Lambda Station

Fermilab

Cal Tech

High-Performance Network
Research PI Meeting

BNL

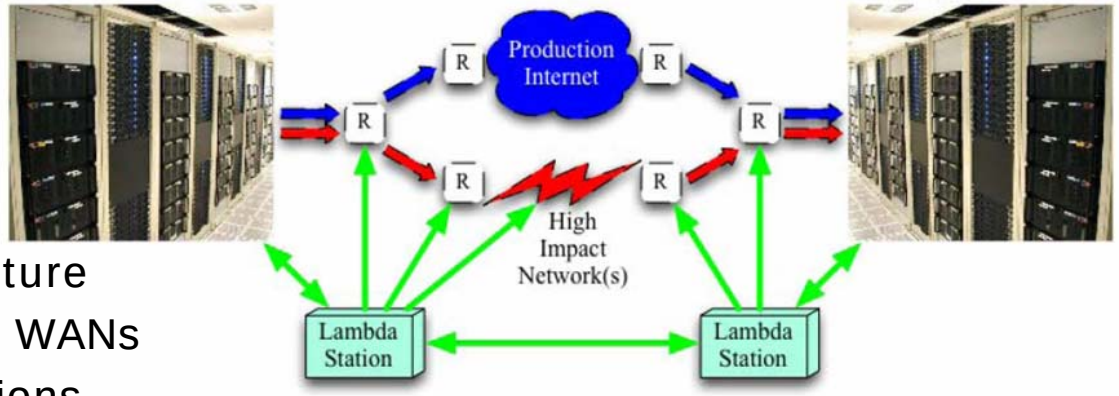
Phil DeMar

September 29, 2005

Lambda Station

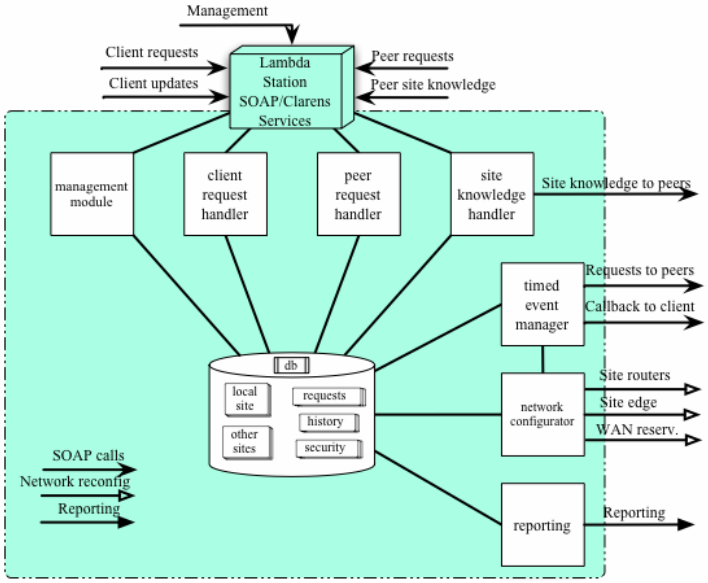
- A network path forwarding service to interface production facilities with advanced research networks:
  - Goal is selective forwarding on a ***per flow*** basis
    - Alternate network paths for high impact data movement
  - Dynamic path modification, with graceful cutover & fallback
  - Current implementation based on policy-based routing & DSCP marking

- Lambda Station interacts with:
  - Host applications / systems
  - LAN infrastructure
  - Site border infrastructure
  - Advanced technology WANs
  - Remote Lambda Stations



Lambda Station Building Blocks

- Soap service (Clarens) supporting:
  - Lambda Station ⇔ client communications
  - Inter-Lambda Station communications
  - Integrated scheduler function
- Network configuration module:
  - Internal path forwarding reconfiguration
  - WAN alternate path coordination
  - Site perimeter modification
- Local Lambda Station configuration:
  - Local site:
    - Alternate WAN path definitions
    - PBR client (internal alternate path definitions
    - Local alternate path clients (host systems/net blocks)
    - Requestors (identity strings for authentication)
  - At least one remote Lambda Station
    - Dynamic exchange of known Lambda Station information



Development Status

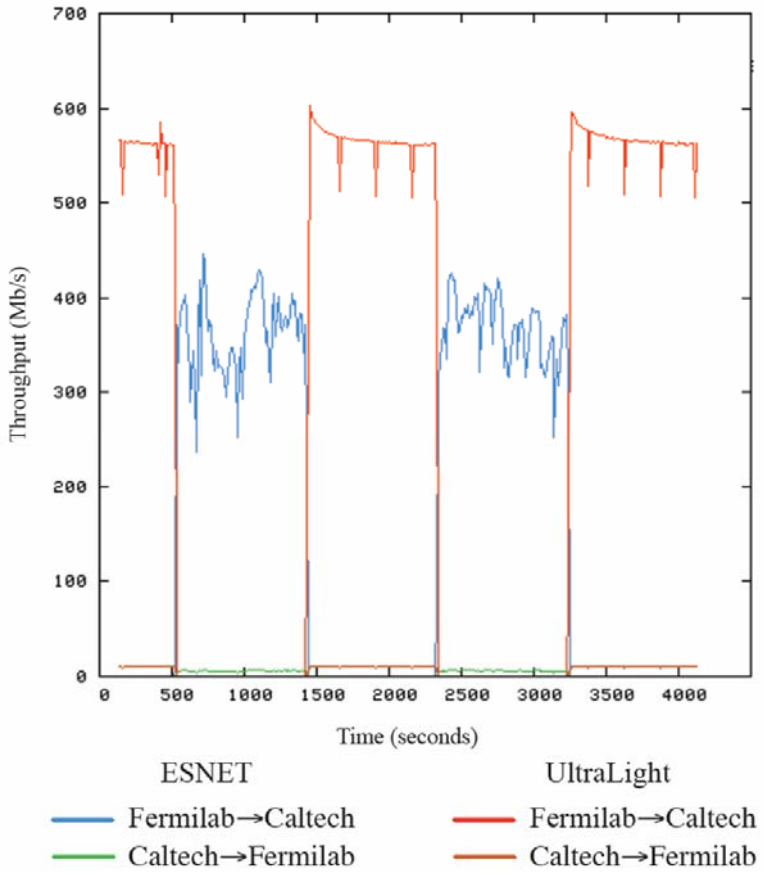
- Lambda Station V1.0 implements full cycle functionality:
  - Functional Soap service:
    - Lambda Station interface protocol defined with 20 primitives
    - Ticket request process operational, including remote end feedback
      - → Peer Lambda Station coordination working…
    - Basic authentication via SSL (today…)
  - Dynamic network reconfiguration working:
    - PBR forwarding set & removed in both directions on multiple routers
    - Concurrent modification of interface ACLs for proper access control
    - Not required. Static PBR forwarding permissible with Lambda Station
  - Lambda Station-aware applications developed:
    - LSIperf (Iperf with Lambda Station awareness)
    - LSTraceroute (Traceroute with Lambda Station awareness)
  - Monitoring capabilities developed:
    - Scripts developed for real time snmp interface displays
    - Quasi-real time analysis of flow data

Deployment & Testing

- Lambda Station servers deployed at FNAL (2) & Cal Tech
- Alternate WAN path testing with UltraScience Net & UltraLight conducted
  - Static scheduling & WAN path setup, at this stage…
- Test network environment developed:
  - Cisco & Force10 (soon…) routers for PBR manipulation
  - Four 10GE systems and 60+ node test farm
- Multiple simultaneous alternate path requests tested:
  - Three in each direction worked fine; scalable to order 10s
- Designed simulation of more advanced reservation system with resource scheduler:
  - Concerns about deadlock conditions & mutual exclusion
  - Four Lambda Stations with three concurrent requests tested OK

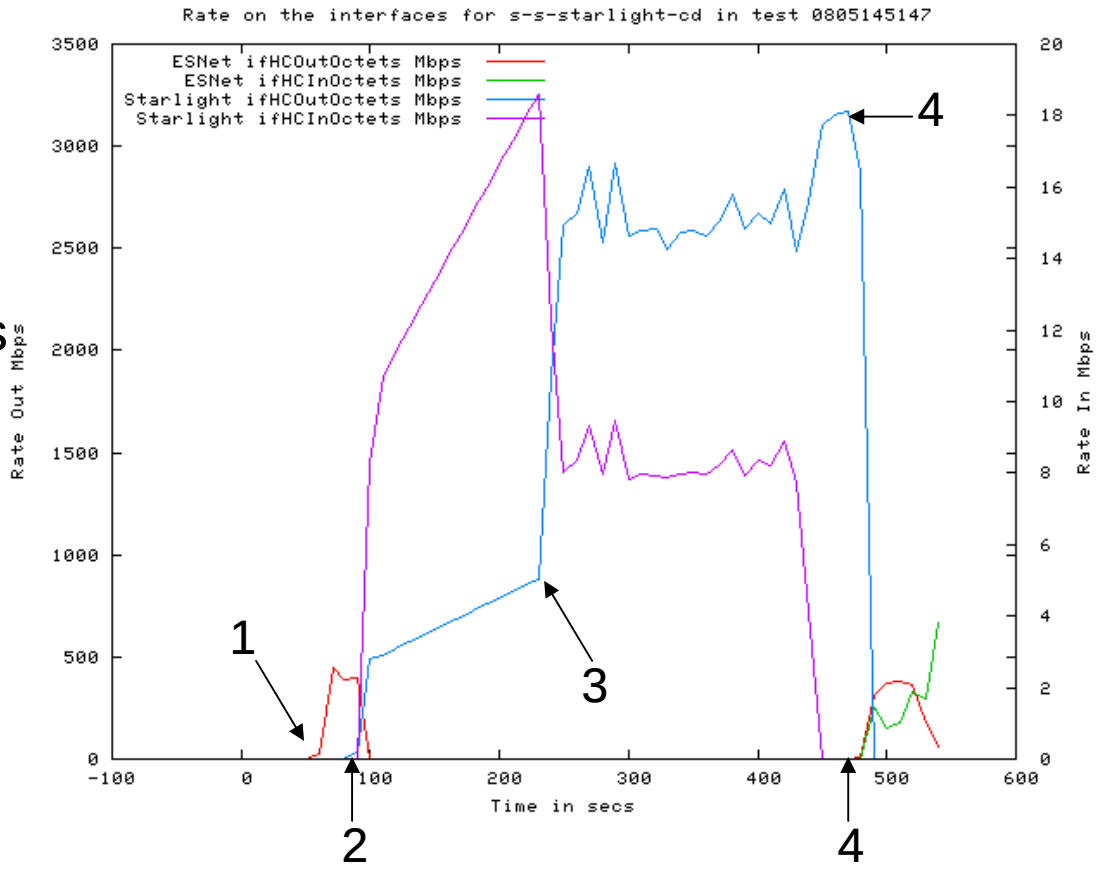
Path Switching Experiment

- Continuous GridFTP transfers:
  - Path toggled at 15 min intervals by Lambda Station servers
  - 1GE NIC, dual 750Mhz source
    - 18 TCP streams
    - Not optimized for performance
- Each path appears reasonably consistent:
  - “Choppiness” of blue (ESnet) path due to OC12 congestion
  - Red (UltraLight) spikes are periodic system resource issue
- Overall, impact of dynamic path changes were not significant



LSiperf End-to-End Test

1. Data transfer started:
   - 10GE host; 5 tcp streams
   - Network path is via ESnet
     - OC12 bottleneck…
   - Path MTU is 1500B
   - Lambda Station service ticket is opened
2. Lambda Station changes network path to USN
3. Host path MTUD check detects larger path MTU
4. Lambda Station service ticket expires:
   - Network path changed back to ESnet



Future Directions

- Work in progress:
  - Lambda Station awareness in Storage Resource Manager:
    - GridFTP DSCP tagging
    - Control channel dialog with GridFTP server
- Development plans for:
  - Automated coordination/setup with WANs
  - Authentication based on GRID certificates
  - Distributable Lambda Station kit
  - Integral path switching monitoring capability
- Testing & Deployment:
  - Deploy Lambda Stations at additional sites (CERN, others?)
  - Test across different advanced technology WANs (OSCARs, others?)
  - Test with different (non-Cisco…) LAN hardware
  - Instigate extensive failure mode testing

Closing Comments

- Lambda Station Project progress:
  - In line with our project plan, and about what was expected
  - Outside interest remains very high
    - Lambda Station deals with a problem space that's still mostly avoided
    - Other uses for Lambda Station have been suggested:
      - → Dynamic bypass of perimeter firewall for high impact traffic
- Issues & Concerns Remain:
  - Complexity
  - Level of customization:
    - May result in high level of local effort to implement
    - Uncertainty about adapting to less traditional (Cisco…) environments
  - Application-awareness is a 'stubborn' issue
  - Avoiding mission-creep…

Supplemental Slide 1: LS Interface Primitives

- Informative Group:

whoami	return identity of local or remote LS
sayHello	a probe message to test a remote LS
ip2client	return site and PBR-client identity based for IP address if exists
getClientInfo	return description of PBR-client
getKnownClients	return the list of known clients (site,client)
getKnownLambdas	return list of all known lambdastations
getStationParameters	return parameters of lambdatstion
NetConfigMode	return current mode of network configuring (static or dynamic)

- Service Group:

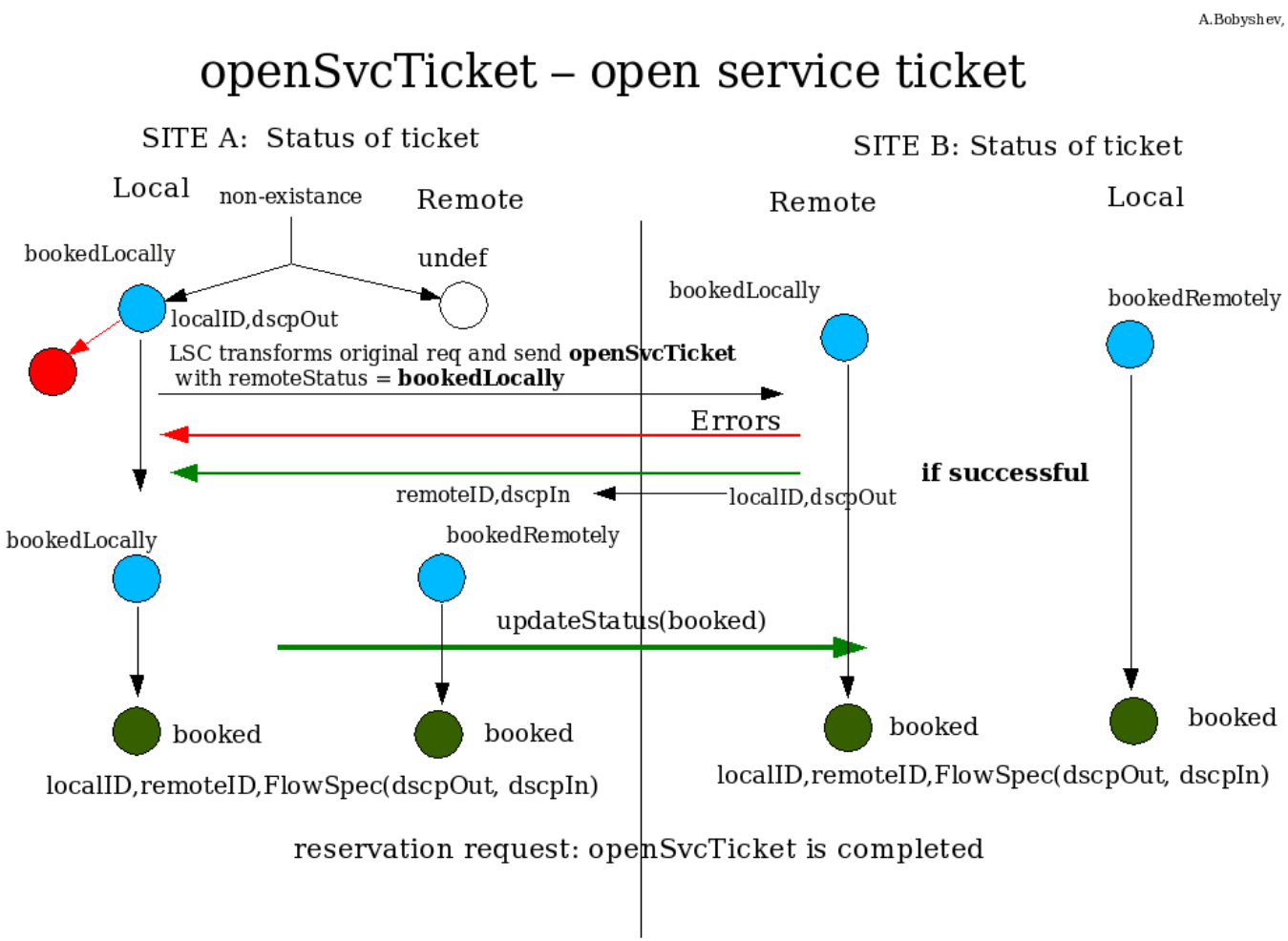
openSvcTicket	place request to establish an alternative path
cancelTicket	abnormal completion of ticket
completeTicket	normal complete of ticket before it expires otherwise expiration process will be applied
getTicket (getTicketStatus)	return ticket's parameters, e.g. its current status
getFlowSpec	return the full description of flows associated with ticket
updateFlowSpec (updateDSCP)	update flow specification
updateTicket	update ticket parameters
getMyRemoteID	get ID of remote ticket associated with localID

- Internal:

getNextDSCP	select the next free DSCP code (in case site use dynamic mode for DSCP assignment)
assignDSCP	assign DSCP to ticket
updateMyLambdaDefs	check for updates (new, remove) PBR clients
localRequest/remoteRequester	determine the authenticated requester of ticket

Supplemental Slide 2: Open Service Ticket

A.Bobyshev, 04/27/2005



Supplemental Slide 3: Ticket State Diagram