

TeraPaths: A QoS Enabled Collaborative Data Sharing Infrastructure for Petascale Computing Research

The TeraPaths Project Team

Usatlas Tier 2 workshop

Outline



- ❄ Introduction
- ❄ The TeraPaths project
- ❄ The TeraPaths system architecture
- ❄ Interoperability with WAN network service
- ❄ Experimental deployment and testing
- ❄ Future work

Introduction



- ❄ **The problem:** support efficient/reliable/predictable peta-scale data movement in modern high-speed networks
 - ❑ Multiple data flows with varying priority
 - ❑ Default “best effort” network behavior can cause performance and service disruption problems

- ❄ **Solution:** enhance network functionality with QoS features to allow prioritization and protection of data flows

The QoS Arsenal



❄ IntServ

- ❑ RSVP: end-to-end, individual flow-based QoS

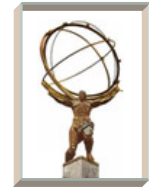
❄ DiffServ

- ❑ Per-packet QoS marking
- ❑ IP precedence (6+2 classes of service)
- ❑ DSCP (64 classes of service)

❄ MPLS/GMPLS

- ❑ Uses RSVP-TE
- ❑ QoS compatible
- ❑ Virtual tunnels, constraint-based routing, policy-based routing

The TeraPaths Project



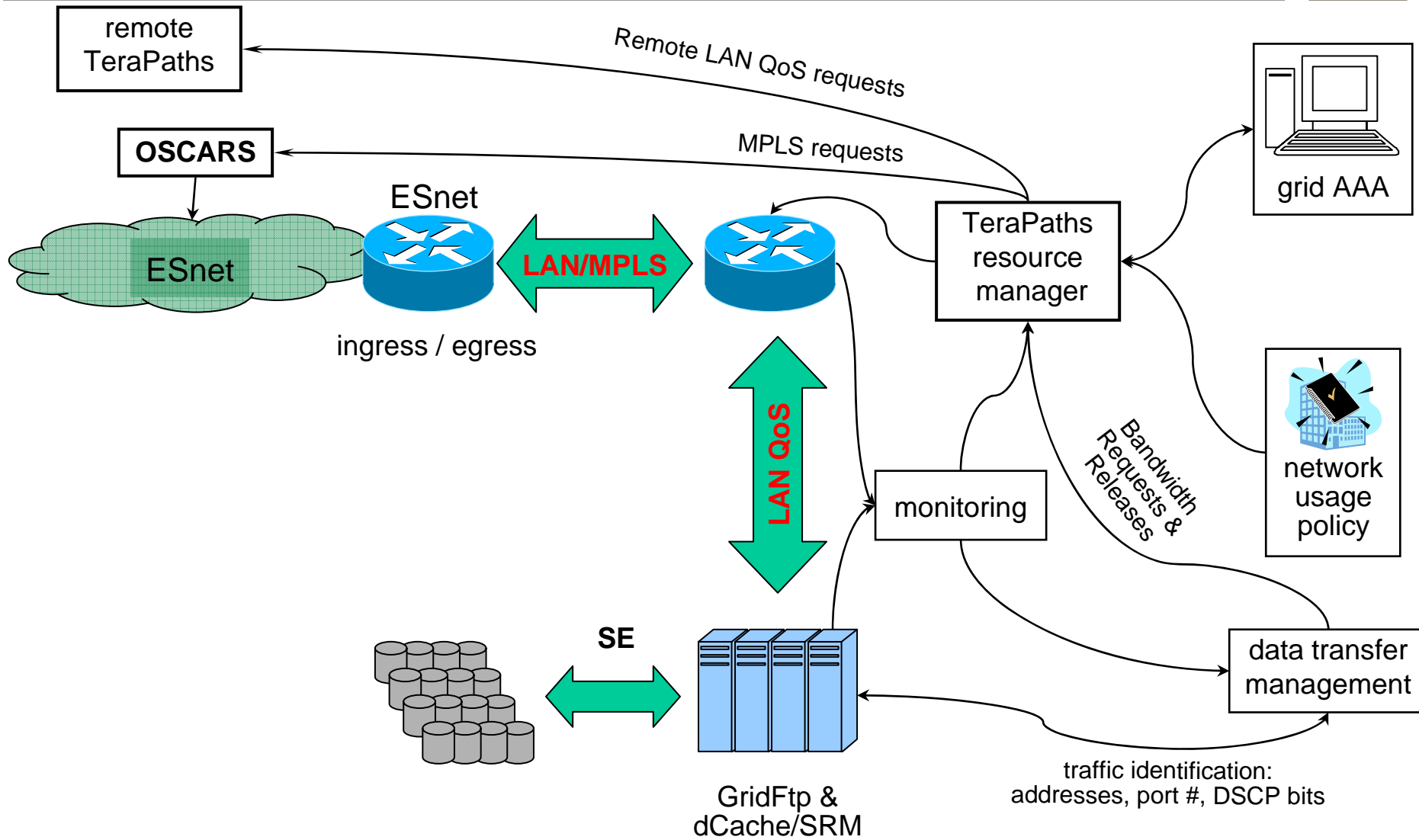
- ❄ The TeraPaths project investigates the integration and use of LAN QoS and MPLS/GMPLS-based differentiated network services in the ATLAS data intensive distributed computing environment in order to manage the network as a **critical resource**
- ❄ DOE: The collaboration includes BNL and the University of Michigan, as well as OSCARS (ESnet), and DWMI (SLAC)
- ❄ NSF: BNL participates in UltraLight to provide the network advances required in enabling petabyte-scale analysis of globally distributed data
- ❄ NSF: BNL participates in a new network initiative: PLaNetS (Physics Lambda Network System), led by CalTech

Newly Development

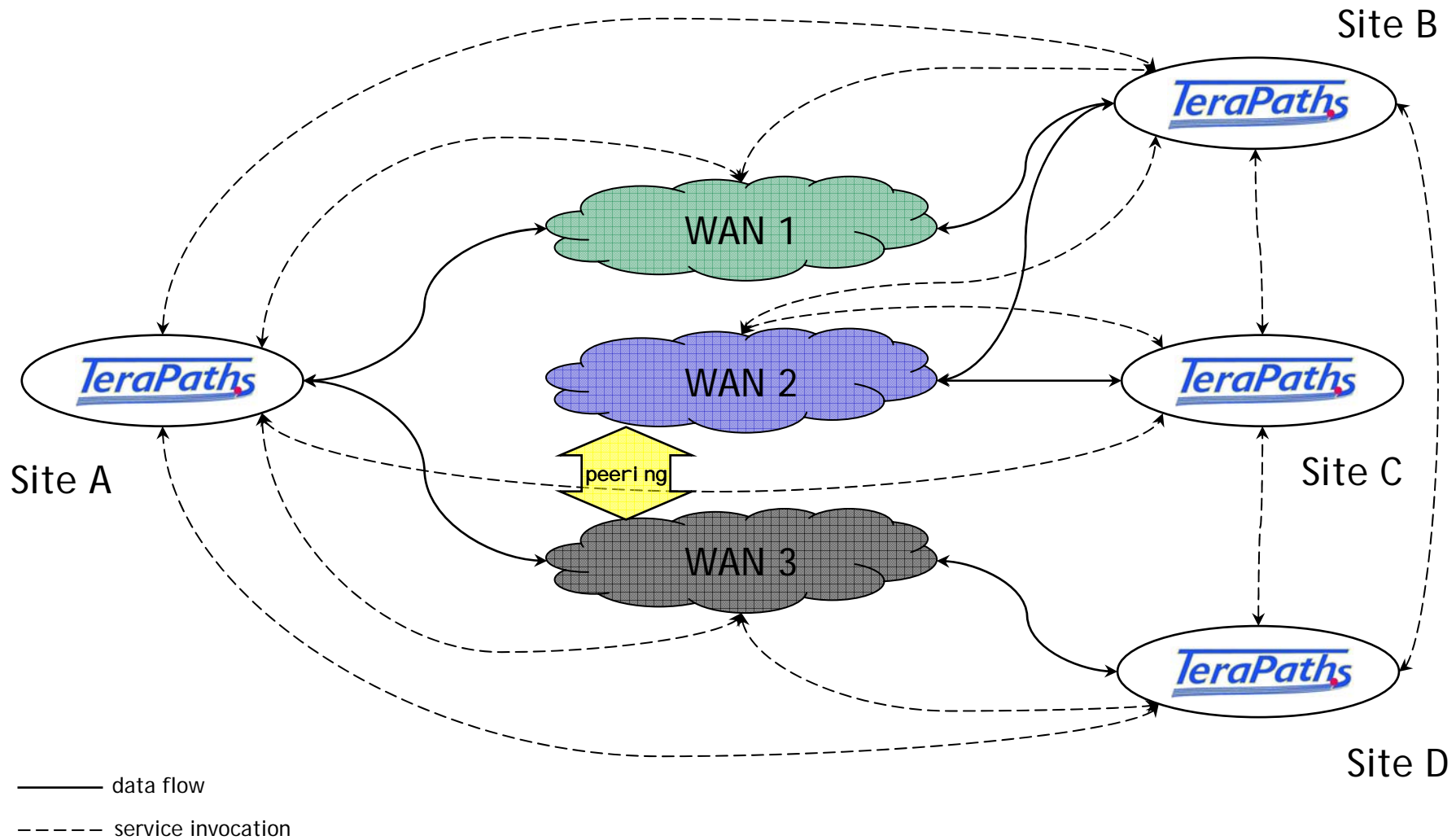
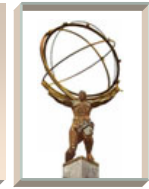


- ❄ TeraPaths is rapidly evolving from a last-mile, LAN QoS provider to a distributed end-to-end network path QoS negotiator through multiple administrative domains.
- ❄ Developed as a web service-based software system.
- ❄ TeraPaths automates the establishment of network paths with QoS guarantees between end sites by configuring their corresponding LANs and requesting MPLS paths through WANs on behalf of end users.
- ❄ The primary mechanism for the creation of such paths is the negotiation and placement of advance reservations across all involved domains.

BNL Site Infrastructure



Envisioned Overall Architecture



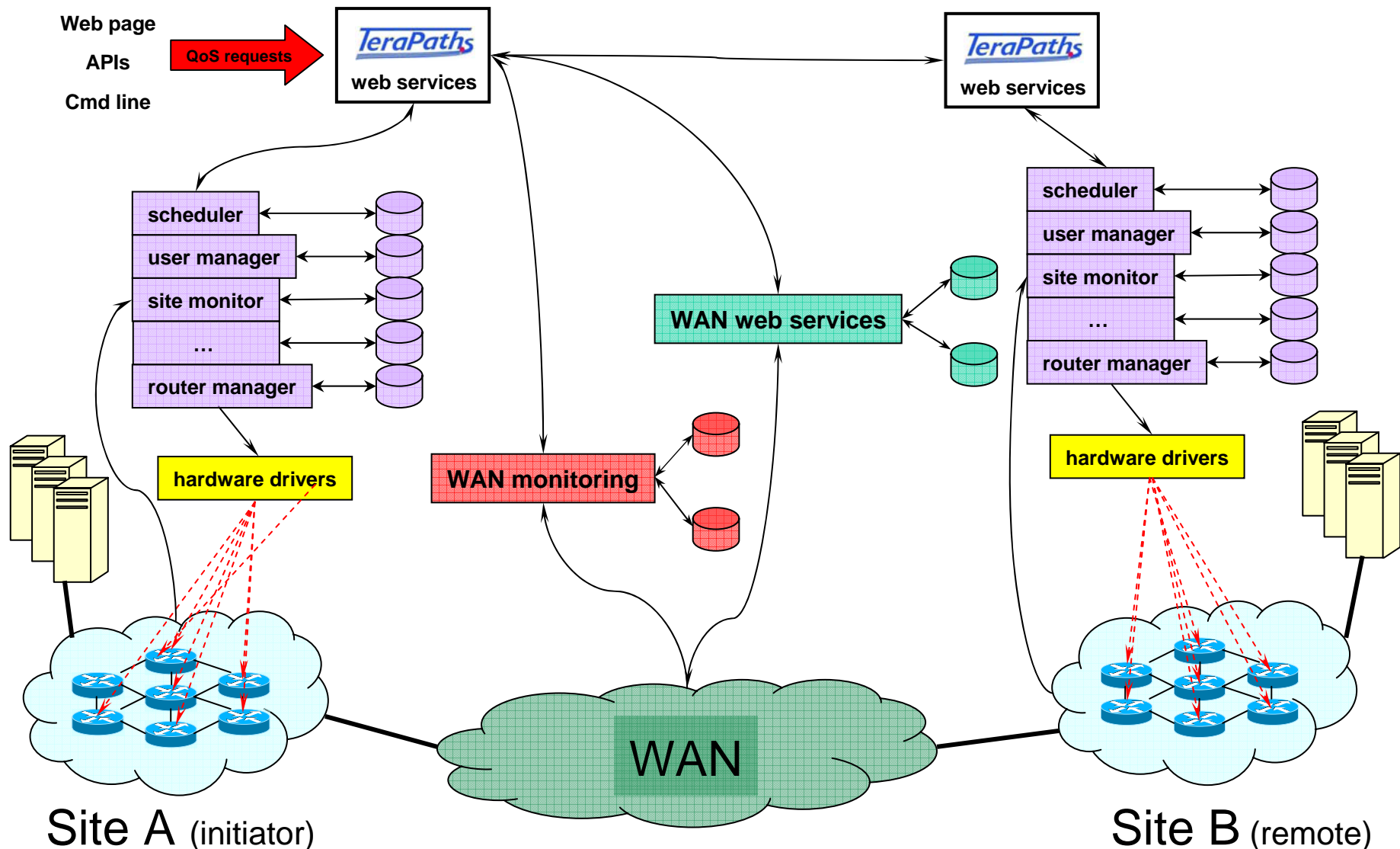
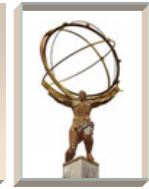
Automate MPLS/LAN QoS Setup



- ❄ QoS reservation and network configuration system for data flows
 - ❑ Access to QoS reservations:
 - ⌘ Manually, through interactive web interface
 - ⌘ From a program, through APIs
 - ❑ Compatible with a variety of networking components
 - ❑ Cooperation with WAN providers and remote LAN sites
 - ❑ Access Control and Accounting
 - ❑ System monitoring

- ❄ **Design goal:** *enable the reservation of end-to-end network resources to assure a specified “Quality of Service”*
 - ❑ User requests minimum bandwidth, start time, and duration
 - ❑ System either grants request or makes a “counter offer”
 - ❑ Network is setup end-to-end with one user request

TeraPaths System Architecture



TeraPaths Web Services



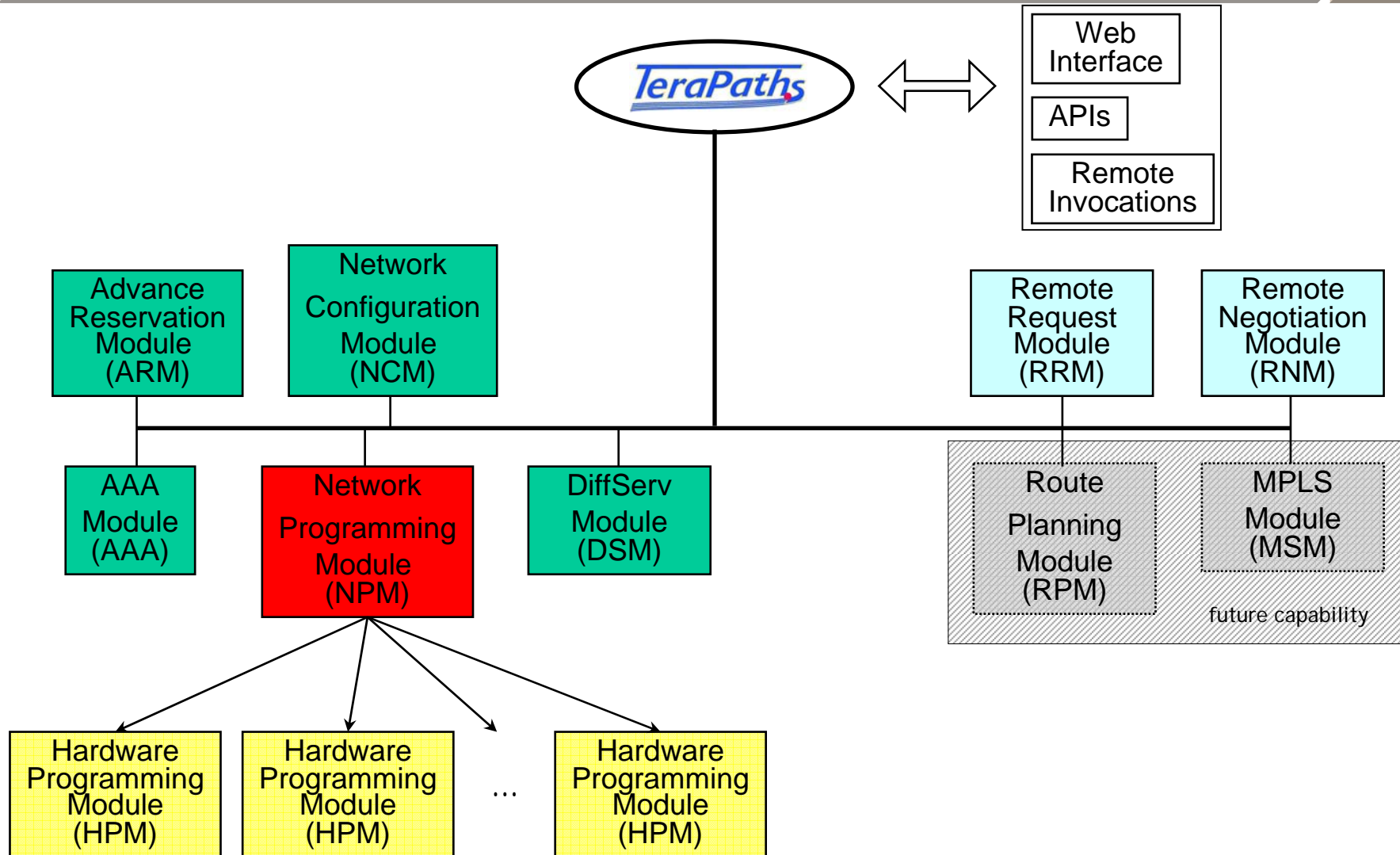
❄ TeraPaths modules implemented as “web services”

- ❑ Each network device (router/switch) is accessible/programmable from at least one management node
- ❑ Site management node maintains reservation etc. databases and distributes network programming by invoking web services on subordinate management nodes
- ❑ Remote requests to/from other sites invoke corresponding web services (destination site’s TeraPaths or WAN provider’s)

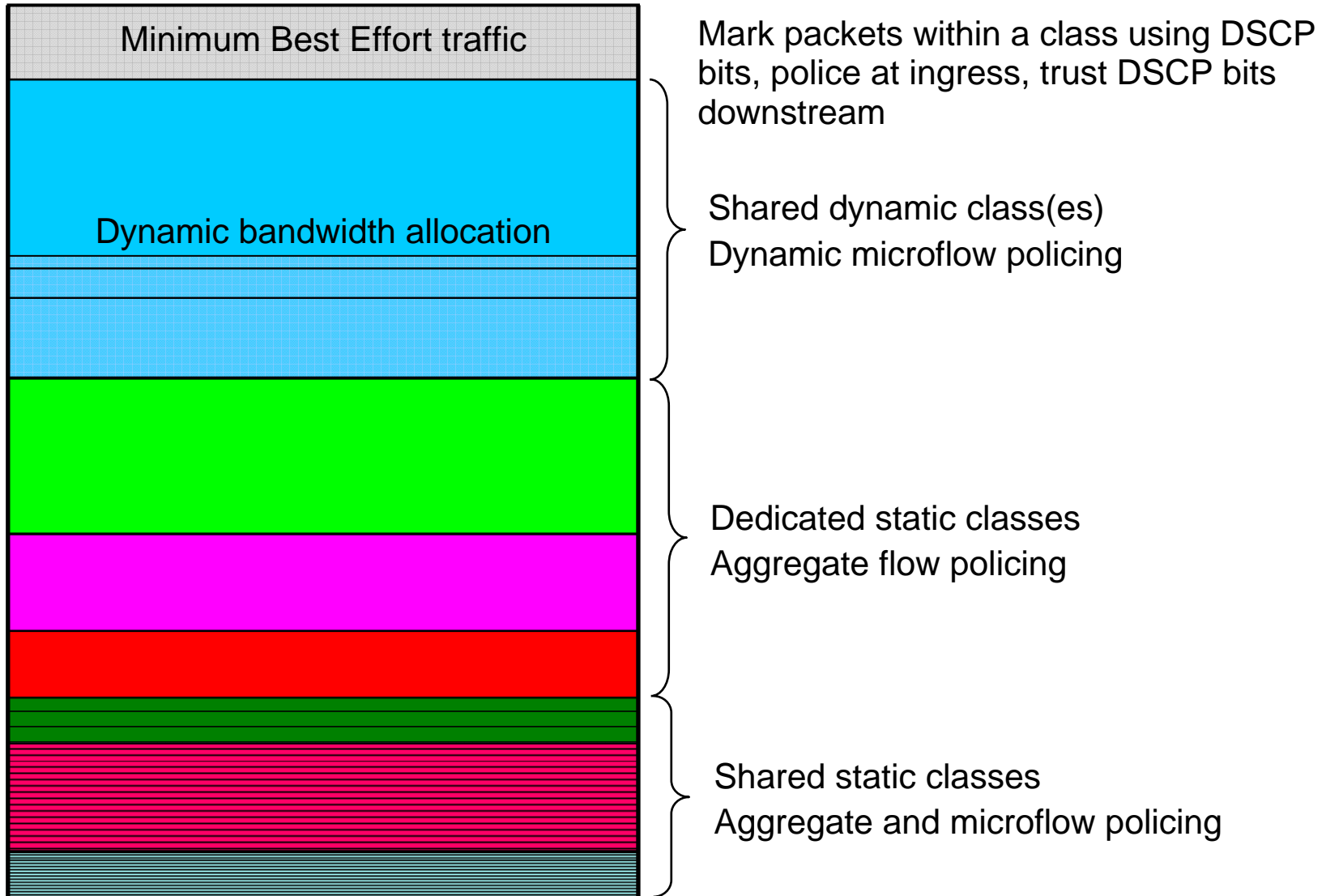
❄ Web services benefits

- ❑ Standardized, reliable, and robust environment
- ❑ Implemented in Java and completely portable
- ❑ Accessible via web clients and/or APIs
- ❑ Compatible and easily portable into Grid services and the Web Services Resource Framework (WSRF in GT4)

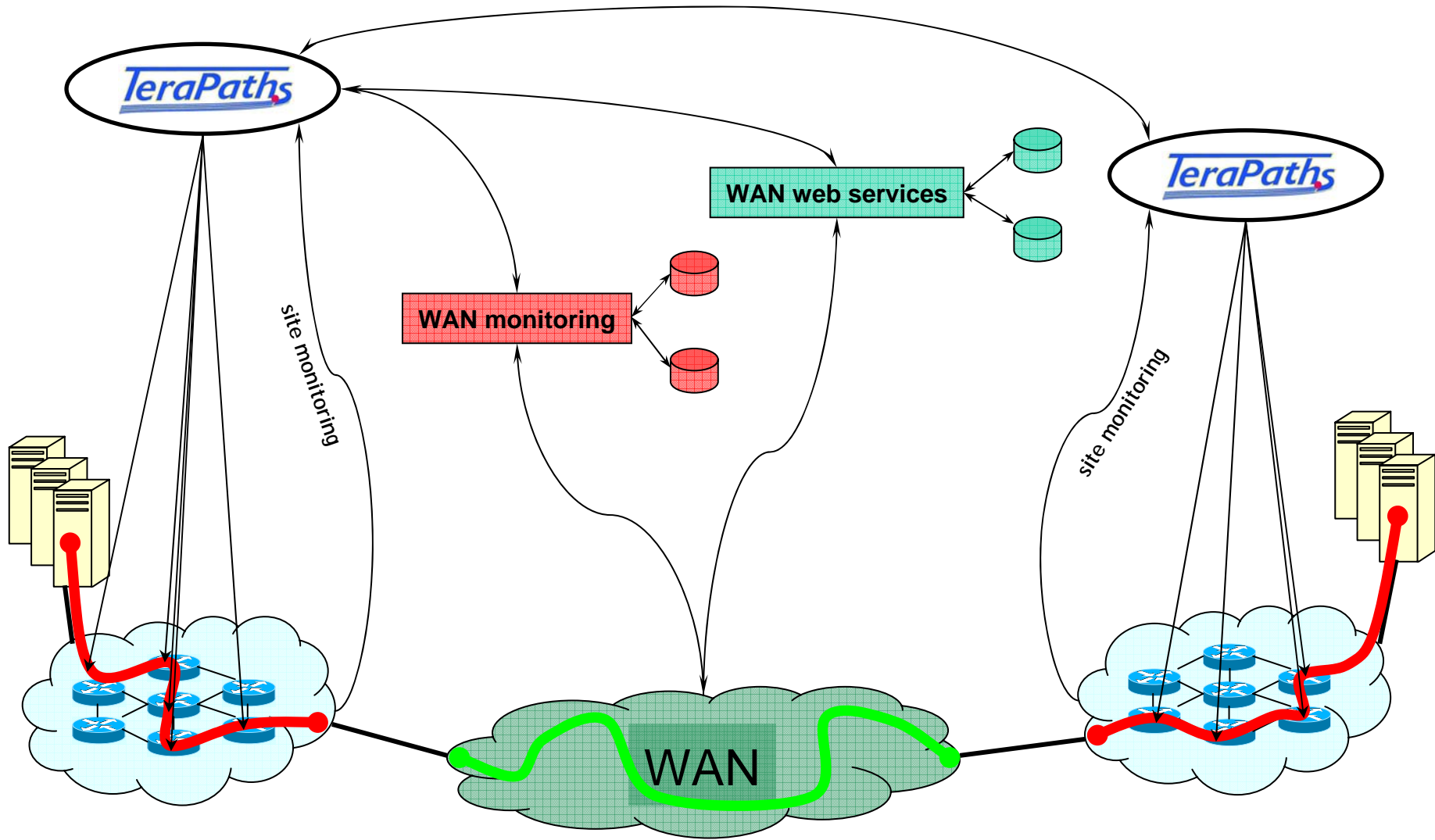
TeraPaths Web Services Structure



Site Bandwidth Partitioning Scheme



Route Planning with MPLS

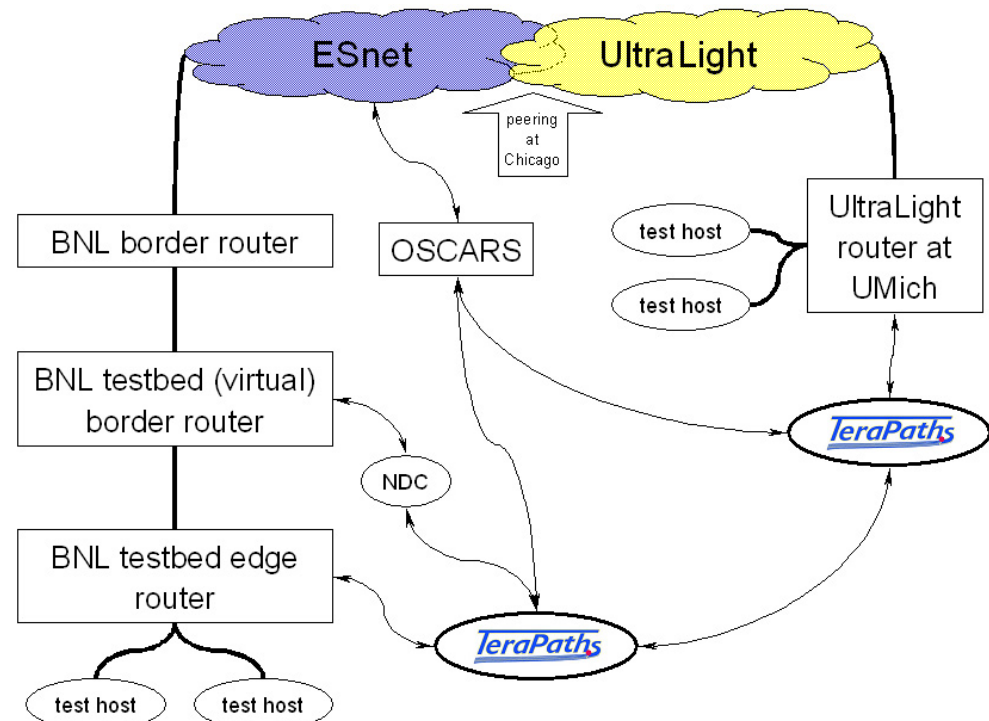
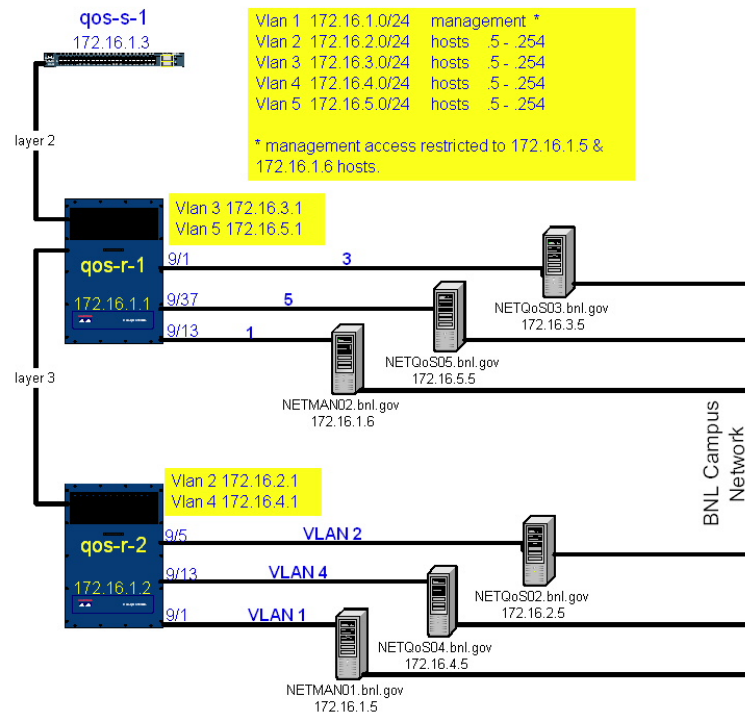


Experimental Setup

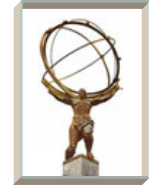


- ❄ Full-featured LAN QoS simulation testbed using a private network environment and WAN network testbed:

QoS Testbed



Acquired Experience



- ❄ Enabled, tested, and verified LAN QoS inside BNL campus network
- ❄ Tested and verified MPLS paths between BNL and University of Michigan.
- ❄ Integrated LAN QoS with MPLS paths reserved with OSCARS
 - ❑ Fully interoperate between Web services of these two projects.
- ❄ Installed DWMI network monitoring tools
- ❄ Examined impact of prioritized traffic on overall network performance and the effectiveness and efficiency of MPLS/LAN QoS
- ❄ Weekly meeting between ESnet, SLAC, University of Michigan and BNL.
- ❄ Developed and deployed remote negotiation/response, etc. services to fully automate end-to-end QoS establishment across multiple network domains
- ❄ Dynamically configure and partition QoS-enabled paths to meet time-constrained network requirements
- ❄ Integrate with software from other network projects: OSCARS, lambda station, and DWMI
- ❄ A peer-reviewed paper will be published in GridNet 2006.

In Progress / Future Work



- ❄ Add GUMS based Grid AAA into the TeraPaths.
- ❄ Develop site-level network resource manager for multiple VOs vying for limited WAN resources
- ❄ Support dynamic bandwidth/routing adjustments based on resource usage policies and network monitoring data (provided by DWMI)
- ❄ widen deployment of QoS capabilities to USATLAS tier 2 sites
- ❄ TeraPaths addendum was approved and new additional funding will be provided for Tier 2 deployment.
 - Each interested Tier 2 cluster will get moderated funding for this activity.
- ❄ End-to-end network meeting will be in BNL around September/October.