Ultra High Definition Media over Optical Networks (CineGrid)

Cees de Laat

University of Amsterdam
Contents

1. Use cases CineGrid & Networks
2. Formats - Numbers - Bits
3. Global Lambda Integrated Facility
4. A LightPath
5. Transport Protocol issues
6. End System Issues
7. Network Storage
8. Q/A
CineGrid Mission

To build an interdisciplinary community that is focused on the research, development, and demonstration of networked collaborative tools to enable the production, use and exchange of very-high-quality digital media over photonic networks.

http://www.cinegrid.org/
Keio/Calit2 Collaboration: Trans-Pacific 4K Teleconference

Like High-Def? Here Comes the Next Level

By JOHN MARKOFF
Published: September 26, 2005

The New York Times

Keio University
President Anzai

Sony
NTT
SGI

UCSD
Chancellor Fox

Used
1Gbps
Dedicated

iGrid 2005
First Remote Interactive High Definition Video Exploration of Deep Sea Vents

Source John Delaney & Deborah Kelley, UWash
US and International OptIPortal Sites
The “Dead Cat” demo
SC2004 & iGrid2005

Produced by:
Michael Scarpa
Robert Belleman
Peter Sloot

Many thanks to:
AMC
SARA
GigaPort
UvA/AIR
Silicon Graphics, Inc.
Zoölogisch Museum
1. Use cases CineGrid & Networks
2. Formats - Numbers - Bits
3. Global Lambda Integrated Facility
4. A LightPath
5. Transport Protocol issues
6. End System Issues
7. Network Storage
8. Q/A
Note for space reasons the following formats are not shown:

- 4096x2160 (DCI specification for 4K)
- 7680x4320 (NHK specification for 8K)
### Format - Numbers - Bits (examples!)

<table>
<thead>
<tr>
<th>Format</th>
<th>X</th>
<th>Y</th>
<th>Rate /s</th>
<th>Color bits/pix</th>
<th>Frame pix</th>
<th>Frame MByte</th>
<th>Raw Stream Gbit/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>720p HD</td>
<td>1280</td>
<td>720</td>
<td>60</td>
<td>24</td>
<td>921600</td>
<td>2.8</td>
<td>1.3</td>
</tr>
<tr>
<td>1080p HD</td>
<td>1920</td>
<td>1080</td>
<td>30</td>
<td>24</td>
<td>2073600</td>
<td>6.2</td>
<td>1.5</td>
</tr>
<tr>
<td>2k</td>
<td>2048</td>
<td>1080</td>
<td>24</td>
<td>36</td>
<td>2211840</td>
<td>10</td>
<td>1.2</td>
</tr>
<tr>
<td>SHD</td>
<td>3840</td>
<td>2160</td>
<td>30</td>
<td>24</td>
<td>8294400</td>
<td>25</td>
<td>6.0</td>
</tr>
<tr>
<td>4k</td>
<td>4096</td>
<td>2160</td>
<td>24</td>
<td>36</td>
<td>8847360</td>
<td>40</td>
<td>7.6</td>
</tr>
</tbody>
</table>

Note: this is excluding sound and these numbers are raw uncompressed data rates (frames*pix*bit/pix) excluding framing overhead!
Formats - Numbers - Bits

• Formats:
  – uncompressed from camera - UMF 3/4 GBytes/sec
  – jpeg2000 NTT 300 - 700 Mbit/s
  – uncompressed TIFF 1.2 GB/s, 4.3 TB/h
  – compressed in DXT 300 - 800 Mbit/s

• Warning; do not compress away the science!

• Storage requirement
  – Holland festival shooting uncompressed about 12 TByte
Number, numbers and more numbers!

• **Digital Motion Picture for Audio Post-Production**
  - 1 TV Episode Dubbing Reference 1 GB
  - 1 Theatrical 5.1 Final Mix 8 GB
  - 1 Theatrical Feature Dubbing reference 30 GB

• **Digital Motion Picture Acquisition**
  - 6:1 up to 20:1 shooting ratios
  - 4k @ 24 FPS @ 10bit/color: ~48MB/Frame uncompressed
  - ~8TB for Finished 2 Hr Feature

• **Digital Dailies**
  - HD compressed MPEG-2 @ 25Mb/s
  - Data Size: ~22GB for 2 Hours

• **Digital Post-production and Visual Effects**
  - Terabytes, Gigabytes, Megabytes To Select Sites Depending on Project

• **Digital Motion Picture Distribution**
  - Film Printing in Regions
    - Features ~8TB
    - Trailers ~200GB
  - Digital Cinema to Theatres
    - Features ~200 - 300GB DCP
    - Trailers ~2 - 4GB DCP

• **Online Download**
  - Features ~1.3GB
  - TV Shows ~600MB
Contents

1. Use cases CineGrid & Networks
2. Formats - Numbers - Bits
3. Global Lambda Integrated Facility
4. A LightPath
5. Transport Protocol issues
6. End System Issues
7. Network Storage
8. Q/A
GLIF Mission Statement

• GLIF is a world-scale Lambda-based Laboratory for application and middleware development on emerging LambdaGrids, where applications rely on dynamically configured networks based on optical wavelengths

• GLIF is an environment (networking infrastructure, network engineering, system integration, middleware, applications) to accomplish real work
CENIC Connects to 10Gb Research and Education Networks Nationwide and Worldwide
Contents

1. Use cases CineGrid & Networks
2. Formats - Numbers - Bits
3. Global Lambda Integrated Facility
4. A LightPath
5. Transport Protocol issues
6. End System Issues
7. Network Storage
8. Q/A
What is a LightPath

• A LightPath is a circuit like connection that connects end systems to each other. This uses usually the same infrastructure as the Internet, but a LightPath gets dedicated resources next to Internet.

• A LightPath can be a combination of:
  – A color in a fiber (Lambda)
  – Sonet/sdh circuit in a sonet infrastructure
  – Vlans and dedicated ports in an ethernet switch
  – Etc.

• Aim is to get predictable and knowable connection characteristics

• Let us look at examples setups used recently!
Very constant and predictable!
Network for “4K Pure Cinema” Trial

DCP is directly transferred from GDMX in LA to distribution centers in Japan via fiber network. Within Japan, DCP is distributed from the distribution centers to TOHO theaters. Key is distributed from Osaka center, based on the contract between WB Japan and TOHO cinemas.

**Diagram Details:**
- DCP transferred from GDMX in LA to NTT distribution centers in Japan via fiber network.
- Key distributed from Osaka center based on contract with TOHO cinemas.
- NTT’s Fiber Network transmits DCP from Osaka to Yokosuka, then to Tokyo.
- From Tokyo, DCP is distributed to TOHO theaters via NTT’s Fiber Network.
- Distribution centers and theaters are connected with 1Gbps speed.

**Key Areas:**
- **Tokyo:** Distribution Center (NTT), Theater (TOHO), NTT’s Fiber Network.
- **Osaka:** Distribution Center (NTT West), Theater (TOHO).
- **Yokosuka:** Distribution Center (NTT), Theater (TOHO).
- **Seattle:** CENIC CineGrid: 1Gbps.
- **Los Angeles:** GDMX (WBEI), Color Management QC, J2K Compression, Encryption, Wrapping.
- **Burbank:** WB Studio (WBEI), *Global Digital Media Xchange*

**Effective Dates:**
Effective 2005 - 2007
Holland Festival  
CineGrid 2007  
19-21 June 2007  
Drawing by Alan Verlo, et al.
Current Links & Available Links for Kyoto Prize Events

[Network diagram with various nodes and links, including specific IP addresses and labels like WIDE, IEEAF, JGN2, CANARIE, and others.]
Contents

1. Use cases CineGrid & Networks
2. Formats - Numbers - Bits
3. Global Lambda Integrated Facility
4. A LightPath
5. Transport Protocol issues
6. End System Issues
7. Network Storage
8. Q/A
Internet Transport Protocols

• **IP = Internet Protocol**
  – Connectionless packet transport service
  – Datagrams of max 64 kByte
  – Can be fragmented down the way
  – Packets can get lost, duplicated or out of order!

• **TCP/IP = Transmission Control Protocol**
  – Reliable byte-stream over potentially unreliable packet service
  – Connection oriented, exactly once and in order, end to end duplex

• **UDP = User Datagram Protocol**
  – Packet service up to 64 kByte
  – Connectionless, unidirectional, L2 switches may start flooding
  – Unreliable delivery, can get out of order, duplicated, lost
Flow control vs Congestion control

• Flow control
  – To prevent a fast sender overflowing a slow receiver
  – Receiver signals sender so it can adapt

• Congestion control
  – Traffic jams in the Internet: packets may get lost
  – For TCP protocol control loops via ack’s and ICMP packets
  – TCP is friendly protocol, can adapt but performance usually takes severe hit
  – RTT is reaction and recovery time
Windows and buffering for reliable protocols

• Round Trip Time (rtt) is time it takes to send a shortest message and get the answer back (unix tool ping)
• That is the shortest time the sender can know that traffic arrived at the other end
• Sender can only discard old data after receiving ack’s
• Lightspeed in fiber = 200000 km/s
• 100 km = 200 km round trip = 1/1000 sec = 1 ms rtt
  – Amsterdam - Geneve ≈ 20 ms
  – Amsterdam - Chicago ≈ 90 ms
  – Amsterdam - San Diego ≈ 160 ms
  – Amsterdam - Tokyo ≈ 250 ms
  – Amsterdam - Sydney ≈ 300 ms
# Buffer space

\[ \text{Window} = \text{RTT} \times \text{BW} \]

<table>
<thead>
<tr>
<th>RTT</th>
<th>100 Mbit/s</th>
<th>1 Gbit/s</th>
<th>10 Gbit/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12.5 kB</td>
<td>125 kB</td>
<td>1.25 MB</td>
</tr>
<tr>
<td>2</td>
<td>25 kB</td>
<td>250 kB</td>
<td>2.5 MB</td>
</tr>
<tr>
<td>5</td>
<td>62.5 kB</td>
<td>615 kB</td>
<td>6.15 MB</td>
</tr>
<tr>
<td>10</td>
<td>125 kB</td>
<td>1.25 MB</td>
<td>12.5 MB</td>
</tr>
<tr>
<td>20</td>
<td>250 kB</td>
<td>2.5 MB</td>
<td>25 MB</td>
</tr>
<tr>
<td>50</td>
<td>625 kB</td>
<td>6.25 MB</td>
<td>62.5 MB</td>
</tr>
<tr>
<td>100</td>
<td>1.25 MB</td>
<td>12.5 MB</td>
<td>125 MB</td>
</tr>
<tr>
<td>200</td>
<td>2.5 MB</td>
<td>25 MB</td>
<td>250 MB</td>
</tr>
<tr>
<td>500</td>
<td>6.25 MB</td>
<td>62.5 MB</td>
<td>625 MB</td>
</tr>
<tr>
<td>1000</td>
<td>12.5 MB</td>
<td>125 MB</td>
<td>1250 MB</td>
</tr>
</tbody>
</table>
TCP Tuning (if not auto-tuning)

- 1 Gbit/s on 160 ms RTT (= Amsterdam - San Diego):
  - `sysctl -w kern.ipc.maxsockbuf=50000000`
  - `sysctl -w net.inet.tcp.sendspace=21000000`
  - `sysctl -w net.inet.tcp.rcvspace=21000000`
  - `sysctl -w net.inet.udp.maxdgram=57344`
  - `sysctl -w net.inet.udp.rcvspace=74848`
  - `sysctl -w net.local.stream.sendspace=32768`
  - `sysctl -w net.local.stream.recvspace=32768`
  - `sysctl -w kern.ipc.somaxconn=512`
  - `sysctl -w net.inet.tcp.mssdflt=1460`
  - `sysctl -w net.inet.tcp.delayed_ack=2`
  - `sysctl -w net.inet.tcp.rfc1323=1`
  - `sysctl -w net.inet.tcp.rfc1644=1`
  - `sysctl -w net.inet.tcp.newreno=1`
Other issues & protocols

- When using UDP watch for bottleneck!
- About 10 other non-standard protocols
- FAST TCP
  - Modified receiver algorithms
- RBUDP
  - Runs on top of UDP, simple back-off and retransmission scheme
Contents

1. Use cases CineGrid & Networks
2. Formats - Numbers - Bits
3. Global Lambda Integrated Facility
4. A LightPath
5. Transport Protocol issues
6. End System Issues
7. Network Storage
8. Q/A
End System Issues

• Ethernet card interface to computer bus system
  – PCI-X
    • 32/64 bit 66/133/266 MHZ -> about 8 Gbit/s max in 133 MHZ mode
  – PCI-Express
    • 2.5 Gbit/s per lane, 4, 8, 16 lanes

• Memory organization

• CPU cache
  – Effect when things go out of cache (small windows, etc.)

• CPU core
  – Takes 1 core to handle network (affinity may help)

• Disk raid subsystem
  – raid0 twice as fast as raid5
  – One disk does typically 40 MB/s write, 60 MB/s read
Contents

1. Use cases CineGrid & Networks
2. Formats - Numbers - Bits
3. Global Lambda Integrated Facility
4. A LightPath
5. Transport Protocol issues
6. End System Issues
7. Network Storage
8. Q/A
CineGrid Exchanges

• The Cinegrid community has 3 main exchanges
  – located in:
    • Tokyo at Keio University/DMC
    • Amsterdam at University of Amsterdam
    • San Diego at CALIT2
  – implemented as store and forward servers
  – connected by 10 Gbit/s lightpaths
  – capacities 20 - 100 TByte
  – for persistent store of content for experiments
  – streaming, workflows, processing, dubbing, etc.
RDF describing Infrastructure

Application: find video containing x, then trans-code to it view on Tiled Display

RDF/CG → RDF/CG → RDF/ST → RDF/NDL → RDF/NDL → RDF/NDL → RDF/CPU → RDF/VIZ

content

content

PG&CdL
Contents

1. Use cases CineGrid & Networks
2. Formats - Numbers - Bits
3. Global Lambda Integrated Facility
4. A LightPath
5. Transport Protocol issues
6. End System Issues
7. Network Storage
8. Q/A