GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN
Ciemat
Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

Comisión Nacional de Energía Atómica

this talk

contact me

LAGO INDICA
INfraestructura DIgital de Ciencia Abierta

LAGO NEXT

Hernán Asorey[1,2] for the LAGO Collaboration

BGA, Colombia, 20/Nov/2024

[1] *Departamento de Tecnología, CIEMAT, Madrid, Spain.*

[2] *Departamento Física Médica, CNEA, San Carlos de Bariloche, Argentina*

hernanasorey@cnea.gob.ar

LAGO next

LAGO edge

When? Who?

LAGO plug&play

When? Who?

LAGO data

When? Who?

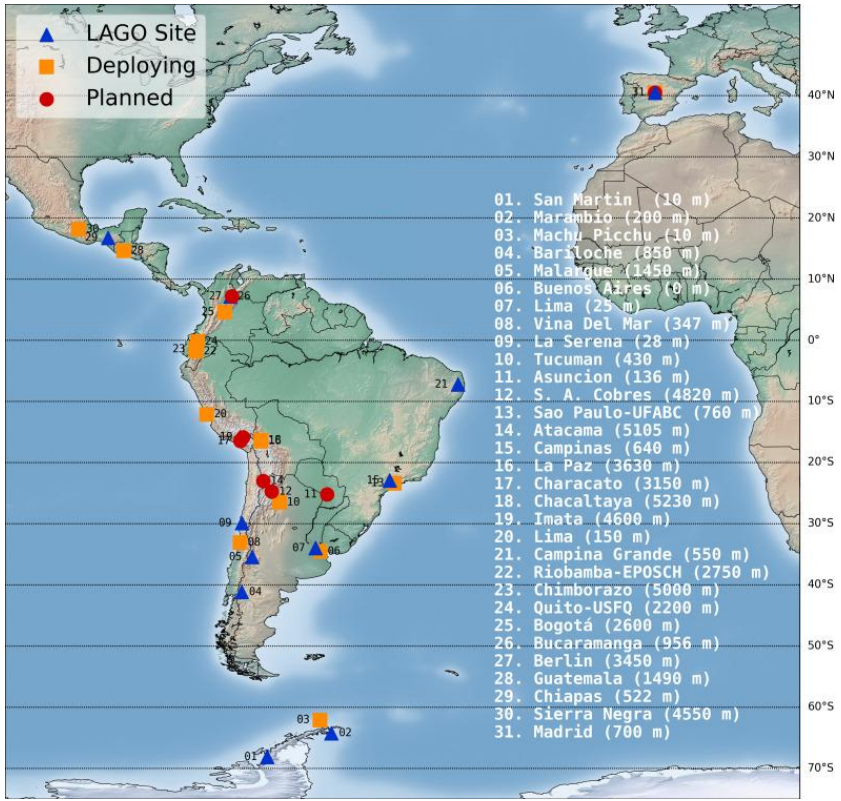
let's talk on this next later today

Conclusions

- **We've come a long way, but there's still plenty left to do.**
- My perspective is that:
  - data should be automatically and meticulously curated;
  - curated data should be fully compliant with FAIR principles;
  - curated data should be open and public (embargo?);
  - operative data should be available in (near) real-time.
- The WCD should be:
  - Simple, reliable, affordable, and intelligent—not just "smart."
  - Designed to be plug-and-play, requiring minimal setup.

It's easy to say… but hey, I'm just a computational physicist

The Latin American Giant Observatory



LAGO is an extended astroparticle observatory at continental scale: from Mexico to Antarctica and Spain!

- **Astroparticle physics to study the extreme universe**
- **Transient and long term space weather phenomena trough Solar modulation of Cosmic Rays**
- **Measurement of background radiation at ground level**
- **Academic goals**
  - **HEP and astro-ph seedbed for Ibero american students**
  - **Build an Ibero-American network of astroparticle researchers**

*I. Sidelnik & H. Asorey, NIM A **876** 173, 2017; H. Asorey et al., PoS ICRC2015 247, 2015*

The Latin American Giant Observatory

LAGO is an extended astroparticle observatory at continental scale: from Mexico to Antarctica and Spain!

- Astroparticle physics to study the extreme universe
- Transient and long term space weather phenomena trough Solar modulation of Cosmic Rays
- Measurement of background radiation at ground level
- Academic goals
  - HEP and astro-ph seedbed for Ibero american students
  - Build an Ibero-American network of astroparticle researchers

Are these still our goals?

*I. Sidelnik & H. Asorey, NIM A **876** 173, 2017; H. Asorey et al., PoS ICRC2015 247, 2015*

The five w's for each section

Why?

Whatfor?

What?

When?

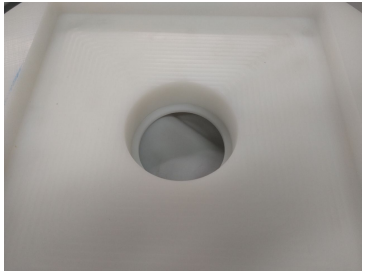
Who?

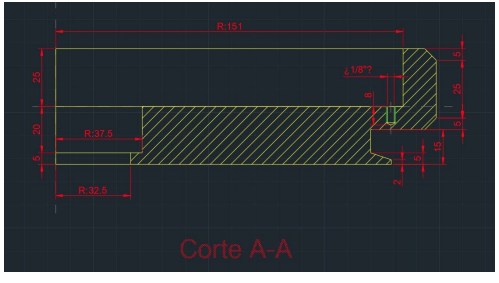
01

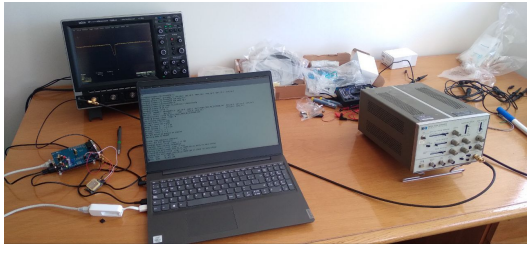
LAGO plug&play

WCD as appliances

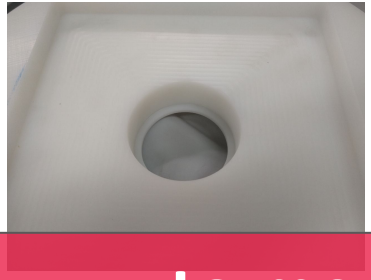
New WCD at Madrid (40.4552N, -3.72501E, 700 m asl)



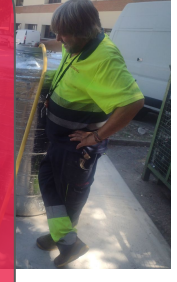


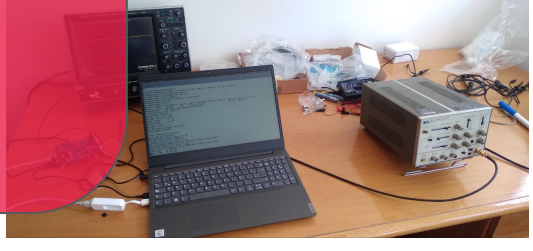


New WCD at Madrid (40.4552N, -3.72501E, 700 m asl)







to me

it was a long, painful and suboptimal process that wastes resources and time, but we had great times doing it.

Hernán, today.

Towards LAGO plug&play

Is it possible to design and build a plug & play fully functional WCD on a …

+ … appliance-like approach?

+ … IKEA-like approach?

Towards LAGO plug&play

Is it possible to design and build a plug & play fully functional WCD on a …

+ … appliance-like approach?

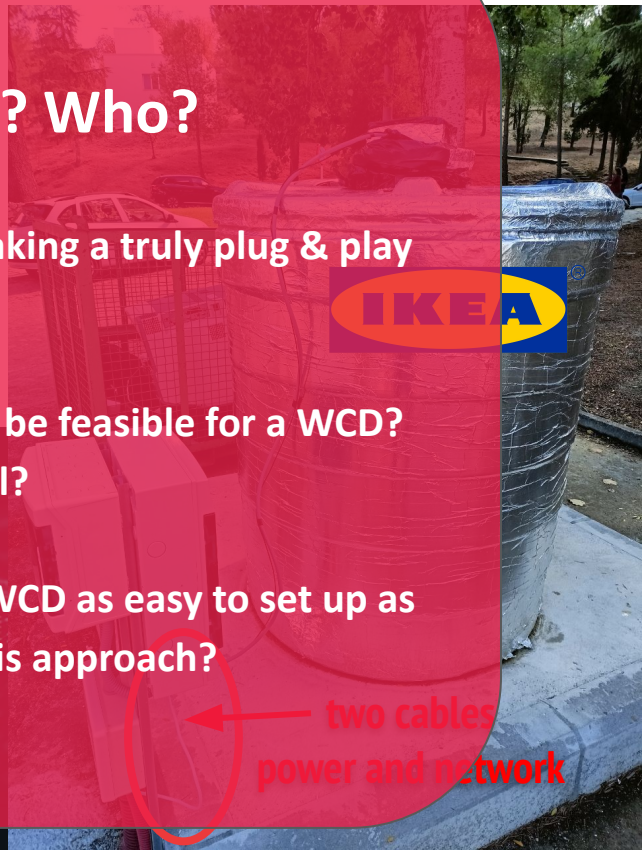
+ … IKEA-like approach?

Why? Whatfor? What? When? Who?

1. **What are the biggest challenges we might face in making a truly plug & play WCD?**

2. **Would an IKEA-like, self-assembly approach actually be feasible for a WCD? What would we need to consider to make it practical?**

3. **What features or design elements would make the WCD as easy to set up as a typical appliance? Are there any trade-offs with this approach?**



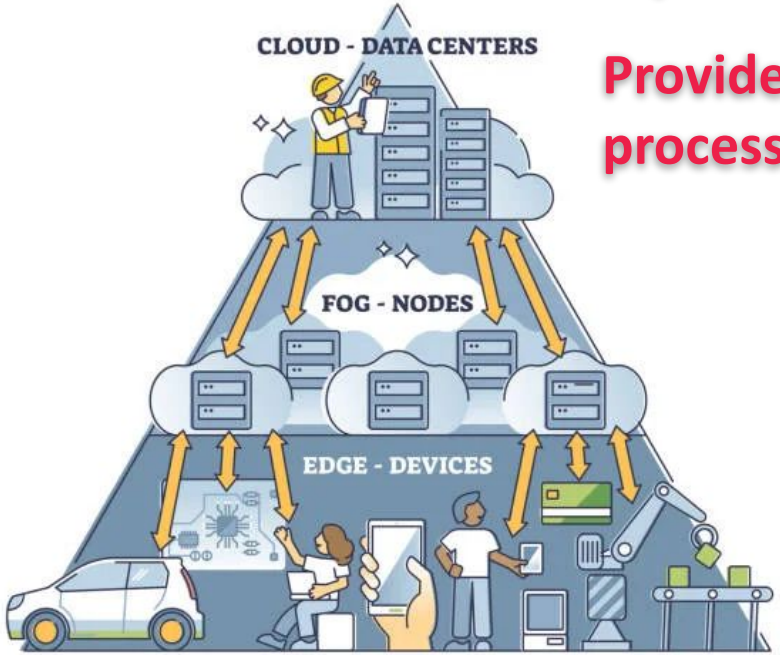
02

LAGO edge

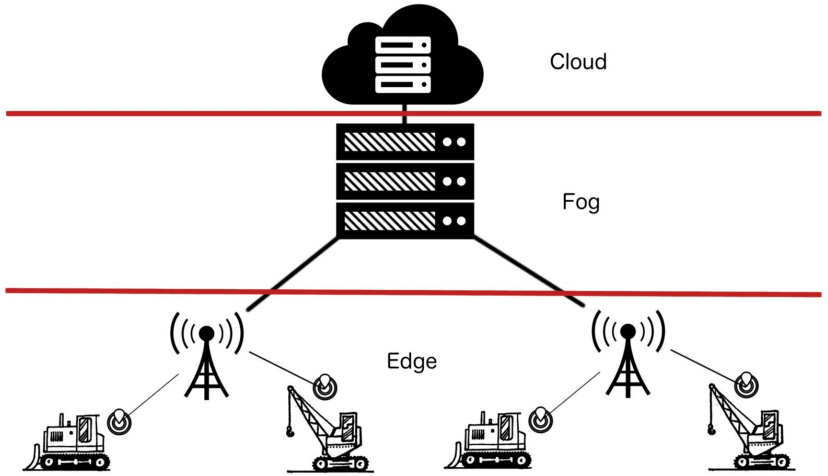
a distributed computing approach

Distributed computing

Allows data processing across multiple locations and layers to **improve efficiency, reduce latency, and optimize resource use**.



Provide faster responses and reduce data load by processing data closer to its source.



Distributed computing

edge

Data processing happens at the device level, close to the data source (e.g., IoT devices like sensors, cameras).

what for?

Minimizes latency, reduces bandwidth use, and allows real-time decisions.

e.g.

modern cars has lot of sensors and onboard computing

fog

Data processing layer between edge devices and the cloud, often within local network infrastructure.

what for?

Aggregates data from edge, performs more processing, and manages local resources.

e.g.

onroute devices aggregate data from multiple cars and process it locally

cloud

Centralized data processing and storage in a remote cloud-based data center.

what for?

Offers scalable high computing power, storage, and advanced analytics.

e.g.

At cloud, intensive data processing and training AI models is performed

Distributed computing

in LAGO

"En la Red Pitaya tenemos una ferrari y la usamos como un cinquecento"

Dennis, LAGO INDICA virtual meeting, Oct'24.

edge

Data processing happens at the device level, close to the data source (e.g., IoT devices like sensors, cameras).

what for?

Minimizes latency, reduces bandwidth use, and allows

fog

Data processing layer between the edge and the cloud, often within local network infrastructure.

what for?

Aggregates data from edge, performs more processing, and manages local resources.

e.g.

route devices aggregate ta from multiple cars and ocess it locally

cloud

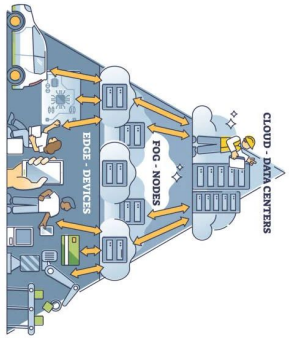
Centralized data processing and storage in a remote cloud-based data center.

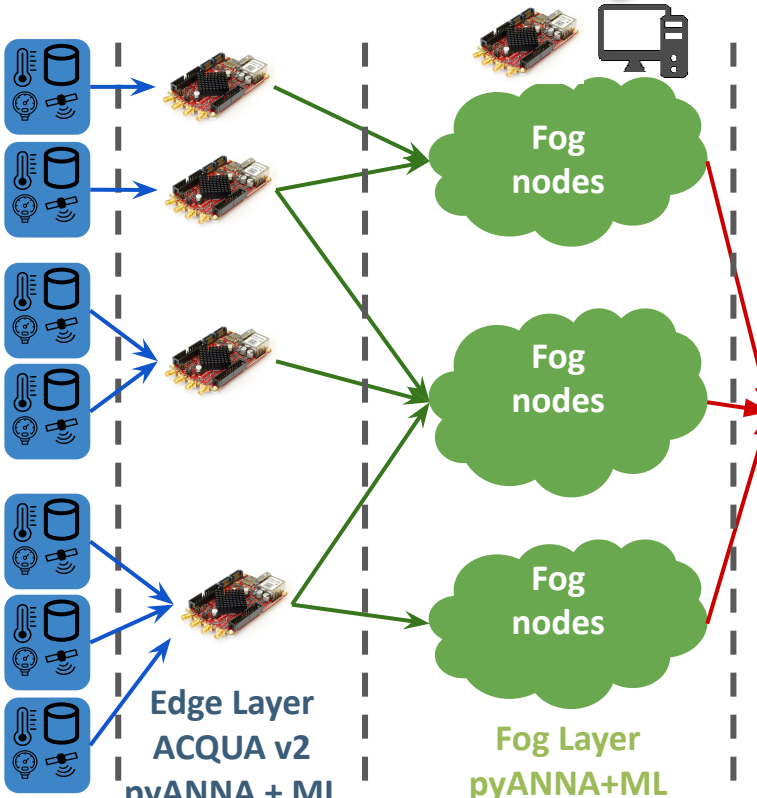
what for?

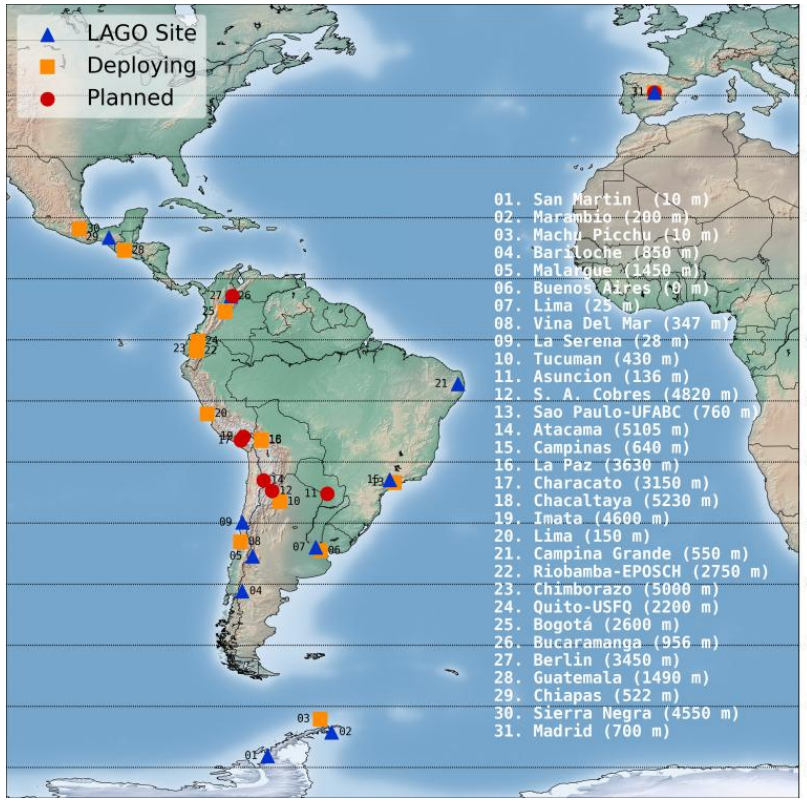
co
an

At
pr
m

Towards LAGO Edge

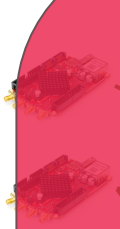


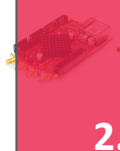


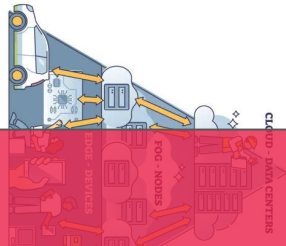


Latin America Giant Observatory
lagoproject.net

Towards LAGO Edge







nodes

Fog nodes

Fog

ONEDATA

B2HANDLE

B2FIND

Why? Whatfor? What? When? Who?

1. What should we think about when deciding where to set up each layer—Edge, Fog, and Cloud—in the LAGO network?
2. How do we make sure our data plays nice across Edge, Fog, and Cloud while keeping it FAIR?
3. What hurdles might we hit running machine learning on the Edge and Fog, and how do we make the most of our resources out there?

Edge Layer
ACQUA v2
pyANNA + ML

Fog Layer
pyANNA+ML
+FAIR?

Cloud Layer
pyAnna

LAGO
Latin America Giant Observatory
lagoproject.net

03

LAGO data

standardizing and optimizing resources

The FAIR paradigm

FAIRSFAIR
Fostering Fair Data Practices in Europe

Findability

discoverable, identifiable and locatable by means of a standard identification mechanism

PiD, Metadata

Accessibility

always available and obtainable; even if the data is restricted, the metadata is open

Cloud storage

Interoperability

syntactically parseable and semantically understandable, allowing data exchange and reuse

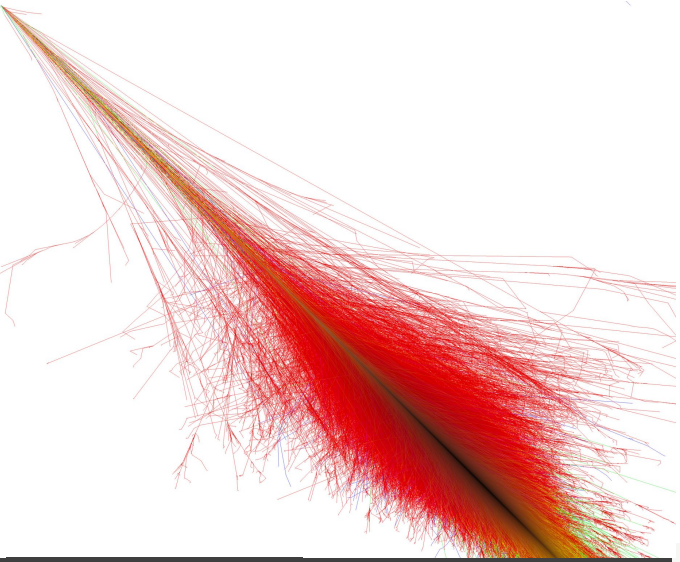
Vocabulary

Reusability

described and shared with the least restrictive licences, allowing the widest reuse possible

Harversters, copyleft

Current LAGO framework



LAGO@github

LAGO ACQUA

WCD control and data acquisition

- Nexys DAQ v1.0
- RP DAQ v2.0

LAGO ANNA

Automated and integrated data analysis

- Adaptable for DAQ
- C++ analysis tools

LAGO ARTI (+Meiga)

Astroparticle simulations

- GCR models
- Atmospheres
- Geomagnetic field
- AP detectors
- + Applications

LAGO onedataSim

LAGO data FAIRificator (currently ARTI data)

- ARTI wrapper
- metadata standardization
- FAIR compliant

Towards LAGO Data

LAGO@github

LAGO ACQUA

WCD control and DAQ
- \+ control
- \+ LAGO **EDGE**
- \+ **AI + FAIR** onboard

LAGO pyANNA

Automated and integrated data analysis
- \+ standard analysis
- \+ central data repo

LAGO pyARTI (+Meiga)

Python-based from scratch
- \+ learning curve
- \+ data analysis
- \+ auto atm & EMF
- \+ detectors
- \+ outputs & visualization
- \+ user friendliness

LAGO onedataSim

LAGO data FAIRificator
- \+ measured data
- \+ open data
- \+ FAIR compliant data

Towards LAGO Data

Why? Whatfor? What? When? Who?

1. **What are the main challenges we anticipate in adapting our current tools for distributed computing and AI integration?**

2. **How can we ensure that all data generated by LAGO is FAIR-compliant, particularly with real-time data requirements and distributed processing across Edge, Fog, and Cloud?**

3. **What specific modifications or upgrades are needed in each component (ACQUA, ANNA, ARTI, onedataSim) to fully support a plug & play WCD and integrate seamlessly with the new LAGO framework?**

LAGO@github

LAGO next

LAGO edge

When? Who?

LAGO plug&play

When? Who?

LAGO data

When? Who?

GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN
Ciemat
Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

Comisión Nacional de Energía Atómica

this talk

contact me

LAGO INDICA
INfraestructura DIgital de Ciencia Abierta

Hernán Asorey[1,2] for the LAGO Collaboration

BGA, Colombia, 18/Nov/2024

Thanks!

hernanasorey@cnea.gob.ar