

# **Factoring Lower Network Layers into the Grid**

**Masum Z. Hasan, PhD**

**Nino Vidovic, PhD**

**Wayne Clark**

**Cisco Systems, Inc.**

## TABLE OF CONTENTS

<b>1</b>	<b>INTRODUCTION</b>	<b>3</b>
<b>2</b>	<b>GRID DOMAINS</b>	<b>3</b>
<b>3</b>	<b>DOMAIN AND LAYER SEPARATION</b>	<b>4</b>
<b>4</b>	<b>FACTORING L1/2/3 RESOURCES</b>	<b>4</b>
4.1	CONFIGURATION RELATED RESOURCES	4
4.2	NETWORK PERFORMANCE RELATED RESOURCES	5
<b>5</b>	<b>NETWORK BASED GRID SERVICES</b>	<b>5</b>
5.1	USE CASES	5
5.2	NBGS AND GLOBUS MIDDLEWARE	7
<b>6</b>	<b>CONCLUSION</b>	<b>8</b>

## Abbreviations

L7	ISO/OSI Layer 7
L1/2/3	IOS/OSI Layers 1, 2, and 3 (lower layers)
L7GS	Grid middleware, tools or applications where L7 resources (compute, storage) are first-class resources, are termed as a L7 Grid system
NBS	Network based Services: L1/2/3 Network services (such as Bandwidth/QoS, VPN)
NBGS	Network based Grid Services serving limited and abstracted L1/2/3 resources to L7GS
NSP	Network Service Provider
NRN (also known as NREN)	National Research Network (Internet2, National LambdaRail, GEANT, CANARIE, etc.)
EIT	Enterprise IT, which provides IT and networking services to Enterprise departments or groups

## 1 Introduction

In this article we discuss the issues of incorporating lower layer (layers 1, 2 and 3<sup>1</sup>) network resources into the Grid. While typical Grid resources<sup>2</sup> (compute and storage) can be *first-class* Grid resources in a Grid middleware, tool (such as a global scheduler) or an application<sup>3</sup>, depending on network domains in which Grids are used, the L1/2/3 resources can not be first-class resources. A resource is a first-class Grid resource, if an L7GS can have full access to and control of that resource from the top level to its constituent resources. For example, a server can be a first-class top level Grid resource and also its constituent resources, such as the CPU. On the other hand, a router or its constituent resources may not be a first-class Grid resource. A *limited* and *abstracted* set of non-first-class L1/2/3 resources can be incorporated into the Grid. It is limited, because a subset of the resources is considered. It is abstracted, because associated complexities or exact nature of the resources are hidden from an L7GS. For example, while a top level router, its interfaces (links) or its routing protocol related resources may not be incorporated into a Grid, (abstracted) bandwidth or QoS related resources can be incorporated. The factoring of L1/2/3 resources into the Grid will facilitate the following:

- Network-aware functions in L7GS. For example, a global (meta) scheduler can make job scheduling decisions based on L1/2/3 network related information.
- Access to network-based services by an L7GS. For example, bandwidth, QoS, firewall, or MPLS VPN services.
- *On-demand* and *as-needed* configuration and provisioning of L1/2/3 network resources invoked *indirectly* by an L7GS.
- Network based Grid service<sup>4</sup> offerings by network service providers (NSP) or Enterprise IT (EIT).

Note that, the focus in this article is on the use of existing L1/2/3 networking technologies (which include protocols, network services, features, control and management systems) to facilitate network-aware Grid services, rather than the “invention” of some new networking technology for Grids.

## 2 Grid Domains

There are different network domains in which Grids can be used:

- Commercial
  - Enterprise IT managing network and serving various departments
  - Consortium of Grid Enterprises
  - Commercial Network Service Providers (NSP)

---

<sup>1</sup> Termed as L1/2/3 resources

<sup>2</sup> Termed as L7 resources

<sup>3</sup> In what follows, Grid middleware, tools or applications where L7 resources are first-class resources, are termed as a L7 Grid system (L7GS).

<sup>4</sup> Termed as NBGS

- Non-commercial or not-so-commercial
  - National Research Network (NRN)
    - Production NRN (P-NRN)
    - Research NRN (R-NRN)
  - Research community (University and other)

How the L1/2/3 information is incorporated into a Grid may vary from domain to domain.

### 3 Domain and Layer Separation

Domain and layer separation is predominant in the network world, for example:

- ISO/OSI or Protocol layer separation.
- UNI (User to Network Interface) between network domains (between client and server networks: between a customer and an NSP or between an IP Data and a Sonet/SDH/Optical transport network).
- Separation of management and control between data and transport networks in NSP environments.

For the same reason that there are domain and layer separations in network domains, the Grid domain also has to have the same or similar separations. For example, consider the following questions: Can an L7GS (such as a global scheduler):

- Be allowed to configure full capabilities of a router, such as configure routing protocols or
- Be allowed to get full access to network topology to make scheduling decisions or
- Be aware of wide varieties of network technologies (Ethernet, Optical, MPLS, etc.) on which a Grid can be built?

The general answer is no. While the separation can be ignored in R-NRN or purely research networks, it can not be in commercial networks.

### 4 Factoring L1/2/3 Resources

If the full set of L1/2/3 resources cannot be first-class Grid resources in certain network domains, a limited set of them can be incorporated into the Grid. The resources incorporated can also be abstracted. The following two major types of resources and interfaces should be factored into a Grid:

- Network based performance/QoS metrics and interfaces (RQ).
- Network Configuration, service, protocol, and device feature related (RC).

An NSP or EIT can serve RQ and RC to L7GS via an NBGS resident in NSP or EIT.

#### 4.1 Configuration related Resources

A non-exhaustive list of limited RC resources and interfaces is as follows:

- Bandwidth related
- QoS related
- Certain network service related, such as VLAN, VPN, and Firewall.

A non-exhaustive list of resources or interfaces that may not be factored into a Grid is as follows:

- Routing protocol configuration.
- Configuration of low level network service related features, such as MPLS VPN VRF configuration.
- Configuration of low level QoS features, such as complex policing.
- Discovery of network elements.
- Full network topology.

## 4.2 Network Performance related Resources

The Grid community has focused on host to host measurements, where the (L1/2/3) network is considered a cloud. The cons of such measurements are as follows:

- An NSP may block measurement packets (ICMP/UDP) from outside the NSP network.
- Even if measurement packets are allowed, the host based measurements will not support:
  - MPLS data plane measurements, if the network is [G]MPLS<sup>5</sup>.
  - Measurement accuracy along ECMP (IGP/OSPF Equal Cost Multi Path).
  - Per CoS<sup>6</sup>/QoS measurements.

Hence there is need for network-based measurements, which are invoked from within:

- A router or switch.
- An NSP or EIT.

## 5 Network based Grid Services

Consider Figure 1, which shows a number of Grid sites and L1/2/3 networks. The figure shows the L7GSs resident in Grid sites and NBGSs resident in an EIT or NSP (resided possibly in a Network Operations Center). An L7GS interacts with an NBGS to perform network-aware Grid functions. The NBGS can be a component layered on top of a network management system (NMS), which manages and controls L1/2/3 resources. The L1/2/3 resources are first-class resources in an NMS, but not in an NBGS.

### 5.1 Use Cases

Following is a list of use cases which shows how an L7GS (component) can make use of L1/2/3 resources and interfaces served by an NBGS:

- A global/meta scheduler requests bandwidth and QoS configurations between respective

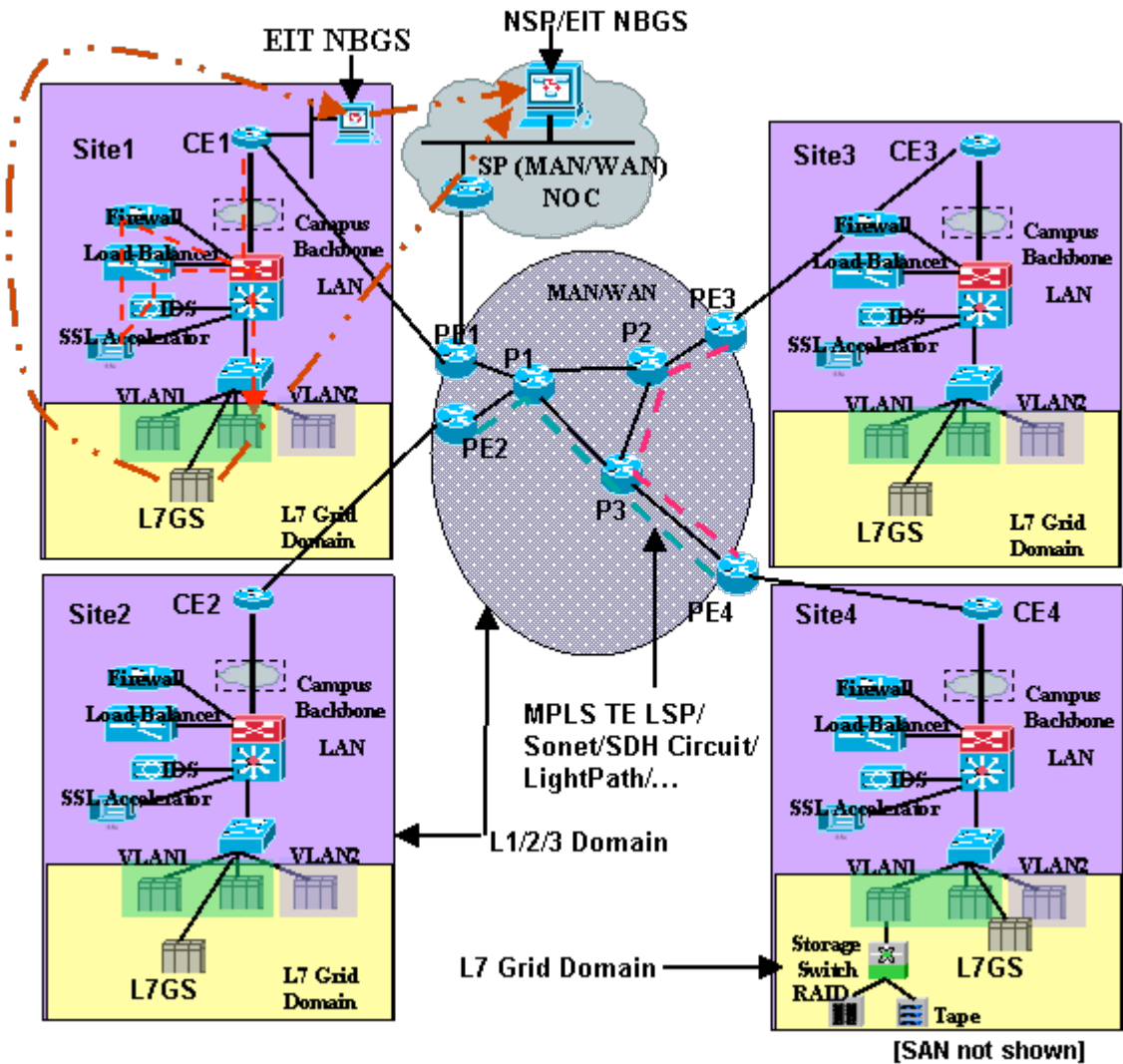
---

<sup>5</sup> MPLS data plane measurements are performed by (UDP based) LSP Pings/traceroute (component of MPLS OAM) initiated from an MPLS router.

<sup>6</sup> CoS: Class of Service

Grid sites before distributing jobs or tasks on Grid sites.

- A global/meta scheduler picks a list of Grid sites for job distribution based on (configured) bandwidth and QoS information.
- A global/meta scheduler picks a list of Grid sites for job distribution based on resource availability in a particular (or a set of) VLAN(s). Available resources in other VLANs are ignored (for security considerations).
- A global/meta scheduler picks a list of Grid sites for job distribution based on resource availability in an MPLS VPN. Available resources in Grid sites accessible via the Internet are ignored. A Grid site belonging to an MPLS VPN can have access to other sites via the public Internet.
- A global/meta scheduler picks a list of Grid sites for job distribution based on network-based performance monitoring information (provided by an NSP or EIT NBGS).
- An L7GS requests an NBGS resident in an NSP addition of a Grid site (the site in which the L7GS resides) to an Extranet MPLS VPN (in which other Grid sites may have already joined, thus creating a secure multi-enterprise Grid).



**Figure 1: Grid Network****5.2 NBGS and Globus Middleware**

Figure 2 shows a (functional) architecture involving NBGS, Globus Middleware (GT4 based), and an NMS. The NBGS can be built using Globus GT4 toolkit. As shown in Figure 2, the NBGS, on one side connects to an NMS or directly to an *intelligent* network, and on the other, to the Globus WS-MDS for resource registration. The L1/2/3 resources in NBGS are specified in WSRF<sup>7</sup>.

With the aid of the NBGS, certain Globus middleware functions, such as RFT or GridFTP can be made network-aware. For example, data transfer can be preceded by any of the following network-aware actions:

- Configure bandwidth and QoS between GridFTP sites,
- Select GridFTP sites based on configured bandwidth/QoS,
- Select GridFTP sites based on network-based performance monitoring information.

---

<sup>7</sup> The WSRF has been adopted by the GGF and implemented in the GT4.

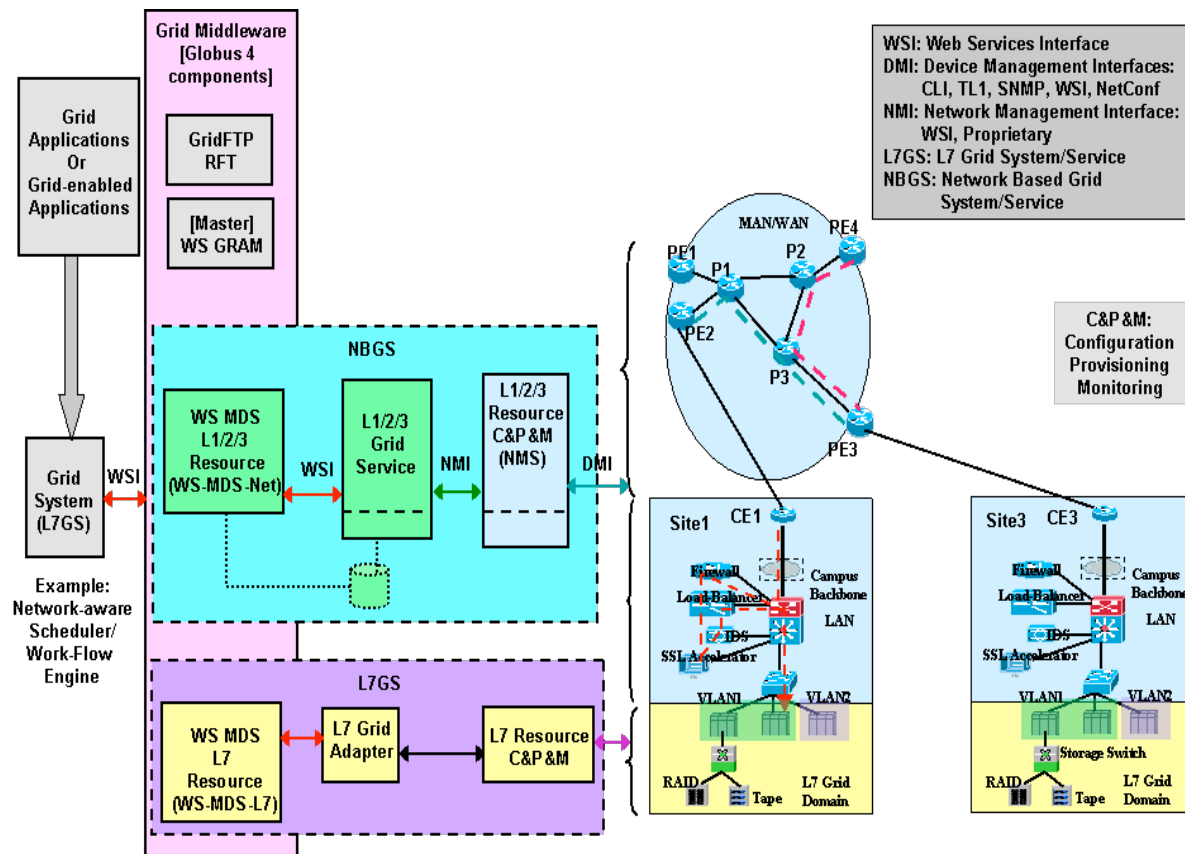


Figure 2: L1/2/3 Grid Service NBGS and Globus

## 6 Conclusion

Factoring limited and abstracted L1/2/3 resources into the Grid and serving these resources via a dynamic L1/2/3 Grid service (NBGS) may have many implications, including:

- The traditional single LAN based Grids are extended to the full domains of LAN, MAN and WAN. Note that, even though a traditional Grid may span across a MAN or WAN, the domain is still L7 in each Grid site (as shown by the yellow region in Figure 2).
- The L7GS systems can provide network-aware Grid functions.
- NSPs can provide new breed of services, which is *dynamic network based Grid services*, to their Enterprise customers utilizing existing networking infrastructures.
- Optimization and cost reduction (better price/performance ratio) achieved through the adoption of Grid is extended to L1/2/3 network resources, and across LAN, MAN, and WAN.

From a standards perspective, there is still substantial work to be done, such as the definition of L1/2/3 configuration and monitoring resources and interfaces using WSRF.

When an EIT or NSP deploys a web services based service like NBGS, security becomes an

issue (like any other dynamic/online service). The non-first-class and abstract nature of resources provides added level security on top of existing security mechanisms (such as GGF/GT4 GSI).