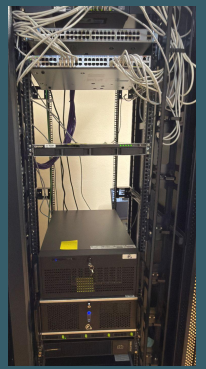
Computing @ ICTEA

Isidro González Caballero

Carlos González Gutiérrez

ICTEA
Instituto Universitario de Ciencias y Tecnologías Espaciales de Asturias

Universidad de Oviedo



Geographical Distribution of Resources

Mieres: Bulk of powerful resources

- 1.5 racks at C3 (Clúster de Computación Científica)
- Well-equipped Data Processing Center (CPD)

Oviedo: F. Geología, Minas,...

- Smaller racks with redundant services…
- … and also desktops in different offices.
- Not so well-equipped facility

⇒ Can we move most of this to C^3?

IMHO, yes (at least main resources)

C3 @ Mieres

Lab @Geol

Office @Minas

Office @Geol

Minas

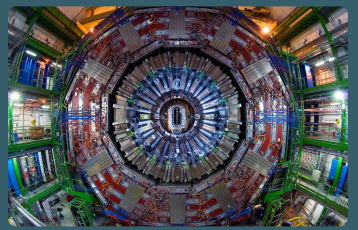
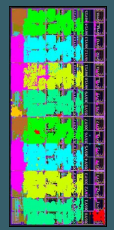
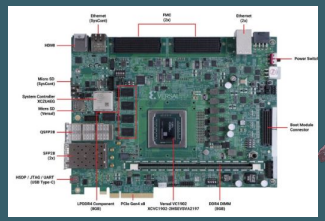
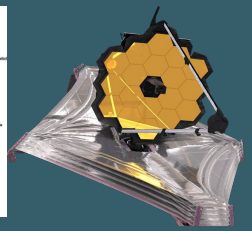
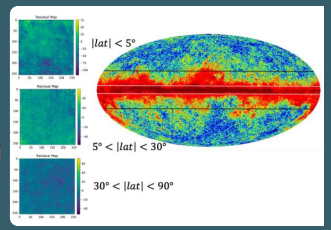
Diverse Scientific Objectives

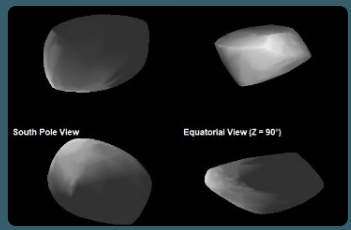
Particle Physics

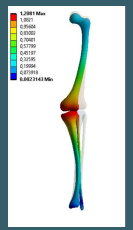
- **Experimental:** Analysis, detector ops, CMS data validation, trigger, FPGA
- **Theoretical:** Emerging needs for the "near" future

MOMA

- **Astrophysics, Astronomy** and **Cosmology**
- **AI & Stats:** Analysis in medicine, economy, and beyond





⇒ Type of computing resources needed may be very different!

Different software solutions

Operating System

AlmaLinux (CERN imposed) vs Ubuntu (moving to AL)

(Windows, Mac for desktops)

Programming languages

Python mainly, but also C++, HLS, CUDA, R, C, matlab

Software packages

CMSSW, ROOT, Vivado, Vitis vs dozens of different packages

How is software installed?

Centrally distributed by CMS vs virtual environments - and containers!

Heterogeneous HW

Resources accumulated over 15+ years by various groups and projects

- Different brands/vendors: HP, DELL, SIE,...
- Multi-generation processor…
- CPU- and GPU- oriented,...

Challenges

- Configuration becomes significantly more complicated.
- Central operation is more difficult to manage.

Advantages

- Access to a large volume of diverse resources.
- Solutions to different problems: memory intensive, pure HPC, HTC, GPU optimized…

HEP resources @ C^3

Storage: ~350 TB - 4 servers

- **Lustre** for main mass storage
- XFS + LVM for legacy data
- NFS for home directories
- LVM for VMs

Computing - 7 servers

- 364 cores / 7**28 threads**
- RAM **2GB/thread**
- Available through **SLURM**

Network

- High speed network through fiber/copper up to 10 GbE
- All nodes see all the data
- All data exported to desktops also

[*Computation*]
Heterogeneous
HP & DELL multicore servers

[*Switch*] - 25/100G

[*Storage*] - DELL
RS650xs + MD1400

[*Switch*] - OS y Mgt

[*Services*]
VM, NFS, Lustre

[*Storage*] - DELL
RS660xs + MD2412

[*Switch*]- Fiber

[*Storage*]
Hitachi HUS 110

[*Storage*]
Hitachi AMS

MOMA resources @ C^3 and Minas

Computing - 6 servers

- [**Urano**] (64GB RAM - 500GB SSD)
  - 1 x NVIDIA GeForce GTX 970 – 4G VRAM
- [**Hyperion/Kronos**] (128/256Gb RAM – 1TB SSD)
  - 2x/4x Nvidia GeForce GTX TITAN X
- [**Gea**] (512 GB RAM - 1 TB SSD)
  - 24 cores Xeon W-3235 CPU
  - 4 x NVIDIA GeForce GTX 2080 – 12Gb VRAM
- [**Poseidon**] (1 TB RAM – 2 TB SSD)
  - 32 cores Ryzen Threadripper PRO 5955WX
  - 4 x Nvidia GeForce GTX 4090 – 24Gb VRAM
- [**Prometeo**] (384 GB RAM - 4 TB SSD)
  - 288 cores AMD Epyc 9825

Storage: ~32 TB - 1 server

- **QNAP TS453BU**
- **44 Tb -> Útiles 32Tb (RAID 5)**

…and local storage

So what is the future…

Current Computing Resource Landscape at ICTEA

Spread geographically

Diverse scientific objectives

Heterogeneous hardware

Many software solutions

Coordination & Centralization

Share computing resources

- Improve CPU/GPU availability by transparently sharing resources
- Scalability for activity peaks

Share financial resources

- Access to better high performance hardware
- and infraestructure

Share human resources

- Through well established deployment and maintenance procedures
- Shared technical expertise
- Faster fix and help times

Design principles

Avoid custom made solutions

If possible use **standards**, mainstream solutions and recommendations from our scientific projects

Flexibility

- Support and **adapt** to different needs
- Be able to quickly install and deploy new requirements

Homogeneous environment

- Transparent usage of nodes and desktops
- Centralized management (accounts, software, updates)

Coordinated purchasing

Coordinate purchasing to avoid overlaps and to improve the common pool of resources

Fair use

Resources (computing and storage) will be monitored and they have to be used fairly

Where are we in this journey?

🛠️ Incorporating MOMA resources into existing HEP cluster environment

Central User Management

FreeIPA Implementation

- Dedicated MOMA group created
- 4 active users onboarded

Computing Resources

SLURM Batch System

- 3 nodes added: 3 GPU-like, 1 CPU-like (288 threads)
- 2 partitions: `momagpu` & `momacpu`
- Remaining nodes will follow soon (8xGPUs)

Storage

HEP Storage Integration

- Users `home`: Shared space in HEP storage (limited but with backup)
  - *Use it wisely*
- Mass storage: Currently separated (no backup)

```
Sun May 31 10:00:15 PM CEST 2026
Total Jobs Executed in momacpu: 90
Total CPU Hours: 39.2636
```

Next steps and open questions

Consolidate at C³

All nodes should be **relocated to C³** to improve connectivity and simplify management.

Desktops

Should we have standard central desktop installation (*a la* HEP)?

Central Services

- **Monitoring:** Prometheus/Grafana, Ganglia
- **Web:** Servers (needs supervision)
- **Wiki/Blog:** Web based/DB backend
- … anything else?

Available funding

Buy shared equipment for the institute. What?

- 2 multicore CPU servers
- 1 (2 x) GPU server
- **~100 TB storage server**

A note of caution

Cybersecurity is a shared responsibility

(Ciber)security is a joint effort between sys admins and users. We are **constantly probed for vulnerabilities**.

⚠ Legal Notice

You are **legally responsible** for damage caused to our infrastructure if safeguards are not implemented.

Rules & Recommendations

- Protect your access credentials → **never share them** → not even with admins!
- Use **strong, unique passwords** or, better, passphrases.
- Avoid public/untrusted networks to access our resources.
- Report any suspicious activity to the technical team immediately.

Summary

Progressing…

Progressing towards a centralized, professional, and coordinated cluster for scientific computing at ICTEA, while still keeping our scientific activity ⇒ Still learning, expect still some changes

Integrating…

Moving from a "sum of parts" to a unified resource pool for all the institute

Moving to C^3

The natural place to place our main computing and storage resources

[Master en Aplicaciones de Supercomputación y IA] - [Jornadas del C^3 → Register now to attend]

First steps towards an ICTEA computing cluster. Stay tuned!

Backup

Máster en Aplicaciones de Supercomputación y Aprendizaje Automático en Ciencia y Tecnología

- Presencial / 20 plazas. 1 curso (60 ECTS).
- Primer **máster universitario en España conectando ciencia/ingeniería e inteligencia artificial**
  - Dirigido a egresados de grados de ciencias, ingeniería y disciplinas transversales.
- **Profesionalizante**: proporciona a graduados en ciencias e ingenierías la formación necesaria para ocupar puestos de trabajo relacionados con la computación científica: análisis de datos, aprendizaje automático, supercomputación, programación científica, etc.
  - **Prácticas externas** en empresas y laboratorios de investigación del Principado de Asturias: Arcelor, INETUM, CINN, INCAR, ISPA, IDONIAL, etc. Posibilidad de prácticas externas en otras comunidades

PLAN DE ESTUDIOS

Primer semestre | septiembre a diciembre

Herramientas Científicas para el Análisis y Modelización de Datos (Programación científica en Python)	6 ECTS	Obligatoria
Computación Científica de Altas Prestaciones	6 ECTS	Obligatoria
Aspectos Legales y Éticos del Tratamiento de Datos Científicos	3 ECTS	Obligatoria
Herramientas Científicas Avanzadas para el Análisis y Modelización de Datos (Programación avanzada en Python)	6 ECTS	Obligatoria
Aprendizaje Automático para Aplicaciones Científicas y Tecnológicas	6 ECTS	Obligatoria
Visualización y Análisis de Datos Científicos	3 ECTS	Obligatoria

Segundo semestre | enero a mayo

Trabajo Fin de Máster	12 ECTS	TFM
Prácticas Externas	6 ECTS	Prácticas Externas

***Cuatro optativas** a escoger entre:*

Aprendizaje Automático Avanzado para Aplicaciones Científicas y Tecnológicas	3 ECTS	Optativa
Optimización de Procesos Asistida por Ordenador	3 ECTS	Optativa
Computación Cuántica	3 ECTS	Optativa
Cribado de Alto Rendimiento en Ciencia de Materiales	3 ECTS	Optativa
Diseño y Simulación de Nuevos Materiales y Nanoestructuras	3 ECTS	Optativa
Simulación y Diseño Computacional de Biomoléculas	3 ECTS	Optativa
Química Computacional de Moléculas y Sólidos	3 ECTS	Optativa
Interpretación y Análisis de Datos Genómicos	3 ECTS	Optativa
Metagenómica y Transcriptómica	3 ECTS	Optativa
Biomarcadores Multiómicos para Medicina Personalizada	3 ECTS	Optativa

Curso 2026-27

Se desarrollará también como máster Ingenium

INGENIUM European University