High scale, computationally distributed landslide modelling

Malcolm Anderson[a], Nick Brook[b], Liz Holcombe[a], Dave Newbold[b], Simon Metson[b], Michael Wallace[b]

[a] School of Geographical Sciences, University Road, Bristol, BS8 1SS
[b] HH Wills Physics Laboratory, Tyndall Avenue, Bristol, BS8 1TL

1. Introduction

Landslides are an increasing problem in developing countries [1]. Multiple landslides can be triggered by heavy rainfall resulting in loss of life, homes and critical infrastructure [2]. The costs of disruption to road networks can be several orders of magnitude higher than direct clean up and repair costs. Such disruption can limit the ability of a country to respond to the disaster. In many cases it is possible to reduce this risk by investigating the underlying risk drivers and investing in appropriate slope management and stabilisation measures prior to disasters [3]. Through computer simulation of individual slopes it is possible to predict the causes, timing and magnitude of landslides and estimate the potential physical impact. This information can form the basis of a cost benefit analysis of different risk reduction strategies, helping to ensure that the limited funds available for disaster prevention are used where they are most needed [4].

Geographical scientists at the University of Bristol have developed software that integrates a physically-based slope hydrology and stability model (CHASM) with an econometric model (QUESTA) in order to predict landslide risk over time [5]. The input data required by CHASM includes slope cross-section geometry, material strength and hydraulic properties, and rainfall data. The model simulates the dynamic physical processes affecting slope stability in order to predict the probability and magnitude of a landslide at that location. When coupled with QUESTA additional inputs are specified to enable modelling of the resulting risk to exposed road networks in terms of network disruption, engineering costs and road user costs. The landslide prediction capabilities of CHASM have been well established through numerous applications. Whilst CHASM and QUESTA remain a platform for research and code development, they are also applied ‘in the field’ to diagnose and solve landslide risk problems in developing countries. In particular these models are being used by government planners and engineers in the Eastern Caribbean to help understand and manage risk to communities and roads [6].

The software allows for a number of scenarios to be evaluated for each slope, accounting for data uncertainties, different engineering interventions, risk management approaches and rainfall patterns. Simulating multiple scenarios for a range of slope properties increases the fidelity of the resulting information. Individual scenarios can be computationally intensive, however each scenario is independent and so multiple scenarios can be executed in parallel.

As more simulations are carried out the amount of data generated increases and the overhead involved in managing simulation data becomes significant. This will become an even greater problem if multiple slopes are to be considered concurrently, as is required for effective disaster planning at national levels. The amount of data that could be generated is in the order of tens of terabytes. The larger number of simulations also results in long overall runtimes; a problem which is exacerbated by the modelling of increasingly complex scenarios.

Governments in developing countries typically lack the resources and infrastructure required to manage these issues. This means that the knowledge that could be gained by aggregating simulation results from many different scenarios across many different slopes remains hidden within the data and unavailable to those who could benefit from it. In order to address these data and workload management issues particle physicists from the University of Bristol are collaborating with the geographical scientists and developing methods for providing simple and effective access to landslide simulations and associated data.

2. Particle physicists and grid computing

Particle physicists have valuable experience in dealing with data complexity and management due to the scale of data generated by particle accelerators such as the Large Hadron Collider (LHC). The LHC generates tens of petabytes of data every year which is stored and analysed using the Worldwide LHC Computing Grid (WLCG) [7]. The WLCG provides the large-scale sharing of storage and compute resources required in order to manage the data and make it available to collaborators across hundreds of different institutions. LHC experiments develop their own software frameworks which sit on top of the WLCG and provide scientists with tools to help manage analysis jobs, simulation jobs and the data itself.

Tools and concepts from the WLCG are being used to drive the development of a Software-as-a-Service (SaaS) platform for providing access to hosted landslide simulation software and data. The platform provides advanced data management features and allows landslide simulations to be run on the WLCG thus dramatically reducing simulation runtimes through parallel execution.

3. Architecture and implementation

Users access a central database via a web browser, where simulation tasks can be composed from available input data. These tasks are broken down into individual jobs and submitted to the WLCG. Simulation results are stored in the database and a number of views allow users to query results and visualise data in such a way that answers high level questions such as "of all slopes where a landslide occurred for a given amount of rainfall, which have the highest economic impact?".

Jobs are submitted to the WLCG using a workload management system called DIRAC, which was developed to meet the requirements of the LHCb experiment [8]. DIRAC provides facilities for scheduling jobs transparently across a diverse set of resources and uses a plugin architecture which allows jobs to be submitted to compute resources including the WLCG, the National Grid Service, Amazon EC2 and local compute farms.

Traditional SaaS deployments require a reliable network connection between the server and the users, however as most landslides in the Eastern Caribbean region are triggered by tropical storms it is likely that telecommunications infrastructure will not be available when access to landslide data is urgently needed. To address this use case the SaaS platform uses bi-directional replication between clients and servers; this allows users of the software to have a personal copy of relevant data and for it to still be accessible in the event of a network failure. Figure 1 shows a conceptual overview of the architecture.

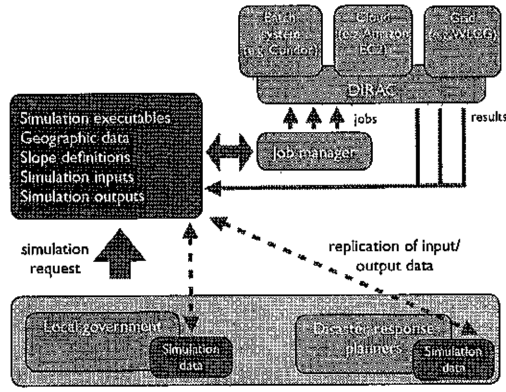


Figure 1: Overview of system architecture

The platform makes use of a number of open source projects which have been successfully used to address data management problems within particle physics. At the core is BigCouch [9], a fork of the Apache CouchDB document database [10] which uses concepts from Amazon's Dynamo database [11] to provide a clustered CouchDB instance that can be used to achieve data partitioning across multiple servers. CouchDB provides support for multi-master replication and is capable of running on a range of deployments from clusters of servers to laptops to mobile devices. All communication to CouchDB takes place over HTTP [12] via a RESTful interface [13].

The BigCouch cluster is used to store simulation input and output data, simulation requests and system state. Users have their own instances of CouchDB which store a relevant subset of the data. Simulation requests are written to these databases and replicated to the cluster. Once received by the cluster the requests are processed by a sentinel, a stateless daemon which creates the individual jobs required by the request. A typical request might be to simulate a particular slope with a range of cut angles, in order to find the optimum angle for reducing landslide risk to an acceptable level.

Individual jobs are processed by another sentinel which is responsible for submitting the jobs to DIRAC. The jobs are then managed by DIRAC, which takes care of scheduling, resource matching, job initialisation, file transfer and failure handling.

A third sentinel monitors system state, closing tasks when all jobs for a task are complete and flushing any temporary state data from the system. An overview of the complete implementation is shown in figure 2.

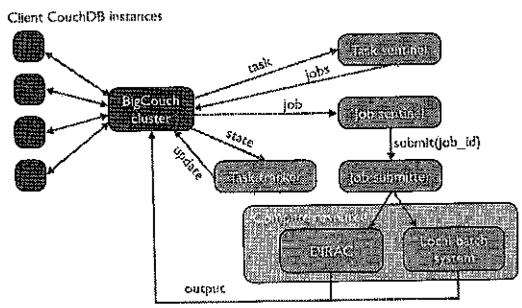


Figure 2. Implementation overview

Each job that DIRAC submits to the grid is a simple wrapper script which knows the unique ID of the job. Once the job wrapper lands on a worker node, the job specification is downloaded from the BigCouch cluster. This specifies the required exectuable as well as any additional input data that is required by the job. Once the executable and input data are downloaded, the simulation is run on the worker node and results files are parsed and uploaded into BigCouch.

The most critical point in the system, in terms of database load, is the download and upload of input and output data by the job wrapper. The jobs for a single task will start and finish at similar times, so read/write requests will hit the database from potentially tens of thousands of nodes almost concurrently. The effect will be similar to that of a distributed denial-of-service attack, therefore it is important to determine the maximum read/write capacity and design around it. Initial testing shows requests start to fail when we exceed 1000 concurrent requests. Early versions of the system are therefore configured to reject any tasks that would generate over 1000 jobs; future releases will implement a retry algorithm that will reduce the likelihood of the concurrency limit being exceeded for larger tasks.

4. Conclusion

Experience gained by physicists in using CouchDB as the core of a workflow management solution for the Compact Muon Solenoid (CMS) experiment [14] has been invaluable in designing this SaaS platform. The architecture and design are heavily influenced by patterns that have been tried and tested in production on real experiments. Through using DIRAC to handle the underlying compute resources we have greatly reduced the complexity involved in porting the CHASM and QUESTA applications to the WLCG.

Although the platform is currently at an alpha stage early usability tests with a prototype have been positive, demonstrating that the collaboration between two unrelated academic disciplines is having a positive effect on the real-world problems faced by developing countries.

References

[1] DfID, Disaster Risk Reduction: A Development Concern, *Department for International Development UK*, 2005

[2] UN, Landslides. Asia has the most; Americas, the deadliest; Europe the costliest; Experts seek ways to mitigate landslide losses, *UN University*, MR/E01/06/rev1, 2006

[3] Dai, F.C., Lee, C.F. and Ngai, Y.Y., Landslide risk assessment and management: an overview, *Engineering Geology*, Vol. 64(1). pp., 2002

[4] Anderson, M.G. and Holcombe, E.A., Sustainable landslide risk reduction in poorer countries, *Engineering Sustainability*, Vol. 159(1). pp. 23-30, 2006

[5] Holcombe, E.A., Modelling Landslide Risk on Highway Cut Slopes in Developing Countries, *Ph. D. University of Bristol UK*, 2006

[6] Anderson, M.G., and Holcombe, E.A., Assessing slope stability in unplanned settlements in developing countries, *Journal Environmental Management*, Vol. 85, pp. 101-111, 2007

[7] Donno, F., Litmaath, M., Data Management in WLCG and EGEE, *CERT-IT-Note-2008-002*, 2008

[8] Paterson, S.K., Tsaregorodtsev, A., et. al., DIRAC Optimized Workload Management, *International Conference on Computing in High Energy and Nuclear Physics (CHEP)*, 2007

[9] http://github.com/cloudant/bigcouch

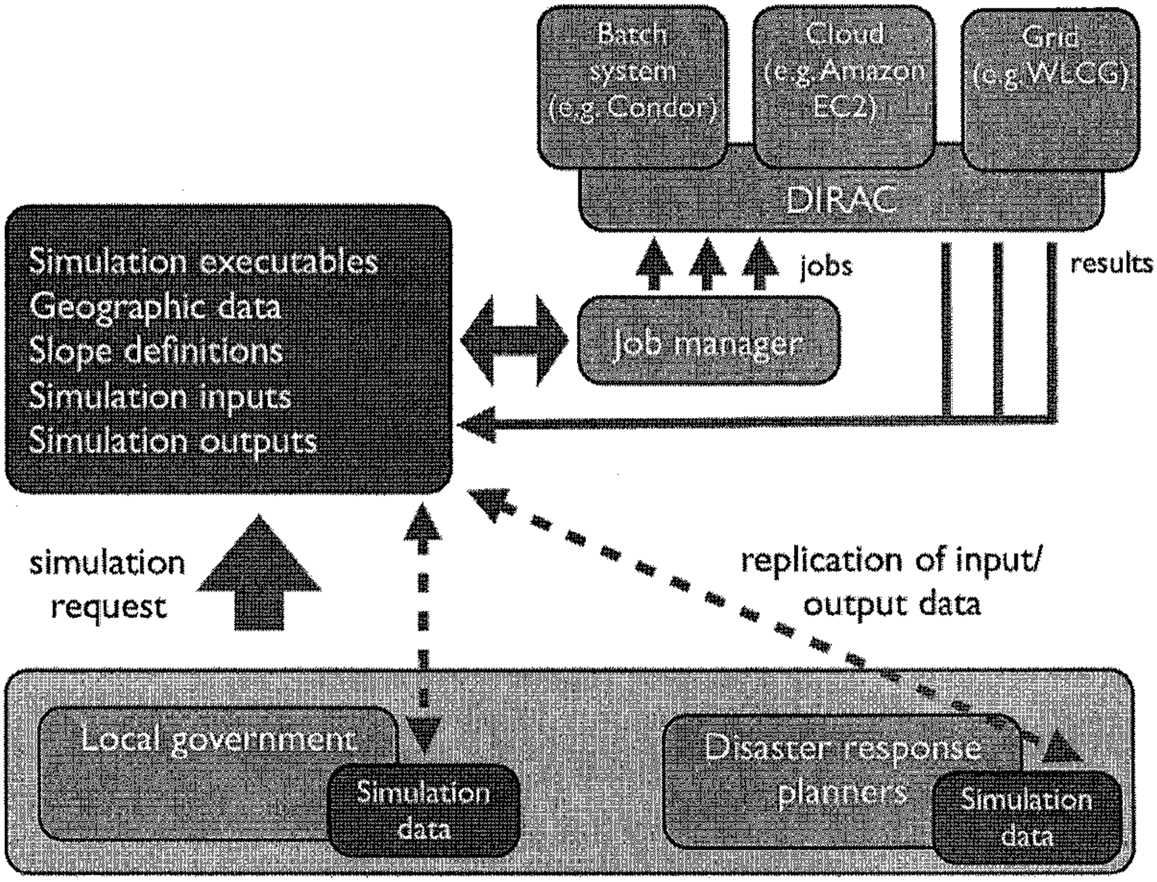
[10] http://couchdb.apache.org/

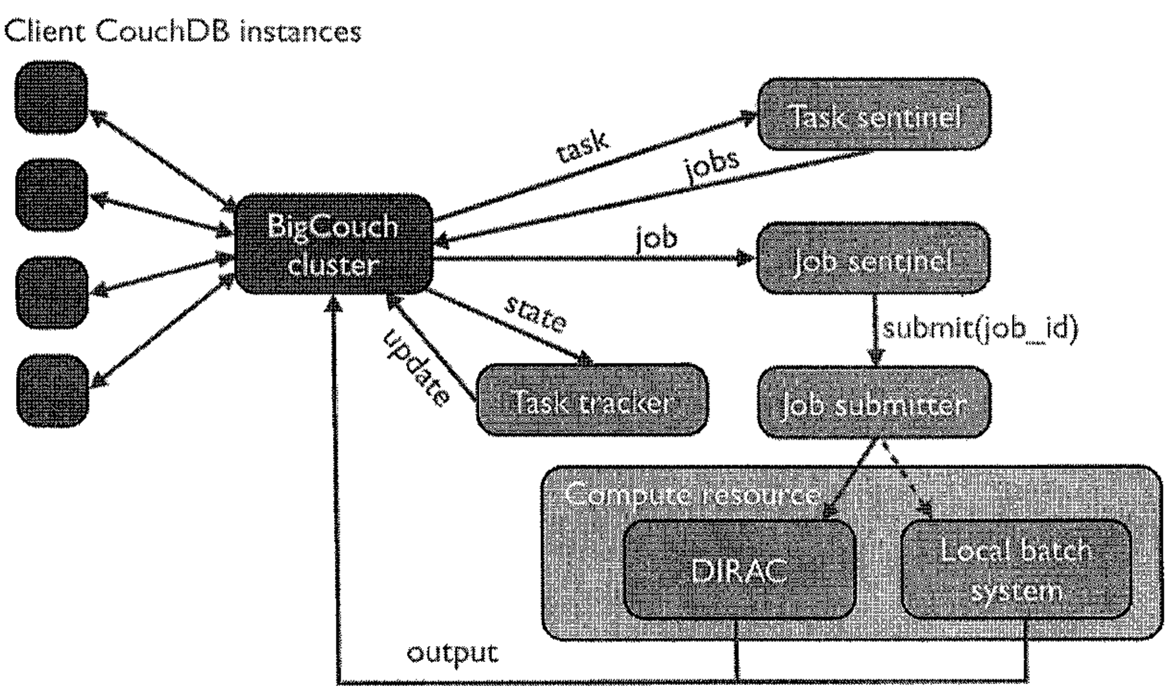
[11] DeCandia, G. et. al., Dynamo: Amazon's Highly Available Key-value Store, *Proceedings of the 21st ACM Symposium on Operating Systems Principles, Stevenson, WA*, 2007

[12] Fielding, R.T., et. al., Hypertext Transfer Protocol -- HTTP/1.1, *IETF Memo*, 1999

[13] Fielding, R.T., Architectural Styles and the Design of Network-based Software Architectures, *Ph. D. University of California USA*, 2000

[14] The CMS Collaboration, et. al., The CMS experiment at the CERN LHC, *JINST 3 S08004*, 2008


Batch system (e.g. Condor)
Cloud (e.g. Amazon EC2)
Grid (e.g. WLCG)
DIRAC
jobs
results
Simulation executables
Geographic data
Slope definitions
Simulation inputs
Simulation outputs
Job manager
simulation request
replication of input/ output data
Local government
Simulation data
Disaster response planners
Simulation data


Client CouchDB instances
task
jobs
Task sentinel
BigCouch cluster
job
Job sentinel
submit(job_id)
state
update
Task tracker
Job submitter
Compute resource
DIRAC
Local batch system
output