CineGrid Networking

Requirements for Digital Media Exchange over WAN

Cees de Laat

University of Amsterdam
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

Keio University
President Anzai

UCSD
Chancellor Fox

Used 1Gbps Dedicated
Sony NTT SGI

Keio/Calit2 Collaboration: Trans-Pacific 4K Teleconference

Like High-Def? Here Comes the Next Level
By JOHN MARKOFF
Published: September 26, 2005

Keio University
President Anzai

UCSD
Chancellor Fox

Used 1Gbps Dedicated
Sony NTT SGI
US and International OptIPortal Sites

SIO
NCMIR
USGS EDC
NCSA & TRECC
SARA
KISTI
AIST
RINCON & Nortel
TAMU
UCI
UIC
CALIT2
Formats - Numbers - Bits
### Format - Numbers - Bits (examples!)

<table>
<thead>
<tr>
<th>Format</th>
<th>X</th>
<th>Y</th>
<th>Rate</th>
<th>Color bits/pix</th>
<th>Frame #pix</th>
<th>Frame MByte</th>
<th>Flow MByte/s</th>
<th>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>170</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>190</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>240</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>48</td>
<td></td>
<td></td>
<td></td>
<td>480</td>
<td>2.4</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>750</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>960</td>
<td>7.6</td>
</tr>
</tbody>
</table>

Note: this is excluding sound!
Note: these are raw uncompressed data rates!
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
Summary

• Different applications, different traffic modes:
  – Conferencing - full duplex
    • typically low latency compressed, low jitter
  – from camera/production to (deep) store/forward
    • rough compression, needs transcoding, near real time
  – from store to theater or tiled display
    • compressed or uncompressed
  – from movie production to editing facility
    • no compression!
  – shared working environments
    • low jitter, no compression
## Buffer space in reliable protocols

$$\text{Window} = \text{RTT} \times \text{BW}$$

<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>
Introducing COCE

CineGrid
Open
Content
Exchange
DAS-3 Cluster Architecture

- **85 (40+45) compute nodes**
- **Fast interconnect**
  - 10 Gb/s Ethernet lanphy
  - 10 Gb/s Ethernet lanphy
  - 8 * 10 Gb/s from bridgenodes to Glimmerglass and NORTEL 8600
  - To SURFnet
- **Local interconnect**
  - 1 Gb/s Ethernet
- **To local University**
- **UvA-node**
- **96 TByte 2 Thumpers**
Amsterdam CineGrid  S/F node

“COCE”

DAS-3 @ UvA

DP AMD processor nodes

comp node

head node

bridge node

bridge node

bridge node

bridge node

bridge node

bridge node

storage node

2 * 48 TByte

suitcees & briefceees

10 Gbit/s

NetherLight, StarPlane
the cp testbeds
and beyond

GlimmerGlass
photonic switch

Rembrandt Cluster
total 22 TByte diskspace
@ LightHouse

Opteron 64 bit nodes

comp node

comp node

comp node

comp node

comp node

comp node

comp node

comp node

comp node

comp node

comp node

comp node

comp node

head node

head node

Node 41

NORTEL
8600
L2/3 switch

F10
L2/3 switch

streaming node
8 TByte

Node 41

suitcees & briefceees

10 Gbit/s

10 Gbit/s
R & D

• interface portal to storage
• interface portal to PBT enabled testbed and Netherlight / SURFnet_6.0
• near real time transcoding on DAS-3
• scalable streaming via bridgenodes
• embedding in semantic web
• Access control / security
• content management / deep storage
• disk 2 network performance
Sequential write performance on Thumper

![Graph showing sequential write performance on Thumper](image-url)
Sequential read performance on Thumper
RDF describing Infrastructure

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

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

see talk of Paola Grosso
TeraThinking

• What constitutes a Tb/s network?
• CALIT2 has 8000 Gigabit drops ?->? Terabit Lan?
• look at 80 core Intel processor
  – cut it in two, left and right communicate 8 TB/s
• think back to teraflop computing!
  – MPI makes it a teraflop machine
• Program desired behavior into the network!
• TeraApps programming model supported by
  – TFlops       ->   MPI / Globus
  – TBytes       ->   OGSA/DAIS
  – TPixels      ->   SAGE
  – TSensors     ->   LOFAR, LHC, LOOKING, CineGrid, ...
  – Tbit/s       ->   ?

ref Larry Smarr & CdL
User Programmable Virtualized Networks allows the results of decades of computer science to handle the complexities of application specific networking.

- The network is virtualized as a collection of resources
- UPVNss enable network resources to be programmed as part of the application
- Mathematica, a powerful mathematical software system, can interact with real networks using UPVNss
Mathematica enables advanced graph queries, visualizations and real-time network manipulations on UPVNs

Topology matters can be dealt with algorithmically

Results can be persisted using a transaction service built in UPVN

Initialization and BFS discovery of NEs

Needs["WebServices"]
<<DiscreteMath`Combinatorica`
<<DiscreteMath`GraphPlot`
InitNetworkTopologyService["edge.ict.tno.nl"]

Available methods:
{DiscoverNetworkElements, GetLinkBandwidth, GetAllIpLinks, Remote, NetworkTokenTransaction}

Getting neighbours of: 139.63.145.94
Internal links: {192.168.0.1, 139.63.145.94}
(...)
Getting neighbours of: 192.168.2.3
Internal links: {192.168.2.3}

Transaction on shortest path with tokens

nodePath = ConvertIndicesToNodes[ShortestPath[g, Node2Index[nids,"192.168.3.4"], Node2Index[nids,"139.63.77.49"]], nids];

Print["Path: ", nodePath];
If[NetworkTokenTransaction[nodePath, "green"] == True, Print["Committed"], Print["Transaction failed"];]

Path: {192.168.3.4,192.168.3.1,139.63.77.30,139.63.77.49}
Committed

ref: Robert J. Meijer, Rudolf J. Strijkers, Leon Gommans, Cees de Laat, User Programmable Virtualized Networks, accepted for publication to the IEEE e-Science 2006 conference Amsterdam.