SNIC National Storage

iRODS User Group Meeting 2017

Dejan Vitlacil KTH-PDC (SNIC)
Ilari Korhonen KTH-PDC (SNIC)
On behalf of the SNIC Storage Group
About

The Swedish National Infrastructure for Computing (SNIC) is a national research infrastructure that provides resources and user support for large scale computation and data storage.

The infrastructure shall meet the needs of researchers from many scientific disciplines and from many institutes for higher education and many research institutes.

The resources provided by SNIC are made available through open procedures such that the best Swedish research is supported and new research is facilitated.

SNAC - Swedish National Allocations Committee - allocates SNIC resources.
About

The SNIC infrastructure is funded by the Swedish Research Council (Vetenskapsrådet) and the partners. SNIC is organized as a long-term project (2012-2016/2018), with its own Board.

SNIC includes six partners. The partners (and their compute centres) in SNIC are Chalmers (C3SE), KTH (PDC), LiU (NSC), LU (Lunarc), UmU (HPC2N) and UU (UPPMAX). The centres provide resources, services, support and expertise and are (local) contact points for the users.
About

SNIC will be reorganized from 2018 in order to achieve a structure more suited for long term governance and funding.

A consortium of ten Universities will govern SNIC to achieve long term sustainability.

SNIC will be funded by 50% from the consortium and 50% from the Swedish Research Council.
Scope

Problematic usage:
- Long term storage
- Secondary copy of active data
- Backup
- ...

- Publications in Nature, with mentioning of SNIC and Swestore predicted using TMHMM v2.0 (ref. 42); Information of the first three best Pfam (v25.0)\(^43\) domains hits by HMMER 3.0 (ref. 44) were included in the table; KO and pathway were predicted by Kyoto Encyclopedia of Genes and Genomes Automatic Annotation Server\(^45\). Raw data from the LC–MS/MS experiment can be found at http://webdag.swegrid.se/snic/bils/uu_sellife/pub/ in the file Jeristrom–Hultqvist_etal_Nat_Commun_2013.zip.
Scope

Data Management Planning:
- design research
- plan data management (formats, storage etc)
- plan consent for sharing
- ...

Creating Data:
- locate existing data
- collect data (experiment, observe, measure, simulate, ...)
- capture and create metadata
- ...

Processing and Analysing Data:
- enter data, digitise, transcribe, translate
- check, validate, clean data
- anonymise data where necessary
- describe data
- data enrichment
- manage and store data
- interpret data
- derive data
- produce research outputs
- author publications
- prepare data for preservation
- document analysis and file manipulations
- manage file versions
- ...

Preserving Data:
- migrate data to best format
- migrate data to suitable medium
- backup and store data
- create metadata and documentation
- archive data
- ...

Giving Access to Data:
- distribute data
- share data
- control access
- establish copyright
- promote data
- ...

Re-Using Data:
- follow-up research
- new research
- undertake research reviews
- scrutinise findings
- teach and learn
- ...

Research data lifecycle diagram
© Jisc CC BY-NC-ND
Archiving (Preserving Data)

Swedish legal framework:

- Freedom of the Press Act,
- Public Access to Information and Secrecy Act (OSL),
- Personal Data Act (PUL),
- Law on Ethical Review of Research
- Archives Act

Research data is to be regarded as a public document since it is stored or was produced at a public authority (University).

First, public documents belong to the public authority and research data is consequently not the property of the individual researcher but of the higher education institution (which normally is a public authority) where the researchers is employed.

Second, everyone who wishes to has the opportunity to request public documents from an authority, which means that we already have a passive form of open access. Research data that includes personal information is protected, however, by legislation regarding secrecy and integrity and any disclosure of such research data may only take place after special review.

Third, the higher education institutions are responsible for archiving and long-term preservation of research data produced by researchers employed by them.
For completed research projects

Raw data, such as data/information that you gathered from experiments, interviews and scientific calculations. The following types of records are raw data and could either be preserved or sorted out.

- Pictures or sound recordings
- Surveys
- Calculations and test results
- Statistics
- Protocols from scientific researches

The following records must be preserved:

- Project diaries
- Final reports, publications and articles
- Theses
- ...

It is up to you to decide whether your raw data should be preserved or not. Take the following criteria into consideration when determining the future value of the raw data: future scientific value, historical value or general interest in the data. The research data that will be preserved must be organized in such a way that makes it as easy as possible for others to search within the material. Place the records in labelled archival boxes/files/electronic folders.

Please note that all raw data that will be sorted out must still be kept for at least 10 years after the financial reporting.
Process

- Check for the open calls

---

**Swedish National Infrastructure for Computing (SNIC)**

The Swedish National Infrastructure for Computing (SNIC) is a national research infrastructure with a threefold mission:

- **Check for the open calls**

1. **SNIC SMALL Allocations**
   - Continuous
   - Link: SNIC SMALL

2. **SNIC MEDIUM Allocations**
   - Continuous
   - Link: SNIC MEDIUM

3. **SNIC LARGE Allocations**
   - The Spring 2017 LARGE call is open. Deadline to apply is 12 April 2017 at 3 PM.
   - Link: SNIC LARGE

4. **Sverstorp SMALL and MEDIUM Allocations**
   - Open for calls.
   - Link: Sverstorp

5. **Sverstorp LARGE Allocations**
   - Not open for calls.
   - Link: Sverstorp

6. **PRACE Call for Tier-0 access**
   - There is presently no open call.
   - Link: PRACE

7. **PRACE Call for Preparatory Access**
   - Continuous. Cut-off date: 9 December 2016, 11:00.
   - Link: PRACE

---

**Important Deadlines:**

<table>
<thead>
<tr>
<th>Call</th>
<th>Upcoming Deadlines</th>
<th>Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNIC SMALL Allocations</td>
<td>Continuous</td>
<td>SNIC SMALL</td>
</tr>
<tr>
<td>SNIC MEDIUM Allocations</td>
<td>Continuous</td>
<td>SNIC MEDIUM</td>
</tr>
<tr>
<td>SNIC LARGE Allocations</td>
<td>The Spring 2017</td>
<td>SNIC LARGE</td>
</tr>
<tr>
<td></td>
<td>LARGE call is open</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Deadline to apply</td>
<td></td>
</tr>
<tr>
<td></td>
<td>is 12 April 2017</td>
<td></td>
</tr>
<tr>
<td></td>
<td>at 3 PM</td>
<td></td>
</tr>
<tr>
<td>Sverstorp SMALL and MEDIUM Allocations</td>
<td>Open for calls</td>
<td>Sverstorp</td>
</tr>
<tr>
<td>Sverstorp LARGE Allocations</td>
<td>Not open for calls</td>
<td>Sverstorp</td>
</tr>
<tr>
<td>PRACE Call for Tier-0 access</td>
<td>There is presently</td>
<td>PRACE</td>
</tr>
<tr>
<td></td>
<td>no open call</td>
<td></td>
</tr>
<tr>
<td>PRACE Call for Preparatory Access</td>
<td>Continuous. Cut-off</td>
<td>PRACE</td>
</tr>
<tr>
<td></td>
<td>date: 9 December</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2016, 11:00</td>
<td></td>
</tr>
</tbody>
</table>

---

**Highlighted Items:**

- SNIC: Acknowledging the use of SNIC resources in publications. [Read more](#)
- PRACE project allocations with Swedish project leader on Europe’s largest HPC (Tier-0) resources. [Read more](#)
- SNIC training events. [Read more](#)
- SciSE: Swedish e-Science Education - courses spring 2017. [Read more](#)
Process

- Check for the open calls
- Register / Login to SUPR

SUPR - SNIC User and Project Repository

SUPR is the SNIC database used to keep track of persons, projects, project proposals and more.

To use SUPR you need to be logged in. If you have been a principal investigator for or member in a recent SNAC project you will be present in the system.

- Login to your existing person.
- Request Password for your existing person.
- Resend Confirmation Email for your recently registered person.
- Register New Person if you are not present in the system.

Proposals Rounds

You can view information about proposal rounds without logging in.

List of Current SNIC Projects

You can view a list of current SNIC projects without logging in.

Current SNIC User Agreement

You can view the current SNIC User Agreement without logging in.
Check for the open calls
Register / Login to SUPR
Choose a round

SUPR - Dejan Vtlacil
You are now logged in.

Undecided Proposal

<table>
<thead>
<tr>
<th>Proposal</th>
<th>PI</th>
<th>Project Title</th>
<th>Round</th>
<th>Centres</th>
<th>Decision Deadline</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNAC 2016/32-13</td>
<td>Sophie Erhardt</td>
<td>Analysing omics in first episode patient...</td>
<td>SNAC Small Swestore 2016</td>
<td>SNIC Storage</td>
<td></td>
</tr>
</tbody>
</table>

The **Decision Deadline** shown is the requested start date (if filled in). Otherwise, it is the end date of the project this proposal is a continuation of (if not more than 90 days in the past).

**Proposals**

You have no proposals in preparation or pending committee decision.

To create a new proposal, select the appropriate open round in the table below and follow the instructions on that page.

**Rounds**

Open for Proposals

<table>
<thead>
<tr>
<th>Round</th>
<th>Deadline</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNAC Large, Spring 2017</td>
<td>2017-04-12 15:00</td>
</tr>
<tr>
<td>SNAC Medium, 2017</td>
<td>—</td>
</tr>
<tr>
<td>SNAC Small C3SE, 2017</td>
<td>—</td>
</tr>
<tr>
<td>SNAC Small HPC2N, 2017</td>
<td>—</td>
</tr>
<tr>
<td>SNAC Small Lunar, 2017</td>
<td>—</td>
</tr>
<tr>
<td>SNAC Small NSC, 2017</td>
<td>—</td>
</tr>
<tr>
<td>SNAC Small UPPMAX, 2017</td>
<td>—</td>
</tr>
<tr>
<td>SNAC Medium Swestore 2017</td>
<td>—</td>
</tr>
<tr>
<td>SNAC Small Swestore 2017</td>
<td>—</td>
</tr>
<tr>
<td>DCS 2017</td>
<td>—</td>
</tr>
<tr>
<td>SNIC SENS 2017</td>
<td>—</td>
</tr>
<tr>
<td>UPPMAX 2017</td>
<td>—</td>
</tr>
<tr>
<td>C3SE Local, 2017</td>
<td>—</td>
</tr>
<tr>
<td>LU Local, 2017</td>
<td>—</td>
</tr>
<tr>
<td>LSENS, 2017</td>
<td>—</td>
</tr>
<tr>
<td>LU Local, 2017</td>
<td>—</td>
</tr>
<tr>
<td>Mosler 2017</td>
<td>—</td>
</tr>
<tr>
<td>PDC Industry 2017</td>
<td>—</td>
</tr>
<tr>
<td>PDC Local 2017</td>
<td>—</td>
</tr>
<tr>
<td>SciLifeLab Storage 2017</td>
<td>—</td>
</tr>
<tr>
<td>UPPMAX courses 2017</td>
<td>—</td>
</tr>
<tr>
<td>UPPMAX Storage 2017</td>
<td>—</td>
</tr>
<tr>
<td>SNIC Science Cloud 2017</td>
<td>—</td>
</tr>
</tbody>
</table>
Check for the open calls
Register / Login to SUPR
Choose a round
Create and submit proposal

Create new Proposal for SNAC Medium Swestore 2017

Checklist
Before creating a new proposal, please consider the points below. Proposals not meeting these requirements cannot be approved.

- You (Dejan Vtlacić) must be the principal investigator (PI) for the proposal.
- The proposal must contain all mandatory information.
- The proposal must be submitted before the deadline (2018-01-01 00:00).

You may choose to create a proposal from scratch or by cloning information from an earlier proposal.

Alternative 1: Create Proposal from Scratch
To create a proposal from scratch for SNAC Medium Swestore 2017, please start by providing a title for the proposal.

Principal Investigator: Dejan Vtlacić
Round: SNAC Medium Swestore 2017
Project Title: SNAC Medium Swestore 2017

[Create New Proposal]

Alternative 2: Create Proposal Based on Earlier Proposal
To create a proposal for SNAC Medium Swestore 2017 by cloning information from an earlier proposal, please choose an earlier proposals to use as a template.

Principal Investigator: Dejan Vtlacić
Round: SNAC Medium Swestore 2017
Proposal to Clone: SNAC Medium Swestore 2017

[Clone Proposal]
Process

- Check for the open calls
- Register / Login to SUPR
- Choose a round
- Create and submit proposal
- DMP
- Formalise SNAC (WG) role

Research data lifecycle diagram ©Jisc CC BY-NC-ND
Process

- Check for the open calls
- Register / Login to SUPR
- Choose a round
- Create and submit proposal
- DMP
- Formalise SNAC (WG) role
- Proposal approved
- User is informed
- Identity Management updated

- 14d for S and M allocation
- L allocation process missing

- Proposal approved
- Confirmation mail about proposal sent
- Change monitoring scripts
  - User
  - SNIC username
  - DN
  - Project
  - Directory (path)
  - E-mail
  - ...

- Swestore information mail
  - Username
  - Project path
  - Temp password
  - Reset password
  - Documentation link
  - ...
Process

- Check for the open calls
- Register / Login to SUPR
- Choose a round
- Create and submit proposal
- DMP
- Formalise SNAC (WG) role
- Proposal approved
- User is informed
- Identity Management updated
- Change password requested
- Storage resource available
Process

SNIC

dCache.org

iRODS

freeIPA

SNAC

SUPR
Process

RT (via support@swestore.se)

Documentation
(Sphinx and GitHub)

Allocation status
(Munin)

GitHub
(configuration files)

IBM Spectrum Protect
(Backup)

Monitoring & Alerting
(Observium, Munin, Ganglia, Nagios)
SNIC (Storage) Infrastructure

- **Swestore dCache**
  - NSC
  - C3SE
  - HPC2N
  - Lunarc

- **Swestore iRODS**
  - PDC
  - NSC

- **Backup & Tape Infrastructure**
  - PDC
  - NSC
  - HPC2N
**SNIC (Storage) Operations**

- **Service Coordinator**
  - SNIC Contact
  - System upgrades
  - Technology expert
  - Investment planning
  - System architecture
  - Advanced debugging
  - Team leader
  - Monitoring / Reporting
  - Needs deputy
  - ...

- **Service Team**
  - Operator on duty
  - User Support
  - Documentation
  - SysAdmin/Ops
  - ...

- **Swestore dCache**
  - NSC
  - C3SE
  - HPC2N
  - Lunarc

- **Swestore iRODS**
  - PDC
  - NSC

- **Backup & Tape Infrastructure**
  - PDC
  - NSC
  - HPC2N

---

**SNIC Storage**

- Dejan Vitlacil
- All Threads
- CHANNELS (13)
  - # backup-and-tape
  - # center-storage
  - # coordination
  - # dcache-core
  - # dcache-sites
  - # edut
  - # general

- **#irods-core**
  - pdc-irods
  - policy
  - # training
  - # user-documentation

- **DIRECT MESSAGES**
  - slackbot
  - Björn Torkelsson
  - Ilari Korhonen
  - Jens Larsson
  - Niklas Edumundsson
  - Peter Gills
  - Pontus Freyhult

- **#irods-core**
  - Ilari Korhonen 12:26
  - no, I like visiting you guys 😊
  - and I'm hoping to get my transport out of stor:
  - Janos Nagy 16:00
  - For the Xen guest pygrub, when counting dow
  - That's /usr/lib64/xen/bin/pygrub, It's a bug, I j
  - Then, there is systemd. He wants to start sshd
  - Dejan Vitlacil 16:03
  - @krishnaveni I forgot name of test SUPR insta
  - Janos Nagy 16:04
  - https://bugzilla.redhat.com/show_bug.cgi?id=3
  - Krishnaveni Chitrantu 16:04
  - disposer.c3se.chalmers.se/supr-test
  - Dejan Vitlacil 16:04
danke
  - Janos Nagy 16:04
  - https://bugzilla.redhat.com/show_bug.cgi?id=3
  - Ilari Korhonen 16:06
  - @vitlacil disposer?
  - Janos Nagy 16:06
  - Now it works with a workaround, if I change t!
  - Janos Nagy 16:45
  - Backups are made, IPA configured with Kerber
  - Jens Larsson 16:56
  - I don't know if I'm around next week. I probab
  - Ilari Korhonen 17:09
  - please, don't let me stand between you and yo
  - if I might visit next week, I was just thinking of
  - Janos Nagy 17:02
  - The question is not that why some people go l
  - Ilari Korhonen 17:02
  - yes, I agree!
Archivematica is a digital preservation system that automates the process of preparing digital objects for ingest into a repository, ingesting them into archival storage and providing access to the archived material. The system is easy to use, though as it draws heavily on the OAIS Reference Model some familiarity with that model is needed to understand the workflows Archivematica supports.

<table>
<thead>
<tr>
<th>Sweetsore dCache</th>
<th>NSC</th>
<th>C3SE</th>
<th>HPC2N</th>
<th>Lunarc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweetsore iRODS</td>
<td>PDC</td>
<td>NSC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Backup &amp; Tape Infrastructure</td>
<td>PDC</td>
<td>NSC</td>
<td>HPC2N</td>
<td></td>
</tr>
</tbody>
</table>
... and now Ilari Kungliga Korhonen
Scope

BE NICE TO ARCHIVISTS
THEY CAN ERASE YOU FROM HISTORY
SNIC (Storage) Activity Plan: 2017

- Upgrade dCache to the latest version, which has many improvements for operational stability and service replication;
- Create an usability workgroup to make Swestore easier to use;
- Create new access methods (suggestions are SFTP and RSYNC);
- Enter more identification methods except user certificates (suggestions are username and password);
- Thorough documentary review;

- Harmonisation of operational procedures and processes;
- Common documentation with dCache;
- Improvement on Monitoring, Alerting and Log Management

- Server SW upgrades;
- Upgrade of Tape Infrastructure (NSC and HPC2N);
- Investigate how off-site copies of backup data will be handled in the future;

**Swestore dCache**
- NSC
- C3SE
- HPC2N
- Lunarc

**Swestore iRODS**
- PDC
- NSC

**Backup & Tape Infrastructure**
- PDC
- NSC
- HPC2N
SNIC (Storage) People

- Swestore dCache
  - NSC
  - C3SE
  - HPC2N
  - Lunarc

- Swestore iRODS
  - PDC
  - NSC

- Backup & Tape Infrastructure
  - PDC
  - NSC
  - HPC2N

People

- Jens Larsson
- Olof Mohill
- Robert Grabowski
- Alex Contis
- Ilari (Kungliga) Korhonen
- Krishnaveni Chitraru
- Janos Nagy
- Peter Gille
- Niklas Edmundsson
- Ragnar Sundblad
- Andreas Johansson
- Dejan Vitlacil
- … and more