ETSI TR 102 767 V1.1.1 (2009-02)

Technical Report

GRID;
Grid Services and Telecom Networks;
Architectural Options

Reference
DTR/GRID-0005

Keywords
architecture, network, service

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Contents

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (http://webapp.etsi.org/IPR/home.asp).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Technical Report (TR) has been produced by ETSI Technical Committee GRID (GRID).

1 Scope

The present document provides an introduction to Grid, Grid Applications, the NGN architecture and provides an analysis of architectural option for combining Grid and NGN.

2 References

References are either specific (identified by date of publication and/or edition number or version number) or non-specific.

- For a specific reference, subsequent revisions do not apply.
- Non-specific reference may be made only to a complete document or a part thereof and only in the following cases:
  - if it is accepted that it will be possible to use all future changes of the referenced document for the purposes of the referring document;
  - for informative references.

Referenced documents which are not found to be publicly available in the expected location might be found at http://docbox.etsi.org/Reference.

NOTE: While any hyperlinks included in this clause were valid at the time of publication ETSI cannot guarantee their long term validity.

2.1 Normative references

The following referenced documents are indispensable for the application of the present document. For dated references, only the edition cited applies. For non-specific references, the latest edition of the referenced document (including any amendments) applies.

Not applicable.

2.2 Informative references

The following referenced documents are not essential to the use of the present document but they assist the user with regard to a particular subject area**.** For non-specific references, the latest version of the referenced document (including any amendments) applies.

[i.1] NESSI: Service Oriented Infrastructure Working Group.

NOTE: Available at http://www.soi-nwg.org/doku.php?id=sra:description.

[i.2] ITU-T Recommendation Y.2001: "General overview of NGN".

[i.3] ITU-T NGN Working definition (2004).

NOTE: Available at http://www.itu.int/ITU-T/studygroups/com13/ngn2004/working_definition.html.

[i.4] TISPAN Terms of Reference (ToRs).

NOTE: Available at http://portal.etsi.org/tispan/TISPAN_ToR.asp.

[i.5] ETSI ES 282 001 (V2.0.0): "Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); NGN Functional Architecture".

[i.6] ITU-T Recommendation Y.2011: "General principles and general reference model for Next Generation Networks".

[i.7] ETSI TR 180 002: "Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); Release 2 definition".

[i.8] ETSI TR 102 659-1 (V1.1.1): "GRID; Study of ICT Grid interoperability gaps; Part 1: Inventory of ICT Stakeholders".

[i.9] ETSI TS 123 002: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Network architecture (3GPP TS 23.002)".

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the following terms and definitions apply:

Grid or Grid computing: Grid is a system that is concerned with the integration, virtualization, and management of services and resources in a distributed, heterogeneous environment that supports collections of users and resources (virtual organizations) across traditional administrative and organizational domains (real organizations) [i.8]

Grid service: service interface associated with a Grid resource

NOTE 1: A resource, logical or physical, and its state (statefulness is the defining characteristic of a Grid service) is controlled and managed via Grid services in a Grid environment.

NOTE 2: This definition is based on IBM Redbook "Introduction to Grid Computing" (http://www.redbooks.ibm.com/redbooks/pdfs/sg246778.pdf)

3.2 Abbreviations

For the purposes of the present document, the following abbreviations apply:

AS	Application Server
ASF	Application Server Function
ETSI	European Telecommunications Standards Institute
HPC	High-Performance Computing
IMS	IP Multimedia Subsystem
IP	Internet Protocol
ITU-T	International Telecommunication Union - Telecom Sector
NASS	Network Attachment SubSystem
NGN	Next Generation Network
PES	PSTN/ISDN Emulation Subsystem
RACS	Resource and Admission Control Subsystem
TISPAN	Telecommunications and Internet converged Services and Protocols for Advanced Networking

4 Introduction to Grid and Grid applications

Grid computing originated in the scientific and high-performance computing (HPC) communities. With availability of additional computing resource leading directly to improved accuracy and quality of results, the demand for specialized computing systems became difficult to satisfy. Techniques for sharing such resources between research groups at different locations and supporting new forms of collaboration were developed. Their adoption and the consequent wider availability of high performance computing led to the emergence of e-Science as a productive approach to research in a number of disciplines such as particle physics, bioinformatics and earth sciences. This suggested the idea of "the Grid" essentially unlimited computing resources being available on demand anywhere - supported by wide area networks, named by analogy with electrical power grids.

Enterprises are increasingly reliant on IT applications to support their business processes and there is a clear trend towards obtaining functionality as services from external providers to provide flexibility. Given the success of HPC Grids in similar scenarios, it is natural to consider how widely applicable the technology could be. Scientific Grid applications generally focus on particular application areas and aim to provide shared access to specialized high performance computing resources or datasets. Typically they involve parallel execution of stateless batch jobs which can be easily distributed. Enterprise IT applications are much more varied. They frequently involve stateful interactions between a potentially complex set of components including databases and application servers. Usage patterns may be either batch processing or interactive. While it is not possible to define a representative business Grid application, it is likely that they will be characterized by intensive use of IT resources, whether this is computational, access to large datasets or demanding constraints on data volumes, request rates or latency, transactional throughput or some combination of these. Simple, undemanding applications are unlikely to justify the additional complexity of sourcing functionality from an independent service provider.

Considerable effort has been devoted in recent years to extending Grid technologies to address business requirements. The NESSI-GRID Strategic Research Agenda [i.1] defines "Business Grids as the adaptive service-oriented utility infrastructure for business applications" and envisions them as "the general ICT backbone in future economies". This gives Grid a far greater significance for mainstream networking than traditional e-Science applications and indicates that it has to be taken seriously. Grid technology can be seen as the basis for a range of approaches to service oriented infrastructures including utility computing, real-time infrastructures and cloud computing.

Networking is a fundamental component of Grid application infrastructures. Scientific Grids have typically used specialized research networks rather than public networks and relatively little attention has been given to operating over public networks. It is very difficult to be specific about the network requirements of business Grid applications as they are very varied. In addition, business applications require end-to-end behaviour (performance, availability, security policy enforcement etc.) to be controlled and predictable. These issues clearly need to be addressed for business Grids. In particular, it is clear that there is a requirement for applications to have a greater awareness of the networks which are involved in their deployment so that they can adapt to available connectivity resources. Similarly, if networks have a greater awareness of the requirements of the applications they support, improved operational efficiency should be achievable.

5 Overview of relevant Telecom reference architecture

5.1 Introduction

A Next Generation Network (NGN) is a packet-based network able to provide services including Telecommunication Services and able to make use of multiple broadband, QoS-enabled transport technologies and in which service-related functions are independent from underlying transport-related technologies. It offers unrestricted access by users to different service providers. It supports generalized mobility which will allow consistent and ubiquitous provision of services to users [i.2].

In a NGN there is a more defined separation between the transport (connectivity) portion of the network and the services that run on top of that transport. This means that whenever a provider wants to enable a new service, they can do so by defining it directly at the service layer without considering the transport layer - i.e. services are independent of transport details [i.3].

TISPAN is the ETSI core competence centre for fixed networks and for migration from switched circuit networks to packet-based networks with an architecture that can serve in both. Therefore TISPAN is responsible for all aspects of standardization for present and future converged networks including the NGN (Next Generation Network) and including, service aspects, architectural aspects, protocol aspects, QoS studies, security related studies, mobility aspects within fixed networks, using existing and emerging technologies [i.4].

5.2 The NGN architecture

The NGN functional architecture as described in [i.5] complies with the ITU-T general reference model for next generation networks [i.6] and is structured according to a service layer and an IP-based transport layer.

The service layer comprises the following components:

- the core IP Multimedia Subsystem (IMS);
- the PSTN/ISDN Emulation Subsystem (PES);
- other multimedia subsystems (e.g. IPTV Dedicated Subsystem) and applications;
- common components (i.e. used by several subsystems) such as those required for accessing applications, charging functions, user profile management, security management, routing data bases (e.g. ENUM), etc.

This subsystem-oriented architecture enables the addition of new subsystems over the time to cover new demands and service classes. It also provides the ability to import (and adapt) subsystems defined by other standardization bodies.

IP-connectivity is provided to NGN user equipment by the transport layer, under the control of the network attachment subsystem (NASS) and the resource and admission control subsystem (RACS). These subsystems hide the transport technology used in access and core networks below the IP layer.

The architecture and related subsystems specifications is a functional architecture. Each subsystem is specified as a set of functional entities and related interfaces. As a result implementers may choose to combine functional entities where this makes sense in the context of the business models, services and capabilities being supported. Where functional entities are combined the interface between them is internal, is hidden and un-testable.

Figure 1 provides an overview of the NGN architecture. An example of realization of this functional architecture.

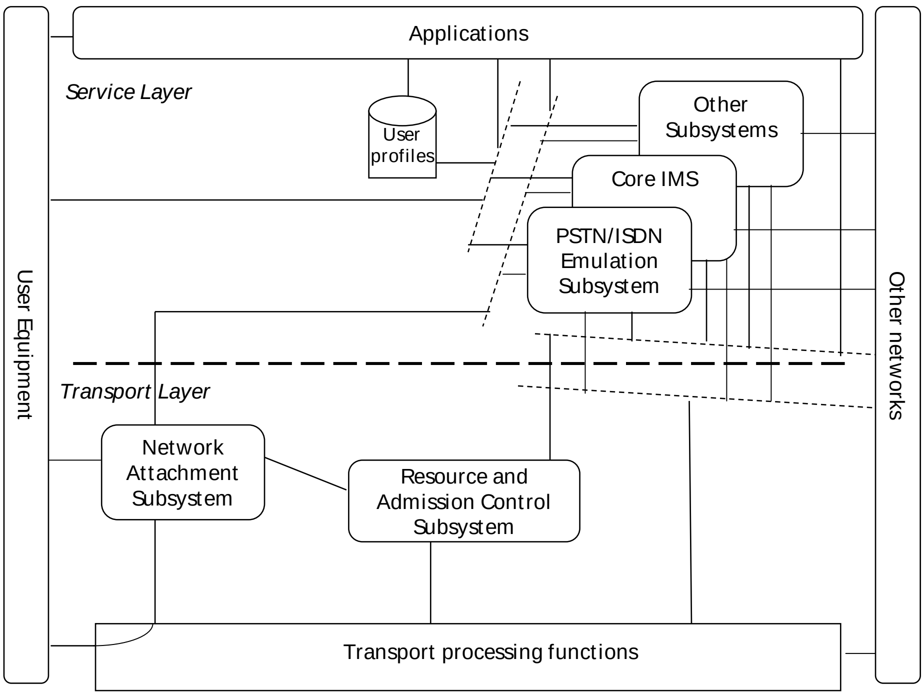

Figure 1: TISPAN NGN overall architecture

6 Grid and Next Generation Network (NGN)

6.1 Introduction

The NGN is an initiative from the Telecoms Industry and provides interoperable, inter-domain all IP-based network solution with enhanced multimedia capabilities. Independently from the access technology, it supports mobility and nomadicity and services including:

- Person-to-Person services: voice mail, voice call, multimedia sessions, etc.
- Messaging services: e-mail, SMS, IM and presence services.
- Content-on-demand services: IPTV, browsing, download, streaming, push, and broadcast services.

The NGN functional Architecture [i.5] identifies 2 NGN layers. A Service Layer and an IP Based Transport Layer (as described in figure 1). The NGN Release 2 definition is contained in [i.7].

As a starting point 4 possible architectural options have been selected to be evaluated. The present document will analyze the following 4 listed options in the following clauses:

- Grid-enabled NGN application.
- NGN subsystems offering Grid Services.
- Grid technology for implementing NGN functionality.
- Combining Grid and networking resources in a new architecture.

The options above result in different impacts on NGN. The implications of these options are for further study.

6.2 Grid-enabled NGN application

In terms of their network requirements, Grid applications are very diverse. In general Grid applications make significant use of computational or storage/data resources in addition to connectivity and require coordination between multiple activities.

Grid applications may be either session-based (e.g. interactive computation steering, low-latency computations, real-time scene rendering in an online game) or not (e.g. complex workflow execution, batch processing, data movement and staging). However, they are all characterized by the need for network connectivity and so it should be possible to regard them as NGN applications. They may have specialized network requirements such as guaranteed low latency or high throughput data transfer, either sustained or intermittent. The ability for an application to interact with NGN to request the network characteristics it requires is desirable.

TISPAN Release 2 [i.5] defines an Application Server Function (ASF) which offers value added services and resides either in the user's home network or in a third party location. The third party could be a network or simply a stand-alone AS.

Application Server Functions may provide standalone services or value added services on top of a basic session. For resource control purposes, the first category of Application Server Functions (ASF Type 1) may interact with the RACS, while the second category (ASF Type 2) relies on the control subsystem that provide the basic session over which the valued added service is built.

Examples of Application Server Functions are SIP Application Servers and OSA Application Servers. Further details about which type of Application Server Function can interact with a particular subsystem can be found in the specification of each subsystem in TISPAN.

NOTE: When sitting on top of the IMS, the second type of ASF is identical to the Application Server (AS) function defined in [i.9], although a network node implementing this functional entity in an NGN network and a network node implementing it in a 3GPP network may differ in terms of supported services.

A Grid-enabled NGN application runs within a Grid Application Server (AS).

It needs to be investigated if there is any impact on already defined NGN reference points between the GRID AS and the NGN Subsystems, such as the ISC reference point (e.g.: between IMS subsystem and application servers).

Also it needs to be clarified if Grid-enabled applications can be either ASF 1 or ASF2 type.

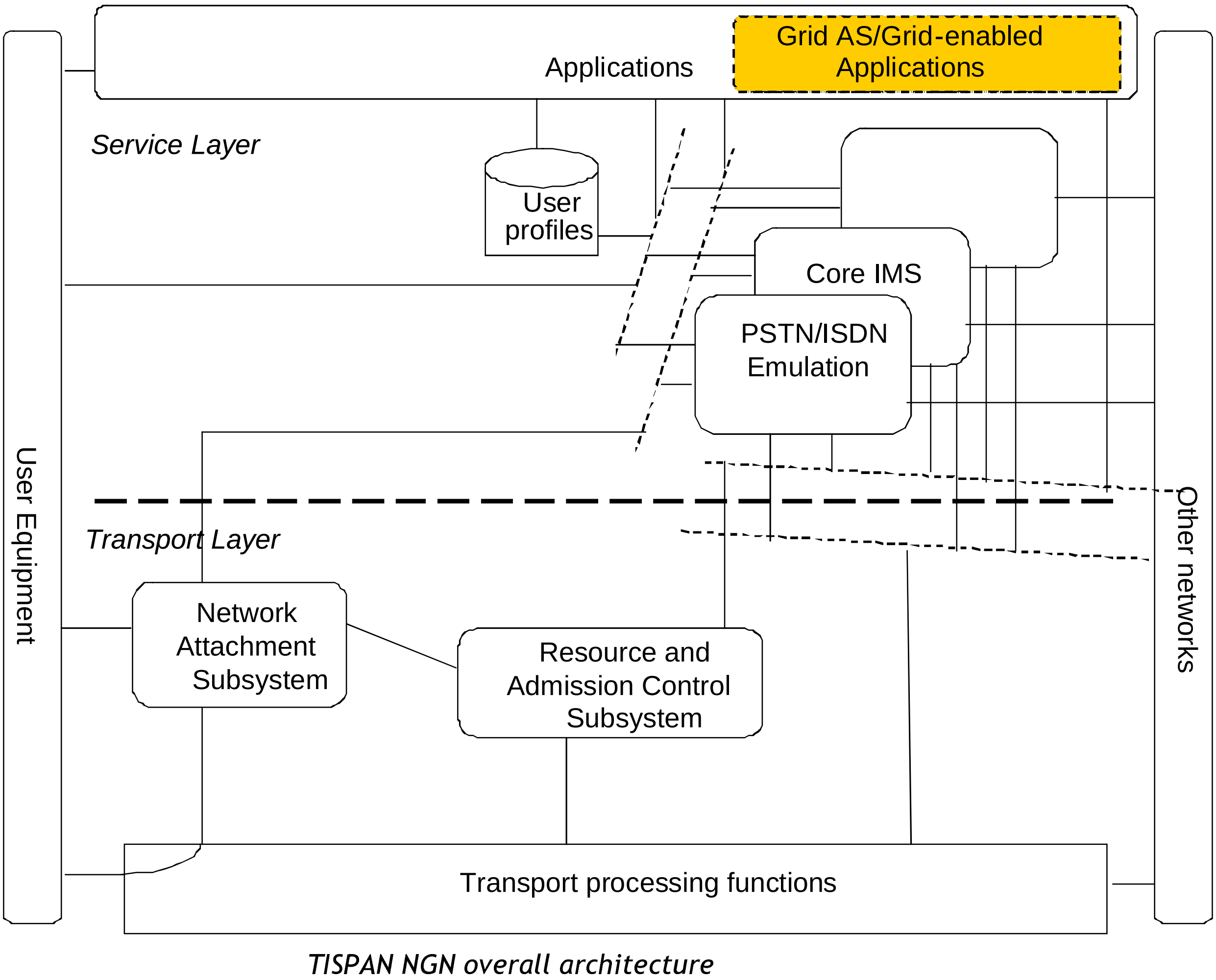

Figure 2: Grid-enabled NGN application

6.3 NGN subsystems offering Grid Services

A new NGN subsystem can support the provision of Grid services using a dedicated service control architecture. This Grid Services subsystem would give access to grid resources, through their virtualization, offering a service interface.

Such an integration of a "Grid Service" Subsystem would enable three different use case options:

1) Offering Grid Services to Grid-enabled applications as defined in clause 6.2 under the scope of the NGN architecture.

2) Offering Grid Services towards end-user applications providing them with resources managed by the "Grid Services" Subsystem.

3) Dedicated grid-enabled functions of other subsystems may use the Grid Service subsystem.

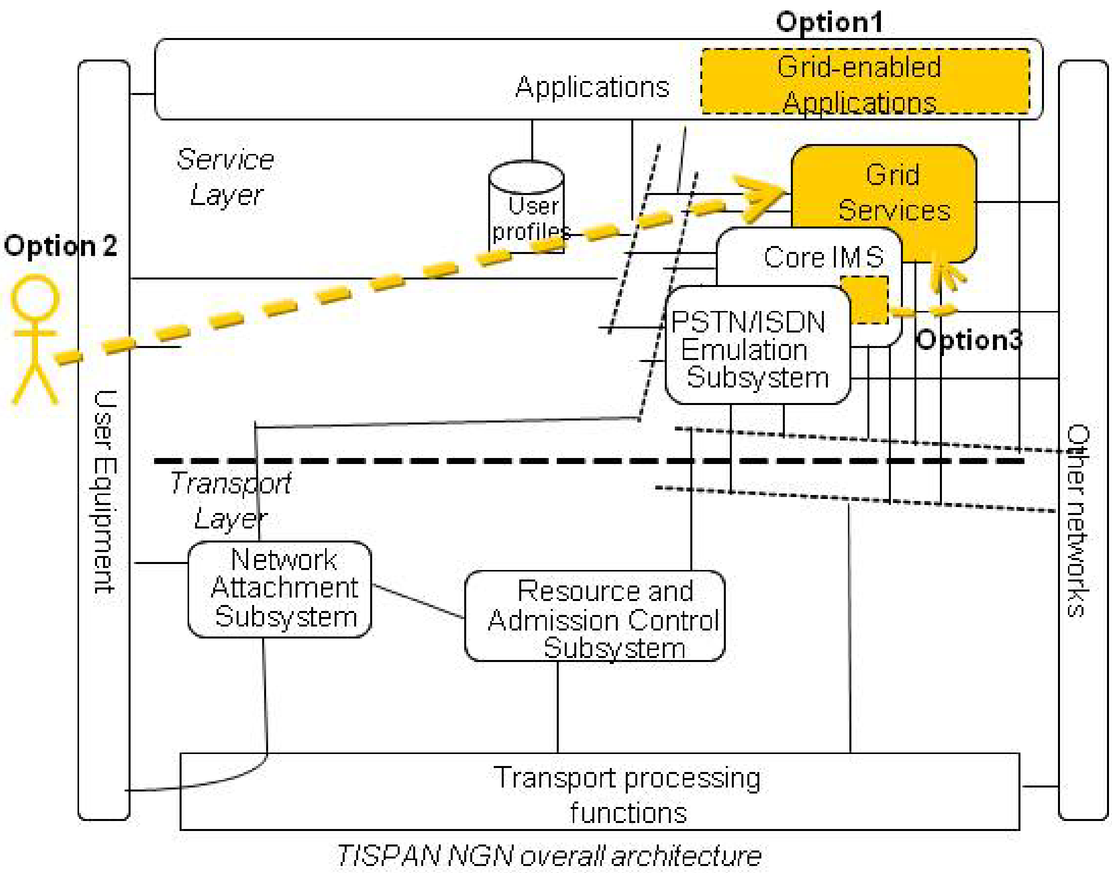

Figure 3: NGN subsystems offering Grid Services

6.4 Grid technology for implementing NGN functionality

Grid technology for implementing NGN functionality means to implement or realize logical NGN functions (from the transport stratum like NASS or RACS up to the Services stratum and the applications) by means of using Grid technology. This would allow the optimization of the resources underneath and a more flexible scalability of the NGN subsystems.

This in only an implementation variant but would need an additional interface to a "Grid Service" entity as depicted in figure 4. As this entity would be required to manage and control the Grid resources.

As this option is only an implementation variant this architectural option is out of scope for work in ETSI Grid.

This does not exclude that NGN architecture or some of its components may need to be adapted if an improvement of functionality or usability can be achieved by usage of utility computing or Grid infrastructure (example peer-to-peer on CSCF and HSS).

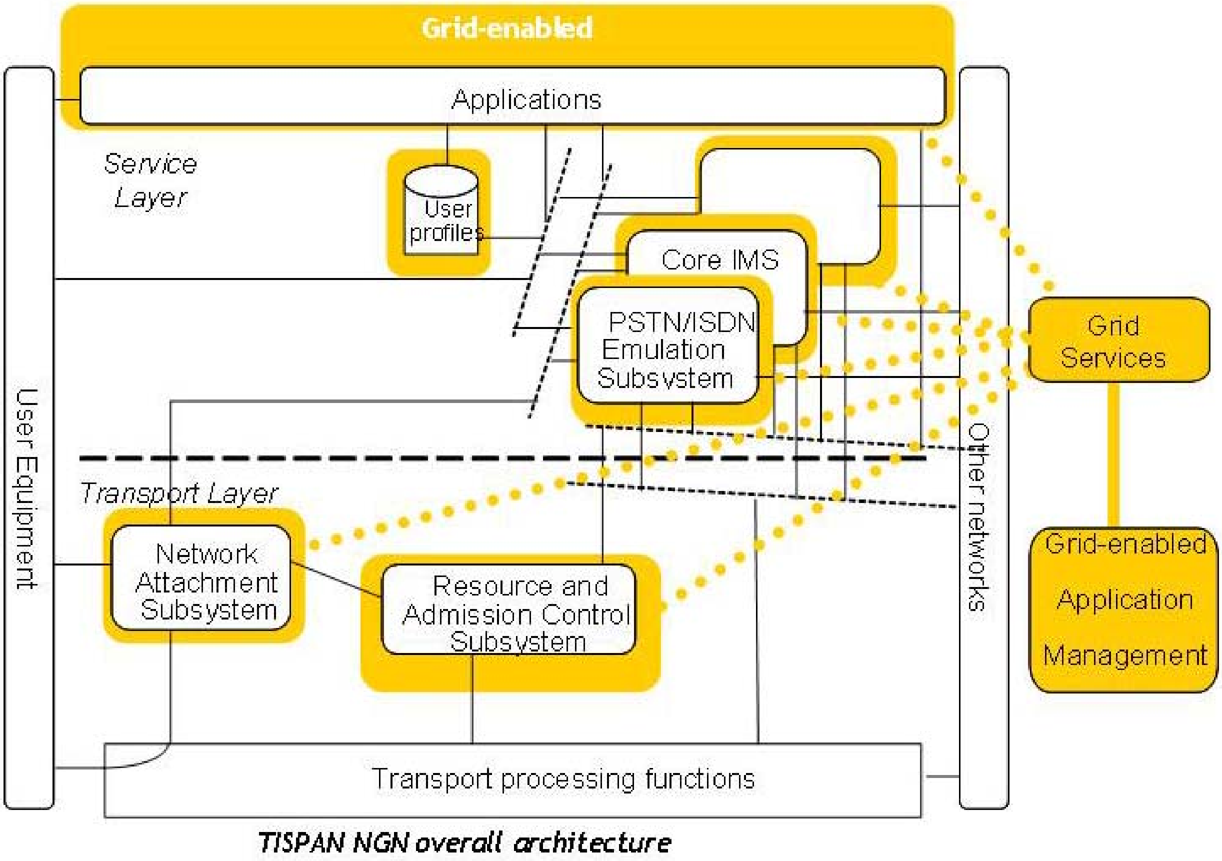

Figure 4: Grid technology for implementing NGN functionality

6.5 Combining Grid and networking resources in a new architecture

The usage of resources like computing power, network and storage resources are orthogonal to the NGN architecture specifying logical functions. This allows defining or adopting standards for the management of the above resources independently of NGN-details.

As such resources will be needed both to run NGN-control, and as resources allocated to subscribers by NGN-control, this should finally lead to a combined architecture, allowing the assignment of all execution, storage and networking resources in a flexible, generic way.

Figure 5 shows the highest level structure of such architecture.

The layer on top of the resources, called "Resource Management" manages all the available resources towards the specific services (e.g. NGN, Grid Service).

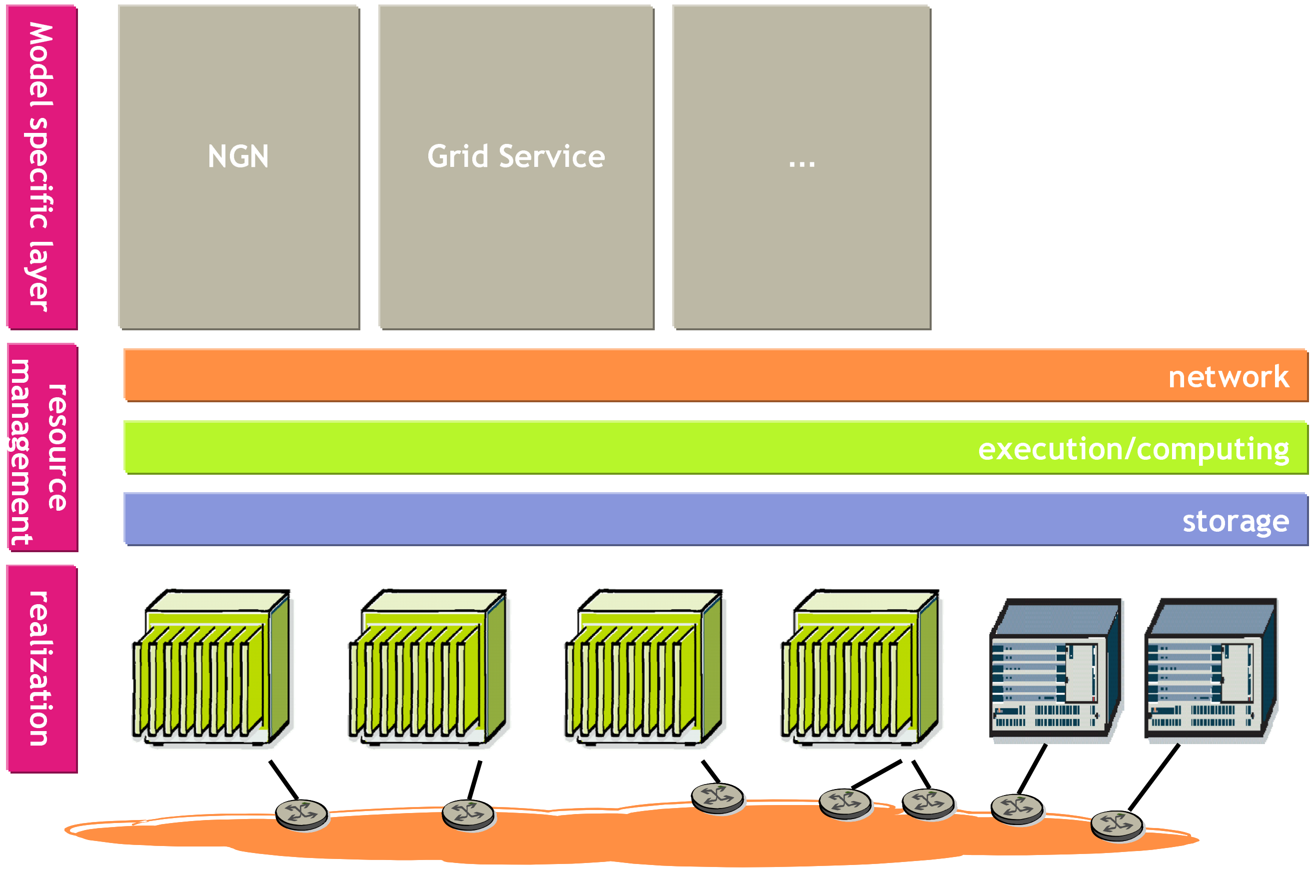

Figure 5: Combining Grid and networking resources in a new architecture

History

Document history		
V1.1.1	February 2009	Publication