AstroGrid Project
AstroGrid-3 Deployment Plan

Authors : Keith Noddle, Andy Lawrence, Nic Walton
(Almost Final) Draft : 2007-12-12

(1) Introduction

This is the deployment plan for the rollout of the *first version* of the operational VO service based on the AstroGrid software infrastructure, VO services run by AstroGrid consortium members, and key UK data holdings. It was endorsed in preliminary form by the AstroGrid Oversight Committee on October 16th 2007. This version is considered the complete initial plan; more detailed planning continues on the project deployment wiki.

The original AstroGrid project began in late 2001, with a second phase "AstroGrid-2" project beginning in 2005. In late 2006 a proposal was made to PPARC to fund an *Operational VO Service for the UK* using AstroGrid software. This was funded for an initial period of two years, to begin *deployment and operations*, with a further review on long term funding expected during 2008-9. This initial two-year period is informally known as "AstroGrid-3".

This Plan consists of several parts :

(1) Introduction
(2) Context and overview
(3) Schedule
(4) Components to be included
(5) Hardware plan
(6) Interactions with data centres
(7) Marketing Plan
(8) Forward Look

(2) Context and overview : VODesktop

Now is the right time to make a definitive release of AstroGrid software and services. (i) The IVOA standards on which all of the VO is based are now (just) mature enough, and there is a large enough set of these to deploy a meaningfully complete set of compliant services. (ii) A large enough number of operating data services is now available to make the VO scientifically useful. (iii) Our core infrastructure software components are now all in place. (iv) During the last year we have made substantial advances in the robustness and speed performance of core VO services based on our software, so that an operating system is reliable and quick enough not to annoy users. (v) The niggling issues of the user interface being somehow not quite right have now been essentially solved, with significant redesigns of all interface components, and we are confident that we have something that will be popular. (vi) A significant suite of third party applications has developed which make use of VO standards in general, and AstroGrid middleware in particular.

A number of "AstroGrid releases" have been made in the past, and promoted at both conferences and roadshow style workshops. These have been crucial learning experiences both for the project, and for the subset of the UK community which has tried us out, and we have real weekly if not daily use of the AstroGrid workbench. We now need a distinct re-branding so that the world knows "this is really it now". However, we need to do so in a way that doesn't lose existing AstroGrid branding.

Although the point of our software is accessing remote services, our intended release is heavily

focused on client side middleware, and third party tools which also run on the client, and which will be installed programmes on a user's own machine. We will refer to this whole package as "VODesktop". This will be marketed as a new product, but from AstroGrid. This will allow later releases with a distinct flavour, such as an enhanced server side product.

From the *user's point of view* we have :

- Standard looking desktop software ...
- ...but which calls remote services
- The toolset has an *AstroGrid core* and a suite of *interoperable tools*

But *behind the scenes*, we have

- Client side middleware (AR and PLASTIC)
- Core AG services (Registry, VOSpace, Community)
- Consortium data services run through DSA
- Rest of world's data services
- Remote applications with a CEA wrapper

(3) Schedule

This is the initial schedule

- Internal VODesktop testing release	01 Aug 2007
- Alpha release to individuals in other VO projects	01 Oct 2007
- plan sign-off by AGOC	16 Oct 2007
- Beta release to tester community	15 Dec 2007
- Metadata review complete	15 Dec 2007
- Design of new web site and look and feel	15 Dec 2007
- Hardware Procurement Plan	15 Dec 2007
- Draft SLAs	01 Feb 2007
- First draft new documentation complete	15 Jan 2008
- First draft new web site	01 Feb 2008
- Help Desk s/w tested	15 Feb 2008
- New hardware procured	15 Feb 2008
- Installations completed	01 Mar 2008
- SLAs agreed	01 Mar 2008
- Updated web site and documentation	01 Mar 2008
- Critical Service Review	03 Mar 2008
- AGAG feedback	03 Mar 2008
- Final testing and debugging complete	15 Mar 2008
- Release announcement to community	31 Mar 2008
- Demo at NAM	31 Mar 2008
- Workshop roadshow begins	01 May 2008

(4) Components to be included

4.1 Infrastructure software components

The following software components and versions are to be included and so completed, debugged, and tested for performance.

- Registry
- VOSpace
- Community,
- Astro Runtime
- Plastic Hub
- CEA
- DSA and sub-parts

4.2 Core VO Services

The following core services will be deployed at the following sites :

To be completed

Bristol ?
Cambridge : Community; Registry; VOSpace;
Edinburgh :
Leicester :
Jodrell Bank : Community; Registry; VOSpace; Application Server
MSSL :
RAL :

Note : DSA services in front of UK datasets are listed separately

4.3 User interface components

From the user's point of view, the following components will be available :

- VOExplorer
- File Explorer
- Task Runner
- Query Builder
- Astroscope on subset
- Python Scripting facility
- Taverna with AG plugins

In addition, if time/IVOA standards allow, we may make available :

- dedicated Helio tool
- improved astroscope-like tool (needs footprint services)
- enhanced QB (views, straight SQL, user-defined fns)
- commissioned image viewer
- VOlookout replacement
- AR C-binding allowing IDL scripting

4.4 Third party software

The following third party tools will be made available in some form as part of the VODesktop package. They will either be bundled in an installation package or linked on a web page for easy download and startup.

- Topcat
- Splat-VO
- Aladin Prototype version
- VOspec

The following will also be listed as recommended compliant tools, with links

- Visivo
- Astroweka
- VOplot
- Astroneural
- Paperscope alpha

In addition, if time/developments allow, we will include

- a new image viewer (Son of Gaia ?)
- the official release version of Aladin, if it has VOSpace facility in time

4.5 Remote applications

The aim here is to have key parameter driven astronomical applications available as remote services, runnable with the Task Runner. AstroGrid is not developing new applications; rather it places a CEA wrapper around existing applications, and hosts them on a application server. (We have also however developed workflows which are also offered as parameterised tasks.) CEA-wrapped remote applications will grow gradually with time, and will be discoverable through the Registry using VOExplorer. We will maintain several application servers (as listed above) and so each application will be available in several places. The deployment wiki will keep a list of target applications to wrap. Some of the most important ones are :

- STILTS
- Sextractor
- HyperZ
- Galexev
- MERLIN imager
- Solar Movie maker
- Pegase
- Schlegel dust maps
- Starburst99

In many cases, the descriptive information in the Registry, and so visible with VOExplorer, is woefully inadequate. We will pick key applications under consortium control and work to make these more useable.

Note-1 : When the UWS protocol is agreed by the IVOA, it should become easier for third parties to run application services that are discoverable and useable by the VODesktop system.

Note-2 : until the TAP protocol is established, queryable datasets (eg SDSS, UKIDSS, 2MASS) are made available in VODesktop as CEA applications.

4.6 Documentation etc

Documentation will be complete and available at the point of release. Several types of documentation are needed :

- public information about the project
- end user documentation for astronomers
- Helpdesk system
- technical documentation for VO projects
- technical documentation for data centres deployers
- technical documentation for third party developers
- guide to software installation for various user types

(5) Hardware plan for AstroGrid core services

5.1 Current hardware deployment

Hardware provision is simpler than anticipated a year ago because the deployment of Virtual Machine (VM) technology is proving very successful.

Note : our proposal assumption was per site 2 dual core CPUs (one for general services, one app server) and 10B storage, estimated at 10K/site. At the recent planning meeting, we agreed that actually we should tailor provision at each site more flexibly; and that we should deliberately hold some funds back until we see the scale of uptake post April 2008.

Edinburgh : current provision
3 single processor 32bit AMD Opteron servers (2 GB RAM each)
2 dual processor 64bit AMD Opteron servers (4 GB RAM each)
Running services :

- Registry
- DSAs
- MySpace
- Community
- Prototype services + development environments
- Wiki, forums, web sites etc

Edinburgh : planned new provision
1 dual processor 64bit AMD Opteron server (8 GB RAM)
1 SAS (Serial Attached SCSI) RAID Storage ~ 2-4 TB of disk space

- VOSpace
- CEA/UWS applications e.g. STILTS etc

Leicester : current provision
dualXeon server, Linux, running :

- filestore
- proxy for registry (for continuity of name)

3 x dualXeon server, Linux VM host running Xen, RAID storage attached
The above host the following VMs in some combination between them, with the failover mechanism to switch from one VM host to another not yet automated:

- Development community
- AstroGrid registry
- Filestore
- AstroGrid community
- AstroGrid JES
- AstroGrid filemanager

1 x dualXeon server, Linux VM host running VMWare for

- Wiki
- cvs mirror - anonymous read only access
- development machine
- wiki upgrade development

1 x Xeon server, Linux VM host running VMWare:

- webpages, http proxy, cvs
- jabber server
- mail server/ email lists
- chunking proxy
- CEA development

1 x dualXeon
server, Linux VM host running VMWare:

- development
- ● plone
- ● remote access gateway

1 x single Xeon server, Linux
1 x dualXeon server, Linux

- development machines

Leicester : planned new provision
dualXeon servers + large SAS/NAS to host

- VOSpace and
- mirrored CEA/UWS applications

Cambridge : current provision
an update has been received from Cambridge which will be inserted

6 servers of various types (SunBlade, Dual Xeon etc), running

- Registry
- DSA
- MySpace
- Community
- CEA services + development environments

Cambridge : planned new provision
Dual Opteron server; Solaris, running

- Community
- VOSpace

Quote from Keith's C6 status report mid-cycle :

"In addition we have Registry DSA services deployed at MSSL along with Solar CEA applications as well as DSA and CEA applications in Jodrell Bank. It is expected that over time, most data centres will want to host VOSpace services of some kind, for result staging if nothing else. Community and local publishing Registries will follow suit but only as local sysAdmin feels rather than is told about the benefits"

Provision and decisions at remaining sites to be completed soon.

MSSL : current provision
xxxx

MSSL : planned new provision
xxxx

RAL : current provision
xxxx

RAL : planned new provision
xxxx

Jodrell Bank : current provision
xxxx

Jodrell Bank : planned new provision
xxxx

5.3 Performance Tuning

Extensive performance testing and tuning has been carried out over the last year, and will be completed during the deployment phase. This will include :

- debugging and optimising infrastructure software
- installation of multiple VMs
- arrangements for failover and load balancing between duplicate services
- stress testing and bottleneck identification
- scaleability tests
- bandwidth testing and tuning between centres

(6) Interactions with data centres

The aim here is to make sure that the initial VODesktop release gives access to as much data as possible, but especially to datasets seen by UK astronomers as being the most important; and to see that the metadata describing the data is as complete and accurate as possible, thus maximising its usefulness.

6.1 Key UK holdings and associated data services

The first priority is data services and applications services under the control of the consortium partners. The following tasks will be completed during deployment

- make a list of key data sets under control of consortium
- specify the type of service for each, e.g. DSA, SIA, CEA-ADQL etc
- work with data centre staff directly to install each such data service
- work with data centre staff directly to publish the data service in the Registry
- make a list of applications to be hosted and CEA-wrapped at each centre
- work with data centre staff directly to set up application server

The next priority is UK data services outwith consortium :

- list target datasets and associated services
- provide information, and assistance where possible, to install and publish data services

6.2 International data holdings

In this initial deployment we do not have time to work directly with data centres worldwide. VO standards are becoming widespread, so we will usually be able to consume their services, but we will have no influence over their quality. However, we will make the DSA software available and documented, and will encourage third parties to publish data services using DSA. Some key international colleagues - eg ESAC and CADC - are likely to adopt DSA, and we will provide assistance if we get time.

A fuller campaign of working with data centres worldwide is intended for 2009.

6.3 Registry metadata campaign

Early use of VOExplorer has revealed a number of problems with the metadata in registry entries - missing fields, unhelpful titles, inadequate descriptions, confusion between fields such as Curator vs Publisher vs Authority, and so on. The problem of inadequate explanation applies to CEA applications as well as data service entries. These metadata entries are generally not under our control, except for data services published by consortium partners, where we must make sure the metadata are of the highest quality. For data services originating outwith the consortium, is not acceptable for us to edit the entries in our own registry; however what we can do is attach annotations.

The following tasks will be completed :

- make list of key data services and CEA services under our control
- find problems with these
- amend entries in co-operation with authors

- make list of top priority target services elsewhere
- find problems with these
- produce annotations as appropriate

- analyse typical problems with registry entries
- write "good practice guide" for data publishers

6.4 AstroGrid linkages through data centre web pages

Astronomers will often start at the web pages for their favourite dataset. For our own data centres, AstroGrid functionality must be easily available to them from this starting point. Initially this means information and links, but in time it will mean embedded functionality. Tasks to complete are :

- install core services (Registry, VOSpace, Community)
- install application services
- install DSA in front of key datasets
- provide clear AstroGrid logo and link to AstroGrid website(s)
- provide brief (one para) information about AstroGrid

As and when it becomes technically sensible, we will also provide

- ability to save generated files etc to VOSpace directly from data centre web page

- ability to launch Plastic apps eg TopCat directly from data centre web page
- ability to launch VODesktop (i.e. AR) directly from data centre web page

(7) Marketing Plan

7.1 Web page

We need a much improved web presence. This is under design at the time of writing and will begin implementation soon. The general goals are that the web pages should

- be professional looking and easy to follow
- provide a launch point for user software installation
- be easy for team members to update
- separate end-users (astronomers) cleanly from deployers and developers
- provide an integrated help desk system
- provide documentation of several distinct types
  - general public information
  - astronomer documentation for using the tools
  - technical documentation for deployers
  - technical documentation for developers

A much more informal project workspace will be maintained through the existing wiki system.

7.2 Announcements and demos

This will involve a sequence of key tasks :

- initial demonstrations to AGOC and AGSAG
- definition of the beta testing community and feedback methods
- release to the beta testers with guidance
- critical service review with AGOC and AGAG
- demo stall at NAM and associated announcement
- planned seminar series by AstroGrid scientists
- dissemination at ADASS, IVOA, VO-SAC

7.3 Departmental Deployment

Initial AstroGrid users can register with one of the communities provided by the six consortium sites. Users will need an up to date JRE, and we will make sure that Departmental sys admins are aware of this. Sys admins may also need to be aware of where cache and profile information etc is kept, to help with occasional user problems. Overall the initial demand on local systems is very light. However our intention is that very soon it will be normal practice for any University Astronomy Department to run a Community Server, and to integrate this with its own local systems, so that when users log on in the morning, they are automatically logged on to AstroGrid, and their data access rights are already clear. Tasks needed are :

- ensure that documentation and help desk system for users minimises burden on local sys admin
- make sure that deployer documentation provides advice for sys admins
- hold meeting for sys admins to brief them on impact of AstroGrid

Then in due course

- provide documentation for Departments to install Community
- make selected deployment visits to provide assistance
- provide advice for Departments wanting to provide other services
  - eg data services, local registry, application server

7.4 Workshops

Following the April release, we will plan a series of workshops. A detailed schedule will be made during early 2008, but it is likely to include L

- two special topic end user workshops
- two resource provider workshops
- deployment visits to user departments (see above)

(8) Forward Look

The April 2008 release of VODesktop is deliberately limited in scope. Enhancements we anticipate in due course include :

- adoption of improved IVOA standards, especially TAP
- enhanced versions of infrastructure components
- IDL scripting ability
- running tools from data centre web pages
- web applications
- server side workflow
- planetarium front end (Gsky / Stellarium / Virgo).
- server side middleware
- more DSA components and publication tools
- in house image viewer
- all-VO scope
- dedicated Helio-tool
- dedicated crossmatch tool
- proliferation of third party tools