

VicNode Research Data Storage Service Offering and Service Level Agreement

VicNode provides Victorian researchers and their collaborators with services to easily store and share research data through an affordable, secure and sustainable service. It is part of a national network of research data nodes funded through the national RDSI (Research Data Storage Infrastructure) program, enabling researchers to share and collaborate with researchers locally, nationally and internationally. VicNode provides storage solutions to suit a variety of research data storage needs. Each storage type comes with its own performance and data protection capabilities and VicNode staff can provide guidance on suitable data storage options.

Context and audience

This document describes the VicNode service and sets out expectations for the delivery and support of the service. The expected audience includes Data Custodians as well as university support staff including IT managers, library staff and research/eResearch support staff. VicNode services are provided by Information Technology Services at the University of Melbourne, and eSolutions at Monash University.

VicNode services will evolve over time according to the needs of the Victorian research community. To reflect these changes, the content of this document will also evolve over time and will be updated and communicated appropriately.

For more information, refer to the VicNode website (<http://vicnode.org.au/>).

VicNode service overview

The VicNode Service Offering comprises of:

- Storage infrastructure that is secure, physically located in Victoria and optimised for storing research data, with a range of performance and data protection options available
- An ongoing commitment to the management and maintenance of the VicNode storage infrastructure. The infrastructure has been set up to target a minimum of 99.5% availability during standard Victorian business hours
- Services and functionality to enable Data Custodians to manage access to research data, including restricting access to authorised users if required
- Reporting on the usage, capacity and performance of storage allocations.

Standard VicNode storage products

VicNode offers the following standard storage products:

Vault 1 (Object) is optimised for frequently accessed research data that do not change very often. It enables fast access to research data with all data stored on disk.

Vault 2 (Tape) is optimised for research data that does not change very often, but primarily utilises tape storage to store data at a low cost, with slower response times.

Market is for research data which is actively used, changing frequently and supports collaboration around data.

Computational is for research data that needs high performance connections to onsite compute facilities, such as NeCTAR. Access to this storage product is via a computational facility external to VicNode. For example, on NeCTAR users will access the computational storage product as an attached volume to a virtual machine.

For all data storage products, the overall speed of retrieval of data is dependent on network conditions and capacity. Additional options are available within the Market and Computational storage products and can be made available as required. Further storage options may be available by negotiation at an additional cost.

More detailed information on the technologies used to provide each storage type is available on the Storage Technologies page of the VicNode website (<http://vicnode.org.au/storage/technologies/>).

Storage types offered

| | Vault 1 (Object) | Vault 2 (Tape) | Market | Computational |
|------------------------------------|---|--|--|--|
| Type of data | Completed research data sets that are regularly accessed | Completed research data sets that are infrequently accessed | Research data is under development and associated with a live research project | High speed storage tightly coupled to specialist high performance computing facilities |
| Data expected to change | No | No | Yes | Yes |
| Data protection | High: data is on dual parity protected disk arrays; offsite replication (2 sets of data stored) | High: Two sets of the data are stored on tape; parity checking [1] | High: offsite replication (2 sets of data stored) or daily back up to tape; daily snapshots; parity checking | Low: storage virtualisation; parity checking [2] |
| Access to previous versions | Versioning is optional and managed by the Data Custodian | Users can request access to last backup | Users can access changes <=30 days | N/A |
| Data Custodian access | Direct | Mediated | Direct (Melbourne infrastructure) Mediated (Monash infrastructure) | Mediated |

| | Vault 1 (Object) | Vault 2 (Tape) | Market | | Computational |
|---|---|-----------------------------|---------------------------|-----------------------------|---|
| Data User access management | Managed by Data Custodian | Managed by Data Custodian | Managed by Data Custodian | | Managed through compute facility |
| Available access methods and protocols [3] | Swift Object store command line 3 rd party object store access products [4] NeCTAR RC dashboard API | NFS (Monash infrastructure) | <u>Monash</u> NFS [5] | <u>Melbourne</u> NFS [5] | Mediated access through the NeCTAR Research Cloud API SFTP SSH |
| Underlying storage medium | Disk (object store) | Tape with disk cache | Disk | | Disk |

- [1] Checksums are applied to data at a block level when data is first ingested. Each block in subsequent copies of the data is also check-summed. These sums are checked periodically.
- [2] Only the inherent data protection capabilities provided by the disk storage technology are available for the computational service, such as RAID and parity checking. These measures are sufficient to allow data to be recovered in case of some infrastructure failure. They do not address the issue of users with appropriate access rights accidentally deleting data, or a catastrophic hardware failure. VicNode cannot add user-specific protection to storage mediated by a compute facility access management layer.
- [3] VicNode does not currently provide a web interface for accessing research data: this service is still under development.
- [4] Examples include CyberDuck, CloudBerry and OwnCloud.
- [5] NFS mount initially available for Melbourne University users using Melbourne infrastructure and Monash University users using Monash infrastructure only

Data protection

VicNode is operated by the University of Melbourne and Monash University. Both institutions have many years of petascale data management experience without any data loss. Both operating centres employ a range of data protection mechanisms. VicNode is implementing high-availability enterprise quality robust storage infrastructure that is designed to automatically recover from local hardware failures where possible.

For Market storage, VicNode maintains at least two copies of the data, housed in different geographical locations within Victoria.

Due to the transient nature of data often stored on the high performance Computational storage service, minimal protections are provided. Users can set-up and maintain additional data protection measures if they are available from within the compute facility.

Managing access to data

VicNode offers a number of access mechanisms for research data. VicNode works with every Data Custodian to ensure the appropriate access mechanism is provisioned for each research data set and that they have the tools to manage access. The current access methods include:

- For data stored on infrastructure at the University of Melbourne, Data Users authenticate to services through the Australian Access Federation (AAF). The user then receives a credential that will enable access to services protected by the OpenStack Keystone Identity Management service. Data Custodians are responsible for managing Data User access to the research data under their custodianship.
- For data stored on infrastructure at Monash University, Data Custodians authenticate to MASSIVE to identify themselves, and can add groups and users.

Data custodians are able to use VicNode storage as a backend to their own access mechanisms. In this scenario, VicNode provides support to create and maintain links between the VicNode storage and the access mechanism. The data custodian is responsible for setting up, maintaining and supporting the access mechanism.

Fees and charges

VicNode operates on a not-for-profit basis and can therefore provide the Victorian research community with cost-effective data storage infrastructure especially tailored to the needs of researchers. Storage fees and charges are based on the products selected. There are three payment models available for VicNode services:

- **Merit:** There are two merit-based schemes that will fund the capital costs associated with storing data with VicNode. Information on applying for these merit-based schemes is available through Merit page of the VicNode website (<http://vicnode.org.au/merit/>).
- **Institutional subscriptions:** Research institutions in Victoria can partner with VicNode to provide large-scale research storage solutions. Research institutions can contact VicNode to discuss options (<http://vicnode.org.au/contact/>). Researchers can contact their local referral

partner to find out more about accessing their institutional subscription

(<http://vicnode.org.au/contact/referral-network/>).

- **Pay-as-you-go:** Researchers can pay VicNode access to storage up to 30 TB for periods of three-months or more. Researchers should contact VicNode to discuss options and prices (<http://vicnode.org.au/contact/>).

Terms and Conditions of use

All Data Custodians using VicNode services must agree to the VicNode Terms and Conditions (for a copy of the Custodian-Node Data Provision Terms and Conditions, see:

<http://vicnode.org.au/about/>). In relation to the Terms and Conditions, Data Custodians should particularly note the following points:

VicNode will:

- Provide infrastructure to enable Data Custodians to enable access to their data.
- Create storage allocations and communicate access information.
- Monitor storage availability and performance to ensure reliability of the service.
- Provide various types of standard data protection approaches depending on the storage product selected (an overview of the data protection options offered on each storage product is outlined in the table on page 3). As such, Data Custodians can be confident that VicNode will safely store their data and it will be accessible to them and their Data Users. However, in the event of user intervention or simultaneous catastrophic hardware failure at multiple VicNode storage locations data may be lost (eg. due to a significant natural disaster or similar).
- Inform Data Custodians of any maintenance or planned outages with at least 7 days' notice.
- Notify Data Custodians if we restrict or remove access to their data for any reason
- Ensure any significant changes to the VicNode terms and conditions are communicated to Data Custodians.

VicNode will not:

- Make any claim over the intellectual property rights over the data it stores.
- Take responsibility for maintaining data.
- Take responsibility for managing Data Users access to data. This remains the Data Custodian's responsibility even if data is accessed via VicNode-provided functionality or services.
- Charge Data Users for access to data.

Data Custodians must:

- Comply with the VicNode Custodian to Node Data Provision Terms and Conditions of Use.
- Ensure all Data Users comply with the VicNode Terms and Conditions of Access, irrespective of the method of access used by the Data User.
- Ensure stored their data does not breach any laws.

- Ensure stored data does not contain any viruses, worms or other malicious software agents.
- Ensure their contact details are up to date.

Uptime and Scheduled Maintenance

VicNode aims for its service to always be available outside planned maintenance windows, but this cannot be guaranteed. The infrastructure providing the VicNode service is targeted to provide more than 99.5% uptime. Support personnel will respond to incidents and unplanned outages during scheduled support hours.

VicNode occasionally schedules planned outages including maintenance windows to allow for optimisation of the systems supporting the service. During a planned outage, a VicNode service may either be unavailable or degraded. In the lead up to any planned outage, Data Custodians will be informed of the outage and of expected effects on the service. VicNode maintenance windows and outages for the VicNode service are published on the VicNode website.

Help and support

VicNode staff can help Data Custodians with:

- selecting the right storage product and funding model for research data sets
- applying for a Merit Storage Allocation through RDSI, if applicable
- ingesting data into VicNode infrastructure
- selecting appropriate access methods for research data
- publishing research data stored on VicNode infrastructure

VicNode continues to automate many of the routine processes related to ingesting and managing research data.

Helpdesk

The VicNode Helpdesk provides support to Data Custodians between 9:00 am to 5:00 pm AEST on Monday to Friday excluding University of Melbourne holidays, public holidays and announced closures (<http://www.unimelb.edu.au/unisec/PDates/holidays.html>).

Data Custodians should contact the VicNode helpdesk via email for support:

support@vicnode.org.au

Data Users should contact the appropriate Data Custodian for all queries related to a research data set.

About VicNode

VicNode is a joint venture between University of Melbourne and Monash University. It is the Victorian arm of the national research data storage capability sponsored by the federal government through the Research Data Storage Infrastructure (RDSI) project.