Applications Grid-Enabling Experience – The Soft Aspects

Thong-Tiong Choo, Hing-Yan Lee, W.C. Wong & Jon Khee-Erng Lau

National Grid Office, 21 Heng Mui Keng Terrace,
Singapore 119613

{thongtiong | hingyan | lwong | jonlau}@ngp.org.sg

Abstract

This paper shares lessons and experience from the grid trenches, the result of over a year long operation of the National Grid Competency Centre. The centre has been set up to help users of the National Grid Pilot Platform to grid-enable their applications in the short-term and raising the users' own skills profile in the long-term. In particular, we focus on the soft aspects such as project management, dealing with user expectations, and promoting usage.

Keyword: grid applications, grid-enabling, gridification, user management

I. Background

The National Grid Pilot Platform (NGPP) is the initial phase of the cyber infrastructure that has been set up in support of the National Grid vision to realise "a Singapore where computer resources can be connected together via a high-speed network such that these resources can be shared in a secure, reliable & efficient manner by authenticated users for education, commercial, entertainment, R&D, national security & other purposes so as to improve the economic & technological competitiveness of Singapore & also the quality of life in Singapore."

The NGPP comprises heterogeneous compute resources made available by research institutes and universities, connected together via 1 Gbps lines. They include Intel Itanium-2, Intel Xeon, Compaq Proliant, SunFire, and IBM Regatta. These compute resources are being made available by their owners, who retain management control over them. Today nearly 500 CPUs are linked to the NGPP. The NGPP Hub has Internet 2 connectivity to the US, Japan and Korea.

Figure 1 : NGPP Infrastructure

No	Organization	Machine Name	Processor	Total CPUs
1	NTU	SUN	UltraSPARC III	2
2	NTU	SUN	UltraSPARC IIIi	4
3	NTU	SUN	UltraSPARCIii	4
4	NUS-SVU	ATLAS	Xeon	32
5	NUS-SMA	HYDRA3	Itanium2	60
6	IHPC	NEUMANN	Power4	16
7	BII	VIPER	Pentium 3/4	128
8	NGPP Hub	IBM	Xeon	4
9	NGPP Hub	HP	Itanium2	4
10	NGPP Hub	SUN	UltraSPARC IIIi	1
11	NGPP Hub	GUAVA	Xeon	14
12	GOG	MELON	Xeon	72
13	GOG	SOURSOP	Itanium2	78
14	IHPC	GENESIS	Opteron	8
15	BII	SUNBEAR	Opteron	8
16	IHPC	LIME	Itanium2	6
17	NTU	SURYA	Pentium 3/4	25
			Total	**466**

Table 1: Resources on NGPP

To manage the NGPP, the National Grid Office (NGO) oversees the following two centres:

a) National Grid Operations Centre (NGOC); and
b) National Grid Competency Centre (NGCC).

The primary objective of the NGOC is to ensure the smooth operations of the NGPP. Its functions include monitoring resource usage, network and grid security, issue of digital certificates, and trouble-shooting. The work of the NGOC does not fall within the scope of this paper, while the experience of the NGCC will be detailed below.

II. Introduction

The NGCC was established in late 2003 with the primary objective to raise the competency in the skills, expertise and experience level in Grid Computing

The NGCC achieves this objective through:

a) Conducting training courses, hands-on workshops, seminars and events that provide opportunities for knowledge and skills transfer; and
b) Providing advice and manpower resources to help users on grid-enabling their applications.

As manpower resources to undertake grid-enabling of applications are limited, the emphasis has been "teaching users how to fish, rather than giving them fishes." In time, we hope that the manpower resources to undertake gridification will diminish and changes from hands-on to an advisory role.

Under the National Grid (Phase 1), an integral part of the strategy is to carry out grid-enabling of applications so that they can run on the NGPP. As the target beneficiaries of Phase 1 is the R&D community, the users typically write or have their own applications. Without grid-enabling these applications, there will be few "vehicles" to run on the grid "highway".

The efforts for grid-enabling consist of two main steps: inviting applications and grid-enabling selected applications. We will elaborate on these steps as part of the NGCC Grid-Enabling Applications Lifecycle (GEAL). This life cycle process has been evolved through our grid-enabling efforts to make it more effective.

III. Inviting Applications

The invitation of applications is undertaken by NGCC staff members. Road shows are conducted at the user organizations (such as research institutions and centres as well as institutes of higher learning) where the benefits of grid computing are expounded. At such forums, invitations for use of the compute resources on the NGPP are extended as well as the availability of assistance in grid-enabling of selected applications.

Within the National Grid framework, Virtual Grid Communities (VGCs) have been formed to bring together like-minded researchers and practitioners in specific domains (such as life sciences, physical sciences, digital media and manufacturing). Each VGC is a focal point to engage in discussions to identify applications of impact to the community. Regular activities are conducted for experts to share their grid experiences. These sessions also allow brainstorming to be carried out to identify more applications to run on the NGPP.

The invitation of applications remains very much evangelistic in flavor. There have been wins along the way as more converts are made.

IV. Managing the Grid-Enabling Projects

Hitherto our experience indicates that such an endeavor is not plain sailing. Despite the known benefits of Grid computing in the form of sharing, aggregation and virtualization of compute resources, users in general are more concerned with having access to more compute resources per se regardless of whether such provision is realized by Grid computing or other means. Although cognizant of the need to grid-enable their applications, the efforts required to do so is yet to be well understood by users. They simply wish their programs can run without any change to their source code. While this is indeed the ideal, incorporating changes has been the rule rather than the exception.

Understanding of users' needs is crucial for the grid-enabled applications to be of beneficial use to the users. In most cases, users just want to run their applications without the need to know about any details of grid computing.

Another key consideration is the management of users' expectation. Users who are new to Grid computing may have unrealistically high expectation. In the course of contact with them, it is important to set their expectations right. Such expectation may take the form of the amount of efforts required, the expected date of application readiness, the amount of compute resources available at the point of running their application, etc. If this is not done properly, it can lead to frustration and negative feedback that may potentially be detrimental to the overall effort to promote the adoption of Grid computing.

Besides catering to the needs of the users, the goals of the National Grid efforts are considered as well. In selecting an application for grid-enabling, due consideration will be given to the efforts needed, relative to the amount of manpower that is available. As resources are limited, priority will usually be given to applications that will enjoy significant benefits from grid-enabling and which, in turn, can become case studies.

V. NGCC Lifecycle for Grid-Enabling Projects

A project management lifecycle has been evolved over time to address the issues. The main objective of the life cycle is to manage the development of the applications. The principles behind the design of the process are to keep it simple and capture only essential documentation.

The GEAL process is depicted in the table below:

No	Activities	Parties	Documents
	Initiation & Requirements Phase		
1	Submission of User Application or mention of trial project at meetings	User	Application Form
2	Meet up with User to better understand project needs	User NGCC	
3	Evaluate the best grid-enabling	NGCC	

No	Activities	Parties	Documents
	approach for code/project		
4	Submit evaluation to NGO for review, together with project schedule	NGCC	Project Plan
5	NGO to review & agree to proceed	DD(NGO) NGCC	Project Plan (revised)
6	Meet up with User to provide feedback upon conclusion of evaluation to understand if that is a satisfactory approach	User NGCC	
7	User is agreeable with the approach	User	Project Plan (revised)
	Development Phase		
8	Handover of codes by User & proceed with Grid-enabling, team to keep NGOC updated	User NGCC	
9	Weekly updates on progress	NGCC	Meeting Notes
10	Regular update on progress with User (at least monthly)	User NGCC	Meeting Notes
11	Upon completion of Grid-enabling effort, report to be submitted to NGO for review, together with proposal on execution on NGPP	NGCC	Porting Report (with Test Results)
12	NGO to review & agree to proceed	DD(NGO) NGCC	
13	Meet User to explain work done & discuss about execution on NGPP	User NGCC	Meeting Notes
	Roll-out & Monitoring Phase		
14	Preparation for execution on NGPP & actual execution	NGCC	
15	Weekly/critical points updates on progress	NGCC	Meeting Notes Problem Tracking Report Usage Tracking Report
16	Upon successful execution, final meeting to gather User's feedback, with instructions on how to run the application	User NGCC	

Table 2: NGCC Grid-Enabling Applications Lifecycle

The GEAL is divided into 3 phases:

- Initiation & Requirements Phase
- Development Phase
- Roll-out & Monitoring Phase

A. Initiation & Requirements Phase

In the Initiation & Requirements Phase, the requirements for the project are elicited from the User. After the NGCC has studied the requirements, it proposes and evaluates various options for grid-enabling of the application. The ability of the grid-enabled application to

meet the User's expectation is also assessed. Issues to consider include whether the application is amenable to be partitioned, whether the framework or model that the application uses is suitable as well as the availability of the compute resources needed by the application.

It is important to ensure that the application can be partitioned and that grid-enabling will benefit the application. Because current efforts has been concentrated on harnessing the compute resources on the NGPP, such an application must also be compute intensive and its data transfer should incur so much overheads that erases the benefits that is derived the grid-enabling efforts.

The requirements for compute resources must be verified. On one hand, an application that needs 20 CPUs can run on the grid, but there are probably more effective and suitable options by, say, providing access it to run on a sufficiently large cluster rather than attempting to grid-enable it. On the other hand, an application that requires 512 CPUs and for which there is no single large enough cluster, may be a better candidate.

The operating system and platform requirements of the application are also assessed. An application that needs compute resources currently unavailable on the NGPP will clearly not be a suitable candidate for implementation.

The proposal and its recommendation are then submitted to the NGO for approval. After taking into consideration the other grid-enabling efforts of NGCC, the timeline may be revised. The approved approach is then shared with the User to ensure that expectations are met. Information such as scope, schedule and likely benefits are discussed. The User's involvement in the various phases is also indicated to ensure his/her commitment.

Sometimes the recommendation is not to proceed with the grid-enabling. There are invariably good reasons for arriving at such a conclusion. However, these reasons need to be communicated to the User in a tactful manner to avoid any misunderstanding or ill feelings.

An example is a class of MPI-based [1] application using MPICH [2]. Even when such applications are migrated to MPICH-G2 [3], the Globus version of MPICH, the applications can only run on different clusters whose compute nodes have public IP addresses. In almost all clusters connected to the NGPP, apart from the head node of each cluster, IP addresses of the other compute nodes are not public. As such, this class of applications cannot run across several clusters and hence the effort to grid-enable may not be worthwhile. Fortunately for us, this specific problem has been resolved using an implementation of the Realm-Specific IP (RSIP) [4] solution from HP and is currently under further testing on several clusters on NGPP.

Other issues include the use of commercial third-party software required by some user applications. If there are insufficient licenses to execute on the resources on the grid, which is typically the case, then the issue of the licenses must be resolved. Or there must be at least a resolution in sight before efforts are undertaken in grid-enabling the application.

B. Development Phase

Upon agreement from the User, the project moves into the development phase. To kick-start the process, the User hands over the application source codes to be grid-enabled. Weekly internal meetings are held to monitor the progress of the effort. Meetings with the User are

conducted at least monthly to provide an update on progress, problems, and obstacles as well as to manage their expectations.

Internal testing is carried out once the grid-enabling effort has been completed. A checkpoint meeting with the management of NGO is held to report the outcome of the grid-enabling efforts and the various test results before the NGCC meets the User to report of the overall effort.

The next step is for the NGOC to plan the execution of the grid-enabled application using compute resources available on the NGPP. After suitable compute resources are identified for the execution of the application, they will be installed with the codes and any other necessary software.

C. *Roll-out & Monitoring Phase*

In the rollout stage, the tested application is rolled out for execution by the User. The NGOC is involved in the monitoring of the execution of the applications. Should there be issues related to the application, NGOC will be report them to the NGCC, which remains the point of contact for the User, to be resolved.

Two main types of monitoring take place. The first is problem tracking. Problems reported will be worked on and the outcome reported to the User. The reported problems will be logged in the problem log. The second type of monitoring carried out is usage monitoring. Should there be low usage, the NGCC will check with the users to find out if there are any issues with the execution. The project will close upon the agreed timeframe of the execution of the application.

As with any methodology, only the relevant portions of the lifecycle will be applied in situations that do not warrant the complete lifecycle. Examples include applications that just need the compute resources on the NGPP and cases where grid-enabling is carried out by the users themselves.

The lifecycle is an evolving one. With the inclusion of industry in the National Grid (Phase 2), we foresee a need to include additional steps in the lifecycle that are important to industries. An example would be to include the testing of the application in a test-bed environment before the User runs it in-house.

VII. Conclusion

In conclusion, the grid-enabling experience by NGCC has been both a challenging and enriching one. It has been and will continue to be one of the key ingredients of the National Grid efforts. Our hope is that the past and valuable experience will enable the NGCC to serve its customers better.

References

[1] MPI: A Message-Passing Interface Standard. http://www.mpi-forum.org/
[2] MPICH: A Portable Implementation of MPI. http://www-unix.mcs.anl.gov/mpi/mpich/
[3] MPICH-G2. http://www.niu.edu/mpi/
[4] RSIP: Realm Specific IP. http://openresources.info.ucl.ac.be/rsip/

Choo Thong Tiong is Head (Promotion) at the National Grid Office. His main role is to promote the adoption of grid computing in Singapore by working with stakeholders, partners and the industry. He undertakes studies of grid initiatives and applications; and identifies potential and innovative opportunities for industrial grid applications. Thong Tiong also organizes seminars and talks to promote grid computing in Singapore. He holds a Bachelor Degree (Honours) in Computer Science from the National University of Singapore.

A National Computer Board scholar, Hing-Yan Lee graduated from the University of Illinois at Urbana-Champaign (USA) with PhD and MS degrees in Computer Science, specializing in artificial intelligence and software reuse. He previously studied at Imperial College (UK) where he obtained a BSc(Eng.) with 1st Class Honours in Computing and a MSc in Management Science.

Hing-Yan is currently Deputy Director of the Singapore National Grid Office where he directs, plans and coordinates the national initiative to realize a cyber-infrastructure for sharing and aggregating compute resources for R&D and industry. He is concurrently Project Director of the National Grid Pilot Platform, besides managing the National Grid Competency Centre and the National Grid Operations Centre. He spends considerably amount of his time promoting Grid Computing to potential users and meeting stakeholders.

Professor Lawrence Wong was appointed Executive Director of Institute for Infocomm Research on 15 November 2002. He is currently also the Director of National Grid Office. He was previously Director of Computer Centre at the National University of Singapore (NUS) and Professor in the Electrical & Computing Engineering Department. He served in various other positions in NUS since 1992 and was also a Member of Technical Staff in AT&T Bell Labs in New Jersey, US from 1980 to 1983.

Lawrence received his BSc (1st Class Honours) and PhD in Electronics & Electrical Engineering from Loughborough University, UK. He is a recipient of the IEE Marconi Premium Award (1989), IEEE Millenium Award (2000) and e-nnovator Awards 2000, Open Category.

Jon Lau is Assistant Head (Technical) at the National Grid Office as well as Technical Manager of the National Grid Pilot Platform. He coordinates the technical issues of the NGPP and virtual grid communities, which spans from network and security to middleware software. He developed the first Access Grid (AG) node in Singapore, and is promoting the deployment of more AG sites in Singapore, through the demonstration of its benefits of AG-enabling conferences and meetings. Jon holds a Bachelor Degree in Computing and a Master of Technology, both from the National University of Singapore.