Tycoon activities at CERN openlab 2007 - 2008

José M. Dana

Jose.Dana@cern.ch

CERN openlab

20 January 2009

Table of Contents

Executive Summary

This document is the final report of the Tycoon two-year project between HP Labs and the CERN openlab. It describes the achievements over this period, summarizes the major outcomes, discusses issues and presents some conclusions.

In summary, the work was successful and the major technical objectives were achieved. The integration of Tycoon with the gLite software suite was successfully carried out, resulting in an interface (”Tycoon-gLite”) allowing easy deployments of Grid elements in a Tycoon environment. Scalability tests permitted to identify and report problems. These tests indicate that, despite certain issues, its design provides Tycoon with a good scalability potential.

On the dissemination front, Tycoon was presented at several Grid-released events and triggered great interest. However, the penetration of Tycoon in the user community may have been hampered by a couple of factors. One is related to technical weaknesses in the field of security which have been documented. The other lies in the fear by certain users of the possible subordination to the bank’s administrators. Both causes seem to have created with some potential users a mix of psychological and technical hesitations to adopt the system.

Due to a recent HP Labs restructuration the collaboration between HP and the CERN openlab regarding Tycoon was stopped in the fall of 2008.

Tycoon is a market-based system for trade of resources mainly developed at HP Labs (Palo Alto).

It is a resource management system that builds on virtualization to implement a specific resource management strategy: resource usage optimization via a bidding system and a global economy. In this aim, it benefits from the resource sharing capability of virtual machines, and in particular the ability of Xen to dynamically adjust the proportion of CPU cycles allocated to different VMs. In addition, as a system that brings up execution environments to remote users, it benefits from data isolation between VMs: a priori untrusted users can access VMs, install and run their own software without threatening each other nor the backing OS.

In 2006, HP proposed to include Tycoon in the projects carried out in partnership with CERN, within the CERN openlab framework. Following sections will give more insights about the computing challenges handled at CERN at that period and the benefits and constraints of working jointly on a project such as Tycoon in this unique environment.

The LHC

The Large Hadron Collider at CERN, near Geneva, is the largest scientific instrument on the planet. When it restarts operations in 2009, it will produce roughly 15 PetaBytes of data annually, which thousands of scientists around the world will analyse.

Discovering new fundamental particles and fields and analysing their properties with the LHC accelerator is possible only through statistical analysis of the massive amounts of data gathered by the LHC detectors ALICE, ATLAS, CMS and LHCb, and detailed comparison with compute-intensive theoretical simulations.

A traditional approach would be to centralize all of this capacity but in the case of the LHC, a novel globally distributed model for data storage and analysis was developed: a computing and data Grid. On 3rd october 2008, the Worldwide LHC Computing Grid (WLCG) consortium announced the readiness of this e-infrastructure conceived and designed to support the data challenge of the LHC, and with it the research of more than 9000 physicists around the globe.

The WLCG

The scheme used to organize and define the Worldwide LHC Computing Grid (WLCG) has three different tiers:

- **Tier-0**: CERN
- **Tier-1**: 11 large centres (dedicated 10 Gbits/sec optical network connections)
    - IN2P3 (Lyon, France), GridKa (Karlsruhe, Germany), TRIUMF (Vancouver, Canada), ASCC (Taipei, Taiwan), Fermilab (Illinois, USA), RAL (Oxford, UK), PIC (Barcelona, Spain), NIKHEF/SARA (Amsterdam, Netherlands), CNAF (Bologna, Italy), Nordic (Nordic countries, distributed Tier-1), Brookhaven (New York, USA).
- **Tier-2**: ~140 smaller centres

These so-called 'Tier-1'centres make the data available to the 'Tier-2' centres for specific analysis tasks. Individual scientists can then access the LHC data from their home country, using local computer clusters or even individual PCs.

The distribution of computing power and storage resources amongst all these tiers is shown in the table and figure below:

Table 1: Planned distribution of computing power and storage resources along the tiers. Note 1) About 100,000 processor cores.

	CERN	All Tier-1s	All Tier-2s	Total
CPU (MSPECint2000s)	25	56	61	142[1]
Disk (PetaBytes)	7	31	19	57
Tape (PetaBytes)	18	35		53

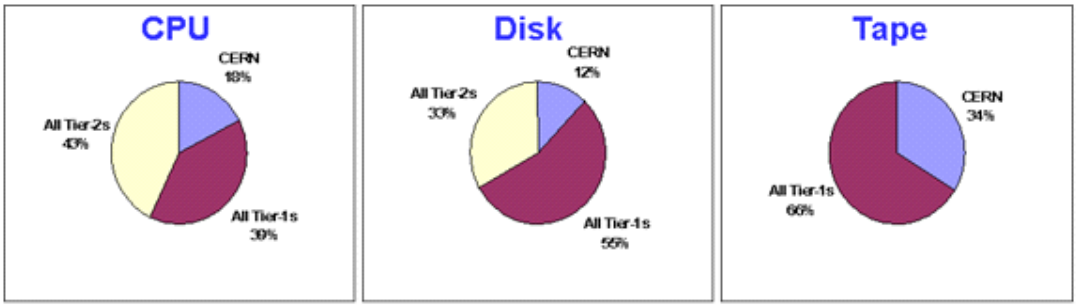

Figure 1: Distribution of computing power and storage resources along the tiers (2007).

The usage in September 2007 is shown below:

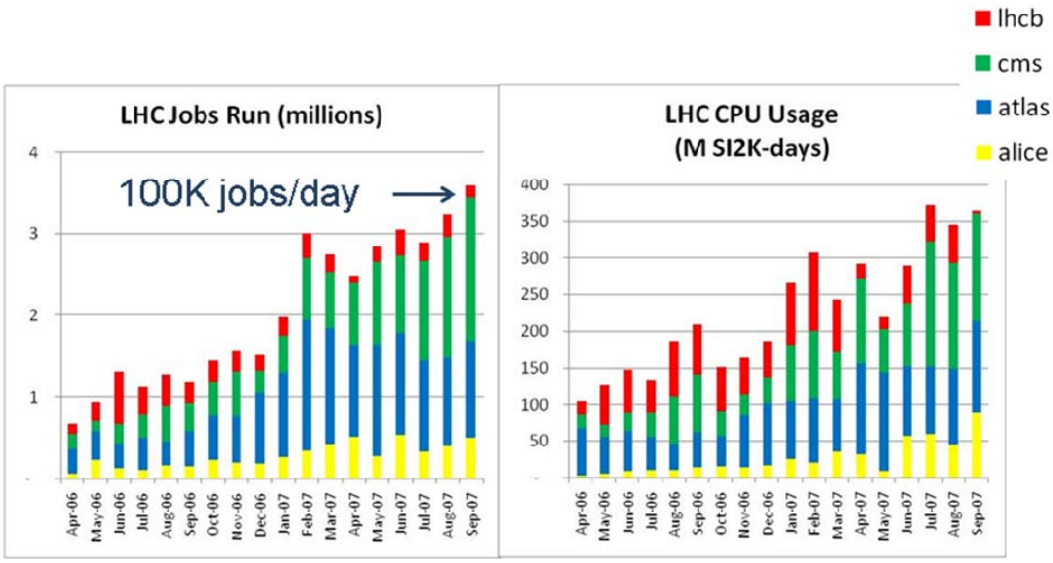

Figure 2: Number of jobs for all CERN experiments.

Figure 3: Real time monitor of Grid activity (http://gridportal.hep.ph.ic.ac.uk/rtm/)

CERN openlab is a framework for evaluating and integrating cutting edge technologies and services in partnership with industry. Being part of this framework and working closely with WLCG teams proved to be a perfect environment to further develop a project such a Tycoon. It was also an incomparable chance to demonstrate Tycoon potential to the academic, scientific and business world.

EGEE

EGEE Enabling Grids for E-sciencE (EGEE) is a project funded by the European Commission's Sixth Framework Programme through Directorate F: Emerging Technologies and Infrastructures, of the Directorate-General for Information Society and Media. It connects more than 70 institutions in 27 European countries to construct a multi-science Grid infrastructure for the European Research Area.

The project was launched using the pre-existing LHC Computing Grid (LCG) project as a springboard.

Building on recent developments in Grid Technology and earlier testbed projects such as EU DataGrid, its main aims are:

1. To build a secure, reliable and robust Grid infrastructure
2. To reengineer a light-weight middleware solution, gLite, specifically intended to be used by many different scientific disciplines
3. To attract, engage and support a wide range of users from science and industry, and provide them with extensive technical and training support.

EGEE was due to end on 31 March 2006, but a follow up project, EGEE-II, was started with the European Commission on 1 April 2006. EGEE-II has a larger consortium, with 91 contracting partners and 48 non-contracting partners from 32 countries, as well as expanded support for non-European participants and application communities. EGEE-II features a refocused middleware effort, increasing the integration of components from outside sources and putting more effort into integration and testing activities. The second phase ended on 30 April 2008.

At present, EGEE-III (that has now secured funding for the next two years) is the largest multi-disciplinary grid infrastructure in the world, which brings together more than 140 institutions to produce a reliable and scalable computing resource available to the European and global research community. Right now, it consists of approximately 300 sites in 50 countries and gives its 10,000 users access to 80,000 CPU cores around the world.

As EGEE promotes the collaboration between public and private institutions, the knowledge transfer and the use of the Grid as a platform for everyone, it is a very good platform to promote Tycoon and show it to the business world, despite the fact no real money can be used in the EGEE context.

GÉANT2

GÉANT2 is the high-bandwidth, academic Internet serving Europe's research and education community. Connecting over 30 million researchers with a multi-domain topology spanning 34 European countries and links to a number of other world regions, GÉANT2 is at the heart of global research networking in Europe.

GÉANT2 is co-funded by the European Commission and Europe's national research and education networks, and is managed by DANTE.

Both WLCG and EGEE use GÉANT2 and, therefore, no commercial transactions are allowed and it is used "only" with educational and scientific purposes. The implication is that Tycoon can be used inside the network as a tool for resource allocation and control of the usage but without the use of real money.

Tycoon activities at CERN openlab

Summary

In this chapter, we would like to describe some of the activities carried out at CERN openlab.

After more than one year of collaboration with HP Labs regarding Tycoon we have, amongst other things, 1) developed a Tycoon-gLite integration, 2) performed scalability tests (done with the involvement of a CERN openlab summer student) and 3) presented Tycoon at three different conferences. In addition, we have been collaborating with several institutions, promoting Tycoon and getting/giving help.

Collaborations

Several collaborations were initiated regarding Tycoon. Some of the most important ones have been described below.

- **HP Labs (Palo Alto)**: This collaboration was of high importance. Amongst all the work it is worth highlighting:
    - Port to SLC4 (Scientific Linux CERN 4): During a visit to HP Labs (Palo Alto) in April 2007, we expressed the need to port Tycoon to SLC4, the

only supported distribution at CERN. The port was finished by the end of September.

- Scalability tests done using HP Labs machines: During the summer, scalability tests were done using HP Labs machines. During those tests, some bugs were discovered and some experiments that will be very interesting for potential Tycoon users were done (the results are described later).

- Support: The Tycoon team helped us at the beginning of our collaboration answering all questions. They also gave us credits and access to HP Labs machines (using Tycoon). After a visit to HP Labs they give us rights in the Tycoon's wiki webpage.

- **EGEE**: EGEE is a European project hosted at CERN and our relations with them is very close. We have had several meetings with EGEE members (during EGEE conferences, informal meetings at CERN, etc.). This collaboration has been crucial during the development of Tycoon-gLite and has helped a lot during the EGEE conferences. We could point out the following:

  - Support with the EGEE and Tycoon integration:

    - Initial set-up and administration of our Grid testbed: During the initial set-up of our Grid testbed (basically a group of machines that give us a minimal configuration for deployment of Grid nodes) gLite developers helped us with their experience in administration of gLite nodes. Due to our special requirements (minimal number of nodes/machines but with real Grid capabilities) several issues were found. A graphical representation of the testbed is shown below. Machines inside the blue box have been deployed using Tycoon-gLite while the rest are standard physical machines with gLite services.

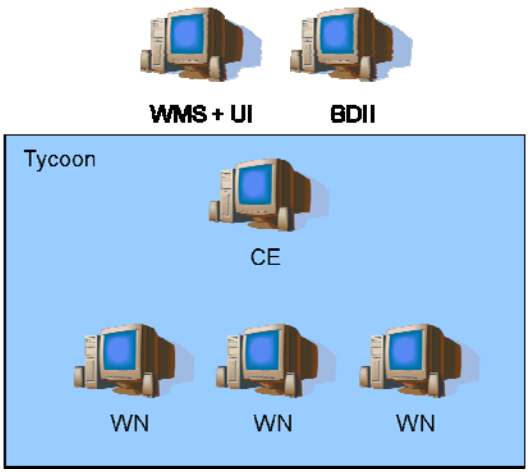

Figure 4: Graphical representation of the Grid testbed at CERN.

- Configuration of the Grid nodes, specially the Worker Node and the Computing Element: Images ready for deployment using Tycoon have been created with gLite installed on them (one for the WN and another one for the CE). During the installation and configuration of those images, several problems were found. Most of them were fixed with the help of gLite developers.

- Support during the EGEE conferences: During the EGEE conferences we were introduced in the Business Track (EGEE Conference '07).

- **BalticGrid**: The initial idea adopted for our Tycoon-gLite integration came from BalticGrid. We received a lot of relevant information and we were planning new scalability tests merging our respective Grid environments using Tycoon. Unfortunately, BalticGrid's Project Director changed during the collaboration when Per Oster moved to EGI (European Grid Initiative). After that we contact Ake Edlund (his successor at BalticGrid) and our collaboration was resumed.

  - Initial idea for Tycoon-gLite integration: Our initial idea for our Tycoon-gLite integration was adopted from a report written by BalticGrid members. During the development we made some modifications but the initial concept remained.

  - Support: During the design work we received some tips and explanations from BalticGrid member that helped to understand their model, advantages and disadvantages.

- **Constellation Technologies**: Constellation Technologies is a company that is interested in auto deployment and configuration of Grid nodes within a market-based environment. They are very interested in Tycoon and we could

begin a closer collaboration with them in the future. The situation with Constellation Technologies will be described more in-depth later.

- Joint development of Tycoon-gLite: Constellation Technologies needs some tool for deployment and control of resources in their market-based Grid environment. They attended one of the conferences where we presented Tycoon and think that there is a good opportunity for merging our efforts and moving Tycoon to the business world. They proved to be extremely interested in our Tycoon-gLite integration and could go one step further moving it to a production environment.

As underlined beforehand, it is important to find more users in order to help find bugs and improve the platform. Attending conferences are a good way to find people and companies interested in the project and willing to collaborate.

Promoting Tycoon is as important as improving it technically. We have found out that this platform awareness is very low: more collaborators would also mean more popularity.

Tycoon-gLite integration

Tycoon-gLite is the name of the platform that has been under development at CERN openlab during the last two years.

Tycoon-gLite has been developed in Python and has an easy-to-use user interface based in GTK (The GIMP Toolkit). It is an interface that merges Tycoon with a gLite based Grid environment (e.g. EGEE), acting as a bridge between Tycoon and the machines (auctioneers) that are ready for deployment of “special” images. It allows the deployment of CEs (Computing Elements) and WNs (Worker Nodes) on-demand using Tycoon’s market-based allocation system. A graphical representation of the system is shown below:

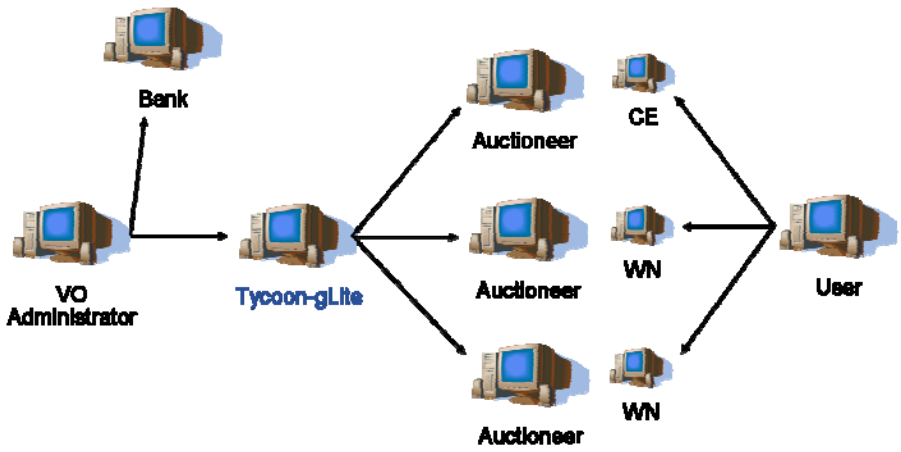

Figure 5: Graphical representation of the Tycoon-gLite platform.

Tycoon-gLite is not a modification of Tycoon. Therefore, there is no need to modify the code every time Tycoon developers change Tycoon's API or internal functionality. Tycoon-gLite must be seen as an interface that offers the possibility of deploying a set of CEs and WNs ready to work for the user. Tycoon-gLite manages the host certificates, creates the necessary configuration files, deploys the nodes and configures them on-the-fly (all the communication between the centralized Tycoon-gLite service and the clients and servers is encrypted using SSL). Once done, Tycoon operates as usual.

For the configuration step, Tycoon-gLite uses YAIM (Yaim Ain't an Installation Manager), which is a well known application used for configuration of Grid sites and developed within the EGEE project. For the development and testing of our platform, a special Grid testbed in our computer center at CERN was established.

The user connects directly to the Tycoon-gLite service (using encrypted communication) in order to get information about available machines. The service maintains an up to date database with information about all the available nodes and their type (CE, WN, SE, etc.).

Afterwards and, as seen in the screenshots of the client of our application shown below, the user can drag and drop hostnames for his CEs and WNs from the boxes on the left to the ones on the right in order to deploy the Grid elements in his favorite physical machines. Afterwards, the Tycoon-gLite platform generates the necessary files and steps for the auto deployment and configuration of the virtual machines.

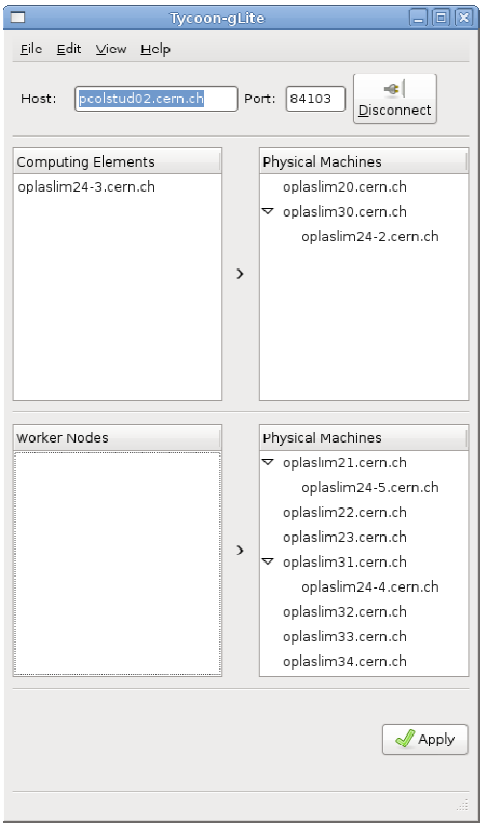

Figure 6: Scheme selection using Tycoon-gLite graphical interface.

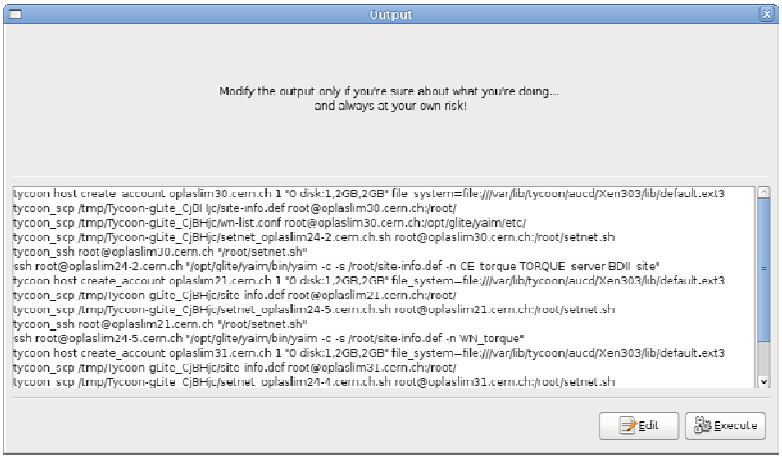

Figure 7: Output generated by Tycoon-gLite for deployment and configuration of the selected scheme.

As we already mentioned before, in order to develop and test Tycoon-gLite, a complete testbed was installed and configured in our cluster in the CERN openlab.

It was necessary to install different Grid services (the WMS, the BDII and a UI were installed for the simplest setup) in order to test our integration within a real Grid environment. For the rest of the necessary services (e.g. VOMS) we used the EGEE infrastructure.

During the first part of 2008 we continued with the development of the Tycoon-gLite integration. We began the support for deployment of more Grid Services (i.e. Storage Elements) and a more elaborated monitoring system.

The development of the platform stopped when HP decided to stop the collaboration. Because of that, the support for Storage Elements has been finished but could not been tested properly and the monitoring system has been improved but not as much as we would have liked to.

The testbed installed at CERN for testing of Tycoon-gLite uses a small part of the WLCG infrastructure. One of the nodes in use is the VOMS (Virtual Organization Management System) server. The VOMS server manages the Virtual Organizations that are allowed to use the Grid (only people within a registered Virtual Organization can ask for user certificates). It means that Virtual Organizations that are not inside WLCG cannot use our platform. Next step would have been to install our own VOMS and to add support in Tycoon-gLite, so that support of non-existent Virtual Organization could be possible.

This Tycoon-gLite integration system was presented during the EGEE User Forum held in Manchester (May 2007), the EGEE Conference held in Budapest (October 2007) and in the Distributed Computing Workshop held in London (May 2008).

More information about Virtual Organizations and WLCG users registration can be found at http://lcg.web.cern.ch/lcg/users/registration/VO.html .

Scalability Tests

Some scalability tests were carried out in order to probe the value of Tycoon within a Grid environment.

Thanks to the CERN openlab summer students programme a student (Andrea Sottoriva) worked on Tycoon during the summer of 2007.

Under my supervision he performed several scalability tests of the Tycoon platform and its bank. Moreover, he discovered some bugs that we sent to the Tycoon developer's team in HP Labs.

The most important bug was related with the bank. The bank was not operating correctly under some situations and users were earning money (see http://tycoon-dev.hpl.hp.com/ticket/345). A ticket was sent to the Tycoon's webpage and an email to Kevin Lai in order to report this important bug that was solved a few days after our submission. He also found some other bugs comprising race conditions that were reported as well.

In summary, the work done by the student comprised:

- Scalability tests (Tycoon platform, Tycoon's bank): Tests of the Tycoon platform were done by deploying a large amount of Tycoon virtual machines in a very short period of time (seconds). The idea was to measure the scalability of the system to see if it was ready for a large scale environment (specially the Grid). The Tycoon's bank was tested separately trying to stress it doing a large amount of requests, again, in a very short period of time. Tycoon's bank is one of the critical parts of the entire Tycoon platform. Tycoon is a distributed system but, however, it has two centralized systems: 1) the SLS and 2) the bank. If either the SLS or the bank were a bottleneck then the scalability of the entire platform would be compromised. The results of the scalability tests are shown below in Figures 6, 7 and 8.

- Bugs report: In addition to the problem with the bank (that was discovered thanks to our scalability tests) there were some problems with servers at HP Labs that were reported immediately. More scalability tests would help to find race conditions and security problems in Tycoon, like the one with the bank. It is why Tycoon really needs more users and companies doing an extensive evaluation of the platform.

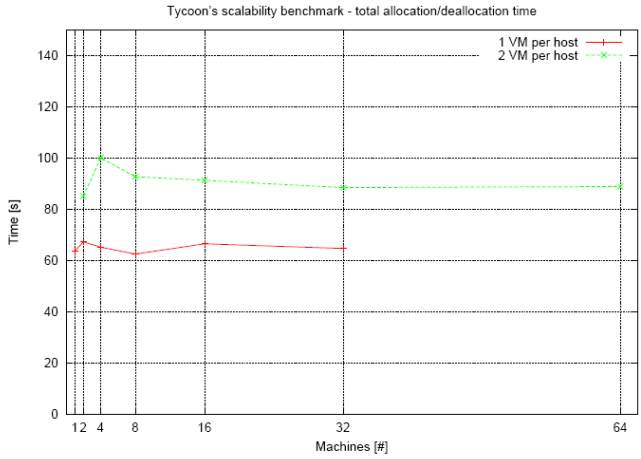

Figure 8: Total allocation/deallocation time of resources using Tycoon.

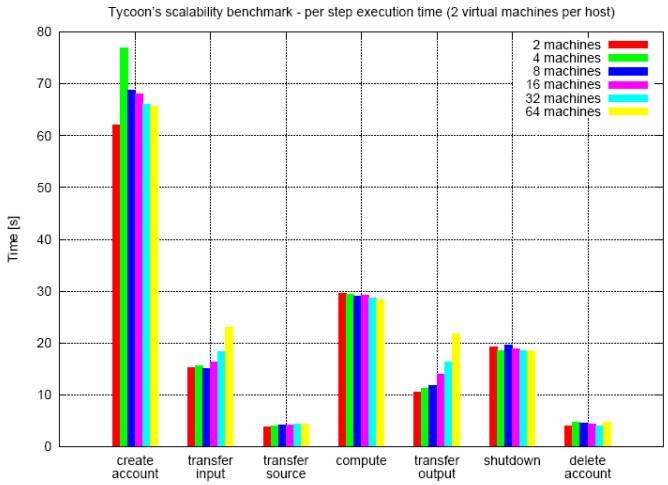

Figure 9: Execution times for deployment of 2 virtual machines per host.

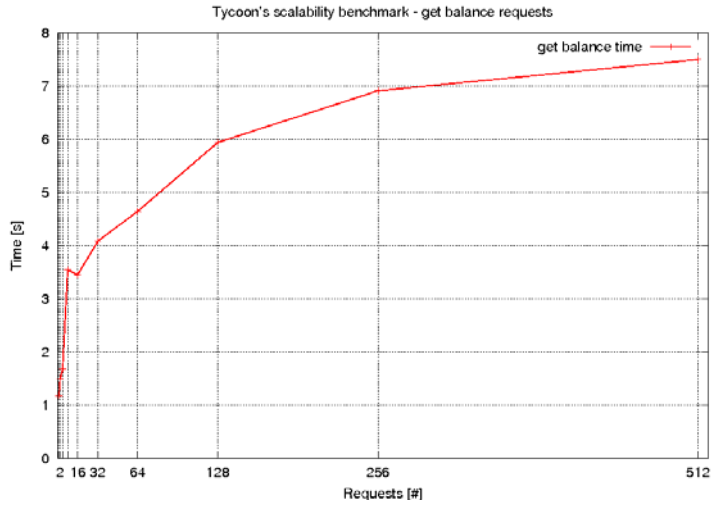

Figure 10: Tycoon's bank scalability.

As we can see in the figures above the system scales well and there are no significant bottlenecks.

However, we can appreciate an exponential increase of the transfer time in columns “transfer input” and “transfer output” in Figure 9. It’s due to our system and not to the Tycoon infrastructure installed in Palo Alto. In this case our system is the bottleneck simply because all the benchmarks were launched in the same source machine.

It was decided to hire a CERN openlab summer student in order to perform further scalability tests of the Tycoon-gLite integration in 2008.

We would have liked to involve HP Labs and BalticGrid in these tests, trying to reproduce a real world situation (three different sites in different geographical locations sharing computational and storage resources). Our initial results (testing Tycoon as a standalone application) are very promising, therefore we expect a similar behavior using Tycoon as a tool for deployment of Grid nodes.

Scalability is a crucial factor when we speak about distributed computing. People in EGEE, specially the gLite developers, care a lot about scalability and the possible bottlenecks. The question *“does it scale?”* has appeared in all the conversations we had with them and also with some companies.

Issues Concerning Security and Trust

Some security issues were identified during this collaboration.

One of these problems is the fact that every port in a Tycoon auctioneer has to be opened in order to share resources with others. The idea of allowing connections to any port of the machines in a local network like the one we have at CERN is unacceptable. This problem is one of the biggest worries of the potential adopters.

There are technical ways to solve the problem (i.e. using a proxy). However, we think that a better solution would be a modification in the design of the platform. The initial impression people have when using a new application is very important and usually decisive. The first time that someone uses Tycoon he will discover that it is necessary to open the entire machine, and this is not always possible or desirable. Even if the user knows how to solve the problem, it will take some time and he could change to another platform. Re-designing Tycoon and solving this problem would help to attract users and would solve one of the biggest worries.

People also express worries about the bank. “Who controls the bank?” and “how are transactions managed?” are both repetitive questions during conferences and meetings. The one that controls the bank will also control the number of credits available. For example, if we install our own bank at CERN, the bank’s administrator will have the possibility of using any “tycoonized” machine (registered in his bank) for free, just “creating” credits. This is somewhat like a National Bank that has the possibility to print money. Obviously, trust is paramount.

Some problems related with race conditions and bugs in the source code were found and Tycoon developers are trying to solve them as soon as possible. Nevertheless, the Tycoon webpage still has a lot of active tickets regarding the bank, auctioneer, etc. that have not been fixed yet.

Conferences

Tycoon was presented at three different conferences:

- *EGEE User Forum '07*, 7-11 May 2007, Manchester (United Kingdom). http://indico.cern.ch/contributionDisplay.py?contribId=92&sessionId=20&confId=7247
- *EGEE Conference '07*, 1-5 October 2007, Budapest (Hungary). http://indico.cern.ch/contributionDisplay.py?contribId=48&sessionId=18&confId=18714
- *Distributed Computing Workshop*, 21 May 2008, London (UK). http://www.qi3.co.uk/events/event.asp?EventID=193

The conferences were good opportunities to officially present Tycoon to the EGEE community. We presented Tycoon inside the Business Track, getting the attention of the business community.

In the last few years, increasing interest in market-based environments has been observed. The sharing and renting of resources are growing and, therefore, an environment where both the use and configuration of resources are well-defined is needed. A lot of people were very interested in the idea of auto deployment and configuration of Grid nodes on-demand using a market-based approach. The interest in market-based environments is growing. In fact, during the EGEE Conference '07, some people did presentations about similar kinds of environments.

Presentations of Tycoon in international conferences are important, given the fact that more and more companies are taking an interest in market-driven approaches.

It is necessary to promote both Tycoon and Tycoon-gLite in order to attract more collaborators, users and companies. International conferences seem to be one of the best ways to promote the Tycoon project, especially when also private companies are present.

Tycoon-gLite was also presented in the Distributed Computing Workshop held in London (UK) the 21st May 2008. One of the topics of the workshop was "Cloud Computing" and a there were a few people interested in market-based allocation and Tycoon.

EGEE conferences are especially interesting. EGEE is the project developing the gLite middleware and the presence of companies at its conferences is increasing

every year. Moreover, EGEE has proved its potential and, currently, it is the biggest Grid project in the world.

One conclusion from the EGEE Conferences is that itis not easy to convince people to embrace market-based models. However, more and more people are becoming aware of the issue, adopting different solutions and understanding that the classical model with a fixed price is outdated. If Tycoon could find a good position within the business world then possibilities could increase exponentially. As said before, the situation is not easy but the market is looking for alternatives and Tycoon is one of them… and it is already working!

Constellation Technologies

Constellation Technologies is a company interested in the deployment of Grid resources (using the gLite middleware) in a market-driven approach.

Their idea is basically the same as the one behind Tycoon: auto-deployment of resources and a dynamic market with bids and clients acting as auctioneers when necessary. Their goal is to integrate Tycoon with gLite (basically what we already did at CERN openlab) in order to deploy and auto-configure nodes on-demand.

However, the academic and business worlds are very different. We cannot always compare the requirements and demands coming from companies with those coming from users and developers within academic environments. This is why we think that this collaboration will help improving Tycoon-gLite integration and Tycoon itself. Constellation Technologies could be the first company using Tycoon as a real business tool, attracting users and other companies to this market-driven approach and helping Tycoon development (by providing more feedback, real life demands, etc.).

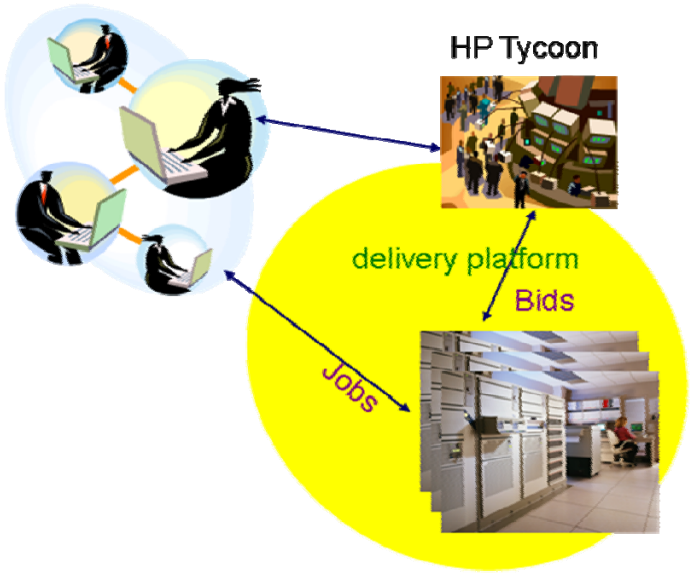

Figure 11: Graphical representation of the system that Constellation Technologies used during their presentation.

The Cloud Computing Initiative

"Cloud Computing" is a style of computing where IT-related capabilities are provided "as a service", allowing users to access technology-enabled services "in the cloud" without knowledge of, expertise with, or control over the technology infrastructure that supports them.

HP Labs is deploying (together with other companies like Intel, Yahoo, etc.) an open, scalable, secure, globally distributed, Internet-scale testing environment designed to encourage research on the software, data center management and hardware issues associated with cloud computing.

There is a federation of "Centers of Excellence":

- UIUC, Singapore IDA, KIT: 3 initial CoE.
- HP, Intel, Yahoo: 3 initial sponsors with CoE.
- Each center: 1000-4000 cores and up to PB storage.

Tycoon is being used as a resource allocation system in "the cloud". We believe that our collaboration with HP helped Tycoon to be part of this new research initiative.

Figure 12: VideoToon demo.

Final Remarks

Tycoon is a platform with considerable potential. During the EGEE User Forum, the EGEE Conference and the Distributed Computing Workshop we had the opportunity to meet many people very interested in a dynamic market-driven approach. However, we realize that Tycoon needs more users and institutions involved: More users would mean more feedback and an increase of the popularity. Regrettably, at the moment, people that are familiar with market-based resource allocation systems know what the other companies are doing (e.g. Amazon) but they do not know anything about Tycoon. We consider that raising Tycoon awareness is vital to its future success.

As already mentioned, Tycoon cannot be used within EGEE for commercial purposes but this does not imply that it cannot be deployed for research or scientific use. In addition, EGEE could work as a bridge between the industry and Tycoon, moving it one step forward.

During the past year, Tycoon developers have ported the platform to modern Linux distributions and hence have solved an outstanding compatibility problem. We strongly believe that Tycoon developers' number one priority should now be the security of the system. This is crucial that people have trust in a system that is dealing with their money and this implies that security is paramount. Some people seem worried about the security of the implementation (i.e. race conditions that could break the bank's transaction system). We know that some developers in BalticGrid are working on these issues, and solving them will be a very important contribution. The problems regarding the ports that have to be opened in a Tycoon auctioneer, the security of the bank's transactions, etc. must also be addressed.

Another important point is support. In the event of strong success of Tycoon, are we ready for a good quality support? This is key to ensure that the users do not waste their time while trying to use this tool: A fast and efficient support is needed.

Even though the amount of users is still limited at the moment (developers and some testers) this question could become a central one for HP Labs as the success of the tool and collaborations (like the one with Constellation Technologies) go further.