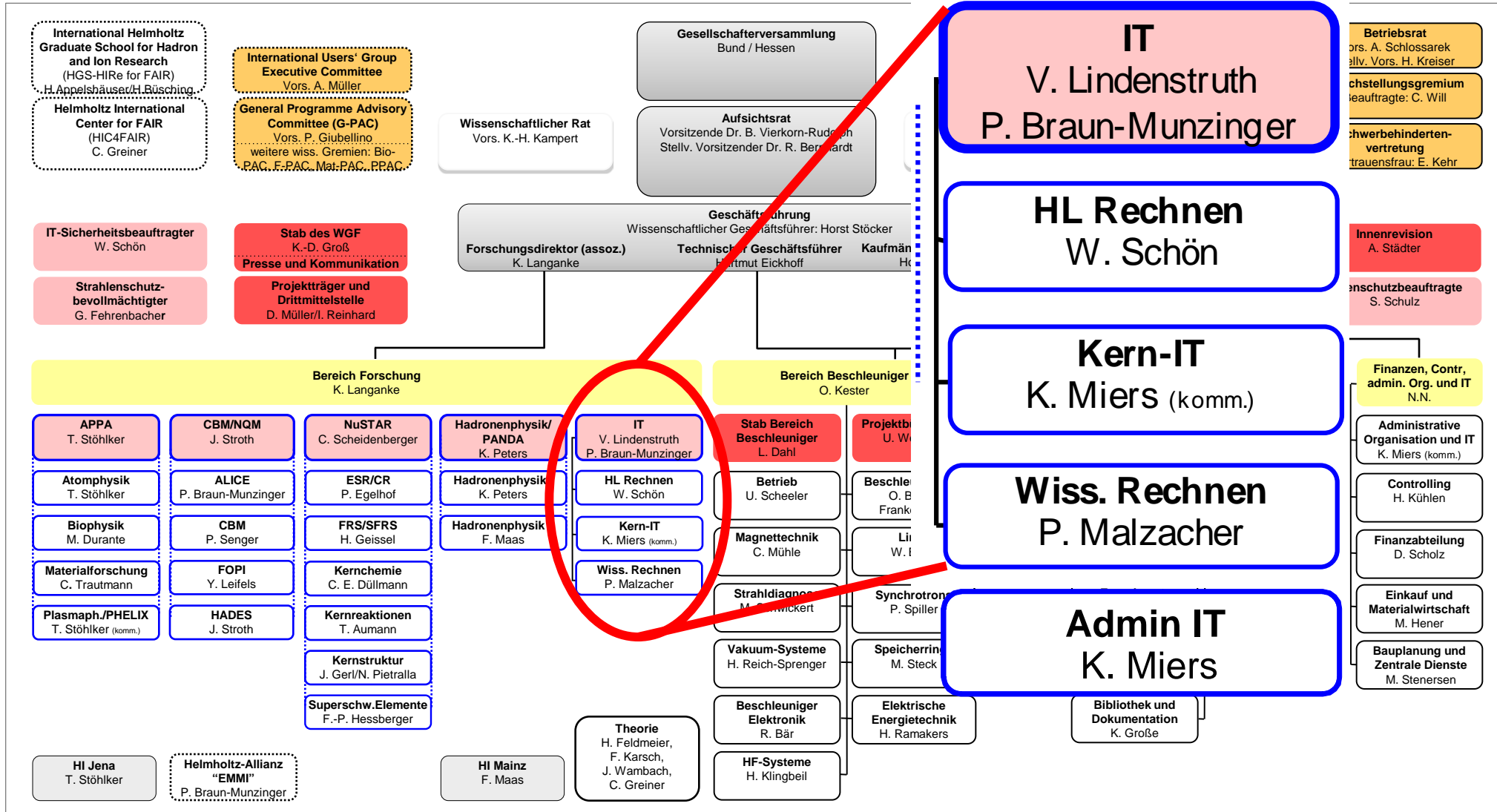




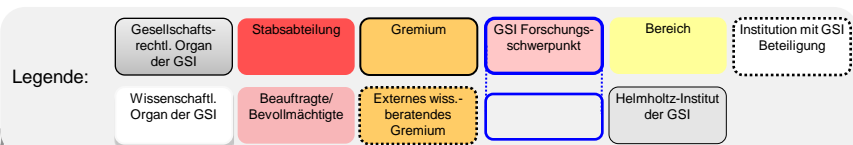
21.5.2011

GSI+FAIR Computing

Volker Lindenstruth
FIAS, IfI, LOEWE Professur
Chair of HPC Architecture
University Frankfurt, Germany
Phone: +49 69 798 44101
Fax: +49 69 798 44109
Email: voli@compeng.de
WWW: www.compeng.de



Stand: Mai 2011.
 Der Geschäftsbereich FAIR ist mit Gründung der FAIR GmbH am 4. Oktober 2010 weggefallen. Voraussichtlich im Sommer 2011 soll der neue GSI-Gesellschaftsvertrag wirksam werden. Die Länder Rheinland-Pfalz und Thüringen werden weitere Gesellschafter der GSI, die wissenschaftlichen Organe WA und WD werden durch den Wissenschaftlich-Technischen Rat (WTR) ersetzt.



Programm	2010			2013			2016		
	CPU # cores	On-line PB	Tape PB/a	CPU # cores	On-line PB	Tape PB/a	CPU # cores	On-line PB	Tape PB/a
APPA	100	< 0,1	0,1	200	1,0	2,0	1.000	1,0	2,0
CBM	250	< 0,2	0,2	7.000	15,0	11,0	60.000	15,0	11,0
NUSTAR	100	< 0,1	0,0	300	2,0	2,0	2.000	2,0	2,0
PANDA	250	< 0,2	0,2	8.000	12,0	12,0	66.000	12,0	12,0
ALICE	3.000	1,5	0,0	6.600	3,5	0,0	14.500	8,0	0,0
FOPI	200	0,1	0,2	400	0,2	0,3			
HADES	1.000	1,0	1,5	1.000	2,0	3,0			
Gitter-QCD	2.500	< 0,3	0,3	20.000	0,5	0,5	100.000	1,0	2,0
Theorie	1.000	< 0,2	0,2	5.000	0,5	0,5	20.000	1,0	2,0
rest	500	< 0,3	0,3	2.000	1,5	1,5	10.000	3,0	3,0

Objekt



USA

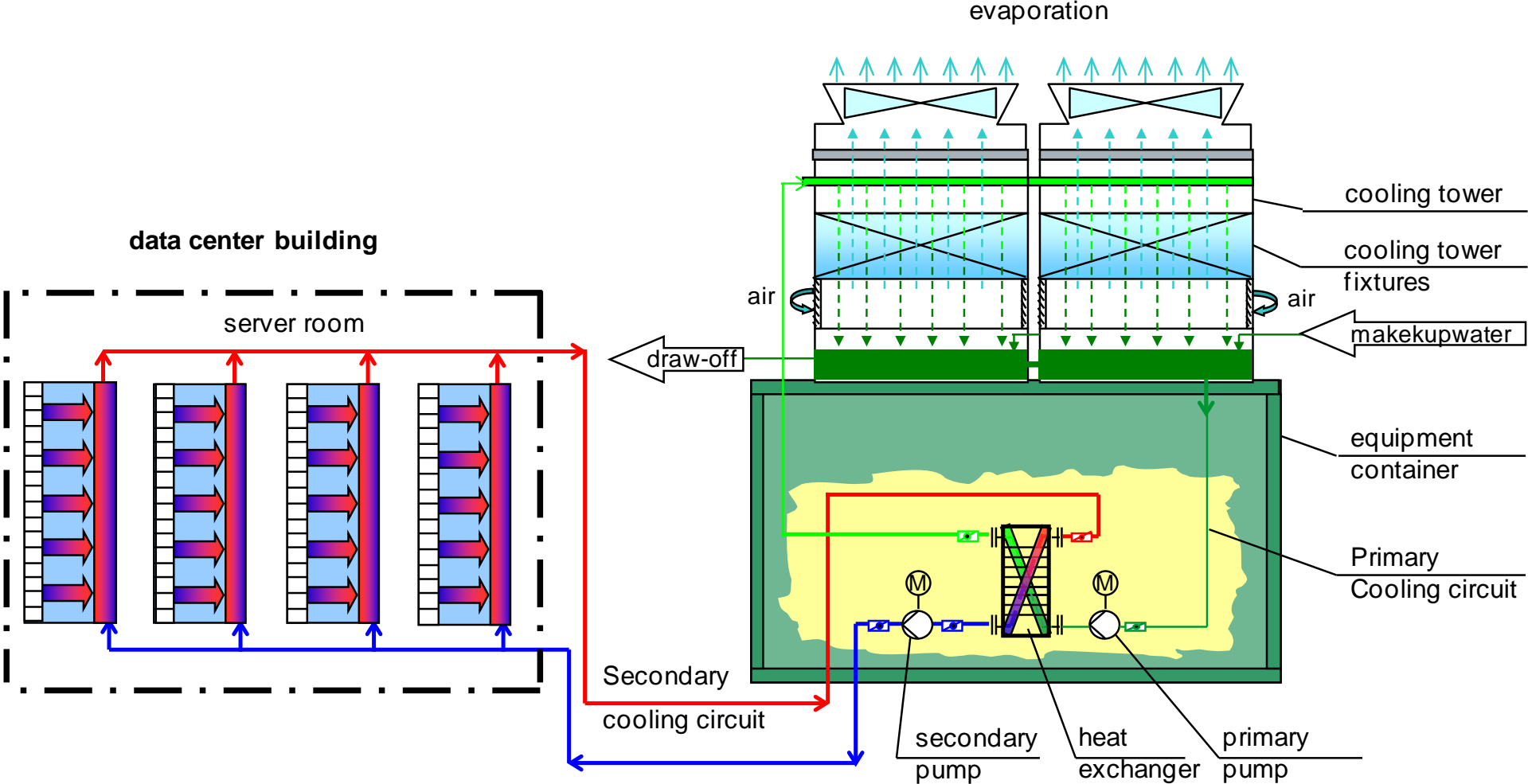


Austria China Finnland France Germany Greece India Italy Poland Slovakia Slovenia Spain Sweden Romania Russia UK

LOEWE-CSC 2011

Green500: 8
Top500: 22 (21)
Cost: 200 €/core
PUE: 1,07
CO₂ neutral

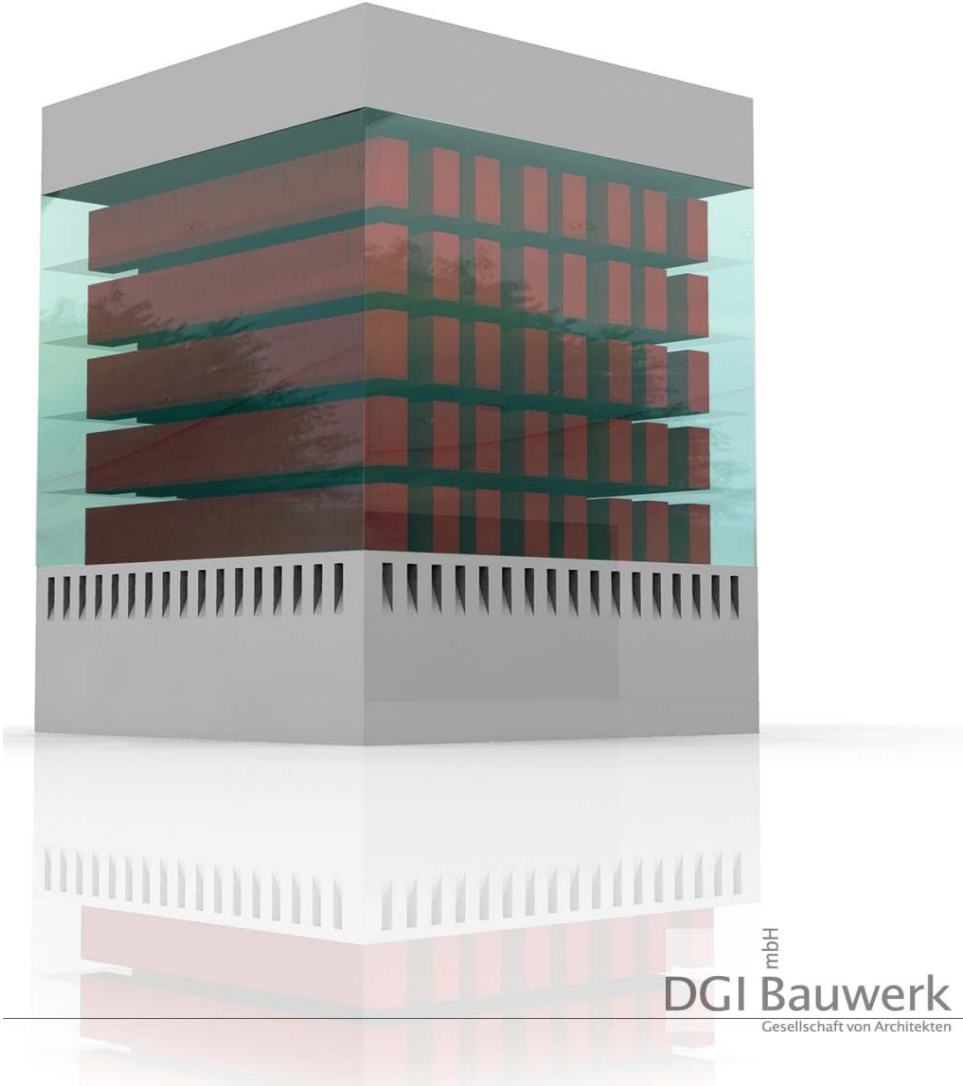
Cooling System Architecture



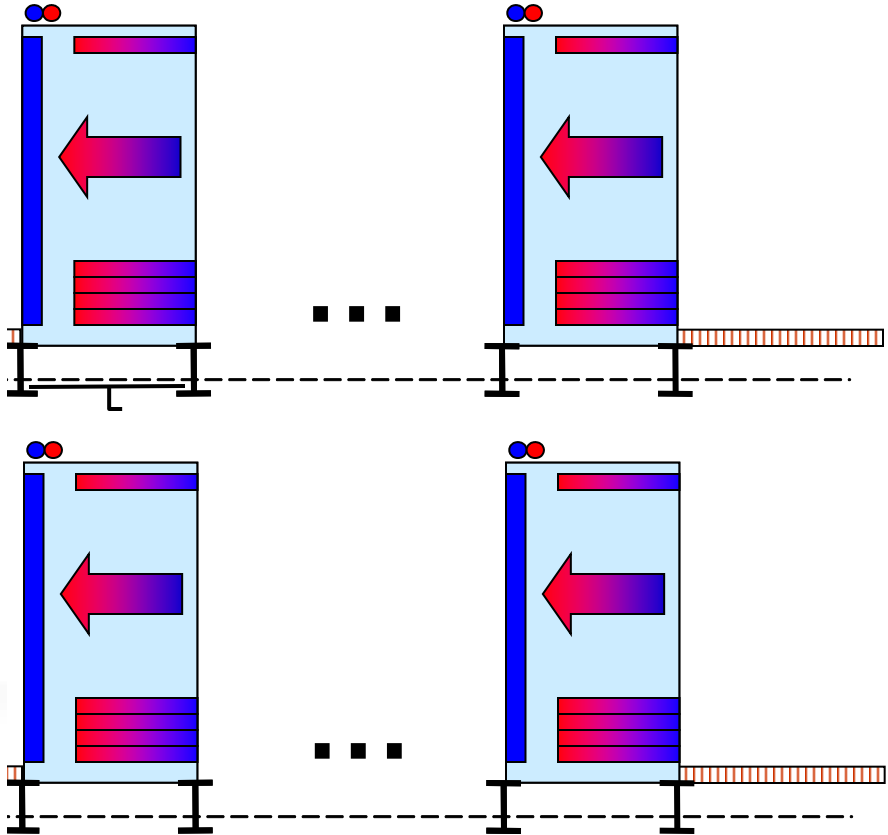
GSI Datacenter Roadmap

- **To date: all resources exhausted**
- **10 19“-rack container hosting Lustre upgrade**
- **100 rack data center (1,5 MW) in Testing hall (completion September 2011)**
- **Begin construction of GSI Fair-IT Cube 2012**

FAIR Green-IT Tier-0 Data Center



mbH
DGI Bauwerk
Gesellschaft von Architekten



**FIAS Patent submitted @ DPMA 5.7.2008
(Lindenstruth, Stöcker)**

FAIR Centers outside GSI

- **KACST Prototype of 10 PF system (2M€)**
- **FRRC HPC system 1M€/a**
- **Acquisition of new computer in 2011 (3M€)**
- **FAIR Tier-0 proposal contains 8 M€ for computers in 2014**

Usage of FAIR Tier-0 center

- **Candidate for possible CERN Tier-0 colocation center**
- **Center for all GSI IT and FAIR IT**
- **Huge upgrade capability (space and power)**
- **Also hosting on-line systems (CBM FLES)**
- **How about PANDA?**
- **Additional requirements?**

Core – IT (Karin Miers)

- **Exchange2010 (this summer)**
- **Windows7/Office2010 (this year)**
- **Sharepoint Introduction (medium-term)**
- **Network design:**
 - **Connection to testing hall (short term)**
 - **Layout for the near future (new technologies, core routing...)**
- **New concepts for mass storage / backup (long-term):**
 - **Lustre, combined hard discs/tape solution...**

Administrative (Organization and) **IT** (Karin Miers)

- **New content management system for web site – typo3 (work in progress, completion hopefully in autumn)**
- **New solution for document management system (mid-term, next year)**

Questions / Input:

- Requirements network?
- Requirements storage and gStore?
- Feature wish list for DMS
- Demand for Sharepoint?
- Service for FAIR:
 - Web site hosting
 - FAIR e-mail addresses

Focus of Scientific Computing (Peter Malzacher):

3 Subgroups: Experiment Simulation and Analysis, e-Infrastructure, DBs

Development and maintenance of the common software framework of the FAIR experiments (FairRoot)

Support of the experiments:

using FairRoot,

in the construction and use of databases,

in software engineering methods and tools,

to adapt to new parallel architectures

and distributed analysis

Participation in national and international research - and development activities for software of the experimental and theoretical particle - and nuclear physics. Community Software Support at GSI / FAIR

Support the software environment of the ALICE experiment at the German Tier-1 center GridKa at KIT and the Tier-2 center at GSI, LOEWE CSC, ...

Strategic planning and design of the e - Infrastructure for the FAIR experiments, development of models for the resource integration of regional universities, national and international FAIR partner with methods of the Grid and Cloud Computing

Short/medium term challenges:

ALICE T2/3 support:

local vs Grid jobs GSI pledged ~900 jobs in parallel for ALICE (since 1. of April)

SE GSI pledged 440TB we provide 300TB: larger SE or SE on lustre

10GB/s to GridKa or into HEPPI

DB Strategy for FAIR experiments:

Oracle is quite expensive, exploration of alternatives (open source SQL or noSQL DBs)

FairRoot

Multi/Many core GPU strategy, integration of OpenCL, new experiments

Community Software Support

Close collaboration with ROOT, AliEn, coordination with G4, Geant5 project, PoD & PROOF

Grid & Cloud for FAIR

PandaGrid, Cloudification of LOEWE CSC, Ffm Cloud, ...

High Performance Computing

Walter Schön, GSI



Walter Schön, GSI



Lustre:

**Additional installation of 1.2 Petabyte
(next days) => 2.5 Pbyte (netto)**

Compute nodes:

2.000 additionalas Cores in compute farm:

SM 2U Twin with 64 Cores (“poor mans blade“)

2GB/Core, 1 Core < 100 Euro

(setup at BG2, Ethernet, Infiniband ready)

Prototype Software-distribution with modules

=> Scalability! → no /u filesystems on farm

GE (openSource) replaces LSF (high license cost)



- „test cube“
- New cooling system
(PUE <1.1)
Production operation with
Lustre

=> „mini cube“ in
Testinghal, September 2011



Cluster for „Mini Cube“: IB based

1 Petabyte Lustre II (Start Version)

=> New tender

**A : „cheap“ solution: SM
+ fast, inexpensive**

- Simple RAID controller

B: „high-end solution“

+ redundant RAID controller,

+ hardware data integrity

- Expensive (> 1.5 more)



**Compute nodes for
„Mini Cube“**

Recommendation:

**2UTwin with FLOP/Euro
Optimized power**

Migration Lustrel → Lustrell

- Without service interruption
- Seperate data sets / compute cluster
- in coordination with experiments stepwise copying of data and subsequent moving of hardware to test cube
 - ⇒ Transfer of (IB) Hardware To Testcube (without service interruption)
 - Connection of remaining nodes via LNET router to new Lustre

