Development Upgrades of the Atacama Large Millimeter/submillimeter Array

ARCADE: ALMA Reduction in the CANFAR Data Environment

Cycle 7 ALMA Development Study Final Report

Principal Investigators: Toby Brown & Helen Kirk

Institution: Herzberg Astronomy & Astrophysics Research Center, National Research Council of Canada

PI Contact Information: Tobias.Brown@nrc-cnrc.gc.ca

Executive Summary

This document is the final report for the Cycle 7 ALMA Development Study, ARCADE: ALMA Reduction in the CANFAR Data Environment. The study funded development to expand the existing ARCADE prototype to enable ALMA data reduction, thoroughly test usability and reliability of the system, and determine the steps required to offer ARCADE to the wider community.

The ARCADE science platform provides a stable, scalable, and interactive computing environment in which users are able to process and analyse ALMA data. ARCADE is built entirely on open source software and designed so that it can be scaled out to meet the storage and processing demands of the user. The software is robust, easy to maintain, and portable across a range of computing infrastructure and storage systems (e.g., local vs. cloud). ARCADE is also fully integrated with existing ALMA value-added software packages, CARTA and ADMIT.

Software and user testing of ARCADE was continuously conducted during the study. Where appropriate, issues raised were addressed by project PIs and/or developers. In general, users are satisfied with the delivered ARCADE beta and, while providing a production-ready platform was not the goal of this study, important lessons were learned for establishing that capability.

The study has concluded on budget and to specification. There was an approved no-cost extension of two months to the schedule. All the technical and user-driven benchmarks established in the proposal have been met.

Along with this report, the study has produced a number of products for delivery to NRAO. The delivered products are: the open-source ARCADE software and documentation for developers, CASA software containers, container maintenance and deployment documentation, user documentation and tutorials.

List Of Images

Background

The ability to efficiently mine the ALMA archive is vital to maximize the community's and the observatory's scientific impact. The archive content continues to grow in size and breadth as ALMA cycles continue, available bandwidth is expanded, longer baselines are incorporated, and the receivers and correlator are upgraded. This requires science users to similarly increase their capacity and computing capabilities so they can properly exploit this increasingly rich resource. Indeed, the development of archive usability and impact is explicitly outlined as a priority by the ALMA 2030 Development Roadmap and the Archive Working Group.

Current barriers to ALMA users successfully achieving their science goals include:
- *Storage Capacity:* Insufficient storage for large ALMA datasets before, during, and after processing, and the inability to handle such data with local computing.
- *Computing Capacity:* Insufficient computing resources for data processing and slow network traffic between the archive and local computers.
- *Software Complexity:* Users are required to install and manage multiple CASA versions, dependencies, and value-added software packages.

The ARCADE Science Platform

ARCADE is an online scientific computing platform specifically designed for reducing and analysing ALMA data. The system is currently in beta deployment with a limited number of simultaneous users.

The ARCADE Development Study goal was to deliver a prototype ARCADE science platform capable of removing the barriers outlined above. The study also set out to determine and quantify the resources and additional development required to be able to offer this science platform to the wider ALMA community.

The platform comprises a collection of open-source software tools, including the *Common Astronomy Software Applications* package (CASA), and provides users with an interactive computing environment on scalable cloud infrastructure. ARCADE handles user accounts and authentication, comes preloaded with software, and is fully interactive via a familiar linux-based desktop environment connected to distributed computational and storage resources.

The system is fully integrated with the Cube Analysis and Rendering Tool for Astronomy (CARTA) and the ALMA Data Mining Toolkit (ADMIT), providing access to all the tools needed for processing and analysing ALMA data from the archive.

Purpose and Scope

Our long-term objective is to provide a secure, stable, scalable science platform on which the community is able to process and analyse ALMA data for its own science.

The Development Study proposal outlines three main objectives.
1. Expand the existing ARCADE prototype to include all necessary components for ALMA data reduction
2. Thoroughly test the usability and reliability of this prototype system
3. Determine and quantify the steps required to be able to offer this service to the wider community (i.e., many simultaneous users).

The scope of objective 1) includes all versions of CASA necessary to ensure that the calibration scripts provided alongside any archival ALMA data can be run successfully without additional installations by the user. We also integrate the computing environment with additional value-added software packages such as ADMIT, CARTA, AnalysisUtils, and Anaconda. Finally, we improve technical aspects of the platform interface, in particular copy/paste functionality and software reliability.

For objective 2), the internal team (NRC) completed testing of the ARCADE system, checking for fundamental issues and bugs. The external team (McMaster University) provides an external review of the usability of the platform and perform tests to quantify the experience of a typical archive user.

Objective 3) determines the technical solutions for optimal storage, data transfer, and resource allocation. These solutions are then developed and tested. This objective connects cloud computing and cloud storage resources reliably with minimum data transfer time while achieving the science use goals.

These objectives are broken down into the nine study deliverables presented in Table 1.

Table 1. Original work-breakdown structure (WBS) number and schedule of the ARCADE Development Study deliverables.

WBS No.	Deliverable	Completion Date
1a	**Research cloud concept deployment** Set up the system for automating deployment, scaling, and management of ARCADE on the research cloud (Compute Canada).	January 16, 2020
1b	**Commercial cloud concept deployment** Set up the system for automating deployment, scaling, and management of ARCADE on the commercial cloud (Google Cloud). This is used as a concept testing platform without interrupting alpha or beta testing.	January 8, 2021
2	**Alpha prototype release** Provide alpha prototype version of ARCADE on scalable infrastructure on the research cloud (Compute Canada). Develop a basic alpha prototype tutorial.	April 6, 2020
3	**CASA container testing on alpha prototype** Add and fully test all CASA versions on the alpha ARCADE prototype. Install other ALMA software (e.g., ADMIT, CARTA).	June 12, 2020
4	**Alpha storage solution** Explore storage solutions to achieve a secure and performant user personal storage space with data sharing.	February 24, 2020
5	**Alpha testing** User testing of the alpha ARCADE prototype. Identify major problems and report them.	June 12, 2020
6	**Beta development** Beta prototype development. Incorporate all easily addressed issues raised by the alpha users testing. Revise the alpha prototype tutorial for the new beta prototype system.	October 1, 2020
7	**Beta user testing** In-depth testing of usability and performance of the beta prototype version of ARCADE for a variety of science use	February 18, 2021

	cases. Includes the introduction of new users to the system to test the beta prototype tutorial.	
8	**Container maintenance** Examine container maintenance and upgrade issues. Document possible solutions.	December 18, 2020
9	**Final Report** Prepare the final report, with input from development, operations, and science team members.	February 19, 2021

Overview
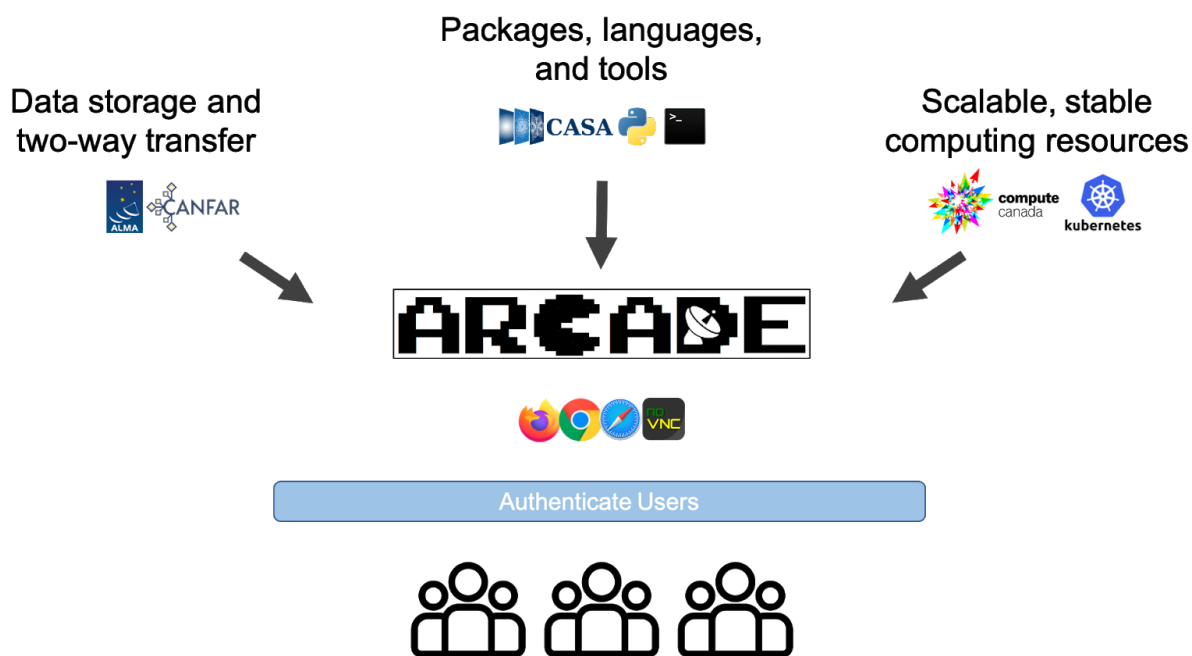
Figure 1. Overview of the ARCADE science platform for ALMA data reduction and analysis from the users' view point.

The diagram in Figure 1 provides an overview of the ARCADE system from the users' perspective. After login (authentication) and permissions check (authorization), users interact with the science platform via interactive desktop interface within their local browser. The authorization system provides the ability to limit access to a particular part of the file system (e.g., home directories, project directories) to particular users or groups (e.g., collaborations). This allows users to store and connect with their own data as well as that of their groups.

ARCADE is deployed in a scalable, stable research cloud computing environment provided by Compute Canada. This provides tools for high-performance data analysis, along with cloud resources to scale computation. ARCADE's scalable storage resources and high-speed network provides streamlined access to large, complex, or otherwise inaccessible data in the cloud. Sessions are preloaded with every version of CASA along with other software designed for ALMA users (e.g., CASA, CARTA, ADMIT). The environment is customisable to suit the needs of users and their groups; if software can be run on a linux desktop, it can be run in ARCADE. Lastly, ARCADE is built entirely on open-source infrastructure, with no vendor lock-in and no proprietary tools.

A screenshot of the ARCADE science platform running in a user's local web browser is shown in Figure 2. Here the user is performing image reconstruction of ALMA data using the task tclean in CASA (v5.7) with the metadata displayed in the CASA Logger graphical user interface (GUI). On the right, the user is editing a pipeline processing script using vim while the content of the $HOME and project/ directories are shown in the terminal. In the ARCADE filesystem, $HOME is used for permanent storage of user files (e.g., scripts, software, README files), while projects/ contains temporary storage allocations (per project) where group data is (temporarily) stored and processed. See the Example ARCADE Use Case section for a more detailed walkthrough of ARCADE.
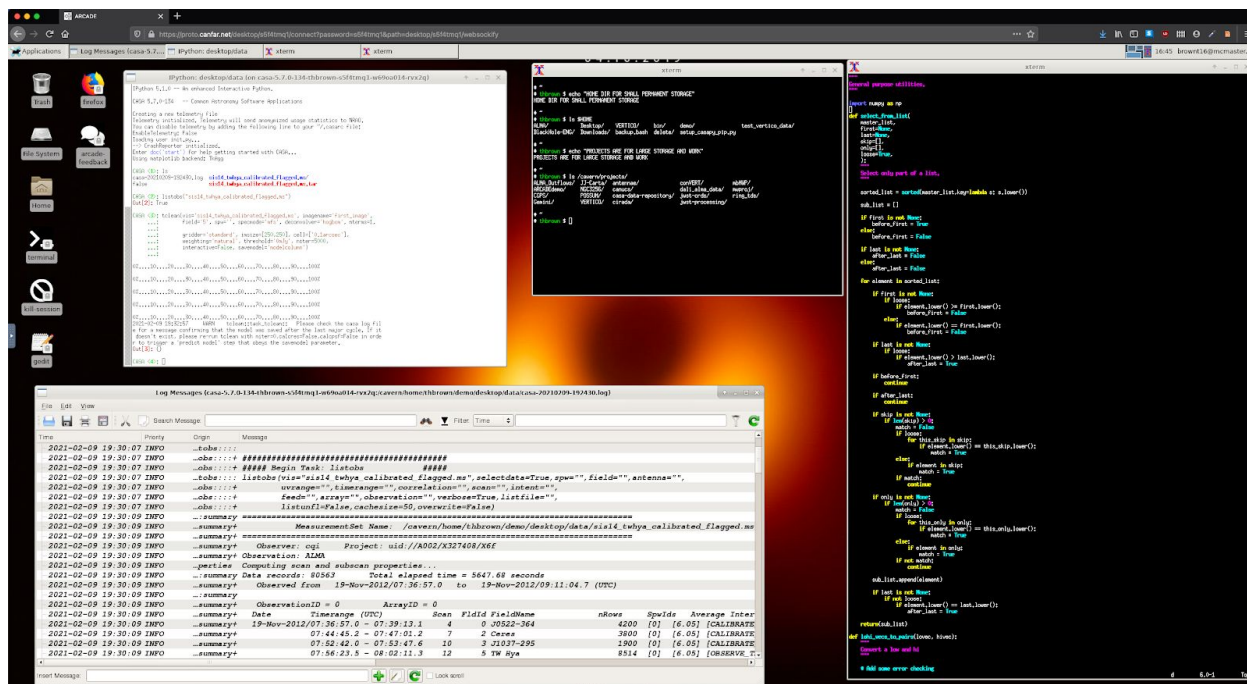

Figure 2. A screen grab of the ARCADE platform desktop as viewed in a local browser. The open software containers are CASA 5.7, the CASA Logger GUI, terminal, and the vim text editor.

Organisational Structure and Personnel

The ARCADE Development Study was completed as a partnership between the National Research Council of Canada (NRC) and McMaster University. The project was managed by NRC, which then provided subcontracts back to McMaster University who provided external science user testing. ARCADE was developed as a new science platform within the Canadian Advanced Network for Astronomy Research (CANFAR) platform, an integrated storage, processing and group management ecosystem deployed primarily on Compute Canada Federation resources and serving the needs of the Canadian astronomy research community. CANFAR is developed and operated by NRC's Canadian Astronomy Data Centre (CADC) which also provided the development effort for the ARCADE project.

Table 2. List of Personnel for the ARCADE Development Study

Title	Key Personnel	Institution	FTE	Duration (months)
Principle Investigator[1]	Toby Brown (July 2020 - February 2021)	National Research Council of Canada	0.17	7
Principle Investigator	Helen Kirk (Dec 2019 - June 2020)	National Research Council of Canada	0.17	7
Science Staff - Internal	Several	National Research Council of Canada	0.08	12
Development Staff	Several	National Research Council of Canada	0.56	12
Operational Staff	Several	National Research Council of Canada	0.19	12
Science Staff - External[2]	Several	McMaster University	0.38	12

Study Management

We assembled an interdisciplinary team of developers, scientists, and operations staff to deliver the ARCADE Development Study. The full team met once per week, led by PI Brown (July 2020

---

[1] PI Kirk went on parental leave starting July 6 2020. Co-I Brown was hired to fill the role and assumed his position as of July 13 2020.

[2] Four part-time (graduate student) and one full-time (co-op student) beta testers were hired to replace co-I Brown's role as a beta tester. They were based at McMaster University.

- February 2021) or PI Kirk (pre-July 2020), with sub-teams meeting on an ad-hoc basis. The development and technical solutions were led by the CADC. Throughout the study, the system was constantly tested by science users external to the development team. To ensure a wide variety of use cases and distance from the development team, system testers were based at both NRC and McMaster University. Project management and scheduling by the operations staff helped ensure the study remained on time, within budget, and in scope.

ARCADE Software

ARCADE is built entirely on open-source source software and designed so that it can be scaled out to meet the storage and processing demands of the user. At a minimum, it has the functionality to calibrate and reduce ALMA data. It should also be stable, easy to maintain, and portable across a range of computing infrastructure. Finally, the ARCADE software should be fully integrated with existing ALMA value-added software, CARTA and ADMIT. This section describes the key software components of the system.

Technologies

The ARCADE beta deployment uses open-source technologies that were tested and successfully incorporated throughout development. Here we summarise the key technologies that underpin the ARCADE system.

Software Containers

A container is a lightweight, isolated unit of software that packages up code and all its dependencies so the application can be deployed quickly and reliably from one computing environment to another. It is this approach that gives ARCADE its stability and portability.

Kubernetes

Kubernetes is a system for automating deployment, scaling, and management of containerized applications. Targeting Kubernetes provides ARCADE with flexibility on the computing environment and storage system of choice, whether that is a local machine (e.g., research cluster) or cloud infrastructure (e.g., Compute Canada, Google Cloud).

CephFS Distributed Storage

ARCADE is built upon the open-source Ceph file system (CephFS) to enable a variety of use cases including shared home directories, high-performance computing scratch space, and distributed workflow shared storage. CephFS is a completely distributed storage solution without a single point of failure and scalable to the exabyte level.

System Components

The ARCADE platform consists entirely of containers executing on a Kubernetes installation. There are two categories: system containers and scientific software containers. The system containers provide users an interface to the various features of the platform, mainly session management, storage access, and access control. The scientific software containers are built to run specific versions of CASA. Since all components are containers additional instances of the computing environment can be quickly supplied with Kubernetes allowing easy scaling to the available compute, storage, and I/O resources to meet user demand.
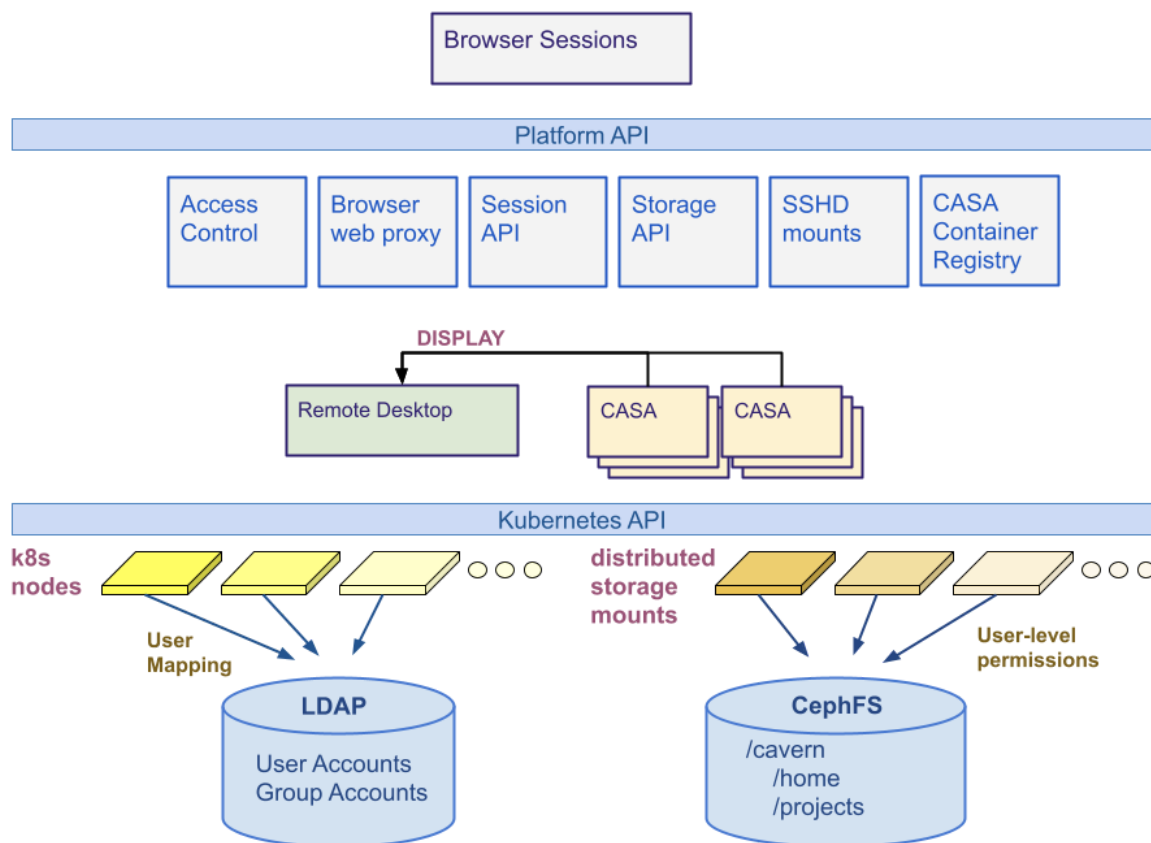

Figure 3. ARCADE software architecture diagram from the implementation view point.

Session API

The CANFAR container Session API module provides the controls for creating, listing, and terminating interactive desktop browser sessions. It also allows users to attach (through remote display) numerous versions of CASA to those desktop sessions.

Browser Web Proxy

The browser web proxy module allows users to connect to desktop sessions with their web browser. The web proxy serves all sessions over the standard HTTPS/SSL port, ensuring access even from institutions with strict firewall rules.

Remote Desktop

The remote desktop is a container representing an ARCADE session. It is built on noVNC and VNC technologies, and has been customized with shortcuts that make attaching CASA containers easy for users.

CASA Container Registry

A software registry exposing Docker container images for every operational distinct version of CASA between 3.6 and 5.7. Users can launch any number of these containers with the displays output automatically routed to the desktop session. Even though the CASA terminal is displayed in the desktop session, it is in fact running elsewhere on the Kubernetes cluster, so does not share resources with the remote desktop or other CASA containers attached to the desktop session.

Storage API

Each scientific container has access to the same common file system, cavern (deployed on CephFS). Within cavern are users' home directories and project directories for sharing data. Data in cavern is persisted between remote desktop sessions and CASA container launching.

The cavern user storage system implements the International Virtual Observatory Alliance (IVOA) VOSpace (distributed storage interface) specification. This means that there is a standard interface for accessing data from outside the CANFAR platform (e.g., from local computing) with standard client data download and upload tools, (e.g., the CADC vos python tools). Additionally, cavern can be mounted on a local machine (with e.g., sshfs) for interaction as if it were a local directory, enabling the use of standard linux tools for data transfer (e.g., rsync).

Access Control

The access control system in the ARCADE platform is fully integrated with the CANFAR access control system. One key feature is the ability to share files and directories through groups and the [IVOA Group Membership Service](). Users have the ability to create and manage groups for assigning read-only or read-write access to data they wish to share.

Deployment

The ARCADE platform consists of a set of system containers that are deployed in the container orchestration system, Kubernetes. The platform itself interacts with Kubernetes to launch the CASA software containers to meet users' requests.

Development Workflow

The CADC was responsible for software development across the different components of the project. Frequent team coordination took place via a weekly meeting, GitHub, and Slack. The project was divided into alpha and beta stages. In each stage, the lead development team (Major, Yeung, Fabbro) created code consistent with the architecture outlined in the study proposal and detailed in Figure 3. Once functionality was in place in the system, the software was tested by the scientific team, providing feedback to the development team.

ARCADE was developed using the Git version control system which provides software project management functionality. The ARCADE development team operated under an "agile" methodology where units of work are scheduled in three week iterations.

Development engineering and design activities were conducted in accordance with established CADC policies, practices, and procedures. All shared documents are dated and bear a revision level number. GitHub (code, documentation) and Google Drive (reports) were used to maintain a revision history for all documents.

Development Phases

ARCADE was developed in two stages, successfully building on the functionality of an existing prototype. The following summarizes the development phases that took place in the main portion of the funded activities:

Alpha

The alpha development and testing was for proof-of-concept. It allowed users to launch different versions of CASA and have them all available on their remote desktop session. The runtime system underneath (Docker) and storage system (single mount volume) was simple and not intended to scale.

Beta

Beta development and testing focussed on the infrastructure underlying ARCADE. A distributed storage with user permissions integrated at the file system level (cavern) replaced the single volume. The runtime execution of containers was migrated from plain Docker to Kubernetes.

Although beyond the scope of this study, the CADC is continuing development of two more Kubernetes clusters with infrastructure improvements that were recommended based on the experiences of Beta testers. These two clusters are isolated from each other so that one can be used for testing new features without disrupting the scientific work happening in the second.

Deployment on the Commercial Cloud

Once the core Beta system functionality was fully developed and the system was deployed on the research cloud (Compute Canada), ARCADE was then deployed on the commercial cloud. This was designed as a test and demonstration of ARCADE's portability.

The chosen commercial cloud provider was the Google Cloud Platform. This choice reflected the ease of account set up and availability of free credits rather than preference of provider. A Kubernetes cluster with storage, network, load-balancing, and public-facing networking was set up to accommodate the ARCADE deployment. A significant number of modifications to the existing deployment scripts and software container build scripts were required, along with modifications of ARCADE's authorisation systems, moving to a token-based authorisation system currently in beta. A working deployment was achieved and used to reconstruct an image from an ALMA measurement set using CASA 5.7 task `tclean`.

Testing Program

The ongoing tests of ARCADE were divided into two categories: software testing and user testing. The former is a verification and validation process that was conducted by developers during development, evaluating all new code to detect differences between the expected output and the actual output. This was managed via the GitHub ecosystem, with development and testing taking place on the development branch and pull requests undergoing review before code was merged into the main branch.

Science user testing began once stable CASA software containers were available on the alpha system and continued on the beta system until study completion. Testing was undertaken by internal and external science teams to ensure a wide range of use cases. The goals were to test the onboarding process, system functionality, reliability, and user experience. User testing was ongoing during the development, allowing the PIs to track feedback as the system progressed from alpha into beta. Issues and feature requests were primarily managed via the GitHub repository Issues page with frequent, less formal communication between users and developers taking place on Slack. Users were asked to complete four user surveys. The first was completed approximately two weeks after account creation, the next two were completed after

approximately several months and six months of ARCADE use. The final survey was completed in February 2021 as the beta user testing phase of study was concluded.

The final user survey found that 100% of ARCADE users would recommend the platform to other ALMA users. Where the processing speed was highlighted as an area for improvement in the mid-term surveys, 80% of users rated it as sufficient in the final survey (note that the beta system was never deployed on infrastructure scaled to production performance). The availability of all CASA versions combined with the large amount (50Tb) of reliable, shared storage was reported as a positive. The main challenges reported by science users were the user experience of the linux desktop environment and the stability of the system in early development stages.

Example Use Case

The data used in this example use case are from ALMA Project 2018.1.00493.S, "Probing CN excitation and abundance in dense gas in luminous infrared galaxies", for which the PI is Christine Wilson. Part of the data for this project has been published in Ledger et al. (submitted 2021).

The objective of this use case is to image the HCN (J=3-2) emission line structure in the Yo-yo Galaxy, IRAS 13120-5453, which is bright and well suited for demonstrating imaging with ARCADE. The data are observations from ALMA's 12-m array taken during Cycle 6.

We calibrate the data using the delivered calibration script (scriptforPI.py) provided alongside the raw data in the ALMA archive. This demonstrates how ARCADE can readily be used to calibrate ALMA data without the installation of additional software because all CASA versions are preinstalled in containers. To make the data set more manageable, we have reduced it in size by averaging in frequency.
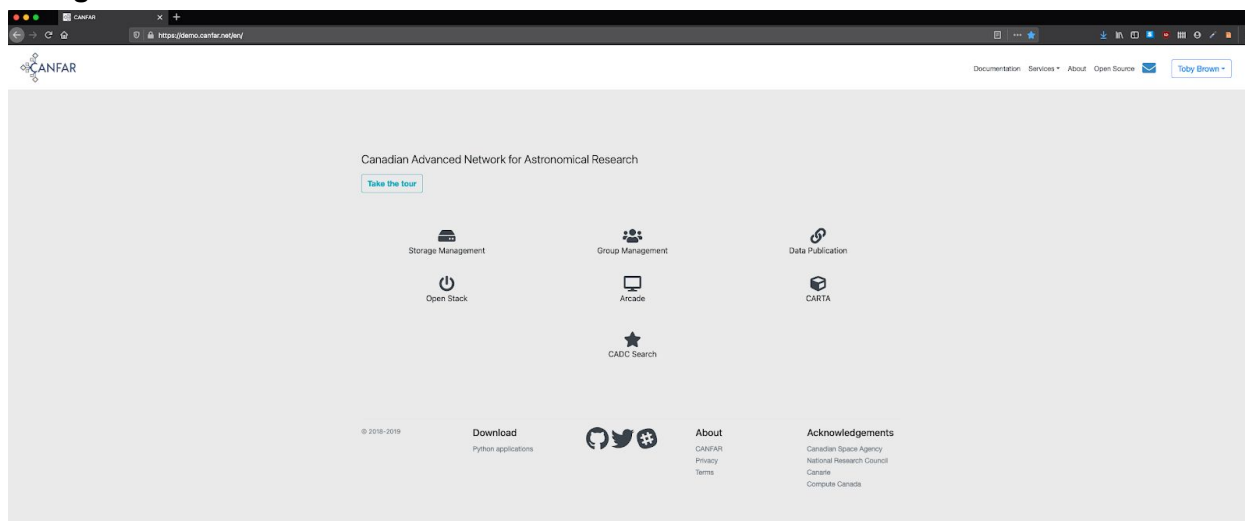
1. Log into ARCADE and download data

Figure 4. CANFAR landing page where the user logs in via the ARCADE icon in the centre of the screen.

Once the dataset is identified in the archive, the user logs into ARCADE and the remaining work takes place entirely within an ARCADE session in the user's local browser.

The dataset is downloaded using the <u>astroquery</u> python module.

```python
from astroquery.alma import Alma

# query archive
tab = Alma.query_tap("select * from ivoa.obscore where proposal_id like
'2018.1.00493.S' and target_name like 'IRAS13120-5453'").to_table()

# download and extract
Alma.download_and_extract_files(list(tab[access_url]), regrex='.*\\.tar$',
include_adsm=True)
```

Note, a common alternative approach is to copy the archive download script into the ARCADE platform via the <u>vos Python module</u> command, `vcp`. Standard linux tools for sshfs mounting and rsync also provide this functionality.

Data are downloaded to `../projects/ARCADEdemo/`. Users can create project directories for themselves, their groups, and their communities. Access and rights are easily customisable by the directory owner using standard linux tools.

2. Calibrate data

For this dataset, calibration takes place in the correct CASA version (in this case CASA 5.4) container launched from a menu in the desktop environment. We are able to seamlessly open and run the scriptforPI.py in the CASA version required by any given dataset.
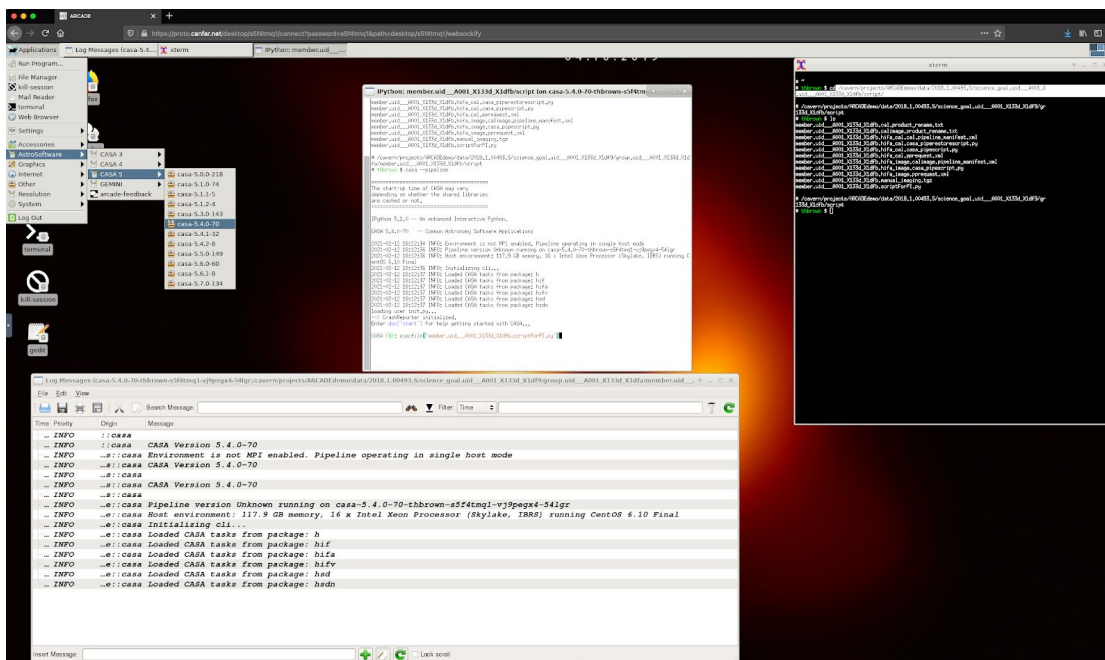

Figure 5. Calibrating data using CASA 5.4 in ARCADE. At this step, ALMA users need the right version of CASA to be installed. ARCADE is preloaded with every available CASA version.

3. Radio interferometric image reconstruction

The calibrated data are then continuum subtracted and imaged in an up-to-date CASA container (currently CASA 5.7) using tasks `uvcontsub` and `tclean`. ARCADE has full support for CASA GUIs (e.g., the viewer, plotms) and CARTA.

Figure 6 shows the user imaging the data (after continuum subtraction), with the CASA Viewer GIU displaying the masked regions during the cleaning process. The imaging script is open for editing on the right hand side of the desktop while the Logger is shown at the bottom.
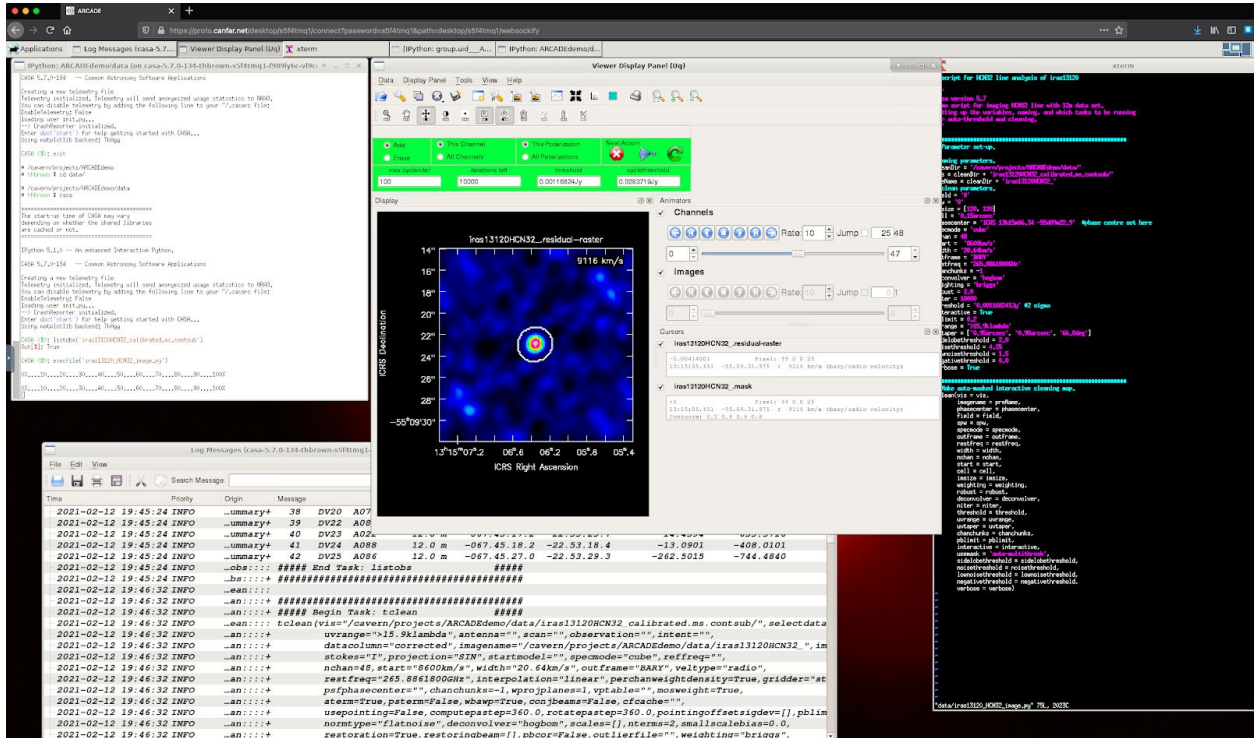
Figure 6. Radio interferometric image reconstruction using CASA on ARCADE.

The resulting data cube can then be explored by the user (or anyone with the correct permissions) in a CARTA session.
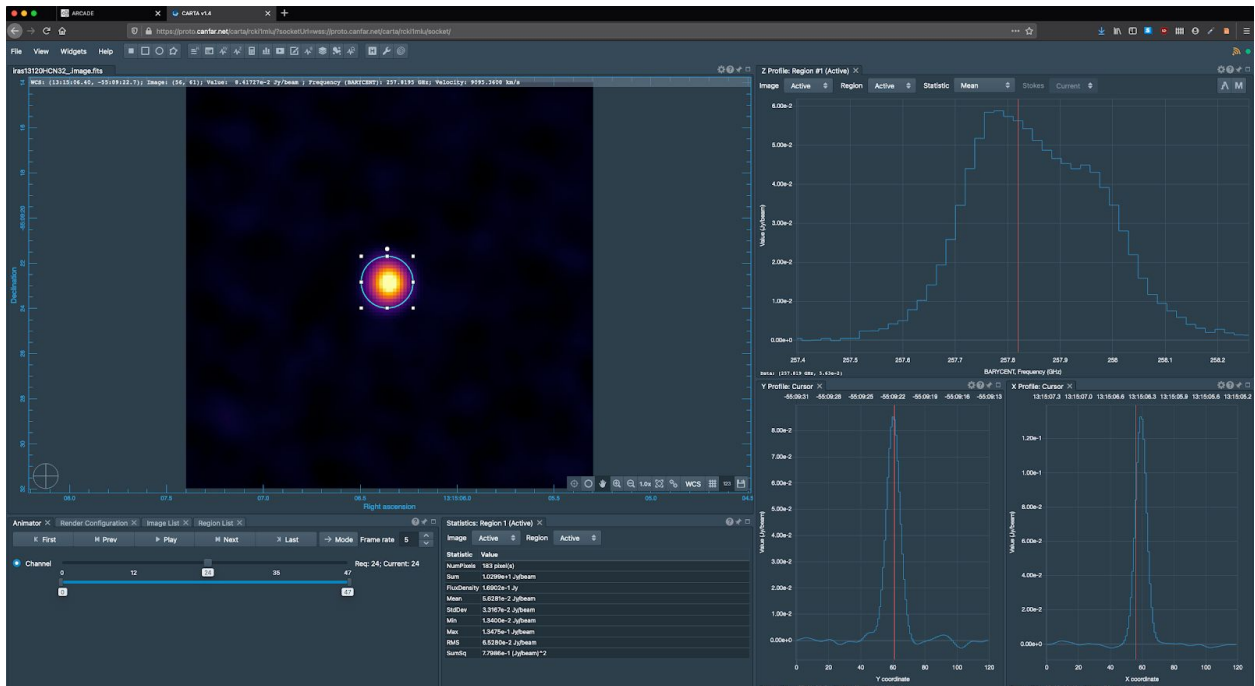

Figure 7. Visualization of the image on ARCADE using CARTA.

It is also possible to conduct imaging and analysis with CASA 6.1 in Jupyter Notebooks.. Note that this functionality is still under development and beyond the scope of the development study. See the Future of the ARCADE Platform section for more details.

Performance Review

This section reviews the performance of the ARCADE development study with respect to original specifications, schedule, and budget outlined in the proposal. Overall, ARCADE has met or exceeded all specifications. The initial choice of storage technology caused stability issues around the midpoint of the development study. To ensure these delays in achieving a stable beta system did not adversely affect the study outcomes, the study was extended by two months with no change to specifications or budget. This decision was taken by PI Brown with approval from NRAO.

Performance to Specifications

As defined in the proposal, we have developed and deployed a beta version of ARCADE. This consists of a web-based portal into a cloud computing environment designed for use with ALMA data. This beta deployment accommodates approximately 12 simultaneous users.

With limited resources and small number of users on the beta system, the decision was taken early on to not enforce independent resource limits, however, this will be straightforward to implement in the next ARCADE installation.

The system is a collection of interacting software containers deployed on the Kubernetes platform in a cloud computing environment. These containers support the system itself and provide users with every available version of CASA (i.e., from CASA 3.2 onward). The CASA containers are publicly available. Additionally, we provide user-controlled secure storage, which integrates with a scalable cloud storage system through CADC's existing user storage architecture, with access managed through the ARCADE user accounts. In a demonstration of portability, the ARCADE platform has been deployed on both a research cloud (Compute Canada) and a commercial cloud vendor (Google Cloud). If desired, we will be glad to provide accounts for the project review panel to access and explore this beta on the research cloud.

The following measures of success defined in the proposal have been either met or surpassed by the study.

A scientifically knowledgeable user can successfully:
1. log in and access a personal storage space where data can be stored in a stable and secure manner for the duration of the project.
2. share datasets in a seamless way between ARCADE users.
3. launch their preferred version of CASA.

4. interactively reduce and image ALMA data, do a basic science examination of the data products, and persist the output files to their secure personal storage space.
5. give controlled access to collaborators to the output files in their secure personal storage space.

The beta prototype is:
1. functional and all available containers in ARCADE are executable (all software runs properly).
2. deemed by beta testers to be straightforward to use, with any major usability or performance issues documented and a procedure identified to rectify or mitigate those issues where possible.
3. deployed in a scalable manner in a cloud environment based on the technical aspects studied (i.e., the user secure storage system, the orchestration system and the container maintenance procedure).

Performance to Schedule

PIs Kirk and Brown held primary responsibility for schedule development and performance to schedule. On December 2nd 2020, the NA ALMA Development Program Office and NRAO approved a no cost extension to the completion date from December 16 2020 to February 19 2021. The propagation of delays to system deployment occurred during the year, compounded by working practices during COVID-19, meant the study would be better served with more time to properly deploy and test the ARCADE system. There were no other scheduling issues to address.

Performance to Budget

PIs Kirk and Brown held primary responsibility for the study budget allocation and cost performance. In collaboration with operations staff at NRC and McMaster, PIs Kirk and Brown prepared and submitted quarterly billing reports to the NA ALMA Development Program Manager. These reports detailed costs, budget load, and budget status. There was no significant cost variance and the study was kept within budget.

Study Closeout

Along with this report, the ARCADE study has produced several products for delivery to NRAO. The delivered products are:

- Full access to the open-source ARCADE software[3]
- Software documentation for developers
- CASA software containers
- Container maintenance and deployment documentation (Deliverable 8)

---

[3] For organisational reasons, the ARCADE software repository is named "skaha".

- [User documentation and tutorials](#)
- Status, billing, and final reports

We have concluded the beta development stage and are at software version 0.3 at time of study closeout.

Lessons Learned

The ARCADE development study highlighted a number of important lessons in the design and development of a scientific computing platform for ALMA users.

Computing Environment

ARCADE provides HTTPS-based VNC access to a desktop computing environment via noVNC. User authorisation and connection are handled by the CANFAR web user interface. Once connected, users see a linux desktop GUI that is preloaded with all the software needed to transfer, download, calibrate, and process ALMA data. Documentation on building and maintaining software containers is also provided.

The primary reason for the desktop environment was support for CASA GUIs and the "interactive clean" functionality which requires the CASA Viewer in all currently released CASA versions. Users' familiarity with linux desktops also helped ease onboarding. However, there are several drawbacks to the desktop environment. Users highlighted issues with writing code, inspecting results, and getting rich outputs from the desktop (see Appendix C), with the lack of a fluid user experience also an issue. The reliance on noVNC software means that noVNC issues automatically become an issue in ARCADE. In particular, the mechanism for copy/paste - a critical part of the ALMA data reduction workflow - within a noVNC session is awkward and outdated.
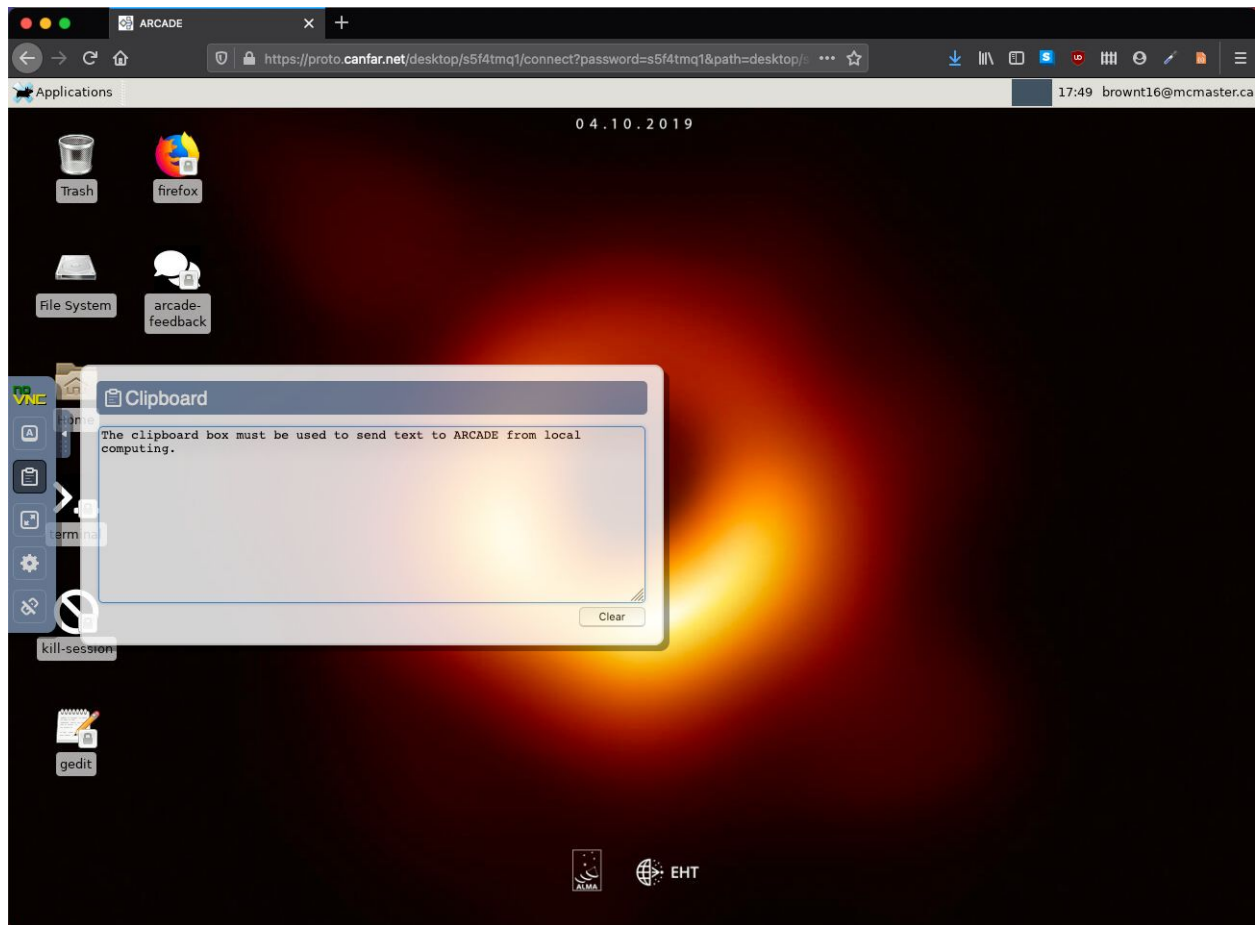
Figure 8. Text copied in the remote desktop will appear in the clipboard box in noVNC's interface. The user must then copy the text from that box to access it in your local clipboard. Local text must be put into the clipboard box to be sent to the remote clipboard.

There are alternatives to running a desktop session on ARCADE infrastructure that would provide almost equivalent functionality in a flexible and powerful web user interface. In particular, ARCADE is capable of running JupyterLab, the next-generation user interface for Project Jupyter. It offers all the familiar building blocks of the classic Jupyter Notebook (notebook, terminal, text editor, file browser, rich file outputs, etc.) while documents and activities integrate with each other, enabling new workflows for interactive computing. Development would need to be expanded considerably in order to fully integrate CASA and CARTA into an ARCADE JupyterLab environment. This functionality is demonstrated in Figure 9.
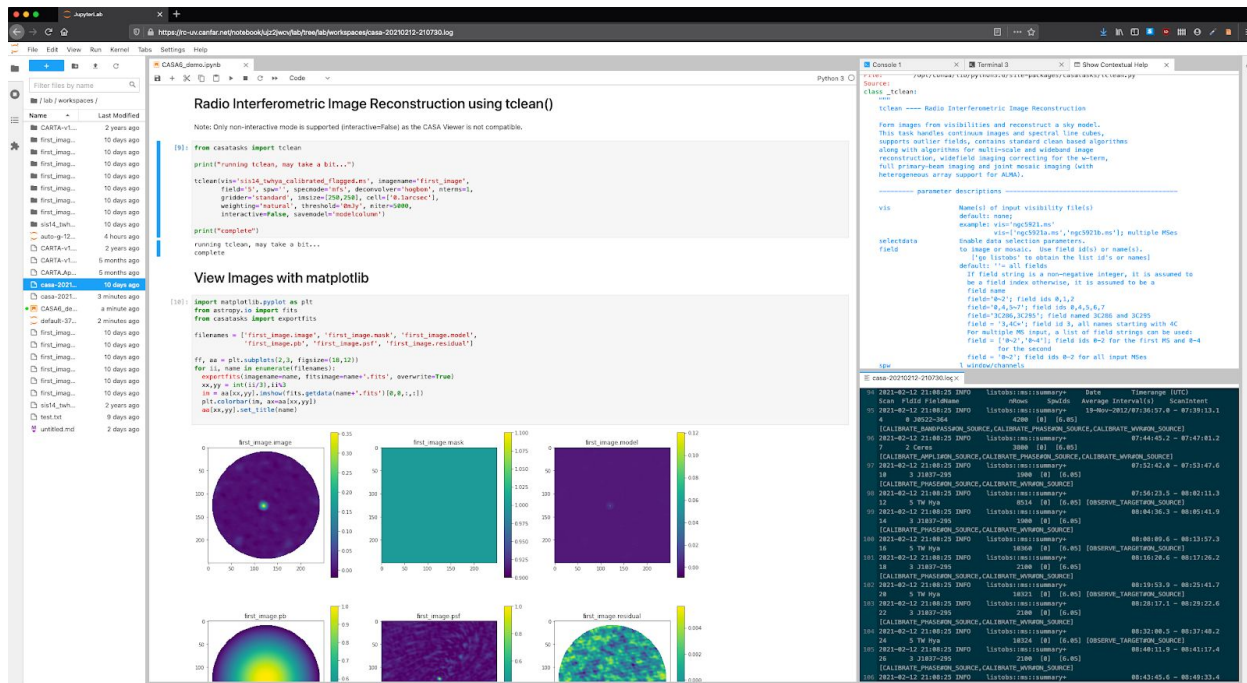
Figure 9. JupyterLab deployed on ARCADE showing image reconstruction with CASA 6.1 in the ipython notebook, contextual help for the tclean task in the top right, logger in the bottom right, and the file browser pane on the left.

Portability

In terms of lessons learned, the effort to deploy ARCADE on the commercial cloud was one of the most valuable to the study.

As long as portability is a requirement for ARCADE, deployment, maintenance, and packaging of software with the commercial cloud in-mind should be part of any solution. One strength of the major commercial cloud vendors is the wide-variety of available solutions, including storage systems, network security, authorisations, and load-balancing. From a practical standpoint, vendors are experienced in servicing existing science platforms (e.g., Pangeo), extensive documentation is always present, issue resolutions are easily searchable, and support teams available. Finally, the reliability and availability of commercial cloud platforms is extremely high, even in comparison to research cloud providers.

Caution must also be exercised before full deployment on a commercial cloud, and issues can arise, usually in relation to financial cost. Targeting a production scale commercial cloud deployment would require significantly more effort on the system architecture to minimise costs and take advantage of existing services. Processing of large Tb scale ALMA datasets in particular would require users to pay close attention to data transfers into and out of the cloud and the limits on automatic workload scaling. If it is not possible to design a system that

minimises data transfer costs, pay-as-you-go payment models and vendor "credits" for research computing would be the most likely option. However, they are unlikely to remove the risk of incurring significant computing costs for all users. For example, at the 2020 price, cube cleaning with 16-cores costs a user approximately $1 per hour, assuming the data remains in the cloud.

Stability

During development and testing the system was unresponsive at times. A large portion of the containers running in Kubernetes would stop responding. After some investigation the reasons behind these issues were discovered.

One problem was with an experimental driver to CephFS, the distributed storage system. This driver, Rook, would consume nearly all the system resources and starve out many of the scientific containers. It was decided to use a different driver in the two new Kubernetes installations being built. Although this was foreseen, it was problematic to do system development on the Kubernetes cluster that was being tested at the same time. Of the two new ARCADE installations beyond the beta, one will be dedicated to development so new features can be tested in isolation before being moved to the stable Kubernetes production installation.

Workflow

The lack of integration between the ALMA archive web interface and the ARCADE platform meant users had to download the archive wget script to their local computer and transfer this back to ARCADE or use the `astroquery` python module. During testing, it became apparent that the ideal ALMA user workflow would be from archive to ARCADE to science, without the need to route via a local machine. In this case, ARCADE integrated with CARTA would provide an end-to-end platform from archive to science.

Group Management

Throughout the study, the potential for collaboration with ARCADE became readily apparent with beta users sharing their data, raw and reduced, and analyses with supervisors and collaborators within the platform. The CADC Group Management System (GMS) allows users to create groups, manage the group's membership and assign groups read or read/write access to storage. For technical reasons, the GMS was only partially integrated into ARCADE and further work is required to complete this integration.

Federated Identities

To use the ARCADE system, a user needs a CADC account, which is distinct from a user's ALMA account giving them access to proprietary data. Another area of improvement is to enable Single Sign-On between CADC and the ALMA Science Archive from within an ARCADE session, thus freeing users of the need to embed their ALMA username and password in

scripts, scripts that are often shared with team members. This would require a technical solution in the ALMA Science Archive to provide that capability and for ARCADE to use that functionality.

Appendix A: Offerer's Cost Breakdown

Table 4. Original offer's cost breakdown for the work breakdown structure

WBS No.	Task Description	Labour (US$)	Materials & Service (US$)	Travel (US$)
1	**Commercial cloud concept deployment.** Set up commercial Kubernetes cloud deployment of ARCADE for use as a concept testing platform to investigate various ideas without interrupting alpha or beta testing.	31,299	1,000	
2	**Alpha prototype release.** Provide alpha prototype version of ARCADE on scalable Kubernetes on the research (Compute Canada) cloud. Develop a basic alpha prototype tutorial.	8,722		
3	**CASA container testing.** Add and fully test all CASA versions on the alpha ARCADE prototype. Install other ALMA software (e.g., ADMIT, CARTA).	20,866		
4	**Alpha storage solution.** Explore storage solutions to achieve a secure and performant user personal storage space with data sharing.	31,299		
5	**Alpha testing.** User testing of the alpha ARCADE prototype. Identify major problems and report them. Includes travel for Co-I Brown to travel to HAA for a week of training.	8,722		3,000
6	**Beta development.** Beta prototype development. Incorporate all easily addressed issues raised by the alpha users testing. Revise the alpha prototype tutorial for the new beta prototype system.	20,866		
7	**Beta user testing.**	52,546		

	In-depth testing of usability and performance of the beta prototype version of ARCADE for a variety of science use cases. Includes the introduction of new users to the system to test the beta prototype tutorial.			
8	**Container maintenance.** Examine container maintenance and upgrade issues. Document possible solutions.	5,217		
9	**Final Report.** Prepare the final report, with input from development, operations, and science team members.	6,955		
Subtotals (US$)		**176,790**	**1,000**	**3,000**
Total Offerer's Cost (US$)				**180,790**

Appendix B: Study Schedule

Table 5. Study schedule showing the planned, revised, and actual date of deliverable completion.

WBS No.	Task Name	Planned Date[4]	Revised Date[5]	Actual Date	Status
1a	Concept Deployment in Research Cloud	November 5, 2019	January 16, 2020	January 16, 2020	C
1b	Beta Deployment in Commercial Cloud	November 5, 2019	January 20, 2021	January 8, 2021	C
4	ALPHA Storage solution	January 27, 2020	February 27 2020	February 24, 2020	C
2	ALPHA Prototype release	December 17, 2019	April 30 2020	April 6, 2020	C
3	CASA Container Testing on ALPHA	November 8, 2019	June 16 2020	June 12, 2020	C
5	ALPHA Testing	February 18, 2020	June 16 2020	June 12, 2020	C
6	Beta development	April 3, 2020	October 09 2020	October 1, 2020	C
7	Beta user testing	August 24, 2020	January 19 2021	February 18, 2021	C
8	Container Maintenance	September 1, 2020	February 04 2021	December 18, 2020	C
9	Final Report	September 30, 2020	February 19 2021	February 19, 2021	C
10	ARCADE Project Complete	September 30, 2020	February 19 2021	February 19, 2021	C

---

[4] Planned Dates are based on the original project proposal.

[5] Revised Dates are based on the actual project start date of 16 December 2019.

Appendix C: ARCADE User Feedback

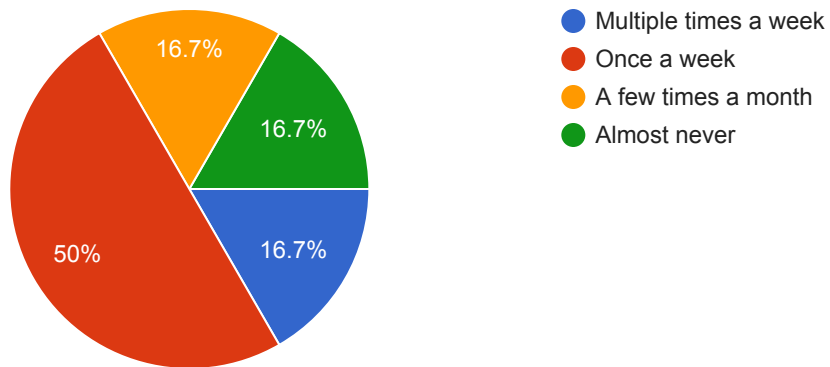Form responses for the final, both midpoint, and initial user surveys are included below.
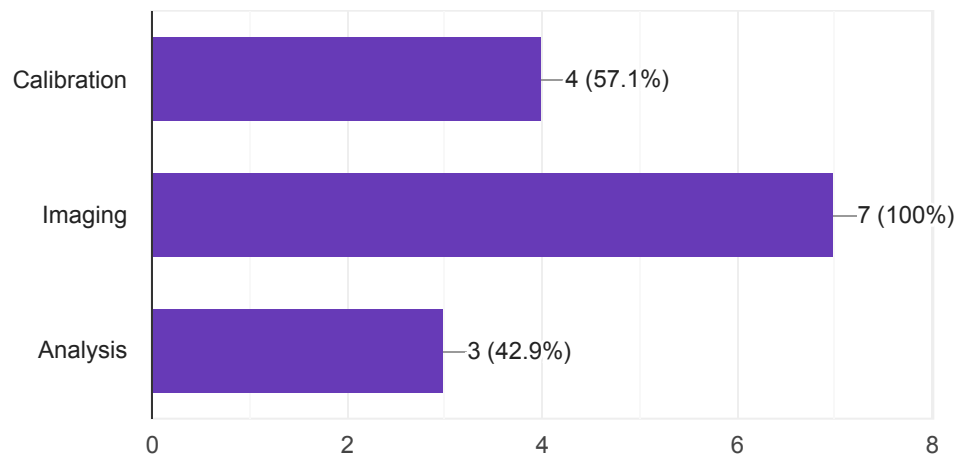
ARCADE User Survey 4

7 responses

Publish analytics

How often are you using ARCADE for your work?

6 responses

- 🔵 Multiple times a week
- 🔴 Once a week
- 🟠 A few times a month
- 🟢 Almost never

What are you using ARCADE for? (select all that apply)

7 responses

Calibration	4 (57.1%)
Imaging	7 (100%)
Analysis	3 (42.9%)

Please provide an estimate of your storage space usage so far

7 responses
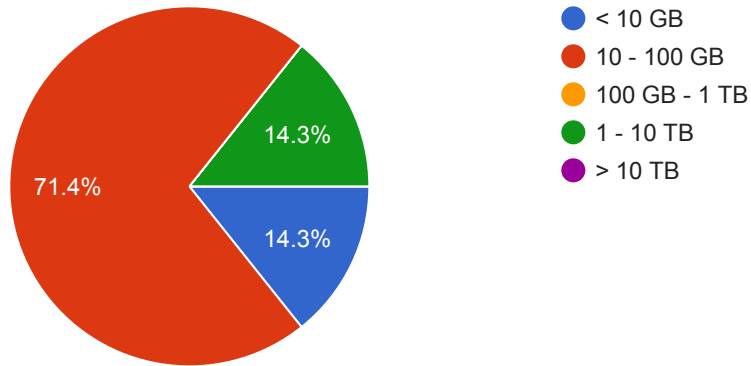
Legend:
- < 10 GB
- 10 - 100 GB
- 100 GB - 1 TB
- 1 - 10 TB
- > 10 TB

Pie chart values: 71.4%, 14.3%, 14.3%

How do you transfer data into and out of ARCADE (select all that apply)

7 responses

Category	Value
ARCADE browser	4 (57.1%)
ARCADE terminal (download scripts)	3 (42.9%)
vos tools (vcp, vsync etc)	5 (71.4%)
File system mount + rsync	4 (57.1%)
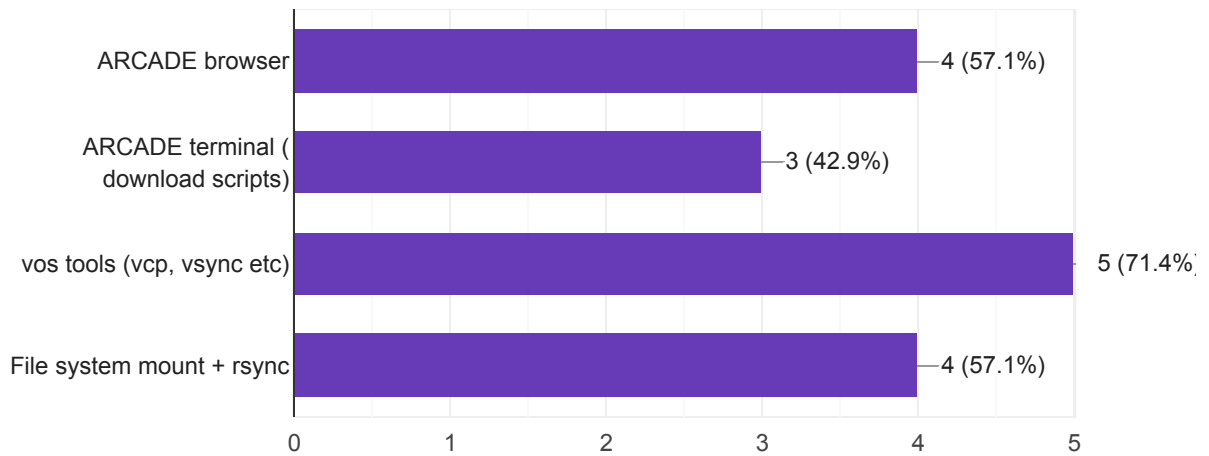
How would you rate the data transfer speed into and out of ARCADE (download/upload/vcp/rsync)

7 responses
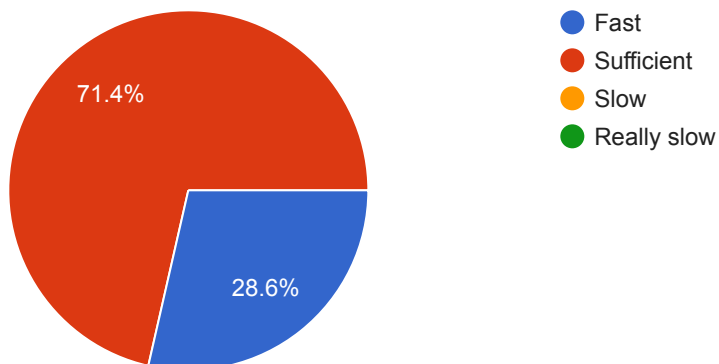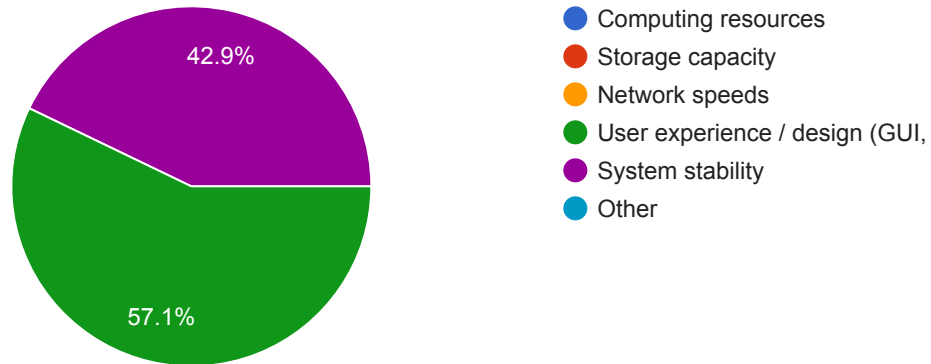
- Fast
- Sufficient
- Slow
- Really slow

71.4%

28.6%

How would you rate the processing speed in ARCADE compared to non-ARCADE systems (e.g., running calibration or imaging scripts)

7 responses

- Fast
- Sufficient
- Slow
- Really slow
- N/A; no comparative experience

14.3%

14.3%

71.4%
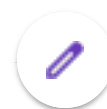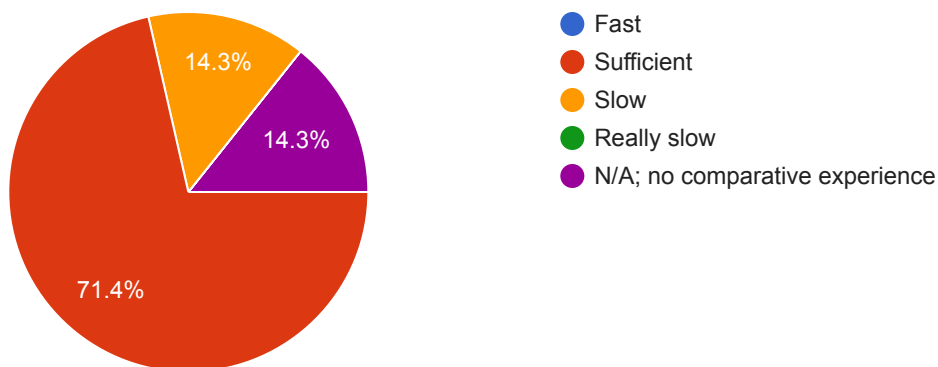
Where do you think ARCADE could improve most?

7 responses

- Computing resources
- Storage capacity
- Network speeds
- User experience / design (GUI,
- System stability
- Other

42.9%

57.1%

Please list any aspects of ARCADE that you find particularly useful / great

6 responses

Great for localizing all data associated with imaging in one place. Multiple versions of CASA.

I love the fact that it's allowed me to be able to switch between multiple versions of CASA without issues!

- ability to access almost every tool needed to process and analyze ALMA data
- sufficient storage
- ability to centralize/organize projects
- easy to access shared data and collaborate

I like how quick and easy it is to use an ARCADE session in combination with a CARTA session. It is very straightforward to have both sessions open and working simultaneously rather than having to navigate using the viewer.

Web browser interface is very convienient and rarely lags.

Would you recommend ARCADE to another ALMA user?

7 responses

- Yes
- No

100%
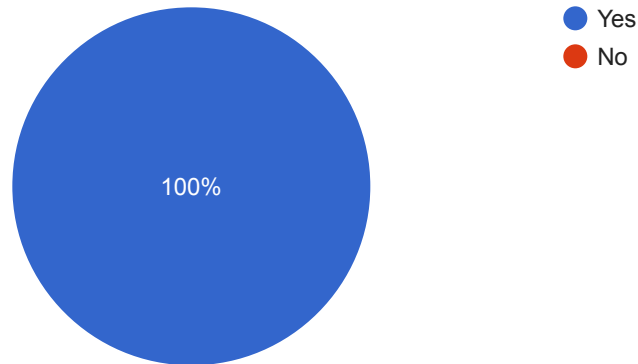
What features would you like to see in the future?

6 responses

Better working copy/paste functionality. More explicit documentation regarding file permissions.

An easier interface for beginners

The ability to work in a JupyterLabs-style environment would be most helpful; being able to combine a terminal, reduction in CASA, and analysis/visualization with Python and CARTA could increase productivity and the overall user experience greatly

I would like to see more options for a text editor i.e. emacs.

Better stability of the web browser, more fully featured desktop applications or the ability to install them.

Thanks for answering our survey! We welcome any additional feedback below.

2 responses

Currently the Jupyter notebook feature of ARCADE draws from a different pool of installed modules than say an IPython console on ARCADE desktop. It would be helpful to have the same modules accessible and/or use virtualenvs in tandem with notebooks

I haven't used ARCADE in a while, but I found that the documentation to transfer data directly to ARCADE via a mount + rsync was straightforward to follow from the GitHub documentation. I would like to see some more simplification and features for a Windows user as well in the future.

Google Forms

ARCADE User Survey 3

5 responses

Publish analytics

How often are you using ARCADE for your work?

5 responses
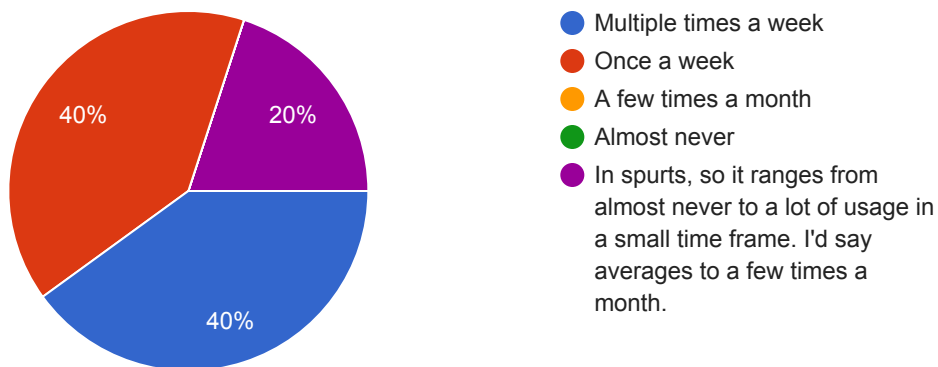
- Multiple times a week
- Once a week
- A few times a month
- Almost never
- In spurts, so it ranges from almost never to a lot of usage in a small time frame. I'd say averages to a few times a month.

40% 40% 20%

What are you using ARCADE for? (select all that apply)

5 responses

Category	Value
Calibration	3 (60%)
Imaging	4 (80%)
Analysis	3 (60%)
Script and paper writing.	1 (20%)
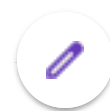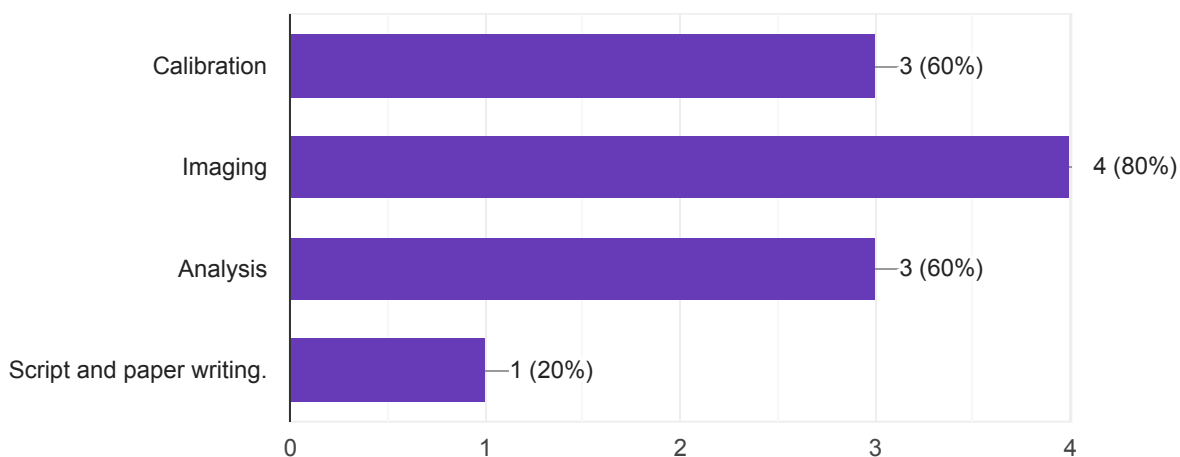
Please provide an estimate of your storage space usage so far

5 responses
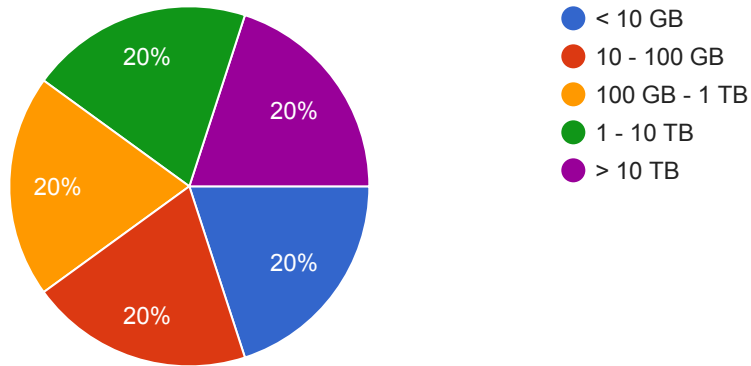
- 🔵 < 10 GB
- 🔴 10 - 100 GB
- 🟠 100 GB - 1 TB
- 🟢 1 - 10 TB
- 🟣 > 10 TB

How do you transfer data into and out of ARCADE (select all that apply)

5 responses

ARCADE browser (GUIs)	2 (40%)
ARCADE terminal (download scripts)	2 (40%)
vos tools (vcp, vsync etc)	5 (100%)
File system mount + rsync	2 (40%)
Through VOSpace	1 (20%)
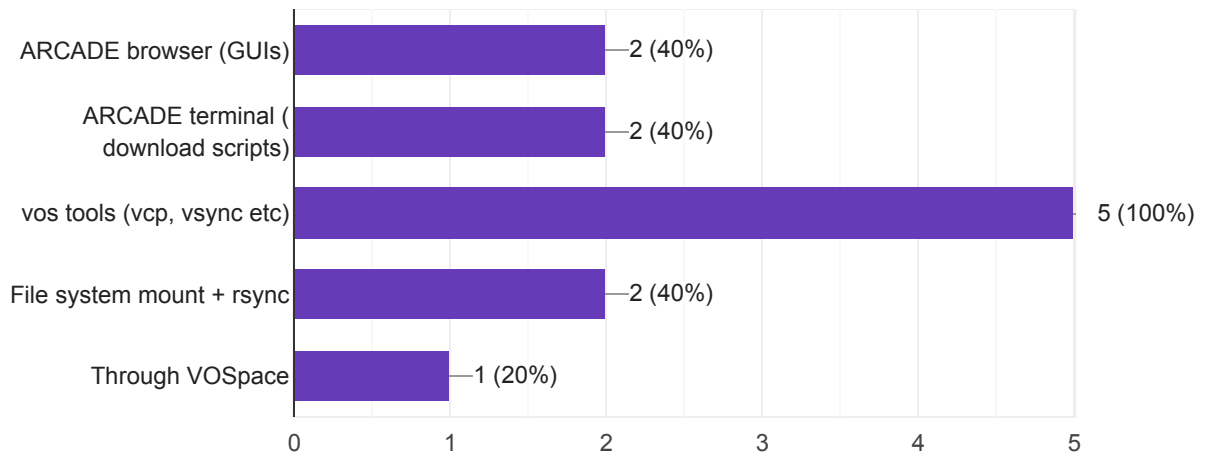
How would you rate the data transfer speed into and out of ARCADE (download/upload/vcp/rsync)

5 responses
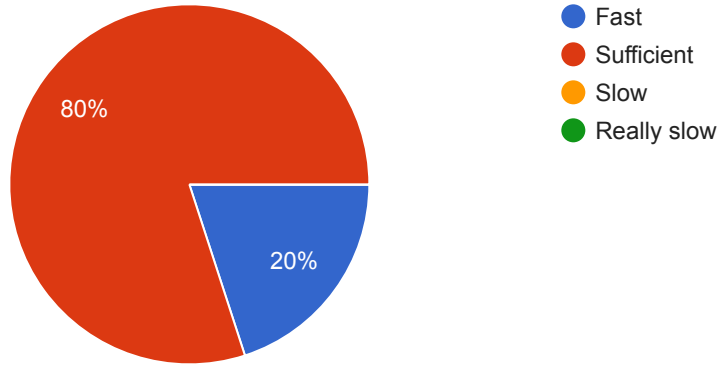
- Fast
- Sufficient
- Slow
- Really slow

80%

20%

How would you rate the processing speed in ARCADE compared to non-ARCADE systems (e.g., running calibration or imaging scripts)

5 responses

- Fast
- Sufficient
- Slow
- Really slow
- Good, but further speed-up is possible with parallelization in CASA for large image cubes.

20%

60%

20%
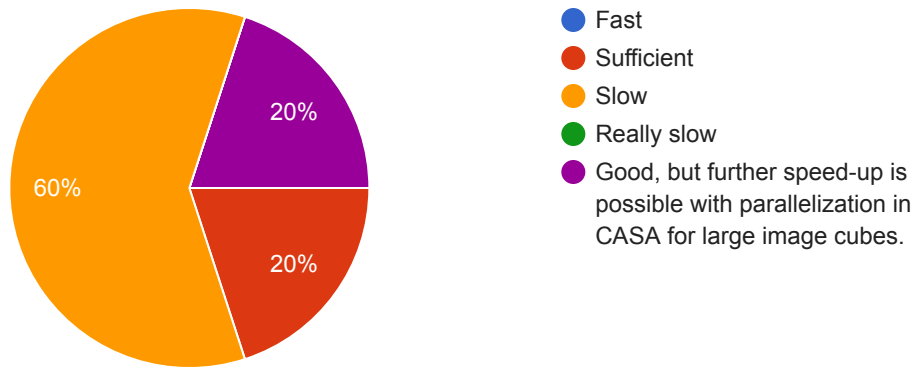
Please list any major challenges / shortcomings you've encountered with ARCADE so far

5 responses

The desktop experience is a bit too minimal. A pdf viewer and image viewer are needed (to view calibration plots and quality assurance reports for ALMA data), for some reason the only program filling this role is GIMP, which is not ideal. Firefox crashes are a huge problem because a web browser is required to access 1) The ALMA archive 2) The html based QA reports in newer pipeline-reduced ALMA projects. gedit is nice to have, but I would add a couple more common text editors, e.g. GUI emacs and VIM.

As a Windows user, there were many challenges that I faced at first in being able to establish my ARCADE and CARTA sessions. Additionally, terminating the sessions or creating new sessions was often challenging with the documented commands not working correctly or needing support from the developers. I also found that the data processing on ARCADE was significantly slower than other machines I was using. Moving around the ARCADE interface, copying pasting, and the actual running of CASA all were slower and required more patience as a user.

Not being able to start up the Jupyter notebook server within my conda environment. At the least, it meant I didn't have access to astropy in notebooks so I can't do the next step of analysis on ARCADE.

Also, I had to move another project off ARCADE because the processing speed and memory available were not as good as what I have available elsewhere.

I also find using CARTA to be a bit slow intermittently. I think it sounded like I/O and memory is shared across users so maybe sometimes this is impacted by whatever other users are doing, but it is frustrating when I just want to change from a per-channel to per-cube histogram (that took 10 seconds on a similar file) and it takes 5 minutes.

ARCADE terminal doesn't have as many functions as general linux machine. For example, I can't copy things on CASA terminal; no multiple tabs in terminal; set up the fontsize in terminal.

Firefox exits randomly sometimes. Copy/paste sometimes doesn't seem to work (specifically the ctrl+shift+c command).

Please list any aspects of ARCADE that you find particularly useful / great

5 responses

Having all the versions of CASA already installed and lot's of storage space is quite nice. The web interface is very convienient and generally quite responsive with a decent internet connection.

I found that there was a lot of storage space and it was a good location to be able to place large amounts of data without having to worry about taking up space elsewhere. The technical and developer support was also very quick and informative. The process of interaction when support was needed was really great. I am sure that having all of the environments and software set up would be quite useful to a brand new user as well, who maybe has less experience downloading and installing software and simply wants to hit the data processing ground running.

Having all versions of CASA easily accessible is great for testing bugs and quickly switching between working on projects from different years. The different options just being in a menu is easy to use, and I think I actually like that versions are compartmentalized in their own windows.

It has large storage for radio data. You can ask for enough memories to process imaging.

Containerization is great, allows for imaging and postprocessing to be kept in it's own place.

What features would you like to see in the future?

5 responses

I reccommend adding Todd Hunter's analysisUtils to the CASA versions available by default since there are hundreds of useful extensions to CASA that are useful to users of all levels of experience.

More choices of text editors (i.e. <insert favourite here> which for me is emacs) would be good. Additionally, having things run a little faster in general would also be a positive feature.

Jupyter notebooks that I can choose the (conda) environment for.

It's great to have basic function working. I would just like some basic functions for the terminal as listed in the shortcomings.

Updated web browser. Multiple desktops.

Thanks for answering our survey! We welcome any additional feedback below.

2 responses

I think that all of the base parts are here to make ARCADE a functional, user-friendly workspace. Having used other things (like Astro at McMaster) to process my data, I don't think that I would rank my experience on ARCADE greater than my experience on Astro, but I think that many of the kinks were worked out throughout the year and could continue to be improved to a point where the experience was comparable. If I reflect on the question of "Would I prefer to use ARCADE over other data processing methods available to me?", I would answer "No." However, I am in the unique circumstance to have other things available to me, like Astro, and therefore ARCADE is a positive thing to have for other scientists without access to large amounts of storage space and computing power in their research groups.

I wanted to add that while I said I use vcp, it is only for small transfers since I haven't bothered to increase the timeout limit for sshfs on my local machine. Essentially anything I transfer that is larger than a Python-script text file, I have had a much better experience with the standard Linux commands (sshfs, rsync, etc.).
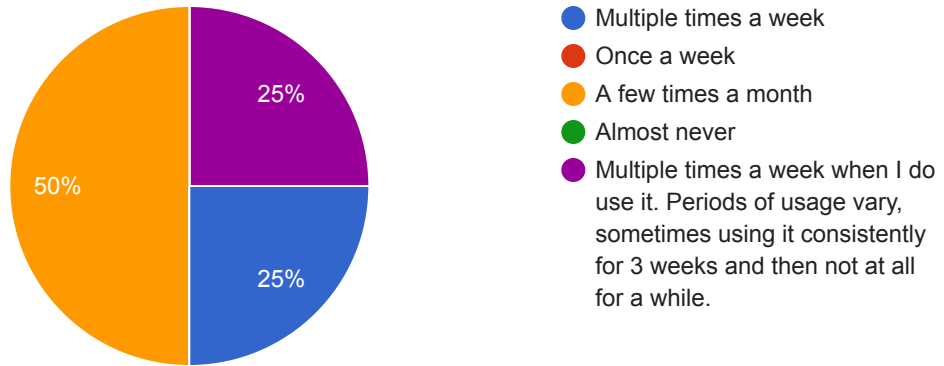
ARCADE User Survey 2

4 responses

Publish analytics

How often are you using ARCADE for your work?

4 responses

- Multiple times a week
- Once a week
- A few times a month
- Almost never
- Multiple times a week when I do use it. Periods of usage vary, sometimes using it consistently for 3 weeks and then not at all for a while.

50% 25% 25%

What are you using ARCADE for? (select all that apply)

4 responses

Calibration	3 (75%)
Imaging	4 (100%)
Analysis within CASA	2 (50%)
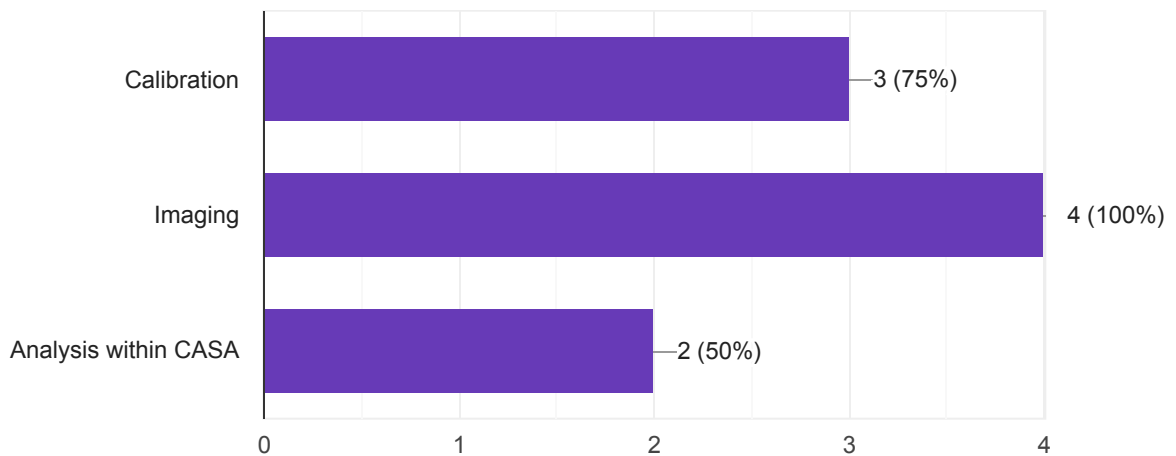
Please provide an estimate of your storage space usage so far

4 responses

- 10's of GB
- 100's of GB
- several TB
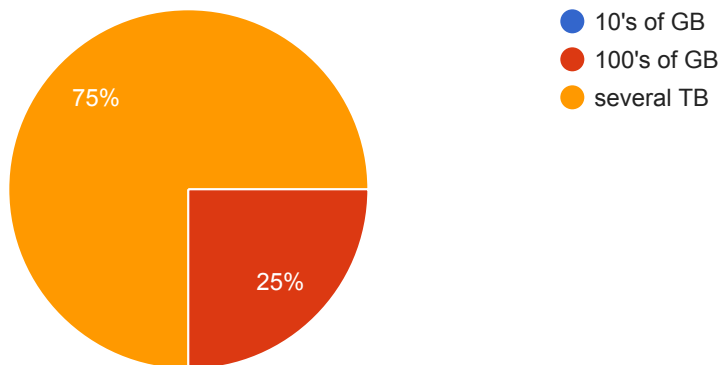
How would you rate the file transfer speed in ARCADE (download/upload/vcp)

4 responses

- Fast
- Sufficient
- Slow
- Really slow
- Downloading from the ALMA archive is fast and works well. Uploading data products to VOspace appears to be broken - transfers of large files with VCP have failed for me, not s…
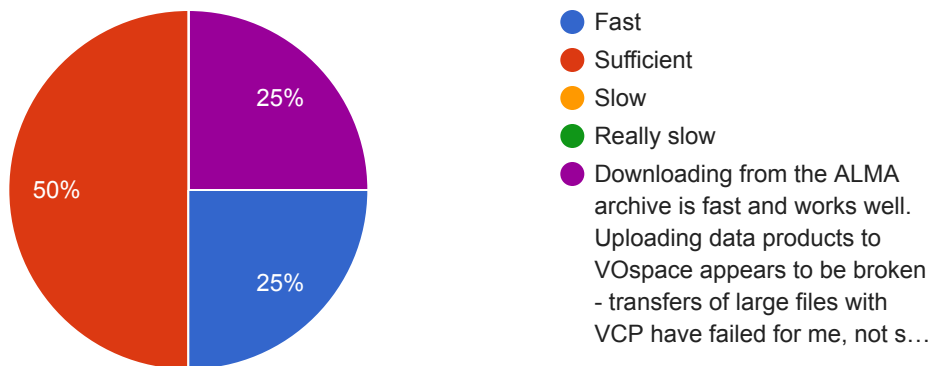
How would you rate the processing speed in ARCADE (e.g., running calibration or imaging scripts)

4 responses

- ● Fast
- ● Sufficient
- ● Slow
- ● Really slow

100%
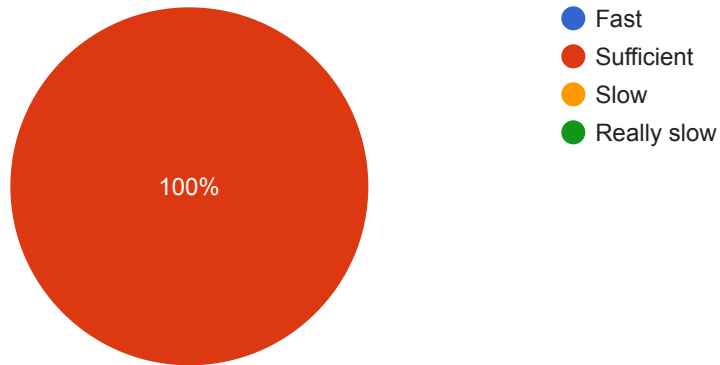
Please list any major challenges / shortcomings you've encountered with ARCADE so far

4 responses

Uploading/downloading all files in a folder using vcp can be finicky, and sometimes results in an empty folder being transferred.

The major challenge were the continuous freezing that ARCADE faced with the CASA terminals every Monday morning. I do think it was just an issue with the infrastructure at the time and it has been fixed(?) but I am not sure.

The lack of a key as to what areas of the interface were good to work with and which ones to stay away from was initially challenging. I've made my account crash once or twice by mistake.

I can't get my data off ARCADES - large file transfers to VOspace have always failed when I transfer to my VOspace so a solution that avoids using it entirely would be be preferred. Needs more basic system applications, especially a GUI text editor, typical ALMA user reduces data while frequently writing and copy/pasting text from scripts so this would be useful.

Please list any aspects of ARCADE that you find particularly useful / great

4 responses

Running ALMA imaging scripts that require more memory than available on my desktop.

I like that I can connect to my VOSpace data on ARCADE through my laptop seamlessly.

I really enjoy the computational power available for students who maybe don't have access to such networks! It is also amazing because it allows us to decentralize this work off our PCs.

I am impressed by the stability/speed of the virtual desktop - very easy to access and no lag, note that I am in Victoria however so other users experience may be more realistic.

Thanks for answering our survey! We welcome any additional feedback below.

1 response

I am so grateful for the existence of ARCADE! Without it my current research project couldn't have gone as far as it has.

Google Forms

ARCADE User Survey 1

13 responses

Was the github tutorial sufficient to get you started?

13 responses

- Yes
- No

23.1%

76.9%
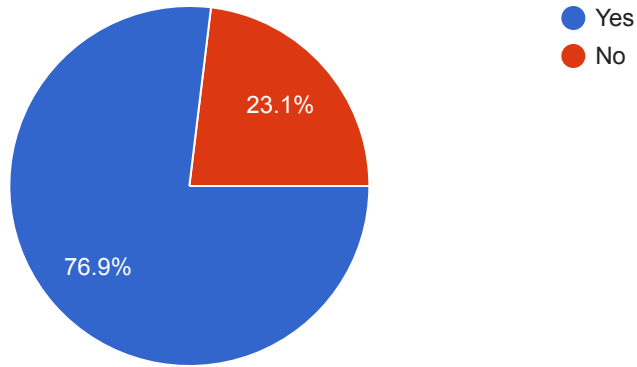
What additional information should be in the tutorial?

11 responses

instructions on using scp/rsync to transfer data - this important as the current process requires moving data to VOspace and THEN another download from VOspace to some other machine.

I would Add a link to the ALMA pipeline page (https://almascience.nrao.edu/processing /science-pipeline), which has links to documentation and indicates which CASA versions were used at ALMA for pipeline operations.

Include some instruction of how to use mpicasa with tclean - this can be a huge timesaver for large images/cubes and is officially supported.

Other useful things: new NRAO archive https://archive-new.nrao.edu/portal/#/ which lets you request calibration at NRAO, and the JVO ALMA fits archive: https://jvo.nao.ac.jp /portal/top-page.do, which is excellent for getting a quick look at product images without having to download them (this is time consuming for high spectral/spatial resolution observations).

It would be nice to have documentation on proper syntax for uploading/downloading files with vcp, in particular transferring folders. I found this to be tricky, and sometimes only empty folders would transfer.

A quick screen recorded video would be excellent. This would give a nice understanding of how things are supposed to be.

I don't think I got a tutorial, so maybe receiving that would have changed my initial perspective and interactions with the platform.

Instructions on how to get data into and out of ARCADE would be great

Setting up local-to-ARCADE vostools transfer. Has since been added. It would also be nice to have an explanation of the different named locations (e.g. cavern, arc, arbutus, vault, vospace, etc.) so we know when we want to put something ARCADE it goes "here" and then not hear or see any of the other names/places.

Start of xterm; Transfer data from local machine to Arcade

I think that example commands on how to copy data into and out of VOSpace and then from VOSpace directly into ARCADE would be helpful. The clipboard tutorial and the VOSpace page directly are both good resources, but joining them together in some way would be good. Also, a bit of information in the tutorial about navigating in and around the ARCADE space and starting terminals and CASA sessions.
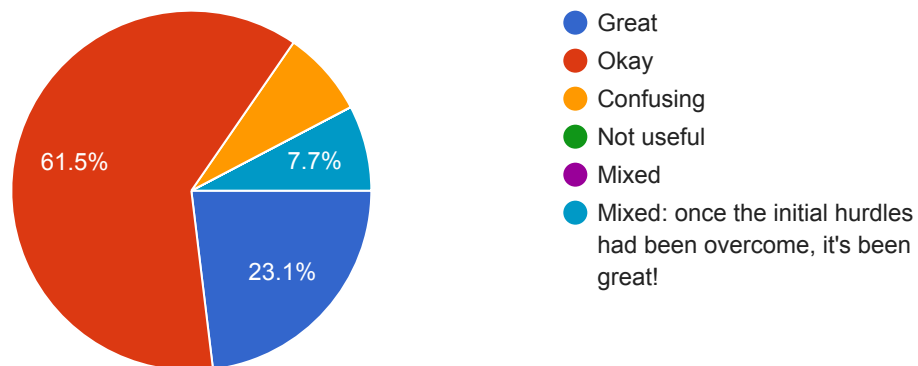
Perhaps include more examples of other vos commands that can be used. Also, maybe acknowledge the occasional proxy errors/loading screen/404's that sometimes appear when connecting and suggest what to do.

It's probably worth highlighting the official pipeline versions of CASA given here: https://almascience.nrao.edu/processing/science-pipeline

Information on the containerization of program installs, such as the unavailability of command-line tools in a CASA terminal, just so the user is aware of how things are set up

What has your overall impression been so far?

13 responses

Legend:
- Great
- Okay
- Confusing
- Not useful
- Mixed
- Mixed: once the initial hurdles had been overcome, it's been great!

Pie chart values: 61.5%, 7.7%, 23.1%

What has your experience been with initiating a session (log-in, etc)?

13 responses

Simple and easy!

This works well.

This step is very easy!

No issues that come to mind. It was a bit slow but otherwise pretty seamless.

Good

Usually no problem. Occasionally I couldn't get the session to respond properly Brian had to kill the session remotely.

I was given a pre-generated URL for my ARCADE and CARTA sessions so I don't yet know what that process is like. I have had a certificate expire for transferring data directly to ARCADE and the message was not immediately obvious (to me) what had changed.

I can't login the session with my McMaster office desktop. I also can't login session with windows computer. It's not user-friendly

As a Windows user, I am having trouble restarting sessions and generating new URLs. Working with the ARCADE team has been really good to re-initiate sessions.
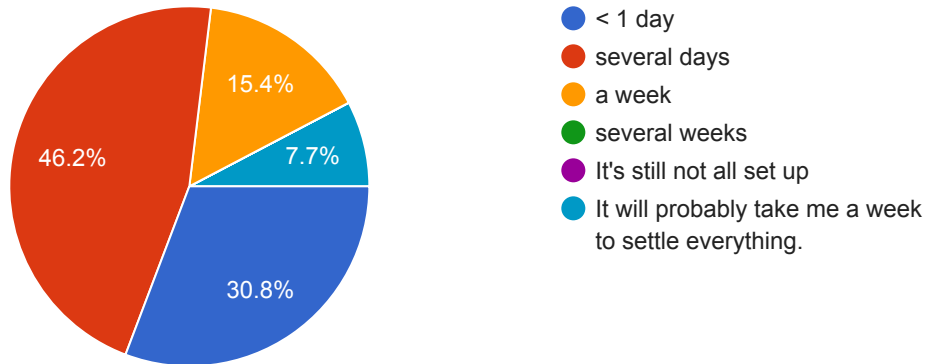
Occasionally will get stuck after connecting to ARCADE on dark screen with three loading squares moving around. But waiting a few minutes and trying again has resolved the problem every time. Also, my first time logging in, I got a Proxy Error, which persisted for a couple minutes, before going away, I haven't gotten another one since.

Works fine

Good experience. The log-in is fast and resumes my session with no issues.

How long did it take you to get everything up and running?

13 responses

- 🔵 < 1 day
- 🔴 several days
- 🟠 a week
- 🟢 several weeks
- 🟣 It's still not all set up
- 🔵 It will probably take me a week to settle everything.

46.2% 15.4% 7.7% 30.8%

What has your experience been with using the system (e.g., starting CASA, terminals and web browsers)?

13 responses

CASA can be finicky but otherwise good

Starting casa is fine. Its unclear to me why so many standard unix tools are missing from the system terminal, but they seem to be available in the CASA terminals for some reason.

These things are great, and feel as if I am using my own desktop.

It was again pretty seamless. The biggest issue was fitting to the size of my laptop screen. I do wish the desktop would scale depending on the size of the screen.

Pretty good.

For a first deployment it's great. There's obviously work to be done but even just having multiple versions of casa will make it super useful. As above, issues that did come up were documented and resolved quickly by developers.

I find the connection timeout to be really short so that iterating to debug something is tedious. I will start a script and it usually crashes quickly the first few times, which means switching over to Google or my text editor, and even quick changes always seem to require me to hit the connect button again.

It's annoying that the terminal emulator is missing typical commands (e.g. less and git). I've seen some discussion around dropping it in favor of only using xterm and I'm a bit surprised. I would prefer the emulator since it has a lot of quality of life features (e.g. tabs, OK default colors, easy to change colors, non-eye-strain-inducing default font).

It would be nice to have emacs as a editor option.

I actually haven't started using CASA yet, just some Python scripts. Will soon.

Haven't done that

In my experience, things run a little slow in the ARCADE interface. In particular, if I am transferring data in from VOSpace and trying to open a CASA shell simultaneously, it takes the shell a few minutes to load and open.

My experience has been smooth in this regard.

Casa/terminal startup is fine. Firefox is very unstable, which is a severe problem for accessing the ALMA archive and viewing ALMA QA weblogs.

- On one occasion, I was unable to launch CASA windows, terminals, or anything; I had to manually kill the session from a local terminal and start a new one

- The web browser is very unstable, but it is still able to handle downloads of files

What has your experience been with data interaction (e.g., downloading, uploading, running scripts, etc)?

13 responses

Good!

Downloading is fine. Uploading ALMA data to VOspace has not worked for me, large files seem to be problematic in general. Support for rsync/scp to external machines needed. Interactive data reduction without a GUI text editor (several should be available) is frustrating - only vi/vim are available and these have a steep learning curve.

Everything is easy, except transferring folders with vcp, as mentioned above.

Really seamless. However, just issues when it came to understanding where everything was stored. I personally felt the core foundations of the directory structure were a bit confusing and complicated. (eg. the reasoning for the 'headless' directory as well as having to go under 'guest' to get the required amount of space didnt make sense to me. The rest of the available directories were seemingly there for no reason.) I do understand however that this was just an issue for the first couple of days and then it became second nature just knowing the location of the data.

Downloading has been good! Some CASA versions sometimes crash when using imaging after having cleaned a dataset.

This is less intuitive but once it was explained to me how to do it, it's fairly straightforward. A walkthrough or instructions would be great as it's such a core use of the system.

I was not able to send a "large" (1 TB) directory with the direct-to-ARCADE tools after trying some different things for a couple days. I was able to upload to my VOspace and then pull that down onto ARCADE with only one interruption in the first step.

Running Python scripts has been fine. I actually installed conda and my environment from another machine pretty easily. That got me going quickly.

Haven't done that

I have not done any data interaction yet.

Scripts have run smoothly. And have had no trouble uploading/downloading data, aside from it being somewhat slow.

Works smoothly. Uploading to VOspace has given me trouble in the past with failed file

transfers, but I will assess how this goes in the next few weeks with some new data sets.

No issues there; everything working well, able to move data in and out of ARCADE with rsync

Thanks for answering our survey! Please let us know below if you have any additional feedback for us.

4 responses

I think it would be better if the little side-tray of buttons (clipboard, fullscreen, etc.) would just pop open and closed instead of slowly sliding in or out. When I want the clipboard I'm trying to do something fast and "efficiently" so it's really painful to watch the tray play its little animation as it slides into view.

I think the clipboard tool works fine. Since it actually works that is already a complete improvement on what I'm used to for VNC. My suggestion would be for some quicker way to access the clipboard, in terms of UI. Having to click that narrow arrow button to open the tray and then click again just to get the clipboard in view is tedious. My suggestion, based on not knowing how this thing works at all, is that it would be best to just have a window that can be minimized and maximized like all the other windows in the desktop environment.

In my experience, things have been going well so far with getting used to using ARCADE. I have not done a lot of data processing or work in the interface so far, but when I am manipulating data or loading things in it seems rather straightforward with a bit of snooping around. I am using a Windows 10 machine, which may offer some difficulty moving forward as I have already experienced some Windows specific issues.

It appears the system does not have a functioning pdf viewer, which is needed for some ALMA QA documents. I am happy to see there is now a GUI text editor available, which makes the workflow for ALMA data reduction easier. Fixing the web browser stability is very important, not just for accessing the ALMA archive but also for. the CASA weblogs contain vital information for data reduction and are made to be opened in a web browser.

Frequently while using ARCADE, I'm taken back to the "Connect" splash screen after only a few seconds of inactivity or even while actively typing. This also happens during a time-out as expected, but it is a small nuisance while working.

Google Forms