# Karaage – Cluster account management and reporting

**Andrew Isaac[1], Sam Morrison[2]**
[1]Victorian Life Sciences Computation Initiative, Melbourne, Australia, aisaac@unimelb.edu.au
[2]Victorian Partnership for Advanced Computing, Melbourne, Australia, sam@vpac.org

## KARAAGE

In just over 12 months, the Victorian Life Science Computation Initiative (VLSCI) [1] has gained over 300 users, 80 projects and collected user feedback and usage data for four quarterly reports and one annual report. This has all been done through Karaage [2], an online account management and reporting system, developed at the Victorian Partnership for Advanced Computing (VPAC) [3]. When a user's access to the supercomputers has been approved, they are given a Karaage account. This allows users to avoid a major impediment to account management – the paper trail required to perform many account management tasks and to collect reporting information – by delegating account management and responsibility to the most suited people – the users. Users are assigned different levels of privilege. Through the user portal, for example, a nominated project manager can invite and approve new users, complete reports and questionnaires, and track project resource usage. Systems' administrators need only monitor user management and collate report information. This simplifies the entire account management and reporting process for users and administrators.

## ACCOUNT MANAGEMENT

Account management is broadly divided into administration and user tasks. Administrative tasks cover all creation, allocation, and authorisation activities, in addition to system monitoring, record keeping, and software and hardware resource management. User tasks can be at an individual level, such as usage checking, shell choice, password resetting, and agreeing to restricted software licences, or at a group level, such as managing group members and group resource usage.

Superficially many of these tasks appear simple, but providing a self-managed, on-line service avoids a great deal of the time and effort required to process these tasks. For example, software often has licence terms and conditions that must be agreed to by the end user prior to use. While allowing an end-user to agree to a software licence may seem like a simple process, it typically involves a sequence of e-mails starting with the user making a request, an administrator advising of, and the user agreeing to, the terms and conditions for that use, and the administrator modifying a database to assign the user to a group of users authorised to use the software, and then notifying the user that the process is complete. Karaage allows a user to identify software that has special licence conditions, review the licence conditions, and agree to the licence terms and conditions all through the user portal. This automatically takes care of assigning the user to the relevant software group and notifying the user. The process requires no administrator intervention. Moreover, a complete transaction history is recorded for audit purposes.

The ability to allow users to manage many authentication and authorisation tasks is achieved through Karaage modules that store account information using LDAP [4], Active Directory [5] or password files. This allows administrators to specify the account management tasks for different levels of users, allowing those with higher level access to manage their own authorisations and authentications for lower-level users. This has significant implications (as illustrated above) since users and project managers are in the best position to manage their own activities without having to wait for system administrators to perform the task for them.

Resource monitoring is another example of a simple feature that provides a great deal of information to the user. Karaage includes modules to interact with cluster schedulers such as PBS Torque [6], OGE/SGE [7], SLURM [8], and Windows HPC [9]. This allows Karaage to log and query database information from external sources, providing important system monitoring and historical information. The information displayed can be retrieved based on user, project, institute, and machine, thereby simplifying accounting, system management and reporting.

The various interactions between the cluster system components are illustrated in Figure 1. Both user and administration interfaces communicate with a central database that stores information from the clusters and resource accounting databases. Authentication and authorisation are controlled through direct interaction with a permissions database.

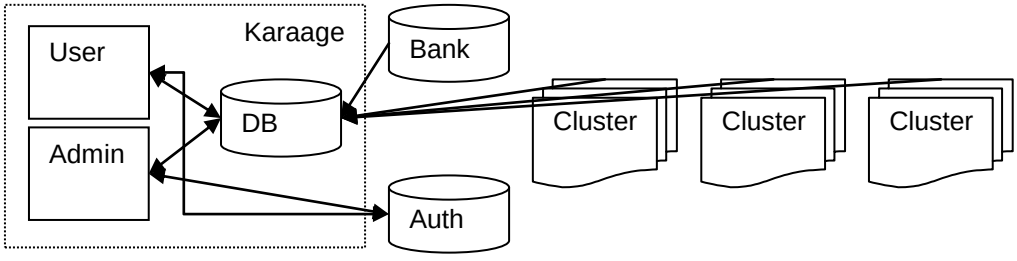


**Figure 1: Karaage interaction with cluster system components.**

## MEETING COMPLIANCE REQUIREMENTS

Peak computing facilities represent a significant investment and the stakeholders expect that their use is governed by appropriate policies and procedures and that reporting is accurate and informative. Karaage's integrated environment and common interface to the account management tools ensures the reliable application of the account management policy and procedure. Moreover, the same interface can be used to collect the reporting information required by users, administrators and funding bodies.

In Karaage, account management activities are divided into categories based on the organisational units that they relate to. These units include: machine categories; member institutes; projects; people; software; usage; and activities such as account applications and project reporting and surveys. This allows workflows to be developed based on higher-level organisational units relating to policy and procedure, rather than the low-level tools. The use of organisational units greatly simplifies the collation of resource usage and reporting information.

Reporting for such facilities can be quite onerous, and tools which assist this can save resources. VLSCI is required to report on the level of researcher interaction. Users submit quarterly usage reports to indicate the amount of 'in-kind' contributions people have made to the Initiative through the associated work being generated by access to the computational resources. These in-kind reports summarise the effort of each project member, detailing their salary level and contribution, delivering a bottom-line dollar value for in-kind contributions to the Initiative. The entire process is done on-line on a pre-populated form, greatly reducing the effort required by a project manager to collate such information. Similarly, end of year progress reports and user feedback surveys are all performed on-line through the same interface. All resource usage and reported content can be extracted and presented in various views (eg. bar graphs, pie charts) based on the organisational unit of interest. This greatly reduces the effort required to collect and extract salient information.

## CUSTOMISATION

Current users of Karaage include: VPAC; the Victorian Government Department of Primary Industries (DPI); La Trobe University; Multi-modal Australian ScienceS Imaging & Visualisation Environment (MASSIVE); VLSCI; and some commercial companies. The majority of these installations have used the inbuilt customisation to tailor the authentication and authorisation and cluster logging backends, in addition to the look and feel of the user portals.

VLSCI has additional reporting and policy and procedure requirements that are not provided by Karaage. These requirements are project-based account application, and quarterly in-kind contribution reporting. The majority of the effort in developing Karaage modules to provide these functions was not the coding, but the development of the policy and procedure and in-kind reporting workflows. Once these workflows were formalised, the customisation process then consisted of relating the organisational units to the underlying tasks, via the Karaage API. Customisation, however, still requires an advanced knowledge of Karaage and the supporting systems.

## CONCLUSION

Karaage is a web-based system that simplifies the workflow of account management. Built on the Django web framework [10], it provides web-based account management facilities to administrators and users. Its design allows users to manage their own accounts and projects. This is achieved by providing modules and middleware to connect the various sub-systems of compute resources and their administration systems. This not only streamlines authentication and authorisation, but also provides interaction with various databases to collect and collate information in a readily accessible manner, all through a single web-portal. Karaage is developed and supported by VPAC and released through the GPL v3 licence [11].

## REFERENCE


1. Victorian Life Sciences Computation Initiative (VLSCI) website: http://vlsci.org.au/, accessed 30th June 2011.
2. Karaage website: https://code.vpac.org/trac/karaage/
3. Victorian Partnership for Advanced Computing (VPAC) website: https://code.vpac.org/trac/karaage/
4. Open Lightweight Directory Access Protocol (LDAP) website: http://www.openldap.org/
5. Microsoft Active Directory website: http://www.microsoft.com/windowsserver2008/en/us/ad-main.aspx
6. Cluster Resources PBS Torque website: http://www.clusterresources.com/products/torque-resource-manager.php
7. Oracle Grid Engine website: http://www.oracle.com/us/products/tools/oracle-grid-engine-075549.html
8. SLURM website: https://computing.llnl.gov/linux/slurm/
9. Microsoft Windows HPC website: http://www.microsoft.com/hpc/
10. Django website: https://www.djangoproject.com/
11. GNU Software licence: http://www.gnu.org/licenses/

## ABOUT THE AUTHORS

Andrew Isaac is a specialist programmer at VLSCI. He has been involved in the development of Karaage modules specific to the policy and procedure of VLSCI. His research interest is in statistical and high throughput computing. He has a PhD in machine learning and has worked as a researcher in bioinformatics, autonomous systems, and human factors engineering.

Sam Morrison is a systems administrator and programmer at VPAC. He created Karaage in 2007 and has been the main developer since. He has assisted in the installation of Karaage at several institutions.