



EGI-InSPIRE

TECHNOLOGY STUDY FOR CTA VIRTUAL TEAM PROJECT

Scientific Gateway User requirements

Date:	4/06/2013
Document Status:	Approved
Dissemination Level:	Public
Document Link:	https://documents.egi.eu/document/xxxx

Abstract

This document has been written to summarize User Requirements for Scientific Gateway and associated Single Sign On functionality.

I. DOCUMENT LOG

Version	Date	Comment	Author/Partner
0.1	4/02/2013	Creation and collect of existing requirements	N. Neyroud, C. Barbier
0.2	14/02/2013	Translation to Science Gateway requirements	N. Neyroud, C. Barbier
0.3	13/03/2013	First addition from CTA scientists	N. Neyroud, C. Barbier, N. Komin
0.4	15/03/2013	Correction from N. Komin	N. Neyroud, C. Barbier, N. Komin
0.5	27/03/2013	SSO new requirements	N. Neyroud, C. Barbier
0.6	3/04/2013	JP. Lenain and N. Komin add-ons. Taking into account G. Castelli User requirements and G. Sipos remarks.	N. Neyroud, C. Barbier, JP. Lenain, N. Komin
0.7	23/04/2013	J. Knödelseder, C. Boisson, J Cohen-Tanugi, JP Lenain remarks	N. Neyroud, C. Barbier
0.8	2/05/2013	J. Cohen-Tanugi, J. Kocot, M. Sterzel, L.A. Antonelli, G. Tosti, S. Vercellone and G. Castelli remarks	N. Neyroud, C. Barbier
0.9	16/05/2013	T. Gurov and Bulgarian CTA colleagues remarks	N. Neyroud, C. Barbier
1.0	4/06/2013	G. Lamanna remarks and final approval	G. Lamanna, N. Neyroud

II. APPLICATION AREA

This document is an internal report produced by the members of the Technology study for CTA EGI Virtual Team project, run under the EGI-InSPIRE NA2 virtual team framework. Further information is available at https://wiki.egi.eu/wiki/Virtual_team.

III. TERMINOLOGY

A complete project glossary is provided at the following page: <http://www.egi.eu/about/glossary/>.



TABLE OF CONTENTS

1	INTRODUCTION	4
1.1	Scope of the document.....	4
1.2	Purpose of the document.....	4
1.3	Definitions, acronyms and abbreviations	4
1.4	Reference documents	5
1.5	Users.....	5
1.6	Overview.....	6
2	GENERAL DESCRIPTION	6
2.1	Product perspective	7
2.2	Operational environment.....	7
3	SCIENCE GATEWAY SPECIFIC REQUIREMENTS	7
3.1	Science Gateway capability requirements	7
3.1.1	Performance.....	9
3.2	Science Gateway Constraint requirements	9
3.2.1	Communications interfaces	9
3.2.2	Hardware interfaces	9
3.2.3	Software interfaces.....	9
3.2.4	Human-Computer Interface.....	9
3.2.5	Adaptability	9
3.2.6	Availability.....	10
3.2.7	Portability	10
3.2.8	Security.....	10
4	SINGLE SIGN ON SPECIFIC REQUIREMENTS.....	11
4.1	Single Sign On capability requirements	11
4.1.1	Performance.....	11
4.1.2	Security.....	11
4.1.3	Standards.....	12
5	APPENDIX	13
5.1	Summary table for Science Gateway requirements.....	13
5.2	Summary table for Single Sign On requirements.....	14



1 INTRODUCTION

1.1 Scope of the document

This document provides the requirements for a Science Gateway Infrastructure product in the context of CTA experiment. The gateway is the software implementation of a web-based service which collects all “User Interface (UI) implementations and tools for science, telescope monitoring, data access, software repository access, etc”, and will provide a CTA-users common working framework. The “data abstraction system” which is beneath the gateway is the synthesis of a series of services and software applications which respond to the overall CTA operational computing model. The Gateway infrastructure should aim to host the CTA global SAS – Scientific Analysis System – in terms of both UI for scientists and web-services for expert-users/operators dedicated to the data pipeline management.

This document is related to two software products, Science Gateway and associated Single Sign On in the context of CTA experiment.

- Web-based Science Gateways are community-specific set of tools, applications, and data collections that are integrated together via a web portal, providing access to resources and services from a distributed computing infrastructure. These gateways can support a variety of capabilities including workflows handling, virtualization of software and hardware, visualization as well as resource discovery, job execution services, access to data collections, applications and tools for data analysis. A Science Gateway enables community members to define and perform custom research scenarios or other types of use cases.
- Single Sign On solutions simplify access control to multiple related, but independent software systems. Through SSO a user logs in once and gains access to all systems he/she needs to carry out activities without being prompted to log in again at each of them. SSO solutions can be relevant for the human and automated workflows of the Observatory that need to use software, data and resources from multiple systems.

1.2 Purpose of the document

This public version of User Requirements document is strongly linked to the CTA requirements already validated by the CTA consortium; it’s completed with more precise requirements collected using a social network between CTA and EGI through the support teams of the National Grid Initiatives. The objective of this document is to be able to study in a second step the most suitable solutions from EGI and its partners that are capable to address the CTA requirements.

1.3 Definitions, acronyms and abbreviations

CTA	Cherenkov Telescope Array
CTAO	CTA Observatory
DCI	Distributed Computing Infrastructure
EGI	European Grid Infrastructure
ESFRI	European Strategy Forum for Research Infrastructures
HCI	Human Computer Interface
LDAP	Lightweight Directory Access Protocol
MC	Monte Carlo simulation
NGI	National Grid Infrastructure
PI	Principal Investigator
SAS	Scientific Analysis System
SSO	Single Sign On
TOO	Targets Of Opportunity
TBD	To Be Defined
UI	User Interface
VO	Virtual Organization in the EGI Grid context.
VObs	Virtual Observatory

1.4 Reference documents

R1: “ESA software engineering standard issue 2”
http://www.esa.int/TEC/Software_engineering_and_standardisation/

1.5 Users

The Science Gateway users will be members of the scientific community including members of the CTA consortium.

CTA consortium users, being identified by a CTA specific authentication mechanism (LDAP directory today), will be able to have different simultaneous responsibilities in addition to their Science tasks, such as Data Archive management, Science user support, Monte-Carlo production, Database management etc... associated with specific access rights; these users are named “privileged users”.

Part of the other users will be authenticated as members of specific research organization through an International Identity Federation system that will allow not only certification of their identity but also use of the same identification data to obtain access to the Science Gateway. Finally all other users will be able to ask for a specific Science Gateway account or decide to access the public Science Gateway only.

In the context of this document, a user is any entity, either a person or a software component that interacts with the system. The following users have been identified:



The Observatory	"The scientific facility with observation site(s), and additional supporting facilities at additional sites, which delivers the required performance for CTA"
User	"A scientific user of the CTA Observatory (CTAO) who may be a Guest Observer, a Privileged User or an Archive User"
Guest Observer	"A member of the scientific community who is granted access to a specific subset of CTA data, associated with a successful proposal to the Observatory"
Privileged User	"A scientific user of the CTA observatory with access to data and/or software and/or services not available to all Users"
Archive User	"A scientific user of the CTA observatory who makes use of archival data, as opposed to data associated with a specific proposal"
Principal Investigator	"A member of the scientific community named in a proposal to be the contact for all interactions with CTA Observatory"
CTA internal user	"A member of the CTA consortium"
CTA Grid user	"A CTA internal user registered in the CTA Virtual Organisation (VO) of EGI"

1.6 Overview

The following sections have been written using a two steps methodology: first phase using the existing CTA requirements then, in a second phase, using the social network to gather additional user requirements concerning two software products:

- Web based Science Gateways operated for the CTA community, making Distributed resources and services from the NGIs accessible for CTA members;
- A Single Sign On (SSO) authentication, internationally federated, mechanism that would make web-based Science Gateways accessible for the CTA community.

This document is organized following the ESA-PSS-05 Software Life Cycle standard and follows the user requirements recommended table of contents.

2 GENERAL DESCRIPTION

By definition, a Science Gateway is a "community-developed set of tools, applications, and data that is integrated via a portal or a suite of applications, usually in a graphical user interface, that is further customized to meet the needs of a specific community". A Science Gateway specific to CTA provides access to, software, services for data access (archive database and dedicated interfaces), tools for user support and data monitoring as well as to storage and computing infrastructures potentially distributed. The distribution of resources could be in the form of EGI Grid computing and storage resources, public and private cloud services, local personal computer resource, user-specific laboratory/institute storage and computing resources and eventually dedicated CTAO resources.

For the CTA community, the personalization of human-computer interface and management of complex access rights is one of the main requirements of this product and explains why this project is composed of two sub-products: Science Gateway and Single Sign On solutions.

Single sign-on (SSO) is a property of access control of multiple related, but independent software systems. With this property a user logs in once and gains access to all systems without being prompted to log in again at each of them. Conversely, Single sign-off is the property whereby a single action of signing out terminates



access to multiple software systems. As different applications and resources support different authentication mechanisms, single sign-on has to internally translate to and store different credentials compared to what is used for initial authentication.

2.1 Product perspective

The Science Gateway software product will require to be strongly linked to the Single Sign On software for the authentication process and management of access rights. Its main functionality will be to integrate external CTA web applications as science tools, data pipeline or monitoring tools, providing data and software access to a large set of users from scientists up to expert-users or operators.

2.2 Operational environment

The user will be connected on its personal device and will be able to access to the Science Gateway through a standard web browser.

3 SCIENCE GATEWAY SPECIFIC REQUIREMENTS

3.1 Science Gateway capability requirements

- UR-SG-0010** The Science Gateway must be able to propose to each user (connected or not) its authorized applications associated with authorized access to the corresponding archived data-sets that he is entitled to.

- UR-SG-0020** The Science Gateway must provide access to the data collections (Archive) resulting from the Science data collection and to the MC simulation results through a data selection application.

- UR-SG-0030** Applications that need to be accessible via the Science Gateway belong to three categories: those enabling the visualization and inspection of nominal CTAO performance as a function of the full possible range of observation modes, observation configuration and pointing direction, ; those more oriented to the targets of the observations, like tool for conditions and sources visibility, for checking already existing information about sources at VHE, minimal services to link to existing archives for astrophysical known sources multi-wavelengths related information , and connection to the Virtual Observatory as well; finally those devoted to Observations proposal handling services including TOO request submission form, High-level Data analysis (Science tools), pipelines (reconstruction, analysis, MC simulation) interfaces and MC simulation tools.



- UR-SG-0040** The minimal Data Management Monitoring applications that need to be accessible via the Science Gateway and more addressed to the guest observers are: status of the execution of scheduled observations; status information on the standard pipeline processing of the subset of data corresponding to the specific executed observation proposal.
- UR-SG-0045** Privileged users (For example Archive scientists, on-site operators) require to have access to additional privileged applications: Telescope monitoring, observations processing, data taking, on-site pipeline execution and calibration, data processing management, science data production, data archive, data transfer, MC simulation production and archive,.
- UR-SG-0050** A centralized access to each application's documentation including FAQ (Frequently Asked Questions) and tutorials must be provided.
- UR-SG-0055** The Science Gateway must provide access to its own documentation including FAQ (Frequently Asked Questions) and tutorials.
- UR-SG-0060** Each application must provide an associated user support with some sort of archive (For example mailing archive in mailman) to allow the knowledge sharing.
- UR-SG-0065** The Science Gateway should provide access to e-mail helpdesk – a single point of contact for issuing requests and reporting incidents.
- UR-SG-0070** The Science Gateway should provide a two-fold forum application, one devoted to the CTA consortium for observatory operation and operators' feedback and one more oriented to a larger and public scientific community for feedback and more effective and prompt scientific dissemination of results.
- UR-SG-0080** The Science Gateway must provide a resource discovery service to be able to propose to the users only their relevant authorized computing and storage services
- UR-SG-0090** The Science Gateway must provide a job execution service able to use distributed resources (see above) for application processing and selected data storage.
- UR-SG-0095** The Science Gateway must provide a service (for example download of applications and data) to enable computing on the local resources if such resources are supported by the application.
- UR-SG-0100** The Science Gateway must be able to transfer information from one application to another one (For example list of observation dataset references from data selection application to data analysis application).



- UR-SG-0110** The Science Gateway must be able to provide application workflow management services, for example: data selection application, followed by a user-specific pipeline of data analysis applications (science tools), then visualization of results.

3.1.1 Performance

- UR-SG-0150** The Science Gateway must be able to manage at least 1000 registered users and 100 simultaneous connections.

3.2 Science Gateway Constraint requirements

3.2.1 Communications interfaces

- UR-SG-0200** The Science Gateway must provide reasonable response time through an Internet standard Service Provider connection.

3.2.2 Hardware interfaces

- UR-SG-0300** The Science Gateway must be supported on various devices: laptop, pads and smartphones.

3.2.3 Software interfaces

- UR-SG-0400** The Science Gateway must be compatible with most popular web browsers: firefox, Internet Explorer, Google chrome and safari. This list could evolve in the future.

- UR-SG-0410** The Science Gateway must be independent of the user device operating system.

3.2.4 Human-Computer Interface

- UR-SG-0500** The Science Gateway must be developed as a web portal accessible from a choice of scientific community currently used web browsers (See UR-SG-0400).

- UR-SG-0510** The Science Gateway language must be English.

3.2.5 Adaptability



- UR-SG-0600** The Science Gateway and related applications must be flexible and modular enough to integrate new features during the period of operations and 10+ years after the CTA decommissioning.
- UR-SG-0610** The Science Gateway must be developed using existing frameworks to minimize the number of specific developments.
- UR-SG-0620** The Science Gateway framework used must follow existing standards.
- UR-SG-0630** Any tool, service, application plugged into the Science Gateway must be designed and developed compliant to a predefined Science Gateway software policy.

3.2.6 Availability

- UR-SG-0700** The availability of the Science Gateway and related applications must be > 98%.

3.2.7 Portability

- UR-SG-0800** The Scientific Gateway and related applications must be portable enough to be maintained over the period of operations and 10 years after CTA decommissioning.

3.2.8 Security

- UR-SG-0910** A specific user can be simultaneously Privileged user, Principal Investigator or Guest observer for one or more specific subset of observation data, Archive user for all the public Archive data.
- UR-SG-0915** Any privileged application could restrict its access to a specific user connected on a specific Internet subnet: for example Telescope monitoring application is accessible only if the user is connected in the site network.
- UR-SG-0916** The Science Gateway must be able to distinguish access rights among applications: e.g. a privileged user for one application could be a standard user for some others.
- UR-SG-0920** All data obtained by CTA observatory must be made public through an archive following a period of proprietary use. The proprietary period could be different for each observation data set.



- UR-SG-0930** Access rights for users and data rights are defined in a CTA data access policy document (To be written). The stakeholders may define or change access limitations to the archive.
- UR-SG-0940** The Science Gateway must be able to provide a public access and applications where the user is not identified.

4 SINGLE SIGN ON SPECIFIC REQUIREMENTS

4.1 *Single Sign On capability requirements*

- UR-SSO-0010** A Guest Observer that cannot be identified by a scientific community must be able to create a local account protected by login/password. The CTAO produces identification credentials for such a user.
- UR-SSO-0020** A member of any scientific community will be identified using either his X509 certificate or his login/password issued by his organization/institute.
- UR-SSO-0030** A CTA internal user will be identified using his CTA login/password or CTA consortium certificates (TBD).
- UR-SSO-0040** A CTA Grid user will be then identified with his GRID certificate to be able to use the CTA VO (Virtual Organization) tools.
- UR-SSO-0050** Once connected, user must be associated with the following attributes: CTA internal user or not, user profile (Standard user or Privileged user for a specific application and specific data), certified user by the Identity Federation (member of the scientific community) or not.
- UR-SSO-0060** Users want to log in once on the Science Gateway and gains access to its authorized applications and dataset without being prompted to log in again at each of them, so that users need to authenticate themselves only once per session.

4.1.1 Performance

- UR-SSO-0100** The Single Sign On process must be completed in a reasonable time.

4.1.2 Security



- UR-SSO-0900** All passwords must be encrypted on network.
- UR-SSO-0910** The identity of users must be checked for each connection request.
- UR-SSO-0920** The Science Gateway automatically disconnects users after a defined period of inactivity (timeout period).
- UR-SSO-0930** The Science Gateway user logout doesn't prevent a specific application to decide to keep user session opened.
- UR-SSO-0940** The user might be able to resume his previous session or start again from scratch.

4.1.3 Standards

- UR-SSO-1000** Standards protocols must be used.

5 APPENDIX

5.1 Summary table for Science Gateway requirements

Req #	Title
UR-SG-0010	Capabilities
UR-SG-0020	Data access
UR-SG-0030	Applications access
UR-SG-0040	Data Management Monitoring Applications access
UR-SG-0045	Privileged Applications access
UR-SG-0050	Documentation access
UR-SG-0055	Science Gateway documentation
UR-SG-0060	User support
UR-SG-0065	Help-desk
UR-SG-0070	Community feedback
UR-SG-0080	Resource discovery service
UR-SG-0090	Job execution service
UR-SG-0100	Input/output transfer between applications
UR-SG-0110	Application workflow management
UR-SG-0150	Simultaneous users performance
UR-SG-0200	Internet Service Providers support
UR-SG-0300	User devices support
UR-SG-0400	Web browser support
UR-SG-0410	User device Operating system dependency
UR-SG-0500	Web portal
UR-SG-0510	Science Gateway language
UR-SG-0600	Maintainability over 30-40 years
UR-SG-0610	Existing framework use
UR-SG-0620	Existing standards compliance
UR-SG-0630	Applications & tools software development policy
UR-SG-0700	Availability>98%
UR-SG-0800	Portability over 30-40 years
UR-SG-0910	Supported simultaneous user profiles
UR-SG-0915	Restricted network access
UR-SG-0916	Specific access rights per application
UR-SG-0920	CTA data collection access rights
UR-SG-0930	CTA access policy
UR-SG-0940	Public access

Table 1: Summary of the requirements in this document for Science Gateway

5.2 Summary table for Single Sign On requirements

Req #	Title
UR-SSO-0010	Local account
UR-SSO-0020	X509 certificate or institute login/password support
UR-SSO-0040	Grid certificate support
UR-SSO-0050	User attributes
UR-SSO-0015	Single Sign On capability
UR-SSO-0100	Performance
UR-SSO-0700	Availability>98%
UR-SSO-0800	Portability over 30-40 years
UR-SSO-0900	Password encrypted on network
UR-SSO-0910	Identity check for each connection
UR-SSO-0920	Connection time-out
UR-SSO-0930	Ability to keep user session opened even if browser is closed
UR-SSO-0940	Session recovery
UR-SSO-1000	Existing standards compliance

Table 2: Summary of the requirements in this document for Single Sign On