DRAGON

Dynamic Resource Allocation via GMPLS Optical Networks

Jerry Sobieski

Mid-Atlantic Crossroads (MAX)

National Science Foundation

MPLS 2004

DRAGON Participants

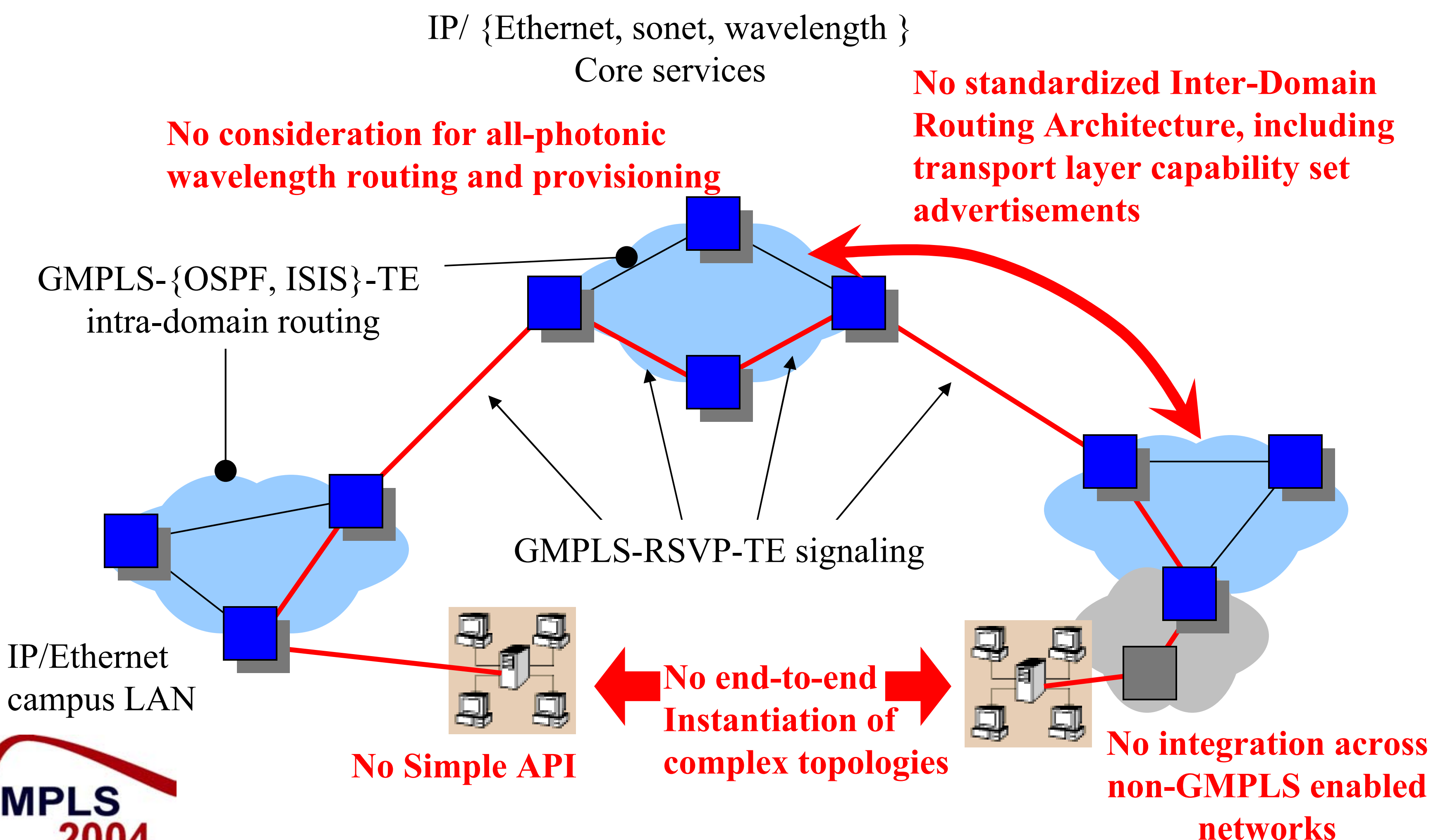
- Mid-Atlantic CrossRoads (MAX)
- University of Southern California Information Sciences Institute (USC/ISI East)
- George Mason University (GMU)
- Movaz Networks
- MIT Haystack Observatory
- NASA Goddard Space Flight Center (GSFC)
- NCSA (National Center for Supercomputing Applications) ACCESS Center
- US Naval Observatory (Wash., DC)
- University of Maryland College Park (UMD)
- Support from:
  - Qwest Communications
  - Nortel Networks
  - Force10 Networks

Funded by National Science Foundation (NSF)

The DRAGON Project
Key Features/Objectives

- Uses all optical transport in the metro/regional core
  - Edge to edge Wavelength switching (2R OEO only for signal integrity)
  - Push OEO demarc to the edge, and increasingly out towards end user
- Standard GMPLS protocols used to dynamically provision intra-domain light paths
  - GMPLS-OSPF-TE and GMPLS-RSVP-TE
- Develop the inter-domain protocol platform
  - Distribute Transport Layer Capability Sets (TLCS) across multiple domains
  - Perform E2E constraint based path computation
  - Resource authorization, scheduling, and accounting
- Develop the "Virtual LSR"
  - Abstracts non-GMPLS network resources into a GMPLS "virtual LSR".
- Simplified API
  - Application Specific Topology definition and instantiation
  - Resource resolution, proxy registration and signaling

End to End GMPLS Transport
What is missing?



MPLS 2004

DRAGON Project
"Advanced Services"

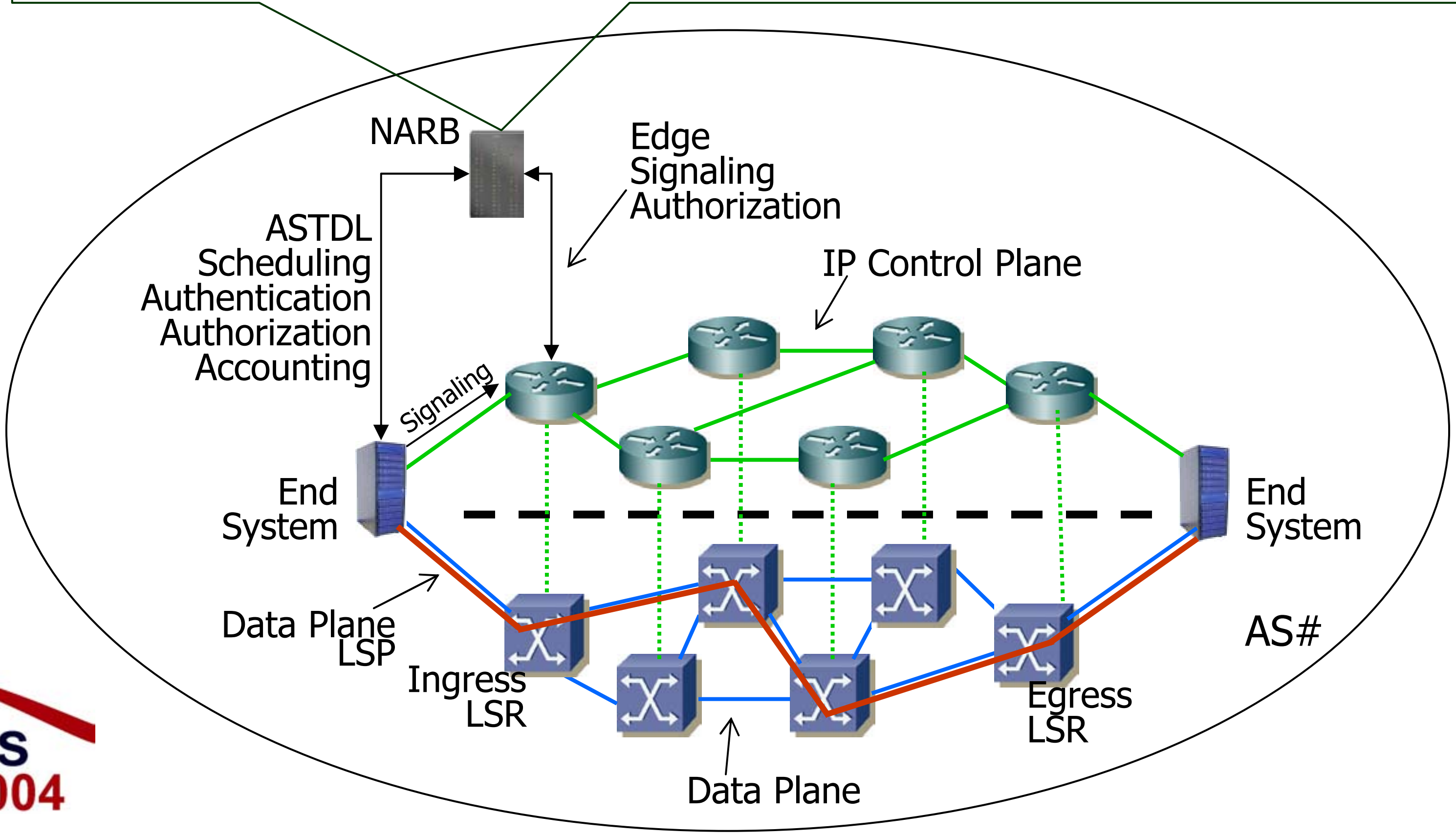
- Dynamic provisioning of dedicated network resources ("Light Paths") end-to-end
  - Deterministic, repeatable, and predictable [network] services
  - Multi-domain, automated, and global provisioning capability
- Rapid provisioning of complex Application Specific Topologies
  - Formalized application<->network requirements
  - Coordination with non-network resource(s)
- Support for:
  - Advanced network resource scheduling/reservations
  - Security and policy (AAA)
- All-photonic metro/regional core:
  - Reduced expense and complexity
  - Protocol, format, framing agnostic: Ethernet/Sonet/FC/SMPTE292

Network Aware Resource Broker (NARB)

- Each NARB agent represents a single domain
- Provides services and functions necessary to address many of the "missing capabilities" required for end-to-end GMPLS scheduling and provisioning
  - Inter-Domain Transport Layer Capability Set (TLCS) exchange
  - Processing of end system topology requests (based on ASTDL)
  - Path Computation
  - Authentication, Authorization, and Accounting (AAA)
  - Resource utilization scheduling, monitoring, and enforcement
  - Edge Signaling Authentication and Enforcement
- NARB is designed to be a robust, distributed agent, managing a persistent traffic engineering and network resource database for the domain.

Network Aware Resource Broker (NARB) Functions – IntraDomain

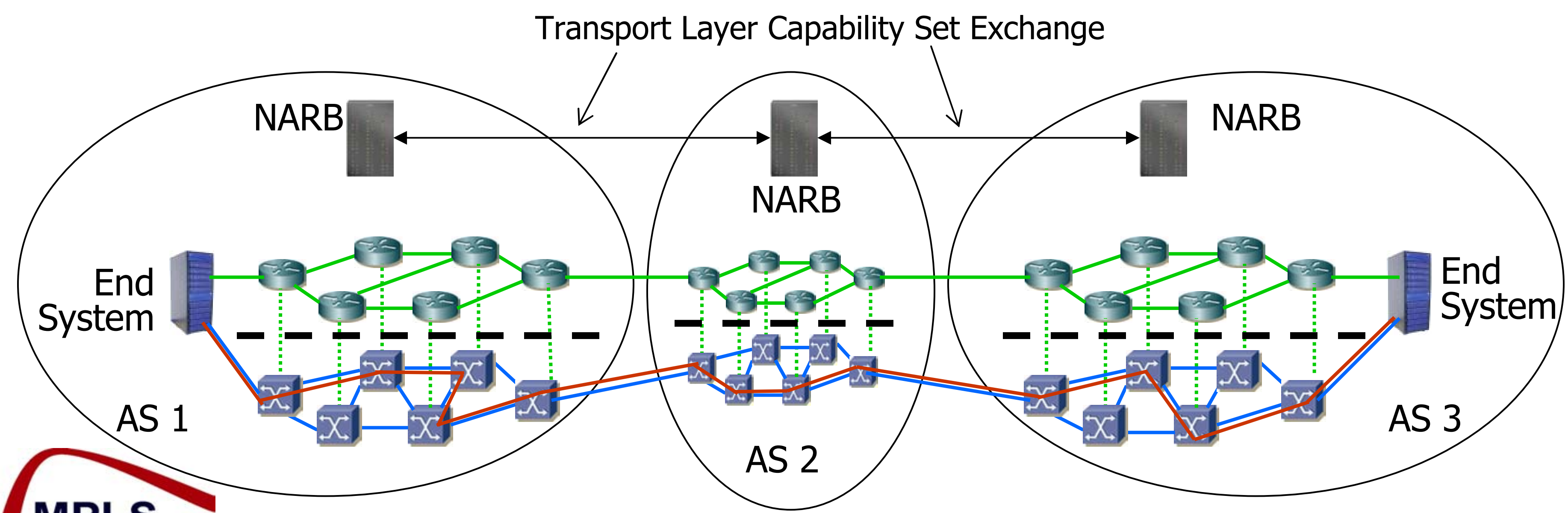
- IGP Listener
- Path Computation
- Scheduling
- Edge Signaling Authentication
- Edge Signaling Enforcement
- ASTDL Induced Topology Computations
- Authorization (flexible policy based)
- Authentication
- Accounting



MPLS 2004

Network Aware Resource Broker (NARB) Functions - InterDomain

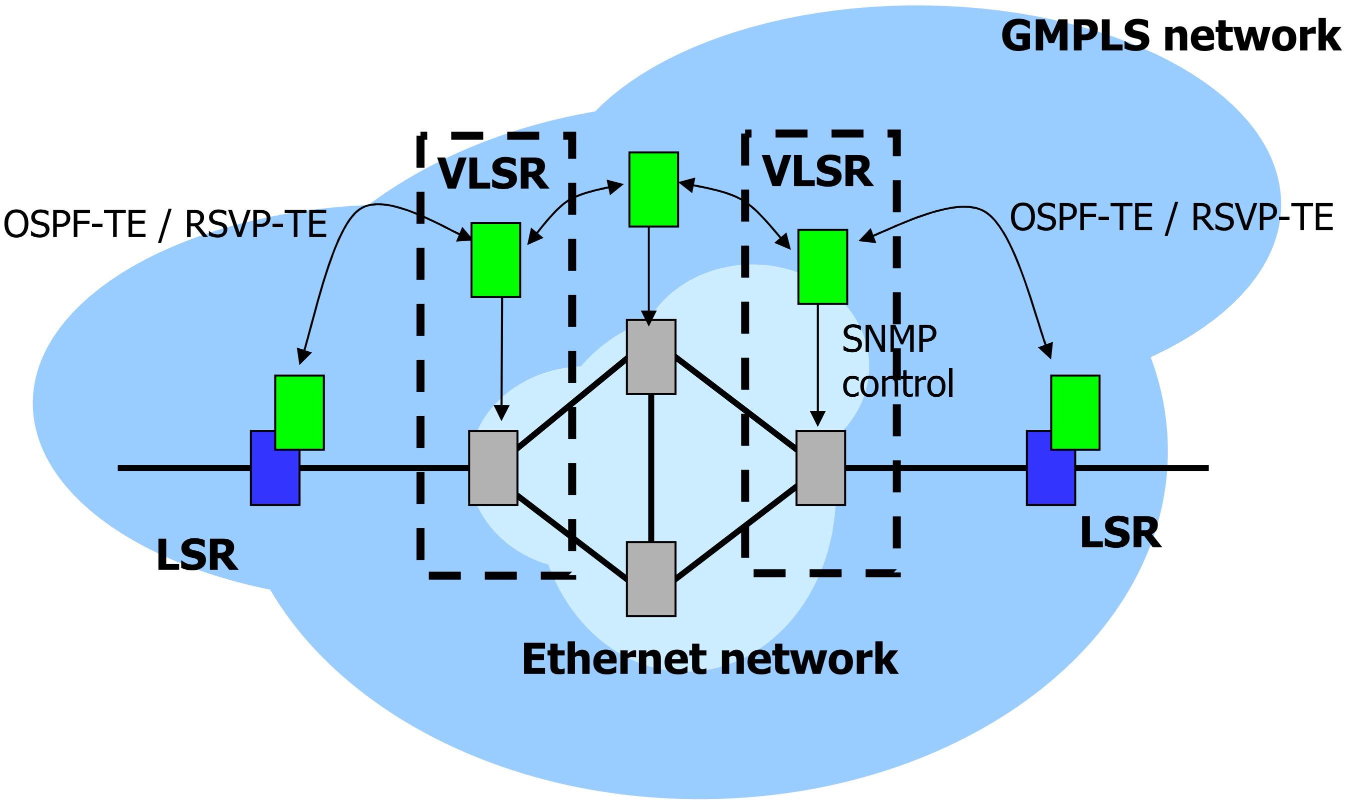
- InterDomain NARB must do all IntraDomain functions plus:
  - EGP Listener
  - Exchange of InterDomain transport layer capability sets
  - InterDomain path calculation
  - InterDomain AAA policy/capability/data exchange and execution



Virtual Label Switched Router - VLSR

- Many networks consist of switching components that do not speak GMPLS, e.g. current ethernet switches, fiber switches, etc
- Such components can be abstracted and "covered" by a Virtual Label Switched Router
  - VLSR incorporates open source implementations of GMPLS-OSPF-TE and GMPLS-RSVP-TE running on a Freebsd control-PC
  - The VLSR interacts with the DRAGON network via GMPLS control plane
  - The VLSR translates GMPLS protocol events into configuration commands to the covered switches via SNMP, TL1, or a similar protocol.

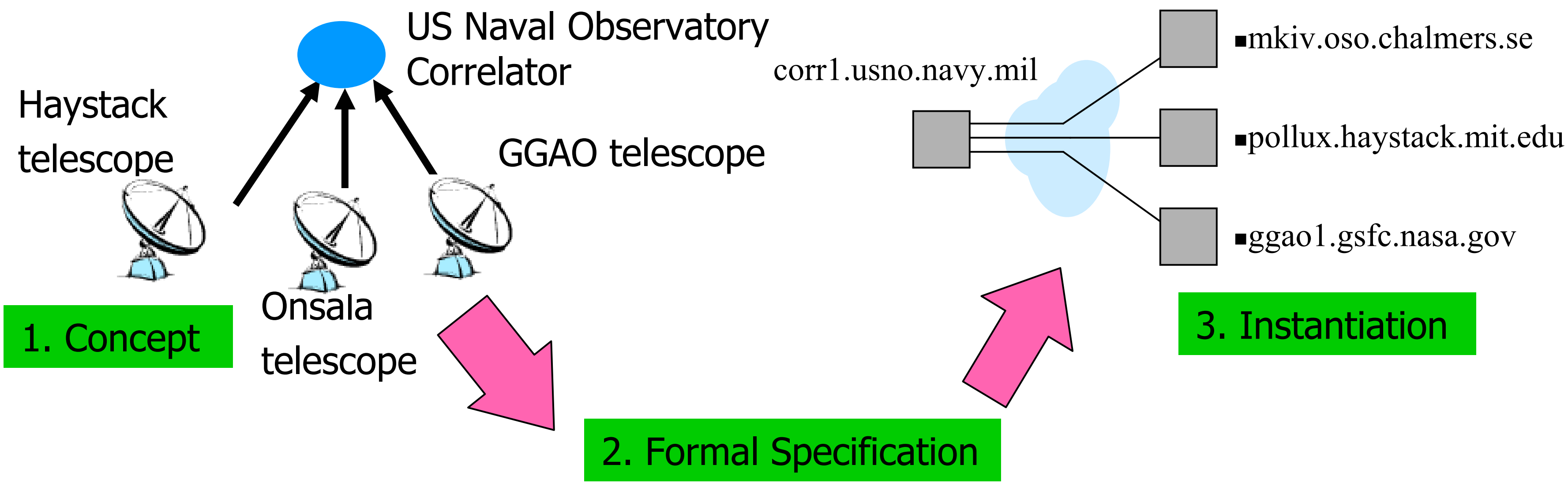
VLSR Abstraction



Application Specific Topologies

- A formalized definition language to describe and instantiate complex topologies
  - Application topologies consist of multiple LSPs ("light paths") that must be instantiated as a whole.
  - Resource availability must be dependable and predictable, i.e. resources must be reserveable in advance for utilization at some later time
  - By formally defining the application's network requirements, service validation and verification can be performed ("wizard gap" issues)

Application Specific Topology Description Language - ASTDL



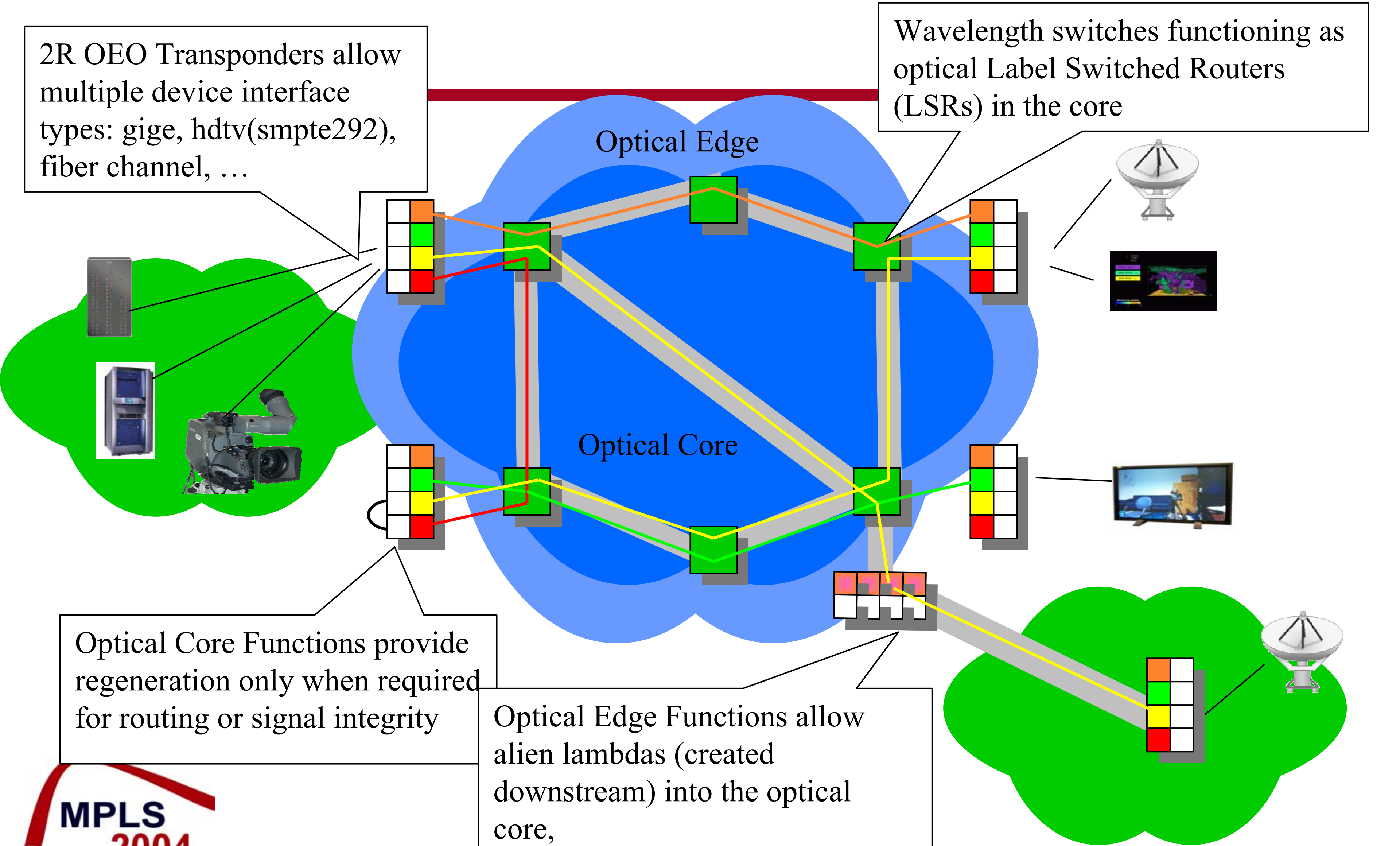

```
Datalink:= { Type=Ethernet; bandwidth=1gbs;
                    SourceAddress=%1::vlbid;   DestinationAddress=%2; }
Topo_vlbi_200406 := {
          Correlator:=corr1.usno.navy.mil::vlbid;         // USNO
          DataLink( mkiv.oso.chalmers.se, Correlator );   // OSO Sweden
          DataLink( pollux.haystack.mit.edu, Correlator );// MIT Haystack
          DataLink( ggao1.gsfc.nasa.gov, Correlator );    // NASA Goddard
          }
```

C++ Code invocation example:

```
eVLBI = new ASTDL::Topo( "Topo_vlbi_200406");  // Get the topology definition
Stat = eVLBI.Create();                         // Make it so!
```

MPLS 2004

All Optical Core and Edge



Routing All Optical Lambdas

- Advantages of “all optical” waves:
  - Framing agnostic: the format of the data modulated onto the wavelength is of no concern (or little concern) to the network
  - Reduced Optical-Electrical-Optical conversion components reducing the cost and complexity of the core
- Challenges
  - Good Optical SNR (I.e. low BER) requires careful attention to fiber engineering, amplification systems, wave band equalization, dispersion management
    - All of these vary with wave path, modulation rates, wavelength, in-path components, etc.
  - Computing optimal paths locally is hard (where you have full visibility of the network characterisitcs)
    - Computing optimal paths across multiple domains is even more challenging

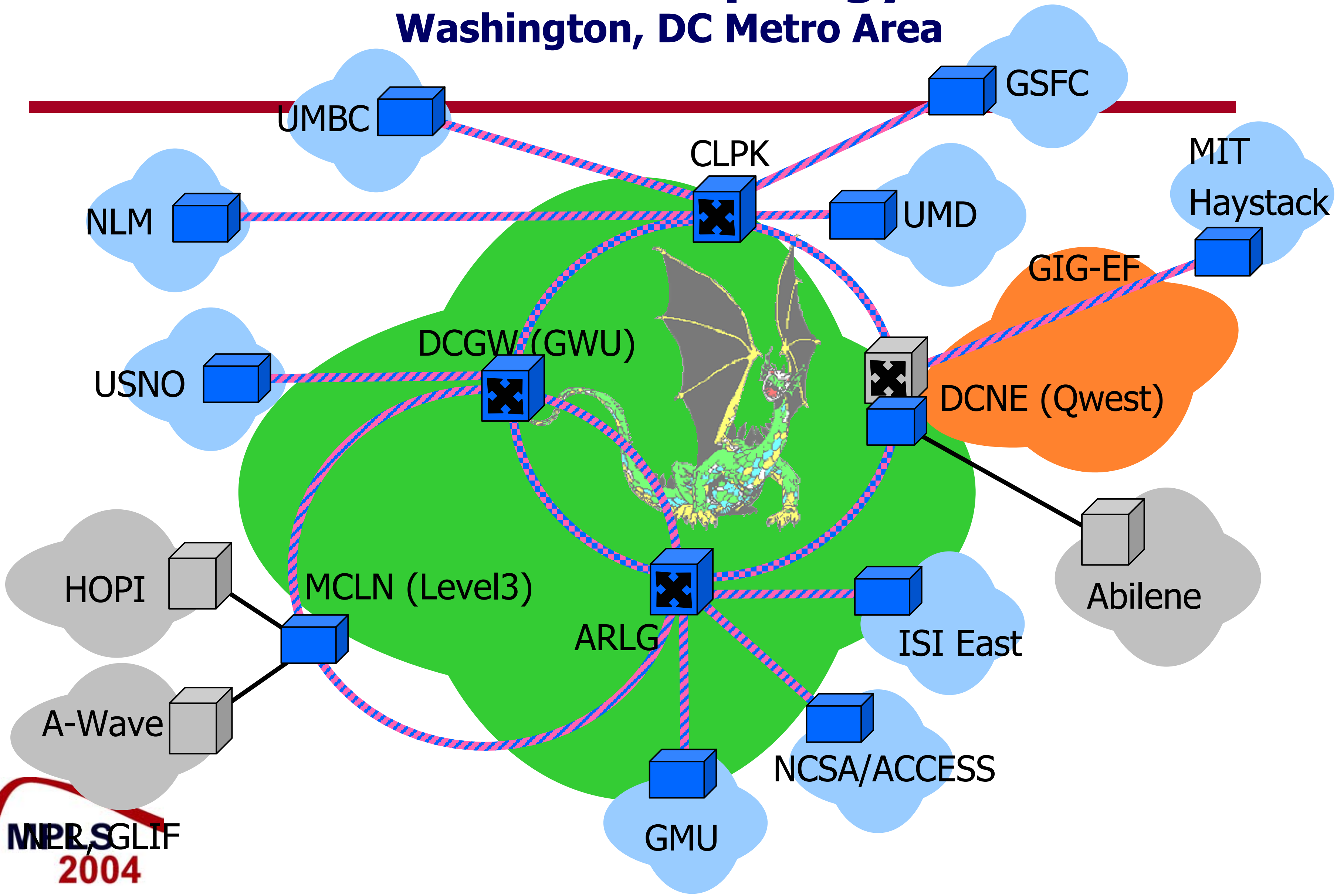
Commercial Partner
Movaz Networks

- MEMS-based switching fabric
- 400 x 400 wavelength switching, scalable to 1000s x 1000s
- 9.23"x7.47"x3.28" in size
- Integrated multiplexing and demultiplexing, eliminating the cost and challenge of complex fiber management

- Dynamic power equalization (<1 dB uniformity) eliminating the need for expensive external equalizers
- Ingress and egress fiber channel monitoring outputs to provide sub-microsecond monitoring of channel performance using the OPM
- Switch times < 5ms

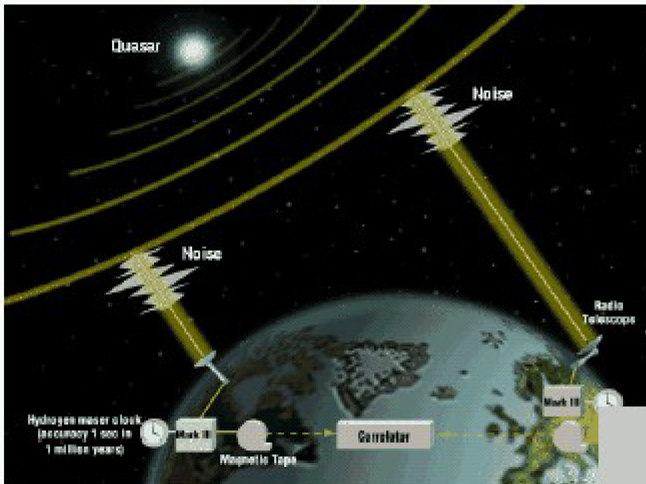
DRAGON Network Topology thru 2006

Washington, DC Metro Area



MPLS 2004

Very Long Baseline Interferometry (VLBI)



The Very-Long Baseline Interferometry (VLBI) Technique
(with traditional data recording on magnetic tape or disk)

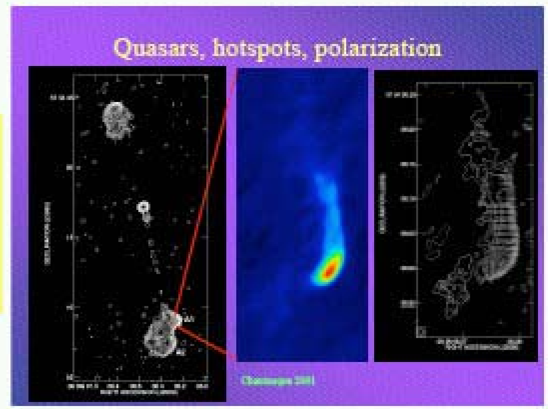
The Global VLBI Array
(up to ~20 stations can be used simultaneously)

MPLS 2004

Very Long Baseline Interferometry (VLBI)

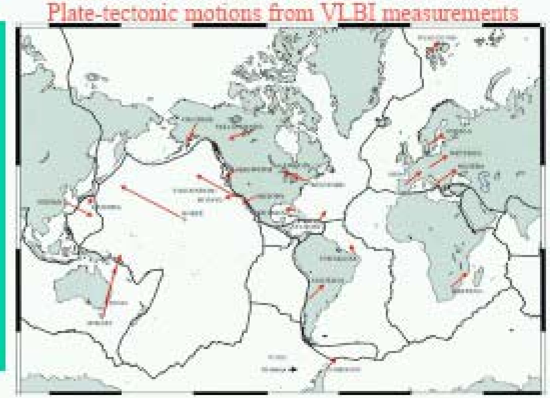
VLBI Science

ASTRONOMY
- Highest resolution technique available to astronomers – tens of microarcseconds
- Allows detailed studies of the most distant objects



GEODESY
- Highest precision (few mm) technique available for global tectonic measurements
- Highest spatial and time resolution of Earth's motion in space for the study of Earth's interior
  - Earth-rotation measurements important for military/civilian navigation
  - Fundamental calibration for GPS constellation within Celestial Ref Frame

Plate-tectonic motions from VLBI measurements



electronic-Very Long Baseline Interferometry (e-VLBI)

- What is it?
  - Radio time series are captured simultaneously by several telescopes around the world (the "Very Long" part)
  - These time series are correlated pairwise (the "Baseline") to identify events occuring within the time series (the "Interferometry") thus allowing the scientist to calculate very accurately where the event occurred.
  - The traditional method for moving the time series data to the correlator sites has been via tapes and jets.
  - Current methods still incur generally two days to capture and move the data to the correlator – that's too long.
  - Why are <Realtime | Near RT | on-demand> resources so important to this application?

MPLS 2004

electronic-Very Long Baseline Interferometry (e-VLBI)

- Because systems such as GPS depend on integrated weather models and VLBI runs, the delay experienced getting the data to the correlators significantly impacts the accuracy and longevity of the predictions:
  - Weather forecast models are typically 5 days
  - Minus 2 days for data transfer = 3 days prediction
  - NRT data transfer will improve predictions by ~40%
- Other celestial events may be transient as well on a scale of minutes to days or weeks
  - Steering the observation NRT allows dramatically more effective use of time on instruments
  - And greatly improved opportunities to acquire useful observations on unpredictable transient events
  - True real-time correlation

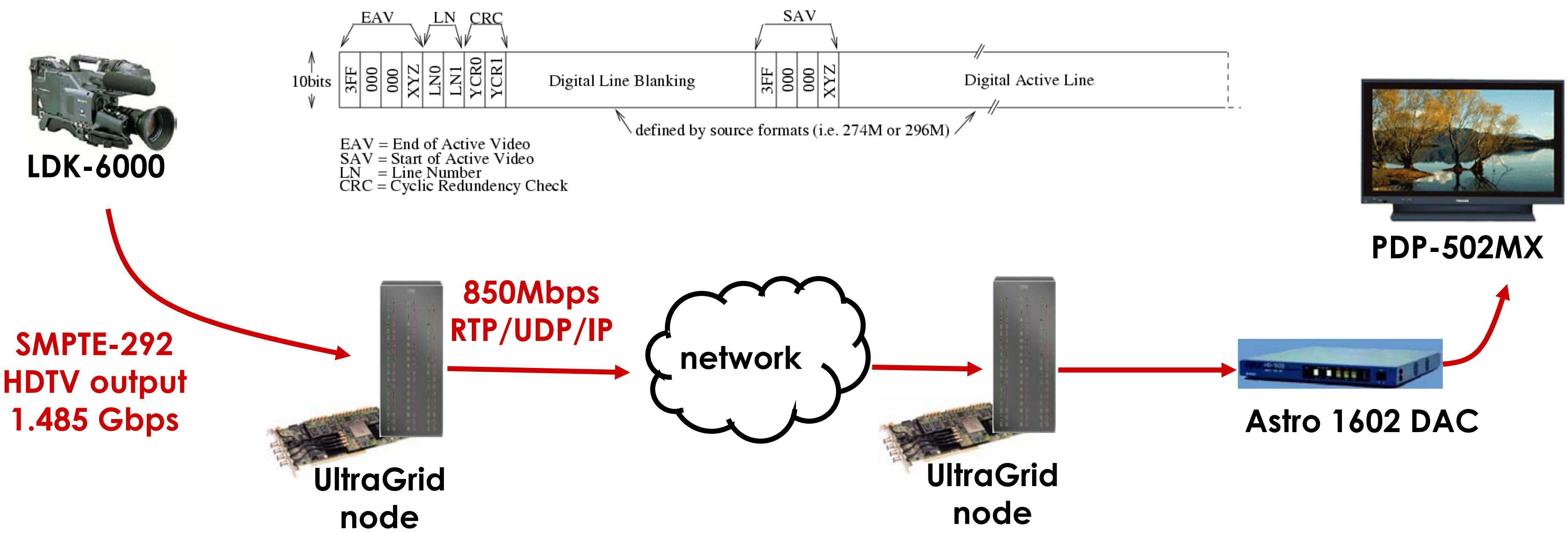
electronic-Very Long Baseline Interferometry (e-VLBI)

- Why is this such an interesting application to demonstrate dedicated, predictable, high performance network topologies?
  - The VLBI process can be used to study the earth.
    - By focusing on very distant very accurately placed objects, scientists can study the changes in the telescope baselines
    - This can provide very accurate information regarding tectonic plate movement, changes in the earth's shape due to glacial rebound from the ice age, etc.
  - Such changes in the Earth's shape change things like the gravitational effects
  - Most notably, VLBI's ability to detect geodetic wobbles in the earth, allow it to predict small but important perturbations in the inertial frame of reference experienced by satellites
    - The Global Positioning System uses VLBI "intensives" to continually recalculate the satellite orbital positions.
  - Interestingly, these geodetic wobbles can be affected by events such as major atmospheric storms such as typhoons or hurricanes.

High Definition Collaboration and Visual Area Networking (HD-CVAN)

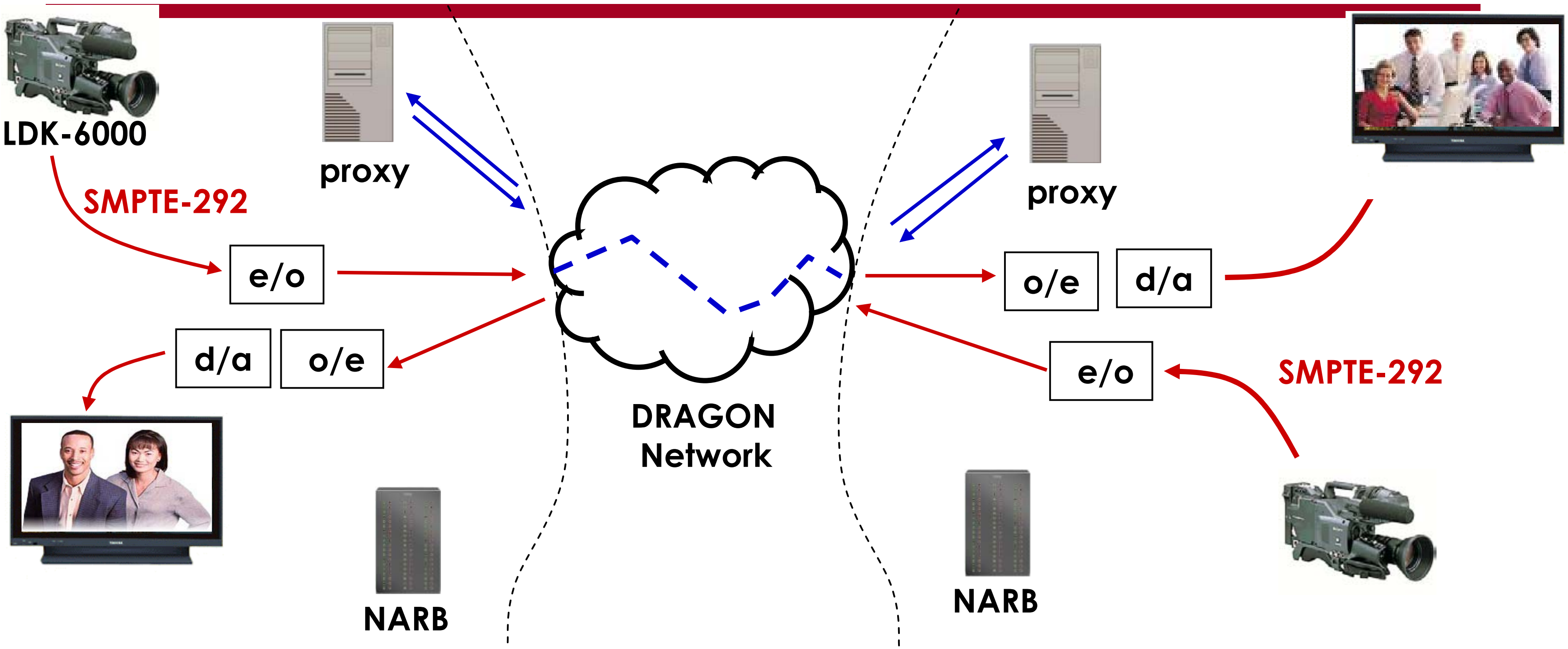
- HD-CVAN Collaborators
  - UMD VPL
  - NASA GSFC (VAL and SVS)
  - USC/ISI (UltraGrid Multimedia Laboratory)
  - NCSA ACCESS
- Dragon dynamic resource reservation will be used to instantiate an application specific topology
  - Video directly from HDTV cameras and 3D visualization clusters will be natively distributed across network
- Integration of 3D visualization remote viewing and steering into HD collaboration environments

Uncompressed HDTV-over-IP Current Method



- Not truly HDTV --> color is subsampled to 8bits
- Performance is at the mercy of best-effort IP network
- UltraGrid processing introduces some latency

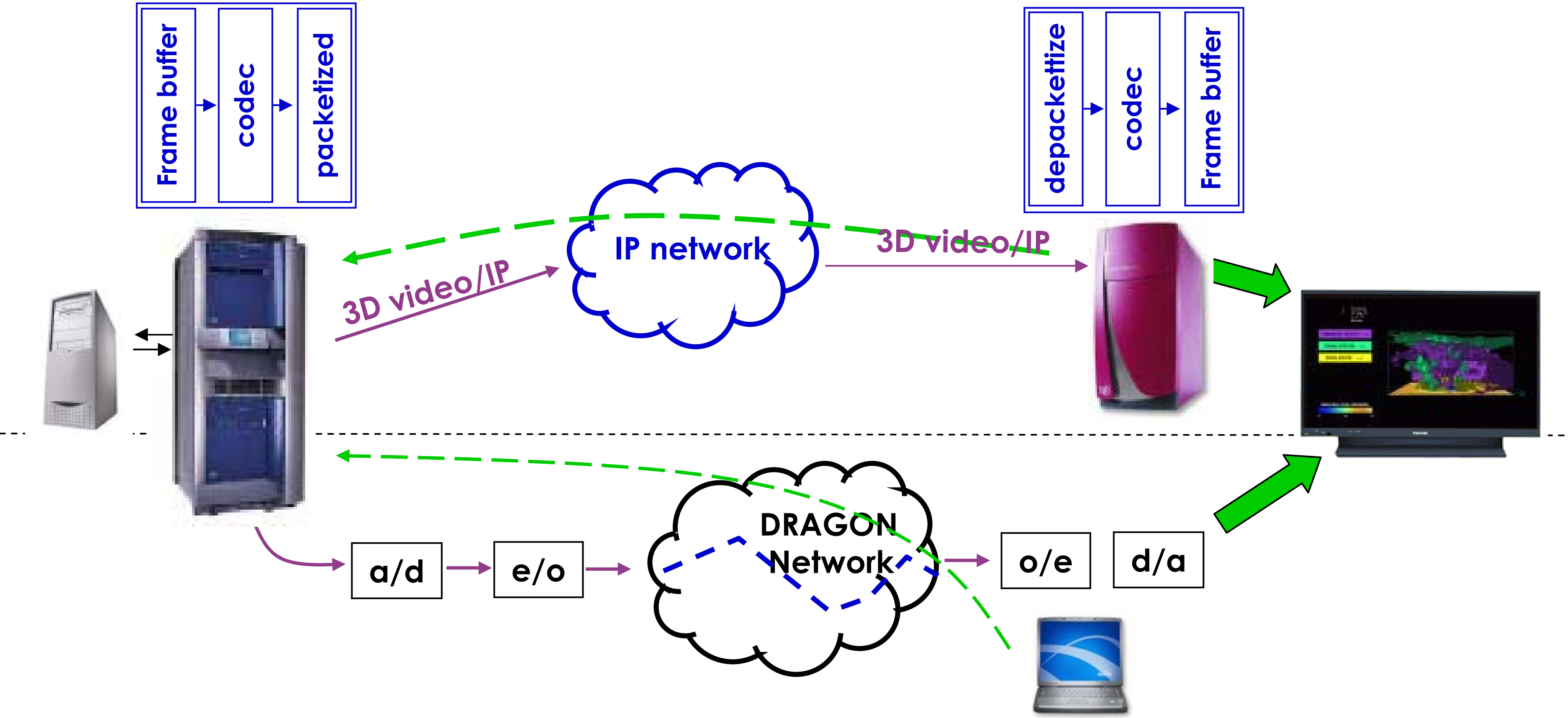
Low latency High Definition Collaboration DRAGON Enabled



- End-to-end native SMPTE 292M transport
- Media devices are directly integrated into the DRAGON environment via proxy hosts
  - Register the media device (camera, display, …)
  - Sink and source signaling protocols
  - Provide Authentication, authorization and accounting.

MPLS 2004

Low Latency Visual Area Networking



- Directly share output of visualization systems across high performance networks.
- DRAGON allows elimination of latencies associated with IP transport.

MPLS 2004

Status to Date

- Wavelength Selective Switch staged at UMD
  - GMPLS control plane operational (Jan 04)
- Ring A operational (Aug 2004)
  - Optical peering issues being worked out with ATDNet.
  - MIT Haystack optically re-attached to DRAGON via 700+ km [unregenerated] OC48 wave (plans to move to 10g)
- Initial VLSR functionality demonstrated
  - Demonstrated signaling setup across ethernet switches (April 04), Movaz DWDM system (June 04), OSPF routing and signaling can be seen at Supercomputing 2004 (Nov 2004)
  - Initial software distribution available at dragon.maxgigapop.net
- Inter-domain testing of NARB is in progress
  - All photonic inter-domain "peering" being tested between DRAGON and GIG-EF (formerly ATDNet/BOSSnet)
  - Plans being made to extend DRAGON to NLR and begin inter-domain experiments with HOPI, GLIF, others

Open Issues

- Common Service Definition(s) needs to be developed
  - What are the semantics of the service parameters communicated between domains? E.g. if I request an “ethernet” label, can I assert spanning tree protocols across it? Or does it simply mean ethernet framing at the handoff?
- Common scheduling and advance reservation protocol across domains
  - Are GRID scheduling algorithms and tools adaptable to network resources? Do we need different tools for the network portion?
- How does the network verify that the desired performance specifications are being provided?
  - Often long distance connections transit multiple NE’s that impose unexpected constraints; How do we verify the integrity of an all-photonic data stream?
- Privacy and security – How do we protect the control plane? What aspects do we need to worry about?

Closing

- DRAGON is an architecture – We hope to develop a model and BCPs for designing and deploying dynamic light path services.
- We welcome inquiries and are always interested in how the project can be improved

Jerry Sobieski
jerrys@maxgigapop.net
301-314-6662

Thank you!