SpaceGRID

C.H. Perry[1], S. Dunkin[1], M. Grande[1], C.D. Pike[1],
B. Stewart[1] and S. Beco[2]

[1] Rutherford Appleton Laboratory, Chilton, Didcot,
Oxfordshire, OX11 0QX, UK, spacegrid@rl.ac.uk

[2] DATAMAT S.p.A., Via Laurentina,
760 - I-00143 Rome, Italy

ABSTRACT

SpaceGRID is an ESA study aimed at identifying the potential benefits of Grids" to the ESA community and defining a road map for the implementation of this technology within ESA.

Grid technology is an emerging computing infrastructure that is intended to provide uniform access to a set of distributed, networked resources that would otherwise be incompatible. Depending on the Grid application, the resources may be large-scale computational systems, data-archives or shared facilities constituting a collaborative working environment.

The SpaceGRID study spans a broad range of domains including Earth Observation, various aspects of spacecraft engineering and Solar System research (solar, STP and planetary). In this presentation we provide an overview of the aims of the SpaceGRID study and report on the requirements analysis for the Solar System research field.

- Grid technology is an emerging computing infrastructure "...*to enable virtual organisations...*"

- It is intended to provide uniform access to a set of distributed networked resources that would otherwise be incompatible.

- Depending on the Grid application, the resources may be large-scale computational systems, data-archives or shared facilities constituting a collaborative working environment.

- A set of standard services provided by software called the "middleware" are used to inter-connect the different resources.

- Solar System Research is a multi-disciplinary science
    - Solar Physics, Solar Terrestrial Physics, Planetary

- Complex 3-D environment
    - Phenomena occur over a range of temporal and spatial scales

- Complex set of instrumentation, data and formats
    - Particles, fields, waves and imagers
    - Scalers, vectors, tensors, images, multi-dimensional arrays

- Data processing is often responsibility of PI
    - Heterogeneous data handling systems

- Researchers need to combine and manipulate multiple data sets
    - This is where a Data Grid and collaborative working environment can help
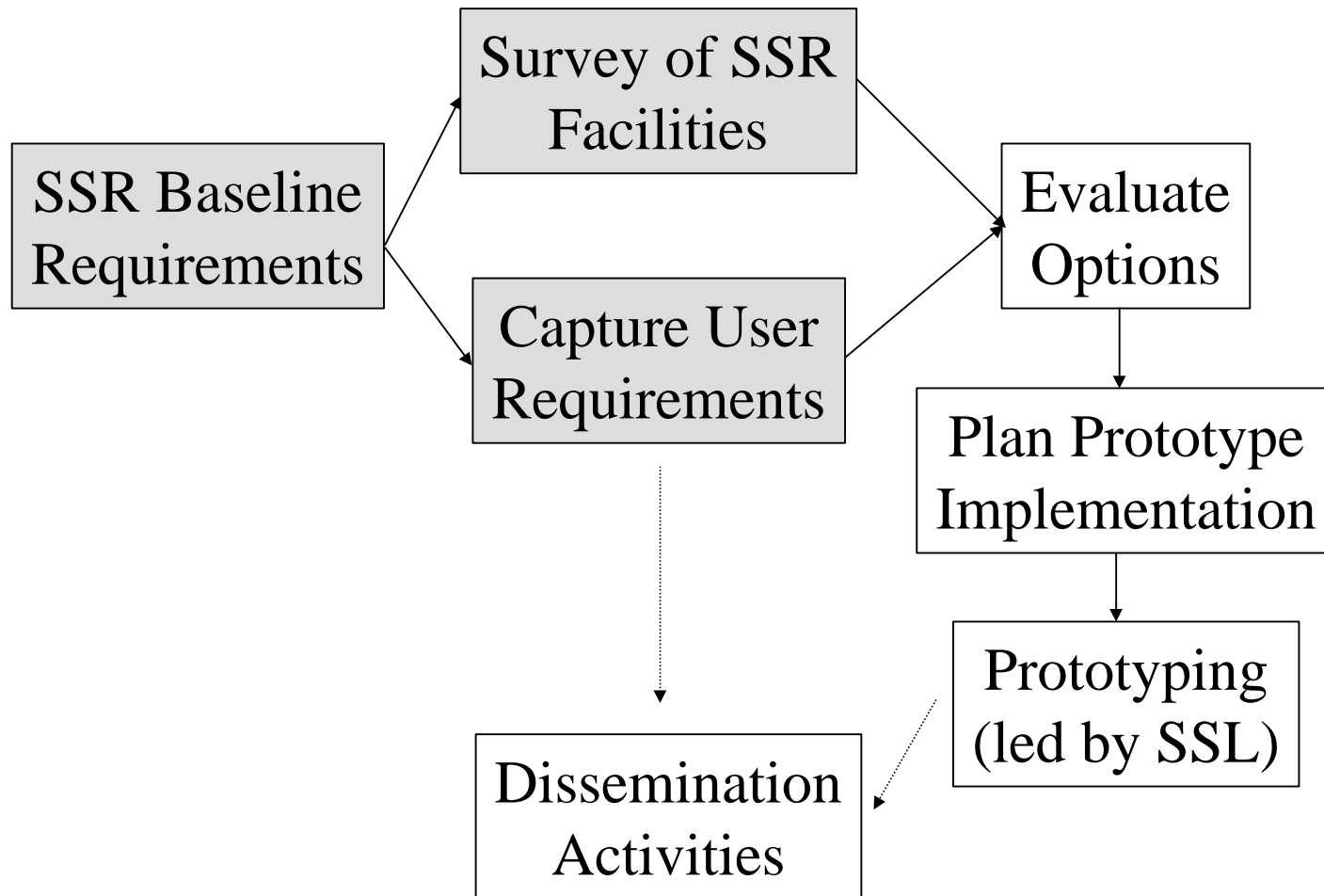
Solar System Research Current Status

- SpaceGRID study started at the end of last year.

- Community consultation and development of User Requirements has been taking place.

- ESA review of User Requirements definition took place at ESRIN 18th April 2002.

- Next step is the development of the system and infrastructure requirements in terms of current GRID middleware technologies such as Globus (http://www.globus.org/) and the Open Grid Service Architecture.

- Several small prototyping projects will start in the summer of 2002 to test the applicability of current GRID technologies to the different domains.

Solar System Research Baseline Requirements

- Facilities for locating online sources of data based on a general query

- Standardisation in the delivery format of data/metadata from different sources

- The ability to query the <u>catalogues</u> of <u>multiple</u> distributed data archives

- The ability to query the <u>data</u> across <u>multiple</u> distributed data archives

- The ability to manipulate and process data remotely prior to download

- A web portal to access distributed resources from a single web site

- A Grid server allowing users to link their own data into SpaceGRID

- A software library allowing programs to access to SpaceGRID facilities

- An online collaborative working environment

Solar System Research Community Consultation 1

- Main community input via online questionnaire.

- 113 responses from 20 countries - Thank You!

- 55% Solar, 36% STP and 9% Planetary

- Broad spread of interests including:

    - planetary interiors & surfaces, planetary and terrestrial atmospheres, ionosphere, aurora, magnetosphere, solar wind, solar terrestrial, interaction, solar flares, corona, solar atmosphere and helioseismology.

- Most people who responded are working on data analysis though significant minorities involved in modelling, operations and software development.

- 60% had not heard or were not sure what GRIDs was and only 5% of the community involved in GRID activities.

- Solar data usage:
  - 70% Space based remote sensing
  - 18% Ground based remote sensing
  - 11% In-situ

- STP data usage:
  - 46% In-situ
  - 34% Ground based
  - 11% Remote sensing

- A few of the difficulties highlighted in data access:-
  - Locating sources of data or models
  - Ascertaining data availability for multiple data sets
  - Limited search capabilities
  - Multiple requests to different archives, interfaces, access control

Solar System Research Community Consultation 3

- Most of the community are involved in collaborations within their area

- Numbers drop to less than a quarter for cross-area collaborations (e.g. Solar and STP).

- Main problems combining data from instrumentation and models:-
  - Lack of standardisation in data specification.
  - Inadequate, incompatible or unreliable metadata.
  - Need for re-sampling temporal or spatial data.
  - Lack of documentation -> incorrect use of available data.

- Three quarters of queries for data are based on specification of a time range…but partly because that is the only option!

- Most common means of collaboration is E-mail. Few respondents mention use of existing, online, concurrent collaboration facilities such as SPARC.

- User prioritisation of baseline requirements
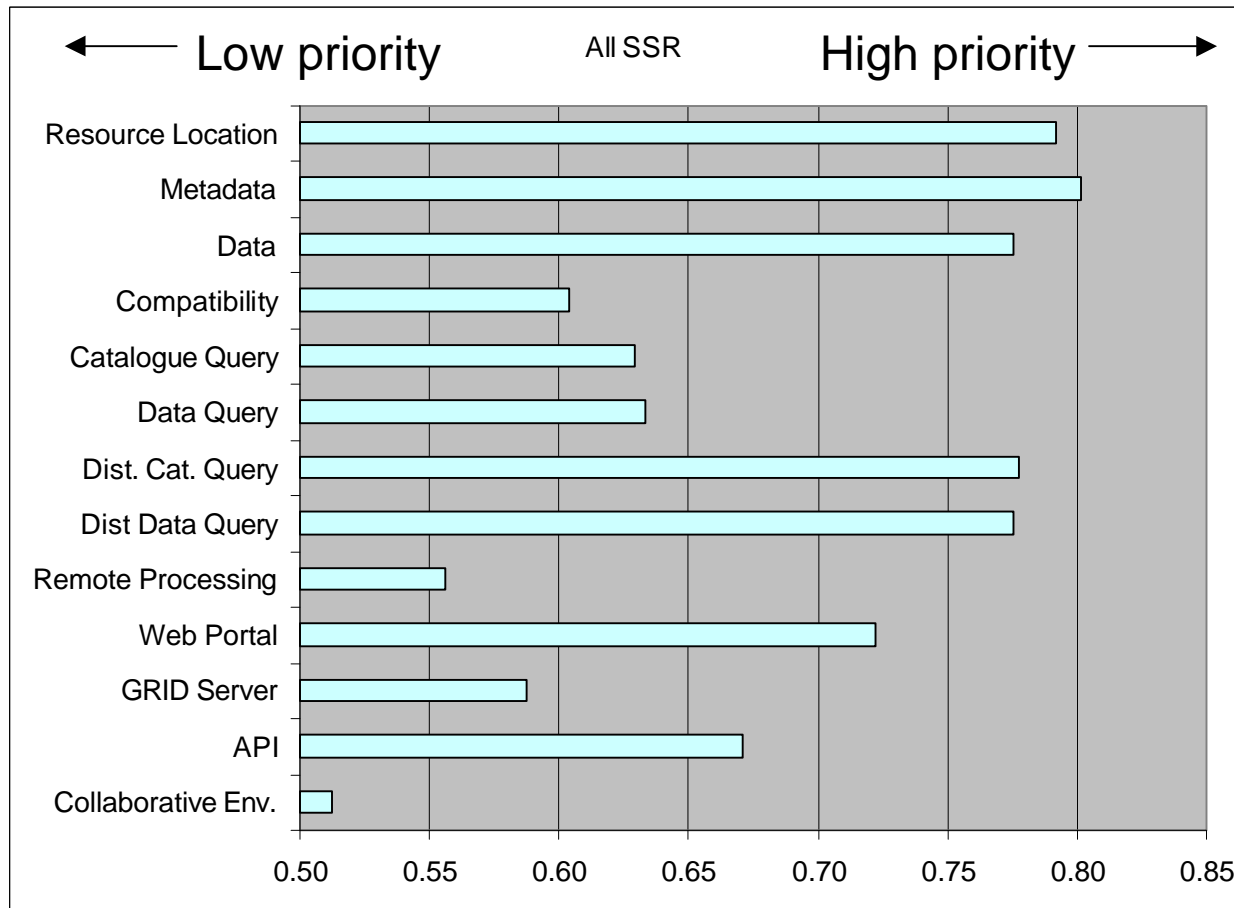
Solar System Research Summary

- GRID technology has the potential of providing an infrastructure to promote efficient use solar system research data.

- Main issues that need to be addressed are:-
  - Improved facilities for locating data held in existing archives
  - Distributed queries and data requests
  - Remote processing of data prior to download
  - Ability to deliver data/metadata in a form that the user desires

- Need to ensure that developments are:-
  - Kept simple; so that facilities are accessible to broad spectrum of the community
  - Easy to upgrade; so that it can take on more functions and abilities as it grows
  - Driven by the needs of the community and not just by what new technologies can provide.

- **Continued input from the community is vital!**
  Visit http://esagrid.esa.int/spacegrid/ for more information.