

PIERRE  
AUGER  
OBSERVATORY

# Manejo de datos de los observatorios HAWC y Pierre Auger

Lukas Nellen

ICN-UNAM

[lukas@nucleares.unam.mx](mailto:lukas@nucleares.unam.mx)

# Con apoyo de muchos

● Laboratorio Nacional HAWC



● Colaboración Pierre Auger



● Red HAWC

● Red Física de Altas Energías



● Red ALICE



ALICE

● UNAM



● CUDI



● CONACyT



● NSF, DOE



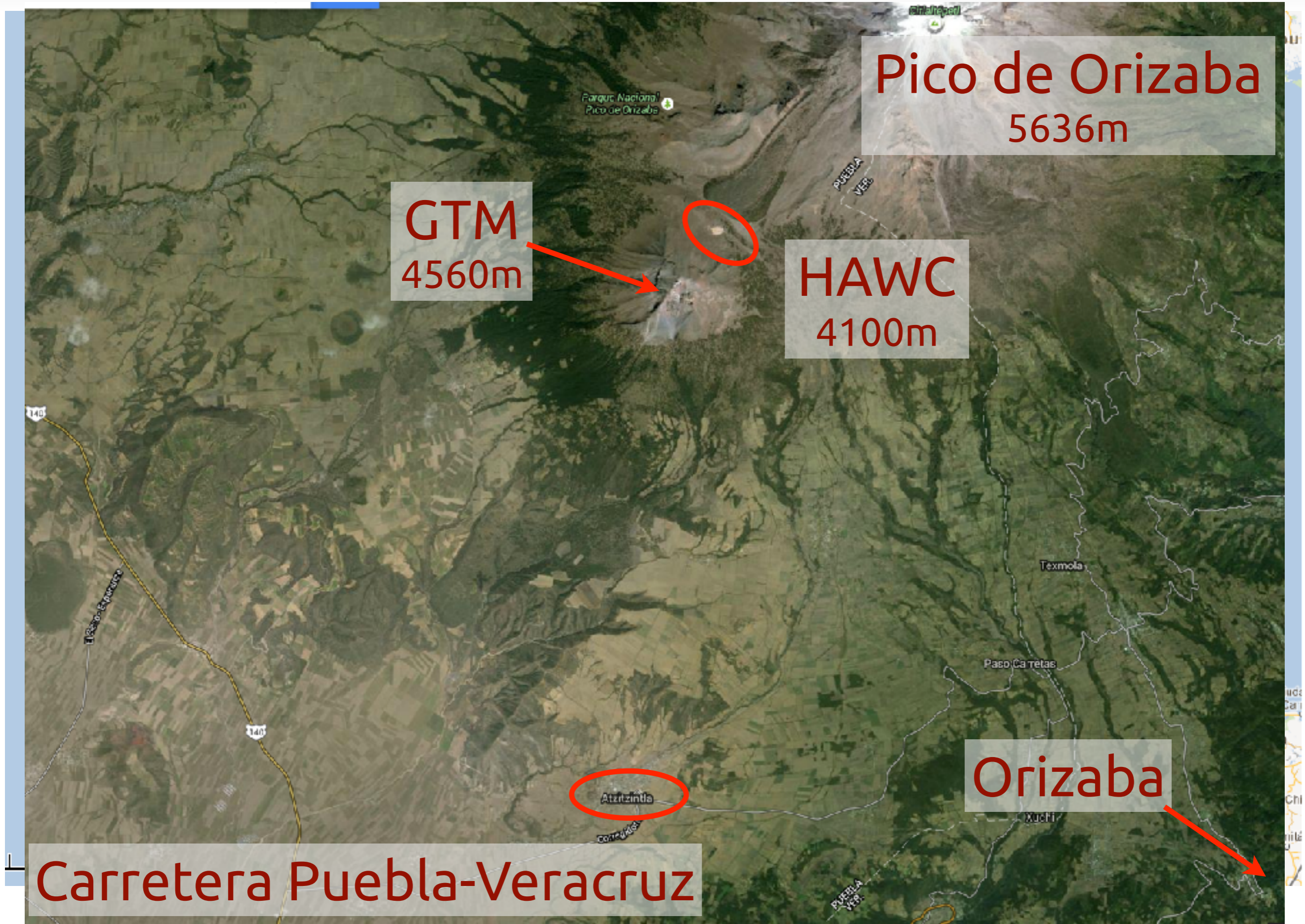
# Tenemos: HAWC – Volcan Sierra Negra



Pico de Orizaba



# Tenemos: HAWC – Volcan Sierra Negra



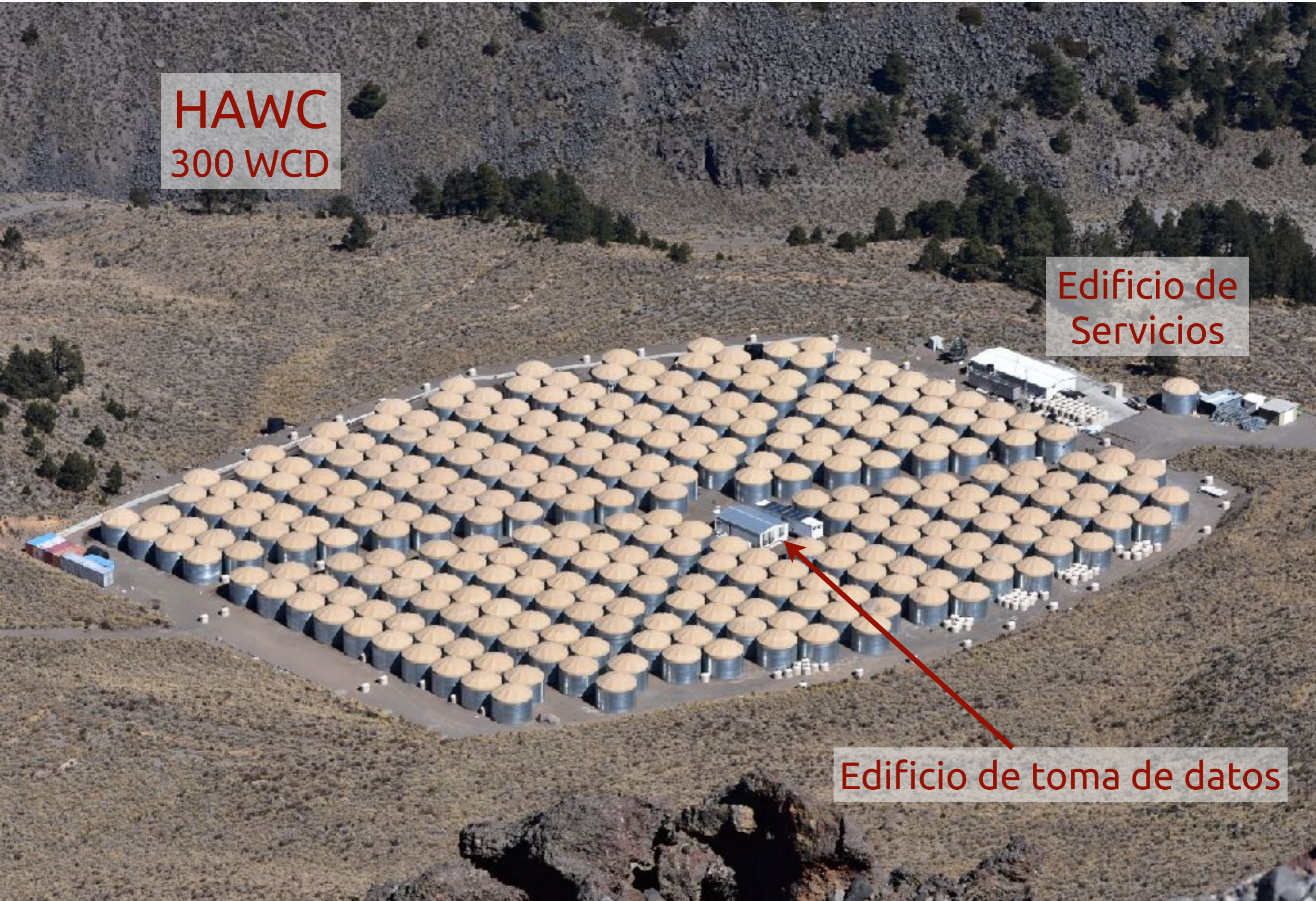


# Tenemos: HAWC – Volcan Sierra Negra

HAWC  
300 WCD

Edificio de  
Servicios

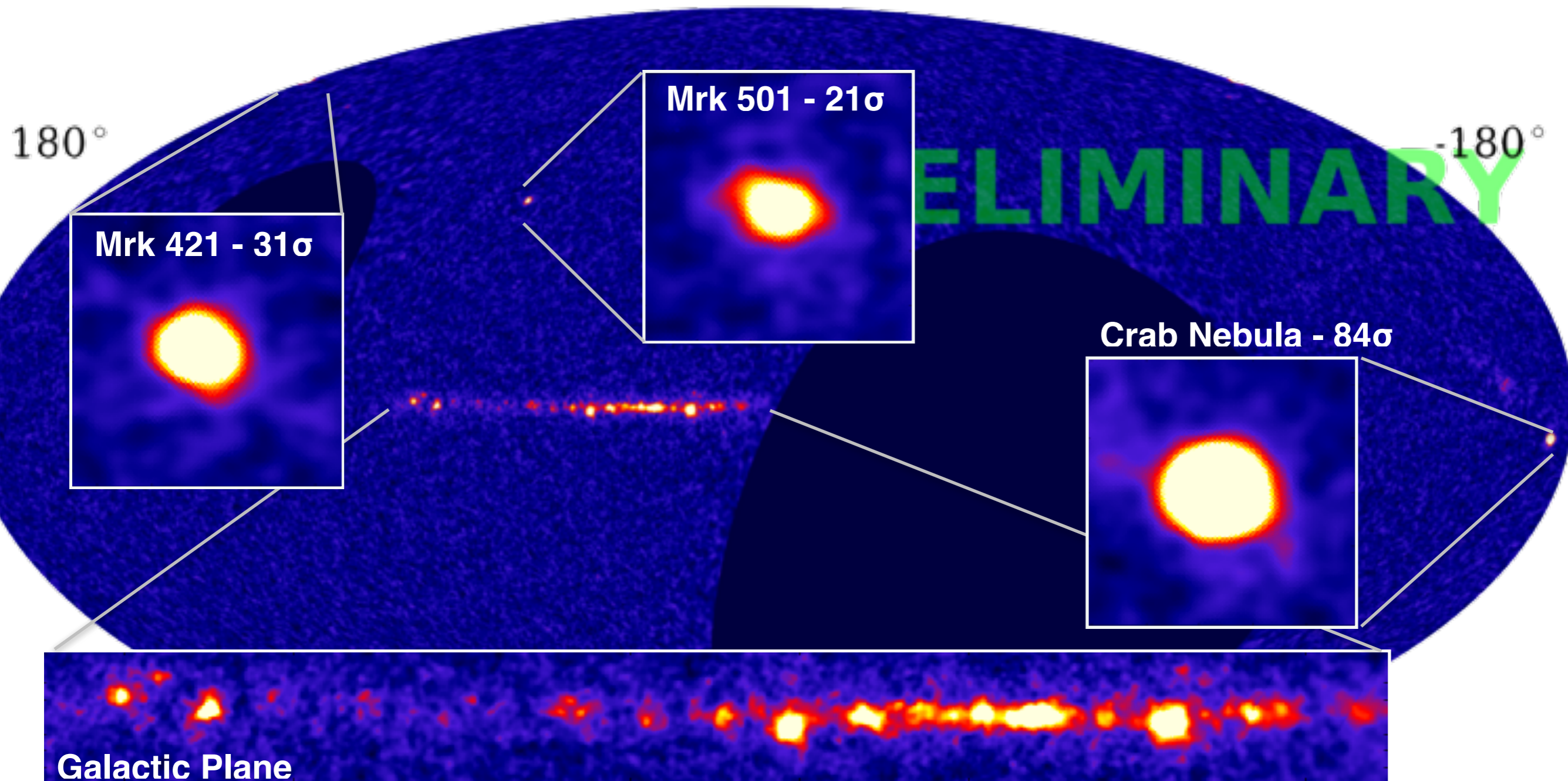
Edificio de toma de datos





# Queremos: Mapa del cielo en gamma


● 341 días de datos con el detector completo



# Solution

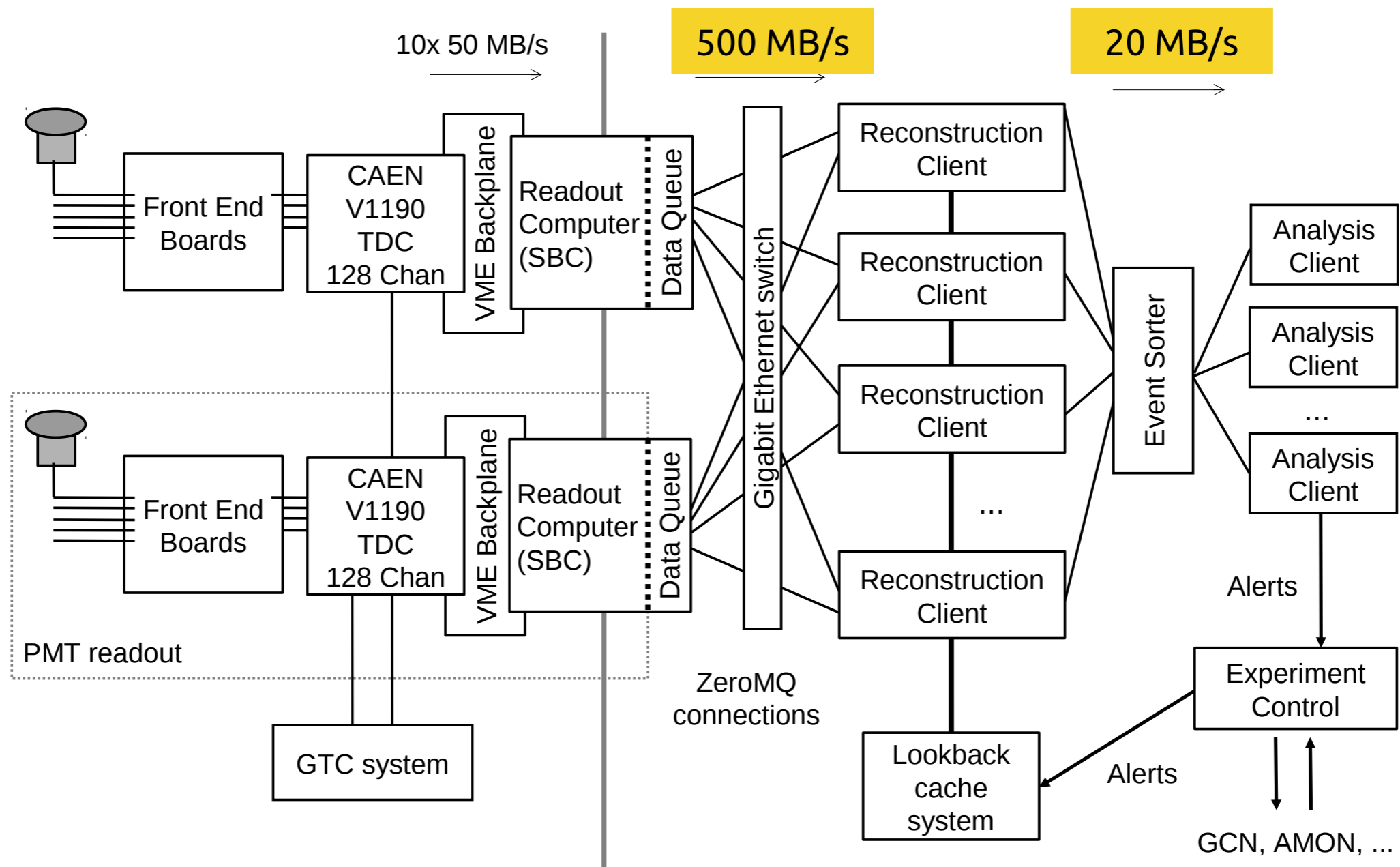
- Dedicated compute clusters
  - 500-2000 cores, 2GB RAM/core
  - Fast interconnects
- Several computing sites
  - Manual coordination of jobs
  - Possible alternative: GRID technology
    - Only time efficient for a larger number of sites
- Distributed file system
  - Multiple servers in parallel
  - Hierarchical storage (under development)
    - active data: reconstructed events
    - “archival” (less frequent access): raw data
  - Continuously growing: ~12TB / week (raw ~600TB / a)

# Solution

- Dedicated compute clusters
    - 500-2000 cores, 2GB RAM/core
    - Fast interconnects
  - Several computing sites
    - Manual coordination of jobs
    - Possible alternative: GRID technology
      - Only time efficient for a larger number of sites
  - Distributed file system
    - Multiple servers in parallel
    - Hierarchical storage (under development)
      - active data: reconstructed events
      - “archival” (less frequent access): raw data
    - Continuously growing: ~12TB / week (raw ~600TB / a)
- Dominant cost
- 



# HAWC online: DAQ y taza de datos



- Procesamiento en servidores estándar
- Reducción en línea y reconstrucción preliminar
- Taza final determinado por costo de operación vs beneficio científico



# Pasos

- Toma de datos
  - reducción
  - primera reconstrucción (➡ alertas, control de calidad)
  - Almacenamiento temporal
- Almacenamiento
- (Re-)reconstrucción
  - Geometría
  - Energías
  - Separación Gamma/Hadron
- Generación de mapas



# Data transfer

## ● HAWC to UNAM

- Currently: portable disk arrays of 12TB & 24 TB
  - one per week
- Bandwidth in national academic internet growing
  - Fibre from HAWC site to Puebla @1 Gbit/s: in progress
- Expect to move data by internet in 2018

## ● UNAM to UMD (Maryland)

- UNAM operates 1Gbit/s to Internet 2
  - We use 60%–80%
  - **Upgrade to 10Gbit/s in process**
- CUDI: 3-10 Gbit/s
- Link tuned with the help of staff from ICN-UNAM
- We can move all HAWC data from Mexico to the US by net



# Storage

- Data volume growing (data and simulation)  
~700-1000 TB per year
- We have already almost 3 PB collected
- Data sets of this size new for Mexico
  - Traditional super-computing CPU oriented
- **Possible to handle in an institute**
- **Share** experience and, if possible, equipment
  - Particle physics
  - Astronomy
  - Geosciences
  - Life sciences, bio-informatics

# El centro de datos en el ICN-UNAM

## ● Proyectos apoyados

- HAWC
- ALICE
- Auger
- DESI
- locales

### Nodos de servicios

Nombre	Tipo	CPU	cores *	Memoria	IB	Ethernet	Disco
xook	Nodo de login	2 x Xeon E5-2640 v3 @ 2.60 GHz	32	128 GB	FDR	1 Gbps	SSD 120 GB
piotl	Nodo de login HAWC	2 x Xeon E5-2620 v3 @ 2.40 GHz	24	128 GB	FDR	10 Gbps	SSD 120 GB
tochtli64	Servidor PBS	2 x Xeon E5520 @ 2.27 GHz	16	16 GB	-	1 Gbps	SSD 120 GB
cihuatontli	VoBox para ALICE	1 x Xeon E3-1240 V2 @ 3.40 GHz	8	8 GB	-	1 Gbps	SSD 120 GB
home	/home por NFS	1 x Xeon E5-2690 v3 @ 2.60 GHz	24	64 GB	-	10 Gbps	RAID 24 TB

\*Todos los nodos tienen HT habilitado



# El centro de datos en el ICN-UNAM

## Nodos de procesamiento

Nombre	Tipo	CPU	cores*	Memoria	IB	Ethernet	Disco	Total nodos	Rendimiento teórico Tflops
tochtli64-[01-02,04,09,15,34-42]	Intel v4	2 x Xeon E5-2680 v4 @ 2.40 GHz	56	128 GB	FDR	1 Gbps	HDD 500 GB	14	15.05
tochtli64-[03,05-08,10-14]	Intel v3	2 x Xeon E5-2690 v3 @ 2.60 GHz	48	128 GB	FDR	1 Gbps	HDD 500 GB	10	9.98
tochtli64-[24-33]	Intel v3	2 x Xeon E5-2680 v3 @ 2.50 GHz	48	128 GB	FDR	1 Gbps	HDD 500 GB	10	9.60
tochtli64-[16-23]	AMD	2 x AMD Opteron(tm) Processor 6378	32	64 GB	QDR	1 Gbps	HDD 500 GB	8	2.45
Total			<b>2000</b>					42	<b>37.08</b>

\*Todos los nodos tienen HT habilitado

## Nodos de almacenamiento

Nombre	Tipo	CPU	cores*	Memoria	IB	Ethernet	Discos	Total nodos	Capacidad teórica TB	Capacidad real TB
mdsA	Servidor de metadatos Lustre (HA)	2 x Xeon(R) CPU E5-2680 v3 @ 2.50GHz	24	128 GB	FDR	1 Gbps	RAID 10	1	7.2 TB	3.6
mdsB	Servidor de metadatos Lustre (HA)	2 x Xeon(R) CPU E5-2680 v3 @ 2.50GHz	24	128 GB	FDR	1 Gbps	Mismo RAID del MDSA	1		
OSS001-A	Servidor OSS Lustre	1 x Xeon E5-2683 v3 @ 2.0 GHz	28	128 GB	FDR	1 Gbps	24 X SAS3 6TB	20	2880	2330
OSS001-B	Enclosure en cascada del OSS tipo A	SIN CPU					24 X SAS3 6TB	20	2880	2330
Total									5760	4660

- El sistema de archivos del almacenamiento es Lustre version 2.8
- El sistema de archivos de cada OSS es ZFS
- El rendimiento probado del almacenamiento Lustre es de:
  - Test IOZONE: 20 GB/s en escritura, 12 GB/s en lectura
  - Medido en producción: 2 GB/s en lectura, sostenido por 5 hrs.

# El centro de datos en el ICN-UNAM

## Nodos de procesamiento

Nombre	Tipo	CPU	cores*	Memoria	IB	Ethernet	Disco	Total nodos	Rendimiento teórico Tflops
tochtli64-[01-02,04,09,15-21,27]	Intel	2 x Xeon E5-2680 v4 @ 2.40 GHz	56	128 GB	FDR	1 Gbps	HDD 500 GB	14	15.05
tochtli64-[14]									9.98
tochtli64-[15]									9.60
tochtli64-[16]									2.45
Total									37.08

Nombre	Tipo	CPU	cores*	Memoria	IB	Ethernet	Disco	Total nodos	Rendimiento teórico Tflops
OSS001-A	Servidor OSS Lustre	1 x Xeon E5-2683 v3 @ 2.0 GHz	28	128 GB	FDR	1 Gbps	24 X SAS3 6TB	20	2880
OSS001-B	Enclosure en cascada del OSS tipo A	SIN CPU					24 X SAS3 6TB	20	2880
Total									5760

### Resumen:

42 nodes

1000/2000 núcleos/hilos (HT)

~ 37.1 Tflop/s (teórico)

4.3PB lustre (5.7PB raw)

Test IOZONE: 20 GB/s en escritura, 12 GB/s en lectura

Medido en producción: 2 GB/s en lectura, sostenido por 5 hrs.

Herramientas para control y monitoreo del lustre desarrollados en el ICN.

\*Todos los

Nodos

- El sistema de archivos del almacenamiento es Lustre version 2.8
- El sistema de archivos de cada OSS es ZFS
- El rendimiento probado del almaceamiento Lustre es de:
  - Test IOZONE: 20 GB/s en escritura, 12 GB/s en lectura
  - Medido en producción: 2 GB/s en lectura, sostenido por 5 hrs.



# Fisicamente: 10 racks







# Auger Computing





# Data set

- About 30TB of data in 10 years of operation
- Upgrade will double data rate
  - 60TB in the next 10 years
- Simulation set ~100TB
  - Reasonably stable
    - software evolves
    - New simulations replace outdated ones
    - Needs management tools
  - Increase due to upgrade possible
- Manageable data set for most institutions
  - Mostly subsets needed

# Laboratorio de Datos y Modelos Científicos

- Colaboración entre expertos
- Involucrar expertos del **dominio** y **en cómputo**
- Apoyo a **diseño y realización** de soluciones para proyectos
  - Computación
  - Almacenamiento
    - Plan de datos
    - Acceso a datos
  - Conectividad
- **Desarrollo**
  - Código
  - Tecnología
- **Entrenamiento**
  - Posgrado(s)
  - Cursos, diplomados, ...
- Independencia de (grandes) proveedores / complementar conocimiento
- Acercamiento a **nuevas comunidades**

# Preservación de datos

- Almacenamiento confiable
  - Raid 6
  - CEPH con multiples copias
  - Respaldo local
    - Cinta
    - Disco economico
- Multiples copias de datos
  - Centros geograficamente distribuidas
  - Beneficio extra: acceso a datos acelerado
- Planes a largo plazo
  - En desarrollo
  - Retos
    - financiamiento
    - Operación



# Acceso a datos

- Abierto para colaboradores
- Publicar datos con publicaciones
- Distribución abierta de sub-conjuntos
  - Datos de N años anteriores
  - Subconjunto de datos
  - Destino:
    - Comunidad científica
    - Divulgación
- Ejemplo claro: Observatorio Fermi
  - Software
  - Datos después de un año



# HAWC datos publicos

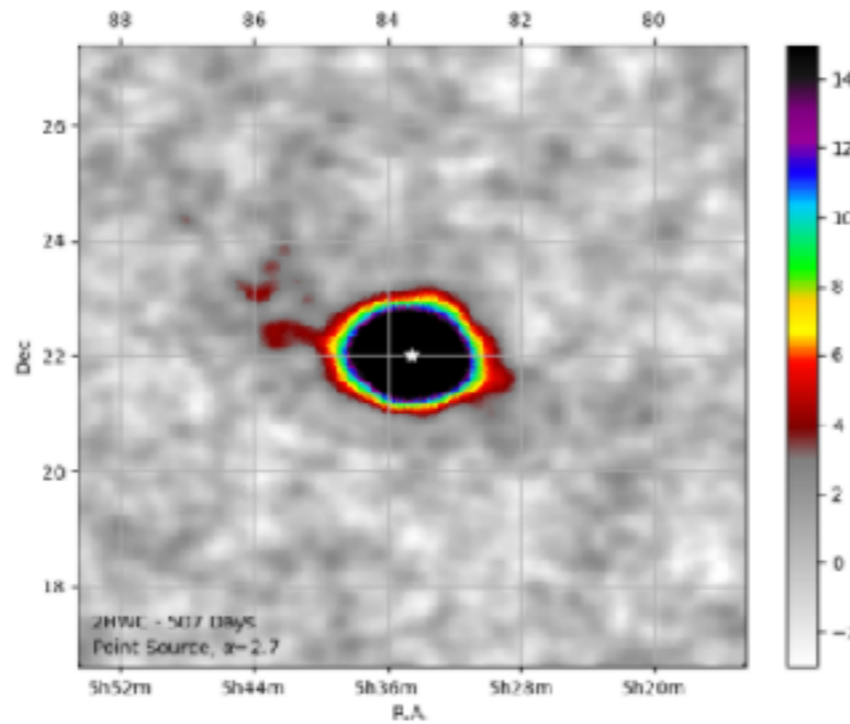
<https://www.hawc-observatory.org/data/>

HAWC Observatory    Publications    Public Datasets    Resources

- Intro
- 2HWC Survey
- Geminga Paper
- Lightcurves

## 2HWC Survey

[Details](#)    [Catalog View](#)    [Coordinate View](#)



Point Source (2.7 Index)

RA (J2000 - Decimal):

03.03

Dec (J2000 - Decimal):

22.01

Significance

104.958443

[sqrt(TS)]

Flux

1.752774e-13

[cm<sup>-2</sup> s<sup>-1</sup>  
TeV<sup>-1</sup>]

Flux Upper

1.796217e-13

[cm<sup>-2</sup> s<sup>-1</sup>  
TeV<sup>-1</sup>]

Flux Lower

1.709344e-13

[cm<sup>-2</sup> s<sup>-1</sup>  
TeV<sup>-1</sup>]

Read the [Details Tab](#) for explanation of these numbers.

# Retos para el futuro

- Generar planes de datos
- Interés en explorar datos propios
  - Tiempo de acceso exclusivo para la colaboración
- Políticas
  - Genera requerimientos
  - Financiamiento
- Infraestructuras
- Discusiones activas en proyectos