Access to High Performance Computing: Tier-2 Technical Assessment Form

Instructions:

- Complete Section 1 below as fully as possible. If you have any questions or require clarification, please contact the centre which you would like to apply for resources on. Contact details for each centre can be found in Appendix 1 of the Access to HPC call document.
- Return the completed form (as a Word document) to the centre which you are applying to by using the appropriate contact identified in the bullet point list above along with a short project description.
- The Tier-2 team will complete Section 2 and will contact you directly for more information if it is required. This may take up to 10 days from receipt of the completed form.
- The Tier-2 team will return the fully completed form to you so you can include it in your application.

Notes for Applications:

- Each service may specific restrictions on the form of the applications it accepts. Details on these can be found in Appendix 1 of the Access to HPC call document.
- Applicants should supply quantitative evidence that the codes to be used scale to cpu/gpu/node counts requested. Details on the evidence required can be found in Section 1, Part 6. Please note this section is not mandatory for Tier-2 systems that have novel architectures.

Section 1: HPC Resources and Case for Support
(*To be completed by the applicant*)

1. Project Information.

1.1. Project Title: LURID2: Further study of URI deployment

1.2. Tier-2 Centre applying to: Cirrus

1.3. PI Name and Contact Details.

Name:	Henry S. Thompson
Department:	Informatics
Institution:	University of Edinburgh
Position Held:	Professor of Web Informatics
Address:	4.22 Informatics Forum; 10 Crichton Street; Edinburgh
Postcode:	EH8 9AB
e-Mail:	ht@inf.ed.ac.uk
Telephone:	+44 7866 471 388
Nationality:	UK

1.4. Contact details for application (if different from PI above)
N/A

1.5. Proposed start date of Tier-2 use: 1 November 2022

1.6. Project length (months) of Tier-2 use: 1 year

1.7. Brief Project Summary

Longitudinal study of usage patterns of URIs in web pages allows us to detect trends that are important for managing the evolution of Web Architecture. The proposed work will extend our work in the LURID (ec184) project (EPSRC Tier 2, 2021-9 thru 2022-9), using Common Crawl to measure e.g. the shift from http: to https: (which reduces web-cache effectiveness) and the uptake of persistent identifiers (which alleviates link rot). We'll expand work done on comparison of links found in HTML vs PDF pages, and use our new approach to validating the representativeness of Common Crawl in order to identify and recommend the best 1% segment of recent months' crawls.

2. Previous Use of HPC Resources.

2.1. Which other HPC services have you used?

Cirrus; Microsoft Azure

3. If you have used other HPC services please provide a brief summary of the core/GPU/KNL node hours used (as appropriate) and the types of jobs run (codes, core/GPU/KNL node counts, typical job lengths):

Cirrus: Something like 650K core hours, jobs ranging from 1 hour of 20 tasks/10 nodes to 10 hours of 5 tasks/5 nodes. Mixture of Hadoop/Nutch web crawler and locally-produced PDF scraper code (shell scripts, Python) using gnu parallel, bash and poppler libraries. Working with multi-TB Common Crawl datasets hosted on /beegfs.

Azure: Several 1000 node hours on 4- and 6-node dedicated clusters, streaming large S3 data-sets through scripts and Python code using either Hadoop, Spark or gnu parallel.

4. Tier-2 Software and Support Requirements.

4.1. Summary of software requirements.

What are the main codes you will be using?

See 3 above. Crawler is based on Nutch, locally patched and built, see /home/dc007/dc007/hst/src/nutch-cc. For sample jobs see e.g. /home/dc007/dc007/hst/bin/cdx.sh and …/bin/bigpdf.sh

Software requirements (e.g. compilers, libraries, tools):

Linux, bash, Python, C, gnu parallel. Nutch/Hadoop will have to be rebuilt, as I used Gentoo Prefix before, but have switched to using Singularity for installing and/or building software not available on Cirrus. Nutch file retrieval will have to be tunnelled via login nodes, throttled if necessary.

Support Requirements

How do you plan to port and optimize your code on Tier-2?

Existing expertise in our group. CSE support not required.

Please summarise any other support requirements for this project:

None

5. Proposed Use of Tier-2 Resources.

5.1. Job size mix for the project

Please see notes at beginning of this document regarding the maximum amounts of time that can be applied for and consult any call guidelines.

	Largest Job	Typical Job	Smallest Job
Number of nodes	5	5	1
Number of cores/GPUs used per node	36	36**	8
Wallclock time for each job*	4 days	5 hours	1 hour
Number of jobs of this type	13	305	20
Memory per node required.	Cirrus standard provisions	Cirrus standard provisions	Cirrus standard provisions

*The maximum permitted wallclock time per job is a function of local Tier-2 centre policy.
** These jobs are IO-intensive, and will typically under-populate the nodes (using only half the cores) but requiring exclusive node use to limit contention.

Amount of compute resource: 499,300 CPU Hours

Notional Cost: £4,594

5.2. Disk space requirements.

You may find it easier to complete this section after completing Section 7 (Data Management and Transfer) below.

	Storage
Core source files and data sets	200TB, on /beegfs, inherited from existing Paracrawl (dc007) project
Working storage	Cirrus standard provision

6. Usage Breakdown

The total number of units requested above must be broken down into 3 month *periods* that span the length of access to Tier-2 that has been requested (e.g. if you have requested 1 year of access in total then the units must be split into four 3 month periods). Please add the correct number of rows to the table below for the total length of your project.

If your application is successful then these period allocations may be enforced on the Tier-2 centre in the following way:

- Any unused allocation at the end of a period may be lost
- You may not be able to move units between different allocation periods

Quarter 1 (months 0-3)	125,000
Quarter 2 (months 3-6)	125,000
Quarter 3 (months 6-9)	175,000
Quarter 4 (months 9-12)	74,300

Note that jobs for the different sizes as tabulated above in section 5 look are based on the following job mix:

Largest	Retrieve and index a month of large (>1MB) PDF files that are truncated in the Common Crawl dataset as distributed
Typical	Process a month of Common Crawl data or data index to tabulate properties
Smallest	Test/debug/develop the software for the above on small subsets

7. Scaling Evidence to Support Proposed Use of Tier-2 Centre

The number of units requested and the job sizes specified in 4.1 above must be backed up by quantitative evidence that the code scales efficiently to the job sizes requested. The evidence must include:

- A graph or table of the *speedup* for a similar problem using the code on another HPC system. The speedup should be provided relative to the smallest number of cores cores/gpus/nodes that can be used feasibly (see examples below).

The primary tasks are all straight-forward data-parallelism tasks, e.g. link and header extraction from pdf files in 10s of thousands of Common Crawl and/or Internet Archive files, each of which is 1GB zipped. The outputs per input file are typical very small by comparison.

The following graph is taken directly from [1], a MInf thesis completed in 2020 under my supervision. It reports on the scaling of a tool which compiles a simple specification of such tasks wrt Common Crawl into gnu-parallel, tested using Microsoft Azure. Over the last year as part of the LURID project I and MSc students have used this approach on Cirrus with comparable results.

The key point here is that the decompression time swamps the actual task-specific computation time, and we see from the graph that decompression scales linearly with processors up to 36, which is the number of cores on Cirrus nodes.

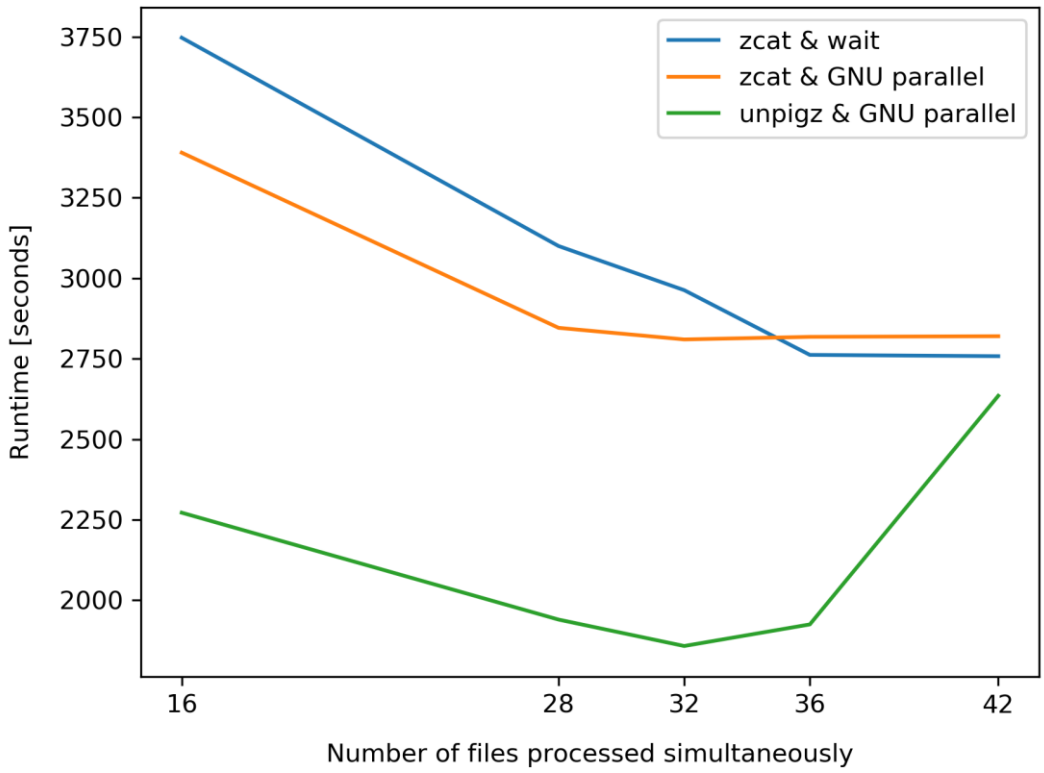

Figure 6.1: Performance comparison of different bash tools. One thousand locally stored WARC files (approximately 850GB of compressed data) were processed using each set of tools. Comparison was repeated for different numbers of files processed simultaneously.

[1] http://www.ltg.ed.ac.uk/~ht/Lukasz_Domanski_MInf_proj.pdf

8. Data Management and Transfer

This section asks some basic questions about the data generated on Tier-2 by the proposed calculations.

7.1 How many files are typically produced by each job?

1000s, divided into 100 subdirectories

7.2 How much data is read in by each job?

60—70,000 gzipped files, each around 1GB == 50—60TB

7.3 How much data is produced by each job?

1GB

7.4 What percentage of the produced data do you expect to transfer?

- To the centres storage facility? None
- Off Tier-2? Negligible, in the form of summary tables and statistics

7.5 How do you plan to transfer data from the Tier-2 centre to the (insert storage facility)?

N/A

7.6 How do you plan to transfer data off Tier-2/Storage facility?

ssh for small summary tables

N.b. Applicants are responsible for ensuring copies of all data. The Tier-2 services are not liable for any data loss on the systems.

Section 2: Technical Assessment (*To be completed by Tier-2 team).*

Date Received by Tier-2 centre: 20th September 2022

Do the applicants have the technical expertise required for the proposed work?	Yes
Applicant is an existing Cirrus user, with a previous project to do similar kinds of computational work.	

Is the software specified technically suitable for the Tier-2 machine requested?	Yes
The main software is already installed and tested on Cirrus. Some of the software (Nutch and Hadoop) will exploit the Singularity container environment on Cirrus.	

Is the compute time requested reasonable and has the job breakdown been technically justified? Are the storage requests reasonable?	Yes
The time request (477k CPU-hours) is based on estimates from previous work on Cirrus and will be used to run three, described classes of workload. Local Cirrus storage requirements are modest (applicant has asked for default quota) though the applicant wishes to make use of a separate BeeGFS system (already connected to Cirrus and used in a previous project (dc007)).	

Has scaling evidence been provide that shows speedup to required job size for the software specified?	Yes
The workload is mostly high-throughput. Performance results are presented from Microsoft Azure, which show a node can be filled with concurrent tasks without significant loss of performance. The applicant confirms the same scaling performance is seen on Cirrus.	

Is the data management and transfer plan reasonable and technically sound?	Yes
Jobs produce a significant amount of data (up to 60TB), which will be written to the BeeGFS system. However, negligible amounts of data need to be transferred off Cirrus.	

Is the application, as outlined above, suitable for access to the Tier-2 service? Yes

Does the project require the technical capabilities of Tier-2? Yes Would a different computing resource be more appropriate? If so, which one? No
The applicant proposes a high-throughput workload which would be challenging to fulfil on a smaller (e.g., institutional) resource. The pre-existing connection to a BeeGFS service, which is being used by the applicant and their peers, makes Cirrus the obvious service on which to run these computations.

Name: George Beckett
Position: Programme Manager
Date: 3rd October 2022