This paper is a pre-print version presented at the ISKO UK 2009 conference, 22-23 June, prior to peer review and editing. For published proceedings see special issue of Aslib Proceedings journal.

Large-scale Grid Computing for Content-based Image Retrieval

Dr Chris Town [1] and Dr Karl Harrison [2]

[1] University of Cambridge Computer Laboratory and Imense Ltd, William Gates Building, Cambridge CB3 0FD

[2] School of Physics and Astronomy, University of Birmingham, B15 2TT (formerly: University of Cambridge High Energy Physics Group, Cavendish Laboratory, JJ Thomson Avenue, Cambridge CB3 0HE)

Abstract

Content-based image retrieval (CBIR) technologies offer many advantages over purely text-based image search. However, one of the drawbacks associated with CBIR is the increased computational cost arising from tasks such as image processing, feature extraction, image classification, and object detection and recognition. Consequently CBIR systems have suffered from a lack of scalability, which has greatly hampered their adoption for real-world public and commercial image search. At the same time, paradigms for large-scale heterogeneous distributed computing such as Grid computing, cloud computing, and utility based computing are gaining traction as a way of providing more scalable and efficient solutions to large-scale computing tasks.

In this paper, we present an approach in which a large distributed processing Grid has been used to apply a range of CBIR methods to a substantial number of images. By massively distributing the required computational task across thousands of Grid nodes, we have achieved very high throughput at relatively low overheads. This has allowed us to analyse and index about 25 million high resolution images thus far while using just two servers for storage and job submission. The CBIR system was developed by Imense Ltd. and is based on automated analysis and recognition of image content using a semantic ontology. It features a range of image processing and analysis modules, including image segmentation, region classification, scene analysis, object detection, and face recognition methods.

Keywords*: Grid computing, content-based image retrieval, image analysis, cloud computing, virtualisation*

1. Introduction and Overview

The Grid processing work discussed in this paper is a collaboration between Imense Ltd., a high-tech company based in Cambridge UK, and the Cambridge University eScience Centre. The collaboration was funded in part through a knowledge transfer grant from the UK STFC (Science and Technology Facilities Council). This has allowed us to deploy Imense's CBIR system to a very large scale (over 120,000 CPU) processing Grid set up by the international particle physics community, although we have thus far restricted ourselves to the UK part of the Grid known as GridPP.

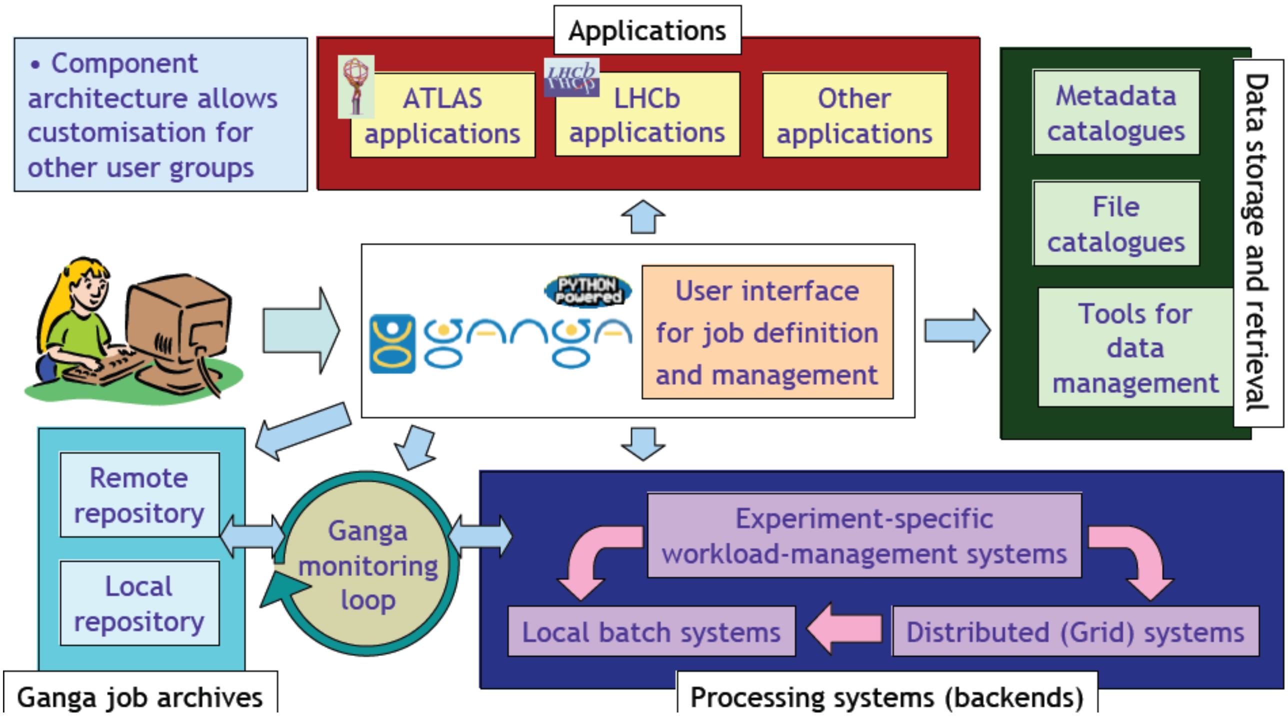

Figure 1: An overview of the components involved in using the Ganga system to perform job definition, submission, monitoring, and control on the particle physics Grid.

The software infrastructure that underpins this massive resource is entirely open source and could therefore be ported to other computing clusters. We are making extensive use of the Ganga job submission and control framework (Brochu, et al. 2009), which is heavily customisable and supports a number of different backends. An overview of the Ganga system is shown in figure 1.

In addition to the gLite and LCG middleware deployed on the Grid, we have also conducted experiments using local Condor processing pools, and we are in the process of measuring the overheads incurred by virtualisation solutions such as VMware. While most GridPP nodes currently run the Scientific Linux operating system and are primarily intended for physics based applications, we have successfully ported the required middleware components to more common Linux distributions and made some additional software modifications and enhancements that could be of benefit to other non-physics applications wishing to make use of Grid computing.

To date we have used the Grid to process some 25 million images and have extensively benchmarked the performance and reliability of the underlying Grid middleware and infrastructure. We are involved in efforts to make it easier to install Grid software such as gLite on commercially owned computing clusters, and to improve the job management facilities provided by Ganga. There is also scope for improving resource utilisation and eliminating network bottlenecks in light of a particular Grid application, and we have some initial results about how this can be done in the case of the image analysis software developed by Imense.

2. Content-based Image Retrieval

The vast bulk of the data on the internet consists of images and video rather than text, very little of which has been adequately described using textual metadata. Many consumers are accumulating thousands of personal images, yet they lack efficient tools to browse, organise, search, and retrieve them. Consequently there is great scope for systems that are able to perform image search on the basis of an automated analysis of the actual content of images, thus allowing users to search “inside the picture” just as they have become accustomed to being able to search within other documents. Unfortunately, most content-based image retrieval (CBIR) systems have failed to gain wide adoption. Furthermore, the high processing and memory requirements involved in automated image analysis have resulted in a wide gap between the relatively small data sets considered by most research projects in CBIR and the needs of users, image archives, and businesses in the ”real world”.

Imense Ltd has implemented a novel image retrieval system, featuring automated analysis and recognition of image content, and an ontological query language. The image analysis undertaken includes recognition of visual properties, such as colour, texture and shape; recognition of materials, such as grass or sky; detection of objects, such as human faces, and determination of their characteristics; and classification of scenes by content, for example beach, forest or sunset. The system uses semantic and linguistic relationships between terms to interpret user queries and retrieve relevant images on the basis of the analysis results. Moreover, the system is extensible, so that additional image classification modules or image context and metadata can easily be integrated into the index. Some of the underlying concepts are discussed in (Town, C.P. and Sinclair, D.A. 2004) and (Town, 2006).

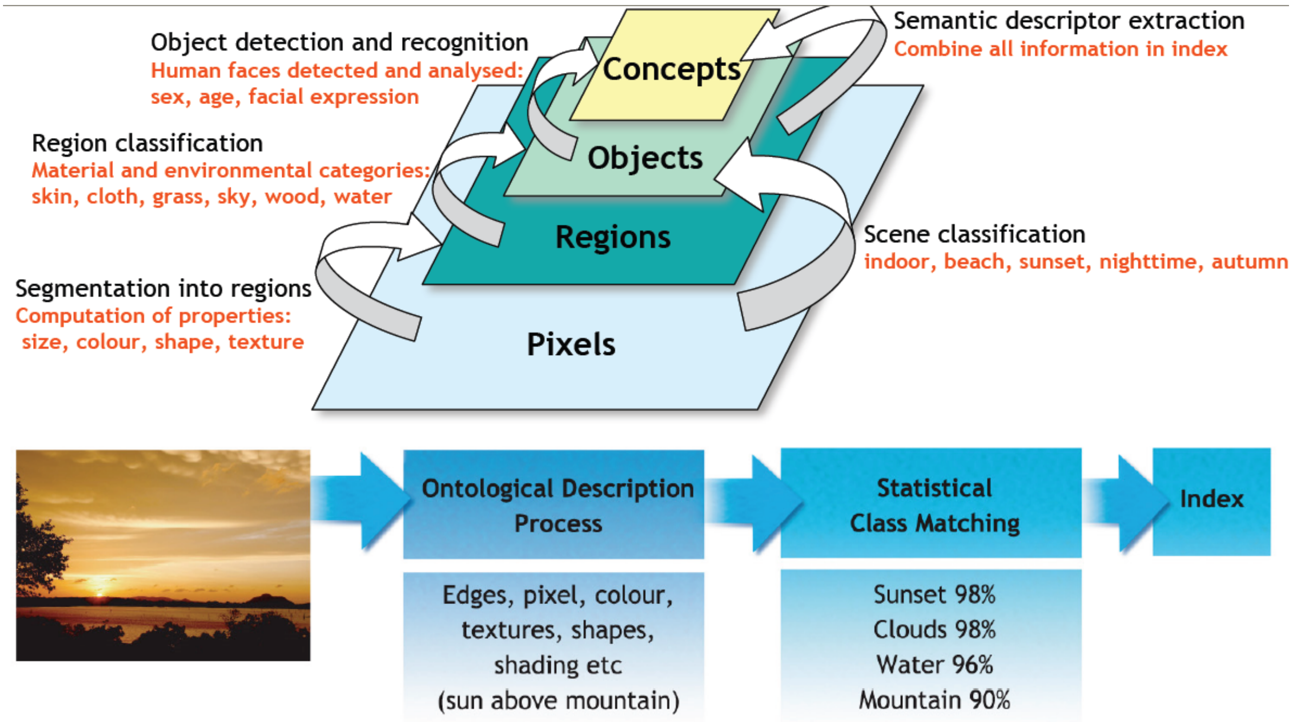

Figure 2: Diagrammatic overview of the image analysis and recognition carried out by Imense Ltd.

As illustrated in figure 2, Imense's technology comprises a number of complex processing stages in order to analyse the content of an image:

- **Image segmentation**: In order to identify salient parts of the image corresponding to objects or object parts, the image is automatically segmented into a covering set of non-overlapping regions and sets of properties such as size, colour, shape, and texture are computed for each region. The number of segmented regions depends on image size and visual complexity, but has the desirable property that most of the image area is usually contained within a few dozen regions which closely correspond to the salient features of the picture.

- **Region classification**: Segmented regions are then automatically classified according to a predefined set of material and environmental categories, such as “grass”, “sky”, “wood”, “water” etc.. Sophisticated statistical machine learning methods are employed to yield a highly reliably probabilistic classification of the image. This may be regarded as an intermediate level semantic representation which serves as the basis for subsequent stages of visual inference and composite object recognition.
- **Scene classification**: A second stage of classifiers is applied to analyse image content at a higher scene level. Examples of scene categories include “indoor”, “beach”, “sunset”, “nighttime”, `”autumn”, etc..
- **Object detection and recognition**: The image analysis also features detectors for common objects. For example, human faces are automatically detected and classified according to personal attributes such as gender, age, and facial expression.
- **Index generation**: Once all image analysis stages have been applied, then all the information from the various classifiers and recognisers is combined into a special indexing format which supports fast content based image retrieval. The searchable index itself is later compiled at a central server, although it may be practical in future to maintain a distributed index across multiple Grid sites to support very large indexes with many millions of images.

3. The Need for Grid Computing

While Imense has developed some highly innovative CBIR technology that allows users to search “inside the picture” more effectively than is currently possible, the image processing overheads involved had previously made it difficult for Imense to test its methods on very large multi-million image data sets. Furthermore, the vast number of images publicly available on the web (estimated at over 20 billion) makes it imperative that any processing must be performed in a highly distributed manner.

The processing stages involved in the Imense image search system, i.e. image analysis and indexing, are intrinsically sequential and take up to 10 seconds for high-resolution images. In order to benefit from parallelisation, it was decided to parallelise at the granularity of single images or small subsets of images. Each image can therefore be processed in isolation on the Grid, and such processing takes no more than few seconds or tens of seconds on a typical GridPP node. In order to minimise overheads, several hundred images are automatically agglomerated into a batch which is then submitted for processing via Ganga, with the results of image processing and analysis being passed back to the submission server upon successful completion.

An initial proof-of-concept exercise was carried out using resources at five UK GridPP sites (Birmingham, Cambridge, Durham, Glasgow, Oxford), and saw about 2 million images successfully analysed. The Imense Ltd software was pre-installed at each Grid site using standard midleware tools. Image-processing jobs were then created, submitted and tracked using Ganga. The archive of test images was stored on a Cambridge-based server, which was also used as the submission machine. Each job was given a list of several hundred images to process, which were downloaded to a worker node. After processing, results were uploaded to a Grid storage element. Ganga, running on the submission machine, continually monitored job status, and automatically retrieved the outputs of completed jobs.

4. Performance Analysis

In this section we report some results arising from a performance analysis of Grid processing of about 18 million high resolution digital photographs using Imense's image analysis algorithms.The processing was carried out at 9 sites across the UK's GridPP particle physics Grid, namely at the Universities of Birmingham, Brunel, Cambridge, Durham, Glasgow, Lancaster, Oxford, and Royal Holloway (RHUL), as well as the Rutherford Appleton Laboratory (RAL).However, the number of images processed at Brunel and Birmingham was relatively small due to problems with older hardware at those two sites, and so results are shown only for the remaining 7 sites.

The processing was carried out over a four-week period between November and December 2008. The first week was dedicated to software deployment and testing at the 9 GridPP sites and the bulk of the processing was performed in the two central weeks, followed by reprocessing of failed jobs during the fourth week. The Ganga job-submission system is highly configurable, and the the strategy that we adopted this time was designed to have large numbers of jobs running in parallel, although with both a self-imposed limit and a limit coming from the fair-share policies in place at Grid sites.

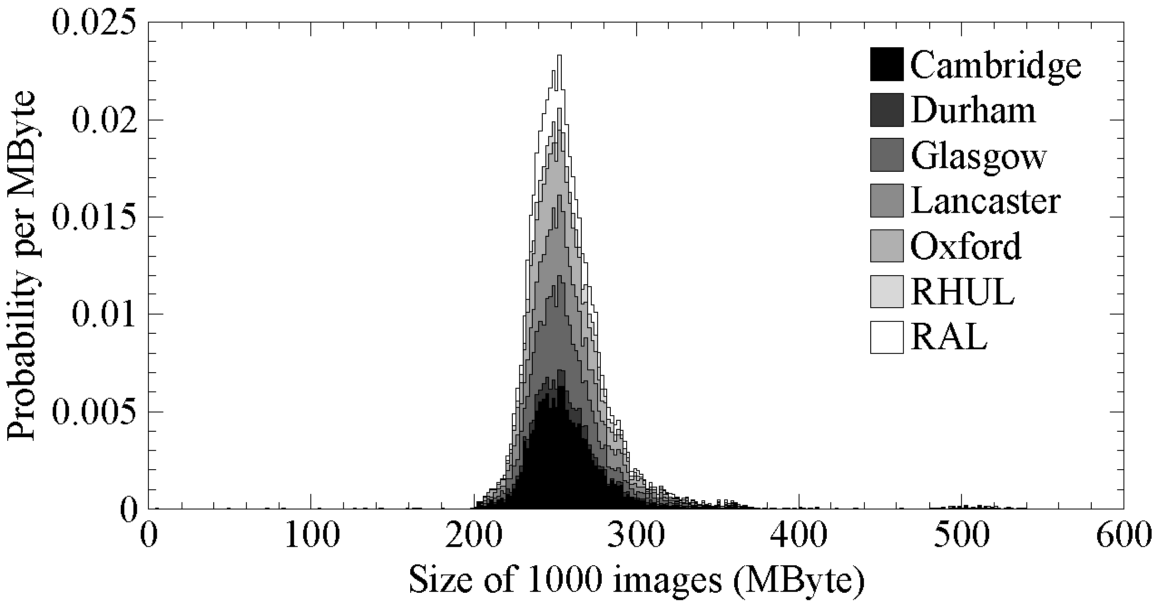

Figure 3: Distribution of image file sizes in each batch of 1000 images at each of the 7 primary sites.

The 18 million images were downsampled to a resolution of about 1 million pixels and stored on two servers provided by Imense, with each server containing 6 1TB disks and a single quadcore Intel Core2 processor. The images were then split into batches of 1000 pictures each for submissions and processing on the Grid. Figure 3 shows the distribution of file sizes for the resulting job batches. This shows that the sample sizes were essentially the same at all sites, and so any variations in processing time reflect variations in hardware specifications at each Grid site.

One Ganga instance ran on each server. The status of jobs at each site was checked every 10 minutes and new jobs submitted via the gLite workload-management system according to the following:

Conditions for submitting a new job (1000 images per job) at a given Grid site:

- Queued or running jobs at site < 100

- Queued jobs at site < 30
- Queued or running jobs at all sites < 400
- Total failed jobs < 100

The aim was to maximise throughput without causing undue inconvenience or delays to other users of computing resources at each site. It should be noted that the LHC (Large Hadron Collider) at CERN, whose experiments form the primary raison d'être for the European and UK particle physics grids, was not yet operational throughout the period these tests were conducted. However, an effective operational constraint of no more than 1% of total Grid resources was voluntarily imposed to minimise any disruptions to normal Grid operations resulting from the image processing tasks. Furthermore, the tests confirm the Grid potential of the Grid architecture, while also showing some potential bottlenecks.

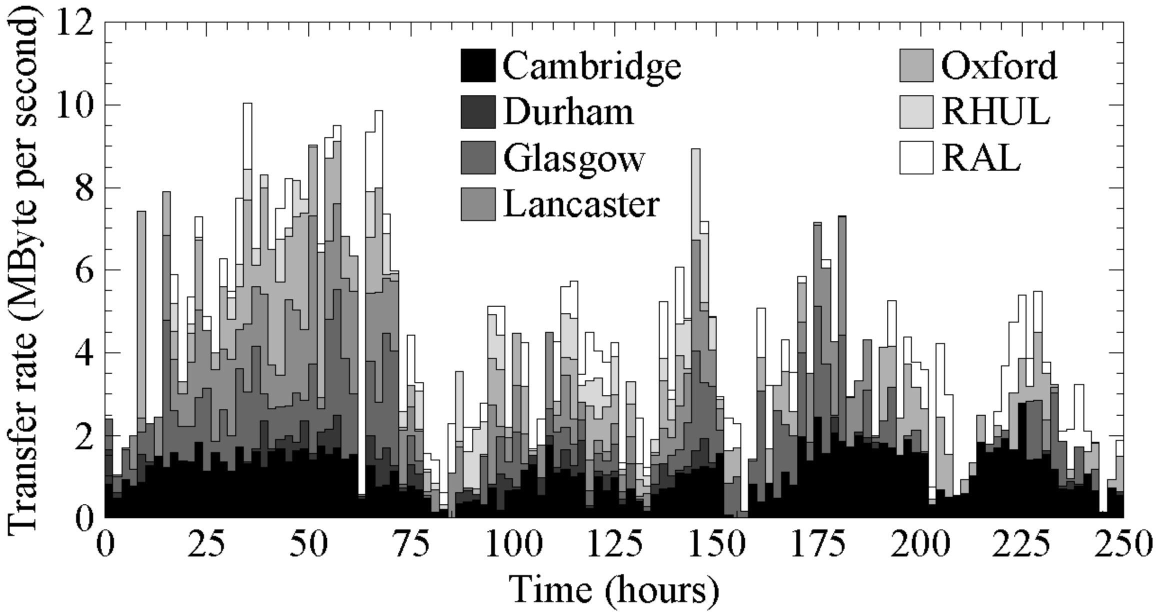

Figure 4: Plot of job transfer rates.

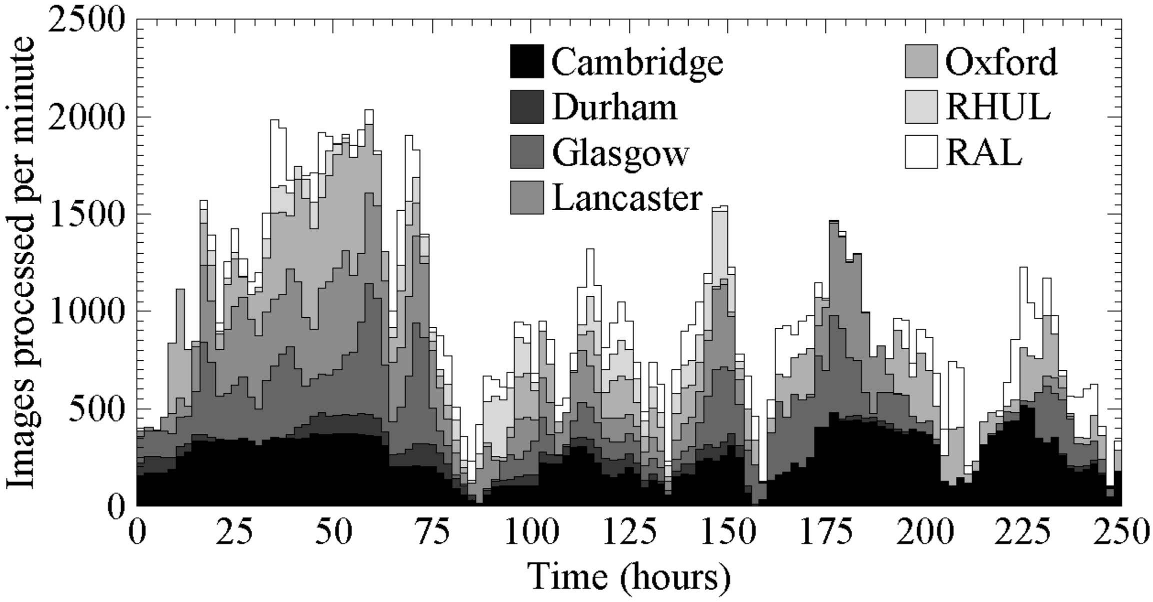

Figure 5: Processing rate as a function of time.

There were up to 500 jobs running in parallel, representing 500,000 images being queued or processed at the same time. Results were returned to the Imense servers at Cambridge via the Cambridge and Oxford Grid storage elements. A check after processing was completed showed that results were obtained for all but 2000 of the images in our sample, meaning a success rate of 99.99%. A sampling of the images for which no results were returned showed that these were due to damaged JPEG image file, and there were zero losses resulting from use of the Grid.

Figure 4 shows the distribution of job transfer latencies, while figure 5 plots the number of images that could be processed at each site per minute. Each bin plotted in these two figures corresponds to an interval of two hours. So the peak of 10 MByte per second means a transfer of about 70 GB or 290,000 images in 2 hours. The typical transfer rate of 4-6 MByte per second is fairly modest, but was achieved over a sustained period: 5 MByte per second for 10 days corresponds to over 4 TB.

It can be seen that the two distributions are similar and heavy-tailed, with some jobs being transferred and processed very quickly (at a rate of up to 2000 images an hour at some sites) while others experience significant delays before processing commences, in some cases of more than a week.

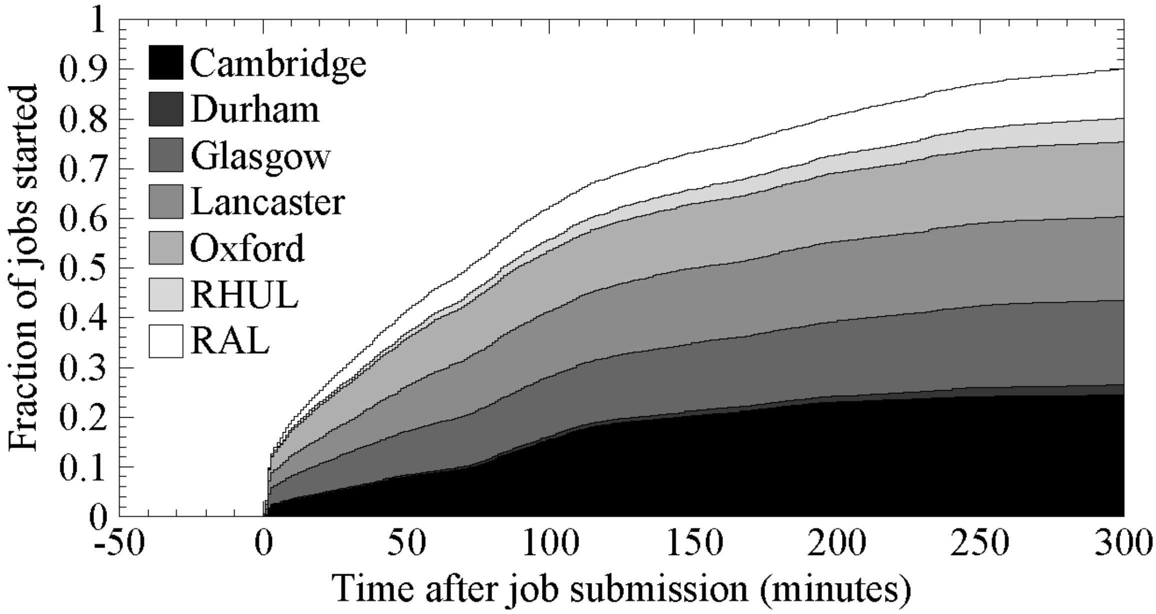

Figure 6: Fraction of jobs started as a function of time since submission.

This is further investigated in figure 6, which shows that for about 50% of jobs processing commenced one hour, whereas 10% of jobs were still queued after 5 hours had elapsed after initial submission. This statistic would be worrying in the case of an application requiring particular guarantees on maximum completions times for each job, but in the case of image analysis we are only concerned with total throughput.

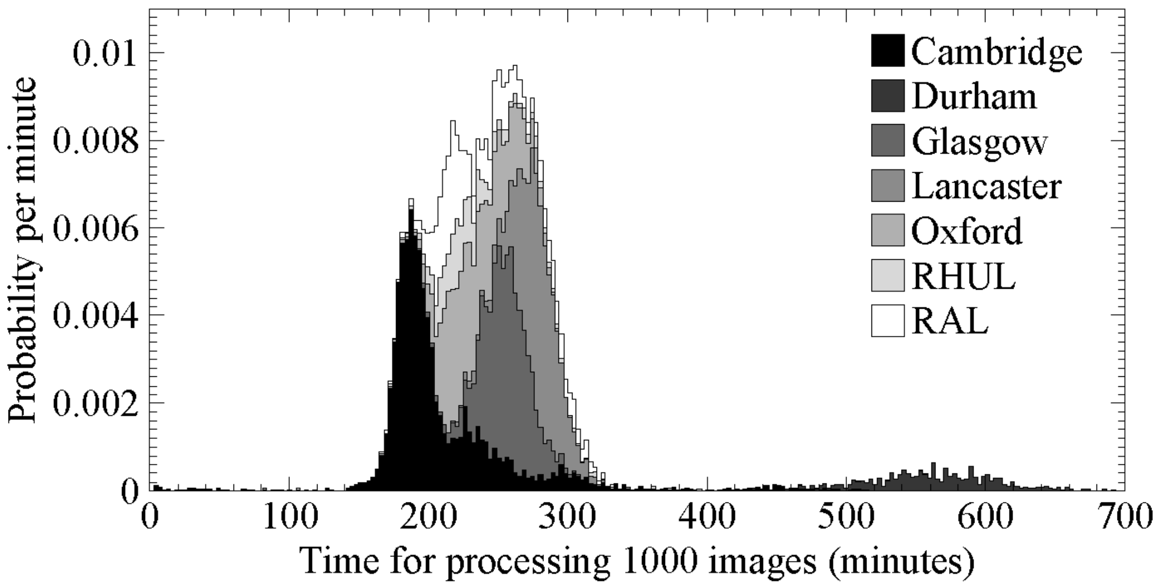

Figure 7: Distribution of time for processing a job containing 1000 images.

Actual processing times are plotted in figure 7 This shows that the time taken to process all images in a given job is similar across the 7 Grid sites, with a mean of just over 4 hours per job. However, the distribution again shows that a very small number of jobs can take considerably longer than this, most likely due to disruptions at a particular Grid node and peaks in demand due to higher-priority tasks been carried out on the Grid. We have also observed delays caused by temporary network problems and occasional issues with the Grid's proxy authentication mechanism (i.e. the users' proxy credentials at a given site expire before the job is executed).

5. Future Outlook

In the process of updating the Imense image analysis software, we encountered some problems with differences in the availability of standard software libraries at different Grid sites, even those having the same version of Scientific Linux (the version of Linux developed at CERN and widely deployed across the particle physics Grid). While we were able to overcome these issues, this experience has given us further impetus to evaluate the potential of virtualisation in Grid computing. Virtualisation is very attractive for service providers as a way of offering a variety of operating systems and configurations using a single set of machines. This ensures that the Grid environment executing a given user's jobs is identical to the user's local environment, thus overcoming many incompatibility and linking problems.

However, we are interested in understanding the overheads of the virtualisation layer. With this in mind, we are measuring the performance cost for the Imense software when running on a hosted virtual machine, as compared with running on a native operating system. We intend to evaluate this both with respect to hardware overheads (CPU and IO usage) and in terms of power consumption. Initial tests from running the Imense software on samples of several thousand images give encouraging results, suggesting that the overhead from the virtualisation solution is at the level of 10-15%. The tests were designed to measure both overall performance as well as throughput on particular processing tasks with different CPU and IO utilisation characteristics.

6. Conclusions

We described how a large distributed processing Grid (GridPP) has been used to apply a range of CBIR methods (provide by Imense Ltd.) to a substantial number of images. By massively distributing the required computational task across thousands of Grid nodes, we have achieved very high throughput at relatively low overheads.

By using the Ganga framework for job submission and management, it has been possible to port and deploy a large part of Imense's image analysis technology to the Grid. In this way we were able to harness the immense computational power of GridPP to analyse the content of and build a searchable index over 20 million high resolution photographic images, in addition to some 5 million images that were processed using Imense's own computing infrastructure.

However, our analysis also shows that the particle physics Grid architecture suffers from some performance bottlenecks which can cause per-job execution time to greatly exceed the mean. This implies that further work would be necessary to make it suitable for applications which require certain maximum per-job completion times. In the case of content based image analysis, the primary performance criterion is the overall throughput achieved by the system in terms of the number of images that can be processed over a given time frame irrespective of the time taken to process any given image. As such Grid processing has great potential for massively parallel content-based image retrieval and other tasks with similar performance requirements.

References

Brochu, F. et al. (2009) “Ganga: a tool for computational-task management and easy access to Grid resources", published in arXiv:0902.2685v1, submitted to Comp. Phys. Comm, 2009

Town, C.P. and Sinclair, D.A. (2004) "Language-based Querying of Image Collections on the basis of an Extensible Ontology", International Journal of Image and Vision Computing, Vol. 22/3, pp 251-267, 2004

Town, C.P. (2006) "Ontological Inference for Image and Video Analysis", International Journal of Machine Vision and Applications, Volume 17, Number 2, pp 94-115, 2006