T-LEX

Akira Kato

The Univ. of Tokyo/WIDE Project
kato@wide.ad.jp

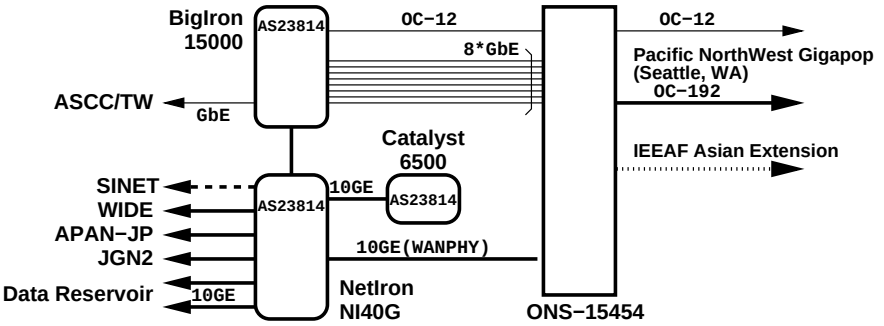
T-LEX : Tokyo Lambda Exchange

Tokyo Lambda EXchange

☆ **Operated by WIDE Project**
☆ **IEEAF Pacific circuits to Seattle are terminated**
- http://www.ieeaf.org/
- An unprotected OC-192c with a protected OC-12c
- Donated by VSNL; NTT Com. for backhaul in Tokyo
- Working closely with PNWGigapop/U of Wash.

☆ **Interface with Japanese R&E Networks**
☆ **A base for possible extension to Asia**
- Not available at this moment

Configuration of T-LEX



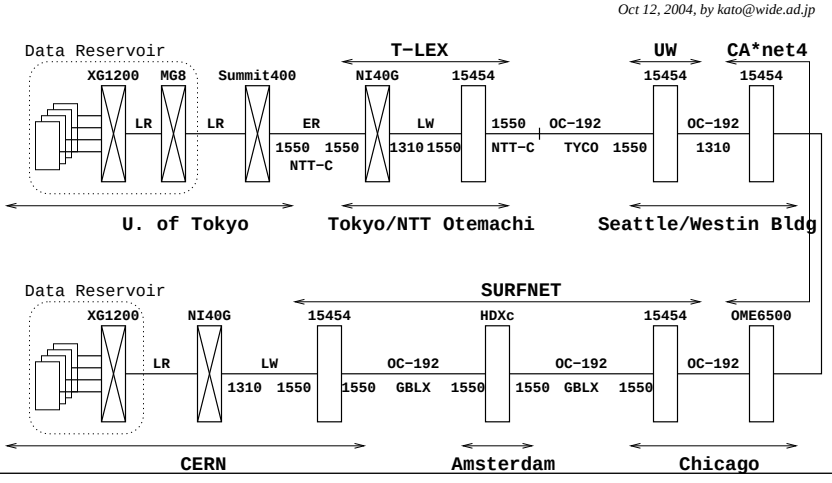
☆ **Japanese domestic access**

- An L2 switch converts WANPHY/LANPHY
- 10GE LANPHY
  - WIDE, APAN-JP, JGN2, (SINET)

The first activity in T-LEX

☆ **The first international OC-192c to Europe**

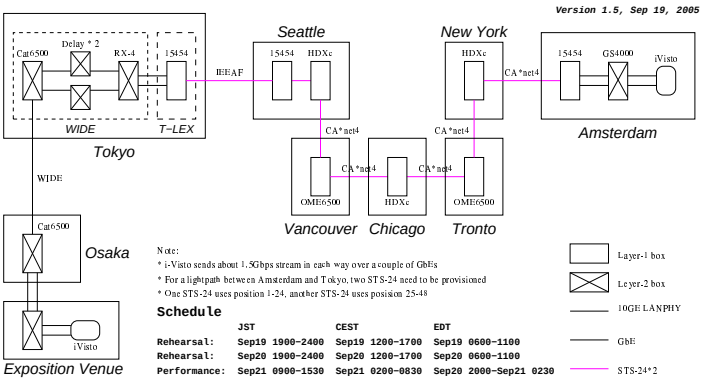
- Oct 2004, Univ. of Tokyo to CERN
- 10GE WANPHY between Tokyo and Geneva



Recent Activities in T-LEX (1)

☆ Aichi World Exposition Closing Event

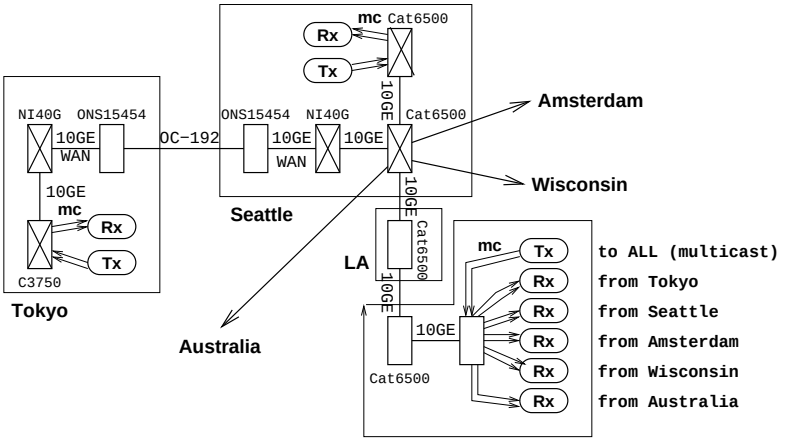
- Details was given by Hiroshi Esaki



Recent Activities in T-LEX (2)

☆ iGRID2005

- US118 N-way uncompressed HDTV transmission
- Uncompressed HDTV from 5 locations to San Diego
- Uncompressed HDTV IPv4 Multicast to 5 locations



CHEP06

☆ **A High Energy Physics conference**
- Held in Mumbai, India in February 2006

☆ **VSNL International donated 4 * OC-3c**
- Between Tokyo and Mumbai for 2 weeks

☆ **Various demonstrations was planned**

☆ **It was not successful in networking**
- Gear delivery to Mumbai was delayed
  - Due to custom reason
- No time to debug the networks
  - Observed packet loss, etc

CineGrid

☆ **4K-Cinema transmission**
- As described by Tom DeFanti
- Layer-2 path between UCSD/CalIt2 and Keio/DMC
  - Via P-Wave, Seattle, IEEAF, T-LEX, and WIDE
  - Jumbo frame enabled
- One demonstration has done in Aug 2006
- Several other trial is planned

Internet2 LSR

☆ **Land Speed Records**
- http://lsr.internet2.edu/
- Partial rule says
  - Distance measured by L3 points
  - Must spawn at least an operational network link
  - IPv4/IPv6 single TCP/multiple TCPs
  - At least 10% greater than previous record
  - Maximum distance is 30,000km

☆ **A Light Path itself doesn't contribute to LSR**
- It is a layer-1 (or layer-2) service
- Need to put layer-3 devices in the middle

Data Reservoir LSRs

☆ **Data Reservoir chaired by Prof. Hiraki**
- http://data-reservoir.adm.s.u-tokyo.ac.jp/

☆ **Original motivation**
- Provide a system for efficient data transfer
  - for scientific applications
  - over a long-fat pipe
- iSCSI based system was developed
- Its record in 2003
  - 6.8Gbps disk-to-disk over 9800 mile
  - With a large number of TCP sessions

☆ **How much performance single TCP marks?**
- Now practical as 10GE is getting popular

Data Reservoir LSRs

☆ **First LSR approved : Nov 9, 2004**
- CERN to Pittsburgh through Tokyo
- IPv4 single TCP, TOE
- 7.21Gbps over 20,645km, 148.8 Pbm/s

Data Reservoir LSRs

☆ **Second LSR : Dec 24, 2004**
- Tokyo-Chicago-Amsterdam-Chicago-[Seattle]-Tokyo
- Starlight performed L3 function in one-way
- 7.21Gbps over 30,000km, 216.3 Pbm/s
- Xmas Eve : minimum congestion in Abilene

☆ **First IPv6 LSR : Oct 29, 2005**
- Tokyo-Chicago-Tokyo-Seattle-Tokyo
- 5.58Gbps over 30,000km, 167.4 Pbm/s
- No TOE was available

☆ **Second IPv6 LSR : Nov 14, 2005 during SC|05**
- Seattle-Tokyo-Chicago-Amsterdam-Seattle
- 6.96Gbps over 30,000km, 208.8 Pbm/s
- No TOE was available

Data Reservoir LSRs

☆ **Latest IPv4 LSR : Feb 20, 2006**
- Seattle-Tokyo-Chicago-Amsterdam-Seattle
- NIC with PCI-X2.0
  - Bandwidth of PCI-X doubled
- 8.80Gbps over 30,000km, 264.1 Pbm/s
- This was the last LSR single IPv4 TCP
  - 9.68Gbps required to break the record
  - It's impossible with a single OC-192c

☆ **Things to be done**
- IPv6 single TCP, >7.66Gbps
- Disk-to-disk transfer for real applications

LSR Trophies

The lessons

☆ **While DR got LSRs, they are still "art"**
- Reproducibility is low
  - Part of the links is operational network
  - Background traffic cause additional jitter

☆ **Microscopic traffic monitoring in the middle required**
- Examine the flow
- Packet loss (non-contiguous seq#), jitter

☆ **TGNLE-1 developed by Prof. Hiraki's group**
- A 10GE box with FPGA
- It was used for "pace" the traffic in LSRs
- Cisco URP awarded as Monitorable Lambda Exchange
- Development is in progress

Other issues in T-LEX

☆ **ONS-15454 has only 4 high-speed slots**
- Additional capacity might be required
- Especially for Asian extension
- An optical switch may be necessary

☆ **No sophisticated control plane is installed**
- UCLP
- TL1 toolkit through web

Summary of T-LEX

☆ **An IEEAF circuits termination**
- Used for various experiments and demonstrations
- LSRs with extensive help from GLIF community

☆ **Microscopic traffic monitoring is being developed**
- At 10GE with FPGA supported by Cisco URP

☆ **Introduction of control plane is planned**
- TL1 toolkit through web?
- Need to configure the L2 switch as well

☆ **Possible future experiments**
- 20Gbps by collaboration with KRLight
  - when JP-KR circuit upgraded to 10GE