

An Innovative Proposal for the CTA Science Gateway

<http://cta-sg.oact.inaf.it/>

EGI Community Forum 2014
May 2014 - University of Helsinki, Finland

A. Costa, U. Becciani, P. Massimino, E. Sciacca, F. Vitello
INAF-OACT

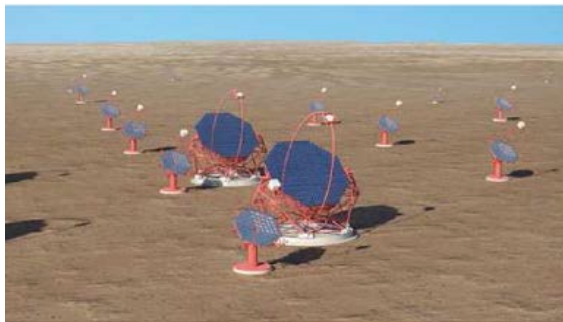


The Cherenkov Telescope Array Concept



- 2006: The CTA concept was first proposed to the ESFRI committee
- CTA Consortium : consists of more than 1000 scientist and engineers.
- More than 160 institutions from 27 countries

The CTA Concept



Current systems of Cherenkov telescopes use at most four telescopes, providing best stereo imaging of particle cascades over a very limited area, with most cascades viewed by only two or three telescopes. **An array of many tens of telescopes will allow the detection of gamma-ray induced cascades over a large area on the ground**, increasing the number of detected gamma rays dramatically, while at the same time providing a much larger number of views of each cascade. This results in both improved angular resolution and better suppression of cosmic-ray background events.

In a possible design scenario, the southern hemisphere array of CTA will consist of three types of telescopes with different mirror sizes in order to cover the full energy range. The northern hemisphere array would consist of the two larger telescope types.



ASTRI SST-2M at Serra La Nave (CT)

Work in progress...

Inauguration on September 24, 2014

LST designed to detect showers at low-energy range < 100 GeV mirror diameter ~ 23 m

MST core energy range 0.2-10 TeV mirror diameter ~ 12 m

SST high-energy range > 10 TeV

<https://portal.cta-observatory.org/Pages/Home.aspx>

The CTA Concept

Table 1: Cumulated data for the construction phase (scenario close to minimum with zero suppression from second year)

Year	2017	2018	2019	2020	2021
Raw data (PB)	2.8	8	1.4	1.9	2.5
Cumulated raw data (PB)	2.8	3.6	5	6.9	9.4
Processed and reprocessed data (PB)	0.5	2.3	4.5	6.8	9
Cumulated Observation data (TB)	3.3	6.4	12.3	21	32.5
Monte-Carlo data (PB)	3.3	6.4	1.23	20	20
Cumulated Data (PB)	6.7	12.8	24.6	41	53.5

Table 2: Cumulated data for the operation phase (scenario close to minimum)

Year	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031
Raw data (PB)	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8
Cumulated raw data (PB)	12.2	15	17.7	20.5	23.3	26.1	28.9	31.6	34.9	37.1
Processed and reprocessed data (PB)	19.7	25.4	31	36.6	42.3	47.9	53.6	59.2	64.8	70.4
Cumulated Observation data(TB)	46	60	74.1	88.1	102.1	116.2	130.2	144.3	158.3	172.3
Monte-Carlo data (PB)	20	20	20	20	20	20	20	20	20	20
Cumulated Data(PB)	66	80	94	108	122	136	150	164	178	192

Table 3: Computing needs in construction phase. All values are in (Milliard kHS06 · sec)

Year	2017	2018	2019	2020	2021
Data pipeline needs	19.7	59.1	98.5	138	177
Simulation needs	360	360	360	360	360
Re-processing needs	0	19.7	78.8	177	315
Cumulated needs	380	439	537	675	853

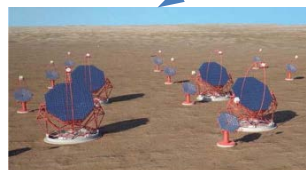
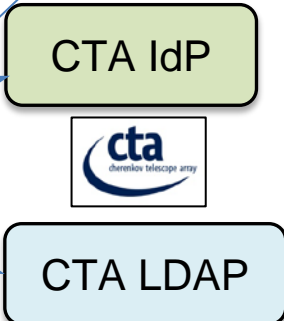
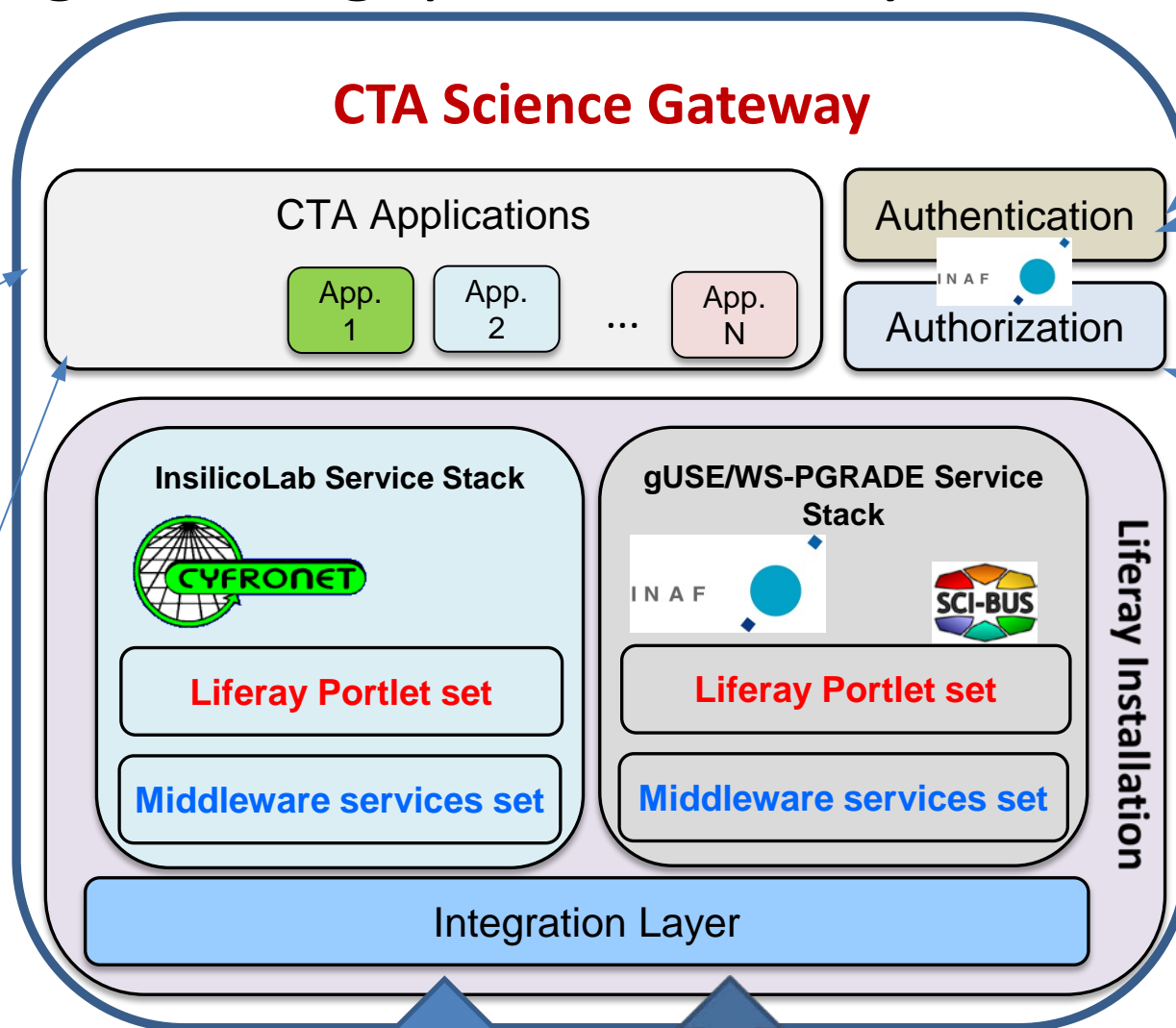
Table 4: Computing needs in operation (from t0 to t0+10 years). All values are in (Milliard kHS06 · sec).

Year	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031
Data pipeline needs	197	197	197	197	197	197	197	197	197	197
Simulation needs	360	360	360	360	360	360	360	360	360	360
Re-processing needs	493	690	887	1084	1281	1478	1675	1872	2070	2267
Cumulated needs	1050	1247	1444	1641	1838	2035	2232	2430	2627	2824

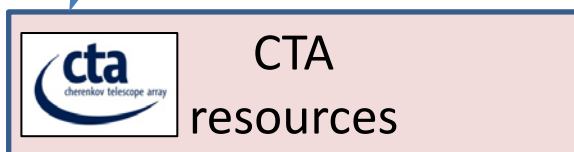
HEP-SPEC06 (HS06) is the new HEP-wide benchmark for measuring CPU performance.

One HS06 is providing a processing capacity of $3600 \times 24 \times 365 = 31536$ kHS06 · sec · year⁻¹.

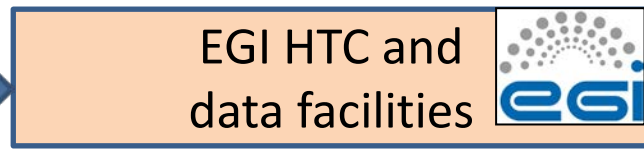
CTA Science Gateway



CTA sites



computation





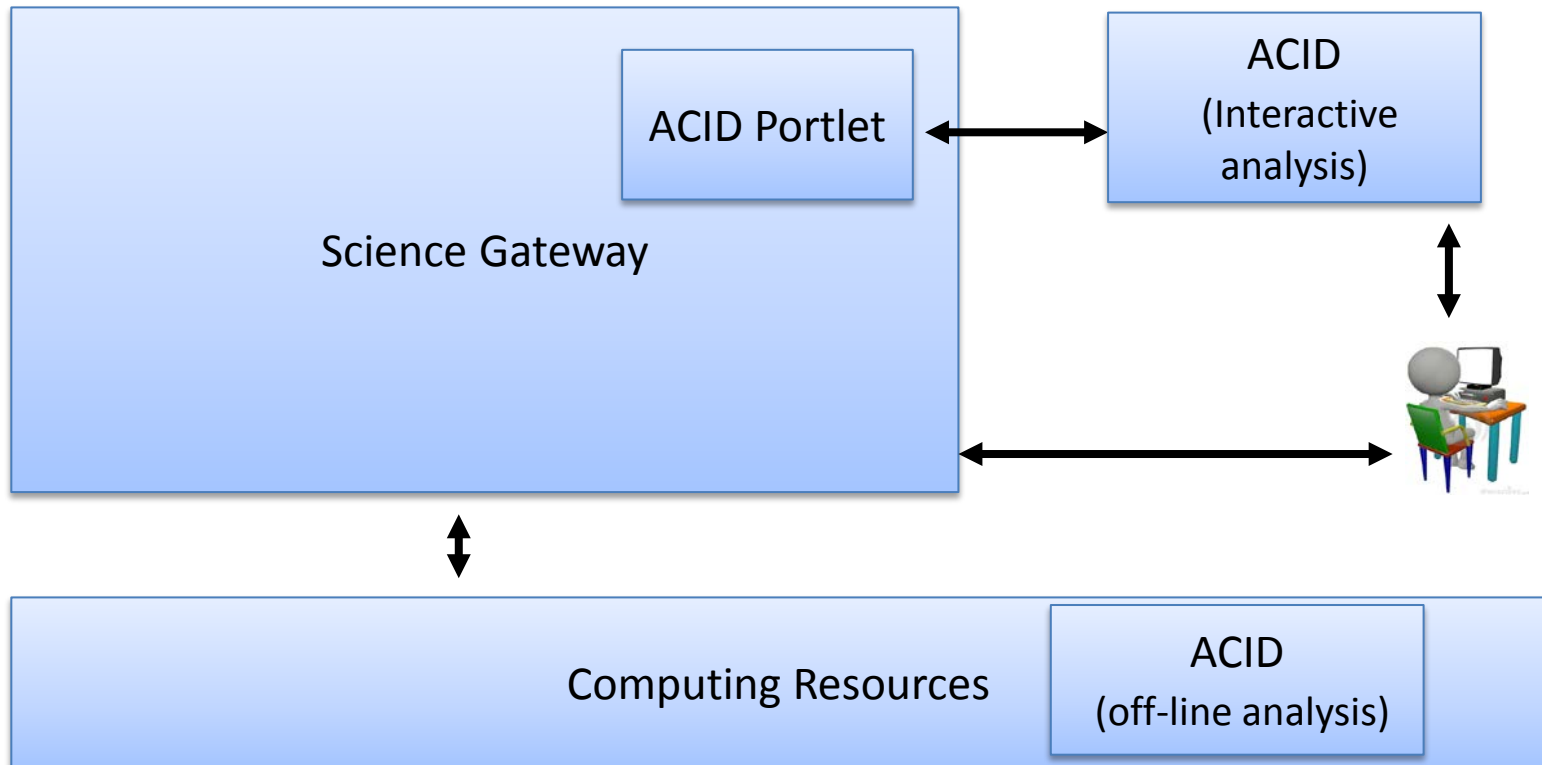
CTA Science Gateway Prototype



- A CTA gUSE/WS-PGRADE Science Gateway implemented by INAF.
<https://cta-sg.oact.inaf.it/>
- Workflows Sharing (SHIWA repository)
- **Single Sign On: SAML 2, Shibboleth.**
- **CTA Identity Provider**
- Easy integration with **Identity Federation**
- **Astronomical & physics Cloud Interactive Desktop (ACID)**
- ACID is an “**Application As A Service**” & “**Data As A Service**”
- Cloud Data through the use of OwnCloud
- More than 150 astronomical and physics tools!
- **ACID can be accessed by mobile devices**

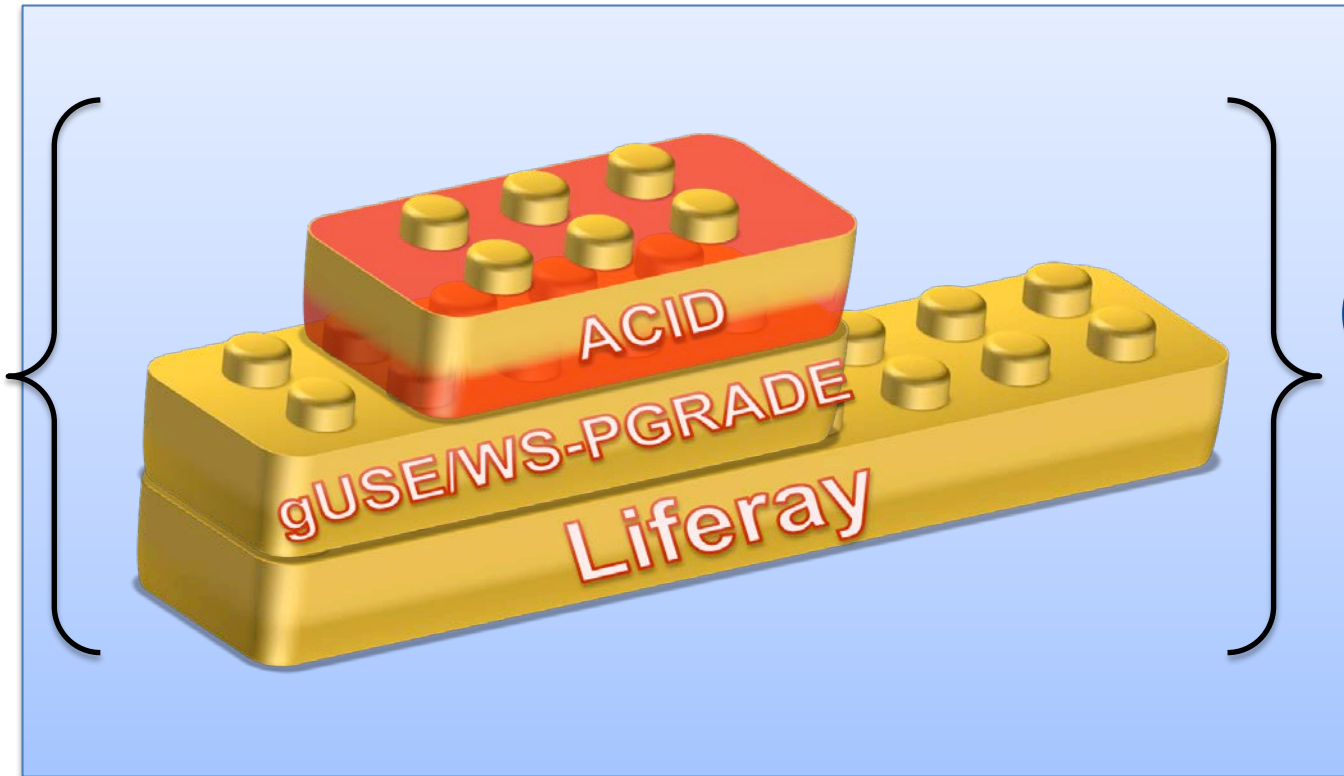


Overall Architecture





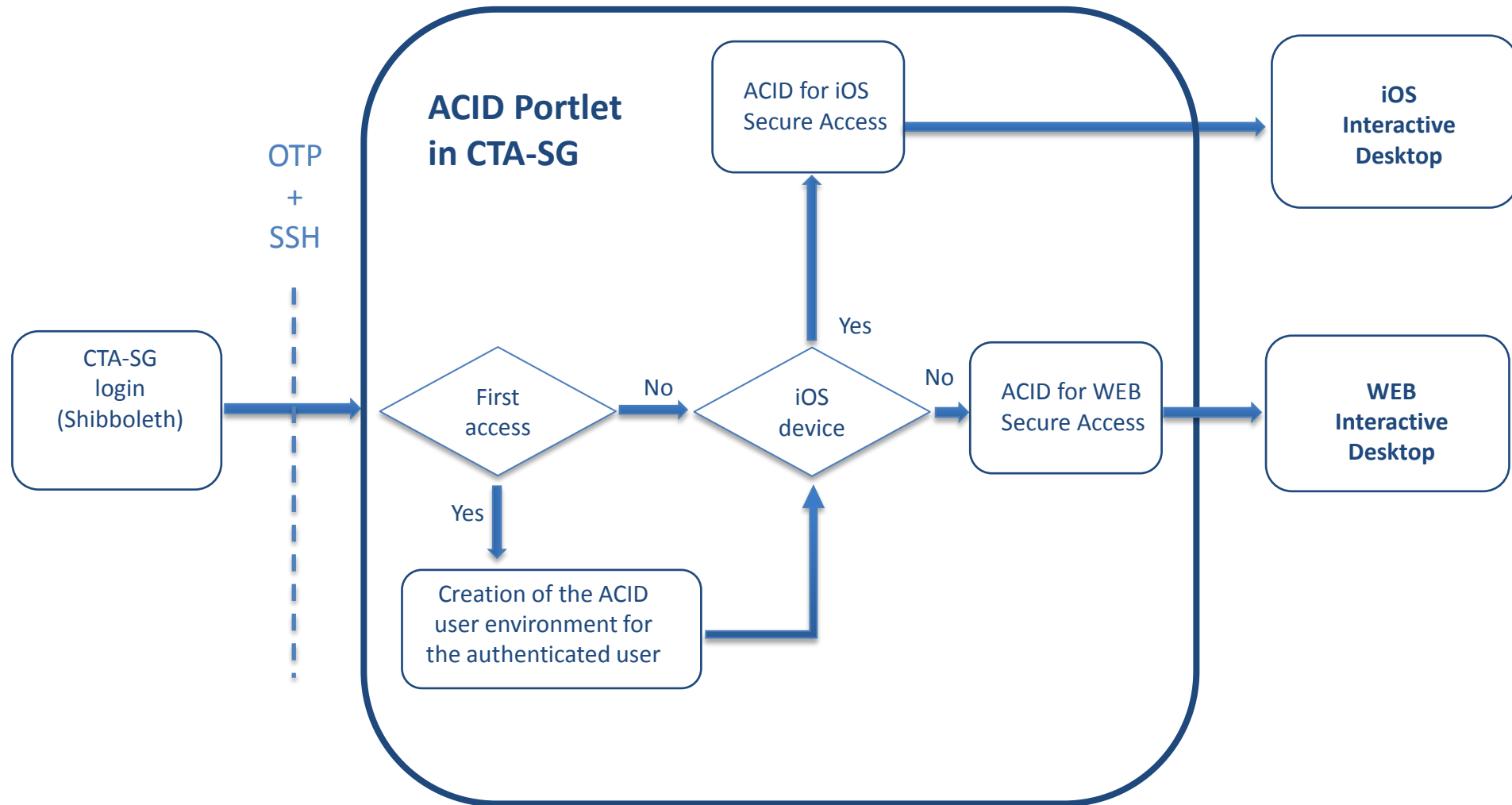
Overall Architecture

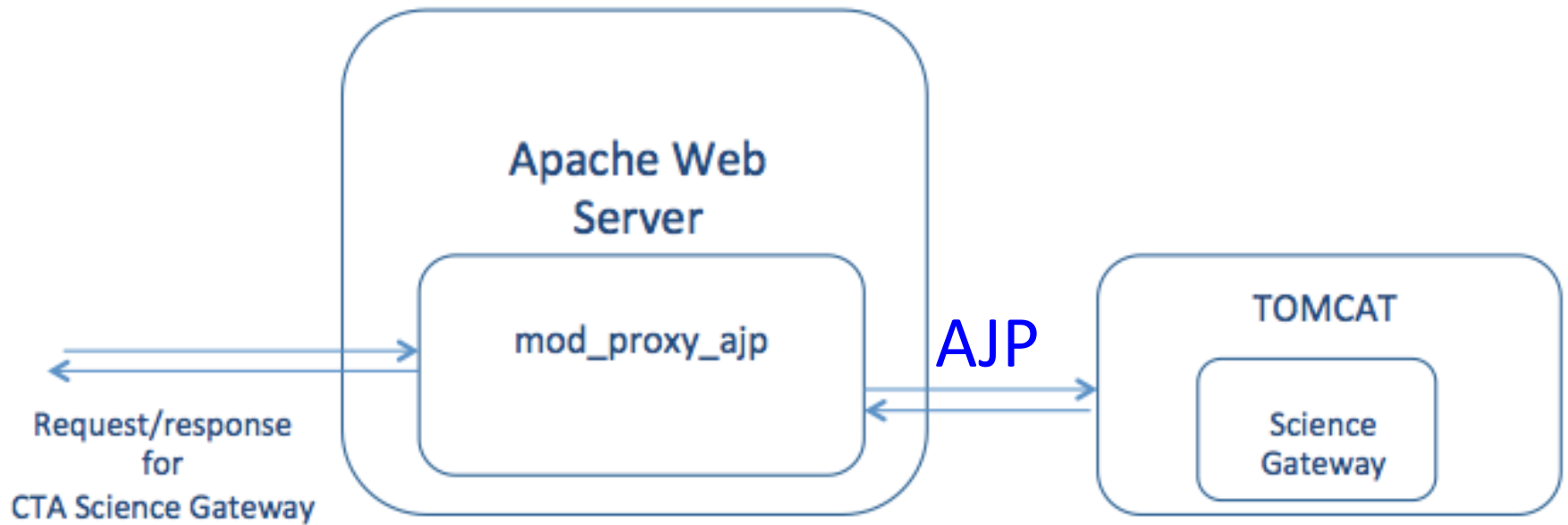


SSO
(Shibboleth)



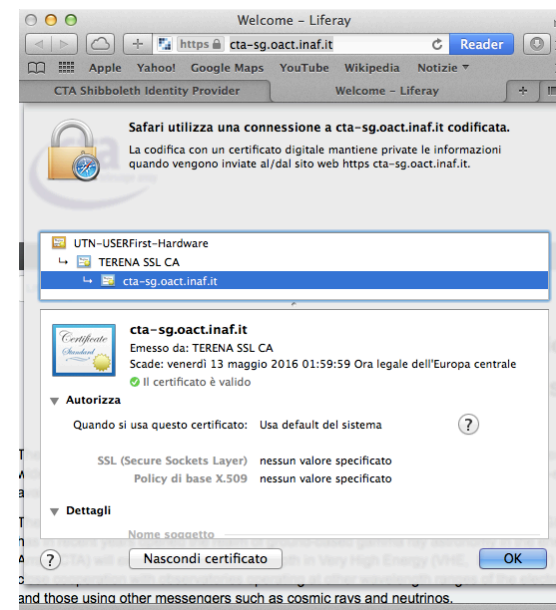
Overall Architecture





- Tomcat runs the Scientific Gateway.
- Tomcat is integrated with Apache by AJP
- This Architecture allows **Load Balancing**

- SAML 2 (*Security Assertion Markup Language*): *The standard.*
- *Shibboleth*
- Most widely used in education, government
- Broadly adopted in Europe
- The new 2.0 release implements SAML 2 !
 - Backward compatible with 1.3
- Free & Open Source
 - *Apache 2.0 license*
- **Identity Provider**
- **Service Provider**
- **Liferay AAI (Authentication Authorization Infrastructure)**
- **HTTPS via ssl Certificates issued by a trusted Certificate Authority's Root Certificate: TERENA (Trans-European Research and Education Networking Association)**





Single Sign On



(Liferay Attribute Management)

- *Auto creation of a new LIFERAY user using Shibboleth attributes (email, first name, last name)*
- *Auto update of user information upon login*
- *Role mapping*



Supported Cloud Standards



- Since **Jan 2014** Direct Cloud Access is provided (*gUSE version 3.6.2*)
- By this feature the user can easily submit jobs directly to an accessible cloud.
- **Any clouds that implements the Amazon EC2 interface**
- **(e.g. OpenNebula) are accessible CTA Science Gateway**

Overall Architecture

- ACID

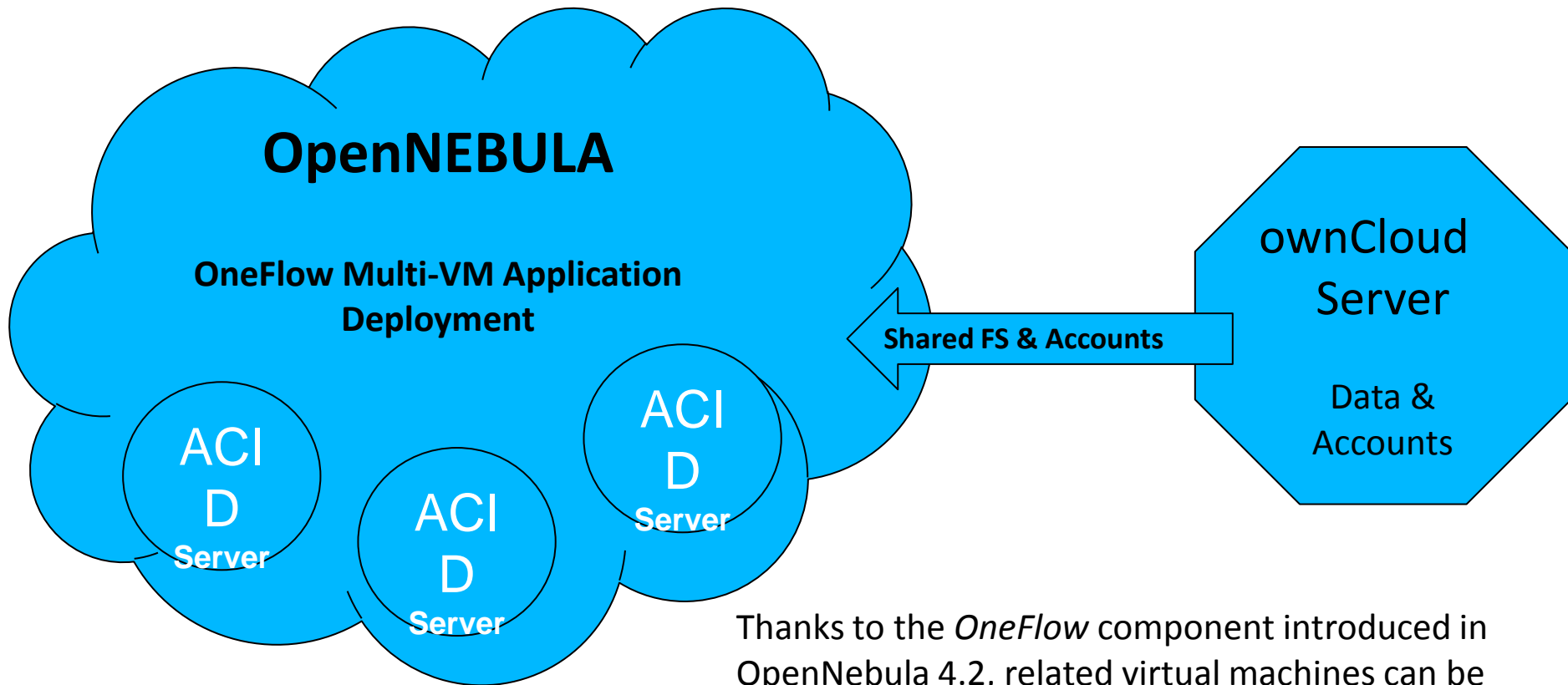
- gUse/WS-PGRADE

- Liferay



LIFERAY®





Thanks to the *OneFlow* component introduced in OpenNebula 4.2, related virtual machines can be grouped into a Service

- **Application Service Provider :**
 - It offers more than 150 applications as a service accessible by the CTA INAF prototype
- **Data Cloud environment**
 - Data Cloud is available to Applications and it is shared with your local computer



SSO
(Shibboleth)

Storage (data cloud)



Thanks to **Unison** and **ownCloud** the user will find his data on CTA Gateway



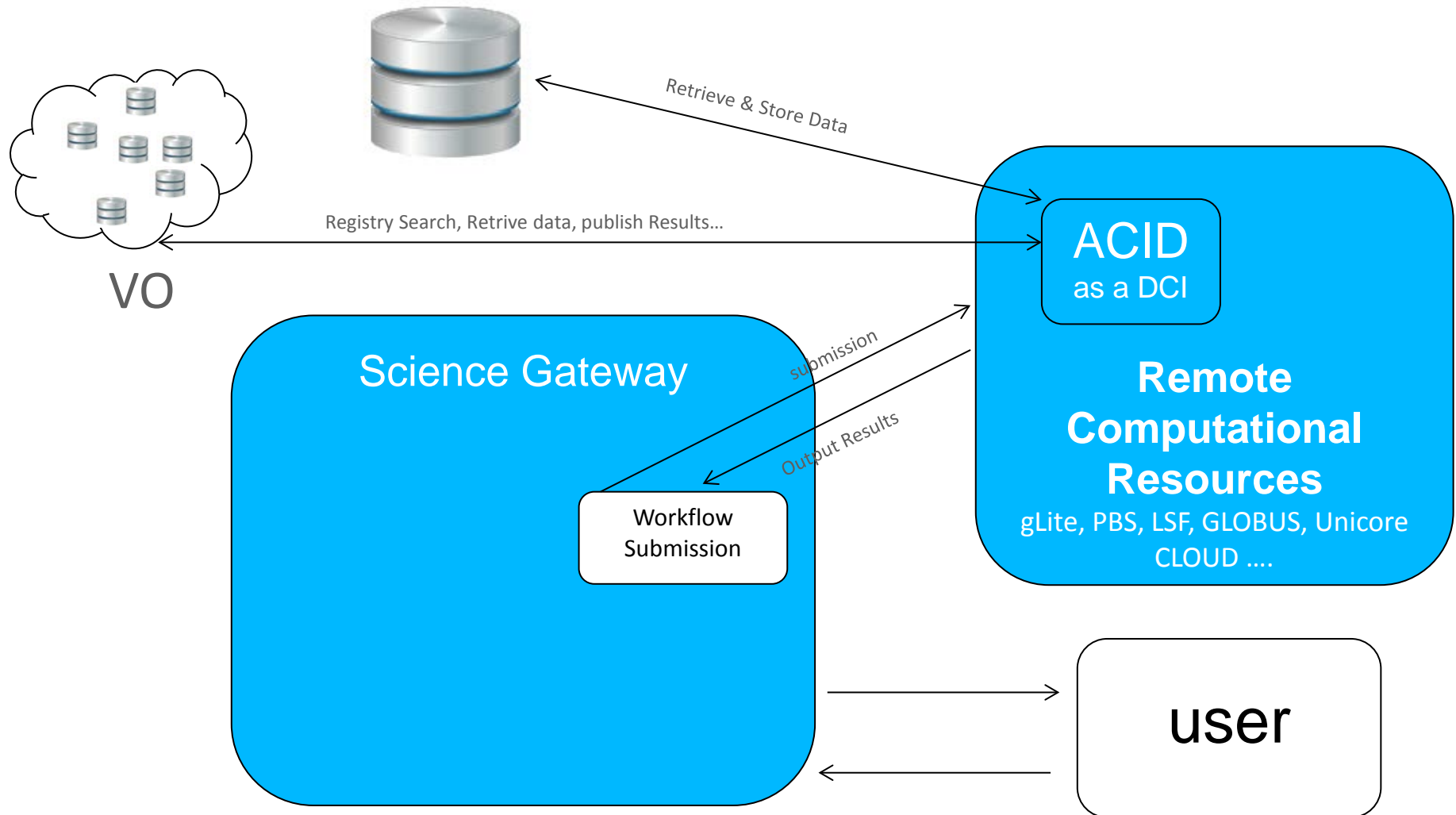
ownCloud client allows end user to share files on his desktop or smartphone

Function and Purpose

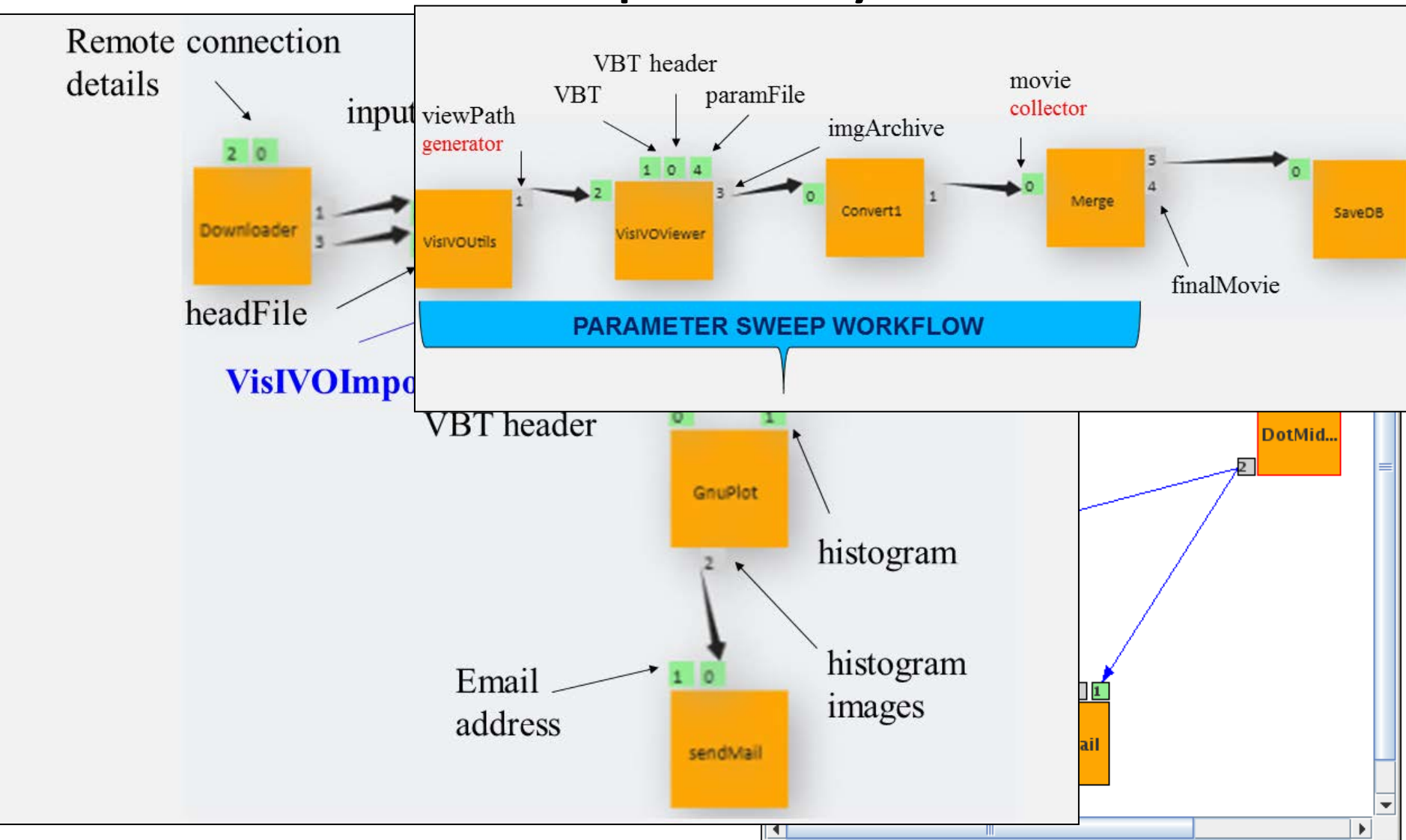
- This CTA Science Gateway is a **web environment**
- It is an environment: **Gateway + Data Cloud**
- It includes different tools that are necessary for the science analysis of CTA data.
- It allows to implement
 - **Standard reduction pipelines (OffLine mode)**
 - **Real-Time analysis pipelines (Interactive mode)**

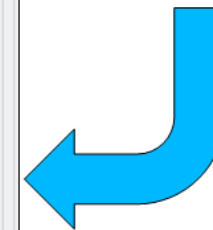
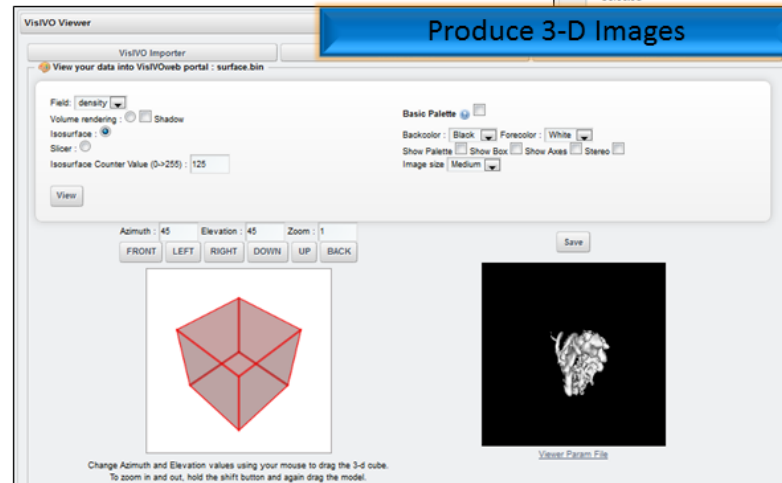
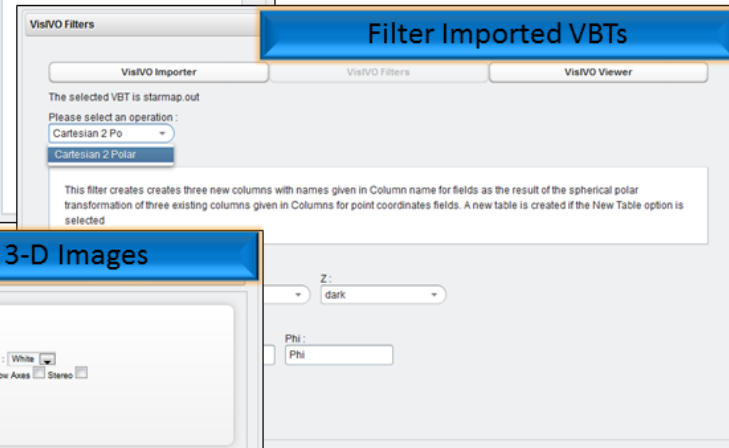
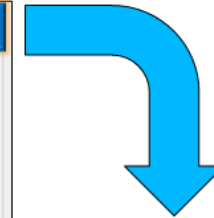
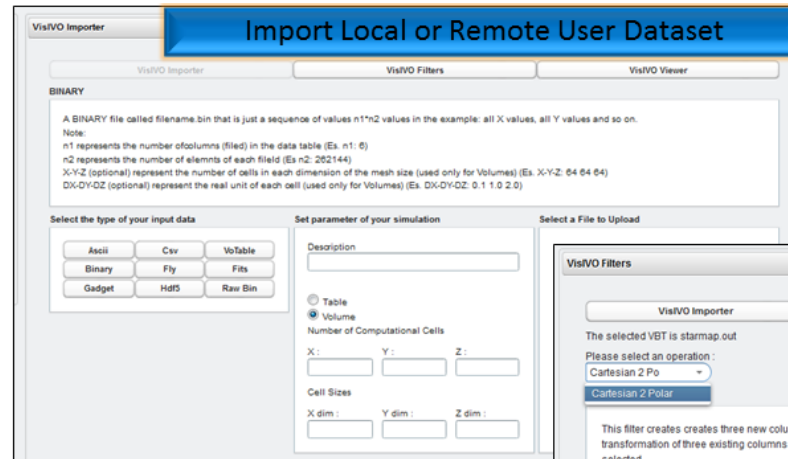
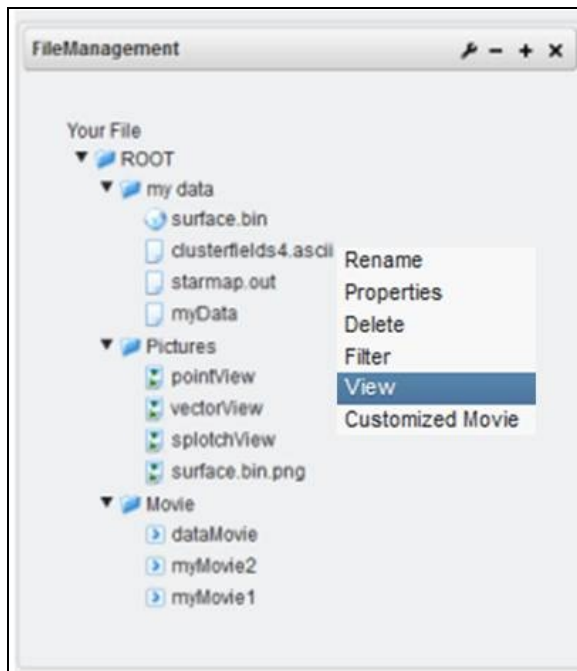
Offline analysis

Archive



Workflow Editor & Repository

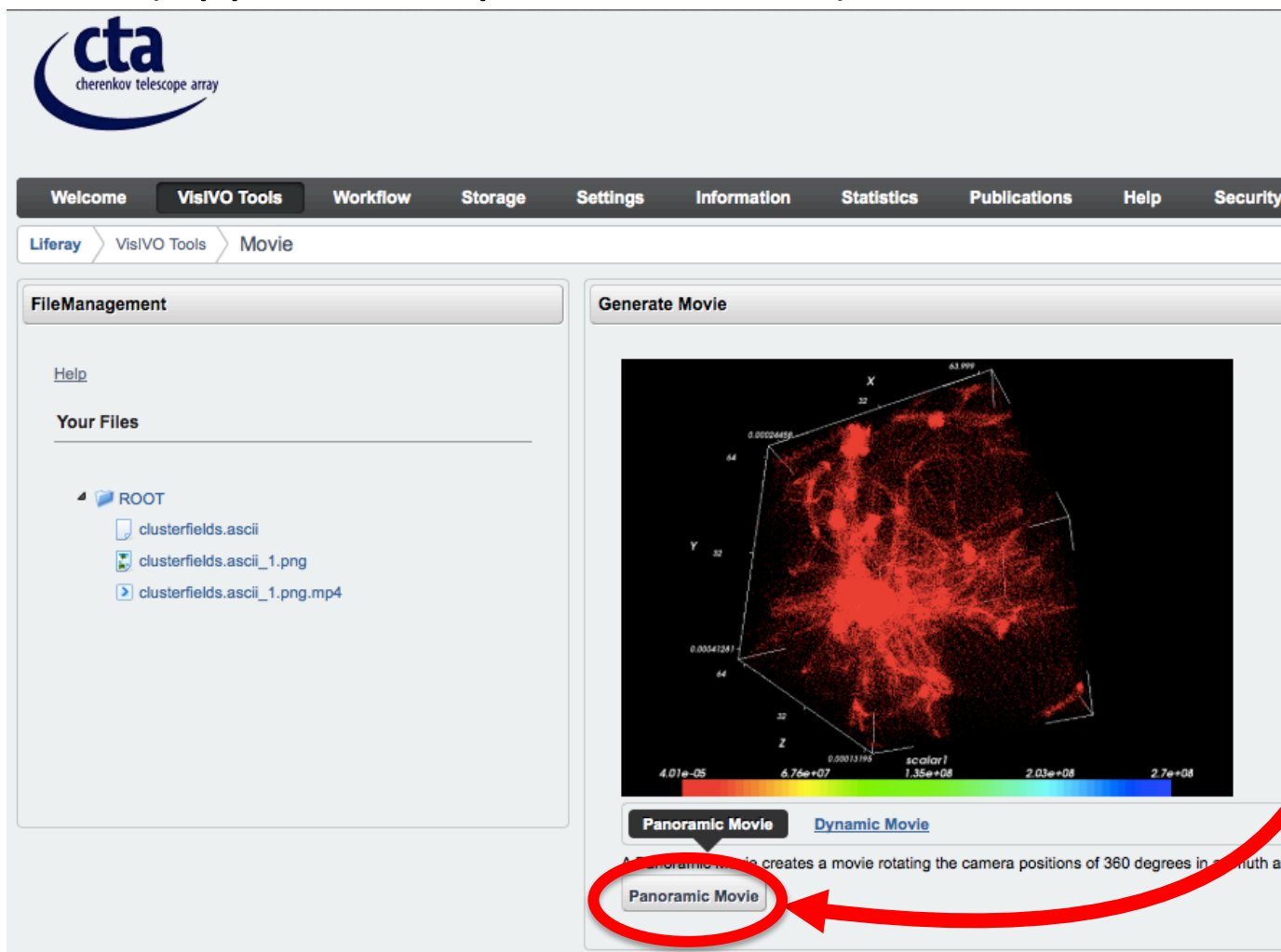




Workflows through a Simple Interface

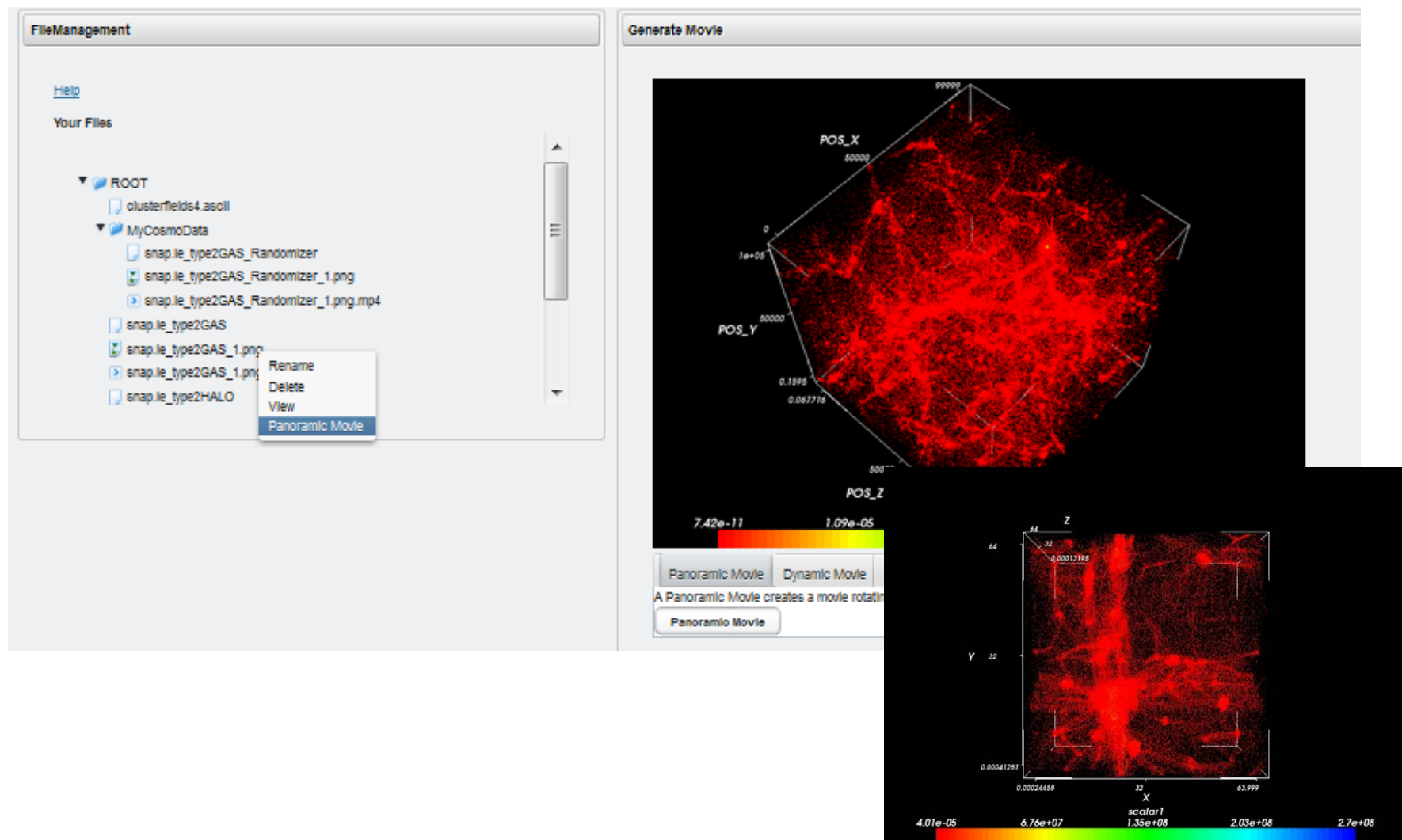
Workflow submission

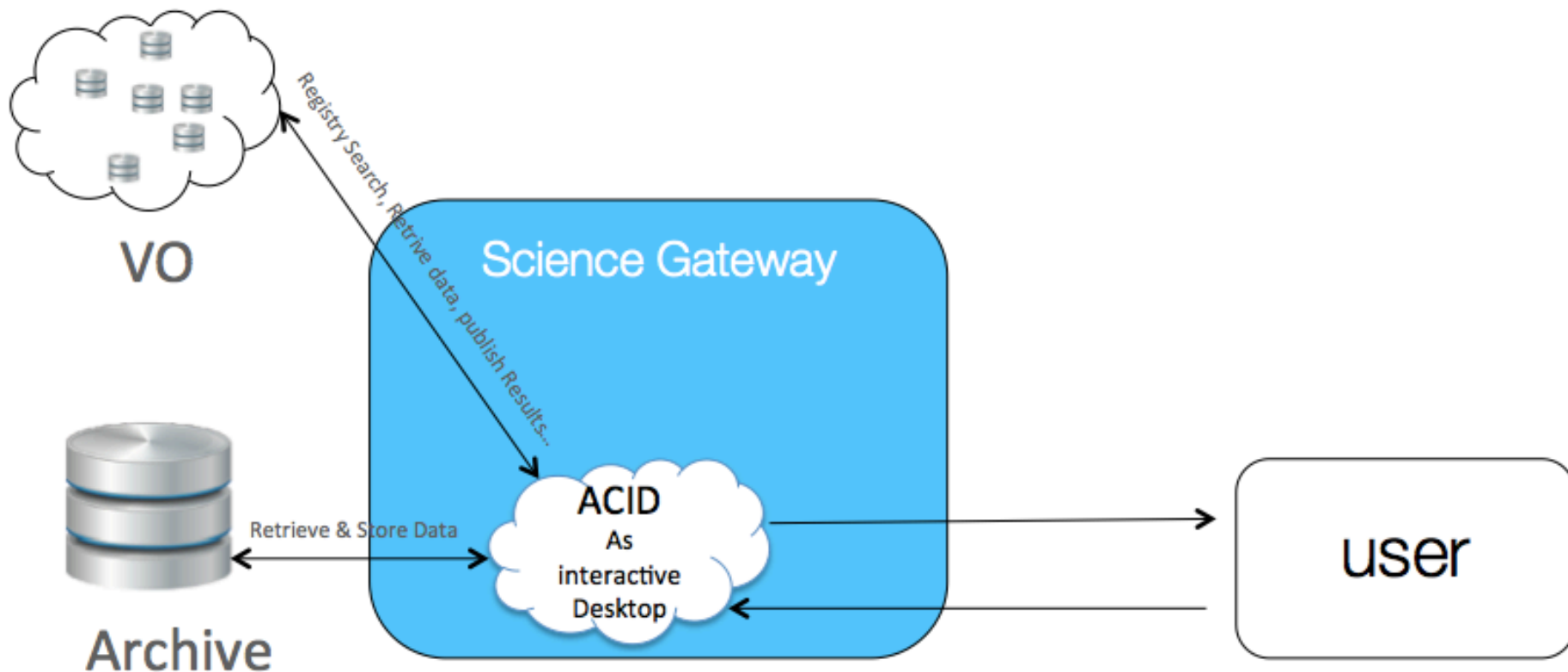
ASM (Application Specific Module) in ACTION

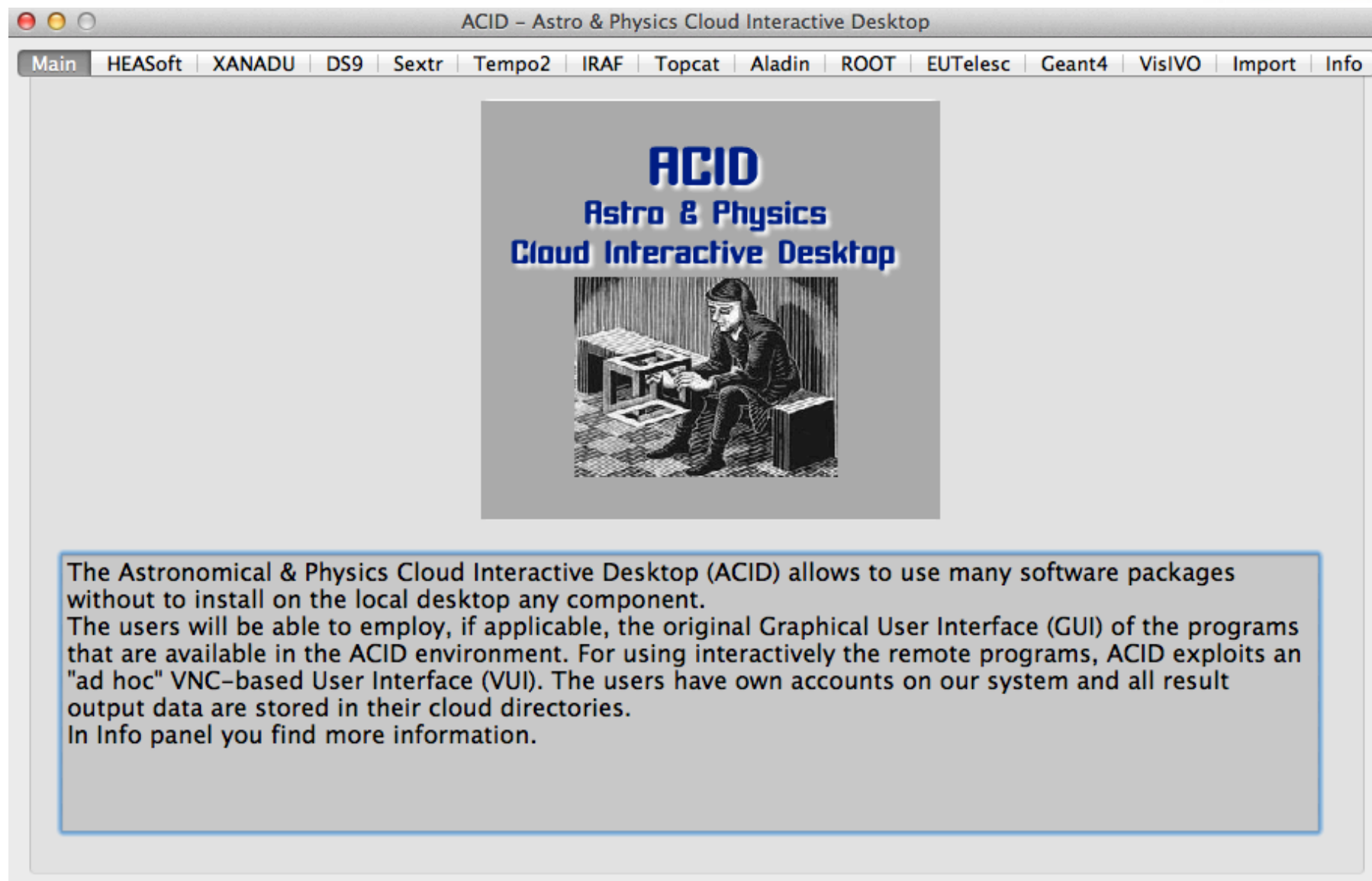


The screenshot displays the CTA VisIVO Tools web interface. The top navigation bar includes links for Welcome, VisIVO Tools, Workflow, Storage, Settings, Information, Statistics, Publications, Help, and Security. The 'VisIVO Tools' tab is active, and the 'Movie' sub-tab is selected. On the left, the 'FileManagement' panel shows a file tree under 'Your Files' with a 'ROOT' directory containing three files: 'clusterfields.ascii', 'clusterfields.ascii_1.png', and 'clusterfields.ascii_1.png.mp4'. The main content area is titled 'Generate Movie' and features a 3D visualization of a cluster field. The visualization shows a dense distribution of red points forming a complex, filamentary structure. A color bar at the bottom of the plot indicates a scalar value range from $4.01e-05$ to $2.7e+08$. Below the plot, there are two buttons: 'Panoramic Movie' and 'Dynamic Movie'. The 'Panoramic Movie' button is circled in red, and a red arrow points to it from the right side of the image. A tooltip is visible over the 'Panoramic Movie' button, stating: 'A Panoramic Movie creates a movie rotating the camera positions of 360 degrees in 1 minute and'.

Workflows through a Simple Interface







The screenshot shows a web browser window titled "ACID - Astro & Physics Cloud Interactive Desktop". The browser's address bar and menu bar are visible. The menu bar includes links: Main, HEASoft, XANADU, DS9, Sextr, Tempo2, IRAF, Topcat, Aladin, ROOT, EUTelesc, Geant4, VisIVO, Import, and Info. The main content area features a large graphic with the text "ACID Astro & Physics Cloud Interactive Desktop" and an illustration of a person sitting at a desk with a computer. Below the graphic, a text box explains the ACID system.

ACID
Astro & Physics
Cloud Interactive Desktop

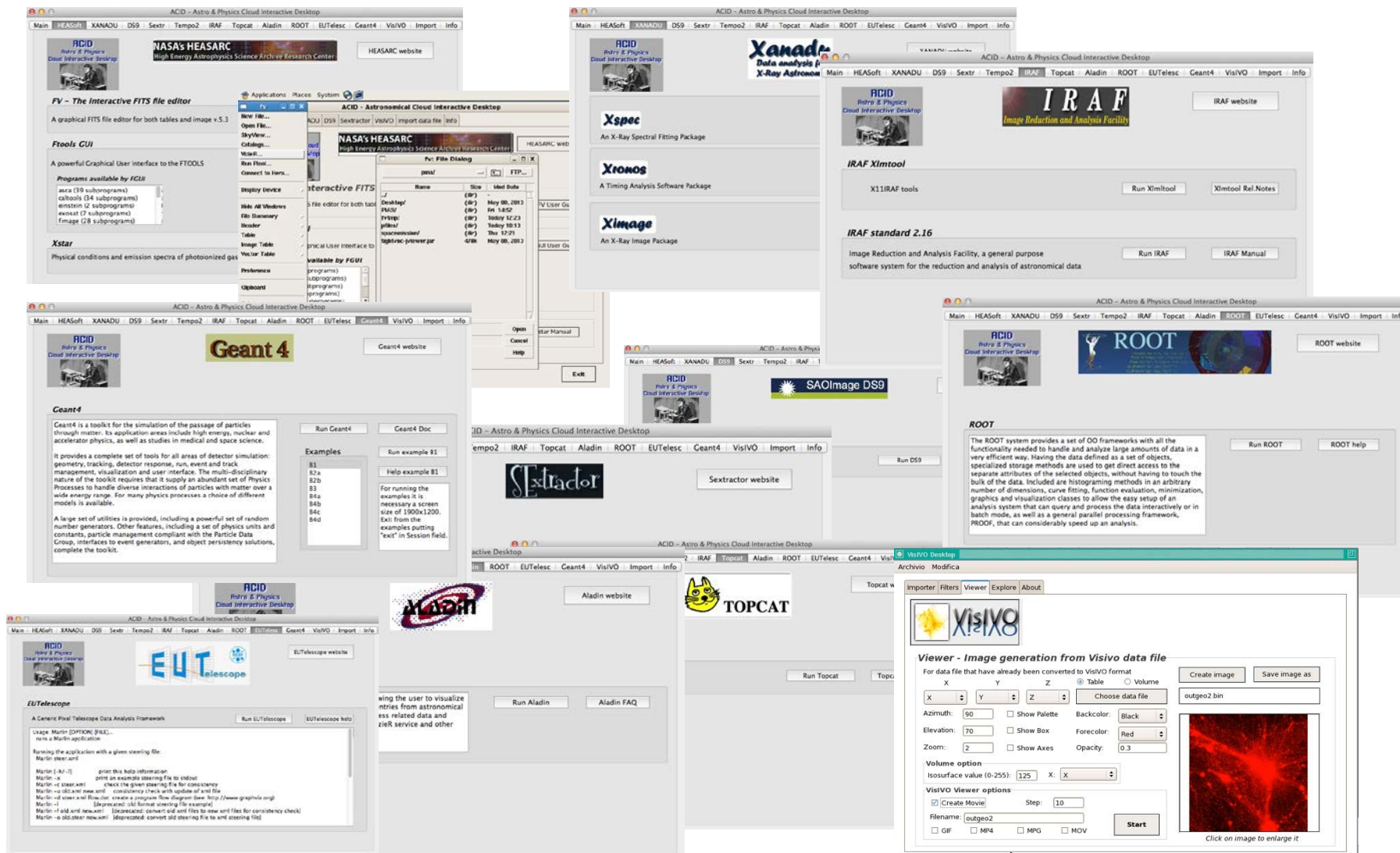
The Astronomical & Physics Cloud Interactive Desktop (ACID) allows to use many software packages without to install on the local desktop any component. The users will be able to employ, if applicable, the original Graphical User Interface (GUI) of the programs that are available in the ACID environment. For using interactively the remote programs, ACID exploits an "ad hoc" VNC-based User Interface (VUI). The users have own accounts on our system and all result output data are stored in their cloud directories. In Info panel you find more information.



ACID

Astronomical & Physics Cloud Interactive Desktop





The collage displays several key components of the ACID desktop environment:

- Navigation and Main Menu:** Multiple windows show the top navigation bar with links to Main, HEADSITE, XANADU, D59, Sext, Tempo2, IRAF, Topcat, Aladin, ROOT, EUTelescope, Geant4, VisIVO, Import, and Info.
- Tools and Applications:**
 - IRAF (Image Reduction and Analysis Facility):** Includes IRAF Ximtool and IRAF standard 2.16 for image reduction and analysis.
 - Xspec and Xronos:** Xspec is an X-Ray Spectral Fitting Package, and Xronos is a Timing Analysis Software Package.
 - Ximage:** An X-Ray Image Package.
 - Geant4:** A toolkit for the simulation of the passage of particles through matter, used in detector simulation.
 - SAOImage DS9:** A multi-platform astronomical image display package.
 - Extractor:** A tool for extracting data from images.
 - Aladin:** A tool for visualizing astronomical data.
 - TOPCAT:** A tool for analyzing and visualizing astronomical data.
 - EUTelescope:** A Generic Pixel Telescope Data Analysis Framework.
 - VisIVO:** A tool for visualizing and analyzing data from the VisIVO framework.
- Data and File Management:** Windows like 'File Dialog' and 'File List' show the management of data files, including their names, sizes, and modification dates.
- Documentation and Help:** Various windows provide links to websites (e.g., HEASARC, XANADU, IRAF, ROOT, Aladin, TOPCAT) and local help files.



Conclusions

- Standards
 - Java Portlet Specification (JSR168)
 - SAML 2.0
 - WS-PGRADE/gUSE
- Enlarge the developer community
- Improve the sustainability of the software
- An ecosystem of new technologies:
 - Mobile access
 - Federated authentication
 - Workflow engine
 - Cloud services
- Tailor a product that suits the present and future requirements of the CTA community.