GFD-I.155
ET-CG

Rüdiger Berlich, Forschungszentrum Karlsruhe
Kilian Schwarz, GSI
Kathryn Cassidy, Trinity College Dublin
October 12, 2009

Towards Professional Grid Certification

Status of This Document

This document provides information to the Grid community on policies for moving towards a professional Grid certification program. It does not define any standards or technical recommendations. Distribution is unlimited. This is a draft version of the document.

Copyright Notice

Important Disclaimer

OGF is *not* a certifying body nor is it endorsing any certification or training programs, this is merely an information document designed to stimulate discussion and act as a starting point which others can use to develop their certification processes. OGF can act as a forum for these stakeholders to meet and discuss the issues but it is not involved in the certification process itself.

Abstract

This is an OGF informational document outlining the need for, and some possible approaches to developing, a Professional Grid Certification Program. It is aimed at the stakeholders who might be interested in developing such a certification programme, including potential employers of certified Grid professionals as well as the certification and training industries and middleware and technology providers.

The document outlines a possible structure of a Certification Program along with some options for how to develop and sustain such an effort. Finally, some possible future directions are outlined in the Future work section.

Contents

1. Introduction and Goals

1.1 Purpose of this document

This is an OGF informational document outlining the need for, and some possible approaches to developing, a Professional Grid Certification Program. It is aimed at the stakeholders who might be interested in developing such a certification programme, including potential employers of certified Grid professionals as well as the training industry, certification and accreditation professional bodies and middleware and technology providers.

The document is designed to encourage further discussion and puts forward some possible approaches but these should not be seen as recommendations rather as a starting point for further discussion by the various stakeholders and training experts who might move forward with a certification program.

1.2 The need for Grid Certification

We are facing a crisis due to a shortage of skills and knowledge in high-tech industry[1], in particular where distributed systems are concerned. We believe that this skill shortage can be partially addressed by the development of a formal training and certification approach. Addressing the problem has the potential to create a "win-win-win" situation for the students, trainers and employers involved. Students gain employability. Employers gain skilled staff, and educators and trainers gain a market.

Strengthening and Unifying the Grid market

A Professional Grid Certification could be one approach which could have the following results leading to "positive feedback" in not only the training market, but the Grid field as a whole:

- a Grid certification accepted by major stakeholders could encourage training providers to develop their own curricula and training courses for the certification, in effect, creating a market for courses;
- the availability of more courses would make it easier, and also perhaps more attractive, for individuals to enter this new market and to broaden their skills;
- the greater availability of trained individuals would help the adoption of Grid computing by providing knowledgeable employees who can bring Grid technologies into their work-places;
- common standards would help employers to compare the skills of employees applying for a Grid job;
- certified professionals could rely on a generally accepted certification to prove their value to employers.

It is thus clear that a professional Grid certification could become a powerful tool, helping to unify and extend the Grid market.

Addressing the gap between Academia and Industry

Grid Computing is today characterised by a remarkable gap between initiatives and products with industrial and with scientific origin. Frequent attempts have been made to market techniques and organizational schemes developed in the context of academic research projects to industry. At the same time, academic research institutes, in particular in the clustering context, often rely on commercial products to drive their installations (the batch submission systems LSF and PBS are prime examples). In spite of these areas of commonality and efforts specifically aimed at strengthening the ties between industry and academia, the gap seems to be widening. There appear to be a number of interconnected reasons for this:

- While the fundamental technical requirements of Grid Computing are the same in industry and academia, their priorities in developing and deploying solutions are not. Where industry tends to focus on tough security measures and is hesitant to make the step from Intra-Grids to open Grid installations, the prime focus of academic research is often in the field of scalability and the orchestration of virtual organisations spanning the entire globe. It is thus no surprise that development communities working on entirely different usage scenarios rarely mix.
- While business models exist that could help market solutions originating in academia to enterprises on all levels, a lack of commercial-grade support and training for these systems hampers their adoption as a viable business platform.
- Similarly academia, with its abundance of Grid solutions, is not an easy target for doing business. While scientific communities work hard on the interoperability of their very diverse installations, the point where a single commercial solution can be deployed across a single scientific World Wide Grid is far away.

- As both industry and academia tend to use systems developed in the proximity of their core fields of expertise, and development communities, let alone management circles, rarely mix, there is a striking lack of information on both sides.

Education and Training can act as a catalyst for a new level of mutual recognition and cooperation. Likewise, the understanding of each others' needs can help to exploit synergies and adopt technical and management solutions already available in the other camp.

But while numerous academic initiatives have in the past strived to offer education and training to interested parties on all levels (among them ICEAGE and EGEE), uptake on the side of commercial training providers is low.

Possible reasons may include the diversity of the environment, which makes a "one size fits all" approach difficult and hence involves higher investments. On the trainee-side, up-take may be limited by a lack of standardized options.

This proposal wants to pave the way towards the introduction of a vendor-neutral, peer-reviewed professional Grid certification, based on the contributions from major players in the market.

2. Scope and Definitions

2.1 Scope

OGF is not attempting to become a certifying body or to endorse Grid Certification programmes. Rather this document is an informational document aimed at stimulating discussion on this topic.

Today's Grid environment is a patch-work of different technical and organisational solutions, ranging from local clusters to open grids involving thousands of users around the globe. Nevertheless it is to be expected that, over time, many "local" Grid types will become multi-institutional or at least span multiple sites of the same organization. Hence administrators and users of local installation will soon have to face the same complex trust issues and technical challenges inherent to Grid types of a geographically distributed nature.

Thus the assumption in this document is that a Grid certification would target Grids of the “wide-area” type rather than local Grid installations. It must be clear, however, that a certification spanning the entire industry must ultimately include components covering as many types of Grid tool and middleware as possible. The initial focus on wide-area Grids alone may also be revised, depending on the type of contributions available and the areas of expertize of contributors.

The authors would like to emphasize that this proposal addresses the question of professional Grid certification only. It does not try to make statements about the courses

given by commercial or academic training providers. Under this proposal, it is the sole responsibility of a training provider to make sure that courses enable their participants to take part in an examination. Examination and courses are seen as two distinct entities.

2.2 Context

The development of a professional Grid certification has to take into account many issues relating to the current environment, for example:

- the diversity of the Grid environment;
- market acceptance and trends;
- vendor-acceptance and acceptance by training providers;
- quality assurance procedures;
- identification, implementation and maintenance of generic parts of the certification;
- involvement of middleware-authors and vendors for implementation-specific parts;
- legal issues related to intellectual property rights;
- sustainability;
- funding issues;
- continuum of qualification with academic Grid teaching.

All of these issues, and perhaps many more would need to be addressed in order to develop a successful broad-based certification program, we have here attempted merely to identify what some of these issues are.

2.3 Definitions

The terms used in this document are, where possible, taken from the ET-CG Grid Education and Training Glossary, the latest version of which is available at https://forge.gridforum.org/sf/wiki/do/viewPage/projects.et-cg/wiki/Definitions

The term "Grid" is used throughout this document to refer to the infrastructures used in eScience and CyberInfrastructure.

3. Stakeholders

Contributions from a number of different entities are required for a certification of this type, and the participation should be on a global scale. Companies in the training and area, middleware providers, along with governmental bodies and national Grid initiatives, are the most likely candidates to contribute.

Possible participants include (in no particular order):

- Providers of Grid technologies;
- Training companies wishing to offer courses;

- Companies that want their in-house internal courses and exams to be certified;
- Companies offering tests for trainees (should be independent of the training providers!);
- Accreditation bodies responsible for accreditation of the organisations carrying out the actual certification procedure.

4. Possible Structure of Certifications

4.1 A modular Certification

There are many different Grid middleware and technologies so creating a single exam to certify users in Grids is quite a difficult prospect. One approach might be to base the exams on the model developed by the Linux Professional Institute (LPI)[2].

Aspects of Grid Computing (in the sense of wide-area Grids) today loosely resemble the situation of Linux in the late 1990s. Linux had already proven its usefulness. It enjoyed wide-spread support in academia, and commercial adoption was underway. First commercial training providers began to offer courses, based on their own curricula and test procedures. A large number of commercial- and non-commercial vendors began to offer what was then called a „Linux distribution“ - collections of Linux software, together with easy installation procedures and often bundled with support and documentation. Users of different distributions thus saw very different environments, although they were technically working on the same Operating system.

In 1999, the Linux Professional Institute was incorporated as a non-profit organisation. As quoted from http://www.lpi.org, „LPI shall promote and certify essential skills on Linux and Open Source technologies through the global delivery of comprehensive, top quality, vendor-independent exams.“ .

The initial training model was to identify generic components of all Linux distributions and to devise a certification based on these components. Training providers were then free to create course material for this certification.

Specific exams offered the user the ability to gain knowledge in subjects particular to a given Linux distribution. This model appears to be workable for Grid Computing as well.

It might be possible to create a similarly structured certification with core exams covering topics inherent to all Grid types, and plug-in exams developed by the different technologies. Possible core exams might include networking, security, Grid history or an overview of general Grid categories. Certification „plug-ins“ developed by each contributing vendor or middleware initiative, and peer reviewed by a neutral body, can build on this framework to test skills particular to a given Grid implementation.

It is also possible to evaluate existing certification procedures of various vendors and certify them if they fulfil the needs. This will also help with gaining acceptance of the procedure within industry.

The generic base exam would likely consist of relatively static components with a long lifecycle. “Quick moving” parts of the certification would be handled by the experts in this subject, i.e. the organization offering a particular technology.

4.2 Levels of qualification

Looking at the market requirements there are a number of different employee types each with different skill level requirements. A modular exam structure could allow learners to gain the skills relevant to their level. Ideally there would be some intermediate level qualifications which the learner could earn on their way to the full qualification and there should be a clear migration path from one to the next.

In discussions at OGF workshop sessions a three-tiered certificate structure was proposed, with each level having a different focus and covering different topics, or similar topics to different depths. Some ideas about how these different levels might be broken up and some of the types of knowledge which would be tested at each level are given below.

1. Certified Grid Technician („CGT“, works supervised)

- Base Exam / Technician
- 1 specialisation and/or overview exam
- More geared towards practical than conceptual skills
- Example requirements (horizontal vs. vertical approach)
    - EITHER: Detailed Know-how in one specific middleware
        - "find the logs"
        - "use the administration console to perform task x"
    - OR: Basic knowledge in a variety of different technologies
        - "look at queue status in middleware x,y,z"

2. Certified Grid Professional ("CGP", works unsupervised)

- Needs CGT exam
- Needs n>1 additional exams, but should include in any case installation/administration of one Grid type
- Conceptual and practical skills
- At least one extended project should be included in the education
- A CGP works within a functioning Grid environment

3. Certified Grid Architect ("CGA", has the high level view of Grid Architecture)

- Needs CGT and CGP exams
- CGA should be able to design new Grid environments that will use standard Grid components and can therefore start working already before a functional Grid exists
- A CGA is not necessarily included anymore in the day-to-day work of administering a Grid. Thus planning and management skills (such as project

management) should be part of the certification
- A CGA should be able to solve conceptual problems, like adding 500 new nodes to the Grid or do disaster recovery
- A practical project, possibly via work experience, involving design and planning is required in order to demonstrate application of the student's knowledge

It was suggested that, initially at least, much of the material necessary to pass the higher-level exams would taught within workshops rather than training events. In time training events might become more common as training organisations develop training tailored for the exams.

One important consideration which was stressed at repeated OGF workshops was the need for a migration path from CGT to CGP and CGA. This criterion was based on an understanding of how existing certification programs work. This would probably mean that it was necessary to take the CGT exam before you could progress to the CGP and so on. Similarly the CGA would have to build upon the topics covered in the CGP curriculum.

4.3 Accreditation

Companies wishing to offer Grid certification should be accredited, existing mechanisms should be investigated to identify a suitable model for accreditation of any Grid certification program.

5. Sustainability

The majority of stakeholders from the Grid area, while they may be very keen to see a professional certification program be developed for Grid, have neither the time nor the resources to be heavily involved in such an endeavour in the long term. While these Grid experts are needed to develop initial demand and enthusiasm for certification, as well as to help determine the criteria for worthwhile, reputable certification exams, etc. the long-term maintenance of any Grid certification program should not be expected to be handled by these Grid organizations and companies.

5.1 Existing certification bodies

Many certification bodies exist which handle a range of technical certification programs in the ICT domain. It may be possible to interest one of these bodies in taking on the Professional Grid Certification program if the Grid community can show that a market exists and can provide a possible curriculum and point to some existing training efforts which already cover part or all of that curriculum. To do this the Grid community should continue to develop and document these efforts.

Possible certification bodies to approach include

- the British Computer Society (BCS)
- the Association for Computing Machinery (ACM)
- the Institute of Electrical and Electronics Engineers (IEEE).

Unfortunately these bodies are often region-specific, for example, the BCS is mainly active in the UK and Europe while the ACM is active mainly in the US.

5.2 The Linux Professional Institute model: the Grid Professional Institute?

We have previously alluded to the Linux Professional Institute as a possible model for the structure of the Grid certification exams. It might also be possible to use the model of the Linux Professional Institute for the ongoing sustainability of a Grid certification. This would mean that a non-profit organization, possibly called „Grid Professional Institute", GPI, would assume responsibility for the entire process, in the same way that the non-profit LPI handles the Linux Professional Institute certification.

This would probably require commercial contributors willing to fund the work of the GPI, e.g. through membership contributions and it is not clear whether the organization could provide enough value to attract these contributions. Further investigation would be required to determine whether the model of the LPI could be successfully applied to the Grid world.

Forming such a GPI is not the role of an Open Standards Development organization such as OGF; therefore, a prerequisite of the formation of the GPI is for some other body or group of bodies to take the initiative in establishing the GPI.

6. Future Work

A detailed analysis of the current market should be conducted to show that there is a need for professional Grid certification. This could include surveying participating companies and academic institutes using the Grid. As part of this study, it should be ensured that there is indeed sufficient commonality between different Grid implementations, particularly amongst industrial providers, so that a joint certification process makes sense. The OGF Education and Training Community Group is already working on a Training Requirements Document[3] which captures the output of some such surveys.

It is necessary to convince as many major players as possible to participate in the field of developing a general education and training process leading to a professional grid certification. This certification needs to be globally visible and acceptable. The OGF and the Education and Training Community Group might prove a useful forum for ongoing developments, providing a contact and discussion forum, and working as a catalyst to trigger continuous work in the field. Other fora should also be used with the aim of getting inputs from as diverse a group of stakeholders as possible.

The general content of the curricula and exams must be further defined. There exist other

efforts in the area of curricula, in particular the ICEAGE[4] project has run workshops on Curriculum development and is developing a document in conjunction with the OGF Education and Training Community Group which focuses on curricula for eScience University Education. Other commercial middleware providers have their own existing training programs, as do many Research-based middleware and Grid projects. Out of these a general curriculum could be developed which might be used as the basis of a professional certification program. It is important that this work be undertaken by professionals from both industry and academia in a joint effort.

Once a curriculum has been developed, exams for different Grid implementations can be defined. Input should come from corresponding technology providers. Alternatively, in-house internal exams from various Grid providers could be certified as fulfilling the common standard if they cover the concepts required by the agreed curriculum.

Once the above work is underway the shape of the certification can be further defined in conjunction with certification bodies. For example issues of how a certified person stays current (must the exam be repeated in regular intervals or can a mechanism be devised via which a certificate holder can show that his/her skills are still up-to-date) and other practical issues would need to be addressed.

7. Contributors

- Adriano Rippa, Engineering Ingegneria Informatica SpA, Italy.
- Alex Voss, NCeSS, University of Manchester, UK.
- Amy Krause, University of Edinburgh, UK.
- Anand Patil, Dante.
- Anitha Orhi, Platform Computing Inc.
- Anurag Shankar, Indiana University, US.
- Ben Clifford, University of Chicago, US.
- Bernhard Schott, Platform Computing GmbH.
- Boon Low, National eScience Centre (NeSC), Edinburgh, UK.
- Borja Sotomayor, University of Chicago, US.
- Chris Higgins, EDINA, University of Edinburgh, UK.
- Christoph Erdmann Pfeiler, Forschungszentrum Karlsruhe, Germany.
- Daniel Templeton, Sun Microsystems.
- David Fergusson, National eScience Centre (NeSC), Edinburgh, UK.
- Donal Fellows, University of Manchester, UK.
- Doris Heathman, Forschungszentrum Karlsruhe, Germany.
- Elizabeth Vander Meer, National eScience Centre (NeSC), Edinburgh, UK.
- Erwin Laure, CERN.
- Feikje Hielkema, University of Aberdeen, UK.
- Gabriel Zaquine, CS SI, France.
- Geoffrey Fox, Indiana University, US.
- Gerd Behrmann, Nordic DataGrid Facility (NDGF).
- Gergely Sipos, MTA Sztaki, Hungary.
- Gian Luca Volpato, RRZN, Leibniz Universitaet Hannover, Germany.
- Gillian Law, University of Edinburgh, UK.

- Giuseppe Andronico, INFN Sezione di Catania, Italy.
- Huimin Lin, Academia Sinica, Taipei, Taiwan.
- Ian Frame, NIEeS, UK.
- Igor Tkachev, Joint Inistitute for Nuclear Research, Russia.
- Israel Hernandex, University of Edinburgh, UK.
- Jennifer Schopf, Argonne National Laboratory (ANL), US.
- Junseok Hwang, Seoul National University, Korea.
- Jysoo Lee, KISTI, Korea.
- Kathryn Cassidy, Trinity College Dublin, Ireland.
- Kilian Schwarz, GSI Gesellschaft für Schwerionenforschung mbH, Germany.
- Malcolm Atkinson, National eScience Centre (NeSC), Edinburgh, UK.
- Mathias Dalheimer, Fraunhofer ITWM, Germany.
- Matthew Dovey, JISC, UK.
- Merry Rabb, Sass ISP.
- Morgane Artacho, National eScience Centre (NeSC), Edinburgh, UK.
- Nishadi De Silva, University of Southampton, UK.
- Oscar Corcho, Universidad Politécnia de Madrid, Spain.
- Peter Brezany, University of Vienna, Austria.
- Puri Bangalore, University of Alabama, US.
- Rüdiger Berlich, Forschungszentrum Karlsruhe, Germany.
- Shahbaz Memon, Forschungszentrum Juelich, Germany.
- Steinar Henden, University of Tromsø, Norway.
- Stephanie McLean, Renaissance Computing Institute (RENCI), NC, US.
- Steve Brewere, OMII Europe.
- Suhen Hammoud, University of Brunel, West London, UK.
- Tatiana Strizh, Joint Inistitute for Nuclear Research, Russia.
- Thomas Prokosh, GUP, Joh. Kepler University Linz, Austria.
- Torsten Antoni, Forschungszentrum Karlsruhe, Germany.
- Vladimir Koren'kov, Joint Institute for Nuclear Research, Russia.
- Weilong Ueng, Academia Sinica, Taipei, Taiwan.
- Wolfgang Gentzsch, D-Grid, Germany.
- Yang Lui, University of Brunel, West London, UK.
- Yehia El Khatib, Lancaster University, UK.
- Zack Kertcher, University of Chicago, US.

8. IPR statement

The OGF takes no position regarding the validity or scope of any intellectual property or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; neither does it represent that it has made any effort to identify any such rights. Copies of claims of rights made available for publication and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the OGF Secretariat.

The OGF invites any interested party to bring to its attention any copyrights, patents or

patent applications, or other proprietary rights which may cover technology that may be required to practice this recommendation. Please address the information to the OGF Executive Director.

9. Disclaimer

This document and the information contained herein is provided on an "As Is" basis and the OGF disclaims all warranties, express or implied, including but not limited to any warranty that the use of the information herein will not infringe any rights or any implied warranties of merchantability or fitness for a particular purpose.

10. Full copyright notice

11. References

[1] Policy for Supporting Grid Education and Training https://forge.gridforum.org/sf/wiki/do/viewPage/projects.et-cg/wiki/NationalAndInternationalGridEducationTrainingPolicy

[2] Linux Professional Institute http://www.lpi.org/

[3] Training Requirements Document https://forge.gridforum.org/sf/wiki/do/viewPage/projects.et-cg/wiki/TrainingRequirementsDocument

[4] The ICEAGE Proejct http://www.iceage-eu.org/