

Pasos estratégicos para el desarrollo de Centros de Supercómputo Nacional

Centro de Matemáticas Aplicadas y Cómputo de Alto Rendimiento del Cinvestav (Unidad Estado de México)

Jaime Klapp
ININ - Cinvestav

Objetivos

Contar con una supercomputadora en la fase inicial de aproximadamente 150 Tflops (Top 100) o Top 10 a 40 posteriormente. Fuerte componente de GPUs, 500-1000 TB HD, 10,000-20,000 núcleos, 100-150 GPUs.

Por comparación, rendimiento de supercomputadoras de UNAM, UAM-I y Cinvestav (8,18, 25 Tflops). UNAM a corto plazo se espera llegue a 100 Tflops.

Ofrecer facilidades reales de supercómputo a los investigadores del país.

Que el proyecto sea de clase mundial.

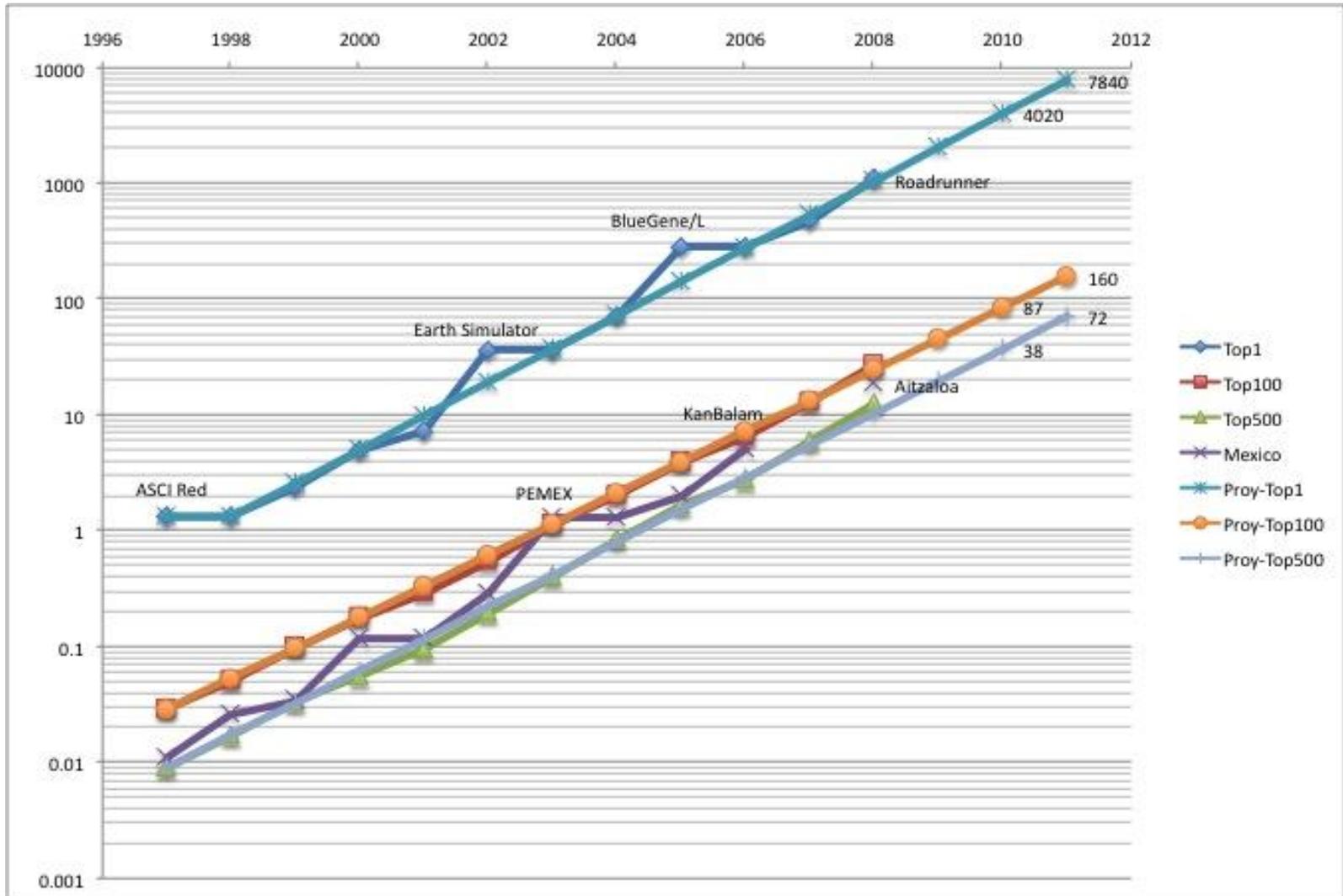
Que el centro sea sustentable.

Aplicaciones en la industria, empresas y gobierno.

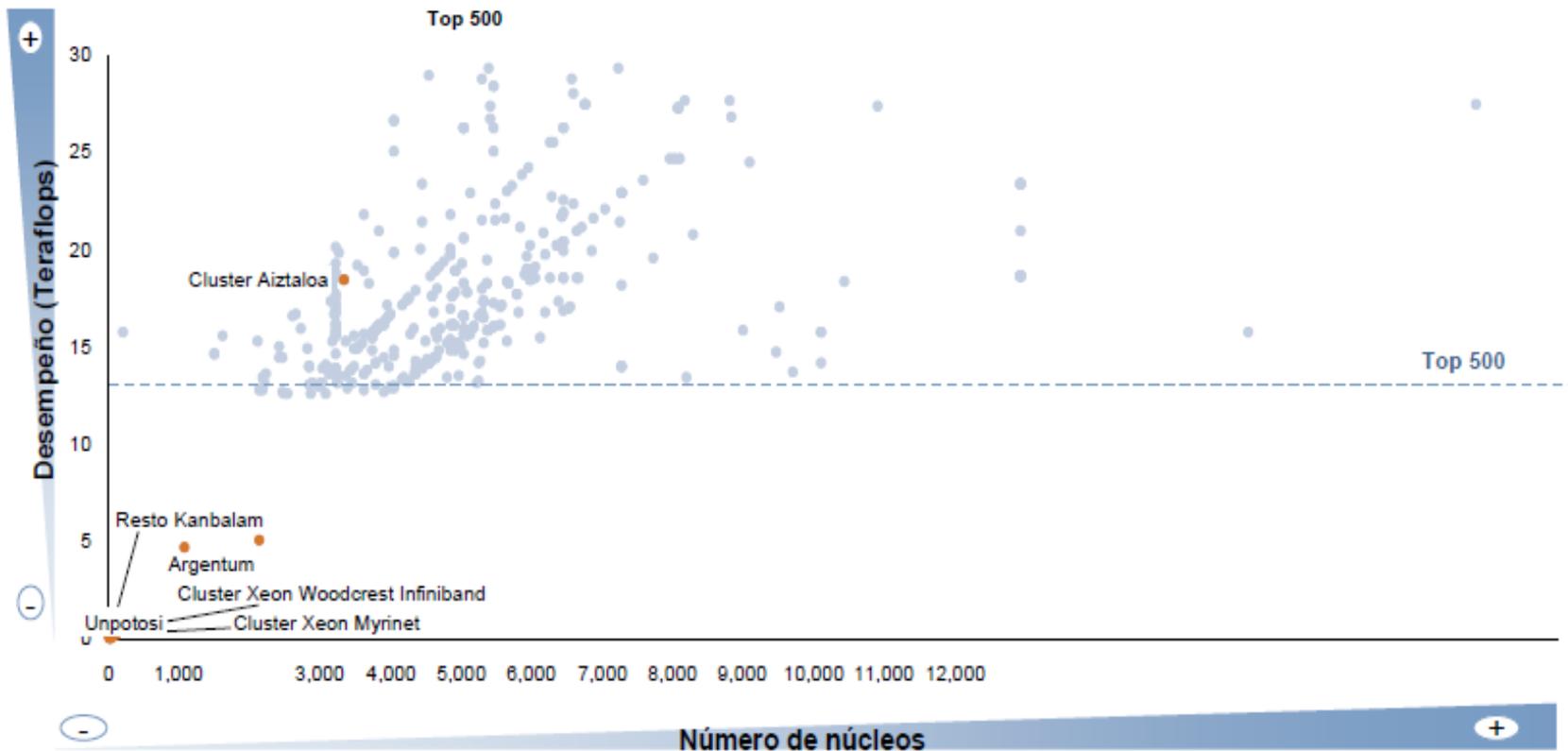
Proyecto de conectividad muy ambicioso.

Servicios de Data Center.

Evolución del Supercómputo



Top 500 (2009): prácticamente México no aparece en el Top 500



Realmente ha habido supercómputo en México: NO

- Ejemplo: Kambalam (UNAM) tiene 1400 núcleos, UAM 2,000 .
- $2,000/100$ investigadores = 20 núcleos por investigador.
- Supercómputo real requiere poder ejecutar en 1,000 – 5,000 núcleos por semanas o meses (proyecto tipo Grand Challenge).
- La comunidad académica nacional requiere del orden de 50,000 – 100,000 núcleos (y cientos de GPUs) localizados en varios centros de cómputo de alto rendimiento.
- **Contar con esto debe ser el objetivo de un Plan Nacional de Supercomputo + conectividad.**

Como hacerlo?

- Se requieren 100 millones de USD entre infraestructura de cómputo y comunicaciones.
- En USA y europa existen fondos para este tipo de proyectos (Jaguar (Oak Ridge), Ranger (TACC), etc.).
- En México esto hay que hacerlo de recursos propios o CONACYT.
- En nuestro caso se nos asignó recursos suficientes pero el compromiso fuerte es hacer sustentable el proyecto.
- El otro problema es la conectividad a los centros de cómputo y CFE no se ha mostrado muy positivo a ayudarnos.

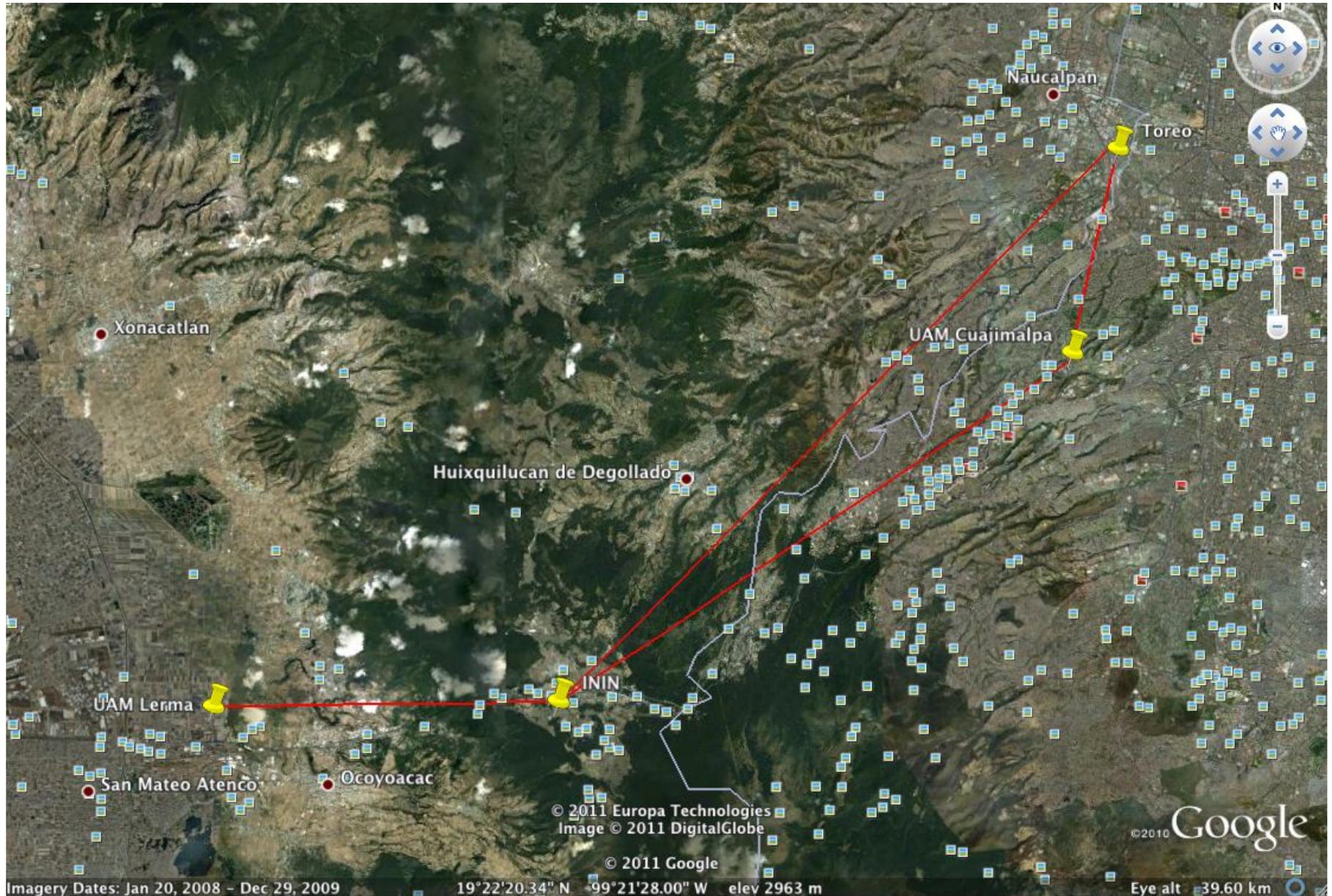
Proyecto: Desarrollo de un Centro Nacional de Computo de Alto Rendimiento

Objetivo: colocar al menos 100,000 núcleos (dentro del Top 10) en 3-4 centros, buena conectividad a ellos (CUDI) y compartir recursos.

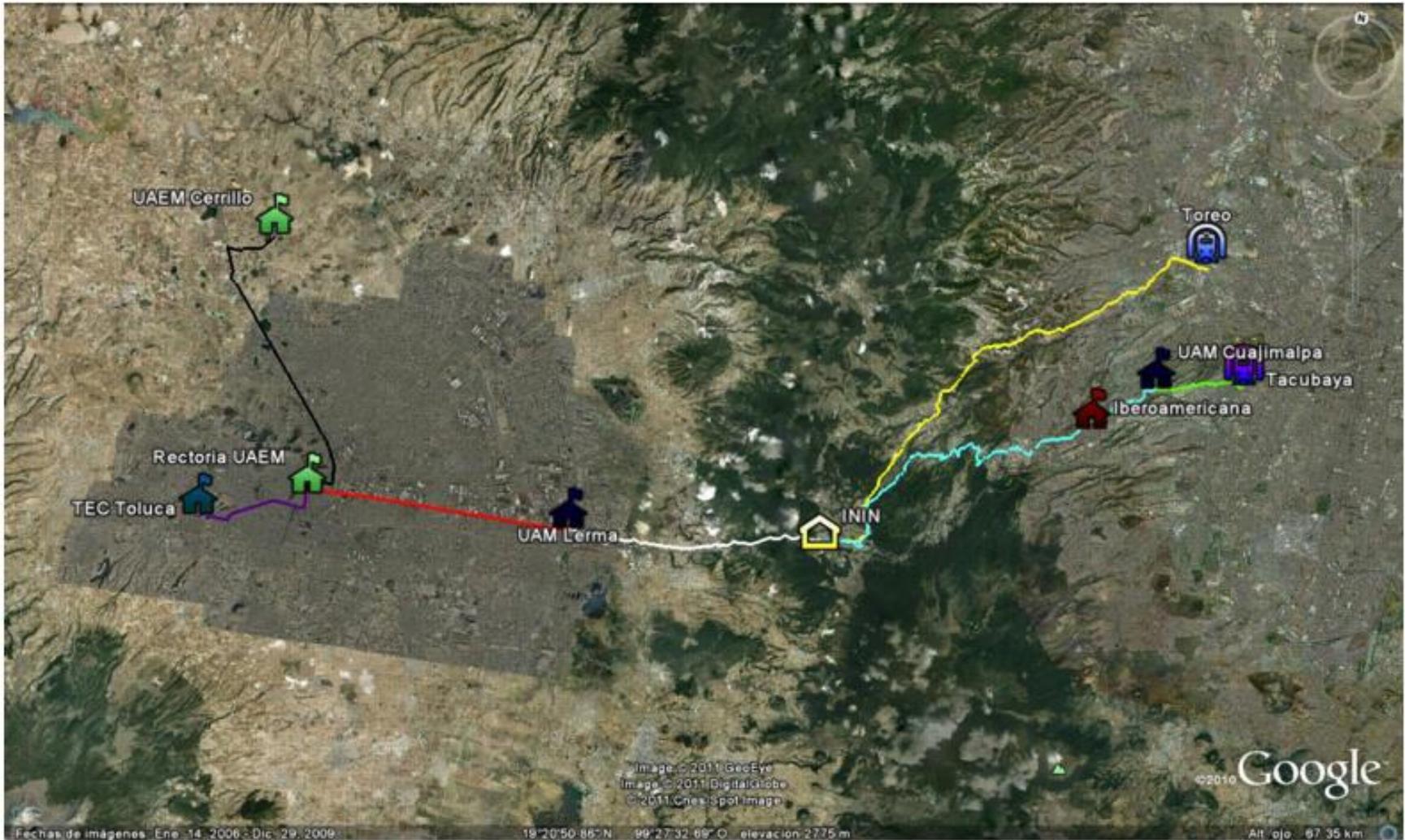
Se esta elaborando el proyecto:

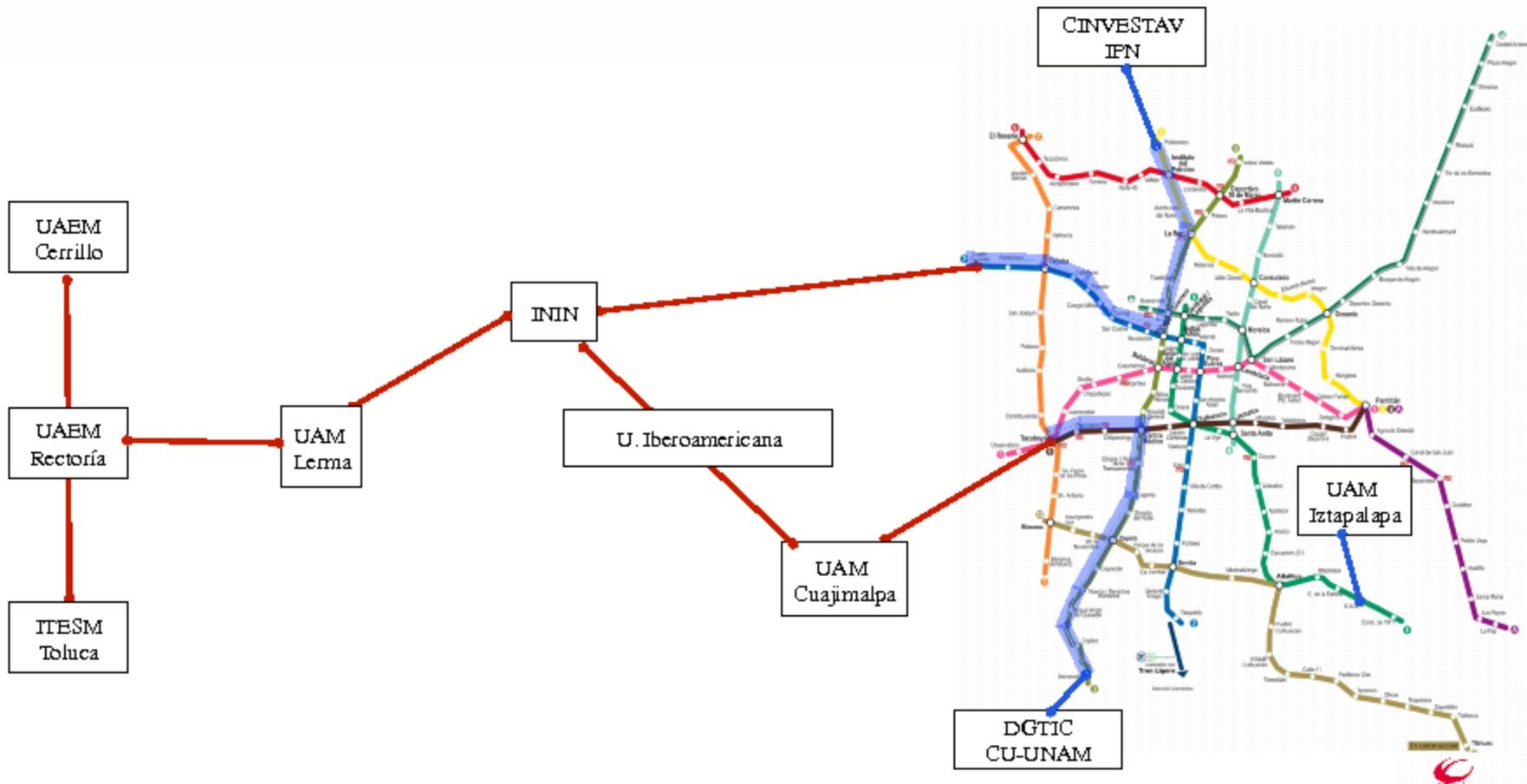
- UNAM (Felipe Bracho).
- UAM (José Luis Gazquez)
- Cinvestav - Zacatenco (Mariano Gamboa)
- Cinvestav - Estado de México (Jaime Klapp).

Localización y conectividad



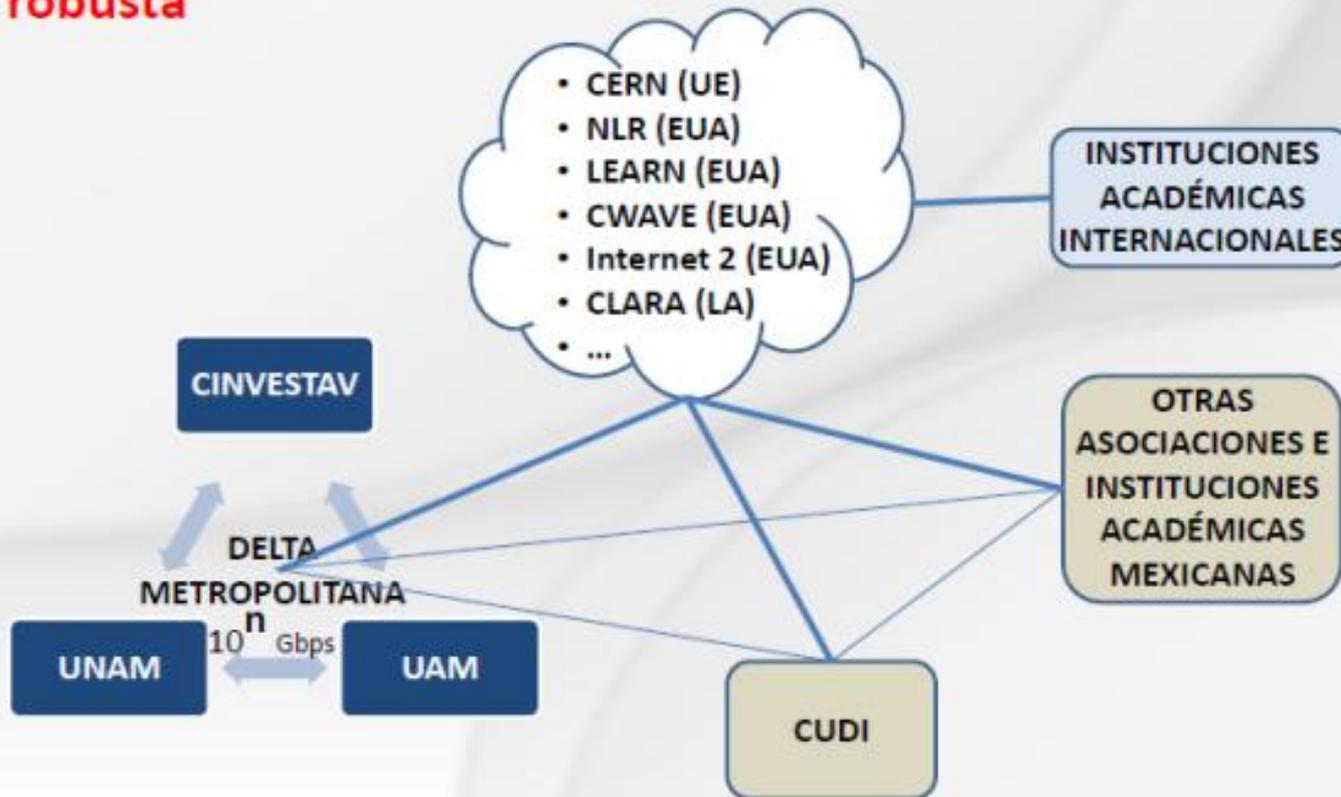






El futuro de LANCAD

Visión de LANCAD: Las instituciones académicas de México se desarrollan y se fortalecen colaborando en proyectos que hacen uso de una infraestructura compartida y robusta



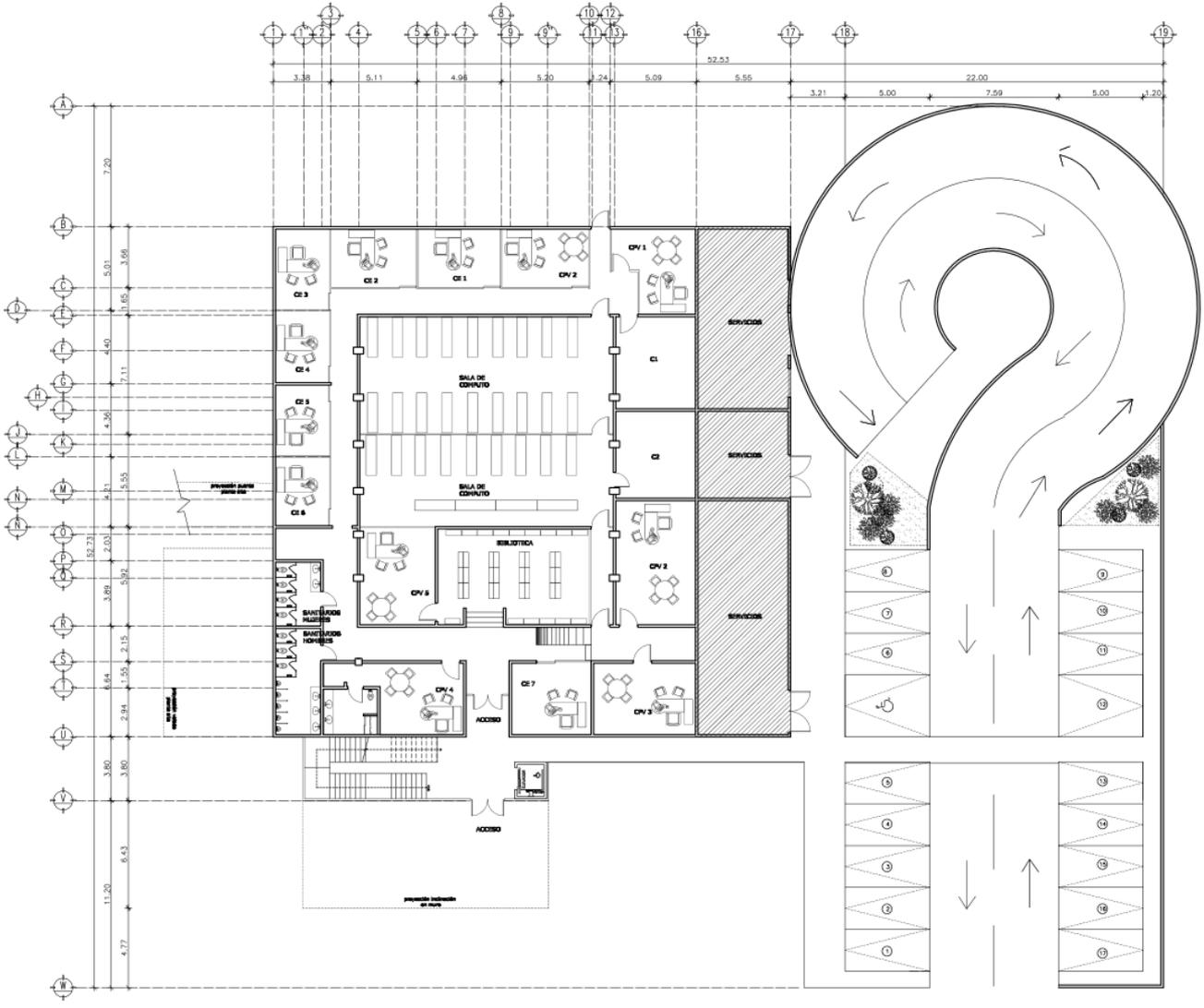
- CENTRO NACIONAL DE MATEMATICAS APLICADAS



AREAS

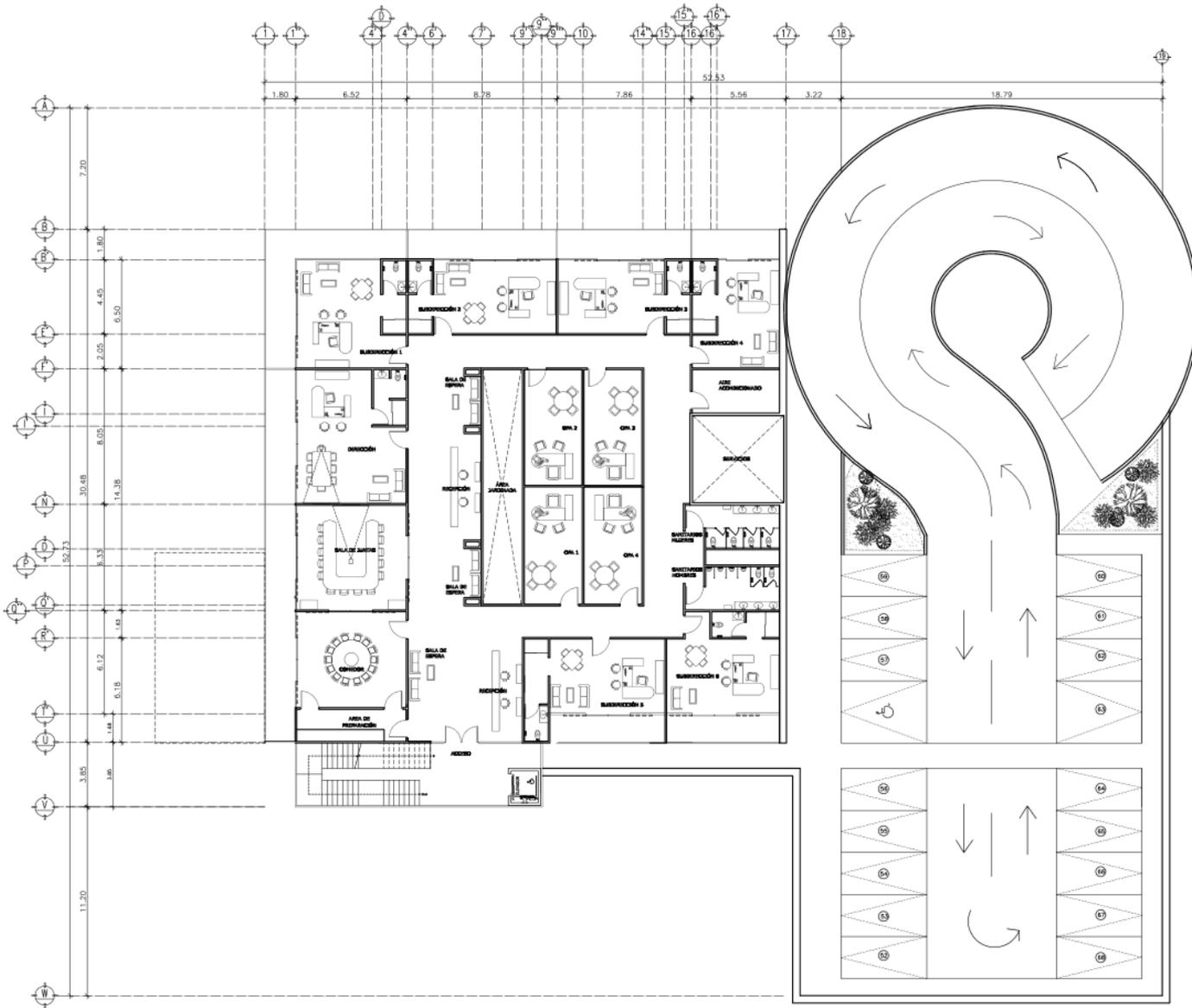
ANTEPROYECTO ABACUS			
ACUMULADO CUADRO DE ÁREAS:			
NIVEL	ESPACIOS	CANTIDAD	M2
PLANTA BAJA EXISTENTE			
	CUBÍCULOS DE ESTUDIANTES (CE)	7	115.75
	CUBÍCULOS PARA PROFESORES VISITANTES (CPV)	5	135.70
	BIBLIOTECA	1	54.57
	SALA DE COMPUTO	1	187.57
		SUBTOTAL P. BAJA EXIST.	493.59
PRIMER NIVEL			
	SALONES DE SEMINARIOS	3	159.43
	SALONES DE CLASE	3	199.36
	AUDITORIO	1	167.57
	JARDÍN INTERIOR	1	36.73
		SUBTOTAL PRIMER NIVEL	563.09
SEGUNDO NIVEL			
SEGUNDO	CUBÍCULOS DE INVESTIGADORES (CI)	23	378.08
	SNACK/CAFETERÍA	1	71.12
		SUBTOTAL SEGUNDO NIVEL	449.20
TERCER NIVEL			
	OFICINAS PARA PERSONAL ADMINISTRATIVO (OPA)	4	100.27
	SUBDIRECCIONES	6	233.58
	DIRECCIÓN	1	147.57
		SUBTOTAL TERCER NIVEL	481.42
GENERALES			
	MÓDULO DE BAÑOS	4	181.19
	ÁREA DE SERVICIO (BODEGA, ALMACÉN, SITE, AIRE ACOND., ETC.)	4	310.98
	PASILLOS, ACCESOS, VESTÍBULOS, SALAS DE ESPERA, ETC.		789.39
1 ° AL 3°	ESCALERA	4	220.40
SEGUNDO Y TERCER NIVEL			
	TERRAZAS		299.96
		TOTAL	5776.52
ESTACIONAMIENTOS			
	CUATRO NIVELES		4051.36

P B A J J A P L A N T A





P L A N
N I V E L 2
A



PLANTA 3

- CENTRO NACIONAL DE MATEMATICAS APLICADAS





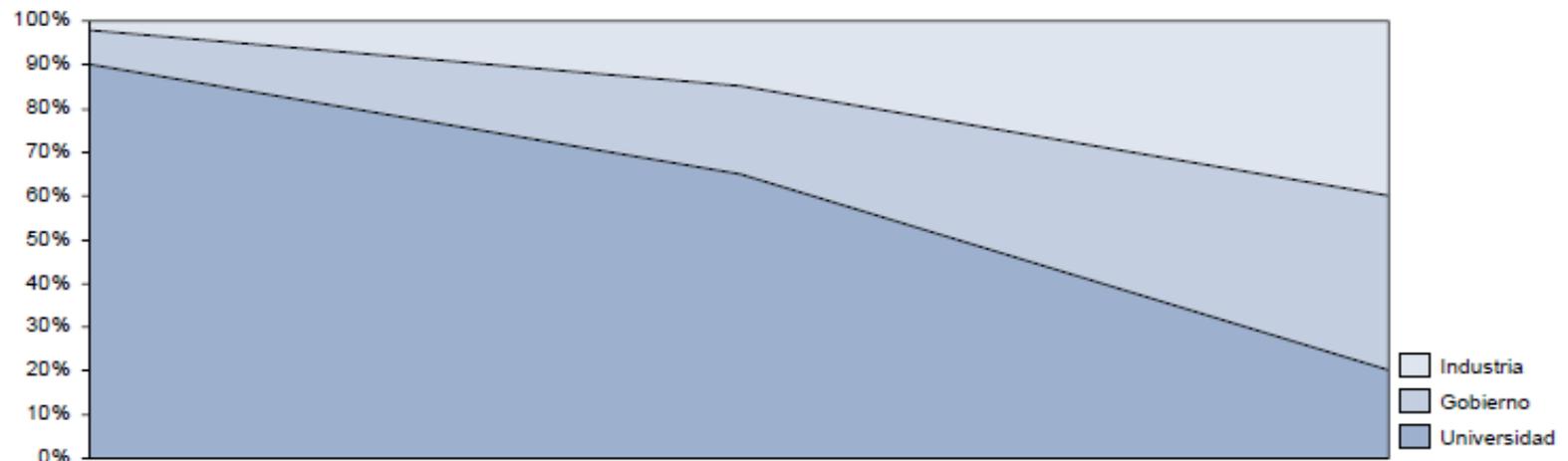






En función de la ambición que se tenga para el CCAREM y el volumen de la apuesta relacionada (inversión) se podrá esperar una evolución de la utilización del Centro de lo más abstracto (investigación) a lo más concreto (aplicaciones industriales)

Perspectivas de utilización



Corto Plazo

- La investigación es capaz de llenar los recursos de cómputo de alto rendimiento existentes

Mediano Plazo

- Las iniciativas del gobierno se ponen en marcha y comienzan a demandar recursos de cómputo de alto rendimiento
- Determinadas investigaciones pueden ser transferidas a productos exitosos dentro de la industria

Largo Plazo

- La industria reconoce las necesidades de cómputo de alto rendimiento para competir y empieza a demandar participación.
- Las iniciativas de gobierno se consolidan
- Se mantiene la Universidad con Cómputo de Alto Rendimiento más aplicado

Tipo de sistema (HPC+Cloud computing)



Provides Up to 115% Heat Extraction, and can even cool the room!



Water flow off



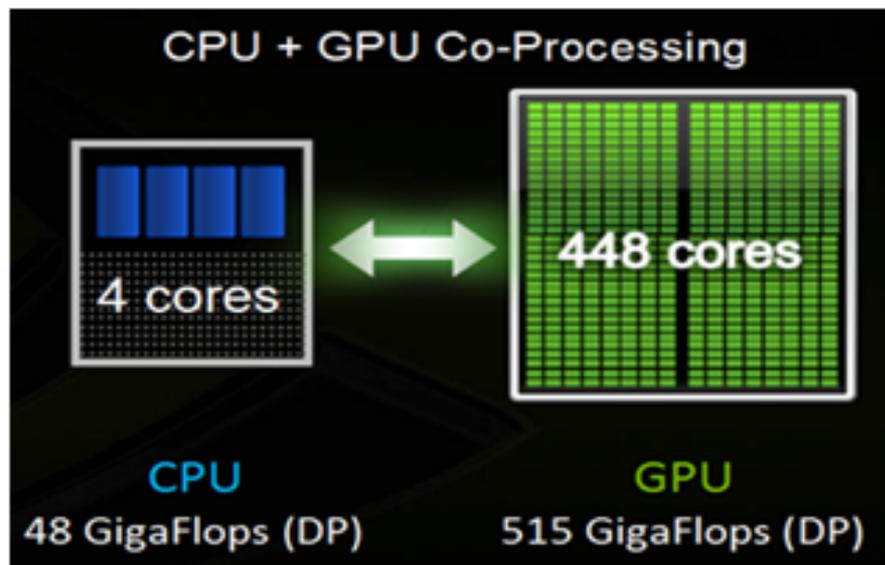
Water flow on

Why GPU's?

Speed
Performance



- GPU = Graphics Processing Unit
- GPU's do general purpose scientific and engineering computing
- CPU and GPU work together in heterogeneous computing model
 - The sequential part of the application runs on the CPU
 - The computationally-intensive part runs on the GPU.



Key Benefit

The application runs faster because it is using the high-performance of the GPU to boost performance.

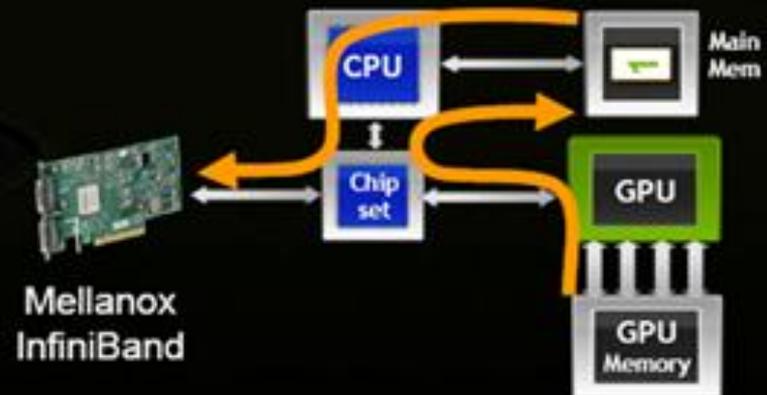
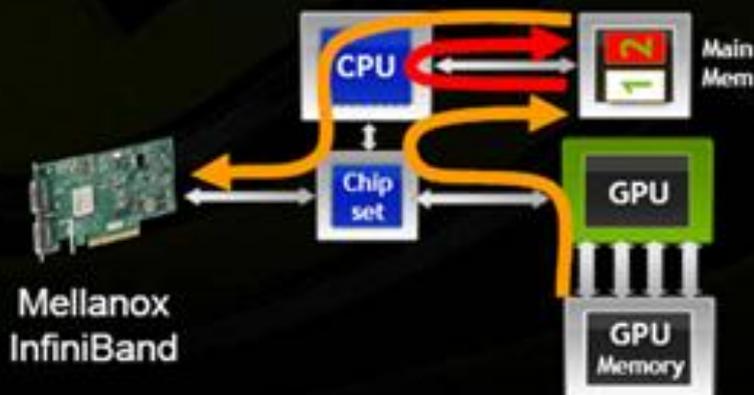
NVIDIA and Mellanox collaborate to further improve performance!

New GPU – InfiniBand Faster Communication



- Saves 30% in communication time
- Software available as CUDA C calls in Q2 2010
- Works with existing Tesla S1070, M1060 products

1. GPU Writes to Pinned Main Mem 1
2. CPU Copies MMem 1 to MMem 2
3. INF Reads MMem 2



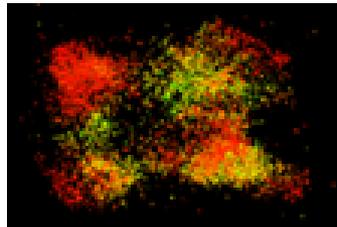
Key Research Applications Available on GPUs



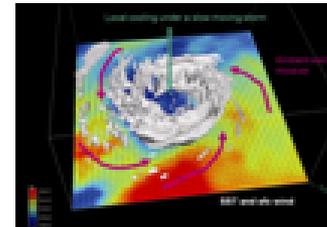
Molecular Dynamics / Quantum Chemistry



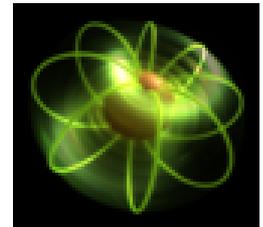
Computational Fluid Dynamics



Astrophysics



Weather & Climate Modeling



Many More

- AMBER
- ABINIT
- DL_POLY
- GROMACS
- LAMMPS
- MADNESS
- NAMD
- Q-Chem
- TeraChem

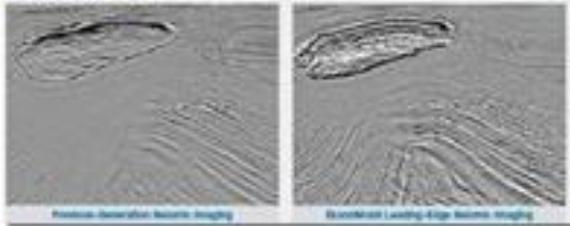
- OpenCurrent
- BAE Systems
- ANDSolver
- Euler Solvers
- Lattice Boltzman
- Navier Stokes

- N-body
- Chimera
- GADGET2
- Many published papers

- ASUCA (Japan)
- CO2 Modeling (Japan)
- HOMME
- Tsunami modeling
- NOAA NIM
- WRF

- Materials Science
- DCA++
- gWL-LSMS
- Combustion
- S3D
- Lattice QCD
- Chroma (QUDA)

INDUSTRIAL



Oil & Gas/
Seismic

Energy Discovery
Reservoir Simulation

Automotive

Computational Fluid Dynamics
Stereo Lithography modeling

FSS



Financial Markets

Monte Carlo Simulation
Market Prediction
Option Pricing, Data Mining

Insurance

Risk Analysis, Trending
Asset Liability Modeling

COMMUNICATION



Media &
Entertainment

Imaging & Computer Animation
Digital Content Creation

Gaming

Timesliced Gaming
Multiplayer Online Gaming
Video Streaming, Rendering
Virtual Desktop Decompression

PUBLIC



Life Sciences

Biology-Informatics

Higher Ed/
Research

Computational Chemistry
Quantum Mechanics Simulation

Healthcare

Diagnostic Imaging
Digital Radiography

Defense/
Government

Atmospheric, Ocean Modeling
Military & Space Science

Ejemplos de proyecto específicos: Etapa I

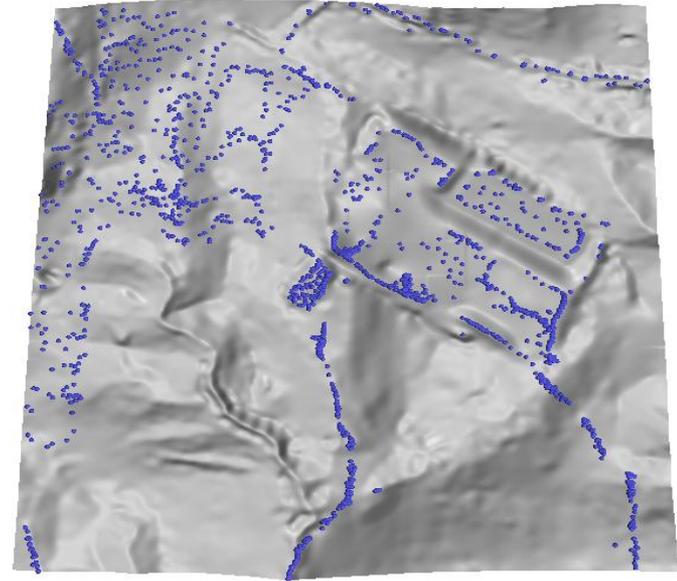
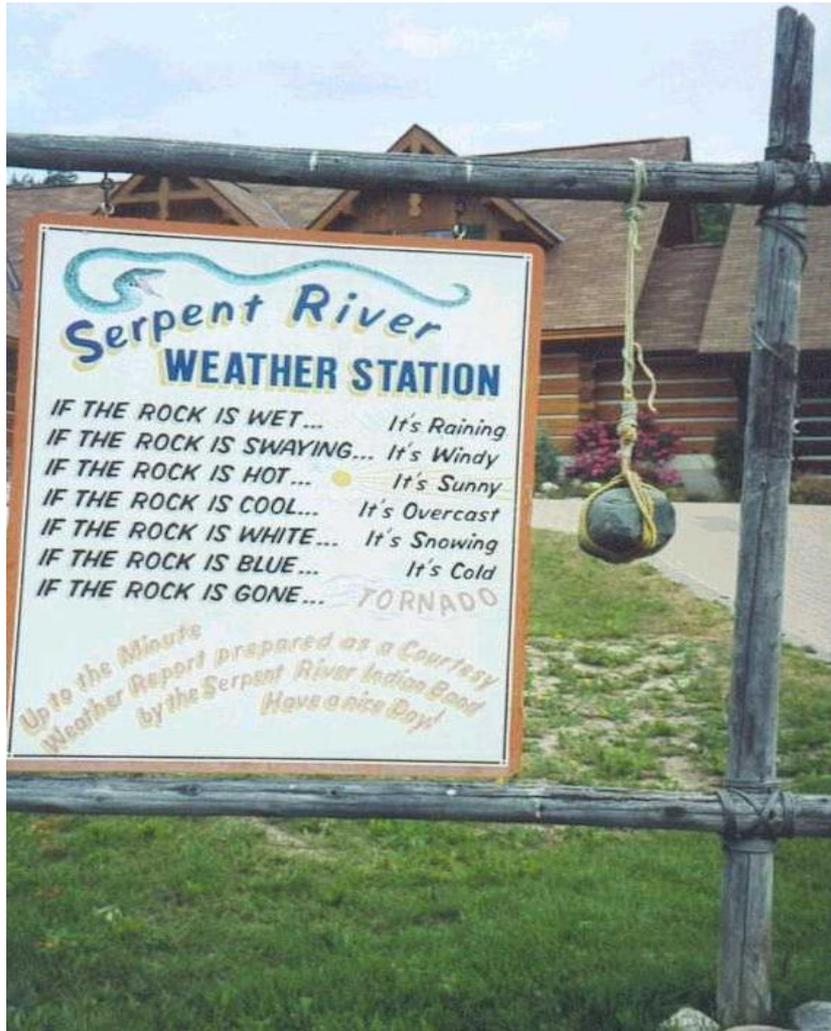
SISTEMA DE MODELACION NUMERICA DEL TIEMPO Y EL CLIMA BAJO DEMANDA

APLICACIONES:

1. Pronóstico del tiempo bajo demanda (cloud computing o cómputo en la nube), público en general.
2. Bancos de Niebla de Aeropuertos (pronósticos y servicios web). Se puede dar servicio a 50 aeropuertos.
3. Energía Eólica (simulación a largo plazo años). Determinación de sitios óptimos para Aerogeneradores.
4. Seguros Agrícolas (Modelación Climática) y préstamos bancarios. Agroasemex.
5. Modelación del tiempo por regiones.
6. Contaminación por regiones.
- 7.- Pronostico de clima por estado.

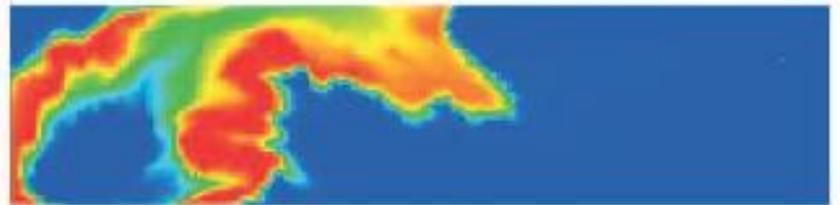
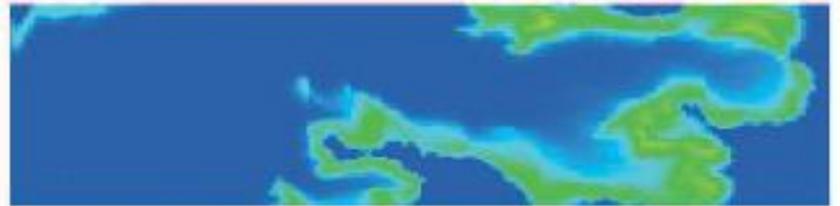
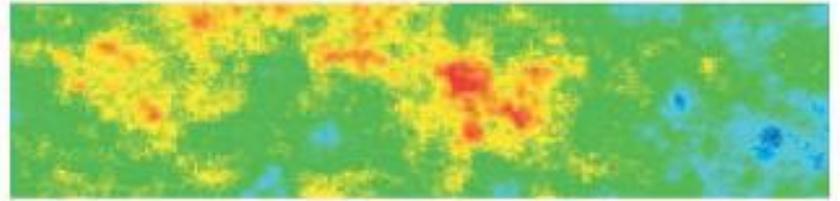
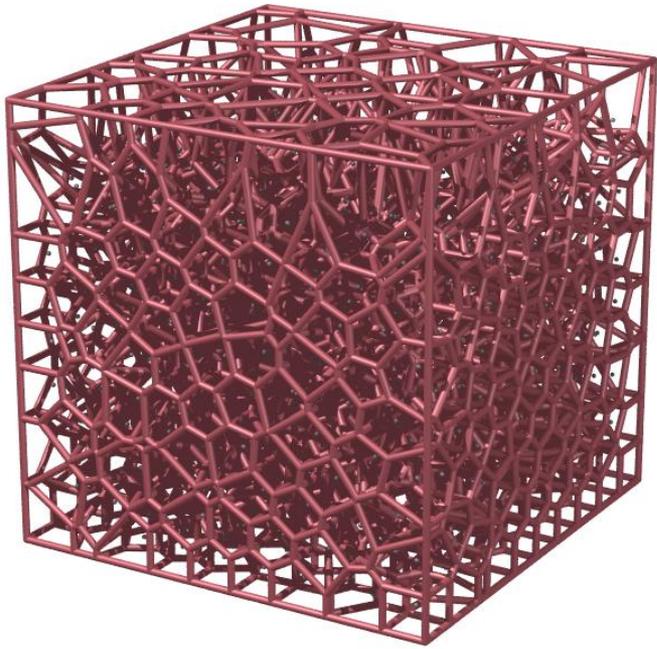
Sistema de alerta temprana para pronósticos meteorológicos y de dispersión de contaminantes para la Zona Metropolitana del Valle de Toluca (ZMVT)



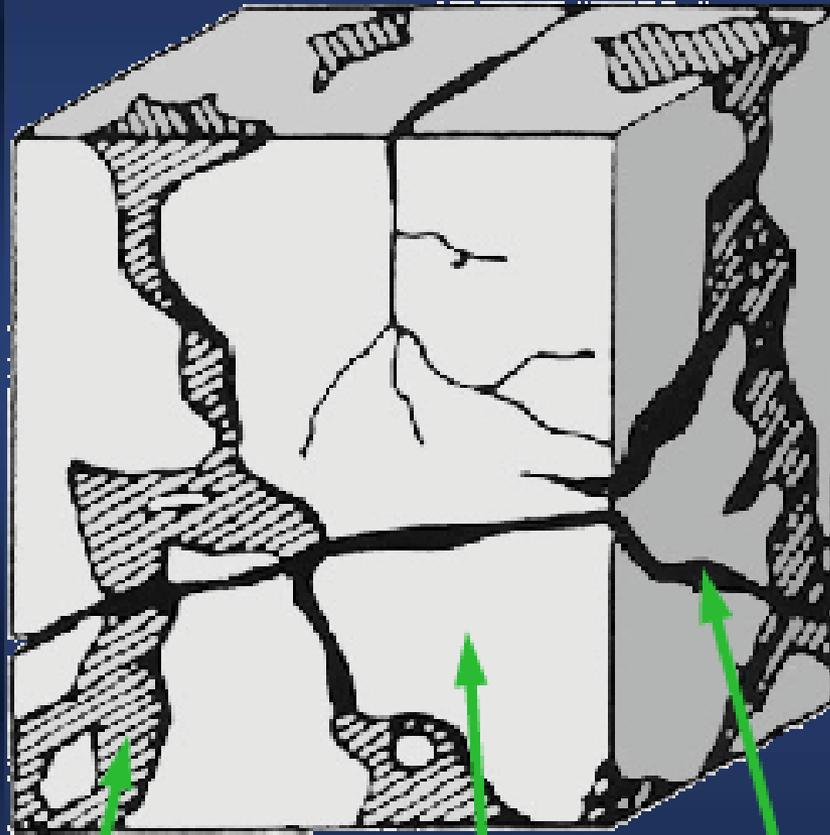


Desarrollo de un simulador de yacimientos naturalmente fracturados basado en SPH sobre procesadores CPU-GPU.

1. Desarrollo de un simulador para yacimientos fracturados con un modelo matemático con visión fractal, triple porosidad, doble permeabilidad, compresibilidad moderada y deformación del medio poroso. El método numérico estará basado en la técnica conocida como Hidrodinámica de Partículas Suavizadas o SPH, que es adaptiva en tiempo y espacio, acepta geometrías altamente irregulares y físico-química sofisticada. El sistema permitirá modelar la matriz, fracturas, vórgulos e incluir CO₂ y nitrógeno.
2. Desarrollo de un código numérico SPH para ser ejecutado en programas masivamente paralelos que utilicen tanto procesadores tradicionales CPU como procesadores gráficos GPU, lo que permitirá la ejecución de modelos de alta resolución en tiempos moderados.
3. Desarrollo de interfaces para acceso, uso y análisis de resultados del simulador.



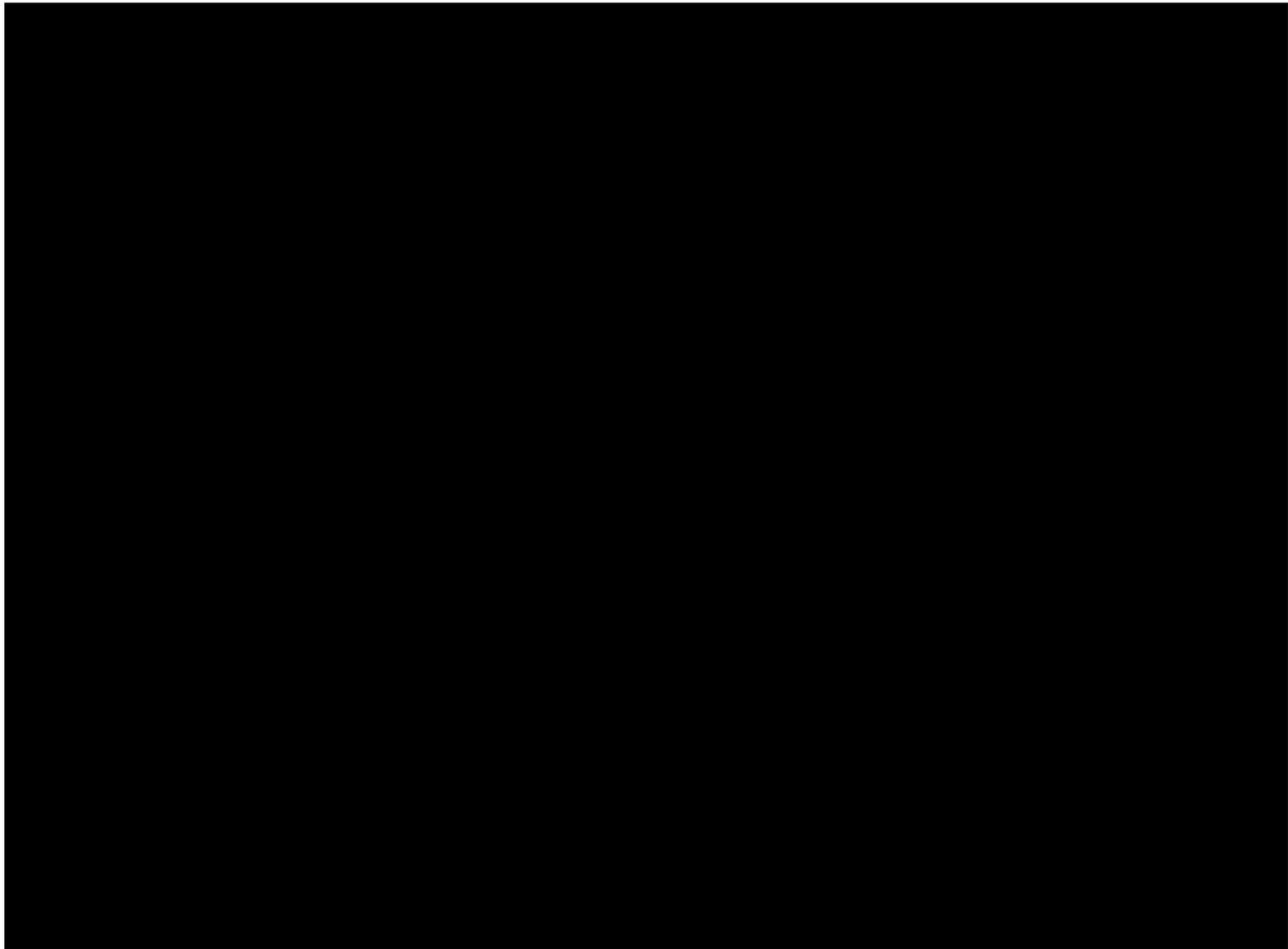
Yacimiento Real



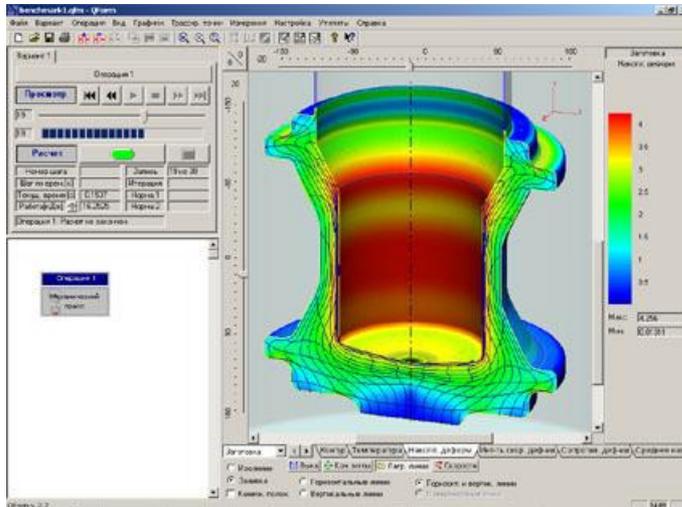
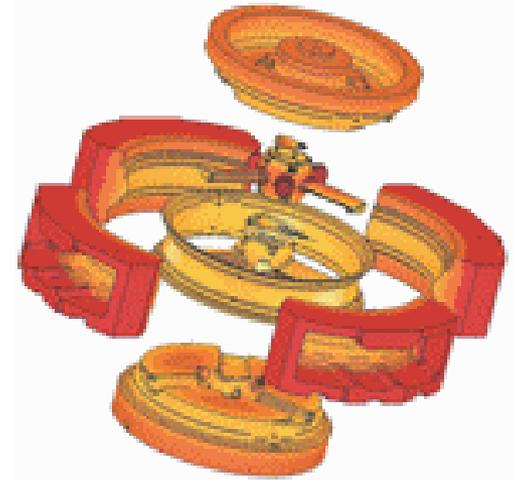
vugulus

matriz

fractura

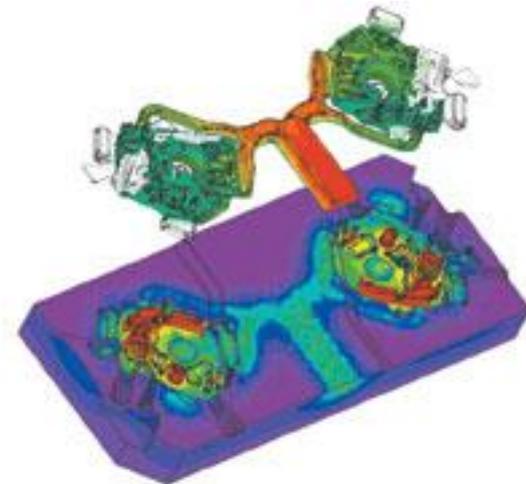


Industria metalúrgica y del plástico



QForm2D interface.

Workpiece with strain distribution and flow lines are shown.



Industria de recubrimientos (Sherwin Williams y Comex)

Diseño mediante modelado molecular de nuevos materiales para mejora de propiedades en recubrimientos y pinturas. Se propone el diseño de nuevos compuestos químicos que permitan la mejora en propiedades específicas en la industria de recubrimiento como brillo, adhesión, resistencia, color y estabilidad de las formulaciones. El diseño de nuevos materiales mediante simulación molecular permitirá reducir el costo en el diseño y desarrollo experimental de los mismos haciendo un análisis teórico previo a la síntesis de éstos que permita decidir sobre la factibilidad de la producción de ciertos materiales para cubrir las especificaciones esperadas en distintas propiedades. Este tipo de modelado se llevará a cabo empleando diversas técnicas de simulación según las propiedades que se deseen estudiar.

Apoyo de actividades bursátiles de la Bolsa Mexicana de Valores.

Optimización de procesos y formulación. Diseño de formulaciones haciendo análisis de las propiedades de mezclas complejas mediante simulación computacional y procesos de optimización numérica eligiendo las mejores formulaciones que optimicen alguna propiedad en función de tiempo de producción y costo de la misma.

Determinación de propiedades fisicoquímicas para análisis de viabilidad de producción y estabilidad de los sistemas formulados. Obtención de relaciones de reactividad para procesos de polimerización, estudios de estabilidad coloidal mediante isotermas de adsorción y de presión de disgregación, y construcción de diagramas de fases.

Predicción de propiedades de materiales mediante simulación numérica. Por ejemplo propiedades reológicas, resistencia, color, brillo y adherencia.

Desarrollo de software que permita de manera adecuada el ajuste de color en base a las propiedades electrónicas de pigmentos.

Ejemplo de aplicaciones

DINAMICA DE FLUIDOS COMPUTACIONAL

Existe una gran variedad de posibles servicios que involucran Dinámica de Fluidos.

PEMEX

Proyectos de exploración de nuevos yacimientos. Ingresos podrían ser significativos.

IMP

Si nos conectamos a la delta, el IMP como institución podría ser un gran cliente.

EDUCACIÓN A DISTANCIA PARA PROGRAMADORES EN JAVA, JRAD.

Aplicaciones de programación rápida para convertir desarrolladores de bajo perfil en desarrolladores productivos.

EDUCACIÓN A DISTANCIA

La demanda de estos servicios es gigantesca.

INDUSTRIA FARMACEUTICA

Ejemplo de aplicaciones

INTELIGENCIA DE NEGOCIOS PARA EL EDOMEX.

Proyecto Data Mining para el Edomex. Se requiere de acuerdos a nivel estado. Ingresos significativos.

DATA CENTER

0. Cloud Services (procesamiento comercial, científico, aplicaciones, almacenamiento, comunicación).

- Proveedor autorizado de Certificación (SAT), factura electrónica.
- Ser proveedor.
- Vender infraestructura a los proveedores.

1. Servicios de Gobierno

- Aplicaciones de conectividad en puntos remotos.
- Kioscos de pago de servicios.
- Educación a distancia.

2. Servicios industriales (Carrier de Aplicaciones).

- Industria del plástico y metalurgia.
- Automotriz
- Farmacéuticas.
- Fabrica de software.
- GPS.
- ERP micro.
- DRP Recuperación de desastres.

3. Bases de Datos

- Información de personal, centros de consumo, sectorización de mercados.
- Minería de datos comerciales.
- Seguridad.
- e-commerce.

Ejemplo de aplicaciones

4. Procesamiento de imágenes.

- Flujo vehicular
- Envío de multas electrónicas por reconocimiento de placas (radares).

5 Telecomunicaciones (con conexión a la delta).

- VoIP, voz IP (empresas, Gobierno)

6. internet

- Hosting y colocacion

7. Seguridad y Policia. Globo 5555149668

Conclusiones

- El Centro de Matemáticas Aplicadas y Cómputo de Alto Rendimiento se pretende sea sustentable y con alto impacto científico internacional.
- Se esta proponiendo la creación de un Centro Nacional de Cómputo de Alto Rendimiento con al menos 100,000 núcleos, un numero importante de GPUs y buena conectividad (dentro del Top 10).