Dynamic Resource Allocation via GMPLS Optical Networks

Overview and Status Presentation at
ONT3, Sept 7 2006 Tokyo, Japan

- **Jerry Sobieski**
  Mid-Atlantic Crossroads (MAX)
- **Tom Lehman**
  University of Southern California
  Information Sciences Institute (USC ISI)
- **Bijan Jabbari**
  George Mason University (GMU)
- **Don Riley**
  University of Maryland (UMD)

National Science Foundation

Outline

- DRAGON Project Overview
- DRAGON Network Status
- DRAGON Control Plane Status
- Next Steps: Key Focus Areas

Project Overview
Single Slide Overview

- Principal Investigators
  - Jerry Sobieski - Mid-Atlantic Crossroads (MAX)
  - Tom Lehman - USC/ Information Sciences Institute (ISI East)
  - Bijan Jabbari - George Mason University (GMU)
  - Don Riley – University of Maryland
- Commercial Partner – MOVAZ Networks
- NSF Funded program
- All Optical Metropolitan Area Networking
  - Testbed deployed in the Washington DC region
- GMPLS based control plane
  - Dynamic provisioning across heterogeneous network technologies
  - Fiber (FSC), Lambda (LSC), SONET (TDM), Ethernet (L2SC), Packet (LSC)
  - Multi-layer Traffic Engineering
  - Open Source Software
  - Interdomain Provisioning (routing, path computation, signaling)
  - Authentication, Authorization, Accounting (AAA)
  - Scheduling
- Application Support
- http://dragon.maxgigapop.net, http://dragon.east.isi.edu

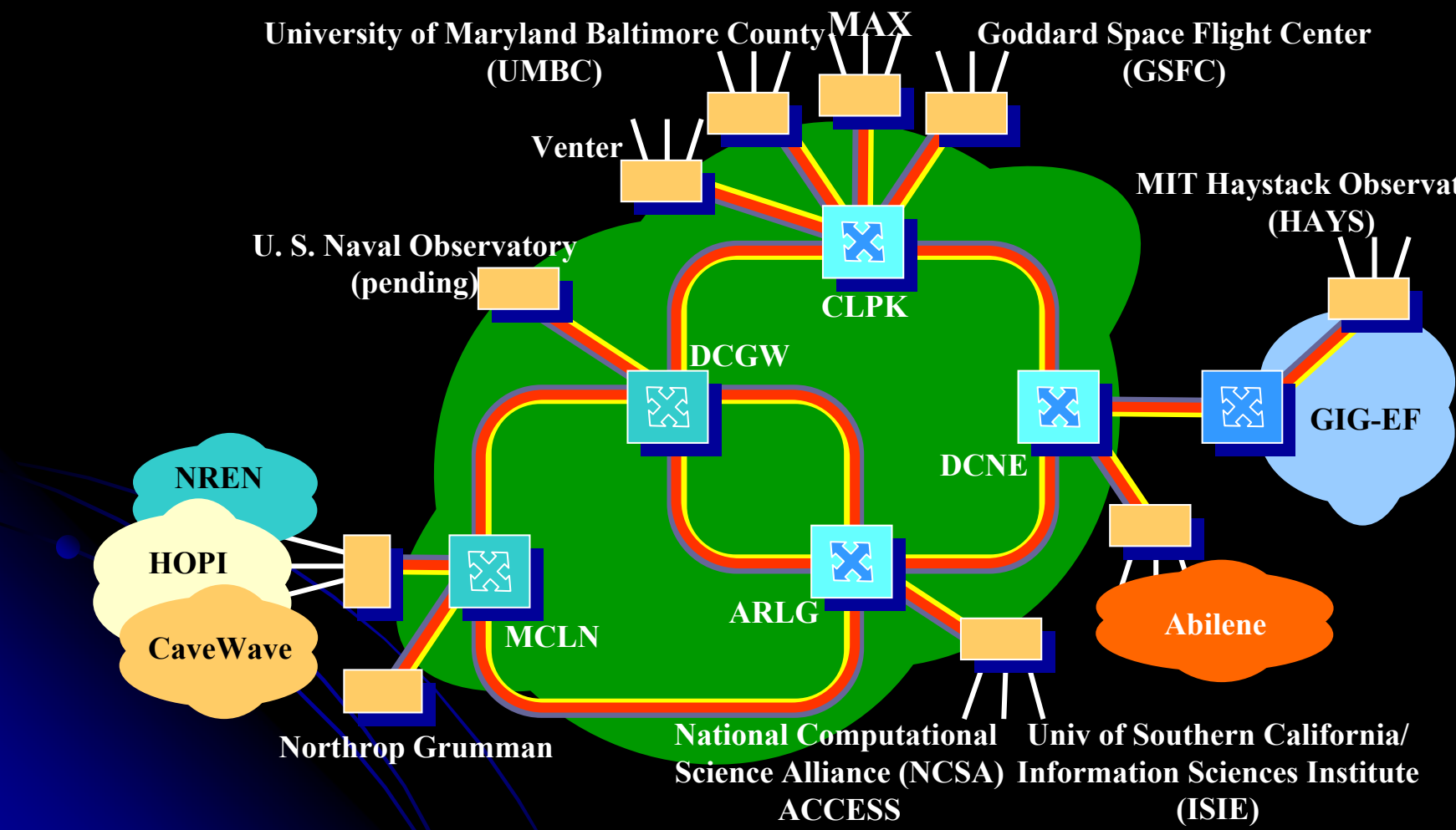
DRAGON
Initial Collaborators

- Mid-Atlantic Crossroads
- USC / Information Sciences Institute – East
- George Mason University
- University of Maryland
- Movaz Networks
- MIT Haystack Observatory
- NASA Goddard Space Flight Center
- NCSA ACCESS Center
- US Naval Observatory

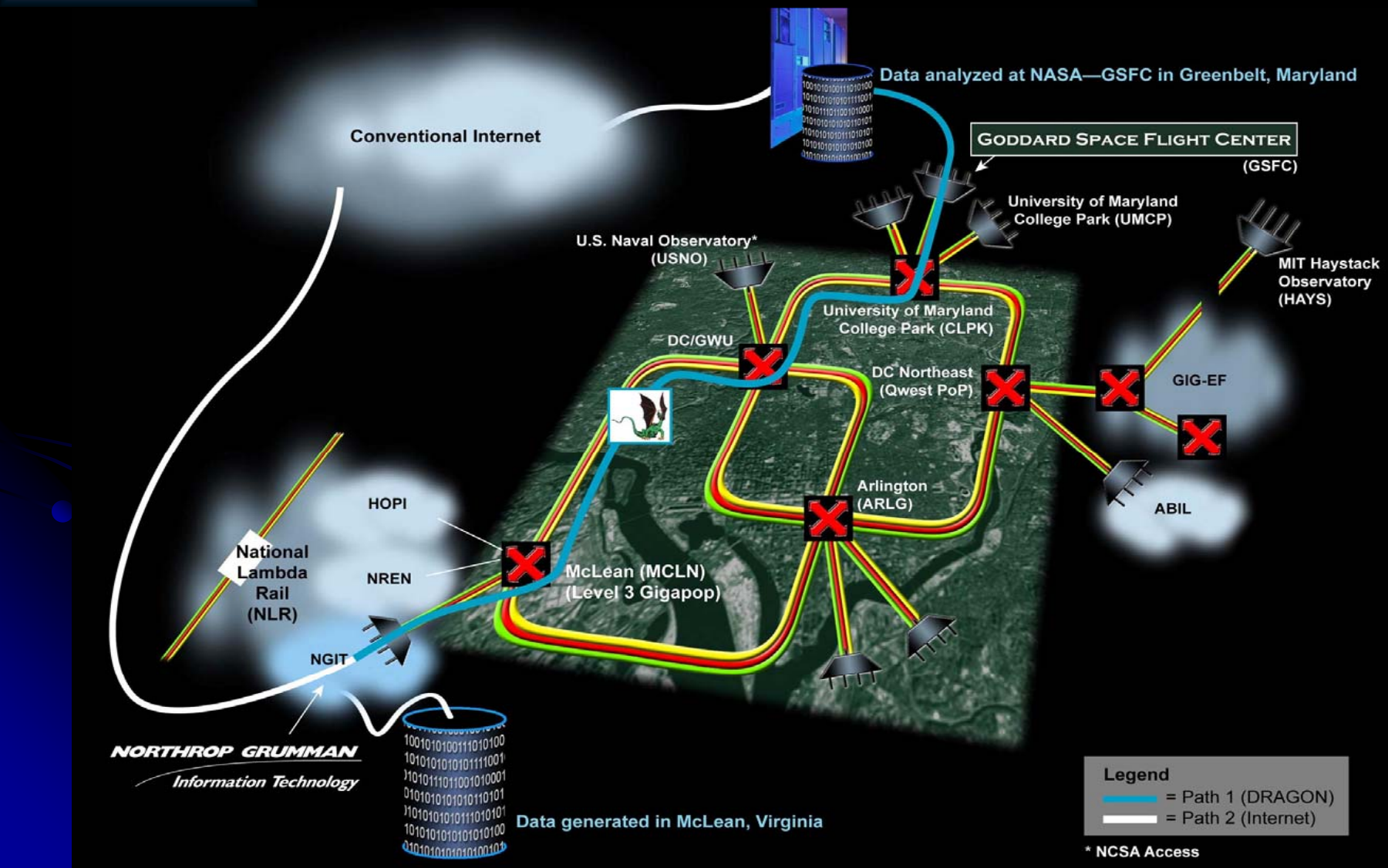
DRAGON Collaborators Today (September 2006)

- Additional Collaborators:
  - Internet2 / HOPI
  - University of Maryland Baltimore County
  - Laboratory for Telecommunication Science
  - Raptor Networks
  - Force10 Networks
  - NASA Ames
  - Northrup Grumman Corp.
  - Naval Research Lab
  - Ciena
  - Others in the works…
- International:
  - KTH Stockholm (SE)
  - Univ of Amsterdam and JIVE (NL)
  - Univ of Manchester (UK)
  - NICT/JGN2 Tokyo (JP)

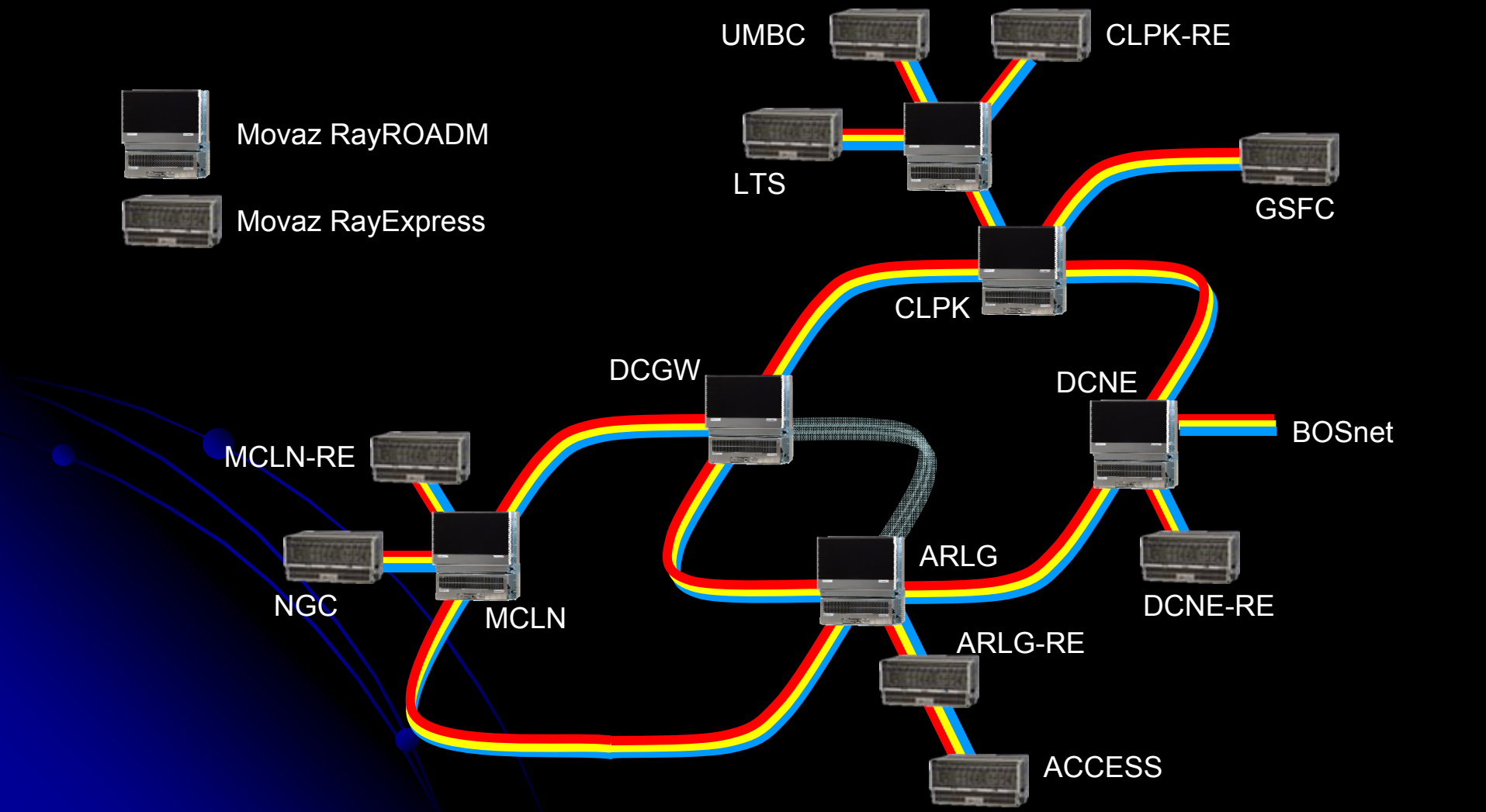
DRAGON Network



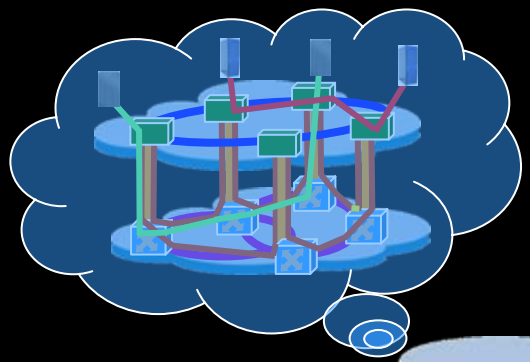
DRAGON Network



DRAGON Testbed Optical Layer



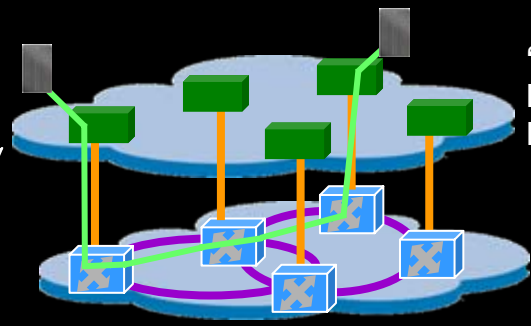
The Vision: One Infrastructure Multiple Topologies/Services



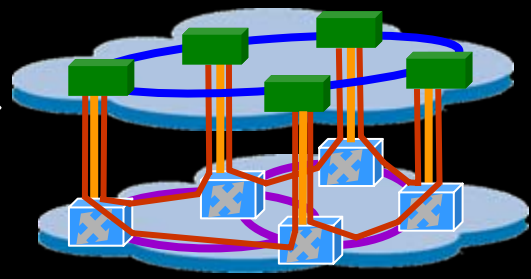
Ethernet Layer

Switched WDM Optical Layer

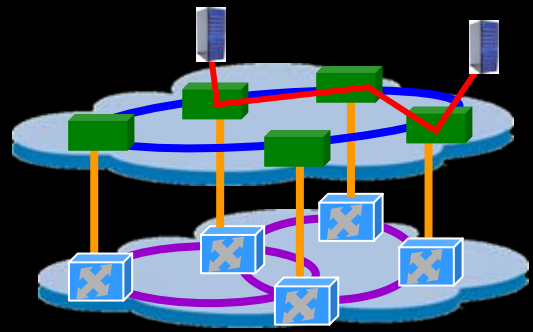
Provisioned Topologies



“ Ethernet Framed Lambda”

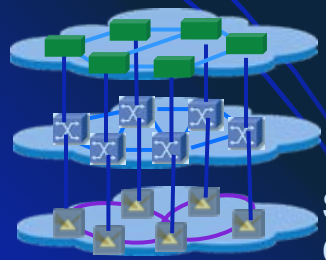


“Basic Ethernet Service”



“Dedicated VLAN Connection over Ethernet”

Multi-Layer GMPLS Networks



Ethernet Layer

Switched SONET Layer (vcat, lcas)

Switched WDM Optical Layer

DRAGON Control Plane R&E "Hybrid" Networks

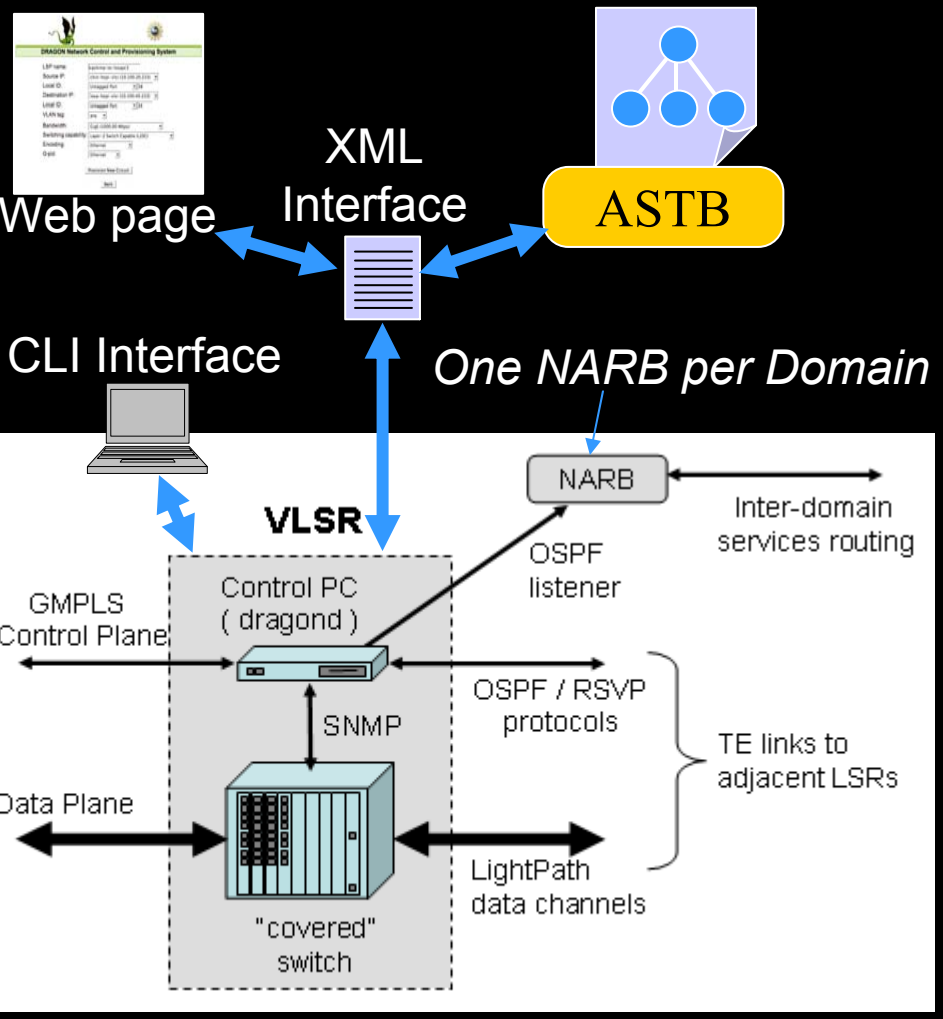
- Multi-Service, Multi-Level, Multi-Domain
- One "infrastructure" which provides basic IP routed service as well services at lower layer
  - i.e., connectionless and connection oriented services
- Services could be point to point circuits or application specific layer2 multipoint broadcast domains
- Interoperable architectures & control planes needed
- Integration challenges (control, data, management planes)
- Multi-layer adaptations "horizontal" for multi-domain
- Multi-layer adaptations "vertically" for traffic grooming
- Key control plane functions: routing, signaling, path computation
- Scheduling and AAA functions also needed
- Integration of (G)MPLS and Web Services

DRAGON Control Plane Key Components

- Network Aware Resource Broker – NARB
  - Intradomain listener, Path Computation, Interdomain Routing
- Virtual Label Swapping Router – VLSR
  - Open source protocols running on PC act as GMPLS network element (OSPF-TE, RSVP-TE)
  - Control PCs participate in protocol exchanges and provisions covered switch according to protocol events (PATH setup, PATH tear down, state query, etc)
- Client System Agent – CSA
  - End system or client software for signaling into network (UNI or peer mode)
- Application Specific Topology Builder – ASTB
  - User Interface and processing which build topologies on behalf of users
  - Topologies are a user specific configuration of multiple LSPs

VLSR
(Virtual Label Switching Router)

- GMPLS Proxy
  - (OSPF-TE, RSVP-TE)
- Local control channel
  - CLI,TL1, SNMP, others
- Used primarily for ethernet switches
- Provisioning requests via CLI, XML, or ASTB



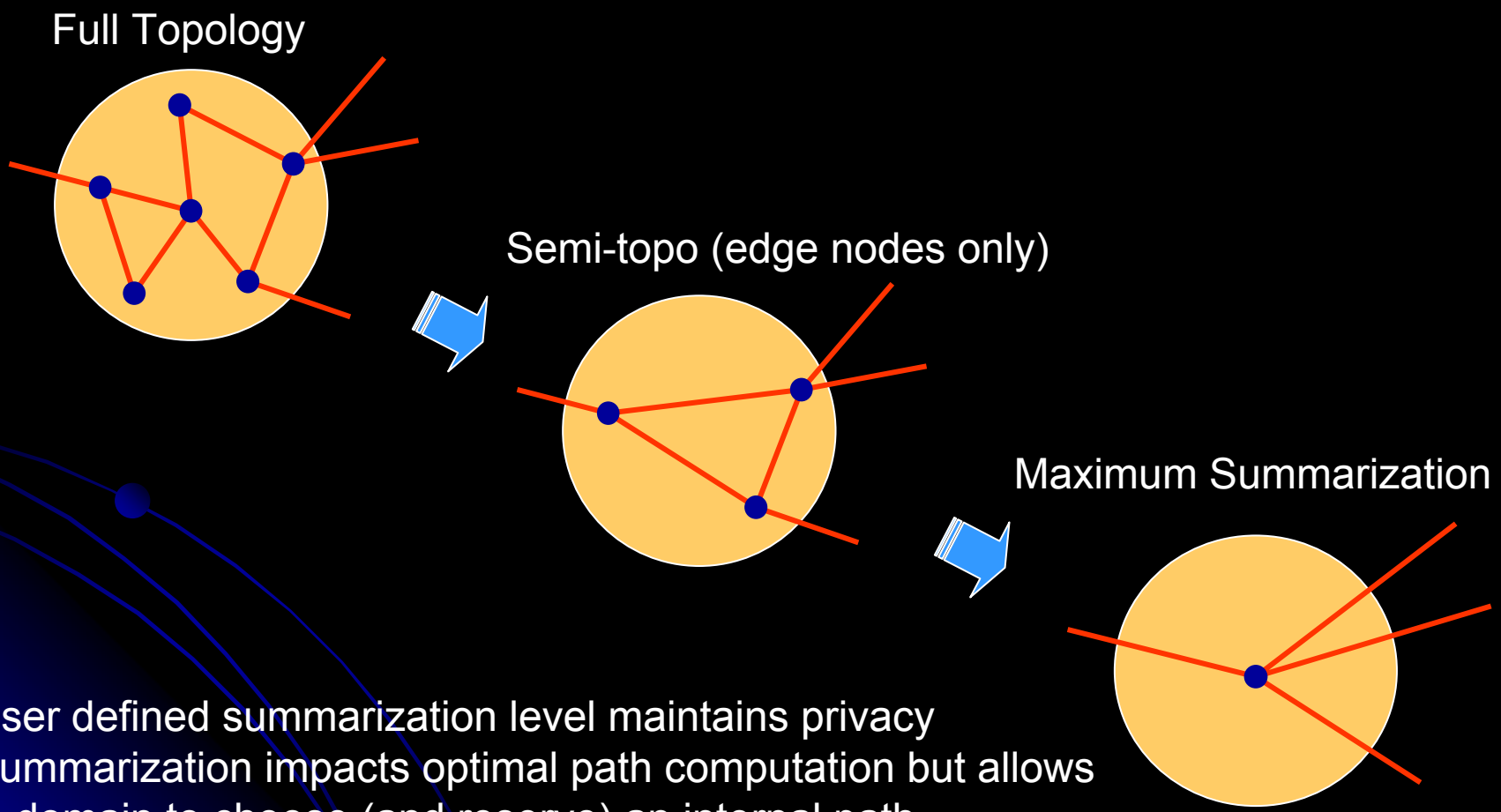
VLSR
(Virtual Label Switching Router)

- RSVP Signaling module
  - Originated from Martin Karsten's C++ KOM-RSVP
  - Extended to support RSVP-TE (RFC 3209)
  - Extended to support GMPLS (RFC 3473)
  - Extended to support Q-Bridge MIB (RFC 2674)
  - For manipulation of VLANs via SNMP (cross-connect)
  - Extended to support VLAN control through CLI
- OSPF Routing module
  - Originated from GNU Zebra
  - Extended to support OSPF-TE (RFC 3630)
  - Extended to support GMPLS (RFC 4203)
- Ethernet switches tested to date
  - Dell PowerConnect, Extreme, Intel, Raptor, Force10

NARB (Network Aware Resource Broker)

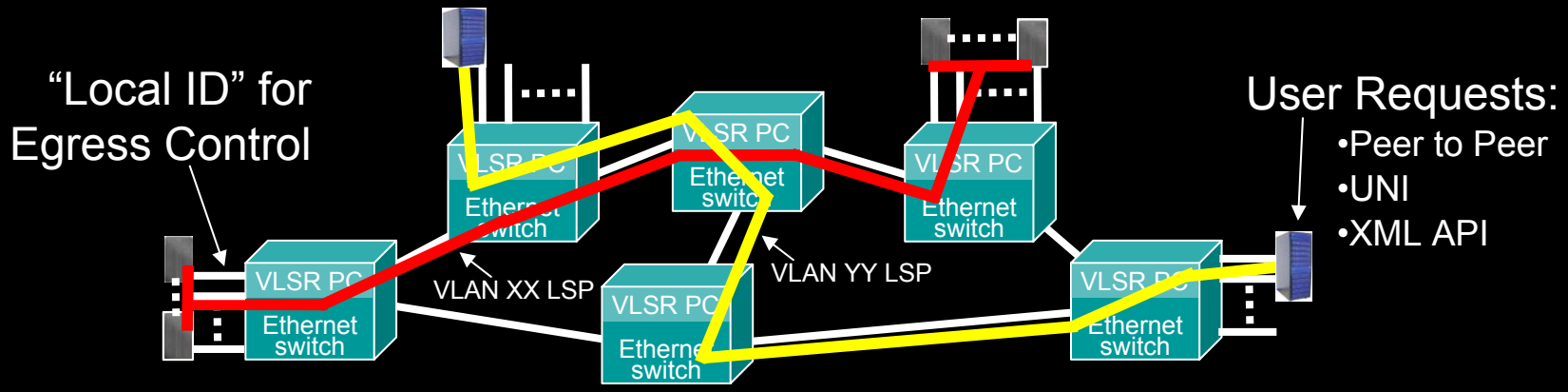
- NARB is an agent that represents a domain
- Intra-domain Listener
  - Listens to OSPF-TE to acquire intra-domain topology
  - Builds an abstracted view of internal domain topology
- Inter-domain routing
  - Peers with NARBs in adjacent domains
  - Exchanges (abstracted) topology information
  - Maintains an inter-domain link state database
- Path Computation
  - Performs intra-domain (strict hop) TE path computation
  - Performs inter-domain (loose hop) TE path computation
  - Expands loose hop specified paths as requested by domain boundary (V)LSRs.
- Hooks for incorporation of AAA and scheduling into path computation via a "3 Dimensional Resource Computation Engine (3D RCE)"
  - The Traffic Engineering DataBase (TEDB) and Constrained Shortest Path Computation (CSPF) are extended to include dimensions of GMPLS TE parameters, AAA constraints, and Scheduling constraints.
  - 3D RCE is the combination of 3D TEDB and 3D CSPF
  - http://dragon.east.isi.edu/data/dragon/documents/dragon-infocom-APBM-workshop-apr282006.pdf

Inter-Domain Topology Summarization



- User defined summarization level maintains privacy
- Summarization impacts optimal path computation but allows the domain to choose (and reserve) an internal path

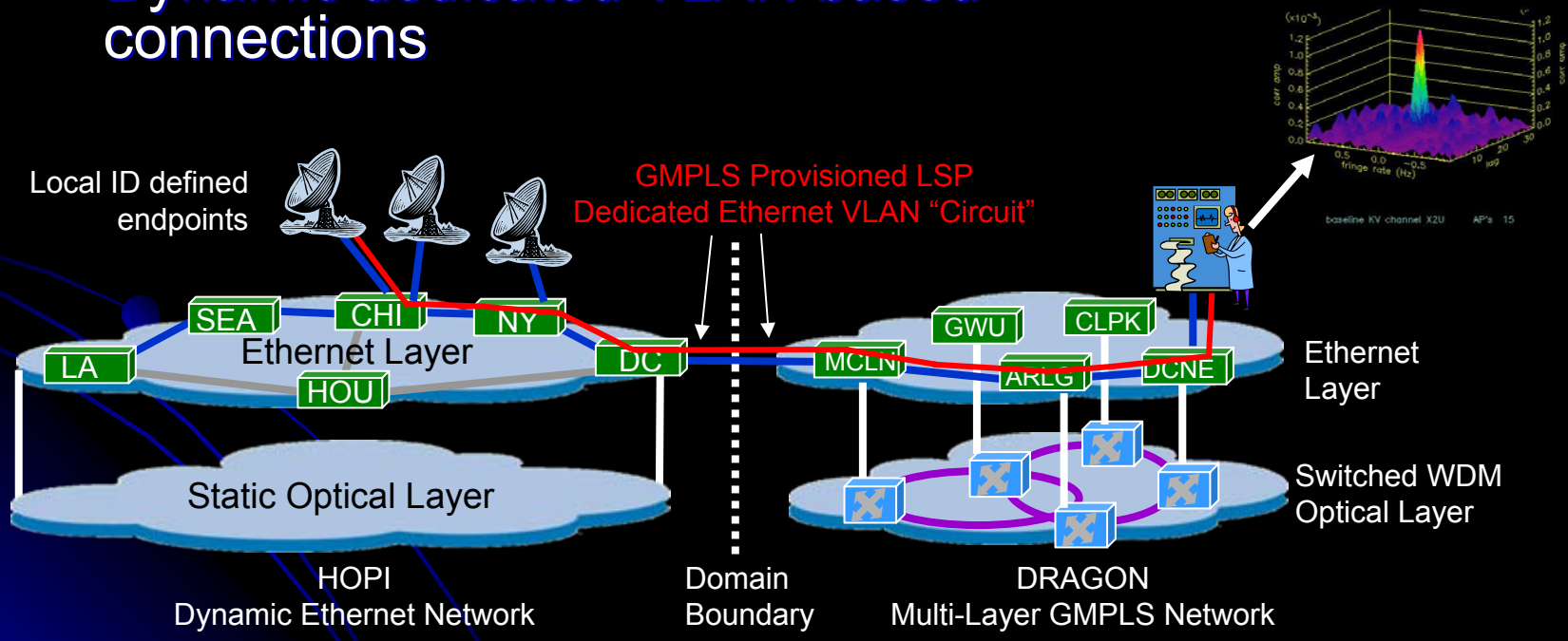
GMPLS Provisioned Ethernet Services



- Multiple Ethernet Provisioning Options
- Point to Point Ethernet VLAN based LSPs
- Ethernet switch (vendor specific) features applied to guarantee LSP bandwidth in increments of 100 Mbit/s
- Edge connection flexibility provided by use of "Local ID" feature which allows flexible combinations of one port, multiple ports, tagged ports, and untagged ports to be glued on to end of LSP. Can be dynamically adjusted.
- Users can request services via Peer to Peer GMPLS, UNI style GMPLS, or via an XML application interface
- Ethernet VLAN space is "flat" across provisioned space. Constrained based path computation utilized to find available VLAN Tags.
- VLAN tags treated in a similar manner to wavelengths

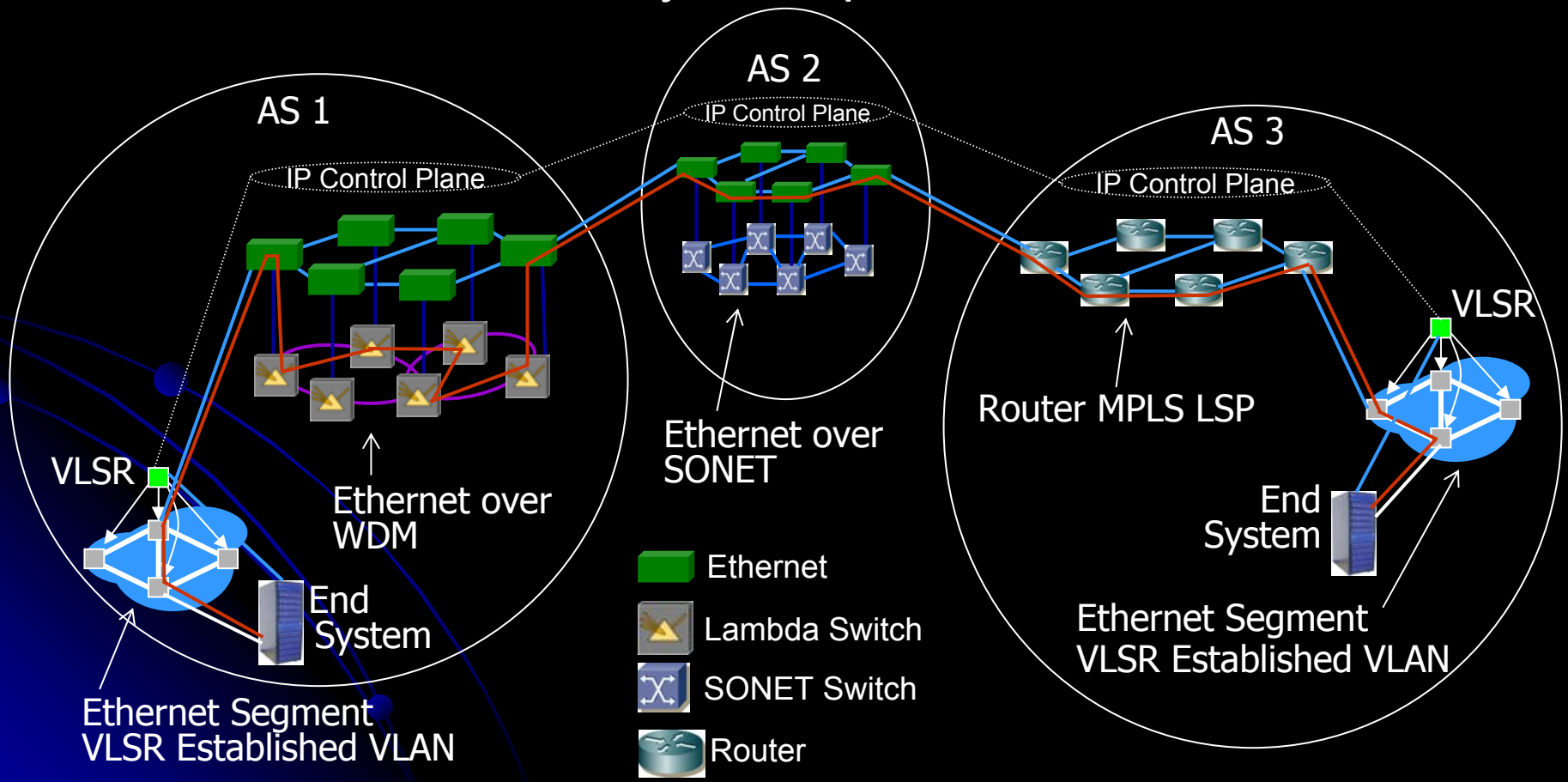
DRAGON/HOPI Control Plane Provisioning Environment

- GMPLS Multi-layer, Multi-Domain
- Ethernet Service Provisioning
- Dynamic dedicated VLAN based connections



Heterogeneous Network Technologies Complex End to End Paths

"horizontal" multi-layer adaptations for multi-domain

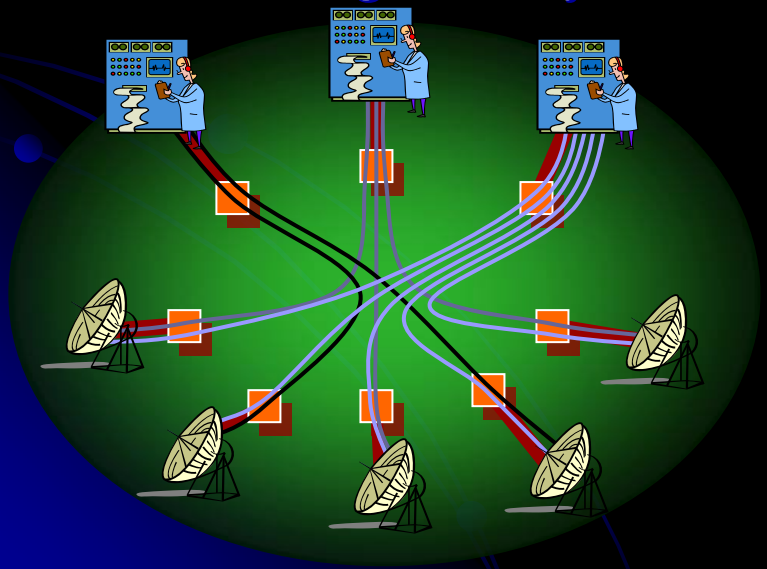


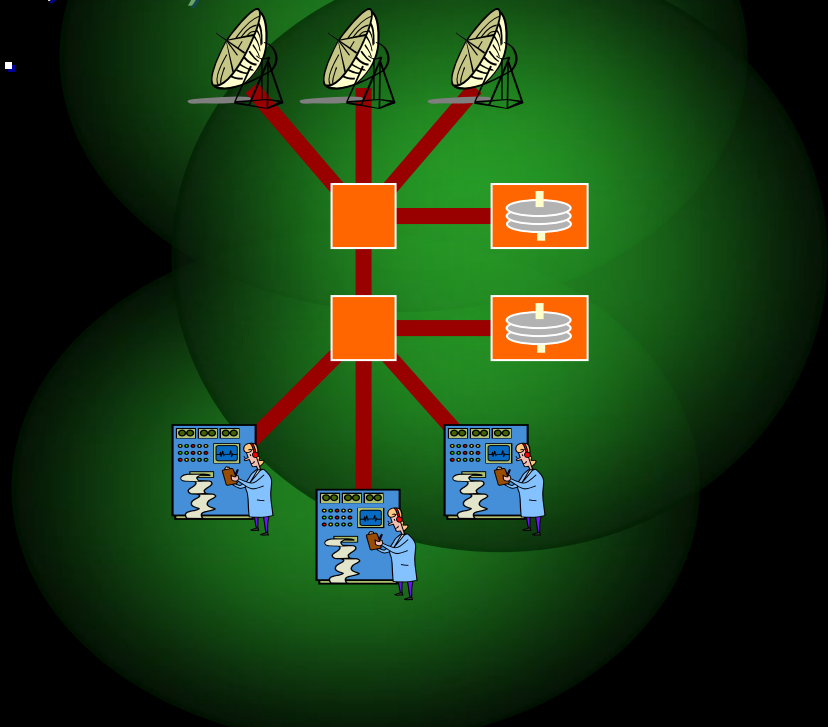
InterDomain (G)MPLS and Web Services

- Currently working on interdomain virtual circuit provisioning between:
  - ESnet
  - Abilene
  - HOPI
  - UltraScience Net
- Focusing on how to accomplish routing, signaling, path computation in a mixed (G)MPLS and Web Service environment

An “ eVLBI” Application Specific Network

- Telescopes connect to intermediate realtime storage/spooling facilities
  - These storage facilities may be a) at the telescope, b) at the correlator, or c) somewhere else logistically useful.





Application Specific Topologies using XML


```
<topology>
   <resource>
        <resource_type> eVLBI.Mark5a        </resource_type>
        <name>    Haystack.muk1             </name>
        <ip_addr> muk1.haystack.mit.edu     </ip_addr>
        <te_addr> muk1-ge0.haystack.mit.edu </te_addr>
        <appl>    /usr/local/evlbi_script   </appl>
   </resource>
   <resource>
        <resource_type> eVLBI.Mark5a        </resource_type>
        <name>    Westford1                 </name>
        <ip_addr> wstf.haystack.mit.edu     </ip_addr>
        <te_addr> wstf-ge0.haystack.mit.edu </te_addr>
        <appl>    /usr/local/evlbi_script   </appl>
   </resource>
   <resource>
        <resource_type> EtherPipeBasic      </resource_type>
        <src>           Haystack.muk1       </src>
        <dest>          Westford.muk1       </dest>
        <datarate>      1 Gbs               </datarate>
        </resource>
</topology>
```

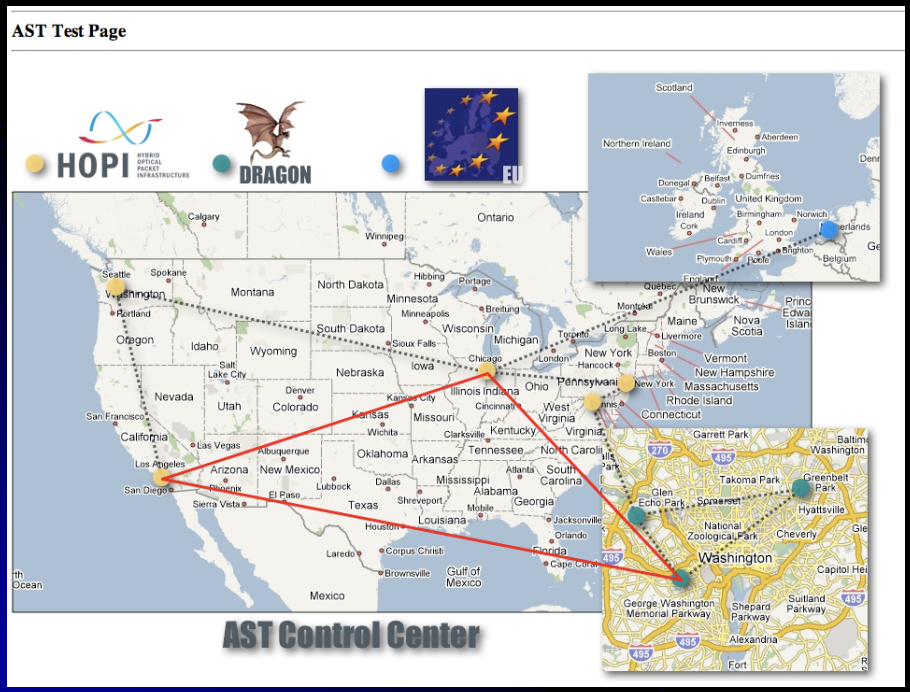
A

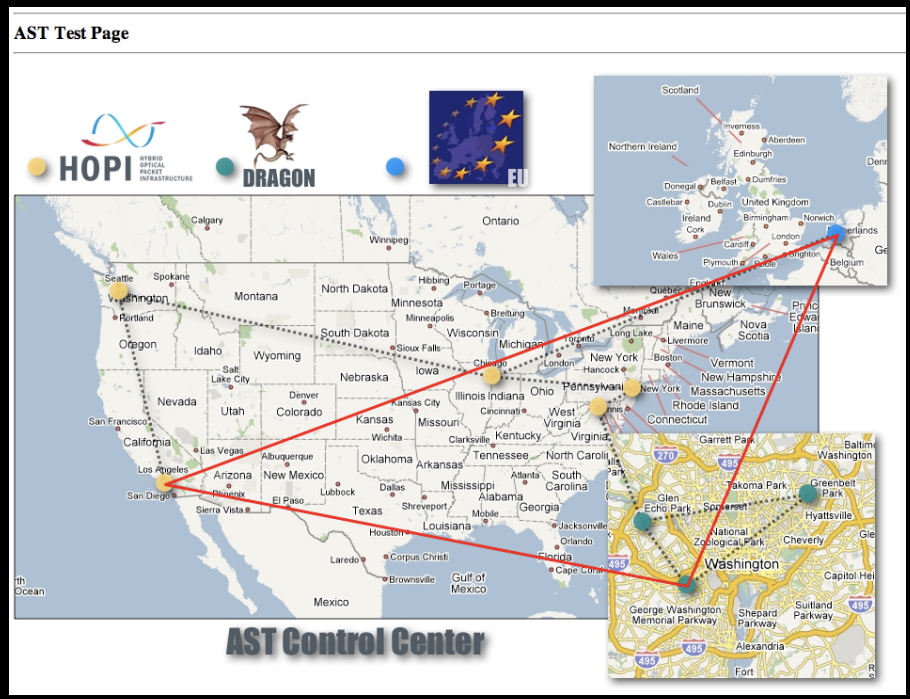
B

C

Application Specific Topologies

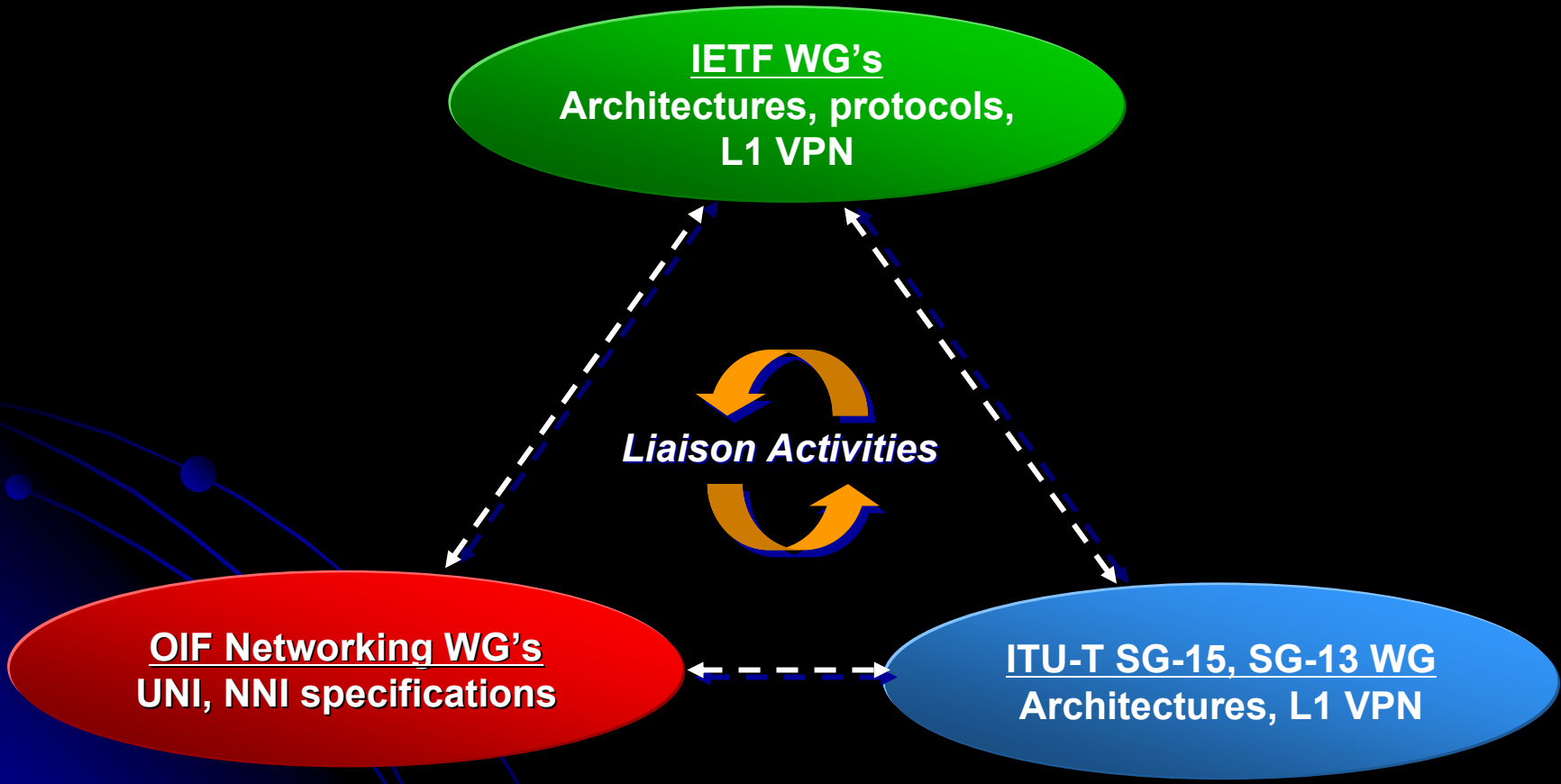
- Live demonstration at Internet2 Spring Member Meeting (April 2006, Washington DC)
  - See www.internet2.edu for webcast of “HOPI update” presentation.
- Set up global multi-link topologies
  - ~30 seconds





Standards Tracking

Multi-Layer / Multi-Domain Activities



Current DRAGON Deployment Status

- DRAGON (in Washington metro area is fully operational
  - ROADMs deployed – wave layer constantly growing and in flux (LSC working, interoperability testing in progress)
  - Multi-layer topology; Ethernet (L2SC) over Lambda (LSC)
  - VLSR deployed (ethernet VLAN based “circuits”)
  - NARB deployed (interdomain routing, path computation element)
- HOPI has deployed VLSR + NARB
  - Operational since fall 05
- Working on integration of international VLSRs
- Application support for eVLBI, HD Video Services, others

Continuing Work
Key Focus Areas

- GMPLS Control Plane
  - Inter-domain routing and signaling agreements
    - R&E community should make this a priority
  - Advanced path computation techniques
  - Inter-operability with vendor stacks
  - Multi-layer stitching
- AAA and Scheduling Control Plane Features
- Web Service based control planes
- Application Specific Topologies
  - Integration/reconciliation of AST, Network Description Language, Common Service Definition specs
  - Integration with applications

Thank You

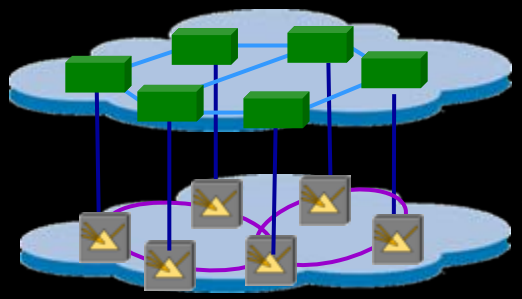
Questions/Comments?:

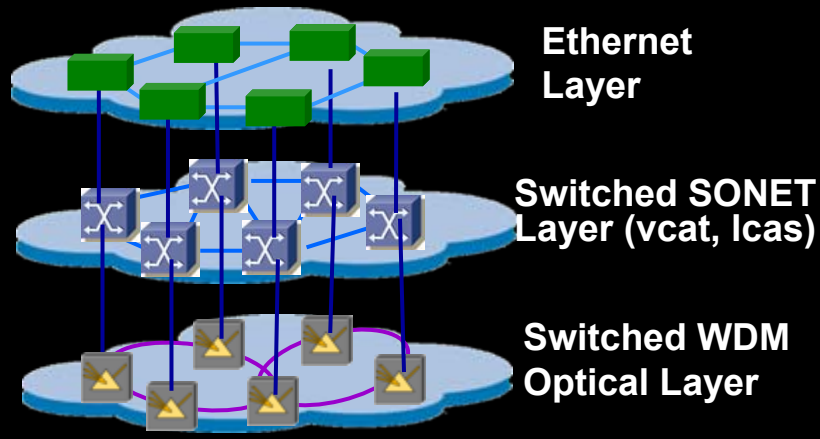
Tom Lehman
tlehman at isi.edu
http://dragon.east.isi.edu
or
Jerry Sobieski
jerrys at maxigapop.net
http://dragon.maxgigapop.net

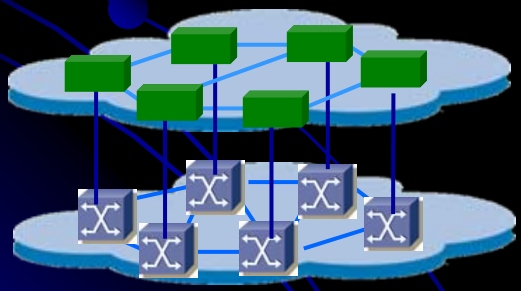
Extra Slides

Multi-Layer GMPLS Networks

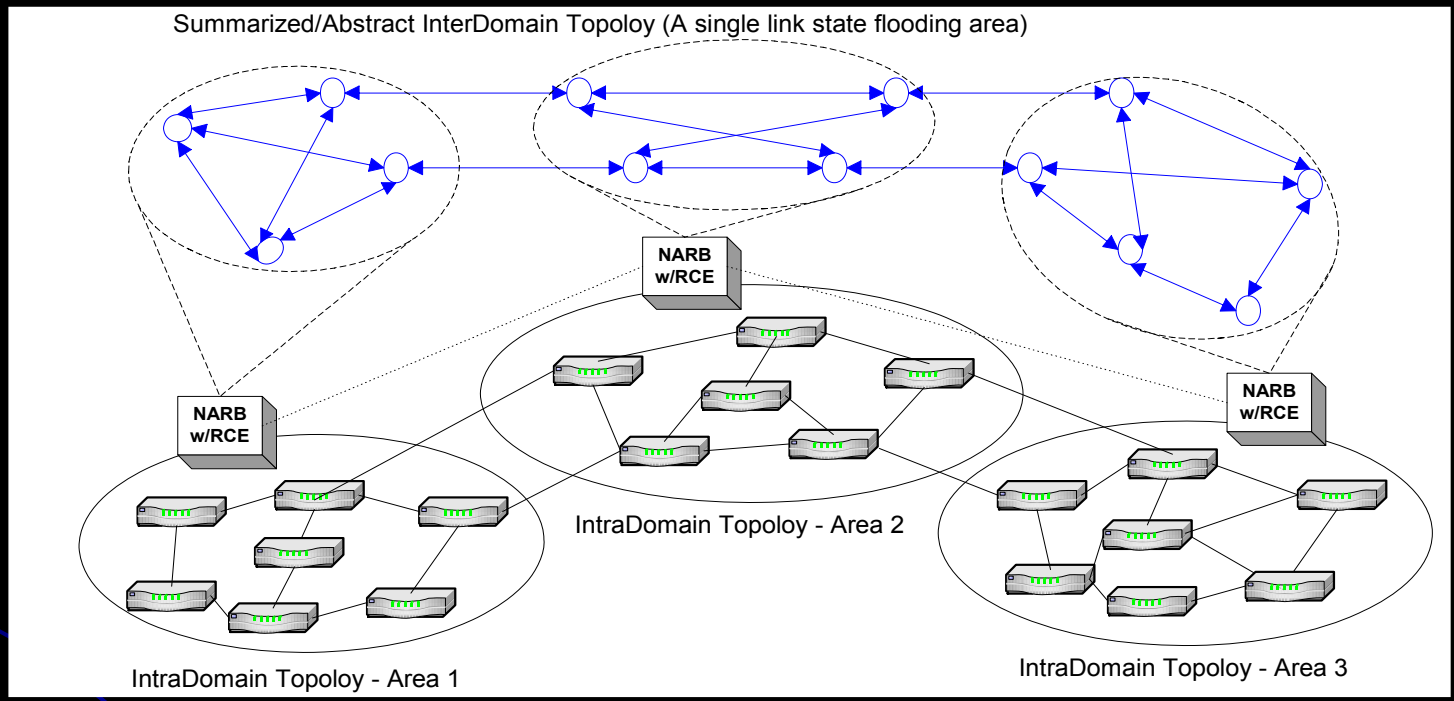
"vertical" multi-layer adaptations for traffic grooming, multiple services, multiple "virtual" networks







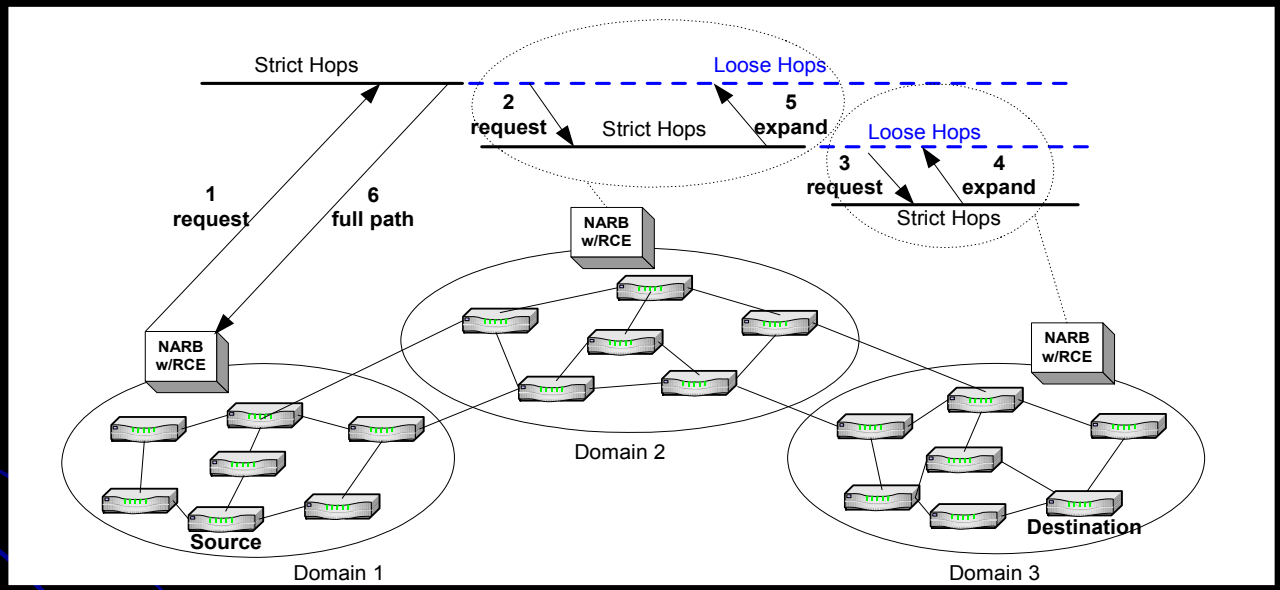
Interdomain Path Computation A Hierarchical Architecture



- NARB summarizes individual domain topology and advertise it globally using link-state routing protocol, generating an abstract topology.
- RCE computes partial paths by combining the abstract global topology and detailed local topology.
- NARB's assemble the partial paths into a full path by speaking to one another across domains.

E2E Multi-Domain Path Computation Scheme

DRAGON mainly uses Recursive Per-Domain (RPD) interdomain path computation

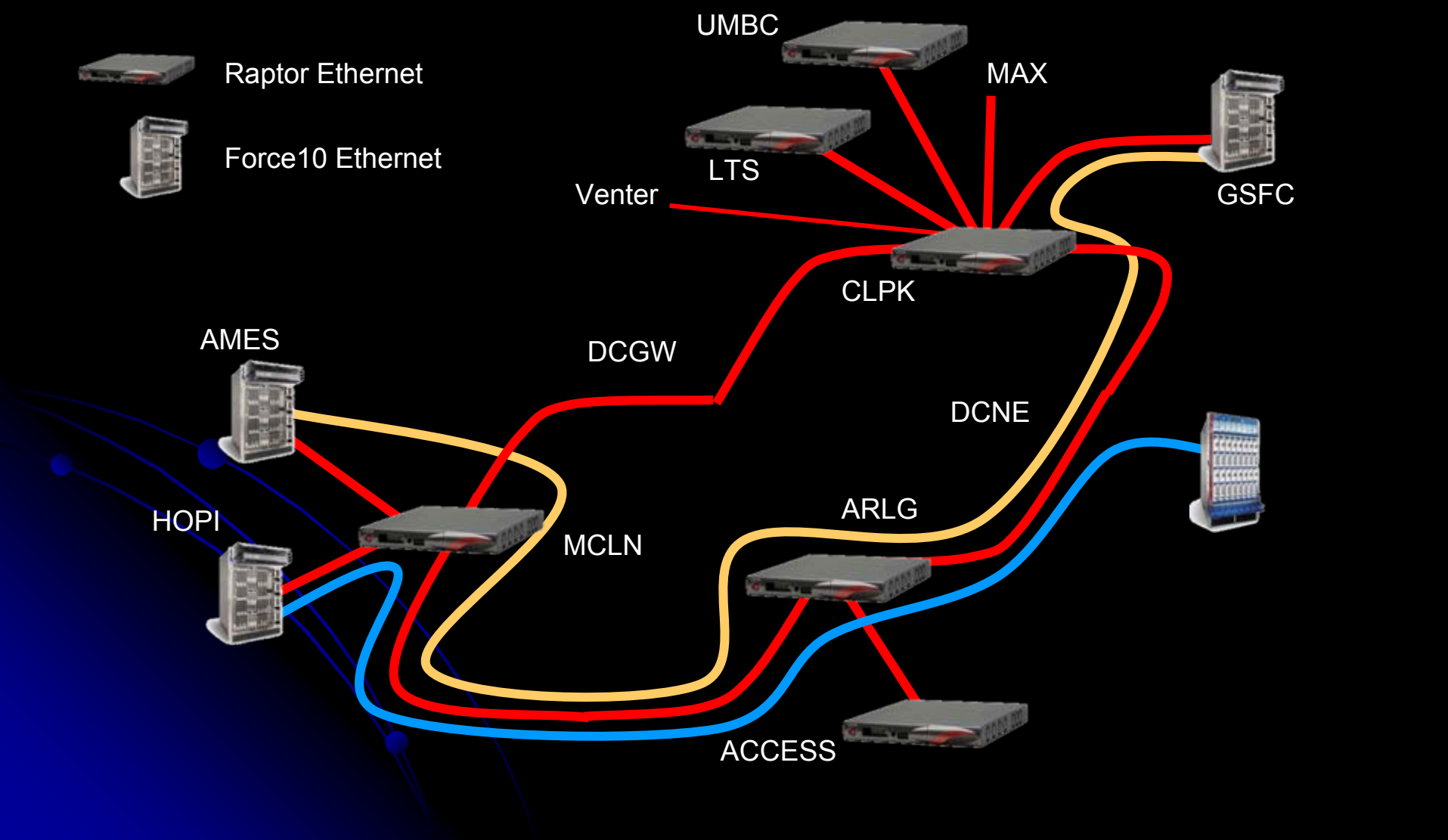


- Full explicit path is obtained before signaling.
- Other supported schemes include Centralized path computation and Forward Per-Domain (FPD) path computation.

DRAGON CSPF Path Computation Heuristics

- A breadth first search based CSPF heuristic in deployment
  - Takes flexible combination of various constraints, such as bandwidth, switch cap., wavelength, VLAN tag and add-on policy constraints.
  - Supports multi-region networks using configurable region-crossing criteria
  - Reliable results; probably time-consuming in large networks (~30ms in the 12-node HOPI+DRAGON network)
- Other heuristics under research; one is based on a channel-graph model in combination with K-shortest path routing.

DRAGON Ethernet Layer



Ethernet VLAN based Provisioning

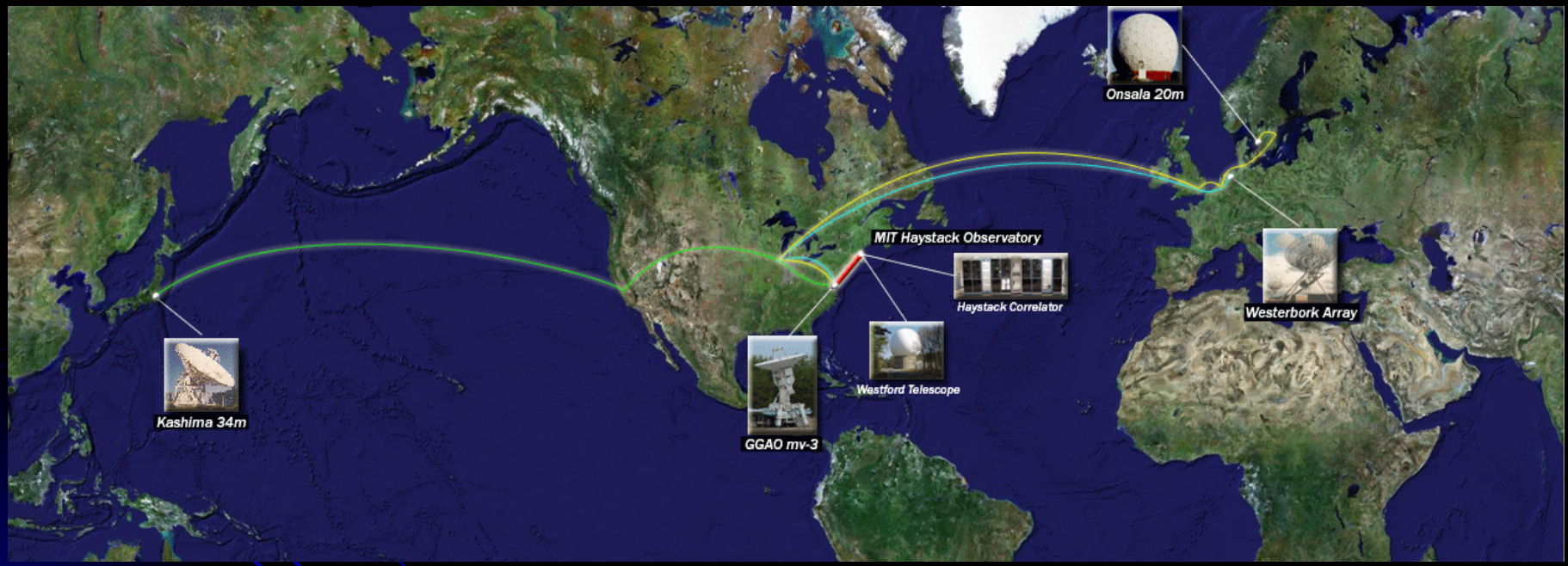
- Local ID defines the VLAN tag/edge port mapping
  - Several options; tagged, untagged, single port, port groups, automatic
  - Local ID definitions can be adjusted dynamically
- OSPF
  - configure vlans on each interface
  - advertise out in IfSwCap Descriptor TLV inside a TE Link LSA
  - update vlans availability and bandwidth in response to provisioning
  - similar to the existing ifswcap-specific-psc and ifswcap-specific-tdm
- RSVP ERO
  - proprietary Unnumbered Interface ID Subobjects (UnNumIfID) used to encode VLAN information in ERO
  - 32-bit UnNumbered Interface ID: type(1byte):value(24bits, vlan tag info)
- NARB/RCE
  - listen to OSPF
  - path computation with bandwidth and vlan constraints
  - create EROs with UnNumIFID objects
- Driven by need to provision across HOPI (10 gigabit interfaces)

Collaborations with European Research Teams

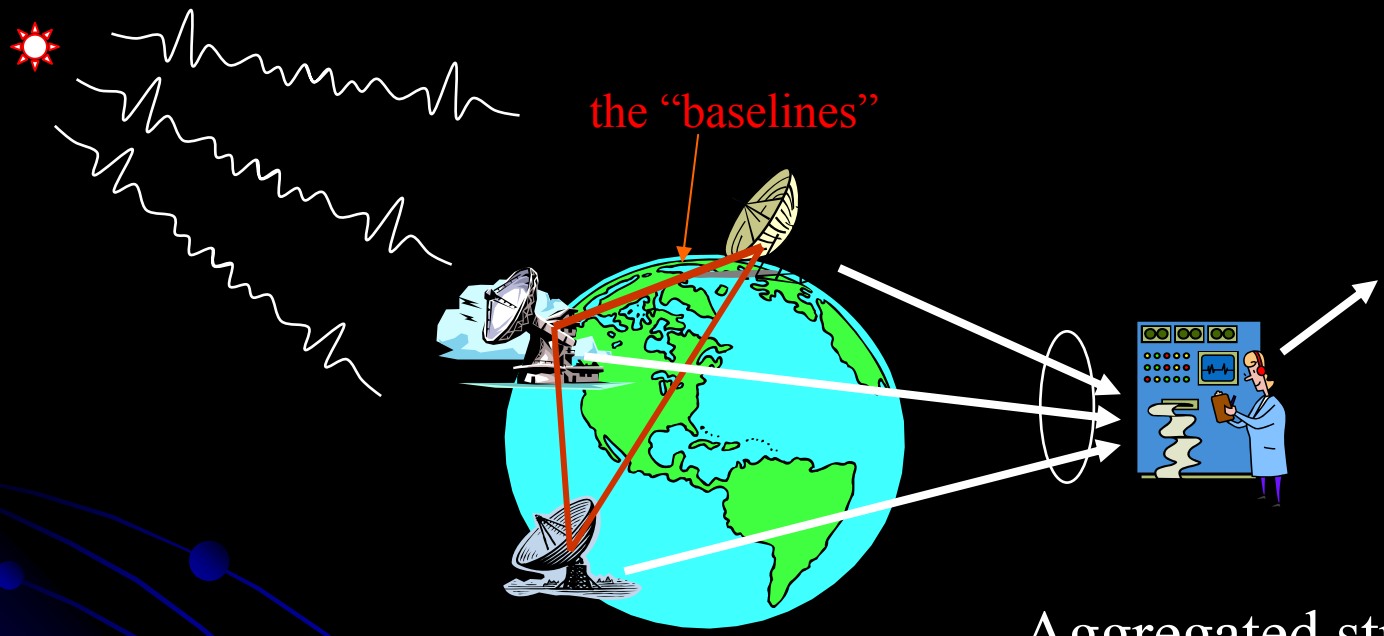
- NetherLight is hosting a VLSR + NARB in Amsterdam
  - Operational as of April 2006
  - Peers with HOPI in Chicago via transAtlantic 10G link
- NorthernLight VLSR is in place at KTH – integration with will happen over this summer
- Univ of Manchester has VLSR
- Hopefully will have a VLSR in Tokyo soon

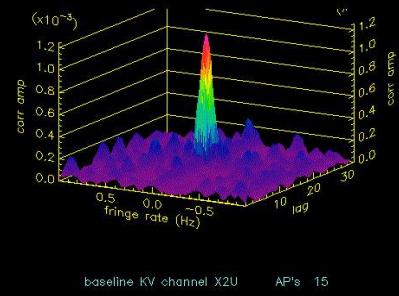
Global e-VLBI

iGrid 2005



Very Long Baseline Interferometry "E-VLBI"





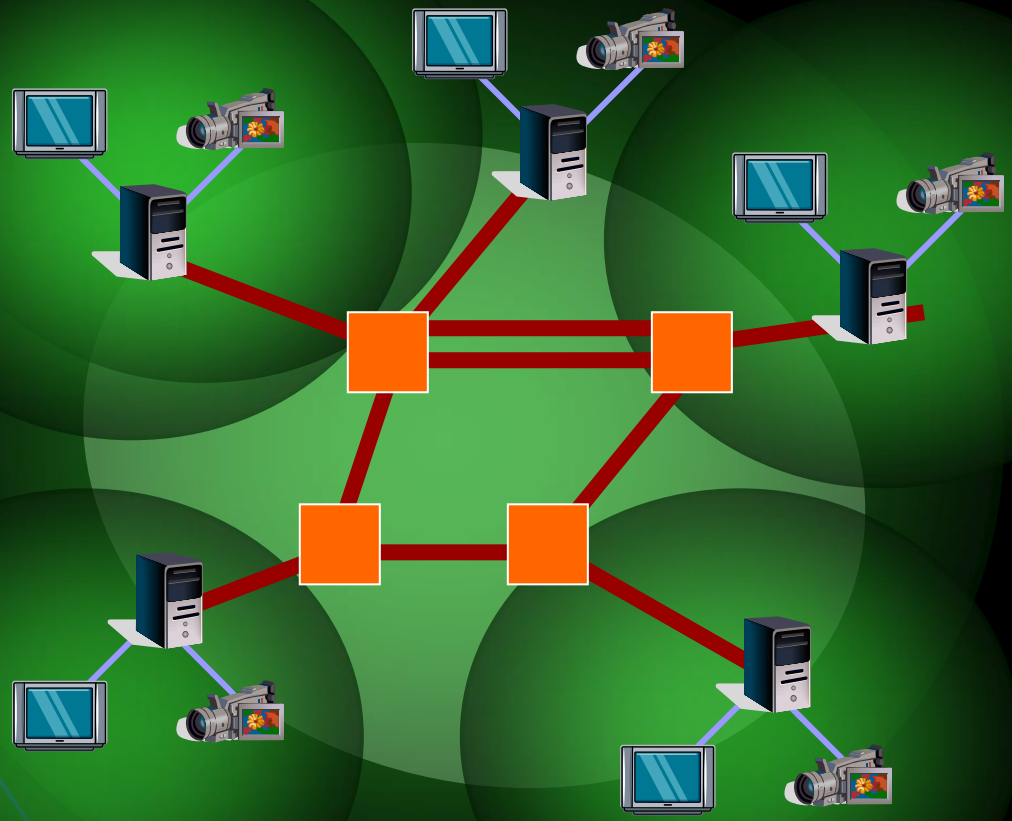
Radio Telescopes
2005 = 512 Mbs
2007 = 2 Gbs
2009 > 4+ Gbs

Aggregated streams at correlator:
2005 > 2 Gbs
2007 ~ 10 Gbs to 20+ Gbs
2009 > 20 Gbs to 40+ Gbs

Video Service Application Controlled Networks

- Video Services
  - Digital video, HD video,
  - Video requires very stringent performance requirements –
    - Compression schemes such as MPEG are extremely sensitive to loss in the network, so engineering long distance video links to eliminate jitter and buffering can reduce loss
    - Compression adds latency, so uncompressed streaming video can significantly improve human factors, but uncompressed requires significantly higher bandwidth and performance
  - Solution: develop video gateways/servers and protocols that know of each other around the world. These servers request specific performance requirements of the network

HD Collaborative
“Video Area Network”



Bulk Data Transfer Application Controlled Networks

- Bulk Data Transport Services
  - Designed to make file transfer work well even when the end systems are not tuned for TCP over long fat pipes
    - TCP sessions can be intercepted (upon user's request) by Generic Session Layer gateway
    - High performance well engineered links, tuned TCP stacks, and TCP proxy processing exist in the GSL gateways distributed around the world
  - GSL gateways know of each other and construct an internal mesh of high speed transport links
    - End systems hosts talk to local gateways and vice versa

The “Black Cloud” project