UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
S E C R E T A R Í A G E N E R A L
Dirección General de Cómputo y de Tecnologías de Información y Comunicación

ICT infrastructure to support scientific research

Dr. Ignacio Ania B.
(ignacio.ania@unam.mx)
February 8th, 2011

DGTIC
tic.unam.mx

Agenda

- Institutional vision, academic requirements and solution design criteria
- Internal strategy
- External strategy

DGTIC
tic.unam.mx

Institutional vision

*"The **main challenge** of our University today is to go forward in order to respond to the world's dramatic transformation,… [and] to the spreading of the new **information and communication technologies**…"*

DR. JOSÉ NARRO ROBLES (January 2008)

Guidelines to articulate an academic proposal for the 2007-2011 period

DGTIC
tic.unam.mx

Basic academic requirements

Communication and interaction
(multiple national and international groups)
[email, phone, telepresence, access to grid and high performance computing resources, storage of high volume data,…]

Enhanced capacity and coverage
(voice and data networks)

Greater mobility
(students and academic staff)
[mobile devices]

Multimedia contents
(publication and online access)
[network of digital repositories: RadUNAM]

Technology in the learning spaces

Distance learning

Safe ICT infrastructure

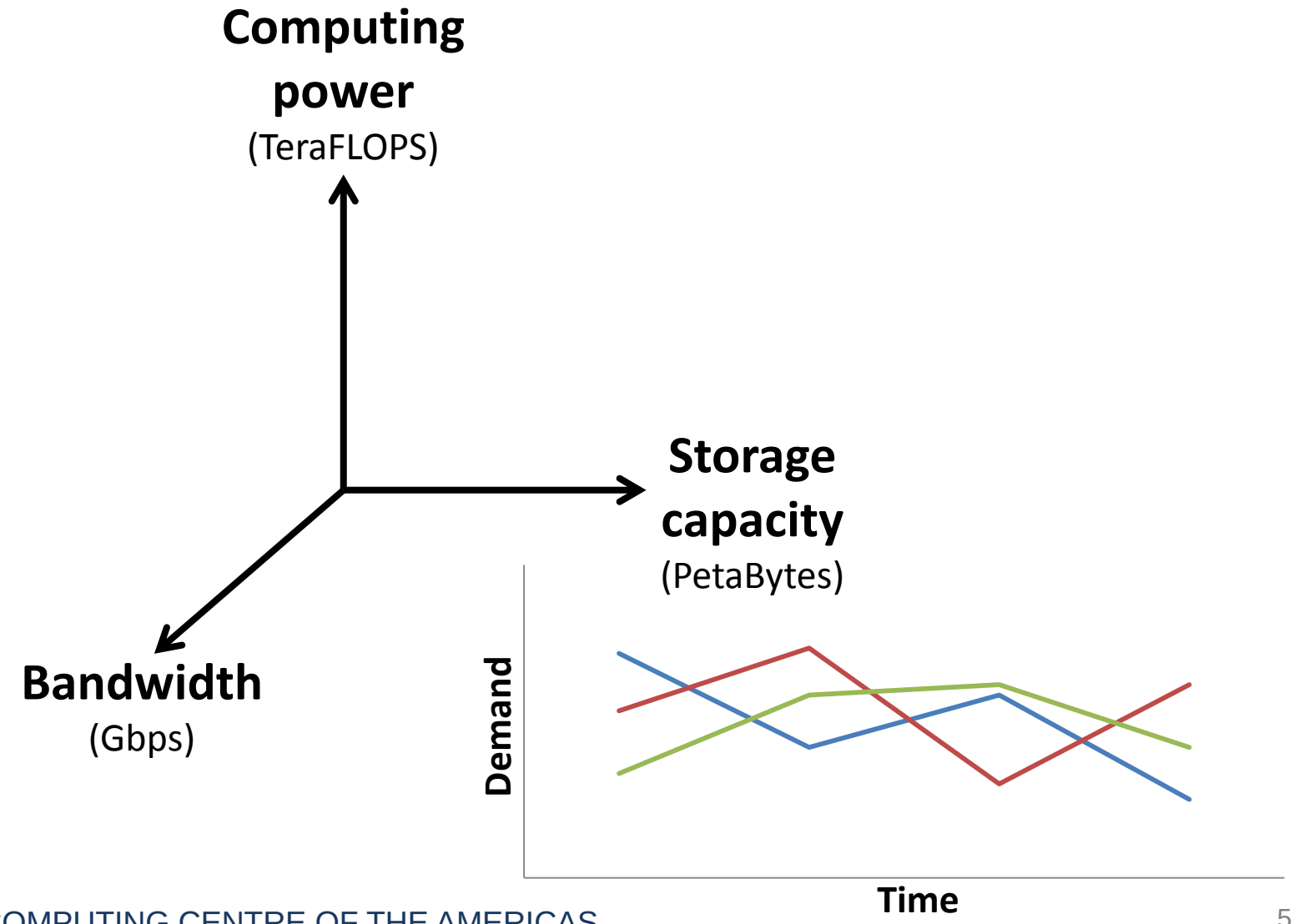
Virtual classrooms and laboratories
[streaming, video-training, podcasts,…]

Operations management continuity
(infrastructure and processes)

Video-surveillance, management efficiency,…

DGTIC
tic.unam.mx

Requirement dimensions



DGTIC
tic.unam.mx

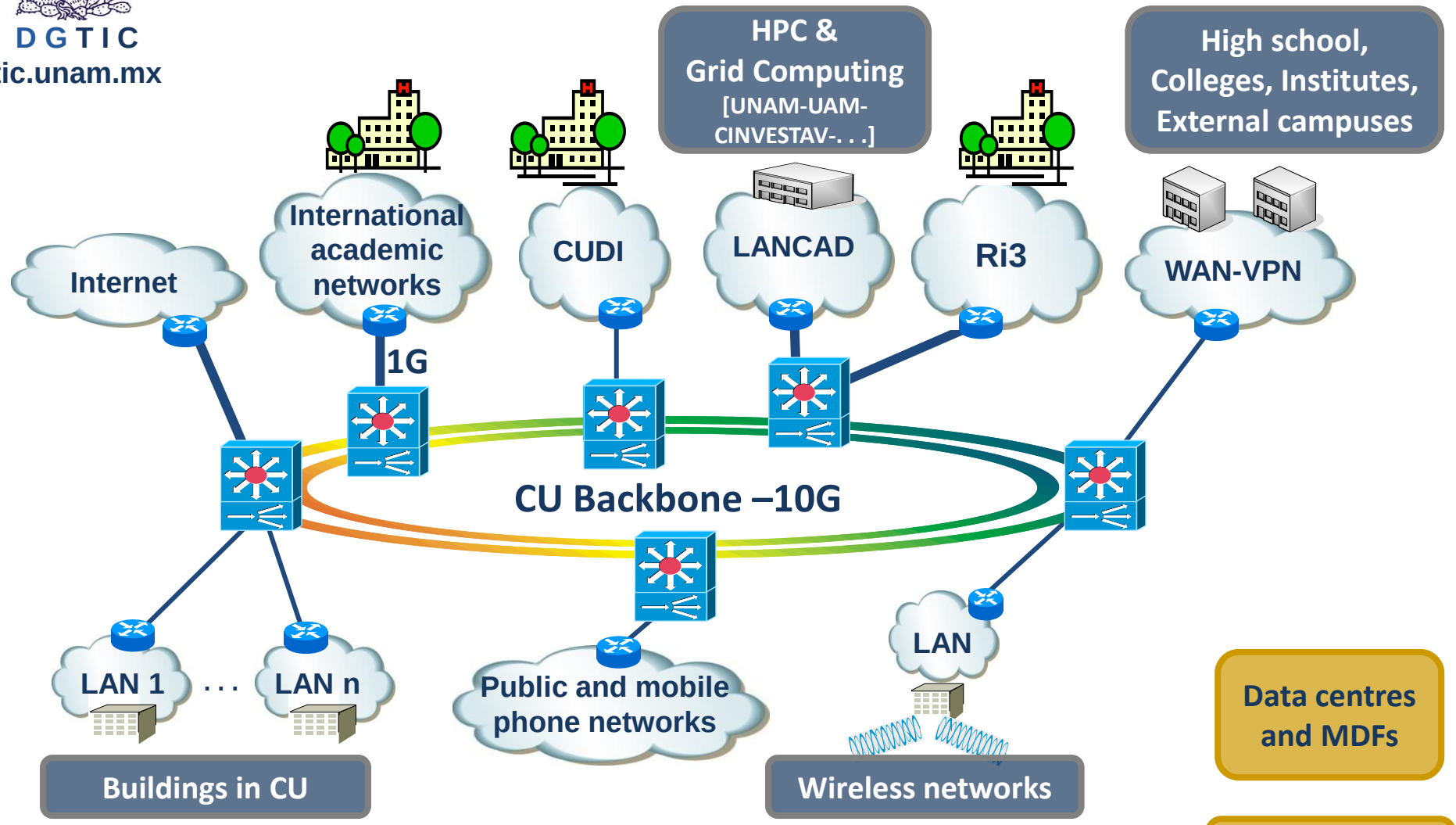
Solution design criteria

Availability	• 7/24
Reliability	• No information loss
Security	• No malware • No spam
Flexibility	• Multiple configurations
Capacity	• ≥ 50 TeraFLOPS • ≥ 10 PetaBytes • ≥ 1 Gbps
Coverage	• Central • Metropolitan • National • International

Internal strategy

Current infrastructure

DGTIC
tic.unam.mx



DGTIC
tic.unam.mx

Current status

- **Main drawbacks**
  - Servers: Widespread infrastructure among several locations
  - Data storage: High disintegration and fragmentation. No datacenter available
  - Services: hardly any virtualization
- **Consequences**
  - Idle capacity (TeraFLOPS, PetaBytes)
    - Storage: mainly 1 to 1 associations between physical processors and data storage resources. Example: 50K PCs x 100GB = 5 Pbytes. (use: 30%)
    - Processing: Underused. High percentage of CPU power on stand by
  - Multiplication of the same information. Example: attachments in emails
  - High management costs: air conditioning, electric power consumption, training,…
  - High software costs
  - Greater information security challenges

DGTIC
tic.unam.mx

New approach: UNAM cloud

- Goals:
  - Investments redeploying to get optimal ICT architecture and use
  - High availability and flexibility in ICT services
  - Increase and diversification of ICT installed capacity and services

Model A

Centralized processing

Cluster @ DGTIC

Hardware:
- Hundreds or thousands of CPU's
- Hundreds or thousands of GPU's
- Tens of TeraFLOPS
- Tens of TeraBytes RAM

Software:
- Servers virtualization
- Operating Systems images (Win, Lnx, Mac)
- Optimizers / Balancers

Centralized data storage

Hardware:
- TeraBytes / PetaBytes
- 15 KRPM / 7 KRPM
- SAN / NAS
- RAID n

Filesystems:
- NFS
- EXT1 / EXT2/ EXT3 / EXT4
- NTFS
- HPFS
- OS independent (RAW)

Datacenter @ DGTIC

RedUNAM

Physical infrastructure layer

Virtual services and infrastructure layer (SaaS, IaaS)

UNAM cloud

TIER 1 / GRID

E-mail

VPN access

ID management

Print services

Webcast / podcast

Databases

WWW services

Administrative apps

Institutional repositories

Other apps

Unified communications

GRID COMPUTING CENTRE OF THE AMERICAS

DGTIC January 2011	Project	UNAM cloud. Centralized processing and data storage
	Dirección General de Cómputo y de Tecnologías de Información y Comunicación	Dirección de Sistemas y Servicios Institucionales

Model B

Centralized processing

Cluster @ DGTIC

Hardware:
- Hundreds or thousands of CPU's
- Hundreds or thousands of GPU's
- Tens of TeraFLOPS
- Tens of TeraBytes RAM

Software:
- Servers virtualization
- Operating Systems images (Win, Lnx, Mac)
- Optimizers / Balancers

Distributed data storage

Hardware:
- TeraBytes per venue
- 15 KRPM / 7 KRPM
- SAN local / NAS global
- RAID n

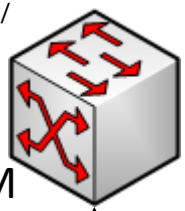
Filesystems:
- NFS
- EXT1 / EXT2/ EXT3 / EXT4
- NTFS
- HPFS
- OS independent (RAW)

Datacenter DF

Datacenter MOR

Datacenter QRO

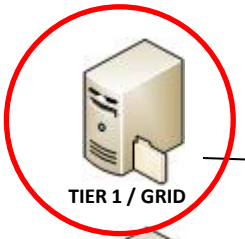
Datacenter MICH



RedUNAM

Physical infrastructure layer

Virtual services and infrastructure layer (SaaS, IaaS)



TIER 1 / GRID

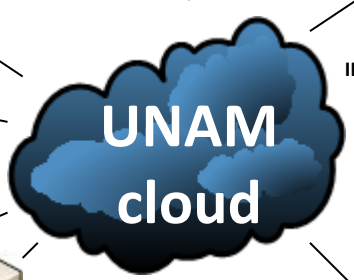
E-mail

VPN access

ID management

Print services

Webcast / podcast



UNAM cloud

Databases

WWW services

Administrative apps

Institutional repositories

Other apps

Unified communications

GRID COMPUTING CENTRE OF THE AMERICAS

DGTIC January 2011	Project	UNAM cloud. Centralized processing. Distributed data storage.
	Dirección General de Cómputo y de Tecnologías de Información y Comunicación	Dirección de Sistemas y Servicios Institucionales

Model C

Distributed processing
UNAM GRID

Cluster DF

Cluster MOR

Cluster QRO

Cluster MICH

Hardware:
- Hundreds of CPU's per venue
- Hundreds of GPU's per venue
- Local TeraFLOPS
- Local TeraBytes RAM

Software:
- Local services virtualization
- Operating Systems images (Win, Lnx, Mac)
- Local optimizers / balancers

RedUNAM

Distributed data storage

Hardware:
- TeraBytes per venue
- 15 KRPM / 7 KRPM
- SAN local / NAS global
- RAID n

Filesystems:
- NFS
- EXT1 / EXT2/ EXT3 / EXT4
- NTFS
- HPFS
- OS independent (RAW)

Datacenter DF

Datacenter MOR

Datacenter QRO

Datacenter MICH

Physical infrastructure layer

Virtual services and infrastructure layer (SaaS, IaaS)

UNAM cloud

TIER 1 / GRID

E-mail

VPN access

ID management

Print services

Webcast / podcast

Databases

WWW services

Administrative apps

Institutional repositories

Other apps

Unified communications

GRID COMPUTING CENTRE OF THE AMERICAS

DGTIC January 2011	Project	UNAM cloud. Distributed processing and data storage
	Dirección General de Cómputo y de Tecnologías de Información y Comunicación	Dirección de Sistemas y Servicios Institucionales

DGTIC
tic.unam.mx

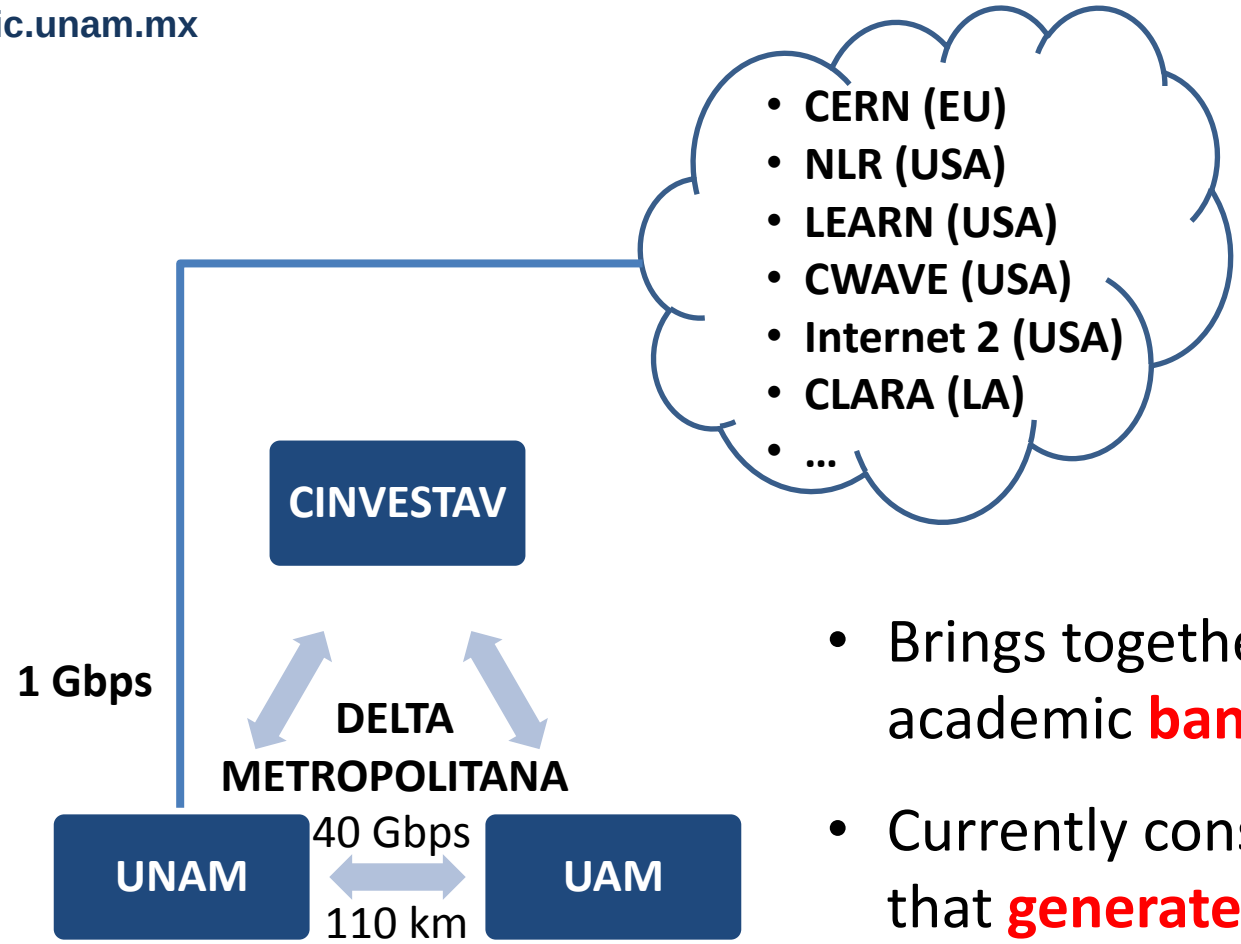
Central network

External strategy

DGTIC
tic.unam.mx

LANCAD

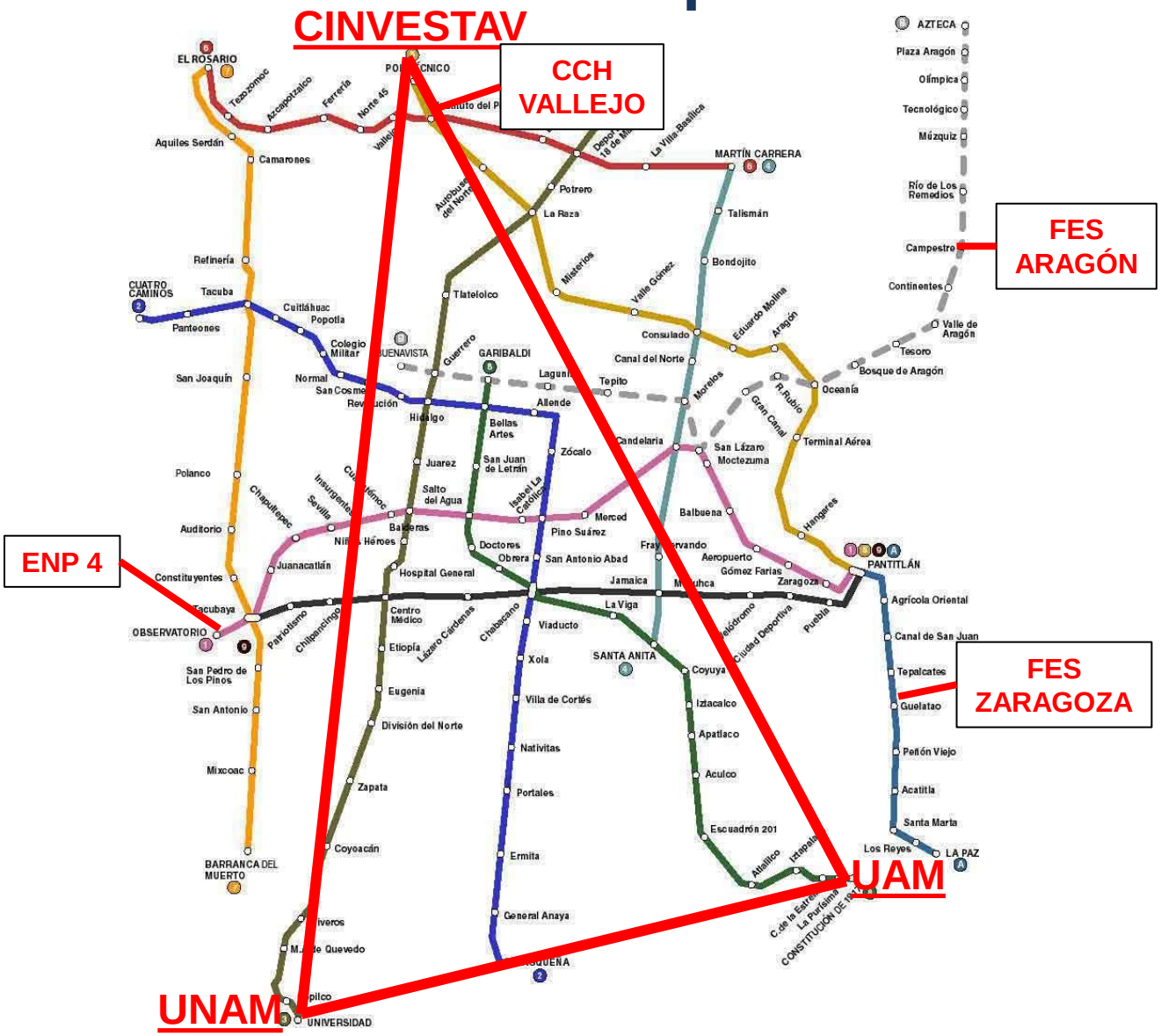
Laboratorio Nacional de Cómputo de Alto Desempeño



- Brings together the **main** Mexican academic **bandwidth consumers**
- Currently consists of 3 institutions that **generate more than half of the scientific papers** produced in Mexico

DGTIC
tic.unam.mx

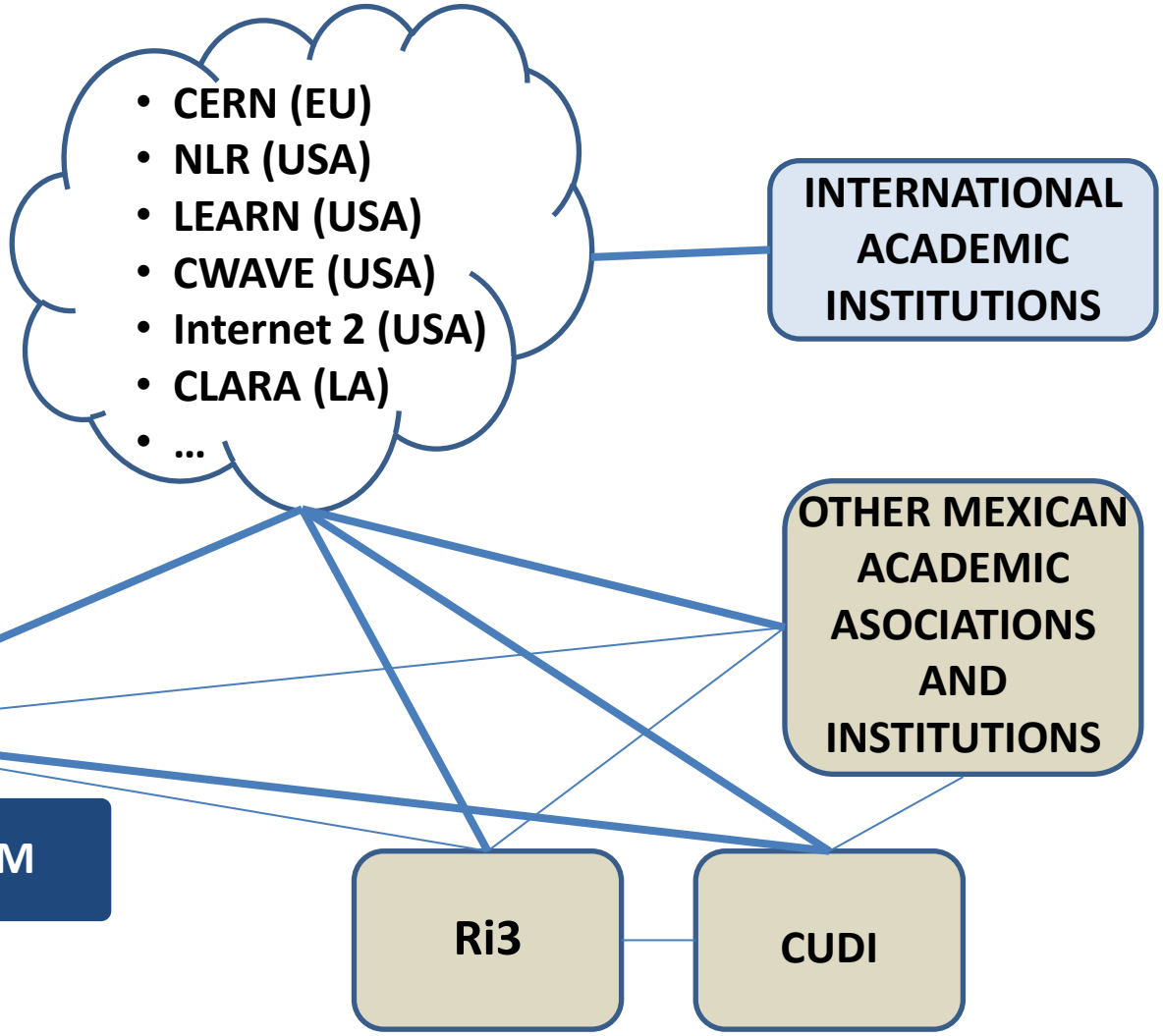
LANCAD's metropolitan network



DGTIC
tic.unam.mx

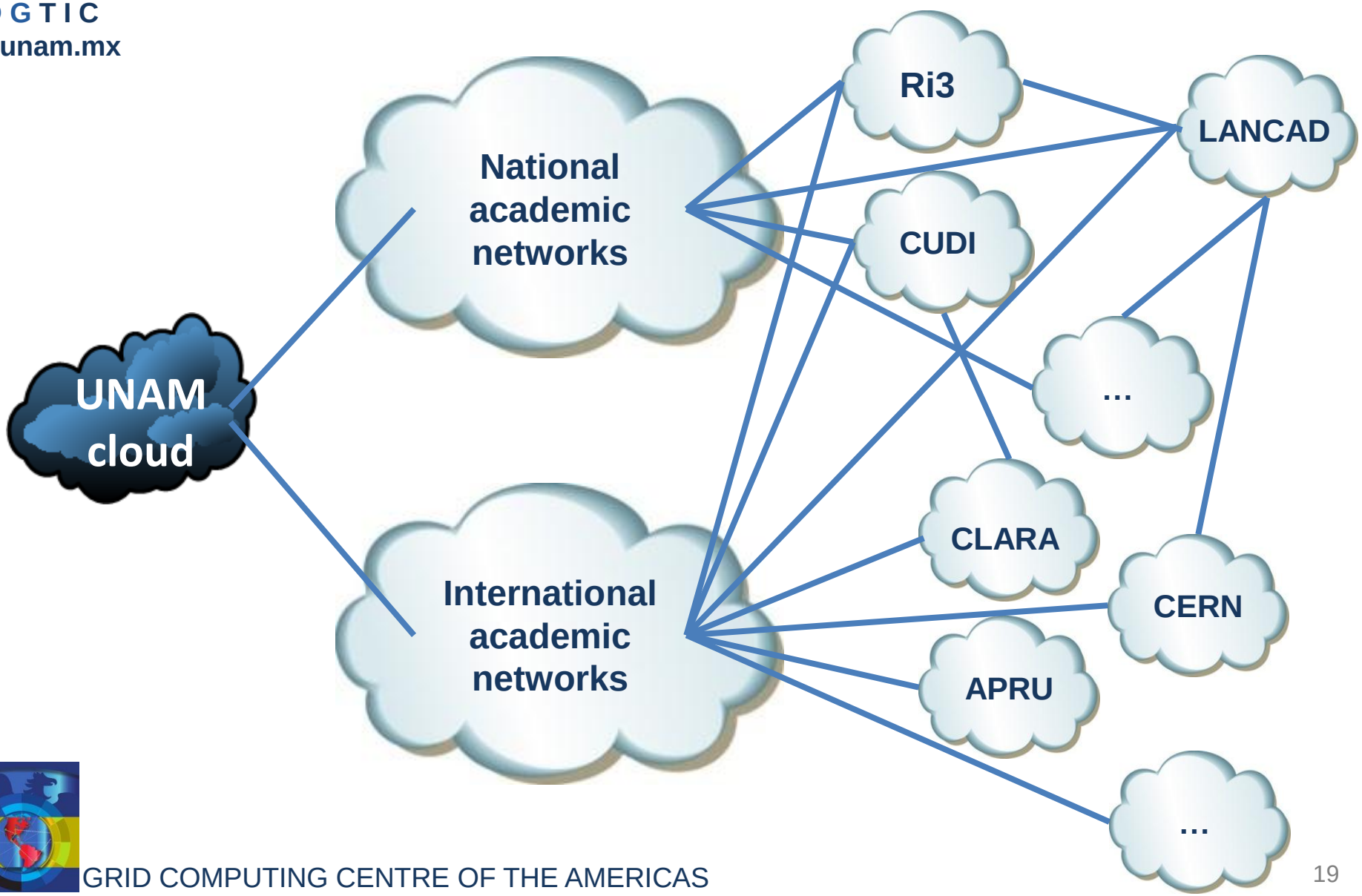
Envisioned network

The Mexican academic institutions **evolve and become stronger** by collaborating in projects that use a **shared and strong infrastructure**



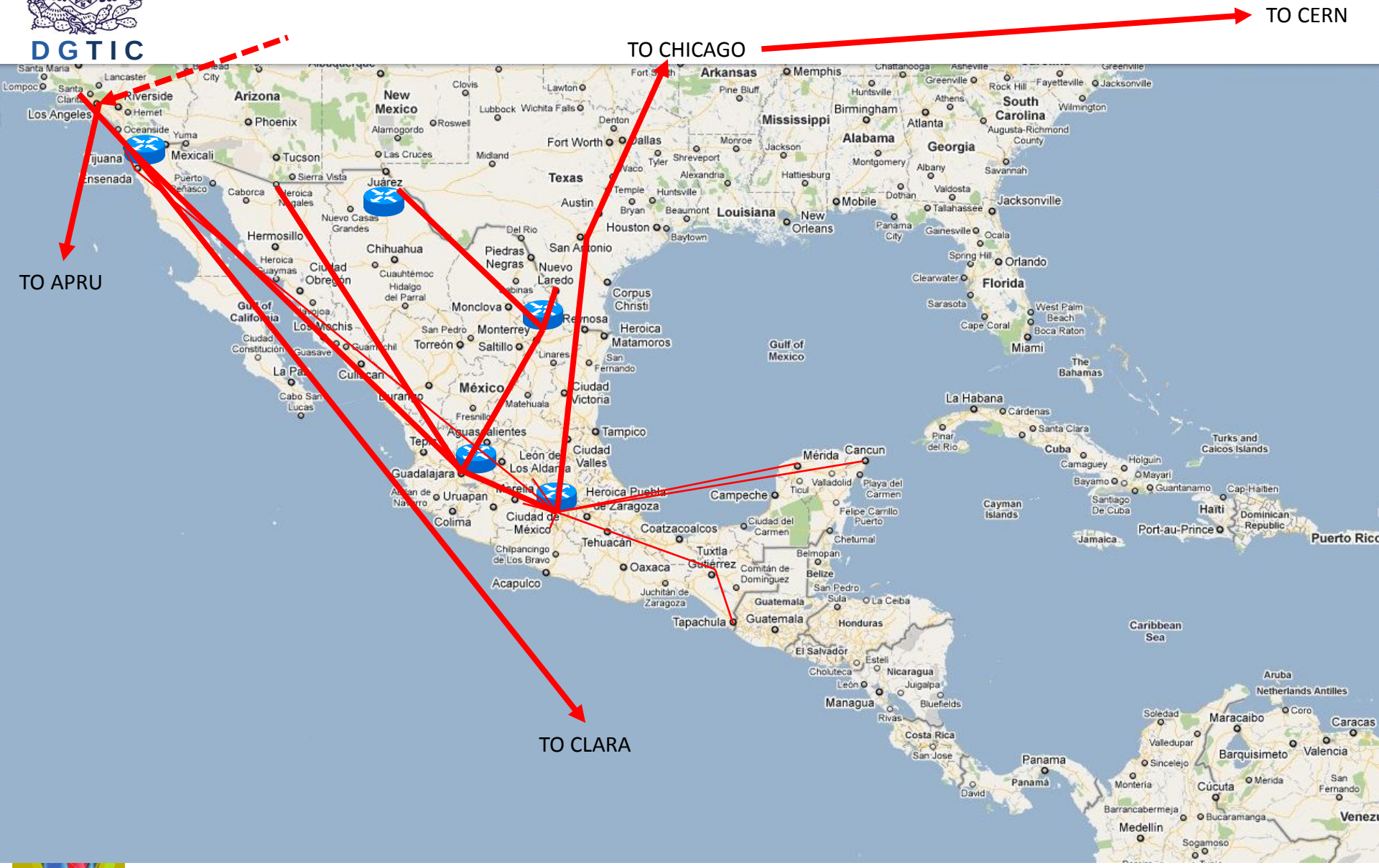
DGTIC
tic.unam.mx

Collaboration and interoperation



UNAM's (inter)national connectivity

DGTIC



DGTIC
tic.unam.mx

Grid Computing Centre of the Americas

- UNAM has many resources and characteristics that make this project feasible:
  - People
  - Infrastructure
  - Knowledge
  - Experience
  - Strategy
  - Leadership
  - Vision
  - Interest

ICT infrastructure to support scientific research

¿Questions, comments?

Dr. Ignacio Ania B.
(ignacio.ania@unam.mx)
February 8th, 2011