Towards GRID Operating Systems: from GLinux to a GVM

Domenico Talia
DEIS, University of Calabria
talia@deis.unical.it

Abstract

This short position paper proposes a feasible approach for the development of operating system facilities to integrate machines in GRIDs. We propose a two-step strategy that in a first stage aims to integrate existing operating systems with a GRID through an interface based on a ***Grid Virtual Machine*** (***GVM***); in this phase Linux can be used as a use case for developing a Gridified operating system. The second stage should provide different versions of the ***GVM*** that can be configured on different devices implementing GRID nodes.

Background and interest

Today grid middleware is used to address the complexity of GRID environments and to help users in using GRID resources in an integrated way. This role in conventional computers is played by operating systems. Now it is time to develop a GRID operating system that may offer an integrated support for efficient management of local and remote resources available on a GRID environment to which a machine is connected. Without an operating system, Grids can fail the goal to enter mainstream computing and will not exploit all their functionality. As a conventional operating system provides an abstraction layer on top of the underlying physical resources of a computer, a GRID operating system must be designed to provide a virtual machine interface layered over the distributed, heterogeneous, autonomous, and dynamically available resources that compose a GRID. Resource sharing is the main objective of GRIDs and an operating systems is the more appropriate environment for providing GRID users access to resource sharing facilities in a secure and transparent way.

The research team of DEIS at the University of Calabria is actively involved in several projects related to GRID and is active in research activities in the Grid computing area. In particular, the related research areas in which we are active are: P2P computing, scalable information systems for GRIDs, self-organizing systems, distributed middleware, Grid metadata management, Grid-based problem solving environments. Many of these research areas are related to GRID operating systems and several solutions defined in those fields can be exploited in GRID operating systems for achieving scalability, global naming management, resource discovery and heterogeneity handling, etc. We have strong research partnerships with other research centres such as ICAR-CNR, ISTI-CNR and SPACI, and with spin-off companies such Exeura. All those institutions are working in national and European projects and are active in research work on GRID.

What do you expect from a GridOS?

A developer and an end-user would like to connect her/his machine to the GRID and access the GRID facilities as today a computer/pda/mobile phone connects to the Intranet/Internet and accesses to their services. A GRID OS should

- provide simple connection to the GRID,
- offer access to GRID resources, and
- define policies for providing local resource to a GRID.

In the development of internetworking facilities in traditional computers they were initially implemented as a middleware on top of the existing operating systems and successively were embodied in the new releases of operating systems. According to the same approach, facilities for integrating computers in GRIDs should move from the middleware layer to the operating system layer.

This approach could be feasible for the development of operating system facilities to support GRIDs through a two-step strategy:

1. The first stage should integrate existing operating systems within a GRID through an interface based on a *Grid Virtual Machine* (***GVM***). The ***GVM*** must be defined as a layered set of components that implement the services that applications needs to access and use GRID resources. A very interesting use case to implement this phase could be the integration of a GVM as a kernel Grid OS into Linux. That kernel Grid OS must be small, light, and efficient to be configured on different Linux machines The kernel should be based on a Grid Virtual Machine that integrates a Linux-based PC in the GRID (***GLinux*** as a Gridified Linux). A Grid-aware file system (***GFS***) based on a global name space should be implemented as a component of GLinux. Other modules can be put on top of the kernel and can use the services of the GFS.

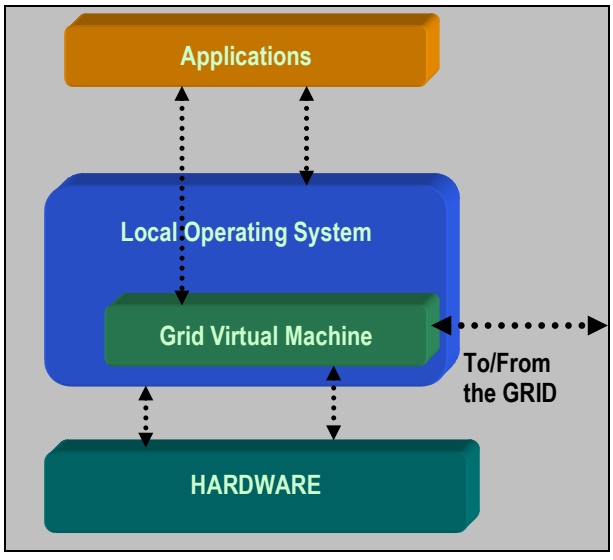

Figure 1. A GVM embedded in a local operating system.

2. The second stage should provide an operating system for a pervasive GRID composed of different grained computing nodes. The modular design of the ***GVM*** will support its total/partial implementation of different GRID nodes from cluster computers, to PDA and sensor devices resulting in different version of the ***GVM***. Those different versions of the ***GVM*** can be configured/adapted on different devices representing GRID nodes. Different instantiations of the ***GVM*** result in different GRID OS kernels depending on the hardware available. A minimal kernel should be configurable on every GRID node; it will provide different implementations of a GRID file system (***GFS***), of a *global naming system*, and defines a *GRID horizon* composed of all the GRID environments accessible from a machine are represented and accessed through publication and discovery mechanisms.

What kind of expertise can you supply?

We are interested to study the basic mechanisms for the development of a Grid Virtual Machine. Moreover, we can supply our research background and knowledge on P2P techniques, models and services for scalable GRIDs, metadata management techniques and models, architectures for GRID-based PSE toolkits, scalable resource discovery and GRID information systems. The reference section lists some papers published by the group members and related the mentioned expertise areas.

References

1. A. Iamnitchi and D. Talia, "P2P Computing and Interaction with Grids", *Future Generation Computer Systems*, North-Holland, vol. 21, no.3, pp. 331-332, 2005.
2. M. Cannataro and D. Talia, "Semantics and Knowledge Grids: Building the Next-Generation Grid", *IEEE Intelligent Systems*, vol. 19, no 1. pp. 56-63, 2004.
3. C. Mastroianni, D. Talia D. , P. Trunfio, "Metadata for Managing Grid Resources in Data Mining Applications", *Journal of Grid Computing*, Kluwer vo. 2, no. 1, pp. 85-102, 20048.
4. M. Cannataro, A. Congiusta, A. Pugliese, D. Talia, P. Trunfio, "Distributed Data Mining on Grids: Services, Tools, and Applications", *IEEE Transactions on Systems, Man, Cybernetics*, Part B, vol. 34, n. 6, pp. 1-15, 2004.
5. D. Talia, P. Trunfio, "Toward a Synergy Between P2P and Grids", *IEEE Internet Computing*, vol. 7, no. 4, pp. 94-96, 2003.
6. M. Cannataro, D. Talia, "The Knowledge Grid", *Communications of the ACM*, vol. 46, no. 1, pp. 89-93, 2003.
7. C. Lee, D. Talia, "Grid Programming Models: Current Tools, Issues and Directions", in *Grid Computing: Making the Global Infrastructure a Reality*, F. Berman, G. Fox and T. Hey (eds.), Wiley, chapt. 21, pp. 555-578, 2003.
8. D. Talia , P. Trunfio, "A P2P Grid Services-Based Protocol: Design and Evaluation", Proc. *EuroPar 2004*, Pisa, Springer-Verlag, LNCS vol. 3149, pp. 1022-1031, 2004,
9. M. Cannataro, C. Comito, A. Congiusta, G. Folino, C. Mastroianni, A. Pugliese, G. Spezzano, D. Talia, P. Veltri, "A General Architecture for Grid-based PSE Toolkits", *Proc. International Workshop on State-of-the-Art in Scientific Computing (PARA'04)*, Copenhagen, LNCS, Springer-Verlag, 2004.

10. C. Mastroianni, D. Talia, O. Verta, "A Super-Peer Model for Building Resource Discovery Services in Grids: Design and Simulation Analysis", Proc. European Grid Conference – EGC 2005, LNCS, 2005. (Best paper award).