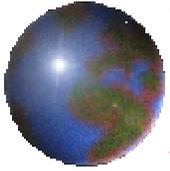


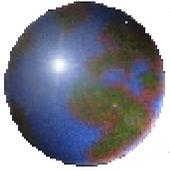
*D0 Grid plans for D0
Distributed (Remote)
Computing*

I. Terekhov



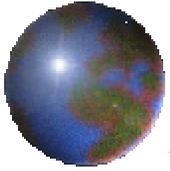
SAM and D0 Grid projects

- ⊕ These overlap (SAM is de facto a data grid by its design, people – Lee, Igor, Sinisa) but are different
- ⊕ Project type and members
 - ⊞ SAM is large, mature, strictly CD staff with some D0 (and CDF) participation
 - ⊞ D0 Grid is emerging, has newly funded PPDG component, has active offsite D0 collaborators (you)
- ⊕ Foci
 - ⊞ SAM focuses on central issues, from Enstore to FNAL security issues
 - ⊞ D0Grid is strictly distributed (and remote) issues



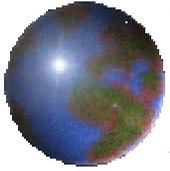
D0 Grid plans

- ✚ Adopt the standard Grid technologies, to “grid-enable” SAM and D0
 - ▣ Inter-operability with other Grids
- ✚ Job management
 - ▣ Logical: specification, description, structuring (MC, chained processing)
 - ▣ Physical: scheduling and Resource Management
- ✚ Monitoring (and Information Services)
 - ▣ Individual job status
 - ▣ Grid system as a whole (what’s going on?)



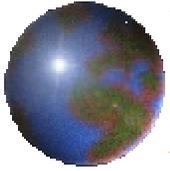
D0Grid Deliverables for Analysis

- ✦ Job Description Language (6 months)
 - ▣ Describe application in a portable way
 - ▣ Multiple stages chained (DAG)
 - ▣ Parallelism
 - ▣ Handling of output (store, define a dataset, etc)



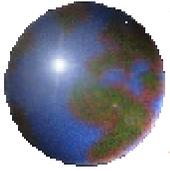
D0Grid Deliverables for Analysis

- ✿ Submit and execute job (months)
 - ❑ Sam submit (0) -> d0grid submit(6) -> Grid submit(24)
 - ❑ System will co-locate jobs and data
 - ❑ Execution at an island site (network disconnect from FNAL/the rest of D0 grid) (9 months)
 - ❑ Whole job brokering (system selects based on data availability, resource usage etc)
 - Advisory service only (6 months)
 - Automated global dispatch (station selection – 9-12 months)
 - ❑ Decomposition and global distribution of job parts for ultimate throughput (24 months)



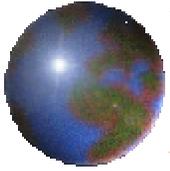
D0Grid Deliverables for Analysis

- ❖ Information Services (SAM internal metadata) decentralized – 9 months
- ❖ Monitor jobs and the whole system
 - ❑ Submitted->pending->done->etc progress (6m?)
 - ❑ Where it was dispatched, remote monitoring (6m?)
 - ❑ Historic mining (what did I do 2 weeks ago?), including that at an island – 9-12 months
 - ❑ Monitor system performance and resource usage (24 months)



How YOU can contribute

- ✿ Use cases, use cases, use cases
 - ▣ How you want to define your analysis job
 - ▣ How you want it executed (local/global, parallel, etc. Can you handle distributed job output? Can you combine output? Can you checkpoint your analysis program?)
 - ▣ How you want to see what's going on with the job or with the (sub)system (site)
- ✿ Hands-on contributions – see the plan



Summary

- ❖ D0 Grid aims at enabling globally distributed computing and analysis, on top of SAM as the data delivery system
- ❖ Analysis on the Grid is a priority
- ❖ Will need your input at earlier stages