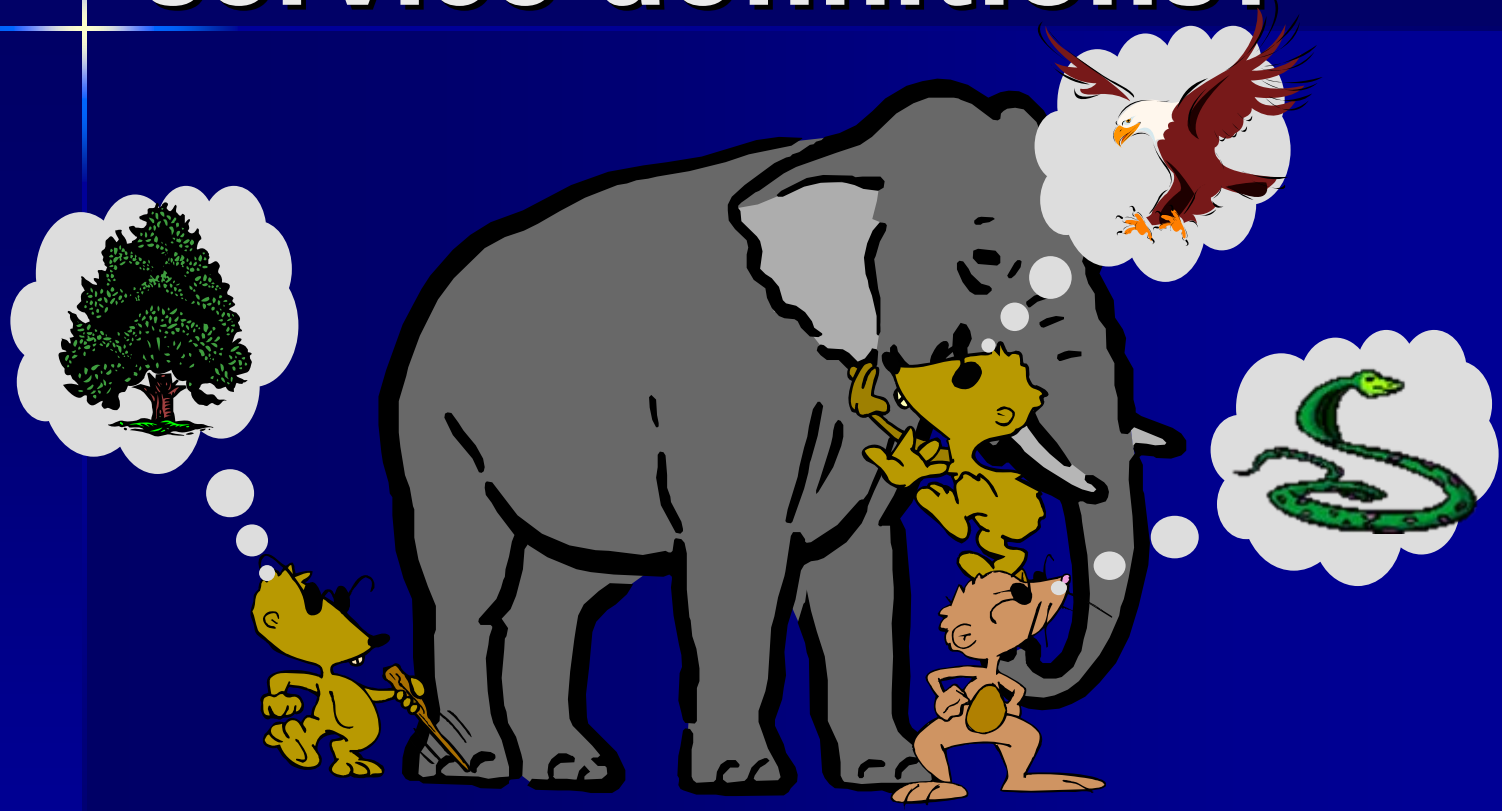
Common Service Definitions

Jerry Sobieski
Director, Research Initiatives
Mid-Atlantic Crossroads

Presented to The GLIF
ABQ (that city in New Mexico, USA)
February 8, 2006

Why do we need formal service definitions?



The three blind application developers and the network elephant…

The Problem:

- We are building networks with new capabilities that go well beyond conventional IP services
  - “non-routed” or “light path” services are a new twist on well established ideas but have never been deployed on the scale we see now
  - Capacities requirements for e-science are exceeding existing technology insertion models...
- These network capabilities are *not* standardized [yet]
  - Even “similar” services exhibit significant differences that prevent interoperability
- How do we create a true global quilt of interoperable and end-to-end lightpath services that provides the consistency and reach of the current R&E reach?

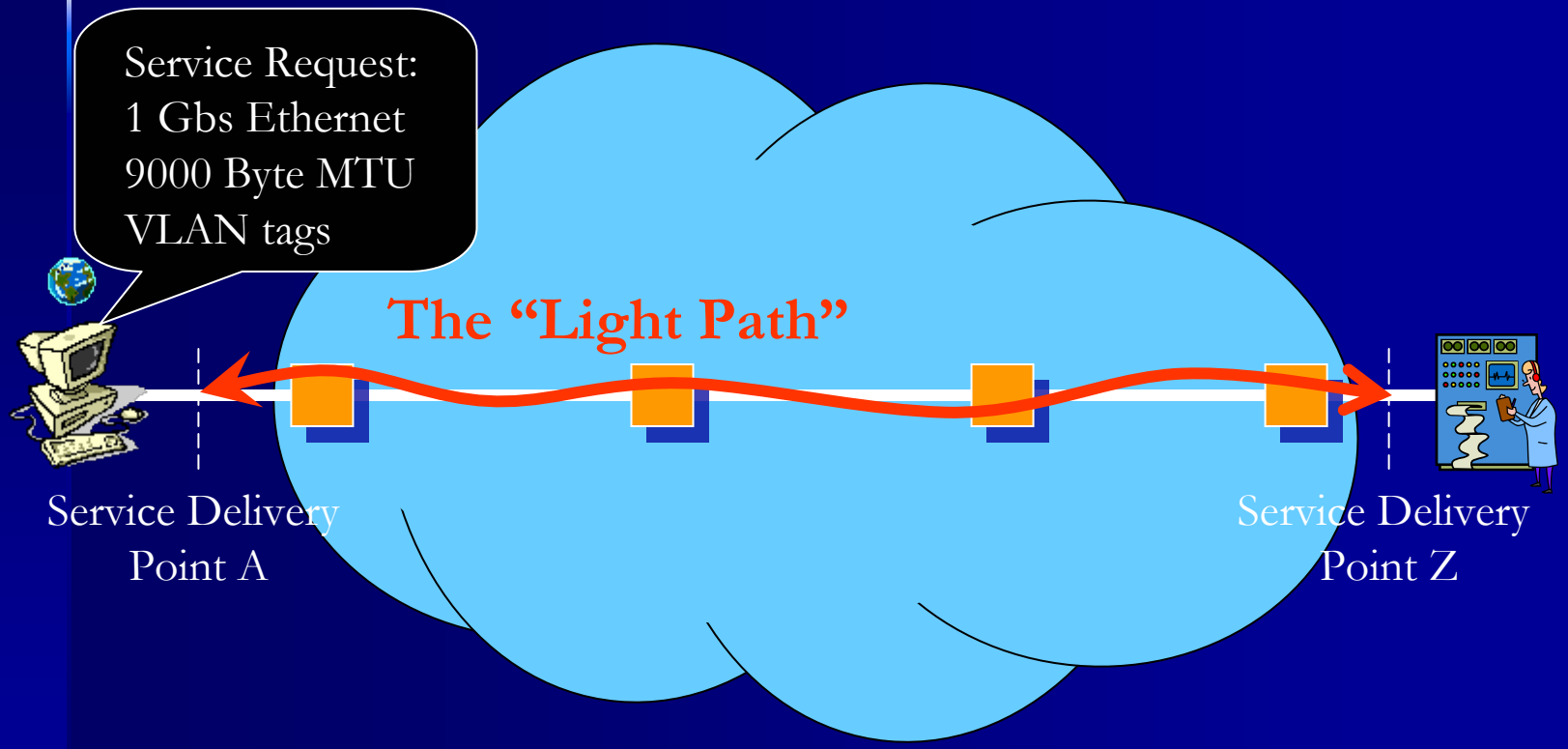
Define the Service

- **The "Service Definition" specifies the service characteristics experienced by the consumer.**
  - It does not specify *how* the service is engineered or provisioned, it simply specifies what is delivered to the end user. How the service provider decides to construct the service infrastructure is not part of the Service Definition.
- A well defined service should be ***predictable*** :
  - The user knows in advance what to expect of the service in terms of performance or other characteristics
- A well defined service is ***verifiable*** :
  - It can be measured at the service delivery points and found to conform to some set of prescribed service characteristics
- A well defined service is ***repeatable*** :
  - It behaves the exact same way every time it is invoked
- A well defined service is ***end to end*** :
  - It should not matter which provider(s) or which networks are involved in delivering the service for the end users – they will all result in conforming capabilities as experienced by the end user.

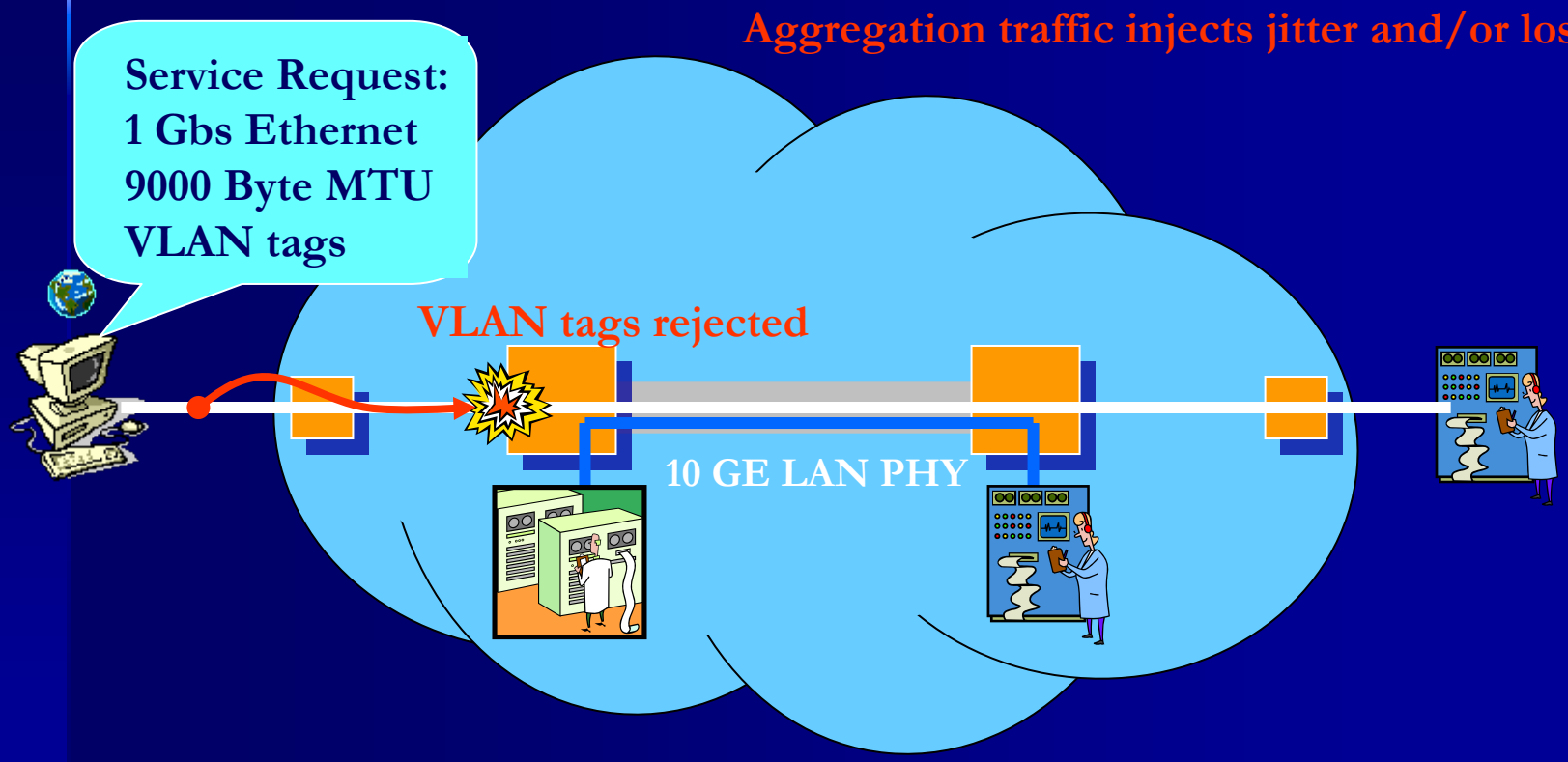
Why the R&E networks must formally define our services…

- **Users need to know exactly what they can expect from these new services**
- **Network engineers need to know exactly what capabilities they need to design and deploy within their networks – and how to interoperate with their peers**
- R&E networks represents the sum of many different networks around the world, employing many different types of hardware, with varying capabilities and service models.
  - In order to deliver wavelength services across these heterogeneous and globally distributed facilities, we ***must*** be able to engineer and concatenate the services across any combination of these networks into *deterministic* data paths that are
    - Predictable
    - Verifiable
    - Repeatable
    - End to end

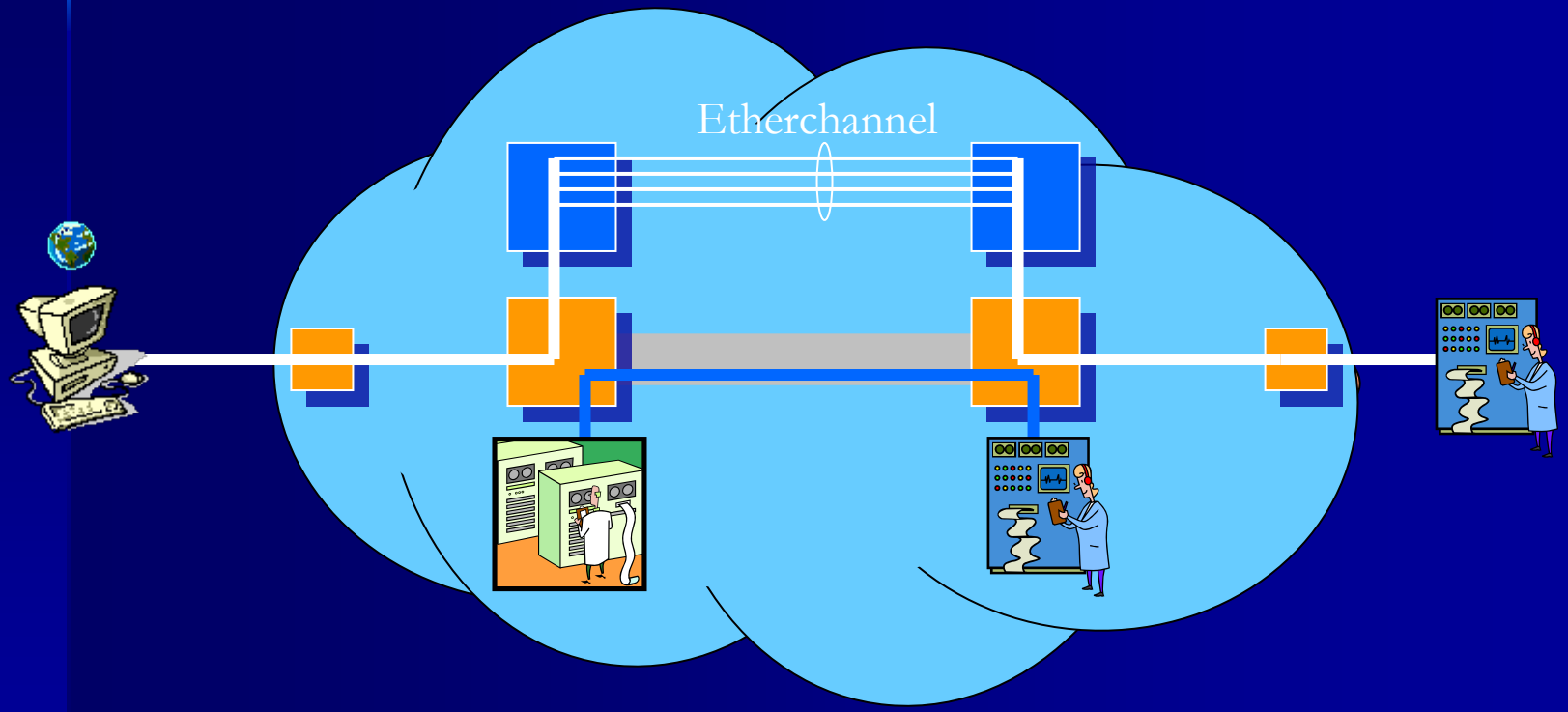
An Ethernet Light Path (on Tuesday)



An Ethernet Light Path (on Thursday)



An Ethernet Light Path (on Friday)



Examples of [bad] assumptions: “Ethernet Light Path”

- **An Ethernet Light Path has only two end points (i.e. only two sources of traffic)**
  - What happens if it does not? (e.g. VLAN linking, broadcast storms, etc.)
- **Large MTUs are supported on all big fat pipes**
  - Not a safe assumption (even today) – often limited to 1500 Byte MTUs
- **“10Gigabit” means 10,000,000,000 bits/sec …always…end to end**
  - Transoceanic links only support 9.4 Gbs Sonet/SDH WANPHY
- **A point to point ethernet Light Path will not drop packets**
  - Any LightPath aggregated with other traffic is subject to queuing…and by implication may in certain conditions exceed available buffering for those queues. And ethernet switching particularly depends upon buffering which may be overrun by bursty traffic…
  - Does the path infrastructure support flow control?
- **Ethernet frames will be delivered in order and transparently**
  - VLAN tags are often not carried end to end. Intermediate Ethernet infrastructure often uses VLANs themselves for provisioning Light Paths through the network
  - Bonded etherchannel configurations reorder packets and/or constrain bandwidth
  - Is Spanning Tree supported? Is it intercepted by transit switches?
- **An ethernet light path does not introduce jitter**
  - Ethernet aggregation almost always causes queueing and thereby introduces jitter – which can be compounded by multi-hop paths with raw aggregation provisioning…
  - Flow control/shaping introduces jitter end-to-end

Moving from 1000 emails to 1000 milliseconds…

- Current processes for creating end-to-end light paths is fraught with extreme communication challenges between network operators, and between operators and applications developers/scientists. Service instantiations are slow and often variable.
- Early adopter tools for service establishment and agreement on service definitions
- Fully automated provisioning and a wide range of service capabilities…in seconds (or less)

Service Definitions

- **Service Definitions are arbitrary.**
  - Each Service Definition specifies a specific set of service parameters, and associated values/defaults for those parameters, that form the complete set of measurable service characteristics.
    - E.g. MTU <= 9000 Bytes; or
    - SpanningTreeProtocol := notTransported;
  - Any characteristic of a service that is not explicitly defined in the service definition is explicitly undefined
    - i.e. the user can neither expect it to be present nor expect it to be absent.
- **Service Definitions are flexible**
  - They can be modified, augmented, enhanced, refined, etc from time to time as the community sees fit.
    - In the near term, the GLIF should adopt several basic broad and inclusive definitions that allow the process of define-deploy-review-refine to begin.
- **There may be multiple Service Definitions** (e.g. Ethernet, Sonet, Infiniband, Packet)
  - Differentiating services is somewhat arbitrary, but should reflect fundamentally different network transport capabilities.
  - It is conceivable that some services may be subsets of others, or offer a base set of characteristics inherited by an entire family of service dialects.

An Experiment: Defining HOPI Services

- **In an attempt to eat our own dog food…**
  Can we define the service(s) deployed over the Internet2 HOPI infrastucture?
  - **HOPI provides layer2 Ethernet connections today**
  - **We are exploring ways to integrate this with Abilene's packet/MPLS capabilities**
  - **We anticipate TDM sonet/sdh capabilities**

- **So several engineers from the HOPI project, DRAGON, and other various ad hoc discussion have begun drafting three HOPI Service Definitions:**

  - **Ethernet LightPath**– Transport of high capacity flows, and transparent to [most] layer2 services.
  - **TDM LightPath** – Sonet/SDH service. Deterministic latency/jitter transport, intercontinental and commercial interworking
  - **Packet LSP** – IP/MPLS service to complement the other service definitions.

Ethernet LightPath v1.1

- **Service_Name** := Ethernet_LightPath;
- **Version** := 2006.01.10_1.1beta;
- **Description** := The “Ethernet_LightPath” service is a point to point service that is to be used by large capacity users with certain performance criteria. ;
- **Access mode** :=
  { 1GE | 1GE_tagged | 10GE | 10GE_tagged };
- **Data Rate** := 1 Mbs to 10 Gbs by 1 Mbs;
  // 1Mbs granularity up to 10 gigabits/sec
- **Pacing Window** := 1 second; // used to constrain buf req
- **Framing** := IEEE_802.3; // plain ol’ ethernet
- **Frame Size** <= 9252 Bytes ;
  // Force10 E600 limit for MTU
- **Frame Loss Rate** <= 1E-8;
  // Assumes BER=1E-12 and 9000B datagrams for approximately 1 frame/100 million loss rate. Note that this assumes zero loss due to congestion.
- **VLAN_transport** = False ;
  // HOPI does not support VLAN stacking [yet]

Ethernet service parameters

- Data rate
  - The amount of user payload transported
  - Stated in bits-per-second reflecting the total user data to be transported / expected lifetime of the light path
- Pacing window
  - The data rate x pacing window = maximum burst size (the total amount of user data that can be presented to the network within a moving window of time=pacing window.)
  - In some sense this is the duty cycle

TDM LightPath 1.0beta

- **Service_Name** := TDMBasic
- **Version** := 2006.01.10_v1.1beta
- **Description** := The TDMBasic service is intended to be a Sonet/SDH point to point capability compatible with Next Gen Sonet/SDH features.
- **Framing** := ITU G.707 | G.708 ; // Sonet SDH framing
  // Any multiple STS1
- **Access mode** = { FE_GFP-F, GE_GFP-F, TGE_GFP-F,
  OC48, OC192 } // We can accept ethernet via GFP-F or native Sonet/SDH
- **Frame Rate** = 1 … 192; // any multiple of STS1 corresponds to 51.840 Mbs increments… might be 155mbs increments
- **Frame Loss Rate** <= 1E-10 frames/sec;
  // Assumes BER=1E-12 and user datagram of 9000B and then some…(think of Ethernet over sonet combined FLR)

PacketLSP 1.0 beta

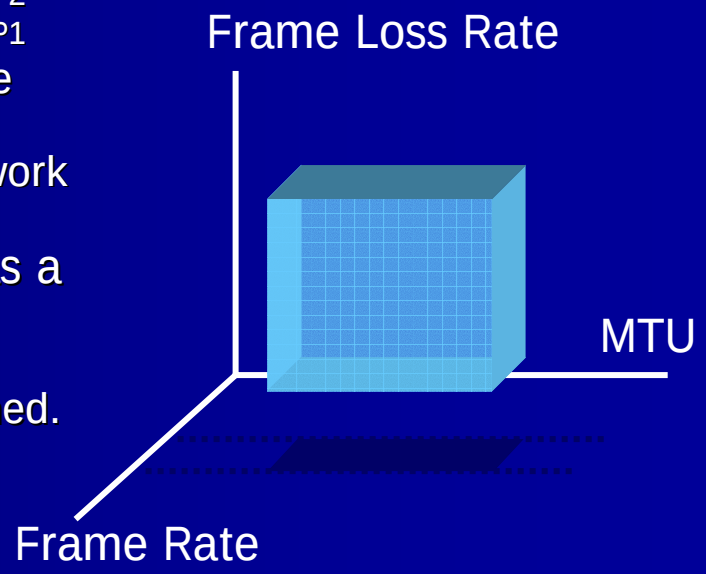
- **Service_Name** := PacketLSP;
- **Version** := 1.0beta 2005.09.30
- **Description** := Packet LSPs are intended for relatively small flows that can be handled in large numbers over an MPLS/IP backbone, or for larger flows (~1 Gbs) required to access or egress other lower layer Light Path infrastructure.
- **Data Rate** := 1 Mbs to 1 Gbs by 1 Mbs;
- **Packet_Sequence** := Stable;
- **Framing** = IPv4 | IPv6 | MPLS ;
- **Frame Loss Rate** <= 1E-8;
  // assumes ~BER 1E-12 (@10Gbs) and 9KB datagrams

LightWaveBasic 1.0beta

- **Service_Name** := LightWaveBasic 1.0beta
- **Version** := 0.1 alpha 2005.09.30
- **Description** := This service is a true ITU compliant photonic wave transport. This service is framing agnostic but assumes certain basic encoding and modulation techniques that allow for engineering of service infrastructure. This service does not do wave translation or regeneration.
- **Wavelength_Spacing** := ITU_100GHz
- **Modulation_Rate** := 100Mhz … 13 Ghz;

Inter-domain Service Matching

- Since peering networks may have slightly different service definitions, there is a need to “compare” service requests for compatibility:
- For instance:
  - NSP1 offers ethernet with 9258B mtu to NSP2
  - NSP2 offers ethernet with 1500B mtu to NSP1
  - A service request for 1400 Byte mtu can be provisioned either direction…
  - But a service request for 4000 B mtu will work across NSP1 only.
- A Service Definition can be conceived of as a multi-dimensional volume in n-space.
  - Service requests that project inside this n-dimensional service space can be provisioned. Service requests that lie outside cannot be established.



Cautions, Caveats, and Open Issues…

- We don't want to get lost in the details
  - Think globally, but act locally – i.e. Common Service Definitions should allow us to deploy new infrastructure and services within our own regions and know it will integrate and interoperate with the service environments deployed globally.
  - So we should strive to keep the service definitions as simple as possible and as high level as possible. We aren't a standards body…
  - We need to find the common interoperable service dialects…
- For the services where asynchronous framing is used (Ethernet, packets, etc) there are engineering and provisioning constraints that must be refined:
  - Any aggregation of async lightpaths is subject to burst transients that overrun buffers (and create packet loss). So specification of allowable burst characteristics (sampling period and duty cycle) is necessary to guarranty frame/packet loss rates.
- Other services/capabilties need defining:
  - Wavelength light path – i.e. framing agnostic, ITU compliant, digital modulation to some rate…
  - OTU2 capabilities
- The service definitions discussed in this presentation are examples only – more focused discussion is needed and broader involvement of the lightpath netwokring community is neceaary.

Cautions, Caveats, and Open Issues…

- **A Common Service Definition says *nothing* about how the services are or should be architected, engineered, or provisioned inside a network domain**
  - The service provisioning between the end points may take a number of different paths over potentially many different technologies (e.g. Ethernet over GFP encapsulation over a SDH link…) Indeed, the intermediate networks may request lower layer Light Paths between themselves to support upper layer Light Path service requests from other users.
  - This engineering/provisioning is left to the individual service domains to implement as they see fit – as long as the service characteristics are maintained end-to-end.
- **Where CSD parameter offers the user a choice or range, the user must explicitly select one (or alternatively, the GLIF community can adopt an explicit default value.)**
  - This is important in that the first hop provider may need to assert the request downstream to other providers who may have a different default.
- **Looking forward, the GLIF Common Services Definitions and the iterative process of refining them should enable automated service routing techniques, scheduling, authorization/accounting, etc.**

Notes and Issues on Provisioning

- The user should only have to ask their first hop service provider to establish a service instance.
  - The provisioning process downstream should be opaque to the end user so that the NSP has full flexibility (within the scope of explicitly specified service parameters) to negotiate and fulfill the service request end-to-end.
  - This downstream service routing process is a Hard Thing and an open research and experimental networking topic. Keeping this process opaque to the user will allow the GLIF the greatest flexibility to implement and evaluate many different management models and user interfaces.
- Actual Service Definitions are being developed in XML format (rather than BNF) so that they can be readily integrated into web services and other applications..

Thanks!

- The “Common Services Definition” white paper (Sobieski/Lehman April 05) can be found at:

  dragon.maxgigapop.net

  Or contact

  Jerry Sobieski (MAX)

  or

  Tom Lehman (ISI East)