J Grid Computing (2010) 8:61–75
DOI 10.1007/s10723-009-9143-2

A Grid-enabled CPU Scavenging Architecture and a Case Study of its Use in the Greek School Network

**Fotis Georgatos · Vasileios Gkamas ·
Aristeidis Ilias · Giannis Kouretis ·
Emmanouel Varvarigos**

Received: 24 February 2008 / Accepted: 8 December 2009 / Published online: 7 January 2010

Abstract In this paper we present a CPU scavenging architecture suitable for desktop resources, and we study its appropriateness in exploiting the PC Laboratory resources of the Greek School Network and their integration to the existing HellasGrid national infrastructure. School laboratories form an extensive network equipped with computational systems and fast Internet connections. As this infrastructure is utilized at most 8 h per day and 5 days per week, it could be made available during its remaining idle time for computational purposes through the use of Grid technology. The structure and organization of the school laboratories and backbone network enables the CPU scavenging service, as an independent and additional service, which will not violate the operational rules and policies of the school network, while it will add additional resources to the current HellasGrid infrastructure with low adaptation cost.

Keywords CPU scavenging · Grid computing · School network

1 Introduction

Grid computing has emerged as an important new field, distinguished from conventional distributed computing by its focus on integrated large-scale resource sharing, innovative applications, and in some cases, high-performance orientation. Grids introduce new ways to share computing and storage resources across geographically separated sites, by establishing global resource management architecture [10].

Recent Grid-related technologies like Condor and gLite, combined can do more than effectively

---

F. Georgatos
CSCS, ETH Zürich, Manno, Switzerland
e-mail: fotis@mail.cern.ch

V. Gkamas · E. Varvarigos
Research Academic Computer Technology Institute,
Rio, Patras, Greece

V. Gkamas (✉) · A. Ilias · E. Varvarigos
Computer Engineering and Informatics Department,
University of Patras, Rio, Patras, Greece
e-mail: gkamas@ceid.upatras.gr

A. Ilias
e-mail: aristeid@ceid.upatras.gr

E. Varvarigos
e-mail: manos@ceid.upatras.gr

F. Georgatos · G. Kouretis
National Technical University of Athens,
Athens, Greece

G. Kouretis
e-mail: ge01275@mail.ntua.gr

manage dedicated computer clusters [11]. Condor can scavenge and manage wasted CPU power from otherwise idle desktop workstations across an entire organization with minimal effort. For example, Condor can be configured to run jobs on desktop workstations only when the keyboard and CPU are idle. If a distributed Grid job is running on a workstation, when the user returns and hits a key, Condor is even able to migrate the job to a different workstation and resume its execution from the point it was interrupted. Many institutes have already proceeded in such a deployment, defining and tuning a configuration policy for multi-thousand node resources [39].

Extensive national Grid infrastructures have been installed and developed in many countries (for example, in the context of the Enabling Grids for E-sciencE—EGEE—European project [1]), including the HellasGrid infrastructure [2] developed in Greece. The main objective of such efforts has been the development of Grid infrastructures that include heterogeneous distributed computational and storage resources in order to provide Grid services to the research and academic community. Such infrastructures are typically under the control of a National Research and Educational Network (NREN), Greek Research and Technology Network (GRNET) [3] in the case of Greece. Similarly to other European NRENs, GRNET offers high speed Internet connections and advanced next generation network services to Greek academic and research institutes. The applications that have been or will be developed at HellasGrid, exploit the advanced network infrastructure of the GRNET-3 network, which is based on Dense Wavelength Division Multiplexing technology. The value of the HellasGrid infrastructure will be accented from pilot and innovative applications [6, 7], being developed from other related efforts.

The research and academic community is increasing its use of Grid technologies. Grid computing presents various advantages: overall improvement of the computational efficiency, cost reduction, usage of underutilized resources, activation of research collaborations in the context of virtual organizations, increase of the storage capacity, capability to perform high performance parallel computations, access to more resources, and finally, drastic reduction of the time needed to derive the desired results, which is in many cases the main objective.

In this context, Greece, a country with moderately developed infrastructures, must give added value at the existent computational infrastructures, by considering CPU scavenging models, like those of Hungary [18, 23, 24, 28, 29], *XtremWeb* [31], SETI@home [15] or other commercial Desktop Grids solutions for enterprises [27, 32, 33] which permit the usage of under-utilized personal computers, (mainly in nightly hours) from communities with increased needs and requirements. CPU scavenging or cycle-scavenging is a special case of distributed computing that harnesses unused PC resources worldwide, generally for research purposes, at nights, weekends, and other idle times. This technique is also known as shared computing cycle stealing. More details on related efforts for CPU scavenging solutions are given in the Related Work section.

The Greek School Network (GSN) [4] is an ideal case for this purpose. GSN is the educational intranet of the Ministry of National Education and Affairs [5] that interlinks all school laboratories and provides basic and advanced telematic services to the educational community. GSN has available continuously updated computational infrastructures, mostly comprised of personal computers, with rather low utilization percentages. Personal computers at the school laboratories are used mainly during day time, normally with an upper limit of 8 h per day, 5 days per week, so the percentage of the underutilized computational power is rather high.

In order to exploit the underutilized resources of the school laboratories, we propose an architecture for their integration to the existing HellasGrid infrastructure in the form of Desktop Grids [17]. We address the issue not only at the *technical* level, but also at the *structural* and *administrative* level. Any proposed solution will have to take into account the school network's special characteristics and should not violate its policies of operation. Our objective is not to substitute the computational resources dedicated to HellasGrid, but to complement them by creating Desktop

Grids, that will operate in addition to the main infrastructure by adding significant computational resources with minimal cost.

After evaluation of the possible alternative solutions, the proposed CPU scavenging technology is LiveWN [21], a fully functional desktop Grid environment, compliant with LCG/EGEE Grids [8, 9] which provides an easy and versatile way to use underutilized computational resources without the need of any special operation, system installation or middleware configuration. The LiveWN technology, which has been developed originally by some of the authors for the needs of a High Energy Physics team within the National Technical University of Athens (NTUA), gives the ability to setup a Worker Node (WN) and a User Interface (UI) in a quick and simple way, under diverse environments (Virtual Machines, dialup lines, firewalls, etc) without the need of installation or special configuration. As we will argue later on, this technology is suitable and meets the special characteristics of school laboratories, as well as the needs of the special fields for which we own applications: High Energy Physics (low-data volume simulations), Digital Media (Rendering and complex image processing) as well as Data Mining (very wide workflows, jobs' data in memory).

The rest of this work is organized as follows. CPU scavenging initiatives and projects similar or related to LiveWN are presented in Section 2. Section 3 describes the HellasGrid infrastructure, while the Greek school laboratories and network infrastructure is described in Section 4. The proposed technical architecture for school labs CPU scavenging is represented in Section 5. Section 6 addresses structural and administrative issues while an estimation of the added computational power is performed in Section 7. Finally, Section 8 concludes the study.

2 Related Work

Many related CPU scavenging/Desktop Grid projects exist. This section discusses various Desktop Grid solutions that are in use today. These projects have either similarities or partially common functionality with our proposed approach, but we believe our solution compares favorably to most of these approaches in terms of total capabilities or is complementary to them.

The E-Grid live cd for instance [12] works well in many cases where fixed resources are provided for the Grid, but does not provide for software updates, distributed user filespace, or dynamic allocation of Worker Nodes (WNs). It also requires fully qualified hostname/domain names (FQDNs) and static IP addresses preconfigured right in the end-user's environment, which is hard in practice to achieve.

The most well-known Desktop Grid solution is the SETI@home [15], in which approximately 4 million PCs have been involved. Condor [13] and BOINC [14] are other well-tested and viable solutions for desktop resource scavenging, but they are technically not by themselves Grid solutions, since they support only single administrative domain functionality. This is a view with which the author of the BOINC system and project SETI@home [15], David Anderson, appears to agree and doubts on calling them Grid systems, although there is a great deal of technical merit in considering them as complementary solutions [16]. Grid indeed is not a panacea.

SZTAKI Desktop Grid, [23, 24, 28, 29] is a BOINC project initiated in Hungary, run by the Computer and Automation Research Institute (SZTAKI) of the Hungarian Academy of Sciences. It has an impressive proven scalability of more than 30,000 nodes. Work units are assigned by a Master process to Workers, through a Distributed Computing API. This model of execution is sufficient for SETI@home style of applications, but just like BOINC or Condor, it is not a fully implemented Grid model, since it does not allow direct communication between work units. Also, ClusterGrid [18] is another project from researchers in Hungary, which aims at integrating x86 processor based PCs into a single, large, countrywide interconnected set of clusters. The PCs are provided by participating Hungarian institutes, such as high schools, universities, or public libraries, and central infrastructure and coordination is provided by NIIF/HUNGARNET.

Similarly to BOINC, XtremWeb [31] is a research project, which aims to serve as a substrate for global computing experiments. Basically, it supports the centralized set-up of servers and PCs as Worker Nodes. In addition, it can also be used to build a peer-to-peer system with centralized control, where any worker node can become a client that submits jobs; a functionality existing also at the LiveWN solution.

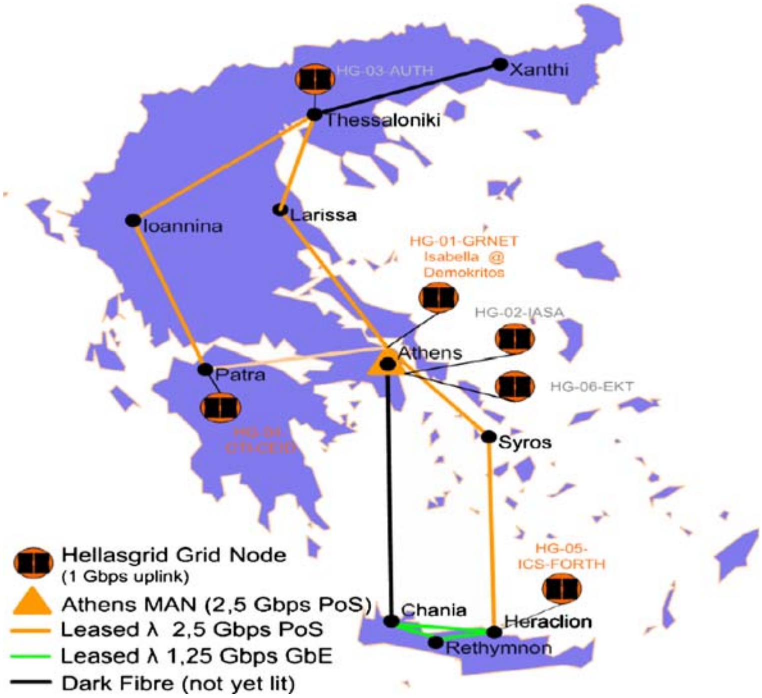
EDGeS [25] is another project with the aim of creating an integrated Grid infrastructure that seamlessly integrates a variety of Desktop Grids with EGEE type of service Grids. In [30] the authors focus on bridging from desktop Grids towards service Grids, that is, making desktop Grids able to utilize free service Grid resources, and they present and compare three approaches that solve interoperability between desktop and service Grids. Also, the authors in [26] present an interface between the Globus Toolkit and the BOINC, in order to expand the reach of Grid Computing. This solution is interesting but there is no special provision for maintaining the end to end IP capabilities, that is, nodes that are in private address space are not reachable from outside.

There are also a number of companies providing Desktop Grid solutions for enterprises [27, 32, 33]. The most well-known examples are the Entropia Inc, and the United Devices. Those systems support the desktops, clusters and database servers available at an enterprise. Entropia can completely seclude the execution of the desktop Grid applications from other processes running on the PC thus, ensuring that Grid applications cannot access data on the client machines.

3 HellasGrid Infrastructure

The HellasGrid infrastructure (Fig. 1) consists of six computer clusters with a total of more than 700 64-bit CPUs, and on-line storage (disks and tapes) of more than 100 TBs, interconnected over an end-to-end Gigabit backbone network. The HellasGrid infrastructure is fully integrated with the pan-European Grid infrastructure, EGEE, currently offering more than 75,000 CPUs and 300 PetaBytes of storage. This infrastructure is available free of charge for the research and

Fig. 1 HellasGrid infrastructure



academic community, for their day-to-day application needs, big or small research projects, etc.

The HellasGrid infrastructure was developed in two phases. Phase 1 involved the deployment of HG-01-GRNET cluster node composed of 32 Dual CPU 1U servers, a 10TB Fiber Channel (FCAL) Storage Area Network (SAN) and a 10 TB (expanded to 30 TB) Tape Library, located at NCSR Demokritos, Athens. Phase 2 involved the expansion of the Grid infrastructure with larger clusters, in the following institutes:

- HG-02-IASA: At the institute of Accelerating Systems and Applications (IASA) of the University of Athens: 64 Dual CPU cluster, 4 TB SAN Storage.
- HG-03-AUTH: At the Aristotle University of Thessaloniki (AUTH): 64 Dual CPU cluster, 4 TB SAN Storage.
- HG-04-CTI-CEID: At the Department of Computer Engineering and Informatics (CEID) of the University of Patras, in collaboration with the Research Academic Computer Technology Institute (RACTI) in Patras: 64 Dual CPU cluster, 4 TB SAN Storage.
- HG-05-FORTH: At the Institute of Computer Science—Foundation for Research and Technology Hellas (ICS-FORTH): 64 Dual CPU Cluster, 4 TB SAN Storage.
- HG-06-EKT: At the National Documentation Center (NDC-EKT): 134 Dual CPU cluster, 12 TB SAN Storage and 50 TB of tape library.

All the above nodes use GRNET as the core network for their interconnection.

4 Greek School Network and School Laboratories Infrastructure

GSN is hierarchically structured into three layers (Fig. 2):

- Backbone Network: The Greek School Network interconnects with GRNET at eight main points (Athens, Thessaloniki, Patras, Heraclion, Larisa, Ioannina, Xanthi and Syros), using it as its backbone network.
- Distribution Network: GSN has installed network and computational equipment at the capital of every prefecture, thus ensuring optimal access of the prefecture's schools to the network and its services. The distribution network consists of 51 nodes, one at every prefecture and is divided in two levels: (1) first level, where the eight main nodes of the backbone network are, and where the interconnection with the GRNET network is done and (2) second level, which is comprised by 42 nodes installed at the remaining prefectures, which are the secondary nodes of the distribution network.
- Access network: It is used to directly and efficiently interconnect the schools to the prefecture's node. The interconnection technologies used to interconnect each school are selected on the basis of financial and technical criteria from an array of available options: Digital ISDN circuit (bandwidth: 64–128 kbps), Analog leased line (128 kbps–2 Mbps), Public Switched Telephone Network dialup circuit (56 kbps), Wireless link (10 Mbps), VDSL circuit (10–15 Mbps) and ADSL circuit (2000/256 kbps).

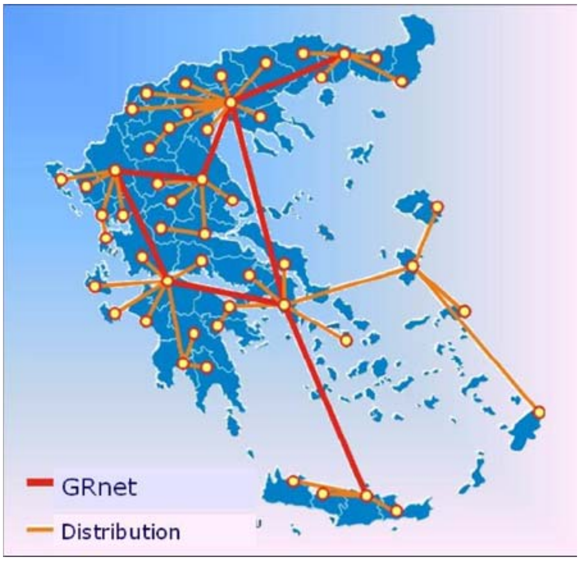


Fig. 2 GSN's infrastructure, Tier 1 & 2

Note that both the HellasGrid nodes and the school labs use GRNET as the core network for their interconnection, a feature which is architecturally exploited in our scenario.

Each school laboratory consists of several personal computers connected at a local network in accordance to the client–server model. The servers have Windows Server (NT/2000/2003) operating system installed, while the personal computers use Windows (2000/XP) operating system. The PCs can be configured to operate at dual-boot mode, where the second operating system is Linux, in order to support additional services. Linux also works in collaboration with Windows Server for the users' authentication and offers services like file sharing, e-mail, etc. The Internet connection of the school labs is performed with various and complementary network access technologies (analog leased lines, ISDN, dial-up, wireless, ADSL, VDSL).

According to information from the Ministry of Education, 13,863 schools, 2,631 administrative units, 494 school, 71 public and 29 municipal libraries, and 60 State General Archive are connected to GSN. Table 1 shows the available computational infrastructure of the GSN. This table describes the computational infrastructure of the public schools only (it does not contain private schools) and only of those public schools for which we have available information. According to Table 1, the number of personal computers that could be offered as computational resources to HellasGrid currently exceeds 60,000, a rather large number, in relation to local needs in Greece.

Table 2 describes the CPU version of the 62,807 PCs located at the various public schools connected to the GSN. The available CPU types are: AMD, Celeron, Pentium II (233 MHz to 450 MHz), Pentium III (450 MHz to 1.4 GHz), Pentium IV (1.3 GHz to 3.8 GHz) and other.

The design of the GSN infrastructure depends heavily on Network Address Translation (NAT) technology in order to conserve the small IP address space assigned to GSN. Specifically, every school laboratory is assigned a 4 IP address subnet. One IP address is assigned to the router and another one to the school laboratory's server (by using static NAT). School LANs use private addressing of the form 10.X.Y.Z/24 or 10.X.Y.Z/28, where the numbers X and Y is a unique combination for every laboratory. The laboratory's LAN router always has the private IP address 10.X.Y.1 and the server the IP address 10.X.Y.10. The workstations are assigned an IP address from the DHCP service running at the laboratory's server. The workstations get Internet access via a proxy server (for WEB, FTP, etc), or by using the router as the gateway of the LAN and NAT technology. The only thing that someone can know is the real address that has been assigned to the server by using static NAT. Any system outside the school laboratory's LAN cannot communicate directly with a school PC. This was deliberate, but now it is a problem for Grid technologies, whereby some applications like Hadoop or MapReduce, assume direct communication using IP addresses, and therefore it has to be considered by the proposed solution. For instance, an application that uses the traditional ftp protocol might not work within private address space and behind a firewall. This has been shown to be a problem in other deployments [39].

Table 1 GSN's computational infrastructure

	Schools	Laboratories	PCs
Elementary school	6,107	5,088	20,453
Junior high school	1,896	1,712	19,413
High school	1,138	1,044	15,587
Professional high school	485	392	7,354
Total	9,626	8,236	62,807

Table 2 Schools laboratories personal computers by CPU generation

	Elementary school	Junior high school	High school	Professional high school
AMD (all)	232	159	51	104
Celeron	1,457	1,551	1,073	970
Pentium II	1,543	2,054	2,671	1,154
Pentium III	3,958	2,858	8,050	1,648
Pentium IV	12,908	12,690	3,668	3,427
Other type	299	98	74	51
Total	20,397	19,410	15,587	7,354

5 Proposed Architecture

Before describing the proposed CPU scavenging solution we elaborate on the type of jobs that would be suitable to be served by the school labs' personal computers.

GSN, in its current form, encourage the transmission of large amounts of data to the school labs, due to the existence of broadband Internet connections at many schools. Also, the school labs' computational resources can be appropriate for CPU intensive applications. This type of Grid Services, allows for relatively low amounts of data transfers but provides important computational power, which can become readily available in a big scale, since the number of personal computers that remain inactive for long durations (nights, weekends, holidays) is substantial.

Also, by the end of 2008, a significant improvement in GSN has been performed: the percentage of schools with broadband Internet connections (ADSL 2000/256 kbps, leased lines 0.5–12 Mbps, Wireless, Satellite) has been increase to 92% [19], as a result of various actions that has taken place for the broadband upgrade of the GSN network. As the GSN's broadband upgrade allows fast transmission of larger amounts of data, the infrastructure of the school laboratories could also be used for Data Grid [10] purposes as well.

The integration of school labs' computational resources will be even more efficient in the case of distributed workflow applications, with resilience requirements (the jobs must be segmented into shorter jobs with smaller data sets). One way to ensure resilience is by executing parts of a job on multiple CPUs in parallel and in the case of an unsuccessful execution, the respective job can be executed at another available computational resource. This technique is better known in the Grid Community as fault tolerance by job replication. This strategy is justified by the following school labs characteristics:

- The standard computational infrastructure of school labs consists of workstations with limited computational power of a single processor and is not comprised by supercomputers or clusters.
- The personal computers of school laboratories cannot be fully dedicated to Grid infrastructure, but are available instead at specific time periods only. Furthermore, 100% availability of the scheduled resources cannot be guaranteed (for example a personal computer may be employed by a task left by a local user).
- The clustered homogeneity of the majority of the equipment and the operating systems running at the machines of the school laboratories, guarantee uniform duration length for the execution of jobs at the various personal computers.

The primary application that is tested to run on the LiveWN infrastructure belongs to the category of High Energy Physics simulations, namely Garfield, Geant4 and some more components of the ATLAS collaboration toolkit, which all require small input/output data. Although the described effort is a test deployment, it can serve the needs of other Virtual Organizations; in particular, the needs of very fast memory-based data mining applications, planned to occur from the South Eastern Europe and Digital Media Virtual Organizations.

5.1 A CPU Scavenging Solution Based on LiveWN

The LiveWN [21] solution is a mixture of three technologies:

- Linux LiveCD technology: The LiveCD technology provides for ease of use and hardware independence. Thus, a user does not have to install or configure his system, while at the same time he has a fully relocatable environment.
- LCG/EGEE middleware: The use of gLite middleware provides full Grid interoperability, since it is a standard middleware stack in wide use by the LCG, EGEE and many other compatible Grids and Grid projects.
- LiveWN Server: The LiveWN server incorporates all the extra services needed by a site in order to properly deploy the LiveWN technology in an easy manner.

The services provided by the LiveWN server are:

- OpenVPN server: A tunneling technique provides for a generic VPN solution under unpredictable network environments (for instance, behind firewalls) and solves the public IPv4 requirement, which is imposed by the middleware, by providing same IP address per Worker Node identity.

- OpenAFS server: The AFS is a highly available and scalable solution for filespace that allows the users to store and retrieve files, even when they work in a totally diskless environment; this proves handy under stateless environments, such as CD/DVD-ROMS, NetBoot, etc.
- Rsync server: The rsync server instance provides the development team with the ability to apply patches and updates during reconfiguration, minimizing the need for frequent software distributions on physical media or other network- or labor-intensive methods.
- PlanetLab Node: This service is optional but it improves, when placed at strategic core locations, the availability of the LiveWN Configuration Management system through the properties of the Coral Content Distribution Network [20]. Ideally, it should be placed near backbone network routers, for optimal reliability.

Upon boot, LiveWN is by default configured to ask an IP address by a DHCP server, and then it configures some initial network access parameters. Once network connectivity is established, the user starts the LiveWN service. The user is authenticated with a login/password in order to be assigned with a unique Worker Node identity, which is configured as part of a Computing Element (CE); it is entirely possible to auto-subscribe so-called anonymous resources, as well; the latter has been considered non-adequate in the case of GSN. Once correct credentials have been supplied, an OpenVPN tunnel is created and the system configures its hostname and domain name to be the proper ones, and then turns to a resource pool in accordance to an associated Computing Element. Indeed, since we have preconfigured correct forward and reverse DNS, ssh keys and other minor configuration bits, it appears as just another Worker Node, so it joins the CE's queues and starts accepting and executing jobs.

The LiveWN solution is engineered to work with most common deployment formats so that it can be used in different scenarios. Different supported usage cases include:

- DVD: The DVD version provides UI/WN functionality and a large collection of user tools for physics, mathematics, video and graphics. The user can use OpenAFS as network filespace for his work or any extra software, placed under the /afs virtual directory tree.
- CD: The CD version also provides UI/WN functionality but the minimum system requirements are significantly reduced, as well as the image footprint; and some of the software is available only through OpenAFS.
- USB key: The USB version is based either on the CD or DVD version depending on the requirements and it can also provide for a permanent local data area for the user.
- ISO file; suitable for both Virtual Machines and netboot.

LiveWN is an adaptable technology, because of the VPN technique used. It works behind firewalls; at systems within private address space networks; even under many unexpected platforms and environments, which we keep discovering every day. This characteristic of LiveWN solves efficiently the problem of the usage of private IP addresses at the workstation of every school's laboratory.

5.2 Testing the Proposed CPU Scavenging Solution

The proposed solution has been tested with emphasis placed on three important criteria: hardware compatibility, scalability and heterogeneity tolerance. It successfully confirmed applicability under all following tested scenarios:

- A bare x86 PC system booting with a LiveWN optical disk, CD or DVD placed in its optical drive.
- A 64bit Linux system running a Virtual Machine with qemu (VM/32bit), and then booting either by CD, USB or local image within the Virtual Machine.
- MacOSX v10.x with either VirtualBox, VMWare or Parallels with both DVD and local .iso images—on Intel platforms.
- Windows XP system, running VMWare's vmplayer or qemu and inside it a local image,

again with success and a recorded efficiency greater than 90%; as that efficiency would be measured for CPU-bound tasks.

Also, we proceeded with the pilot deployment of a virtual cluster consisting of 52 clients (26 dual-core Worker Nodes) serving the ATLAS Virtual Organization. All the machines have been successfully connected to the installed LiveWN server, named livece.gridlab.ntua.gr, and have been successfully executing various jobs.

Furthermore, a LiveWN server of the scalability test machines used (see Fig. 3) has been tested and verified to be sustainable under heavy load as:

- it has been shown to accept 360 jobs per hour (this is also a gLite Workload Management System limitation)
- it has been observed to achieve a job success rate larger than 99.5%
- it has served more than 50 clients (25 dual-core Worker Nodes)

What makes these results very satisfactory is the linear scalability in the gLite stack, so five such LiveWN servers can accept 1,800 jobs/hour and support 250 clients. Under heavy stress load, the outcomes would be even more favorable for jobs that carry small payloads of CPU-intensive activity. Also, the LiveWN's compatibility with the Short Lived Credential Service [45], as WN & UI has been verified.

Fig. 3 Equipment used for scalability tests of LiveWN, located at NTUA, Athens, Greece

LiveWN has also been tested for network diversity and under different scenarios to the highest possible degree, and it has even achieved a functional Grid site, which was spanning CPUs across three continents. Since 2006, LiveWN has passed all tests done, without counter-evidence of any trouble. We expect the performance of the LiveWN solution to be found up to par with what the Virtual Machines themselves can offer, so it should be a good choice for a wide range of cases. It is worthy to note that LiveWN and Virtualization are orthogonal technologies, so any progress in the VM arena, is automatically also progress that can be incorporated in this environment.

The future plan for the LiveWN project is to develop and incorporate solutions that improve and add functionality to Grid infrastructures in order to provide a fully customizable and versatile desktop Grid framework that can be adapted to any usage scenario. The major areas of interest for future work are [22]:

- Configuration Management System (livewnd)
- Authentication (OpenID, ShibGrid)
- Accounting (e.g., DGAS, WMS/RB-PA)
- Resource Characterization (LM-bench, GridBench, etc)
- Check pointing (e.g. based on Condor, or app-specific)
- IPv6 support and Network Resource Load Balancing

On top of these directions, which are LiveWN-specific, extra work has to be done, so that resources of significant size will be manageable with minimal effort, either for the automated installation of clients under Windows OSs [34] or for the uniform application of resource scheduling policies across multi-cluster farms of machines.

5.3 Deploying LiveWN at a Large Scale

The deployment of the LiveWN technology at a scale as large as the Greek School Network environment requires solutions to be given to problems unseen before. The major issues that have to be considered were: (1) resource identification, (2) end-to-end routing and (3) overall topology optimization.

Their treatment plan follows hereby:

- A new identification model has to be introduced, if we are to maximize autonomic aspects, possibly based on MAC addresses; these are unique 48-bits keys that PCs' network cards use in a LAN. Regarding this approach, there are security issues to be considered, along with the administrative benefits it has. The reason we decided to take this approach has to do with the fact that in this way we can ramp-up quickly the resources of GSN, without having to rely on complex and scarce resources like Trusted Platform Module (TPM) that only exist in a subset of the computing resources. Even if we used TPM and other hardware-based tokens for resource identity specification, the Desktop Grid nature would not allow us to consider these presented identities as 100% accurate; the technical explanation can be found in [40]. Despite these problems, there are published and established techniques on how to cope with these issues that appear on volunteer or scavenging-based systems [41].
- The number of resources (>60,000) is vast and if we do not intend to break a fully-featured Grid's end-to-end communication model, we should try to employ a whole IPv4/16 address space (formerly known as a Class B subnet). This is a major technical and administrative undertaking that is being coordinated together with the European Regional Internet Registry, RIPE NCC. We note that this is a constraint imposed not only by applications, but by the gLite middleware stack itself, because it does not currently support IPv6 address space. Experience coming from the EUChinaGRID project [35] suggests that this will need more time to be accomplished, so if we are to fully support Grid capabilities at the Worker Node edge we should deploy public IPv4 address space.
- A new Grid deployment architecture has to be followed that has to be aligned with the existing network infrastructure of GSN, taking into account the previous two issues, in addition to the network topology, as can be seen in Fig. 2. Next paragraph addresses that.

As we have already described, the Greek School Network is comprised of three layers:

- Backbone Network; implied by the GRNET2/3 backbone network with eight nodes at Athens, Thessaloniki, Patras, Heraclion, Larisa, Ioannina, Xanthi and Syros.
- Distribution Network; these are 51 locations at the respective prefecture capital cities, some of them overlapping with the backbone network's primary nodes.
- Access Network; these are the actual client machines, accounted for the purposes of this design to be less than 63,000; withstanding availability factors.

Let us assign the respective Tier1, Tier2, Tier3 abbreviations to Backbone, Distribution and Access networks, respectively. According to the current design of LCG [8], EGEE [9] and LiveWN [21] capabilities, ideally, Tier1 nodes should include PlanetLab node [38] plus Workload Management System (WMS) [37] and Berkeley Database Information Index (BDII) [36] services, Tier2 nodes should include one or more PoPs (where a PoP is comprised by a Computing Element and a LiveWN server) and Tier3 nodes are the actual Worker Nodes at various schools' LANs, running LiveWN clients.

There is also still an open question about where Storage Elements (SEs) are going to be placed. This is a delicate technical decision that may have to be revised later on. The current proposal is to place them at Tier1s, mainly for maintenance reasons, although such a choice does have certain implications regarding Desktop Grid dependability issues. Still, it appears that placing SEs and MONs at Tier1 nodes is a balanced choice, regarding the particular kind of jobs that a Desktop Grid is expected to run. We have experimented with this setup and indeed all current evidence converges that this choice is appropriate.

Another issue of concern is the amount of PoPs a particular Tier2 node must have. We have assumed a design limit of 256 WNs per PoP, which looks tractable; moreover, the numbers agree with existing international experience regarding the capabilities of the gLite middleware stack (256 Computing Elements of 256 Worker Nodes each, is currently feasible). What is under investigation, is how many physical nodes are necessary for providing the required services at the PoPs, but with current hardware it is asserted that this number is eight in the worst case/load, as proven by recent scalability and service performance tests done. On the other hand, a single physical server has been proven to be adequate for running a whole Tier1**,** which implies all the needed LiveWN services, including the VPN, afs, rsync & CE.

It is worth noting that the proposed architecture has certain features of particular interest, not least of them being that LiveWN is very simple to deploy and grow: To begin with, it is not necessary to expand the network to its greatest scale from the first step. We could easily follow a ramp-up deployment and only place PoP nodes as necessary. The expectation is that, if all eligible schools join as planned, we would end-up with at least 1–2 PoPs per Tier2 and at least two PoPs per Tier1 node. Note that having two PoPs will provide increased reliability, which is an important property in a system where components are bound to fail at a significant rate due to multiple reasons, partly related to their huge population. Generic Grid Dependability is a matter that is much more complex than this design could cover, but for the initial needs of a Desktop Grid service for the Greek School Network the current preliminary proposal is deemed sufficient.

Finally, regarding security issues in such a hybrid system, we believe that the issue that has to be addressed in this case is that the scavenged resources are not physically isolated from external entities that could tamper in many ways the intended function of the Greek School Network LiveWN Desktop Grid system, and even more so, its security aspects. This implies that only a certain level of trust guarantees can be provided for these resources that we do not own, and the application developer or Grid user has to be aware of the system's limitations—as well as any other Desktop Grid built on such resources. The proposed approach can make use of existing work in this area, which solves the problem at the application level [43]. At a later stage, it might become interesting as well as to employ appropriate handles for the establishment of trust relationships at the service provision level [44]. Also, we observe that for the use case of the Greek School Network (1) the environment is extremely homogeneous at the same level, from a security point of view, (2) the cost–benefit analysis of involving technologies like Trusted Platform Module indicates that this would not provide benefits and (3) the exact applications that are intended to be executed at this point do not require high level of trust and can be simply verified by applying a subsequent run of the workflow.

6 Structural and Administrative Issues

It is obvious that for the successful integration of underutilized CPU resources to existing Grid infrastructures, the collaboration between all involved parties is required, including:

- The resource providers: The resource provider can be any institute, university, school or other type of organization that is willing to offer underutilized computing resources. In our case, the Greek School Network is the resource provider.
- The National Grid Initiative (NGI) coordinator: The NGI coordinator is the organization leading the development of the Grid infrastructure at national level. For example, in the EGEE projects at many countries (including Greece) the NGI role was taken by the corresponding NREN. In the case of Greece, the coordinator of the HellasGrid initiative is GRNET.
- A CPU scavenging technology provider; in our case the LiveWN technology provider.

This involves administrators of the organizations offering underutilized computing resources, administrators of the existing Grid sites, where the underutilized computing resources will be integrated in the form of Workers Nodes, the NGI coordinator and a research institute to provide technological solutions necessary for the integration. In the case of Greece, from the initial steps of the GSN, there is a successful collaboration between GSN and GRNET, as GSN uses GRNET's network for its backbone network. What are needed to be defined, is the hierarchy and the operational structure of all involved parties, in order to create the expanded Grid infrastructure.

The operational structure and hierarchy will be comprised of the following three levels (Fig. 4):

- Level 1—NGI coordinator (T1 carrier): At the first level belongs the NGI coordinator responsible for the supervision of the expanded Grid infrastructure. In the case of Greece, the NGI coordinator will be GRNET, which already has under its supervision the existing HellasGrid infrastructure. The NGI coordinator will implement the operational rules and policies of the expanded National Grid infrastructure, the standards to be used, the service level agreements and will also check the correct operation of the infrastructure as a whole. The NGI coordinator will be in direct collaboration with the administrators of the existing Grid sites.
- Level 2—Organizations and institutes hosting existing Grid sites (T2 carrier). At the second level of the proposed hierarchy lie the universities and institutes that host the various Grid sites. The administrators of these sites will have the responsibility of the operation and support of the basic services (Computing Elements, Storage Elements, Monitoring Elements). Also, they will be in direct collaboration with the administrators of the computational resources providers, which will offer computational power to the existing Grid sites in the form of Worker Nodes.
- Level 3—Organizations/School laboratories offering computational resources (T3 carrier). At the third level lie the organizations/school laboratories offering computational resources. The entities of this level will have the responsibility for the incessant operation of their PCs and the handling of users' problems.

According to [10], the Grid architecture is comprised of four levels: diverse resources, secure access to resources and services, directory brokering, diagnosis and monitoring, tools and applications. The responsibility of every involved carrier at each level is shown in Fig. 5. T3 type carriers are involved at the first and second level of the Grid architecture, which contains the various diverse resources (computers, storage media, networks and sensors) and the secure access to resources and services. So T3 carries will be responsible for the incessant operation and support of the school laboratories' computational resources according to the defined service level agreements. Furthermore, they will be responsible for the support of the Internet connections of the laboratories and the secure access to the provided computational resources. T2 type carriers are involved to all levels of Grid architecture. So, except from the responsibilities of a standard T3 carrier, their re-

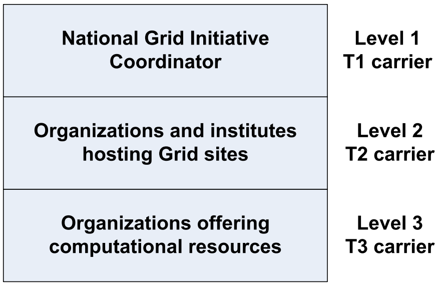


Fig. 4 Proposed operation structure and hierarchy

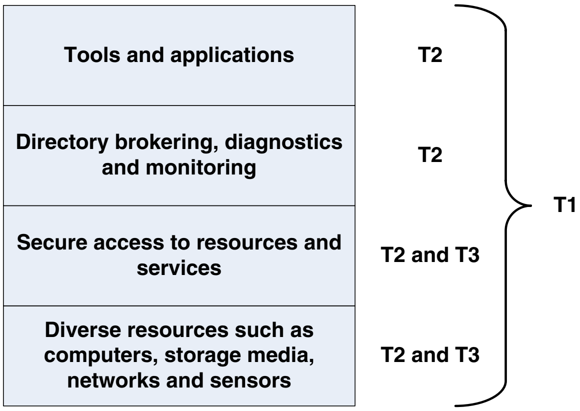


Fig. 5 Responsibility of the involved carriers

sponsibility will include directory brokering, diagnostic and monitoring tools, and applications. Of course, the T1 carrier will supervise and coordinate the actions at all levels.

7 Estimation of the Added Computational Power

At this Section we perform an estimation of the computational power that can be added to the HellasGrid infrastructure from the CPU scavenging of the PCs located at the various school laboratories. The number of available PCs per type of CPU is shown in Table 3.

For demanding user applications, there is a need of computational resources of Pentium III technology, at least. This condition is met by 49,207 PCs spread at various school laboratories. Assuming an appropriate directive/policy is adopted for the operation of each school laboratory; PCs would operate 24 h a day, every day with a satisfactory percentage of them being available in most circumstances.

Malfunctions at the network connections and/or the hardware/software of the PCs are also expected, but it has to be mentioned that the specifications of PCs lying at most school laboratories satisfy the ability of incessant PC operation. Nevertheless, in our computations we must take into account the probability of unavailability of the PCs due to connection problems, hardware/software updates and failures, etc. According to data from the GSN users' helpdesk service, the daily availability of school laboratories is 99% in relation to the access network. For the operational availability (hardware and software) of the school laboratories there is no available data. Considering a pessimistic operational availability of the school laboratories of 50%, (in fact, this number is quite likely under representing the availability of resources [42]) we can obtain a pessimistic estimation of the number of PCs available to the HellasGrid infrastructure:

$$(49,207 \times 99\%) \times 50\% = 24,357 \rightarrow \sim 24,350\,\text{PCs}.$$

It is obvious that these systems will not be available during the hours which the laboratories are used by the students. Excluding holidays, during which the laboratories will be available 24 h daily, we can assume that during workdays they will be available about 16 h per day (two thirds of a full duty cycle). So (24,350 × 2)/3 = 16,233 PCs of technology Pentium III or better, are expected to be available to complement and off-load the existing HellasGrid infrastructure.

Table 3 Number of school laboratories PCs per CPU type

CPU type	Number of PCs
AMD	546
Celeron	5,051
Pentium II	7,422
Pentium III	16,514
Pentium IV	32,693
Other type	522
Total	62,748

8 Summary and Conclusions

We presented a model for the exploitation of the school laboratories' personal computers and their integration to the existing HellasGrid infrastructure, in the form of Desktop Grids. We propose a new CPU scavenging architecture that runs behind firewalls, on systems within private address space networks and, above all, under many unexpected platforms and environments. This solution is suitable for the case of school labs as it is consistent with their special characteristics. We did not address the issue at the technical level only, but at structural and administrative level as well. Finally, we provided estimates on the number of personal computers that can be effectively integrated to the existing HellasGrid infrastructure with a low adaptation cost. Providing an infrastructure of this scale could help significantly the research community. Equally important, new research areas would be opened, as Desktop Grids provide a platform where Computer Science knowledge of multiple sub-disciplines can be put in real-world use and validated.

References

1. Enabling Grids for E-sciencE project. http://www.eu-egee.org
2. HellasGrid Task Force. http://www.hellasgrid.gr
3. Greek Research and Technology Network. http://www.grnet.gr
4. Greek School Network. http://www.sch.gr
5. Ministry of National Education and Religious Affairs. http://www.ypepth.gr
6. Maglogiannis, I., Soldatos, J., Chatzioannou, A., Kanaris, Y., Milonakis, V.: The EKTORAS platform for bioinformatics data analysis in Grid environments. In: 1st HellasGrid User Forum (2008)
7. Kollias, G., Gallopoulos, E., Georgiou, K.: Design and development of Grid ranking tools for web analysis. In: 1st HellasGrid User Forum (2008)
8. Knobloch, J., et al.: Large hadron collider computing Grid technical design report. http://lcg.web.cern.ch/LCG/tdr/
9. Gagliardi, F., Jones, B., Grey, F., Bégin, M., Heikkurinen, M.: Building an infrastructure for scientific Grid computing: status & goals of the EGEE project. Phil. Trans. R. Soc. A. **363**(1833), 1729–1742 (2005)
10. Foster, I., Kesselman, C.: The Grid: Blueprint for a New Computing Infrastructure, 2nd edn. Morgan Kaufmann, San Francisco (2003)
11. Tannenbaum, T., Wright, D., Miller, K., Livny, M.: Condor—a distributed job scheduler. In: Sterling, T. (ed.) Beowulf Cluster Computing with Linux. MIT Press, Cambridge (2002)
12. EGRID Project. http://www.egrid.it
13. Litzkow, M.J., Livny, M., Mutka, M.W.: Condor—a hunter of idle workstations. In: 8th International Conference on Distributed Computing Systems, pp. 104–111 (1988)
14. Anderson, D.: BOINC: a system for public-resource computing and storage. In: 5th IEEE/ACM International Workshop on Grid Computing, pp. 4–10 (2004)
15. Anderson, D., Cobb, J., Korpela, E., Lebofsky, M., Werthimer, D.: SETI@ home: an experiment in public-resource computing. Commun. ACM **45**(11), 56–61 (2002)
16. Anderson, D.P.: iSGTW Opinion—Volunteer computing: Grid or not Grid? http://www.isgtw.org/?pid=1000527 (2007)
17. https://gilda.ct.infn.it/, Kondo, D., Taufer, M., Brooks, C.L., Casanova, H. Chien, A.A.: Characterizing and evaluating desktop Grids: an empirical study. In: 18th International Parallel and Distributed Processing Symposium, pp. 26–30 (2004)
18. Máray, T., Stefán, S., Szalai1, F., Vitéz, G.: The Hungarian clusterGrid project: challenges of a production Grid. DSPN (2004)
19. Paraskeuas, M.: Strategies and actions for broadband at Greek school network. In: 2nd International Conference on Broadband Internet (2007)
20. Freedman, M., Freudenthal, E., Mazi, D.: Democratizing content publication with coral. In: 1st Symposium on Networked Systems Design and Implementation, pp. 18–32 (2004)
21. Kouretis, G., Georgatos, F., Alexopoulos, T., Tsipolitis, Y., Live, W.N.: Scavenging in the Grid era. In: 3rd EELA Conference (2007)
22. Georgatos, F.: LiveWN and gLiteDVD: true Grid scavenging solutions. iSGTW article http://www.isgtw.org/?pid=1000655 (2007)
23. Kacsuk, P., Marosi, A., Kovacs, J., Balaton, Z., Gombas, G., Vida, G., Kornafeld, A.: SZTAKI desktop Grid—a hierarchical desktop Grid system. In: Cracow Grid Workshop. Krakow (2006)
24. Balaton, Z., Gombas, G., Kacsuk, P., Kornafeld, A., Kovacs, J., Marosi, A.C., Vida, G., Podhorszki, N., Kiss, T.: SZTAKI Desktop Grid: a modular and scalable way of building large computing Grids. In: IEEE International Parallel and Distributed Processing Symposium (2007)
25. Montes, M.C., Emmen, A., Marosi, A., Araujo, F., Gombas, G., Terstyanszky, G., Fedak, G., Kelley, I., Taylor, I., Lodygensky, O., Kacsuk, P., Lovas, R., Kiss, T., Balaton, Z., Farkas, Z.: EDGeS: bridging desktop and service Grids. In: 2nd Iberian Grid Infrastructure Conference, pp. 212–226 (2008)
26. Myers, D.S., Bazinet, A.L., Cummings, M.P.: Expanding the reach of Grid computing: combining Globus- and BOINC-based systems, Grids for bioinformatics and computational biology. In: Wiley Book Series on Parallel and Distributed Computing. Wiley, New York (2008)
27. Chien, A., Calder, B., Elbert, S., Bhatia, K.: Entropia: architecture and performance of an enterprise desktop Grid system. J. Parallel Distrib. Comput. **63**, 597–610 (2003)
28. Kacsuk, P., Podhorszki, N.: Scalable desktop Grid system. In: High Performance Computing for Computational Science—VECPAR 2006, pp. 27–38 (2006)
29. Marosi, A., Gombas, G., Balaton, Z., Kacsuk, P., Kiss, T.: SZTAKI desktop Grid: building a scalable, secure platform for desktop Grid computing. In: CoreGRID Workshop on Programming Models Grid and P2P System Architecture Grid Systems, Tools and Environments, pp. 365–376 (2007)
30. Farkas, S., Kacsuk, P., Solar, M.R.D.: Utilizing the EGEE infrastructure for desktop Grids. In: 7th International Conference on Distributed and Parallel Systems, pp. 25–27 (2008)
31. Fedak, G., Germain, C., Neri, V., Cappello, F.: XtremWeb: a generic global computing system. In: 1st IEEE/ACM International Symposium on Cluster Computing and the Grid, pp. 582–587 (2001)
32. P. Computing. Platform LSF. http://www.platform.com
33. D. Inc. DeskGrid. http://www.deskgrid.com
34. Karalis, Y., Kalochristianakis, M., Kokkinos, P., Varvarigos, E., OpenRSM: an open source lightweight integrated remote network and systems management solution. Int. J. Netw. Manage. (2008)
35. EUChinaGRID Project. http://www.euchinagrid.org/
36. Berkeley Database Information Index. https://twiki.cern.ch/twiki//bin/view/EGEE/BDII

37. Workload Management System. http://glite.web.cern.ch/glite/wms/
38. PlanetLab. http://www.planet-lab.org/
39. Jermini, P., Thiémard, M.: EPFL, Condor at the Swiss Institute of Technology of Lausanne, Condor Week 2009. University of Wisconsin, Madison (2009)
40. Martin, A.: Trusted computing for trusted infrastructures. In: 4th International Conference on Grid and Pervasive Computing, Geneva (2009)
41. Domingues, P., Sousa, B., Silva, L.M.: Sabotage tolerance and trust management in desktop Grid computing. Future Gener. Comput. Syst. **23**(7), 904–912 (2007)
42. Kondo, D., Fedak, G., Cappello, F., Chien, A.A., Casanova, H.: Characterizing resource availability in enterprise desktop Grids. Future Gener. Comput. Syst. **23**(7), 888–903 (2007)
43. Sarmenta, L.F.G.: Sabotage-tolerance mechanisms for volunteer computing systems. In: 1st IEEE/ACM International Symposium on Cluster Computing and the Grid, pp. 337–346 (2001)
44. International Science Grid this Week. Opinion—a matter of trust: enabling Grid security through bilateral negotiation. http://www.isgtw.org/?pid=1001540
45. Short Lived Credential Service. http://www.switch.ch/grid/slcs/about/