

# Token Based path authorization at Interconnection Points between Hybrid Networks and a Lambda Grid.

*By Leon Gommans, Cees de Laat, Robert Meijer, University of Amsterdam.*

**Abstract-**In order to provide cost effective transport services for highly demanding data-intensive grid applications, National Research Networks (NRNs) are considering additional types of access to their network infrastructures. Next to traditional IP access, NRNs like to provide automated, grid application driven access to their underlying connection-oriented network infrastructure. This combination is called hybrid networking. Recently, both NRNs and grid communities started to acquire their own global optical network connections. A Lambda Grid was created when these organizations decided to interconnect these links via interconnection points such as Starlight and Netherlight. Apart from supporting scientific applications, the Lambda Grid allows network research. Within this context, this paper will present a novel token-based path selection mechanism that will enable authorized access to Lambda Grid links. The token-based approach allows temporal separation of the path authorization process from obtaining access to a Lambda Grid link. The path authorization process may involve many parties and complex, time consuming decisions whereas path access requires a fast real-time implementation. We will describe the application of the token-based approach for an interconnection point between a Hybrid Network and a Lambda Grid.

## 1.0 Introduction.

Recent regulatory changes enabled NRNs and scientific organizations to acquire their own global transport connections. Such connections are used to by-pass the existing Internet allowing scientific applications to run at lower cost. Collaborative efforts merged these connections into an infrastructure called a Lambda Grid. Allowing owners of Lambda Grid links to

authorize specific user applications to choose a more economic route is an evolution of the Internet that is addressed in this paper. This paper presents an in-packet token-based approach that can be integrated with an automated network path setup-, reservation- and authorization system. In order to position the approach, the paper will first explain the structure of the current Internet. NRNs can make Lambda Grid links accessible via their hybrid networks. We will therefore need to describe the basic concepts behind hybrid networks and show how hybrid networks interconnect. The presented approach performs authorized path selection at interconnection points between a hybrid network and the Lambda Grid. Grid compliant implementations of network path selection- and control functions typically involve web services mechanisms. Recent availability of cryptographic network hardware and disappointing experiences with the performance of web services has led our research to a novel, token-based authorization approach. A high performance switch equipped with the aforementioned cryptographic hardware acts as a real time path selection mechanism, using a token-key to recognize tokens inside IP packets. Before it issues a token-key to a network user, a web services based authorization process could involve several stakeholders and take complex decisions. Next to technological aspects, economic and regulatory factors are used to introduce some of the evolutionary aspects of the Internet. This paper assumes some conceptual knowledge behind authorization such as described in RFC2904<sup>1</sup> and GFD.38<sup>2</sup>.

## 2.0 Internet Interconnection principles.

The evolution of the telephone network in the US created a hierarchical network structure in which the FCC regulated fair competition amongst long distance carriers. In 1982, an anti-trust case<sup>3</sup> was settled where AT&T was forced

to split off its long-distance services and divest its 22 local telephone companies into 7 separate regional operating companies. Legal cases, such as the AT&T case, did influence the structure of the Internet since the Internet backbones were all owned by Telco's. National regulatory bodies such as the FCC closely watched the telecommunications industry structure as to promote competition. This evolved the Internet into a three-tier structure as shown in fig. 1.

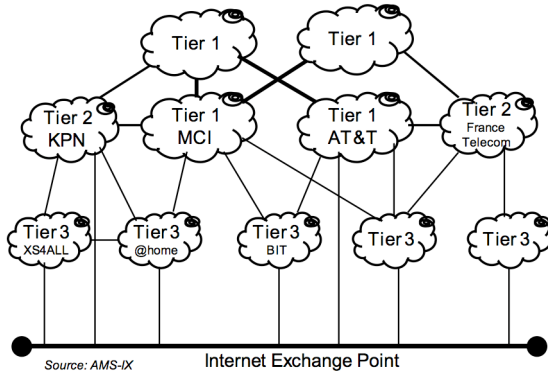


Fig 1. The hierarchical Internet structure

Within this structure, Tier-1 carrier networks such as AT&T and MCI provide both long distance transport and transit services. Tier-2 networks provide regional transport and transit services. Tier-3 networks are national Internet Service Provider (ISP) networks. A Tier-3 network may obtain transit and transport from a Tier-1 or -2 network. For simplicity reasons, we refer to the combined set of Tier-1 and Tier-2 networks as the Transit Network. A Transit Network, by definition, provides connectivity to the entire Internet.

If an ISP discovers that it exchanges a significant amount of traffic with another ISP, both ISPs may find it more economical to create a direct link between each other<sup>4</sup>. Traffic will then be routed via this link, by-passing the Transit Network as shown in fig 2. Policies inside routers that border other networks, will determine the desired routes of the traffic. Border routers use a protocol called the Border Gateway Protocol<sup>5</sup> to communicate available routes. When a group of ISPs see similar needs, ISPs may decide to form an Internet Exchange Point (IXP). At this point ISPs can peer with

each other. The bottom bar in fig.1 shows an Ethernet segment that forms an exchange.

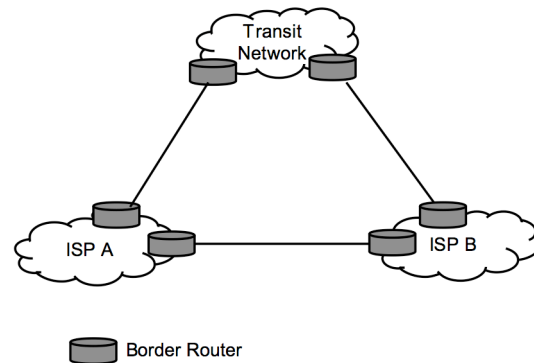


Fig 2. Direct ISP Peering.

It connects Tier-1, 2 and 3 networks via a common network. For simplicity reasons, we again consider the collection of Tier-1 and Tier-2 networks as the Transit Network. This simplification yields fig. 3. According to fig. 1 Tier-1 and -2 networks, also connect to an exchange. Fig. 3 is therefore not entirely accurate. Norton<sup>4</sup>, however, uses a similar model as fig.3 when explaining the business case for peering. We will therefore continue to use the simplified Norton model.

The admission policy to the Amsterdam Internet Exchange (AMS-IX) is called open: any legal entity with a registered Autonomous System Number<sup>6</sup> is allowed to become a member. Other exchanges may not reveal their

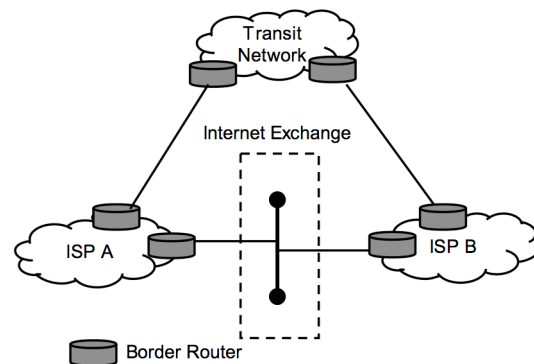


Fig. 3. ISP peering via an Internet exchange point.

admission policy or pose restrictions on data-flows. However, in our paper we assume that an

exchange has an open admission policy and does not impose any restriction. This means that an exchange should allow any member to create a peering relationship with another member autonomously. When peering connection-oriented hybrid network links, one must observe similar requirements. Let us now consider hybrid networks in more detail.

### 3.0 Hybrid Networks.

In traditional IP-networks, the collection of high-capacity links that interconnect routers is called the low-level infrastructure. It uses layer 2 or below technologies to transport data. Regulatory changes, such as the 1996 Telecommunications Act in the US, enabled companies to acquire telecommunication services competitively, including optical fiber links. Tier-1, -2 and -3 networks and even end-users could now start to own fiber-optic links. This enabled IP network operators to offer cheaper lower-level point-to-point transport services, thereby creating hybrid networks. Hybrid Tier-1 or Tier-2 networks could offer both IP transit services and point-to-point transport services. A hybrid ISP could provide point-to-point transport or transport from one point to an upstream network. Alternatively, it could peer with another ISPs. This peering can either be direct, using a peering link, or peering can be done via an exchange point.

As seen in chapter 2, traditional exchanges use a common Ethernet where flows are determined by BGP peering policies between corresponding networks. An exchange, supporting connection-oriented flows from hybrid networks, must perform a switching function to select a certain path. Considering the requirement mentioned at the end of the previous chapter, the path selection should be authorized by a policy-based decision performed amongst the peering entities.

### 4.0 Hybrid ISP peering.

We saw that peering between traditional ISPs is done for economic reasons. We assume the same for hybrid networks based on the following rationale:

1. Lower layer connection-oriented switching between specific data-intensive sites is cheaper than connection-less routing<sup>7</sup>.
2. There is an economic break-even point between peering and transit when transporting increasing volumes of data<sup>4</sup>.

If there is a choice between connectionless and connection-oriented peering, the above rationale will favor connection-oriented peering for high-volume data transfers.

Tier-1 networks offer both transport and transit at global scale. Transit, for connection-oriented flows, must be translated into the ability of a hybrid network to connect to multiple destinations on demand. This on-demand capability can be implemented by offering a control-plane interface into the network. Several web services based control-plane interface implementations are being researched in projects such as UCLP from Canarie and the Generic AAA project from University of Amsterdam. Considering the involved economics, a hybrid ISP must decide, if it connects to a Tier-1 or -2 network or if it peers with another ISP via an exchange. In any case, the involved networks must allow some form of control signaling and some mechanism must exist to select and setup an end-to-end connection. When peering, the exchange must also allow control signaling to select a path that is part of an end-to-end connection. The following consideration, at least from a scientific viewpoint, makes path selection for hybrid networks more complex. We observe in the scientific grid world, that optical global transport connections are scarce resources typically owned by organizations that are either a NRN or a stakeholder in a grid community. Global optical connections are mostly acquired for longer periods and are therefore typically statically provisioned. As more NRNs and scientific communities started to own such connections, the Global Lambda Integrated Facility (GLIF - [www.glif.is](http://www.glif.is)) was formed in 2003. When created, an important aim of the GLIF was the development of functions and services allowing flexible use of what is called a Lambda Grid. Within this network, global transport connections terminate at special

interconnection points such as Netherlight and Starlight. An Inter-Connection Point (ICP) is a general term for the technical facility that allows termination of network connections with the intent to interconnect. The GLIF essentially creates a network of ICPs, allowing scientific applications to peer at a global scale. The current drawing of the GLIF shows that for example CERN and SURFnet own global connections. CERN, when moving terabyte-scale data files on behalf its LHC experiments, might want to restrict other access. SURFnet may decide to allow any application provided it serves a scientific goal. A hybrid ISP network may or may not sit in between a GLIF interconnection point and the user application. NRNs, acting as ISP or Tier-2 network, position themselves to play a role between Grid applications and the GLIF. The network of ICPs and global links, where link owners plays a distinct role in admitting data-flows, creates a new networking situation. We assume that the transit network is the IP network by-passed by the connection-oriented network. Instead of an exchange, a network of ICPs creates the by-pass. A simplification of this situation is shown below.

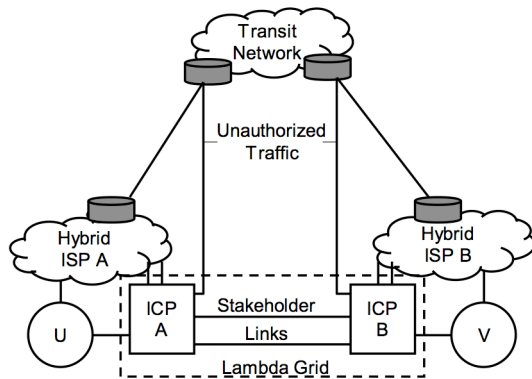


Fig 4. Peering via a Lambda Grid.

Fig. 4 shows that user applications U and V have a choice to peer directly via ICP-A and ICP-B or connect via hybrid ISP-A and ISP-B that connects to ICP-A and ICP-B. Links, owned by one or more stakeholders, interconnect ICP-A and -B. When considering the GLIF, ICP A could be Netherlight in Amsterdam and ICP B could be Starlight in Chicago. Stakeholders, such as CERN or SURFnet, may want to authorize applications U and V to peer via its

Lambda Grid link. Being connection-oriented, ICP-A and ICP-B both enforce access to a Lambda Grid link. ICPs and hybrid networks need to signal each other to create an end-to-end path. In his description of exchange models, Dijkstra<sup>8</sup> recognizes that policy based authorization mechanisms can be used to signal and control optical exchanges. Dijkstra also recognizes that peers are responsible to control their own network interfaces on an exchange. We can extend this observation with the recognition that Lambda Grid link owners have similar responsibilities. We conducted experiments to perform end-to-end path authorization signaling involving multiple autonomous networks. During these experiments, network control interfaces were cast as a web service. Let us now consider some results of these experiments and explain why the token-based authorization approach is helpful.

## 5.0 Web Service performance and Token Based Authorization.

Implementations of modern low-level infrastructures offer elaborate management and control capabilities by means of an Operations Support System (OSS). Examples studied were Nortel's DRAC and Alcatel's 1355 BonD. In general, an OSS performs management, inventory, engineering, planning, and repair functions for communications service providers. These functions represent a single administrative domain. An OSS creates specific connections on behalf of the administrative owner of a network domain. The owner, however, does typically not allow an outside entity to control the provisioning of its connections. We developed an authorization system based on the principles of the Generic AAA Architecture (RFC2903<sup>9</sup>). In several experiments using this system, an AAA server acts for clients as a proxy or broker of an OSS. Oudenaarde<sup>10,11</sup> describes experiments that shows how AAA servers can set up end-to-end path across a set of connection-oriented networks. One conclusion of the experiments expressed concern about the performance of a web services based control interface implementation. Experience from these experiments also showed that the ease of combining a set of services into a distributed

application is the main value of web services. These conclusions create a dilemma when considering authorization mechanisms that involve web services interfaces.

Our research into this issue, and the recent emerge of network switching technology offering cryptographic functions, made us consider the RFC2904<sup>1</sup> push sequence. With the push (or token) sequence, the user makes a request to the AAA server. The AAA server applies a set of policy conditions to authorize the request. Instead of sending commands to the service, the policy action returns a signed list of attributes called a token. The token could be used anytime or at an agreed time. When presented to the service, the authenticity of the token proves to the service that the AAA server has issued the token and the service is expected to act accordingly. This model effectively decouples the taking of a decision from the usage of the decision. A token can be assigned a validity time and can be kept, stored and maybe handed over to another user. If a token contains a token-key, the token-key could be used to generate a new token. These tokens must be securely communicated as signing does not provide confidentiality. If the AAA server acts as a proxy to a resource allocation manager, and if bound to corresponding attributes, a token allows precise resource management and pre-allocation. A token can be requested well ahead of time. These features are expected to help us with solving the timing issues when collecting authorizations from different places using web service interfaces. Once the authorizations are collected, a token could provide real-time access to a network link.

## **6.0 Token model at Interconnection Points.**

Our experiments, referenced in chapter 5, assumed direct peering between networks when they were chained to create an end-to-end path. All control and management functions are implemented inside each network domain. Networks signal path authorization messages using a network of AAA servers. We will now consider situations where a path will explicitly traverse a Lambda Grid and their ICPs. As noted in chapter 4, ICPs need a switch to select a path. A switch needs control input in order to

choose the path. When applying the token model to control an ICP switch, the token can be handed in two different ways:

- 1) Out-of-band – here the switch control function (or higher level function) accepts a token via a separate interface that selects a path within the switch. The control function checks the authenticity and integrity of the token. The embedded information inside the token is used to provision the switch.
- 2) In-Band – Tokens are inserted into the IP-packets that flow through a switch. The switch will base its forwarding decision on the integrity and authenticity of a token. A switch must be pre-provisioned with the right key-material and forwarding information. The key material is needed to allow checking of the token. The forwarding information is needed to determine the output port. A default port can be configured to forward IP packet-flows with no or invalid tokens. A default port could be connected to a best-effort network.

The token must be requested, authorized and created first. A client can make the request using web services mechanisms to an AAA server. This AAA server can act as a broker between a set of other AAA servers. It can collect the proper authorizations from the involved stakeholders in an arbitrary complex way. Recent research of several groups within the GLIF community suggests that workflow mechanisms could be helpful in automating such scenarios. We will not go into further detail here. We assume that an authorization for a path is obtained somehow.

Considering recent advances in high-performance network chip technology used inside switches, the in-band signaling approach deserves a closer look. As said, the in-band approach requires a switch to recognize tokens present in the data-flow. Tokens must be inserted into each datagram at some point in the network, e.g. at the ingress point of the hybrid ISP. The hybrid network will forward the datagrams to the switch of the ICP. Recognizing

tokens requires a switch to perform cryptographic functions such as a Message Authentication Code calculation. These types of calculations can be performed at very high speeds, using the cryptographic functions present in the latest generation of Network Processor Units (NPUs). This feature makes application of a token-based switch feasible for application within an ICP. We use an Intel IXDP 2850 development platform to implement such a type of switch. The token-based switch is both programmed to recognize tokens and to generate and insert a token into a datagram.

There are several options to insert a token into a datagram. A straightforward way is to insert a token into the IP-options field. IP-options is defined in RFC791<sup>12</sup>. This field can be of variable length and resides in the IP header. An IP packet containing options should be forwarded unmodified by a router. A router does not need to understand IP options. This makes the use of IP options transparent to a connectionless routed network in case a token is for example invalid. When a token inside the IP options field is used to make a path decision and when the other IP header information is used to route the packet, the principle elegantly marries connectionless and connection-oriented networking.

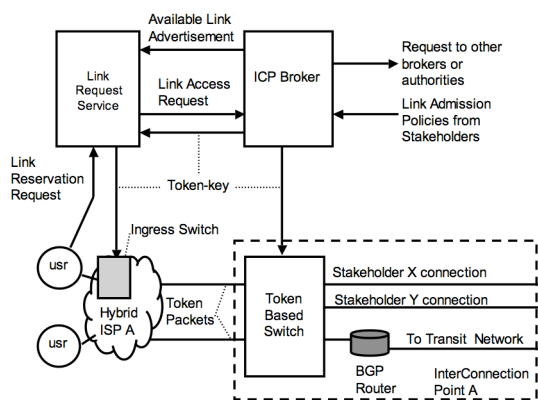


Fig 5. Token based framework for an Interconnection Point. This picture shows a token-based switch that is controlled by an AAA server that both accept link access requests from a user (via a web server) and link access policies from the link stakeholders.

Fig 5. Shows in more detail the network and control plane components around ICP A as shown in fig 4. It is assumed that a user application will use Hybrid ISP A to connect to some service via a peering ISP. The shown control plane elements all use Generic AAA toolkit components to allow distributed policy driven decision taking. A Link Request Service (LRS) generates a Link Access Request (LAR) message to use a particular link. The ISP may for example provide this service. Amongst many possible policy driven scenarios, the LRS could first authenticate the user application and then offer a selection of appropriate links to this application. This information is based on information pushed by the ICP broker to the LRS. Link owners provision the ICP broker with link admission policies that will indicate who may use which links at what times. After the application decided which link it wants to use when, the LRS will compose a LAR defining a period of time and a specific link. This LAR will request one or more tokens from the broker. The tokens may be bound to specific timeslots. Before authorizing the request for the tokens, the driving policy of a broker may define that it needs to contact one or more other brokers along the end-to-end path. The LAR could contain several additional attributes including for example the Source and Destination IP (SIP/DIP) address of the stations or networks that want to communicate across the Lambda Grid link. The token switch can then recognize specific traffic, which it must treat in a connection-oriented fashion. With SIP and DIP, the user essentially asks a connection to be authorized between two classless IP addresses. After the ICP broker receives this information, the next chapter will explain how a token is created and inserted into the packets send to the token-based switch.

## 7.0 Token Creation.

A token is essentially the result of applying a MAC algorithm on a number of fields of the IP-datagram as shown in fig. 6.

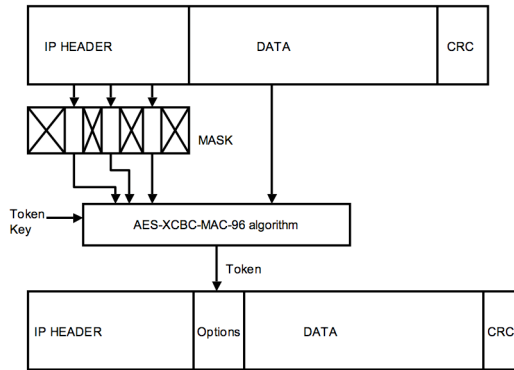


Fig 6. The token creation process masks parts of the IP Header and IP data field to create a unique token for each IP datagram using a key dependant MAC

A MAC algorithm is a key based way to create a one-way hash. The key is called the token-key. As the result of a masked IP header field may yield the same information for consecutive IP-datagrams, one can avoid the generation of duplicate tokens by including (parts of-) the data-field into the MAC calculation. Further research must motivate specific choices. A MAC algorithm provides data integrity and data origin authentication with respect to the originator of the message. An algorithm that is considered safe in creating a MAC for variable length messages is described in RFC3566<sup>13</sup>. If a MAC is created this way on an IP-datagram, the receiver of the MAC can be sure that a device holding the proper token-key generated the packet. We now call the IP-datagram containing a token, a token-packet. The token-key must only be provided to devices authorized to generate or check a token-packet. The token-key must only be valid for the time a link may be used. This will allow a link resource manager to provide precise control of the link usage. The ICP brokers, as shown in fig. 5, must distribute token-keys both to the token-based switches along the path as well as to the devices that insert a token into the token-packets. Fig. 5 shows that the ISP will insert the token in a token-packet on behalf of the user. Insertion could be performed, for example, on the ingress switch of the ISP. This approach is called in IETF terminology the “bump-in-the-wire” approach. An alternative would be the “bump-in-the-stack” approach, where the IP stack of the end-station would insert a token. When packets

arrives at the token-based ingress switch, the token-key is used to generate a token according to fig. 6 and subsequently inserted into the IP options field.

The token-packet is then forwarded to an output port on the switch that (virtually) connects to a path provisioned by some control mechanism inside the ISP. This path leads to the ICP. The token switch inside the ICP will verify the token by applying a similar MAC algorithm with the same token-key on the token-packet. It will apply the same mask to the IP-header to extract the desired field and will combine this with (parts of-) the IP-data field. A MAC will generate a cipher that will be compared with the token inside the IP options field. If a match is found, the token is authentic and the packet can be forwarded to a designated port owned by the stakeholder that was involved in the process that issued the token-key. This process effectively selects the authorized path in real-time. If the token does not match or if no token is present, the token-packet will be forwarded to a default port. This port could be connected to a border router that connects to a transit provider who, by definition, can forward the token-packet to its destination.

A token-switch, using the Intel IXP 2850 employing 16 micro-engines and 2 crypto-units, is expected to handle token generation and recognition at speeds of up to 10 Gb/s.

## 8.0 Conclusions and future research.

We first described the evolution of the Internet into a three-tier structure, where by-passing a transit network is common practice based on economic factors. Regulatory changes allowed any organization to own long distance transport connections, which created, through scientific collaboration, a Lambda Grid. From a scientific viewpoint, we assumed a desire by Lambda Grid link owners to authorize access to its resources for individual user communities. Within this context, we considered that:

- A token-based switch can select an authorized path at an interconnection point between a hybrid network and a Lambda Grid link in real time.

- Web service based components can be used to authorize link access using policy driven scenarios. The obtained authorization can then be transformed into a token-key, which can be used to generate token-packets whenever an authorized path must be accessed.
- A connectionless network is able to transparently forward a token-packet.
- A token-based switch can both connect to a connectionless and connection-oriented network. An IP-datagram without a valid token should be forwarded to the connectionless IP transit network. An IP-datagram with a valid token is switched to a particular connection-oriented link of a stakeholder. As such, we showed that a token switch effectively marries both networking paradigms.
- Current NPU technologies should be able to provide the necessary processing power to handle speeds up to 10 Gb/s.

We need to proof the presented concepts by conducting experiments with the IXDP 2850 NPU development platform. Many questions are still open for research: How could the initial address resolution process be performed best? How should IP fragmentation be handled? What is the optimal granularity for token-validity? Can tokens be re-used, grouped, shared etc. We saw that tokens could be bumped-into-the-wire at the ingress switch of an ISP. This will allow future research into firewall type functions for high volume streams. Can a modern network card, allowing TCP off-loading, be used to bump the token in the stack?

### Acknowledgements.

The research was funded by the GigaPort NG project lead by SURFnet and the NextGRID project (contract number 511563), which is funded by the European Commission's IST 6<sup>th</sup> Framework Programme. The authors like to thank Franco Travostino from Nortel for his significant input and valuable discussions. We like to thank Henk Steenman and Job Witteman from AMS-IX for sharing their view onto the

Internet world, Christophe Diot from Intel for sponsoring the research efforts, Herbert Bos from the VU Amsterdam and Mihai Cristea from Leiden University for their valuable input on the Intel development platform

### Reference

- <sup>1</sup> **RFC2904 AAA Authorization Framework**, J. Vollbrecht, P. Calhoun, S. Farrell, L. Gommans, G. Gross, B. de Bruijn, C. de Laat, M. Holdrege, D. Spence. IETF 2000  
<http://www.ietf.org/rfc/rfc2904.txt>
- <sup>2</sup> **GFD.38 Conceptual Grid Authorization Framework and Classification**, M. Lorch, B. Cowles, R. Baker, L. Gommans, P. Madsen, A. McNab, L. Ramakrishnan, K. Sankar, D. Skow, M. Thompson, GGF 2005.  
<http://www.ggf.org/documents/GWD-I-E/GFD-I.038.pdf>
- <sup>3</sup> **Civil Action No. 82-0192**, United States of America, Plaintiff v.s. Western Electric Company, Inc. and American Telephone and Telegraph Company, Defendants. Plan of Reorganization, Dec. 16<sup>th</sup>, 1982. <http://www.bellsystemmemorial.com/pdf/82-0192.pdf>
- <sup>4</sup> **A business case for ISP Peering**, William B. Norton.  
[www.equinix.com/pdf/whitepapers/Business\\_case.pdf](http://www.equinix.com/pdf/whitepapers/Business_case.pdf)
- <sup>5</sup> **RFC1771, A Border Gateway Protocol 4**, Y. Rekhter, T. Li, IETF 1995, <http://www.ietf.org/rfc/rfc1771.txt>
- <sup>6</sup> **RFC1930, Guidelines for creation, selection, and registration of an Autonomous System**, J. Hawkinson, T. Bates, IETF 1996, <http://www.ietf.org/rfc/rfc1930.txt>
- <sup>7</sup> **The Rationale of the Current Optical Networking Initiatives**, Cees de Laat, Erik Radius, Steven Wallace, iGrid2002 special issue, Future Generation Computer Systems, volume 19 issue 6 (2003)
- <sup>8</sup> **Control Models at Interconnection Points**, F. Dijkstra, B. van Oudenaarde, B. Andree, L. Gommans, J. van der Ham, K. Koymans, C. de Laat, contribution to this issue.
- <sup>9</sup> **RFC2903, Generic AAA Architecture**, C. de Laat, G. Gross, L. Gommans, J. Vollbrecht, D. Spence. IETF 2000.  
<http://www.ietf.org/rfc/rfc2903.txt>
- <sup>10</sup> **Dynamic paths in multi-domain optical networks for grids**, S. van Oudenaarde, Z. Hendrikse, F. Dijkstra, L. Gommans, C. de Laat, R.J. Meijer, Future Generation Computer Systems 21 (2005) pg 539-548.
- <sup>11</sup> **Applications Drive Secure Lightpath Creation across Heterogeneous Domains**, Bas van Oudenaarde et. al., contribution to this IEEE Communication Magazine on Optical Control Plane for Grid Networks.
- <sup>12</sup> **RFC791, Internet Protocol**, J. Postel, IETF 1981,  
<http://www.ietf.org/rfc/rfc791.txt>

---

<sup>13</sup> **RFC3566, The AES-XCBC-MAC-96 Algorithm and Its Use  
With IPsec**, S. Frankel, H. Herbert, IETF 2003,  
<http://www.ietf.org/rfc/rfc3566.txt>