Optical Grid Networks (Invited)

P. Demeester, M. Da Lozzo, C. Develder, I. Stevens, J. Verstuyf, P. De Turck, B. Dhoedt

11th Int. Conf. on Optical Network Design and Modelling (ONDM)
Athens, Greece, May 30th, 2007

---

Evolution Towards Grid Computing:

- Personal Computer
- Cluster

---

Introduction (1):

- eScience:
  - By 2015 it is estimated that particle physicists will require exabytes (10^18) of storage and petabytes per second of computation
  - CERN's LHC Computing Grid (LCG) will start operating in 2007 and will generate 16 petabytes annually (that's ~4 Gbit/s)
Introduction (2)

- Consumer service:
  - E.g. video editing: 2Mbps/frame for HDTV, suppose effect requires 10 frames/frame, then evaluating 10 options for 10s clip is 50 Gbps (today’s high performance PC: ~ 10 Gbps).

Introduction (3)

- Move towards (mobile) thin clients
- Limited processing and storage in the terminals
- Use of server based computing/storage

Introduction (4)

- Grid opportunities ranging from academics over corporate business to home users

Introduction

Network Architecture
Routing
Multiple Domains
Dimensioning
Conclusions
**Optical Network Architecture**

- **Optical Circuit Switching (OCS)**
  - Continuous bit-stream
  - Pre-established light-paths
  - Should be dynamic

- **Optical Burst/Packet Switching (OBS/OPS)**
  - Chunks of bits in bursts/packets
  - Forwarding based on header
  - E.g., label switching, GMPLS

- **Hybrids**

**Optical Circuit Switching**

- **Pro:**
  - Guaranteed service quality once set-up (cf. reserved lambda), thus fixed latency, no jitter, etc.
  - Fixed signaling overhead, independent of (large) job size

- **Con:**
  - Signaling overhead not acceptable for relatively small jobs
  - Requires complex grooming if frequent set-up and tear-downs are to be avoided (i.e., if too slow)
  - Less flexible, dynamic than OBS/OPS, of light-path set-up and tear-down

---

**OBS/OPS**

- **Pro:**
  - Extremely flexible, dynamic
  - Inherent statistical multiplexing of available bandwidth (over multiple wavelengths)

- **Con:**
  - Packet/burst header processing overhead
  - Requires job aggregation if job size too small compared to header overhead
  - Difficult to deliver strict QoS guarantees without 2-way reservation
  - Technology not that mature (hardware)

**Hybrid OCS/OPS**

- Choosing between OCS and OBS depends on...
  - Optical technology (OBS requires faster switches, burst mode Rx/Tx and regenerators)
  - Job sizes

- **Hybrid architectures can offer a compromise**
**Challenges**

- Grid applications pose challenging demands:
  - Connections extend to application end points (rather than traditional network elements)
  - On-demand bandwidth provisioning, both immediate and advance reservations
  - Very dynamic use of end-to-end networking resources
  - Requires fast, real-time feedback for signaling and provisioning
  - Heterogeneous network, computing and storage resources in multiple domains
  - Variety in holding times and bandwidth granularity

---

**Introduction**

Network Architecture
Routing
Multiple Domains
Dimensioning
Conclusions
**Optical Grid specifics**

- Differences with "classical" optical networks or "classical" Grids:
  - Anycast routing: user generally doesn't care where job is executed
  - Burst isolation: not only network contention, also Grid resource contention
  - Failure reservation: some jobs have very loose response time requirements; others are known long beforehand

**Problem Statement**

- Problem:
  - Given a job submitted by a user to an anycast address
  - Find a set of containing at least one Grid preferably one suitable Grid site topology accepting such jobs, a route to reach it

- Sub-problems:
  - Routing/Selection strategies
  - Distributed multi-constrained routing algorithms

**Anycast SAMCRA**

- Problem:
  - Incorporation of other metrics than just Grid resource availability leads to a multi-constraint anycast routing problem
    (unicast multiple-constraint is already NP-complete)

- Our solution:
  - Introduce virtual topology to translate to unicast

**Anycast SAMCRA**

- Problem:
  - Incorporation of other metrics than just Grid resource availability leads to a multi-constraint anycast routing problem
    (unicast multiple-constraint is already NP-complete)

- Our solution:
  - Introduce virtual topology to translate to unicast
  - Use a Self-Adaptive Multiple Constraint Routing Algorithm (SAMCRA)
  - Use a novel path ordering avoiding sub-optimality and loops

---

*UNIVERSITEIT GENT*
Comparison with a shows that even distributed SAGCRA comes very close to (pseudo-optimal) acceptance rate.

Simpler heuristics, taking only 1 measure into account, do not come as close.

Introduction
Network Architecture
Routing
Multiple Domains
Dimensioning
Conclusions
Introduction
Network Architecture
Routing
Multiple Domains
Dimensioning
Conclusions
**Solution**

- **Phased approach**
  1. Determine total server capacity (analytically, Erlang-B)
  2. Determine inter-site bandwidths (simulation)
  3. Dimension link bandwidths (number of wavelengths)

**Algorithm**

- **Step 1:** total # servers $n$ follows from Erlang-B, distribution among sites (N):
  - **uniform** distribution among sites $n / N$
  - **equal proportional** to total traffic rate $n = \text{CalF}(x)$
  - **Max:** try to achieve the same (local) rate rate at each site, $n = n_{\text{max}}$ (with $n_{\text{max}} = \text{Erlang-B}(x, d)$)

- **Step 2:** inter-site rates depend on scheduling algorithm:
  - always try local site (job source) first, if busy
  - **SP:** shortest path, else, choose free server
  - **random:** randomly pick a free site
  - **m/nn:** choose site with most free servers

---

**Processing:**

- achieves highest (local) processing
- reduces server degradation (jitter, logging, overhead, etc.) in total processing

**Link dimensions:**

- intelligent scheduling (variable link bandwidth)

---

**Network Architecture**

**Routing**

**Multiple Domains**

**Dimensioning**

**Conclusions**
Conclusions

- Optical Grid architecture
  - OBS and/or hybrids seem very promising candidates
- Routing
  - Anycast routing requires deployment of new algorithms
- Multiple Domains
  - Proxy architecture to ensure scalability
- Dimensioning
  - Need to be aligned with scheduling algorithms

Challenges

- Integrated OCS/OBS/hybrid control plane
  - Interworking, migration, node architecture
- Advanced dimensioning and network planning algorithms
- Resilience
  - Job migration, protection/restoration approaches...
- Standardisation
  - E.g. GeoOBS architecture, burst format, routing protocols, inter-domain routing
- Partners: AT, CTI, Unitel, iCar, IBBT/UGent
- Supporting studies:
  - Job Demand modeling
  - CDX modeling & routing
  - QoS, multi-carrier routing (Q. physical impairments)
  - Advance Reservations
  - Control plane design
  - Simulation environment
- Joint work (ePhoton/iDEAL) between UEsssex and IBBT
- Work in progress:
  - Dimensioning
  - Scheduling
  - Node & Network design
Home

The 11th International Conference on Optical Networking Design and Modeling - ONDM 2007 - will be held in Athens, Greece on May 29-31, 2007.

ONDM conference is a major event in the rapidly growing area of optical networking addressing recent advances in the design, modeling and implementation of optical networks, including novel switching schemes and paradigms, network optimization and design, new concepts for link and control layer protocols, advanced network subsystems and node architectures, and network inter-working schemes.

A significant number of excellent invited speakers from leading academic and industrial institutions around the world will present their research work and view.

The ONDM 2007 conference program is now available.

Plenary Speakers:

Allan Willner (Univ. of Southern California)
Haruhisa Ichikawa (NTT - director)

Invited Speakers:

Keren Bergman (Columbia University)
Piero Castoldi (SSSUP)
Piet Demeester (GENT)
Andrea Fumagalli (UTD - Texas)
Maurice Gagneure (GET)
Ken-Ichi Kitayama (Osaka University)
Chunming Qiao (University at Buffalo)
Jian Wu (BUPT)
Hans Martin Foisel (T-Systems)
Dimitri Papadimitriou (Lucent-Alcatel)

A significant number of paper contributions from various research institutions around the world will also be presented. The full conference program will be posted once the final camera-ready versions of the accepted papers are received on March 23rd 2007.

The registration procedure will start on March 9th 2007. Late registration will be available after April 30th 2007.

Please visit the relative sections on the left for more updates and information on registration, accommodation and the venue.

You can download the short version of call for participation here.

Sponsors

ONDM 2007 is sponsored by the International Federation of Information Processing IFIP and supported by:

the EU ePhoton/ONE project (http://www.e-photon-one.org/)

and COST 291 action (http://www.ail.edu.gr/cost291/).

Organizer

Athens Information Technology (AIT)
P.O. Box 68
Markopoulo Ave.
GR - 19002 Pefania
Athens (Greece)

Domain Name Provided by Hellas On Line