC
MPCDF/MPP archive	Raw/processed, data/MC	Raw/Processed Data/MC
MPCDF/MPP access	Multiprotocol, worldwide with OPAL VO cert.	Multiprotocol, worldwide with ZEUS VO cert.
Software		
Host reconstruction	Yes	No
Host MC generation	Yes	No
Host analysis	Yes	No
Host user storage	CERN users only	No
Host environment	Default CERN	No
MPCDF/MPP reconstruction	Yes	Yes
MPCDF/MPP MC generation	Yes	Yes
MPCDF/MPP analysis	Yes	Yes
MPCDF/MPP user storage	Yes	Yes
MPCDF/MPP environment	CentOS7 VM available	CentOS7 VM available
Documentation		
Host analysis primer/manual	CERN web-server	No
Host legacy notes, drafts etc.	InSpire+CERN library	InSpire+DESY+J.O.
Host preservation paper/note	Yes	No
MPCDF/MPP analysis primer/manual	Relies on CERN, data description	Yes (Update!)
MPCDF/MPP legacy notes, drafts etc.	InSpire+CERN library	InSpire+DESY library+
MPCDF/MPP preservation paper/note	No	Yes

Host	Finished	Finished, but not optimal	Significant advance	Moderate advance	Will not be done
MPCDF/MPP	Finished	Finished, but not optimal	Significant advance	Moderate advance	Will not be done

Bayesian Analysis Toolkit (BAT)

Oliver Schulz

- Markov Chain Monte Carlo (MCMC) sampling
  - Metropolis-Hastings algorithm
- Sample likelihood (model + data)
  - As function of model parameters
  - Contains prior pdf for model parameters
  - Result is posterior pdf for model parameters given a data set
- Can be computationally costly
  - Many model parameters
  - Large data sets
  - Complex model

BAT

Bayes Theorem:

$$P(\rho \mid X) \sim P(X \mid \rho) \cdot P(\rho)$$

X: data set, ρ: model parameters, $P(X \mid \rho)$ model likelihood, $P(\rho)$: prior likelihood, $P(\rho \mid X)$ posterior likelihood of ρ given Data set X and model in $P(X \mid \rho)$

Metropolis-Hastings Algorithm:

$$P_a(x_{i+1} \mid x_i) = \min(\ 1,\ P(x_{i+1}) P_p(x_{i+1} \mid x_i) \ /\ P(x_i) P_p(x_i \mid x_{i+1})\ )$$

Proposal density $P_p(x_{i+1} \mid x_i)$

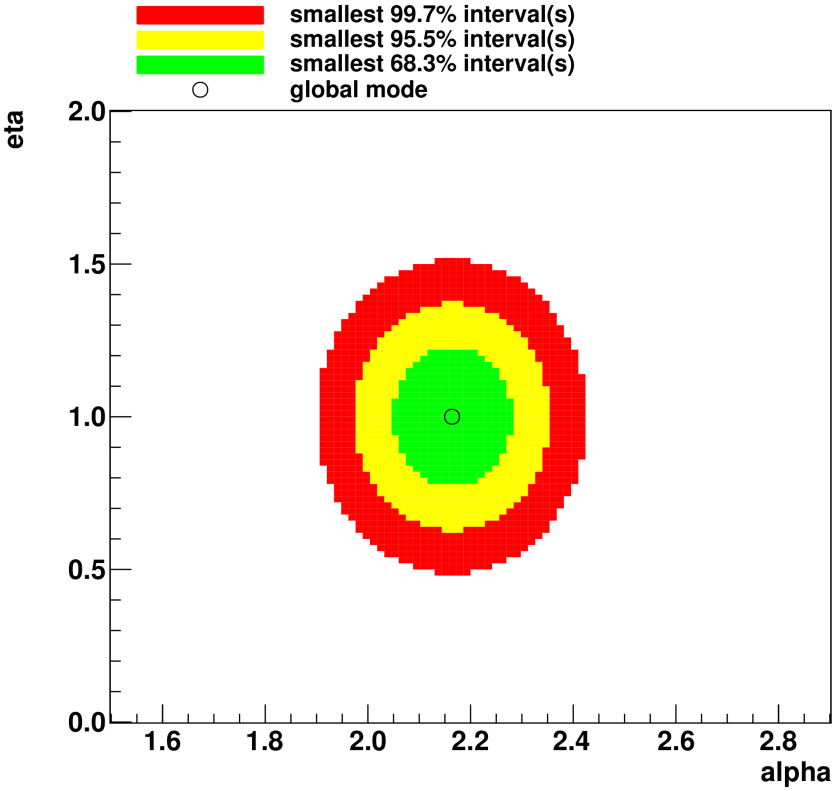
BAT

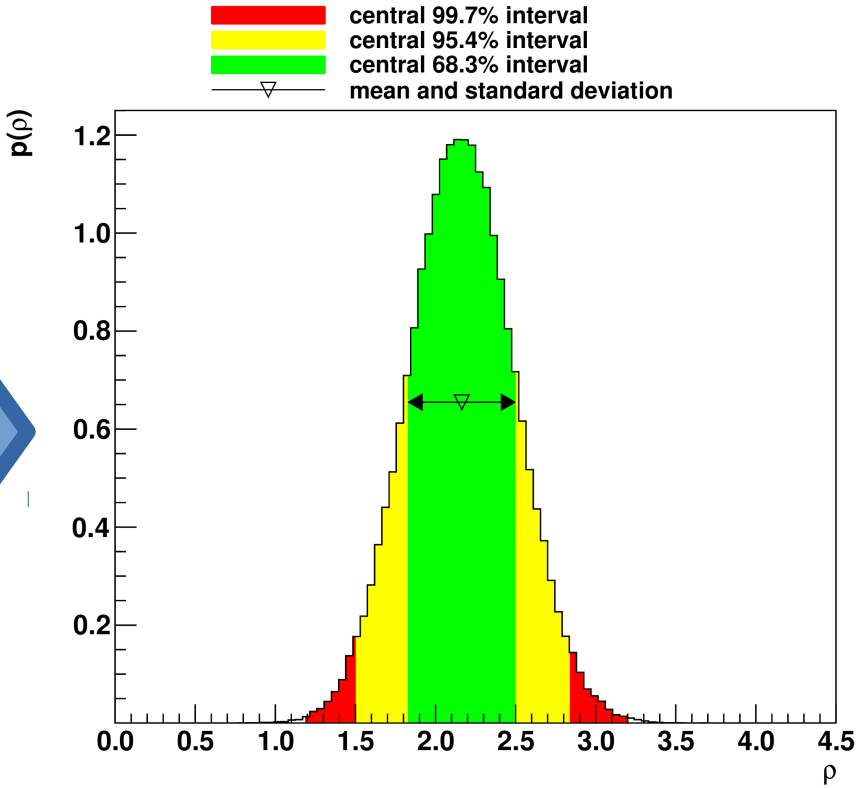
Two results $q_1 = 2.4 \pm 0.12$; $q_2 = 2.0 \pm 0.10$, norm. $N = 1.0 \pm 0.15$

$r_i = Nq_i$ and $\rho = \eta\alpha$ for parameters $\rho \leftrightarrow r_i$, $\eta \leftrightarrow N$, $\alpha \leftrightarrow q_i$

Average of r_i is estimator for ρ

Model likelihood: $P(\{q_i\},N \mid \rho) = \int\int d(\rho-\eta\alpha)G(\{q_i\} \mid \alpha)G(N \mid \eta)\, d\alpha d\eta$





$\langle\rho\rangle = 2.164 \pm 0.334$

BAT

- BAT up to 1.0

bat.mpp.mpg.de
github.com/bat

  - Stable product, large user base, many publications
  - C++ incl. Root
  - BAT 1 not easy to integrate in e.g. python, R, etc.
  - Code not optimal for parallelism
  - Not easy for other sampling algorithms
- BAT 2 project
  - Rewrite in Julia language (first usable release expected in 2018)

Theory

Thomas Hahn

Package Types

'Production'	'Exploration'	'Specific'
		
MG5_aMC@NLO	FormCalc	FeynHiggs
GoSam	FeynCalc	DarkSUSY
OpenLoops	Package-X	Prospino

– p.1

Theory

Packages Developed/Maintained at MPP

'Production' Type: model to events "automatically"

- GoSam – computation of (mainly QCD) cross-sections
- (py)SecDec – evaluation of multiloop integrals
- Cuba – multidimensional numerical integration

'Exploration' Type: toolkit for unusual models, unusual renormalizations, package building etc.

- FeynArts – diagram generation
- FormCalc – 'analytic' computation of one-loop amplitudes
- LoopTools – evaluation of one-loop integrals

'Specific' Type: explicit calculations

- FeynHiggs – computation of Higgs observables

– p.2

Resources: general

- MPCDF
  - Hydra: 338 nodes with dual Nvida Tesla K20X, 2500 new nodes 40 cores arriving
  - Draco: midsize HPC, 880 nodes 32 cores, 106 nodes with GTX980 GPUs
- LRZ
  - SuperMUC: >12.000 nodes, 241.000 cores, fast interconnect
  - To be replaced soon SuperMUC-ng
- Excellence Cluster Universe
  - C2PAP: 128 nodes, >2000 cores, fast interconnect, SuperMUC integration

Ressources: MPP@MPCDF

- Computing
  - 144 nodes, 3.250 cores
  - SLC6, SLURM batch, singularity
  - WLCG
  - User interface nodes mppui[1-4]
  - mppui4 (fat node) has 1 TB RAM
- Storage
  - 4.5 PB storage on RAID arrays
  - IBM gpfs shared filesystem (/ptmp/mpp/...)
  - dCache data storage (xrootd, http, …)
  - Connection to tape library via gpfs possible

Resources: MPP

- Computing
  - > 200 desktop PCs via condor batch system
    - Ubuntu 16.04 or Suse tumbleweed
  - 2 fat nodes with 512 GB RAM (theory)
    - Memory intensive programs e.g. reduze (Feyman diagram to master integral reduction) jobs etc
  - Fat nodes partially with Nvidia GPUs (Gerda group)
- Storage
  - CEPH storage (/remote/ceph/...)
  - Local scratch disks (/mnt/scratch/...)

Virtualisation / Linux containers

- Linux PCs offer VirtualBox
  - Any user able to run VMs, Windows or Linux
  - Behind NAT, IP address on request
  - Host file system access possible
  - Fixed RAM allocation, heavy images
- Singularity (2.4.x, available soon)
  - Run different Linux images in user mode
    - e.g. SLC6 on ubuntu 16.04, Suse tumbleweed on SLC6 on MPP cluster at MPCDF …
    - Must be root to build images → use VMs
  - Share host filesystem e.g. /remote/ceph or /cvmfs

Summary

- Scientific computing essential for our success
- Many activities at MPP
  - From software development to data preservation
- Resources: MPP, MPCDF, LRZ, C2PAP
- All centers provide application support
  - Porting to parallel platforms, performance tuning, …
- Transition to HPC in many of our research areas