

If you don't find the information you need, please email isca-support@exeter.ac.uk for further help or with any direct support issues.

Contents

| | |
|--|-----------|
| 1 About ISCA | 3 |
| 1.1 Meet the team | 4 |
| 2 Where to start? | 5 |
| 2.1 Linux | 5 |
| 2.2 High Performance Computing | 6 |
| 2.3 Other topics | 6 |
| 2.4 Getting training | 7 |
| 2.5 Citing use of ISCA | 7 |
| 3 Access to Isca and applying for an account | 8 |
| 3.1 Access to Isca | 8 |
| 4 Connecting to ISCA and File Transfer | 10 |
| 4.1 Before you start | 10 |
| 4.2 Access methods | 10 |
| 5 Working with Projects | 18 |
| 5.1 Background | 18 |
| 5.2 Project discovery | 18 |
| 5.3 Getting added to a project | 19 |
| 5.4 Funded projects | 20 |
| 6 Using the batch queue | 21 |
| 6.1 Background | 21 |
| 6.2 Queues available for use | 22 |
| 6.3 Writing a job submissions script | 23 |
| 6.4 Cancelling a job | 24 |
| 6.5 Job holds | 25 |
| 6.6 Multi-node MPI jobs | 26 |
| 6.7 Interactive jobs | 27 |
| 6.8 Checking the status of your job | 28 |
| 6.9 Finding out other things about the cluster or jobs | 30 |
| 6.10 Batch priority | 31 |

| | | |
|----------|--|-----------|
| 7 | Filespace | 32 |
| 7.1 | Personal filesystem | 32 |
| 7.2 | Project filesystem | 32 |
| 7.3 | Global scratch | 32 |
| 7.4 | Local scratch | 33 |
| 8 | Software and services | 34 |
| 9 | Problem solving and debugging | 35 |
| 9.1 | Using compiler flags for debugging | 35 |
| 9.2 | General advice | 36 |

1 About ISCA

ISCA is the University of Exeter's flagship HPC facility, which already provides enhanced computing capability to 312 researchers across 165 research projects in Life Sciences, Engineering, Mathematics, Astrophysics, and Computing. The system has already been contributing to research into the modelling and formation of stars and galaxies, using Computational Fluid Dynamics (CFD) within Engineering to understand how flooding affects bridges, as well as being used in the Medical School looking at genetic traits in diabetes using data from the UK Biobank. Abstracts for the research papers can be found on the Research Computing Isca Research Papers¹ page.

The system was designed and integrated by HPC, storage and analytics integrator OCF², using Lenovo, DDN and Mellanox technology and will provide researchers with access to significantly increased levels of resource and further opportunities to undertake ground breaking research.



¹<https://www.exeter.ac.uk/researchcomputing/researchresults/iscaresearchpapers/>

²<https://www.ocf.co.uk/>

1.1 Meet the team

ISCA is supported by two Research IT Officers:



Rosie joined in September 2012 as an IT Assistant in CLES. In Jan 2017 she became an Advanced Research Computing Support Officer, responsible for the day-to-day management and development of the HPC

Rosie's skills are:
Linux, HPC, SQL, Python, Java,
data storage

John arrived here in Jan 2017 from the Met Office where he worked on HPC provisioning amongst other IT duties.

John's skills are:
RHEL System Administrator, ITIL foundation



2 Where to start?

So, you want to use ISCA but you have never used HPC before? Maybe you've never used Linux before? Don't panic! Here are some topics that may help you get started.

2.1 Linux

ISCA runs on CentOS 7 Linux, if you have ever heard of Red Hat Linux they are very similar. The first thing you will see when you log onto ISCA is the shell. Most things on ISCA are done without a GUI (no windows, no clicking with the mouse) so you are going to have to learn to use the command line. Some of the basic commands are:

- **pwd** tells you what directory you are currently in
- **cd** changes your current directory
- **ls** lists the files in your current directory
- **less** views text files
- **nano** is a text file editor (recommended for beginners). Other editors such as **vi** and **emacs** are available
- **cp** copies files and directories
- **mv** moves or renames files and directories
- **mkdir** creates directories
- **man** lets you read the manual

There are many great online resources and books covering all the above commands and more, although please be aware that things like the file structure on ISCA are a little more complex than the average system these often describe.

2.2 High Performance Computing

High Performance Computers are really clusters of hundreds (sometimes thousands) of smaller computers, typically referred to as nodes. Your average desktop or laptop computer has between 4 and 8 CPU cores and you may have heard people refer to ISCA as having over 6,000 cores. However these cores are split up over 350+ nodes of typically 16 or 20 cores each. Each node is likely a lot more powerful than the computer you have on your desk right now, and to look at they are very different, but inside they contain all the same basic components like processors, hard disks, memory, and an operating system. Where the true power of HPC comes in is the nodes' ability to work together. You submit your code as self-contained jobs from a *login* node, and the rest of the nodes, or *compute* nodes, are where your code actually runs. On ISCA a resource manager and batch scheduler handle the distribution of jobs to nodes. In very simple terms all you need to do is write a submission script for your code, stating how many nodes or cores you require and submit your job (Using the batch queue). ISCA takes care of when to execute your job and what nodes to run it on. The concepts of HPC, parallel programming and computer simulations are covered in further depth by various articles, blogs and even online courses.

2.3 Other topics

Other topics of interest for the slightly more advanced user include:

- Openstack
- DCV
- MPI programming
- GPU programming
- Intel Phi programming
- Compiling from source

There are many widely-available resources covering all these topics. If you are interested in a particular area the best place to start is Google. Don't forget to read the rest of the guide, where we cover more of the specifics of ISCA in depth. Any questions regarding ISCA can be directed to isca-support@exeter.ac.uk and we will be happy to help.

2.4 Getting training

Information about training courses run by volunteers or core facilities and links to other resources such as external training courses and online tutorials can be found on the Research Computing Training Home³ page.

2.5 Citing use of ISCA

Please use the following when acknowledging the use of ISCA in research papers:

The authors would like to acknowledge the use of the University of Exeter High-Performance Computing (HPC) facility in carrying out this work.

All papers etc that result from use of the MRC nodes must include the following acknowledgement:

This project utilised high-performance computing funded by the UK Medical Research Council (MRC) Clinical Research Infrastructure Initiative (award number MR/M008924/1)

Also, when you submit your paper or article to Symplectic, please add **ARC - Isca** as an Unclassified Label. Instructions how to do this can be found on the Research IT Sharepoint⁴ pages.

³<https://wiki.exeter.ac.uk/display/researchcomputing/Research+Computing+Training+Home>

⁴https://universityofexeteruk.sharepoint.com/sites/Research_IT/SitePages/Acknowledge-ISCA-on-Symplectic.aspx

3 Access to Isca and applying for an account

3.1 Access to Isca

It is free to use. Isca is available free of charge to all *Research Groups* on all campuses. All *Research Active Staff* are able to access Isca, from any University of Exeter campus network or via a VPN connection. To register simply send an email to isca-support@exeter.ac.uk and your request will be processed as soon as possible.

If you are not part of a *Research Group* but your work is closely related to your supervisor's research then please get them to contact us to ask for a discretionary non research account for you.

If your work is not related to any ongoing research then you will not be permitted to access the Isca HPC. You might wish to investigate the available cloud based, free alternatives. For example Azure⁵, AWS⁶ or Google⁷.

It is priority based. Preferential access will be available for *Research Projects* which have funds allocated on their supporting grant. Full details for this can be found in the HPC Governance Group document titled Principles for allocating priority on Exeter HPC⁸. The HPC is costed as a *Research Facility* and its rates are reviewed annually and benchmarked against national HPC facilities.

| Sponsor | Rate per node-day |
|----------------------------------|-------------------|
| RCUK funded research | £7.17 (exc VAT) |
| Charities and EU funded research | £3.70 (exc VAT) |
| All other sponsors | £7.35 (exc VAT) |

⁵<https://azure.microsoft.com/en-gb/free/free-account-students-faq/>

⁶<https://aws.amazon.com/education/awseducate/>

⁷<https://cloud.google.com/edu/>

⁸<http://login.isca.ex.ac.uk:8080/attachments/41649464/42730954.pdf>

A node-day is defined as the compute capacity of a single 16-cores node running continuously for 24 hours and the differences between the rates primarily relate to whether sponsors are prepared to pay for directly incurred or directly allocated costs. When deciding how much funds to allocate, please bear in mind that Isca is available 24/7 but is only supported during office hours so the actual availability will be somewhere between TRAC methodology and 24/7. Please also be mindful that Isca HPC operates on a *fair share* basis so your funded project may be sharing the 200 available nodes with any number of other funded projects at any given time. Therefore a project which has funds for 200 node-days will probably take more than one day to consume the funding. If you need to discuss the process of allocating funds please consult your *Research Accountant* in the first instance and if they are unable to help, contact the finance helpdesk⁹.

It leverages the Research Toolkit. To see how well Isca is doing, we track usage in node days against the *Research Project* number allocated by the *Research Toolkit*¹⁰ so if your *Research Group* is unfunded or internally funded it is important to let us know in your application otherwise you may experience delay in getting up and running on ISCA.

Glossary

- **Research Toolkit** is the framework in use by the University of Exeter to track Research Projects and their funding
- **Research Project** is the unique number allocated by the T1 or eRAP/PAC process in the Research Toolkit
- **Research Active Staff** are members of one or more valid Research Projects on Isca HPC
- **Research Group** is initially only the PI of the Research Project but they may add more members

⁹<https://www.exeter.ac.uk/finance/financehelpdesk/>

¹⁰<https://www.exeter.ac.uk/researchtoolkit/>

4 Connecting to ISCA and File Transfer

4.1 Before you start

You must be on the campus network via Wifi¹¹, LAN¹², or VPN¹³.

Be aware that the login nodes and virtual desktops are provided for light use tasks only (reading code, checking syntax, light testing).

Read the Software and services page and the Using the batch queue page before trying to use ISCA for computational tasks.

The default timeout for inactivity on the login nodes is 1 day and the session timeout on the virtual desktops is 7 days, therefore to avoid losing any unsaved work you are advised to log out at the end of each day.

4.2 Access methods

SSH

- If you are using a Linux or Apple Mac computer then you already have an SSH client available. To access Isca open a terminal and run the command below, replacing *myUserName* with your actual University username. If you want to enable graphics to be displayed then add the **-Y** flag to the command.

```
ssh myUserName@login.isca.ex.ac.uk
```

- Microsoft Windows does not include an SSH client as part of the operating system so you may need additional software to provide an SSH client (e.g. PuTTY).

¹¹<https://as.exeter.ac.uk/it/network/wifi/>

¹²<https://as.exeter.ac.uk/it/equipmentandsoftware/>

¹³<https://as.exeter.ac.uk/it/network/vpn/>

- If you find your connection attempt fails with a black screen, or is otherwise unresponsive, please check you are on the campus network via Wifi¹⁴, LAN¹⁵, or VPN¹⁶. Check for possible causes for disconnection, such as broadband box reboot, bad Wifi signal, broken LAN cable or loose LAN port. If after that, you still have problems getting a connection, on Linux or Apple Mac computer you can try adding **-v** to the ssh command to view the connection attempt and identify where errors are occurring. If you are using a Windows or other ssh client please refer to your software's help manual for troubleshooting steps.

As soon as you are connected you will find the command prompt saying **username@loginNN** where *NN* is the login server number to which it has connected (there are two at the time of writing). If you are not familiar with the command shell used by ISCA you can type **help** at the command prompt to get help.

N.B. there is a campus firewall between ISCA and the rest of the campus network. To keep SSH sessions open for a long time, you may need to add the following to your local SSH config file (`~/.ssh/config`)

```
ServerAliveInterval 20
ServerAliveCountMax 180
IPQoS=throughput
```

¹⁴<https://as.exeter.ac.uk/it/network/wifi/>

¹⁵<https://as.exeter.ac.uk/it/equipmentandsoftware/>

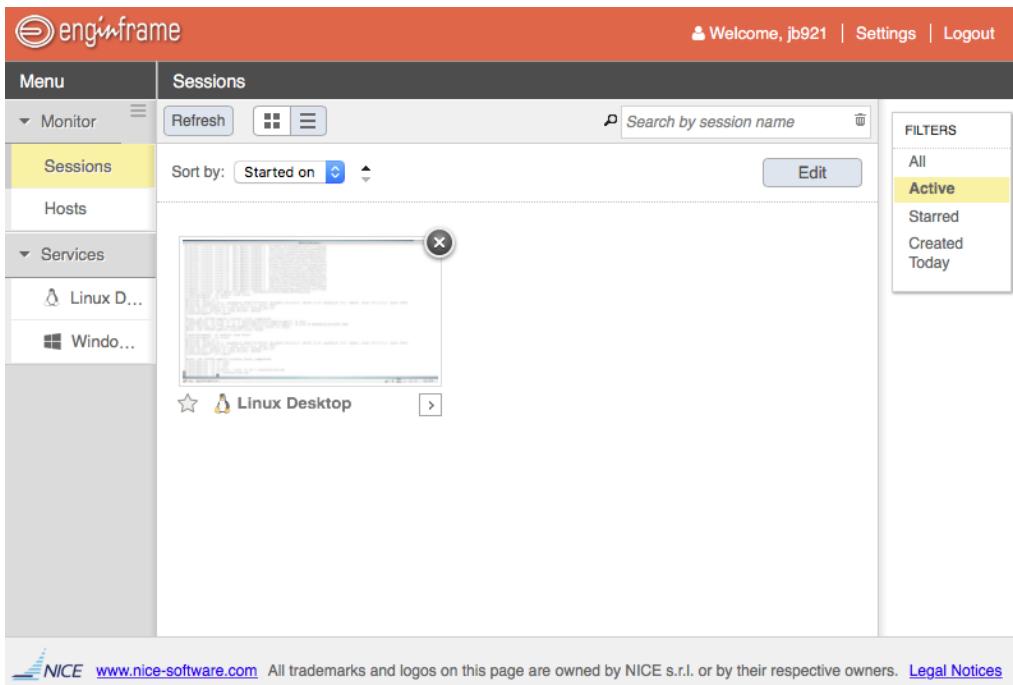
¹⁶<https://as.exeter.ac.uk/it/network/vpn/>

Desktop Cloud Visualisation

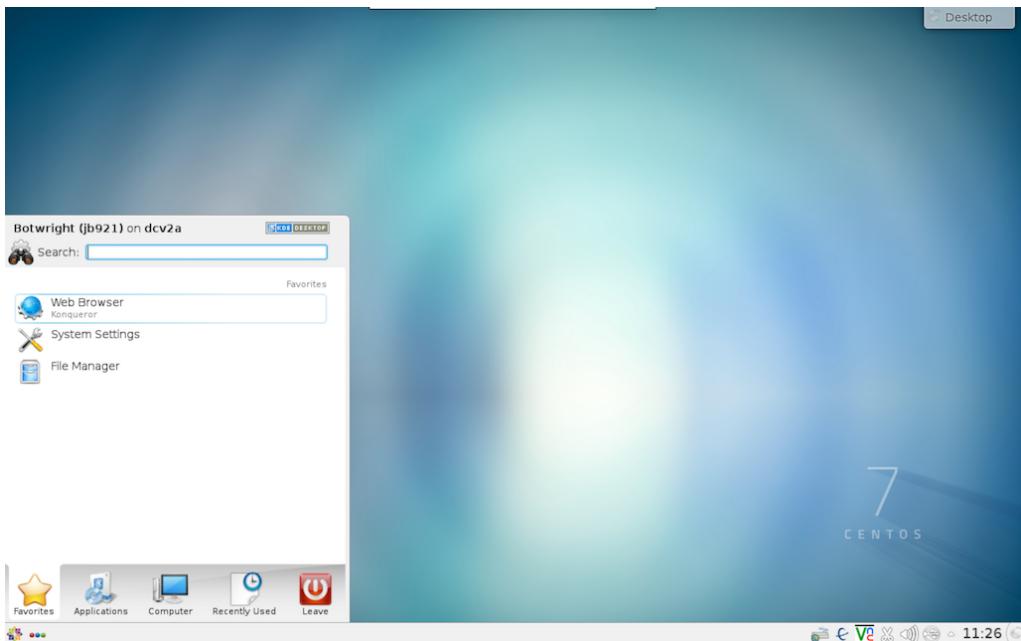
ISCA supports up to ten desktop sessions (KDE or MATE). Point your web browser at <http://login.isca.ex.ac.uk>. This will direct you to a login dialog over secure https. The first time you visit the page you will have to accept an untrusted certificate.

A screenshot of a login interface. It features a light gray background with a central white rectangular input area. Inside this area, there are two text input fields: one for 'Username' and one for 'Password', both preceded by descriptive labels. Below these fields is a single button labeled 'Login'.

Enter your University username and password and click the Login button. This will take you to the portal which will show any existing sessions available for reconnection and the choice of starting a new Linux Desktop or Windows session from the Menu.



When starting a session, it will send you a VNC file which can be opened in your favourite VNC client but it is recommended to use the DCV Endstation client, which is available here: <http://login.isca.ex.ac.uk:8080/endstation.html>. When your session has started you will be displayed a KDE desktop. Use the applications menu to find software you wish to use and click **Leave / log out** when you are finished.



Please note the use of environment modules under DCV is not supported because the virtualisation behind this does not properly support the intel AVX instruction set. Any programs which have been built on a system which supports AVX (i.e. most environment modules) may fail with an *Illegal instruction* error. The solution to this, and a way to get far more performance than is provided by the session, is to submit an interactive job to the batch queue from within your session. See Using the batch queue for more information on how to submit interactive jobs.

Please also note that the KDE environment is different enough to the compute nodes that exporting environment settings between them (#PBS -V) can cause problems.

Viewpoint

Viewpoint (<https://viewpoint.isca.exeter.ac.uk>) is a web interface that lets you:

- Check on the status of your jobs
- View details of nodes and jobs
- View general cluster utilisation
- Browse, upload and download files from the filesystems

The screenshot shows the Moab Viewpoint web interface. At the top, there's a navigation bar with links for HOME, WORKLOAD, TEMPLATES, FILE MANAGER, and SESSIONS. The WORKLOAD tab is currently selected. On the left, there's a search bar and a refresh interval dropdown set to 15s. A 'CREATE JOB' button is also present. The main area is titled 'Workload' and contains a table of completed jobs. The table has columns for Job ID, Job Name, Submitter ID, Start Date, Submit Date, Queue Status, Cores, Nodes, and Wall Clock. The data in the table is as follows:

| Job ID | Job Name | Submitter ID | Start Date | Submit Date | Queue Status | Cores | Nodes | Wall Clock |
|----------|-----------|--------------|------------------------|------------------------|--------------|-------|-------|-------------|
| Moab.113 | paraview | hpotter | 2016-04-19 10:22:13 | 2016-04-19 10:22:06 | COMPLETED | 4 | 1 | 00:12:00:00 |
| Moab.112 | ansys | hpotter | 2016-04-19 10:13:39 | 2016-04-19 10:13:30 | COMPLETED | 1 | 1 | 00:01:00:00 |
| Moab.111 | opendtect | hpotter | 2016-04-13 06:48:14 | 2016-04-13 06:48:06 | COMPLETED | 2 | 1 | 00:01:00:00 |
| Moab.110 | paraview | hpotter | 2016-04-13 06:06:40 | 2016-04-13 06:06:32 | COMPLETED | 4 | 1 | 00:12:00:00 |
| Moab.109 | Nitro_job | hpotter | 2016-04-12 12:25:34 | 2016-04-12 12:25:27 | COMPLETED | 8 | 1 | 00:01:00:00 |
| Moab.108 | Nitro_job | hpotter | 2016-04-12 12:20:10 | 2016-04-12 12:20:00 | COMPLETED | 8 | 1 | 00:01:00:00 |
| Moab.107 | paraview | hpotter | 2016-04-12 12:15:17 | 2016-04-12 12:15:10 | COMPLETED | 4 | 1 | 00:12:00:00 |
| Moab.106 | desktop | hpotter | 2016-04-12 12:12:51 | 2016-04-12 12:12:44 | COMPLETED | 1 | 1 | 00:12:00:00 |
| Moab.105 | ansys | hpotter | 2016-04-12 11:52:04 | 2016-04-12 11:51:57 | COMPLETED | 1 | 1 | 00:00:01:00 |
| Moab.104 | ansys | hpotter | N/A | 2016-04-12 11:50:22 | COMPLETED | 12 | 0 | 00:00:01:00 |

Below the table, there's a 'Show 10 entries' dropdown and a navigation bar with buttons for first, previous, next, and last. To the right of the table, there's a 'Workload Summary' section with a donut chart showing 10 completed jobs. A blue dot on the chart is labeled 'COMPLETED'. Below the chart is a 'View All Workload' button.

Globus File Transfer

Globus file transfer access to home and project areas is available for large managed file transfers. The Globus endpoint name is **exeter#isca**. You will need to set up a Globus account and link it with your University account. The easiest way to do this is via our Globus portal <https://go.exeter.ac.uk>. By default, you will be placed in your homedir. To transfer files to/from your project, browse to it by entering the projects path.

The screenshot shows the 'Transfer Files' interface of the Globus portal. At the top, there's a navigation bar with the University of Exeter logo and links for Transfer Files, Activity, Groups, Endpoints, Bookmarks, Publish, Console, and Account. A banner at the top right says 'Get Globus Connect Personal' and 'Turn your computer into an endpoint.' Below the banner, there are two sets of input fields for endpoints and paths. The left side shows a list of research projects with their IDs and folder icons. The right side is a large empty area with the placeholder text 'Start by selecting an endpoint.' At the bottom, there are sections for labeling the transfer and setting transfer options like sync or delete files.

| Path | Label |
|---------------------|-------------------------|
| /gpfs/ts0/projects/ | Research_Project-106030 |
| | Research_Project-106886 |
| | Research_Project-108389 |
| | Research_Project-108568 |
| | Research_Project-111899 |
| | Research_Project-112761 |
| | Research_Project-114216 |
| | Research_Project-114567 |
| | Research_Project-114691 |
| | Research_Project-115401 |
| | Research_Project-115444 |
| | Research_Project-115838 |
| | Research_Project-116113 |
| | Research_Project-116145 |
| | Research_Project-116175 |
| | Research_Project-116283 |
| | Research_Project-116293 |
| | Research_Project-116804 |
| | Research_Project-116836 |
| | Research_Project-116984 |

You will also need to be added to the list of users who can access the exeter#isca endpoint. The list of users with access can be found on isca at </gpfs/ts0/shared/grid-security/grid-mapfile> or you may view a copy at <http://login.isca.ex.ac.uk:8080/grid-mapfile.txt> (updated overnight). If your userid is not in this file please contact isca-support@exeter.ac.uk

Openstack

Openstack access is available from <https://horizon.openstack.exeter.ac.uk> using your username and password. Projects are automatically created based on data from PAC (<https://as.exeter.ac.uk/it/systems/pac/>). See Working with Projects for details on how to invite or remove other members from your project group

The screenshot shows the Openstack Horizon dashboard with the following details:

- Project:** ISAD - Research_Project-drb
- Compute:** jb921
- Overview:** The main navigation tab.
- Limit Summary:** A section displaying usage statistics for various resources:
 - Instances:** Used 3 of 10
 - VCPUs:** Used 14 of 20
 - RAM:** Used 28,672 of 65,536
 - Floating IPs:** Used 4 of 50
 - Security Groups:** Used 7 of 10
 - Volumes:** Used 0 of 10
- Volume Storage:** Used 0 of 1,000
- Usage Summary:** A section for selecting a time period to query usage.
 - From: 2019-05-01
 - To: 2019-05-01
 - Submit
 - The date should be in YYYY-mm-dd format.
- Active Instances:** 3 Active RAM: 28GB This Period's VCPU-Hours: 200.71 This Period's GB-Hours: 4014.19 This Period's RAM-Hours: 411052.62
- Usage:** A table showing active instances.

| Instance Name | VCPUs | Disk | RAM | Time since created |
|---------------|-------|-------|------|--------------------|
| Dev-R | 4 | 80GB | 8GB | 1 year, 2 months |
| xcat-dev | 8 | 160GB | 16GB | 1 month, 1 week |
| mgmt-dev | 2 | 40GB | 4GB | 1 month, 1 week |

Displaying 3 items
- Download CSV Summary** button.

5 Working with Projects

5.1 Background

To actually submit jobs, use filespace or work with openstack, you will need to be a member of a project in the University's project accounting system (T1). Please ask your Research accountant to provide us (isca-support@exeter.ac.uk) the following information to set up your project on Isca:

- T1 project ID number
- Start and end dates
- PI (Project Instigator, this will probably be you)
- any co-PI's
- amount of priority HPC time you wish to buy (may be zero)

Alternatively, if you know the legacy (PAC) code for your project, it should be already loaded onto the system automatically. See Project Discovery, below.

5.2 Project discovery

In general, projects are named: *Research_Project-[???][number]*

| Source | Format |
|-----------|--|
| PAC | Research_Project-[PAC Project Number] |
| PAC (MRC) | Research_Project-MRC[PAC Project Number] |
| None | Research_Project-[College Abbreviation][Sequential Number] |
| T1 | Research_Project-T[T1 Project Number] |

All projects have a corresponding admin group - this is a list of people allowed to edit membership of a project. (e.g. *Research_Project_Admin-[PAC Project Number]*). By default, projects are setup with PI's and co-PI's as project admins. To find out what you are a member of, on the login node run the **isca_projects** command.

```
[test@login ~]$ isca_projects
Password for testadmin:

ISCA Project Groups
=====
Project Group Name: Research_Project-test
Description: testing
Project End date: 2020-05-31T00:00:00
Disk Quota (GB): default
Priority node-days: none
Members: test
Admins: testadmin
```

If you see your name listed as an Admin of a project, you will have administrative control over the project i.e. you can add new members to the project.

5.3 Getting added to a project

If you feel you should be included in a project, and have already registered on Isca, contact a project admin (e.g. PI) and ask about getting access. If you are an admin for a project, you can add additional members (or admins) to your project.

Examples:

Example a: adding user *testuser* to a project

```
isca_members Research_Project-test add testuser
```

Example b: removing user *testuser* from a project

```
isca_members Research_Project-test remove testuser
```

Example c: enabling user *testuser* as a project admin

```
isca_members Research_Project-test enableadmin testuser
```

5.4 Funded projects

Projects that contribute funding are given an elevated priority on ISCA based on the following rates:

| Sponsor | Rate per node-day |
|----------------------------------|-------------------|
| RCUK funded research | £7.17 (exc VAT) |
| Charities and EU funded research | £3.70 (exc VAT) |
| All other sponsors | £7.35 (exc VAT) |

For enquires regarding including HPC use in new research grant applications, or to discuss funding HPC use from an existing research grant, please contact Research Services¹⁷

If you wish to discuss use of funded time on ISCA and have allocated funds for HPC use in your research grant, please contact isca-support@exeter.ac.uk

To view your remaining node-days run the following command replacing XXXYYY with your research project number.

```
isca_funds Research_Project-XXXYYY
```

See chapter "Using the batch queue" for more information on how to submit interactive jobs using funded priority.

¹⁷<http://www.exeter.ac.uk/research/services/contact/>

6 Using the batch queue

6.1 Background

Jobs on the cluster are under the control of a scheduling system to optimise the throughput of the service. This means that there are some classes of work, for example work that requires repeated access to specified nodes, that are not appropriate for this service. isca-support@exeter.ac.uk may be able to offer alternative facilities for this type of work. The key means by which this is achieved are:

- jobs are scheduled according to the queue, if any, and the resources that are requested.
- jobs are not necessarily run in the order in which they are submitted. The scheduling system runs a fair-share policy, whereby users associated with accounts (projects) that have not used the system for some time will be given a higher priority than accounts who have recently made heavy use of the system. This means that it is not possible for any one account to block the system just by submitting multiple jobs.
- jobs requiring a large number of cores and/or long walltime will have to queue until the requested resources become available. The system will run smaller jobs until all of the resources that have been requested for the job become available - this is known as backfill - and hence it is beneficial to specify a realistic wall time for a job.
- when short jobs are held in the job queue alongside long jobs the priority of these short jobs increases relative to the long jobs over time. This, along with the backfill mechanism described above, means that short jobs cannot be blocked by a large number of long jobs waiting to run.

6.2 Queues available for use

The system is configured with the following queues

| Queue name | Default wall time | Max wall time | Description |
|------------|-------------------|---------------|--|
| stq | 5 minutes | 1 hour | Serial test queue for quick turn around of test jobs. Single node, single processor only. Multiple stq jobs may share the same node. Minimum wall time is 5 minutes. |
| ptq | 5 minutes | 1 hour | Parallel test queue for quick turn around of test jobs. Up to 8 nodes. Multiple ptq jobs may share the same node. Minimum wall time is 5 minutes. |
| sq | 1 hour | 7 days | Serial queue. Single node, single processor only. Multiple sq jobs for the same user may share a node. |
| pq | 1 hour | 21 days | Parallel queue. Up to 200 nodes and at most 1 job per node. Requesting more than a few tens of nodes may mean your job will be queued for a long time before actually running. Please note there is a soft limit of a total of 200 node-days over all your running jobs. |
| mrcq | 1 hour | 21 days | MRC queue. Up to 92 nodes. Multiple jobs may share a node. The MRC queue is reserved for MRC projects only |
| mrchq | 1 hour | 21 days | MRC high memory queue. Up to 2 nodes. Multiple jobs may share a node. The MRC high memory queue is reserved for MRC projects only |

6.3 Writing a job submissions script

To submit a job to the system, you must wrap up your job with a job submission script. These are simply text files so you can use any text editor to create them. For beginners, the **nano** editor on login.isca.ex.ac.uk is recommended as an easy to learn text mode editor. Type **nano** at the command prompt and then paste in the text from the example submission script below.

- Replace *H:MM:SS* with the longest you expect your job to run for, e.g. if you expect your job to finish within ten minutes put **0:10:00**, or if you expect your job to run for seven days put **168:00:00**. You should set this as low as practically possible to help the scheduler get your job run quickly.
- Replace *XXXYYY* with your research project number (you can find your research project number by running the command **isca_projects** on the login node).
- Replace *a.n.other@exeter.ac.uk* with your own email address
- Replace *sleep 100* with the actual command you wish to run.

To save your work in nano, press Ctrl-O and type a name for your file. The name you choose should end with **.sh** to denote it is a **shell** script. Press return to save your file. Press Ctrl-X to exit nano.

Example submission script:

```
#PBS -V # export all environment variables to the batch job.  
#PBS -d . # set working directory to .  
#PBS -q pq # submit to the parallel queue  
#PBS -l walltime=H:MM:SS # Maximum wall time for the job.  
#PBS -A Research_Project-XXXYYY # research project to submit under.  
#PBS -l procs=1 # specify number of processors.  
#PBS -m e -M a.n.other@exeter.ac.uk # email me at job completion  
  
# Commands you wish to run must go here, after the PBS directives  
sleep 100  
echo Hello World
```

- The **#PBS** lines are called **directives**.
- When it is submitted, the script is scanned for PBS directives.
- The initial line in the script that begins with the characters **#!/** or the character **:** will be ignored and scanning will start with the next line.

- Scanning will continue until the first executable line, that is a line that is not blank, not a directive line, nor a line whose first non white space character is #.
- If directives occur on subsequent lines, they will be ignored.
- The lines which start with `#PBS -l` have a lowercase **L** not a **1**.
- If you want to know more about pbs resources please see the **manual** page for `pbs_resources` on the login node

Please note that if you want to use more than one CPU, you should remove the *specify the number of processors* line and replace with the following:

```
#PBS -l nodes=1:ppn=16
# nodes=number of nodes required. ppn=number of processors per node
```

If you want to run your job on the High Memory, GPU, or Phi nodes, add one of the following lines to your job submission script:

```
#PBS -l feature=highmem # High Memory
#PBS -l feature=gpu      # GPU
#PBS -l feature=phi      # Phi
```

If you have paid for node-days, you can enable higher priority for your job by adding the following line to your job submission script:

```
#PBS -l qos=funded
```

When you have saved your submission script you submit it to the scheduler by running either **msub [filename]** or **qsub[filename]**

```
[test@login ~]$ msub test.sh
741.mgmt01.cluster.local
[test@login ~]$
```

Please note: msub will submit jobs directly to moab while qsub uses an intermediate step. In general, msub will provide more detailed diagnostics and will be quicker to submit jobs (particularly array jobs).

6.4 Cancelling a job

You control jobs using the **mjobctl** command. To cancel a job, run **mjobctl -c [job id]** If this fails to stop your job, it might be because the program you are running in your job has crashed and will need an admin reboot the compute node. In that case please send an email to isca-support@exeter.ac.uk stating the Job Id you are trying to cancel and one of the Isca support team will pick it up.

6.5 Job holds

Once in the queue, jobs may be paused for a number of reasons.

Holds

User holds are very straightforward. As would be expected, users can place holds on their jobs to pause their job from being run. Jobs with a user hold in place will have a state of *UserHold* when viewed from **showq**.

To hold a job:

```
mjobctl -h user jobid
```

To release a user hold:

```
mjobctl -u all jobid
```

System Holds

The second category of hold is the system hold. A system administrator can place and release system holds on any job regardless of job ownership. However, unlike a user hold, a normal user cannot release a system hold even on his own jobs. Jobs with a system hold in place will have a state of *SystemHold* when viewed in **showq**.

Batch Holds

Batch holds constitute the third category of job holds. These holds are placed on a job by the scheduler itself when it determines that a job cannot run. The reasons for this vary but can be displayed by issuing the **Some** of the possible reasons are listed below:

- **No Resources** - the job requests resources of a type or amount that do not exist on the system
- **System Limits** - the job is larger or longer than what is allowed by the specified system policies
- **Bank Failure** - the allocations bank is experiencing failures
- **No Allocations** - the job requests use of an account which is out of allocations and no fallback account has been specified
- **RM Reject** - the resource manager refuses to start the job
- **RM Failure** - the resource manager is experiencing failures

- **Policy Violation** - the job violates certain throttling policies preventing it from running now and in the future
- **No QOS Access** - the job does not have access to the QOS level it requests

Jobs which are placed in a batch hold will show up in the state *BatchHold*. However, unlike a user hold, a normal user cannot release a batch held job.

6.6 Multi-node MPI jobs

MPI on ISCA is implemented by two different methods depending on the software modules you have loaded.

Intel MPI (impi)

An MPI job which uses more than one node needs to pass the flag **-bootstrap pbsdsh** to the mpirun command. If you do not pass this flag to mpirun the job will not be able to start up correctly and will hang. For example to run a 32-process MPI job using the intel toolchain:

```
mpirun -bootstrap pbsdsh -f $PBS_NODEFILE -np 32 ./myexecutable > log
```

OpenMPI

An MPI job which uses more than one node requires ssh keys to be set up otherwise the job will hang. For example to run a 32-process MPI job using the foss toolchain:

```
mpirun -machinefile $PBS_NODEFILE -np 32 ./myexecutable > log
```

Setting up ssh keys is fairly straightforward and can be done as follows (on the login node):

```
umask 077; mkdir -p $HOME/.ssh/keys
ssh-keygen -t rsa -b 1024 -N '' -f $HOME/.ssh/keys/id_rsa.isca
cat >> $HOME/.ssh/config << "EOF"
StrictHostKeyChecking=False
UserKnownHostsFile=/dev/null
Host comp* gpu* hmem* phi*
    IdentityFile ~/.ssh/keys/id_rsa.isca
EOF
umask 022
echo -n 'from="10.122.0.0/23"' >> $HOME/.ssh/authorized_keys
cat $HOME/.ssh/keys/id_rsa.isca.pub >> $HOME/.ssh/authorized_keys
```

N.B. do not set a passphrase on your isca key. MPI needs to be able to ssh you around the cluster without user input.

6.7 Interactive jobs

As an alternative to creating a batch script, you may request interactive time on ISCA.

Example command to start an interactive session using one node:

```
msub -I -X -l walltime=0:5:0,nodes=1:ppn=1 -q stq -A Research_Project-XXXXYY
```

Explanation of the command line switches:

- I** = start an interactive job controlled by the current terminal
- X** = forward your local X session to the job as a client
- I walltime=0:5:0** = request 5 minutes time
- q stq** = submit to the serial test queue
- A Research_Project-XXXXYY** = book time to project XXXYYY

After submitting your request, the command will tell you your Moab job number and you will have to wait a length of time depending on the resources immediately available before your job begins and therefore becomes available for interaction. To find out when your job is likely to begin, use the **showstart** command as documented below. Please note that you should not detach processes from the interactive session because they will continue running after your interactive session has ended.

6.8 Checking the status of your job

You can check on the status of your job by running **qstat** or **showq**

```
[test@login ~]$ showq -u $USER
active jobs-----
JOBID USERNAME STATE PROCS REMAINING STARTTIME
741 test Running 13 00:58:25 Fri Nov 11 15:12:44
1 active job 13 of 3162 processors in use by local jobs (0.41%)
 93 of 195 nodes active (47.69%)
eligible jobs-----
JOBID USERNAME STATE PROCS WCLIMIT QUEUETIME

0 eligible jobs
blocked jobs-----
JOBID USERNAME STATE PROCS WCLIMIT QUEUETIME
470 test BatchHold 18 00:01:00 Thu Nov 3 15:26:25
[test@login ~]$ qstat -u $USER
Job ID Name User Time Use S Queue
----- -----
714.mgmt01.cluster.local test.sh test 00:00:00 C pq
741.mgmt01.cluster.local test.sh test 00:00:00 C pq
```

Completed jobs are retained in the queue for one week and with a C status shown in **qstat**. If you do not want to see completed jobs then use **qstat -c** instead of just **qstat**. Alternatively you can set the environment variable **PBS_QSTAT_NO_COMPLETE** e.g.

```
export PBS_QSTAT_NO_COMPLETE=true
```

Much detailed information about your job (such as why it may be blocked) can be found with `checkjob`.

```
[test@login02 ~]$ checkjob Moab.96
job Moab.96

AName: test.sh
State: Idle
Creds: user:test group:domain_users account:Research_Project-test class:sq qos:sqpreemptee
WallTime: 00:00:00 of 00:30:00
BecameEligible: Wed Dec 7 15:51:57
SubmitTime: Wed Dec 7 15:51:53
    (Time Queued Total: 4:53:37 Eligible: 4:53:09)

TemplateSets: DEFAULT
NodeMatchPolicy: EXACTNODE
Total Requested Tasks: 1

Req[0] TaskCount: 1 Partition: ALL
Memory >= 8000M Disk >= 0 Swap >= 0
Dedicated Resources Per Task: PROCS: 1 MEM: 8000M
NodeSet=ONEOF:FEATURE:standard:hmem:phi
NodeCount: 1
    WARNING: empty hostlist specified

SystemID: Moab
SystemJID: Moab.96
Notification Events: JobFail

IWD:          /gpfs/ts0/home/test/.
SubmitDir:    /gpfs/ts0/home/test
Executable:   /opt/moab/spool/moab.job.662IJn

BypassCount: 72
Partition List: pbs
Flags:        RESTARTABLE,PREEMPTEE,GLOBALQUEUE
Attr:         PREEMPTEE,checkpoint
StartPriority: 293
IterationJobRank: 2
comp047      available: 16 tasks supported
comp048      available: 16 tasks supported
comp049      available: 16 tasks supported
comp050      available: 16 tasks supported
comp051      available: 16 tasks supported
comp052      available: 16 tasks supported
comp053      available: 16 tasks supported
comp054      available: 16 tasks supported
<snap>
```

6.9 Finding out other things about the cluster or jobs

pbsnodes -In lists offline or down nodes in the cluster

pbsnodes -a lists all nodes

showbf shows how many idle resources are available at the moment and for how long will they be available

showstart [jobnumber] -e prio shows the *estimated* start time for a job

checkjob [jobnumber] shows detailed information about a job

qstat -i or **qstat -r** show just running jobs rather than the full list from the last 7 days

qsub -I declares that the job is to be run interactively (useful for debugging)

6.10 Batch priority

To view the priority of your queueing jobs use the following command:

```
showq -i -w "class=[queue]"
```

(substitute *[queue]* for the queue you are using i.e. **hq**, **mrcq** etc.)

The following output will be displayed:

```
eligible jobs-----
JOBID      PRIORITY  XFACTOR  Q  USERNAME  ACCOUNT  PROCS    WCLIMIT    CLASS      SYSTEMQUEUETIME
Moab.1*     10000     6.9      -  adam      masters   600  1:00:00:00  pq  Mon Nov 12 13:31:13
Moab.2      10000     1.0      -  adam      masters   600  7:00:00:00  pq  Mon Nov 12 13:31:13
Moab.3      9500      0.9      -  skeletor  villains  16   7:00:00:00  pq  Mon Nov 12 21:51:12
Moab.4      1000      0.7      -  adam      masters   64   1:00:00:00  pq  Sun Nov 18 20:51:02
Moab.5      100       0.0      -  orco     wizards   64   1:00:00:00  pq  Mon Nov 19 11:51:11
Moab.6      100       0.0      -  orco     wizards   64   1:00:00:00  pq  Mon Nov 19 11:51:11

6 eligible jobs

Total jobs: 6
```

We see that Adam's jobs have been queuing since Nov 12 which was 10,000 minutes before today's date (Nov 19). Similarly, the priorities of Skeletor and Orco are how many minutes their jobs have been queuing. The output shows that Adam will likely get his job run next because his job is highest in the queue and it has a star next to it, which means it has reserved the nodes it requires to run.

But the next day...

```
eligible jobs-----
JOBID      PRIORITY  XFACTOR  Q  USERNAME  ACCOUNT  PROCS    WCLIMIT    CLASS      SYSTEMQUEUETIME
Moab.3*     11040     1.2      -  skeletor  villains  16   7:00:00:00  pq  Mon Nov 12 21:51:12
Moab.2      6040      2.0      -  adam      masters   600  7:00:00:00  pq  Mon Nov 12 13:31:13
Moab.4      1640      1.0      -  orco     wizards   600  1:00:00:00  pq  Mon Nov 19 11:51:11
Moab.5      1640      1.0      -  orco     wizards   64   1:00:00:00  pq  Mon Nov 19 11:51:11
Moab.4      -2960     2.8      -  adam      masters   64   1:00:00:00  pq  Sun Nov 18 20:51:02

5 eligible jobs

Total jobs: 5
```

Adam's big job ran and took a large proportion of the resources so fairshare means the priorities of Adam's remaining jobs have been docked. If Adam does not run any further jobs his fairshare will improve within a few days.

7 Filespace

7.1 Personal filesystem

A personal home space of upto 50GB is provided for storing any personal preferences (dot files), or other similar items. Anything larger should be stored within your project folder. To check how much of your home space quota is being used you can run the command:

```
mmlsquota /dev/ts0:homes
```

Please note: this area is currently NOT backed up.

7.2 Project filesystem

Whilst there are charges for using the Facility, **1TB** of disk storage per project is provided free of charge to enable you to use the compute facility. If you anticipate needing more than 1TB of disk storage, a request for funding for any storage requirement above the free 1TB should be included in your grant applications. Project filesystem is removed at the end of a project - warning emails will be sent to all project members to let them know the filesystem is going to be removed at least 3 months prior to removal.

Project space can be found under:

```
/gpfs/ts0/projects/Research_Project-XXXXYY
```

To check how much of a project quota is being used you can run the command

```
df -h /gpfs/ts0/projects/Research_Project-XXXXYY
```

Please note: this area is currently NOT backed up.

7.3 Global scratch

A scratch area has been setup for jobs to work against data outside of project quotas. Files in this area should be treated as transient and are not guaranteed to be preserved - it is best practice to clean up any work from the scratch filesystem at the end of a job.

Scratch can be found under:

```
/gpfs/ts0/scratch
```

You should make a directory for yourself on scratch using the command

```
mkdir /gpfs/ts0/scratch/$USER
```

If you don't want other people to be able to see what is in this directory you should then run the command

```
chmod og-rx /gpfs/ts0/scratch/$USER
```

Scratch has a total capacity of 20TB. To check how much space is being used on scratch you can run the command

```
df -h /gpfs/ts0/scratch
```

Please note: this area is NOT backed up and files may be removed without warning.

7.4 Local scratch

A 700GB per-node /local space has been setup on all compute nodes, which in addition to global scratch space may be used as transient storage needed for a job. At job creation time, the scheduler will create a directory in /local based on your job ID, for example:

```
/local/pbstmp.1234567.mgmt01.cluster.local
```

The job ID is held in the environment variable \$PBS_JOBID so to use the local storage you would put something like the following in your submission script:

```
cd "/local/pbstmp.${PBS_JOBID}"
```

At the end of the job, this job folder is deleted by the scheduler.

Please note: this area is NOT backed up and files may be removed without warning.

8 Software and services

Software on Isca is managed using environment modules. To see the complete list of available and installed modules either refer to

<http://login.isca.ex.ac.uk:8080/modules.txt>

or on a login node simply run the command

```
module av
```

Software is organised into groups called **toolchains** to ensure compatibility. See <http://login.isca.ex.ac.uk:8080/toolchains.txt> for a description of what each toolchain comprises and also see the Easybuild section for further explanation of what a toolchain is used for.

If the software you wish to use is not on the list, or the version for your desired toolchain is not available, please go to the full list of HPC software modules at <http://login.isca.ex.ac.uk:8080/modulesdescriptions.html> and click the software title to send an email to ISCA support requesting the software to be installed. If you require a specific version of the software please include this in the email.

To load a module use the module load command followed by the name of the module.

e.g. to load the intel/2016b module run the command

```
module load intel/2016b
```

To see which modules are currently loaded run the command

```
module list
```

To unload a module run the command

```
module unload
```

You can switch between different modules using

```
module switch
```

e.g. to switch from version 2016b of the Intel compilers to version 2015b run the command

```
module switch intel/2016b intel/2015b
```

9 Problem solving and debugging

9.1 Using compiler flags for debugging

Intel compilers

The Intel compilers have several flags which can help with detecting errors (e.g. out of bounds array access). **ifort** has a *-check all* option which enables run-time checking; this slows the code down significantly but can be very useful when debugging. Adding *-g -traceback* flags will tell you where the error is occurring. See the **ifort** man page for more details of the available flags.

Correctness checking with Intel MPI

The Intel MPI has the ability to perform correctness checking of MPI codes, which can detect errors and resource leaks. Using this option can slow the execution of the program significantly so use with care. To use this capability compile your code with the *-check_mpi* option e.g. `mpif90 -check_mpi myprog.f90`

The program is then run as normal but will produce MPI debugging information in the standard error file. Before using this flag you will need to run the command
`source /gpfs/ts0/shared/software/intel/itac/9.1.2.024/bin/itacvars.sh`

9.2 General advice

Advice to ensure that software works as expected on the HPC

- Avoid changing the \$PATH variable unless necessary. Add entries to the end not the beginning of \$PATH
- Put your personal executables in \$HOME/.local/bin or \$HOME/bin which are already in your \$PATH by default
- Avoid putting the current working directory '.' anywhere in your \$PATH
- Do not set the \$LD_LIBRARY_PATH variable unless necessary

Useful information you should provide to us to help debug problems

- Your user id on the HPC
- Whether you have amended your environment variables
- The working directory you experience the issue in
- The full command line you are running
- The error message which is issued
- The output of the command module list