

## Data Intensive High Performance Computing -Challenges for the Future

Dieter Kranzlmüller kranzlmueller@gup.jku.at

GUP, Joh. Kepler Univ. Linz, Austria & EGEE, CERN, Switzerland



### Still long way to go!!

To digitize the information of the human body one would need several trillions of the fastest hard discs that exist today...



## Questions

- 1. For the next several years, assuming an increasing demand for storage capacity and speed, but no change in the storage device availability and structure, what should be the areas for system research? Assuming that the "preservation" trend continues and its challenges can be met by a new function in storage then how could the HPC community benefit from it?
- 2. Object Storage has for a long time been the "love-kid" of this community. As time passes and the component economics change, how should Object Storage evolve?: To a completely programmable device? Or partly programmable device? And, then what? Can it satisfy other markets?
- 3. Programming models have a very basic support for storage: the best example is, that with all the widespread usage of search programming, ingest of data is still completely tied to precisely defined data set. Is this good enough? How can we bridge the gap between search and data usage by programs?
- 4. As the storage usage volume moves away from computing, do we need to look again beyond commodity, to some forms of storage that are "specialized for HPC"? Should this be an industry effort? (Is there enough market?) Or should this be a government(s) sponsored effort?





## **CGCC** EGEE – Enabling Grids for E-SciencE

Country participating in EGEE

## Enabling Grids for E-science > 60

#### **EGEE Grid Infrastructure**

> 180 sites (24x7) in 39 countries
> 10 000 concurrent jobs per day
> 60 Virtual Organizations (20 App. Domains)

GLite



# My "crazy" Predictions



- Demand for storage capacity and speed will continue to increase
- "Preservation" trend will continue
- Complexity **will continue** to increase
- Uncertainty **will continue** to become more important
- Reliability **will continue** to be more and more important







## What Needs to be Done?



- Hierarchy, Abstraction, Layering of Architecture
- Intelligent, Automatic and Autonomous Behaviour of (Specialized) Components
- Interconnecting Loosly-Coupled Components (Protocols, ...)
- Focus on User Needs Provide Tools (Programming Models, APIs)





### How would HPC Community Benefit?

Lot's of interesting work ahead!

## Most important:

Collaboration is needed to solve these problems

SOS10 has been a Great Event!



D. Kranzlmüller, GUP

SOS 10, Maui, HI

Renaissance Wailea Beach Resort, Maui, Hawaii, USA, March 6-9, 2006 20

Sandia National Laboratories 
Oak Ridge National Laboratory 
Swiss Institute of Technology
10th Workshop on Distributed Supercomputing

So Long, and Thanks for All the Fish ...