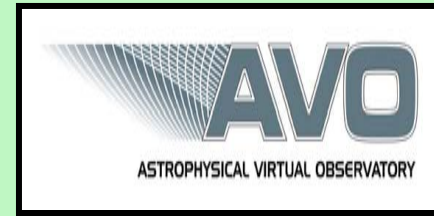
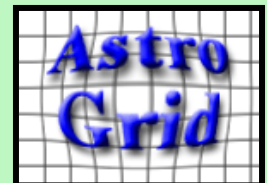
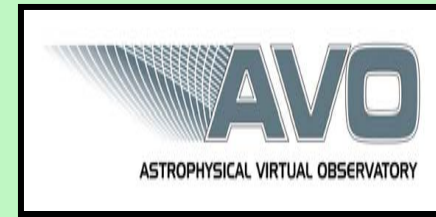


WA3 report

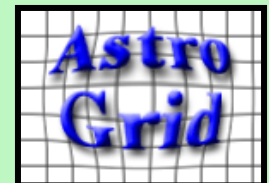


- Personnel
- Milestone progress
- AstroGrid status
- DBMS and Middleware reports
- VOTC plans
- Goals this semester

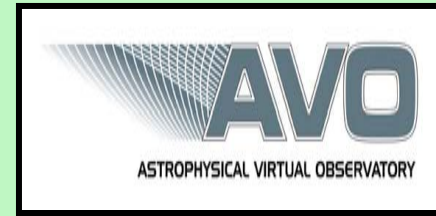




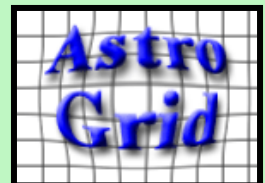
- led by AstroGrid
 - plus effort from TeraPix, CDS, ESO
- current staffing : 6.6 FTEs
 - 3.55 EU funded, 3.05 partner funded
- three developers 100% paid on WA3
 - Martin Hill, John Taylor, Clive Davenhall
- two scientists paid part-time on WA3
 - Richards, Tissier
- several more AstroGrid, CDS, ESO staff



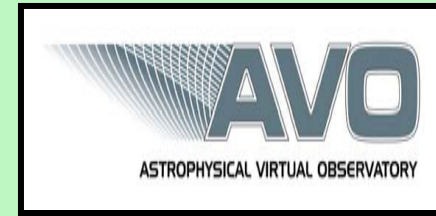
milestones 2002



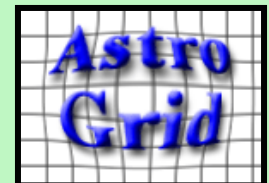
- Nov : deploy trial hardware
 - ABANDONED : under resourced
- Nov : benchmark network tasks
 - DONE : report by Rixon Dec 2002
- Nov : benchmark datamining tasks
 - DONE : part of AstroGrid Phase A report.



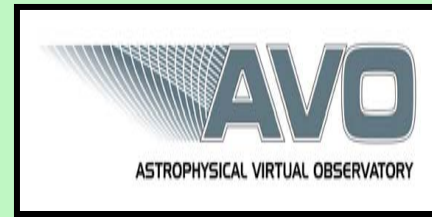
milestones 2003



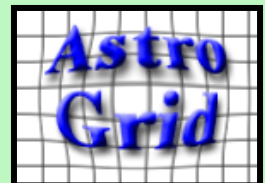
- May : DBMS assessment report and reccs
 - DONE Jan 2004 : report by Clive Page
- Nov : Middleware assessment report and reccs
 - DONE Jan 2004 : report by Guy Rixon
- Nov : demonstration of working datagrid
 - DONE : achieved with other WAs in January 2004 demo
- Nov : working datamining system for use by external users
 - DONE : achieved with other WAs in January 2004 demo



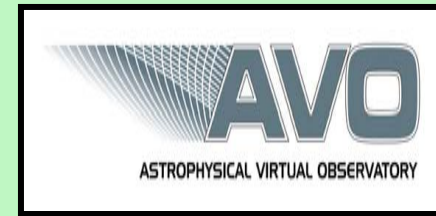
milestones 2003 contd



- Nov : network analysis report
 - MOVED BACK : no spare resource until Davenhall hire
 - now running tests across UK , across Europe shortly
 - issues : TCP buffer tuning; GridFTP speed up; route tracing across partners; bottleneck identification; disc I/O
 - aimed at real practical recommendations
 - target report by April
- Nov : storage assessment report and reccs
 - MOVED BACK : Wicenec collecting more info from partners
 - target draft report January; final report February



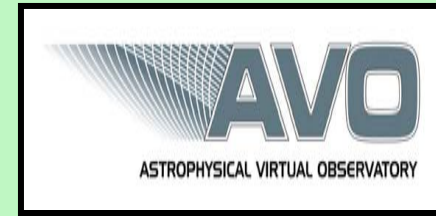
milestones 2004



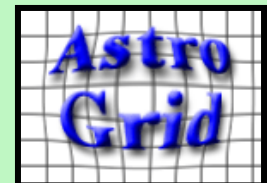
- May : Architecture Assessment Report
 - not a detailed architecture : general principles
 - need a real implementation plan for initial Euro-VO
 - target July 2004



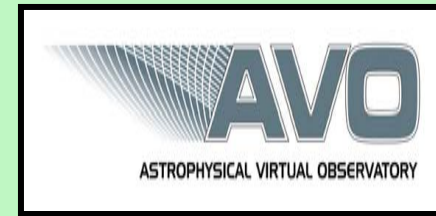
AstroGrid status



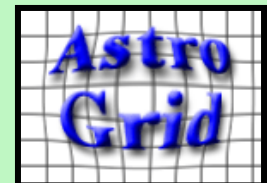
- 4/8 cycles complete
- working infrastructure components
 - Registry, MySpace, Data Access, Portal, Workflow
 - Community soon
 - intention is this is all available to AVO
- continued research
 - grid middleware ; experimental grid services
 - DBMS and datamining tests
 - aimed largely at AVO deliverables
- AG2 proposal 2005-2007 funded
 - more infrastructure; ontology; agents; datamining; visualisation



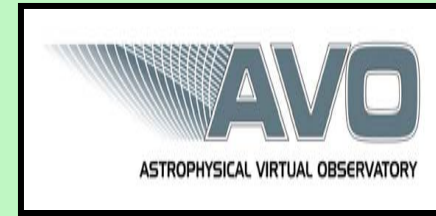
Middleware report



- tested variety of grid middleware
 - OGSI, OGSA-DAI, LDAS, GridFTP, SRB, GSI/CAS
- current m'ware : industry standard web services
 - XML, SOAP, WSDL, HTTP
 - problem : synchronous, stateless, blocks channel, no authentication/
- solution-1 : Globus+WS = OGSA
- solution-2 : WS-CTX

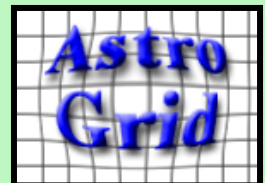


OGSI vs WS-CTX

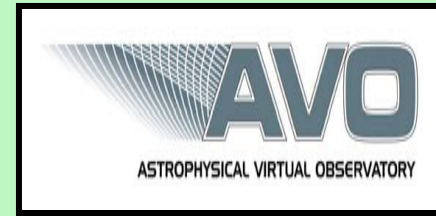


"OGSI-compliant services are SOAP services that distinguish contexts by assigning a new URL to each context; they require that the application software set up this URL explicitly using a factory service. OGSI complicates the semantics of all grid services but leaves the syntax of the invocations simple."

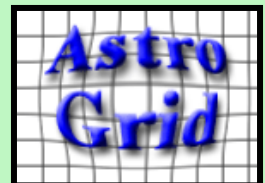
"WS-CTX-compliant services are SOAP services that distinguish contexts by notations in the headers of their messages. They complicate the syntax of messages in order to keep the semantics of the service simple."



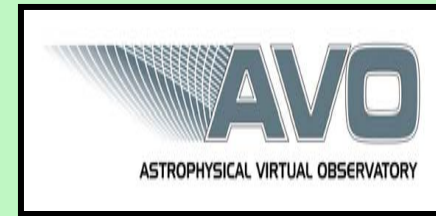
some recommendations



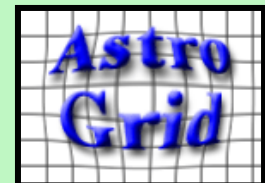
- Don't re-engineer with OGSi : prefer WS-CTX.
- Use GridFTP as the main transport for files in the data grid.
- Do not use SRB as the basis of the data intergrid. It is OK in intragrids.
- Use OGSA-DAI (or ELDAS) to move relational tables.
- Use GSI-compatible X.509 certificates to identify users.
(Portals can still use passwords at the system boundary.)
- IVOA discussions on interoperation of identity certificates
- Until users have certificates, certify software agents to work on their behalf.



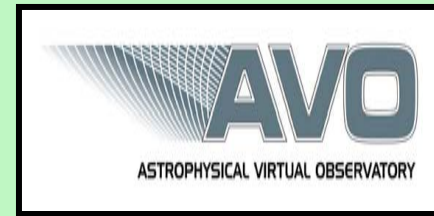
DBMS report



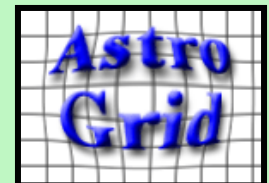
- examined many different DBMS
- quantitative evaluation of four
 - SyBase ASE, Oracle, SQL Server, DB2, MySQL, and PostgreSQL
 - no single clear winner
- also looked at alternatives
 - flat-file systems, XML databases
- and investigated wider issues
 - query language, data mining packages, cross-matching, indexing



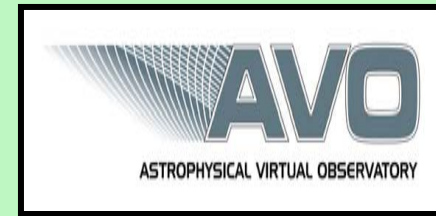
some recommendations



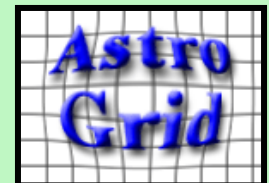
- Many xmatch problems involve big one small DB; move latter to former
- Standard DBMS/SQL problems : coping with large datasets; losing metadata; inadequate mathematical, statistical, and graphical facilities; 1-D indexing
- so.. develop home grown ADQL and 2-D indexing
- Column-oriented storage has advantages for data-intensive querying.
- PostgreSQL had good all-round score :
reliable, well-supported, powerful, easy to install and use, cheap.



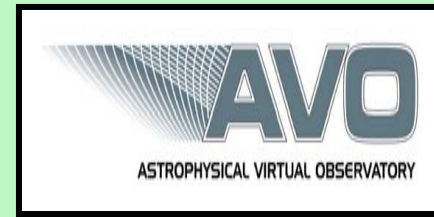
the VOTC



- Euro-VO plan calls for creation of VOTC
- WA3 work segues smoothly into this
- proposal planned for March 2004
- 7 partners each approx 2+2 FTEs
 - Cambridge, CDS, Edinburgh, ESO, (IAP), Leicester, INAF
 - 3 year, 5Me project
- focused on downstream technology needs
 - new infrastructure; user tools; automated resource discovery; data mining and visualisation



goals for next two quarters



- **Jan** : WA3 component of annual report
- **Feb** : hardware assessment report
- **March** : VOTC proposal
- **April** : network analysis report
- **July** : architecture report
- **November** : final report

