Research at LARC-USP
E-Science, Cloud & Big Data Projects

Fernando Redigolo

LARC – USP

Laboratory of Computer Architecture and Networks

Department of Computer and Digital System Engineering

USP University of São Paulo – Brazil

São Paulo

~ 12 Million inhabitants

University of São Paulo

- Created in 1934
- **11** campi (4 – city of São Paulo).
- **89** University Divisions.
- **92.064** students (undergrad, grad and extension).
- **5.860** professors.
- **16.837** administrative staff.
- **249** undergraduation programs.
- **239** graduation Programs

- Consistently the best-positioned Latin America institution on worldwide rankings (eg. 51-60 on The Times Higher Education World Reputation Ranking 2015)

Public University, founded in 1934

Source: Anuário Estatistico 2013

LARC-USP

- Computer Networks and Architecture Lab
  - **Created in 1993.**
  - **8 professors**
  - **50** collaborators, distributed among Doctorate, Master and Undergrad **students and full-time researchers**

- Main fields of interest
  - Security
  - **High-Definition Networked Media & Visualization**
  - Wireless and Sensor Networks
  - **Advanced Internet & Applications**
  - **SDN (Software Defined Network)**
  - High-Performance Hardware For Networking
  - **Cloud Computing**

Main Partnerships

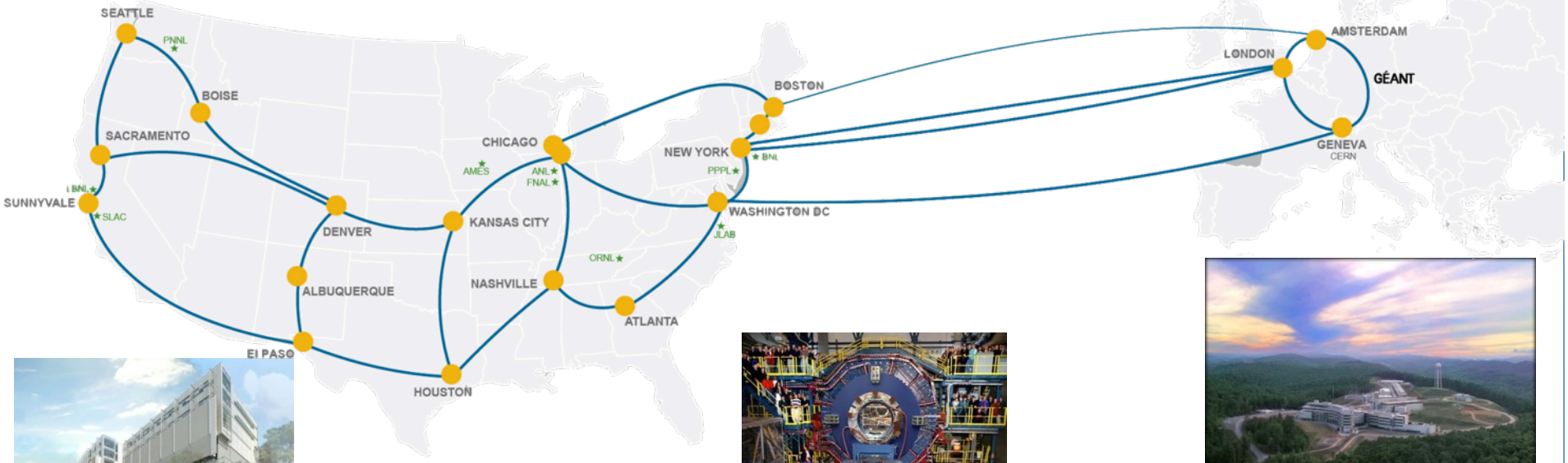
- **RNP** (National Network for Research and Education)
- **ANSP** (Academic Network of the State of São Paulo)
- **FIU / AMPath**
- **Ericsson Research** Sweden, Canada, Finland, Brazil
- **IBM Research** – T.J. Watson
- **Financial Institutions**: Bradesco & Itaú Banks, Scopus Tecnologia
- **Petrobras** (Brazilian Oil Company)
- **Incor** (Heart Institute)

E-Science, Cloud & Big Data Projects

- Characteristiscs
  - **Usually Collaborative Projects** involving multiple Partners
  - **Usually infrastructure + Use Cases / Demos**

- 4 Main Areas
  - New Network Architectures
  - High-Definition Networked Media & Visualization
  - Cloud Computing
  - Big Data

Network as ~~Infrastructure~~ Research Instrument



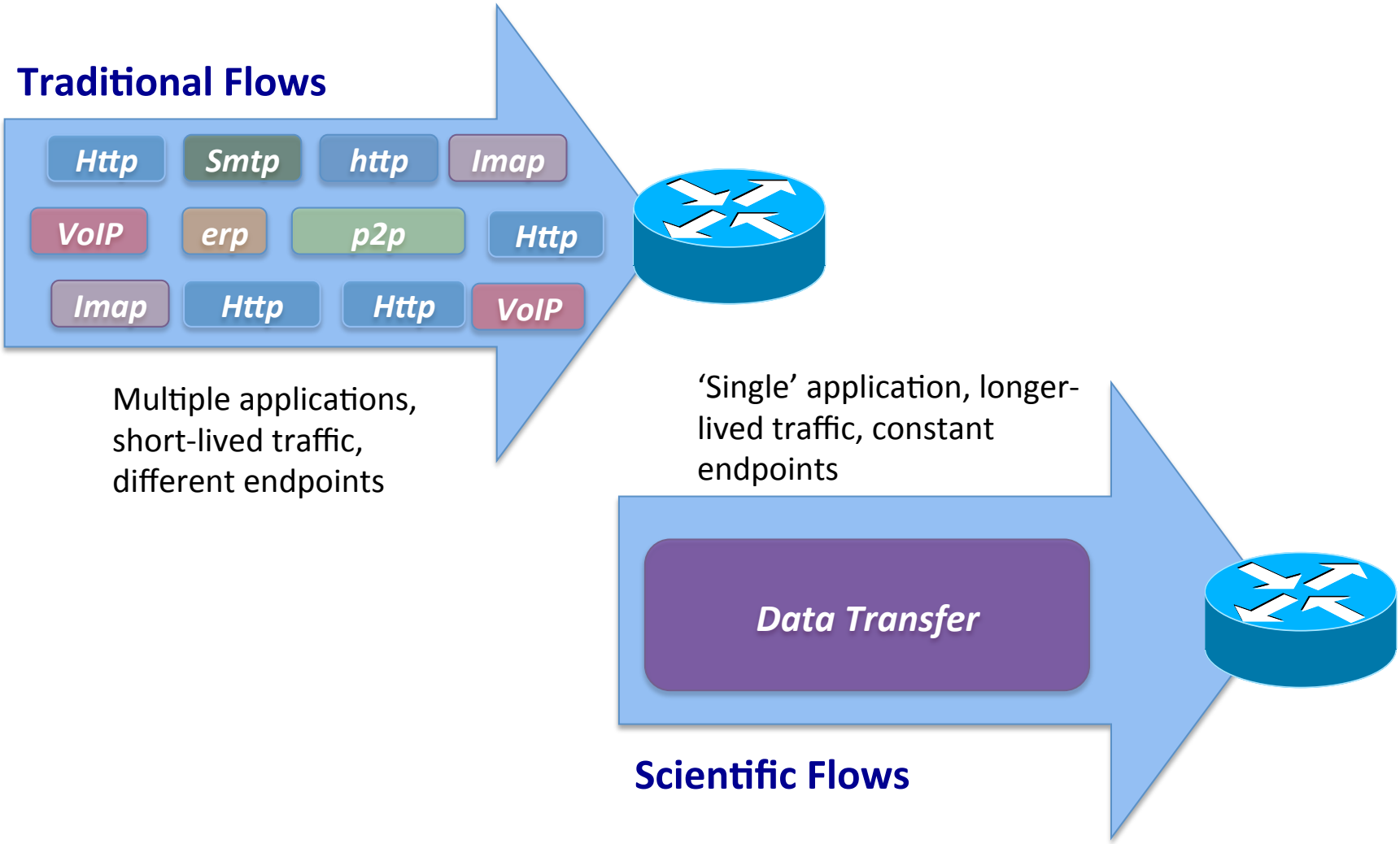
ESNet Vision: Scientific Progress **should not be limited** by physical location of instruments, people, computational resources and/or data

- No longer possible to dissociate network from scientific research process

But ...

Institutions	Application	Data set	Notes
UFRJ	Images and videos Neurology	**200GB/day**	They can not perform the research because they have connection and bandwidth problems.
UFPE	DB Vegetal Genetics	**18TB/experiment**	Wget – **3 to 4 weeks** scp - **4 to 6 weeks**
INPE	Meteorological Data	**240 MB/day**	**24 hours** to transfer
USP	LHC Alice Experiment	-	They use practically only local simulated data, as the connectivity is deemed insufficient for real-time data transfer

Why ?
Traditional Flows x Scientific Flows



Why ?
Improper Data Transfer Tools

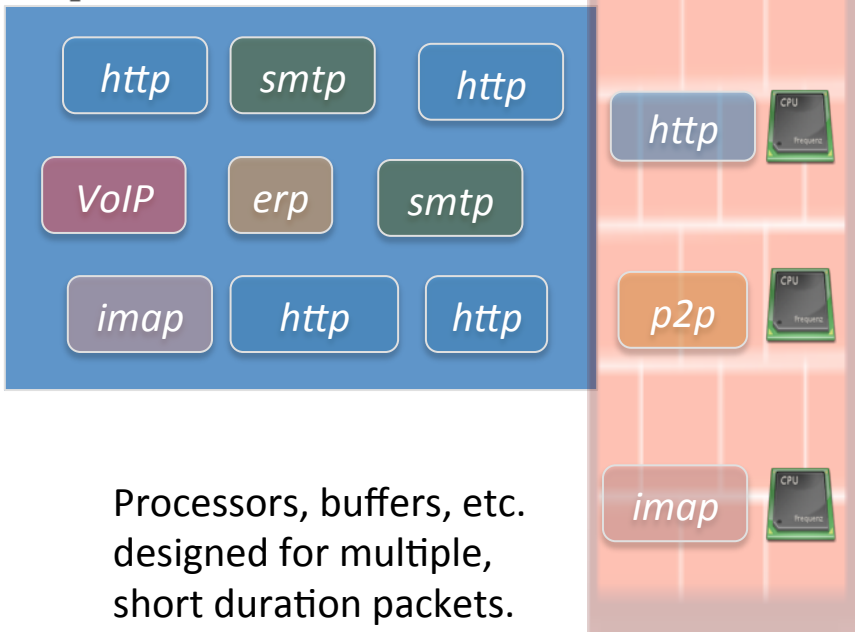
Source - https://fasterdata.es.net/data-transfer-tools/

Data Transfer Tools	Disk Architecture	Throughput
Scp	1	140 Mbps (17.5 MB/s)
HPN-scp	1	760 Mbps (95 MB/s)
HPN-scp	RAID-0	1.2 Gbps (150 MB/s)
GridFTP com 1 fluxo	1	760 Mbps (95 MB/s)
GridFTP com 1 fluxo	RAID-0	1.4 Gbps (175 MB/s)
GridFTP com 4 fluxos	RAID-0	5.4 Gbps (675 MB/s)
GridFTP com 8 fluxos	RAID-0	6.6 Gbps (825 MB/s)

Disk-to-disk transfer from Berkeley, CA to Argonne, IL. RTT = **53 ms**, bandwidth > **10Gbps**, 4-disk **RAID 0**. In order to go over 1 Gbps (125 MB/s) RAID was needed.

Traditional Flows x Scientific Flows: Firewalls

Traditional Flows



Processors, buffers, etc. designed for multiple, short duration packets.

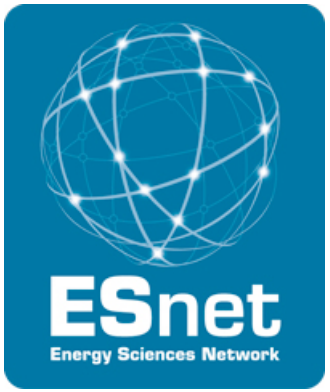
Scientific Flows

Normally, security equipments do not support large flows

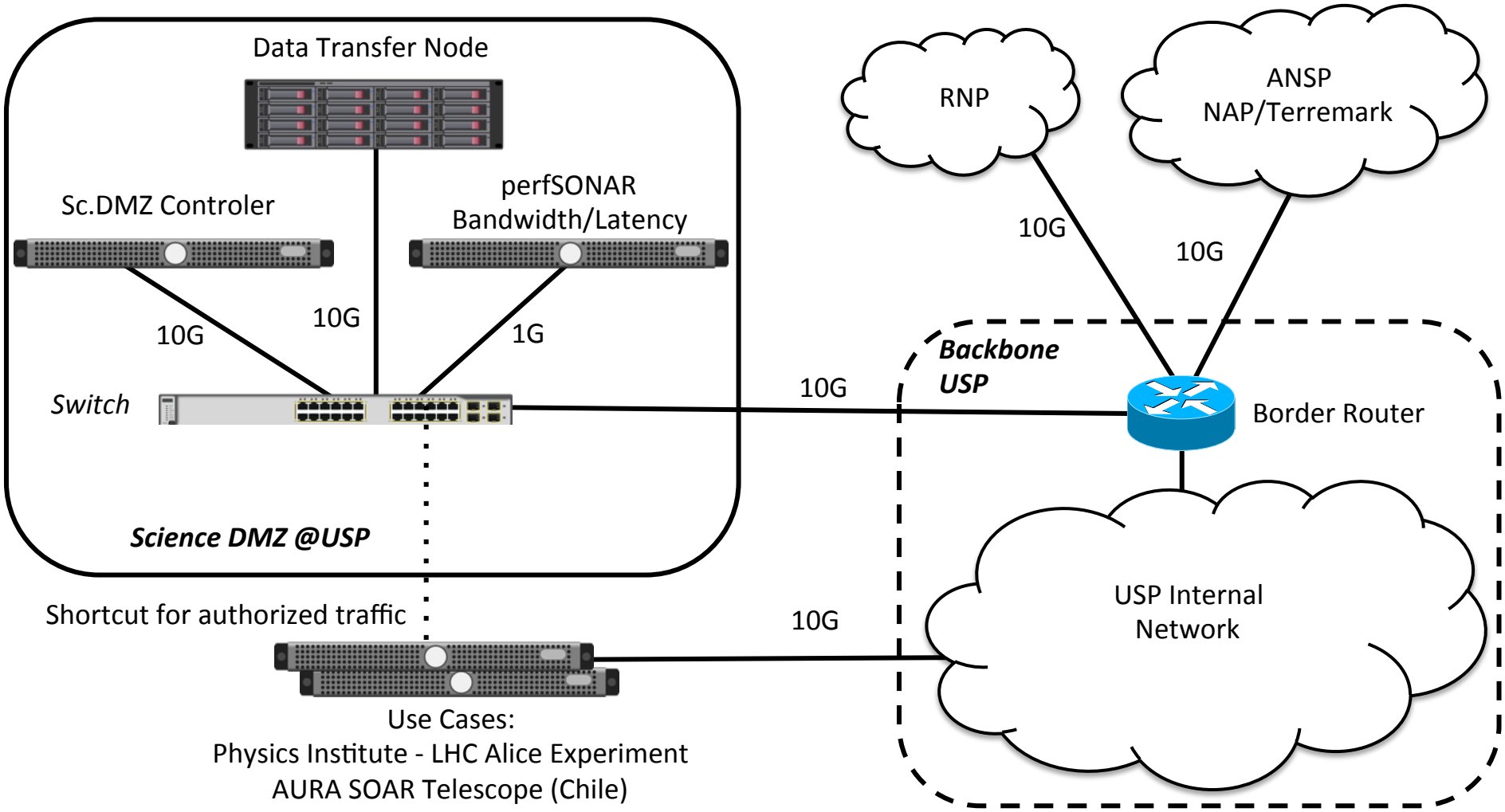
Science DMZ

What is a Science DMZ ?

- **Specialized Network Architecture** for **high-performance scientific computation**, with **differentiated policies and configuration** in comparison to production network.
- Optimized machines for content **transfer and monitoring.**
- **Priviledged Connectivity**
- Concept created by DoE ESNet



Science DMZ @USP - 2015



$(SD)^2$
Software-Defined Science DMZ

- Sponsorship: RNP (Brazilian NREN)
- Main goals:
  - Develop and Deploy a **Science DMZ Prototype with SDN/Openflow capability** at Brazilian universities.
  - Architecture **Evaluation** under **Use Case Scenarios**
    - **Metereology**
    - **Astronomy**
    - **Genetics**
    - **High-Energy Physics**

RNP

USP

CCE
CENTRO DE COMPUTAÇÃO ELETRÔNICA

Observatório Nacional

UFRJ

INPE

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA
PERNAMBUCO

$(SD)^2$ Activities - 2015

- '**Research Network-as-a-Service**'
  - Deploy a **Web Portal** for:
    - **Schedule** / **initiate** data transfer, using **different** data **transfer tools**
    - **Setup monitoring** rules for data transfer troubleshooting
    - Specify **specific network requirements** (eg. Layer 2 circuits)
    - Handle **security** (authentication, open/close TCP/UDP ports, etc.)
  - **Web Services** for mutual Science DMZ negotiations
    - Available Data Transfer Tools
    - Authorized/Unauthorized traffic

Other Networking Projects

- **USP – OSU (Ohio State University) – 2014**
  - Evaluating **long-distance collaboration** using **Science DMZ**

- **FIBRE -** Future Internet testbed for Brazilian Experimentation
  - **Testbed for Future Internet & Software-Defined Networking,** similar to **GENI** Project in US
  - **10 islands** interconnected by means of RNP (Brazilian NRE)
  - **Extensions** for adding **Cloud Computing** and **IoT** (internet of Things) **under evaluation**
  - **Integration** with US & European projects **(ProtoGENI & Fed4FIRE)**

Visualization Portal – SAGE2

Environment for high-definition Collaborative Visualization

- Based on Univ. Illinois / Hawaii **SAGE2 middleware**
- **Multiple simultaneous high-def content** streamed through the network
- **Interaction** with **users devices** and **portals** from **other institutions**
- **HTML/JS-based** application development

3x1 Portal - WRNP 2015

RNP SAGE2 Project

- Goals
  - **Collaborative work** among Brazilian Research Entities with **Tiled Displays**
  - **Disseminate** this technology to Brazilian R&D community
  - **Evaluate/Adapt its** use for Scientific Applications (Use Case Analysis)

RNP SAGE2 Project

- Coordination:
  - **Profa. Tereza Cristina Melo de Brito Carvalho**
  - **Fernando Frota Redigolo**

- Sponsorship: **RNP (Brazilian NREN)**

- Use Case Scenarios
  - Telemedicine
  - Cinema / Video
  - **Metereology**
  - **Geoprocessing**
  - Simulation

16 Mpixels 4x2 Portal – LASSU-USP

Video Streaming

- Projects involving **special audio/video transmissions**
  - Uncompressed Full HD Video
  - Compressed 4K Video
- Demos usually involving international Partners & Arts groups
- Pushing the boundaries on the networking

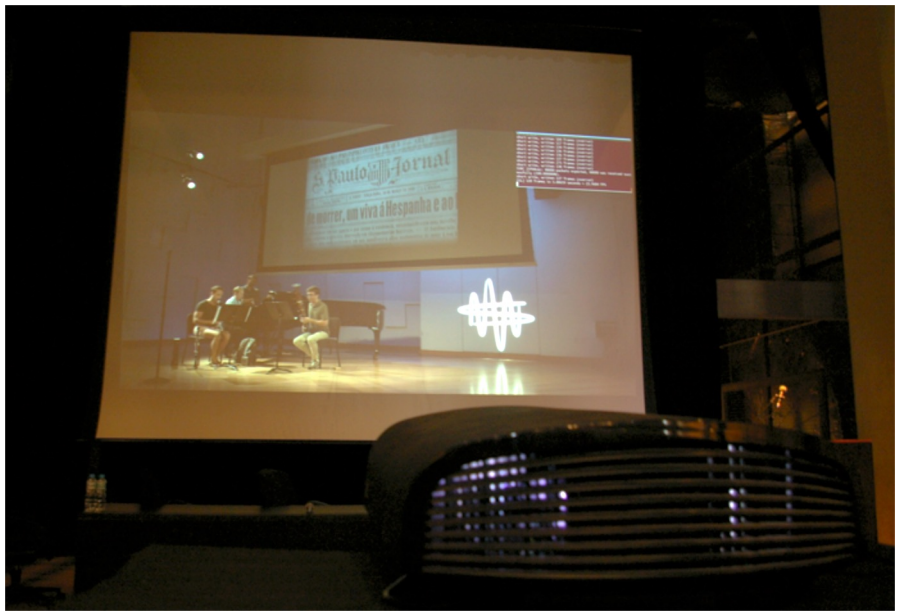
Previous Projects

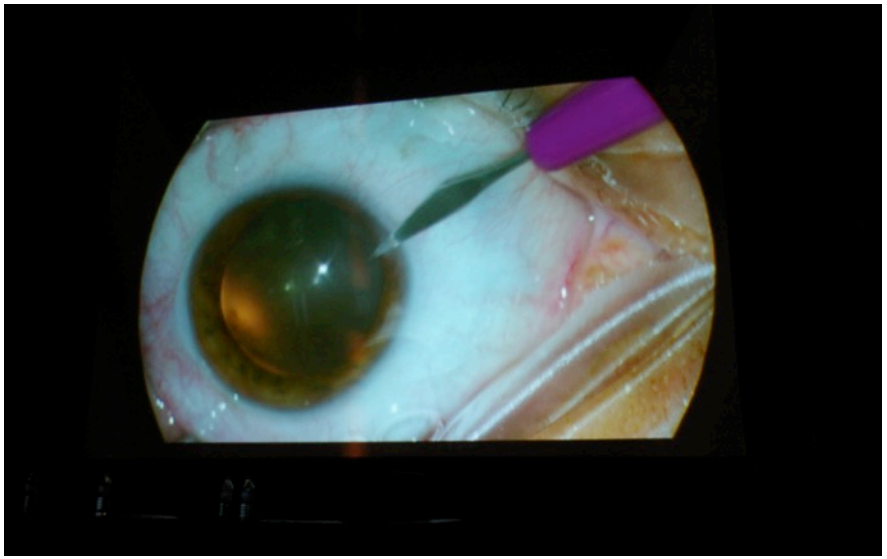
- **4K** Film **transmission** – Brazil – US – Japan (2009) + Q&A session w/ **compressed and uncompressed videoconferencing**

- **Remote Master Class of Piano** – Brazil – Spain (2009) **w/ compressed video** and **uncompressed audio**

Cinegrid Brasil 2014

- **Cinegrid**
  - Community involving Digital Cinema, Arts and Scientific Visualization over networks
- **Cinegrid Brasil – Aug/2014**
  - 2nd Regional Meeting in Brazil

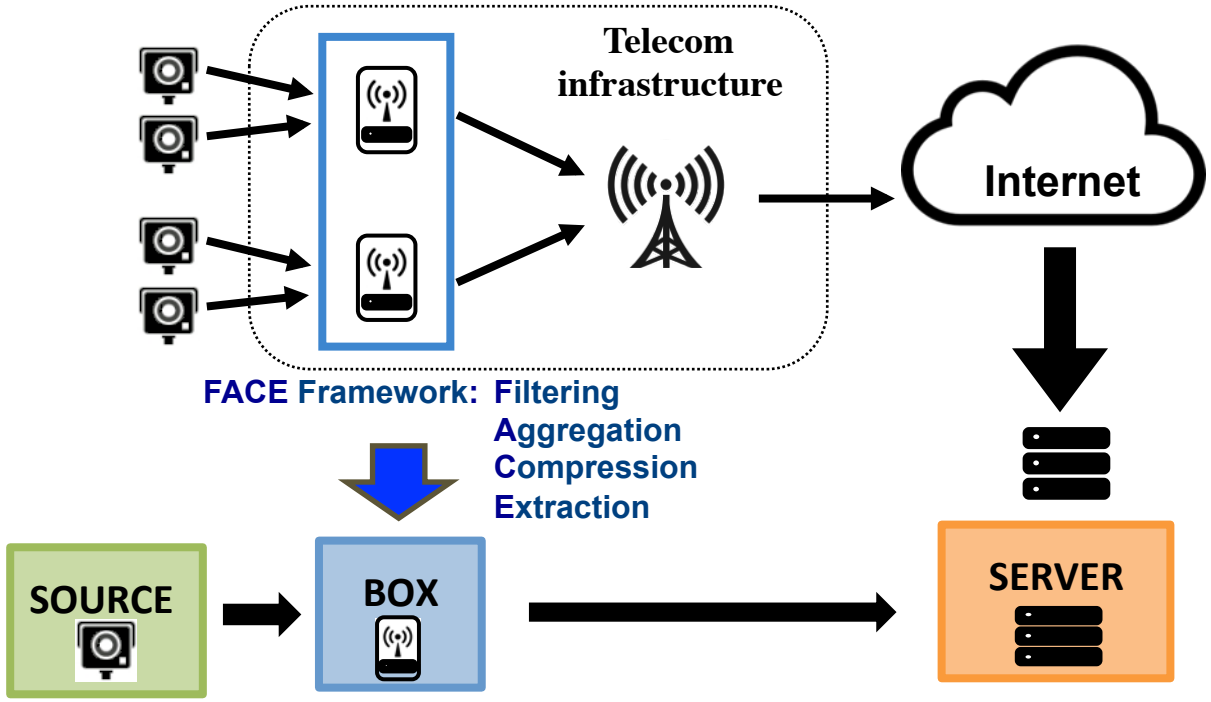




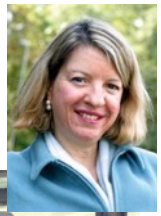
Cloud Computing

- **Projects involving Cloud Computing Infrastructure**
  - Credential Management
  - Secure Virtual Networking
  - Security SLA
  - Security Visualization

Advanced Cloud Computing Services for Telecom



General Coordination & Members



Goals:

- Distributed Cloud Computing Framework for Telco Service offering
- FACE Framework Prototype – Speed Radar / Video Surveillance with Motion Detection

Network Traffic Modelling

- Model a Company-owned satellite network
  - Model based on real captured traffic - Identify applications & traffic patterns
  - Analyze 'What-if' Scenarios
  - 1-2 TB data (packet headers only) for a 1-week capture