Intrusion Prevention and Detection in Grid computing - The ALICE Case

21st International Conference on Computing in High Energy and Nuclear Physics

Outline:

- Introduction
- Threat model
- Intrusion prevention
- Intrusion detection
- Summary

Andrés Gómez, Camilo Lara, Udo Kebschull
for the ALICE Collaboration
IRI - Goethe University Frankfurt

andres.gomez@cern.ch

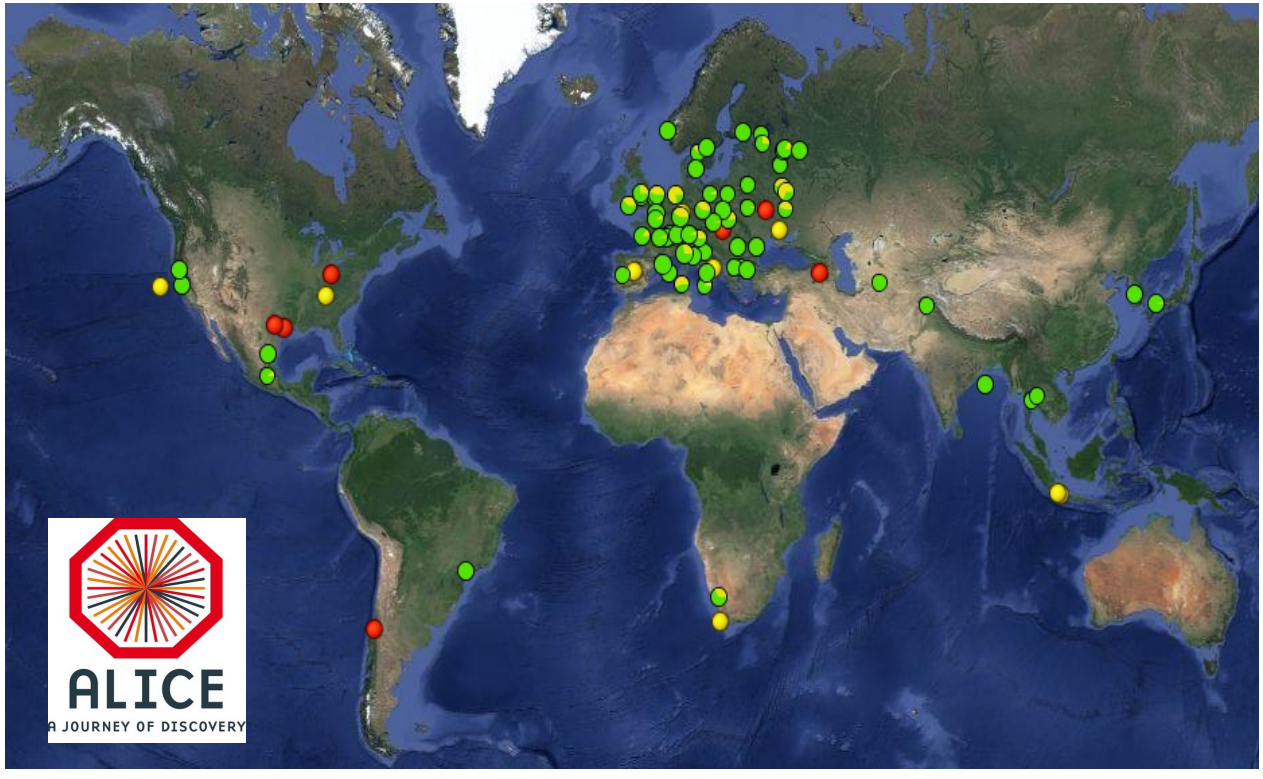
GOETHE UNIVERSITÄT FRANKFURT AM MAIN

SPONSORED BY THE

Federal Ministry of Education and Research

Introduction: ALICE Grid

- > 70 sites
- > 30 countries
- >45000 CPU cores
- > 50PB of storage
- > 1000 users



- Arbitrary code execution by design
- Huge amount of computational power and organization reputation, a goal for adversaries
- Focus on HEP → data is public but integrity is important

ALICE
A JOURNEY OF DISCOVERY

Grid Threat model

The adversary may have one or more goals:

- **Modify** experiment data
- **Attack** experiment infrastructure -> online-offline 2018
- **Abuse** Grid resources

- Steal sensitive data
- Compromise users' machines
- Denegation of service
- Damage the organization reputation

ALICE
A JOURNEY OF DISCOVERY

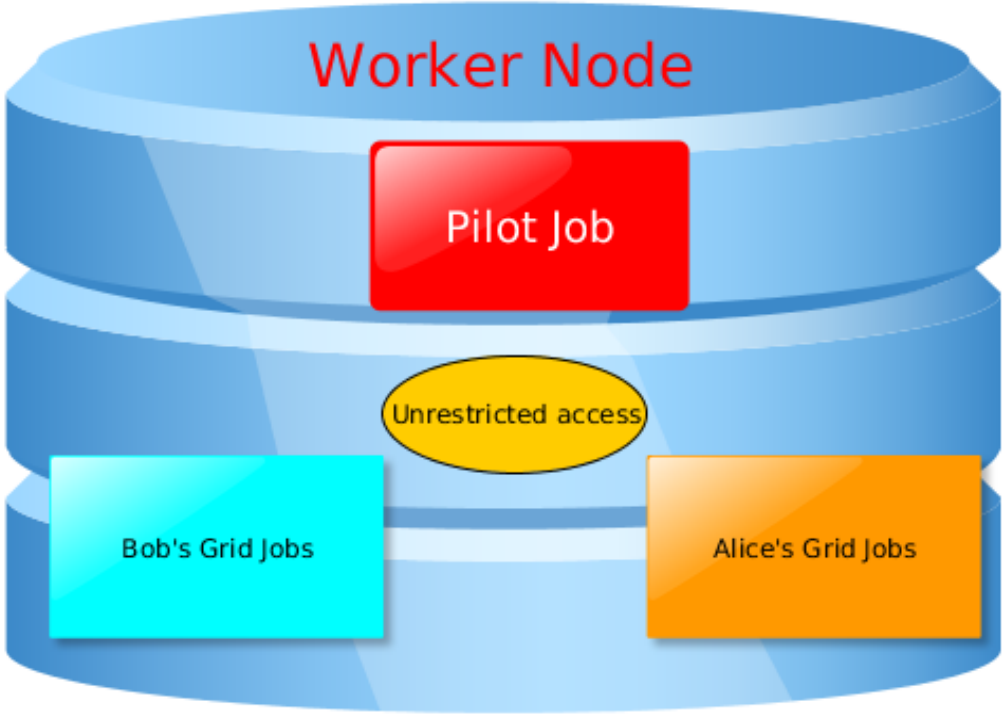
Project main goals

Improve computer security in the GRID by:

- Intrusion prevention
- Security by isolation
- Intrusion detection
- Analysis of Job behavior
- Machine learning

ALICE
A JOURNEY OF DISCOVERY

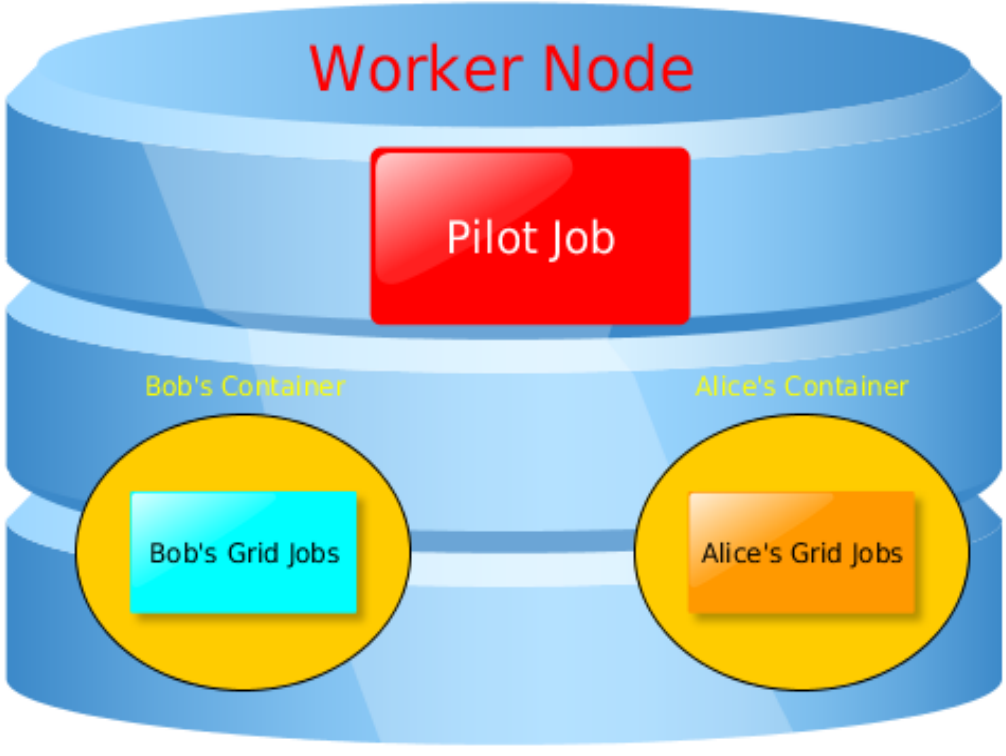
Specific Grid issues we want to address



- No separation between different levels of privileges
- Job execution environment not properly enforced
- No multi user execution
- Sensitive resources not isolated
- No automatic way of preventing and detecting intrusions

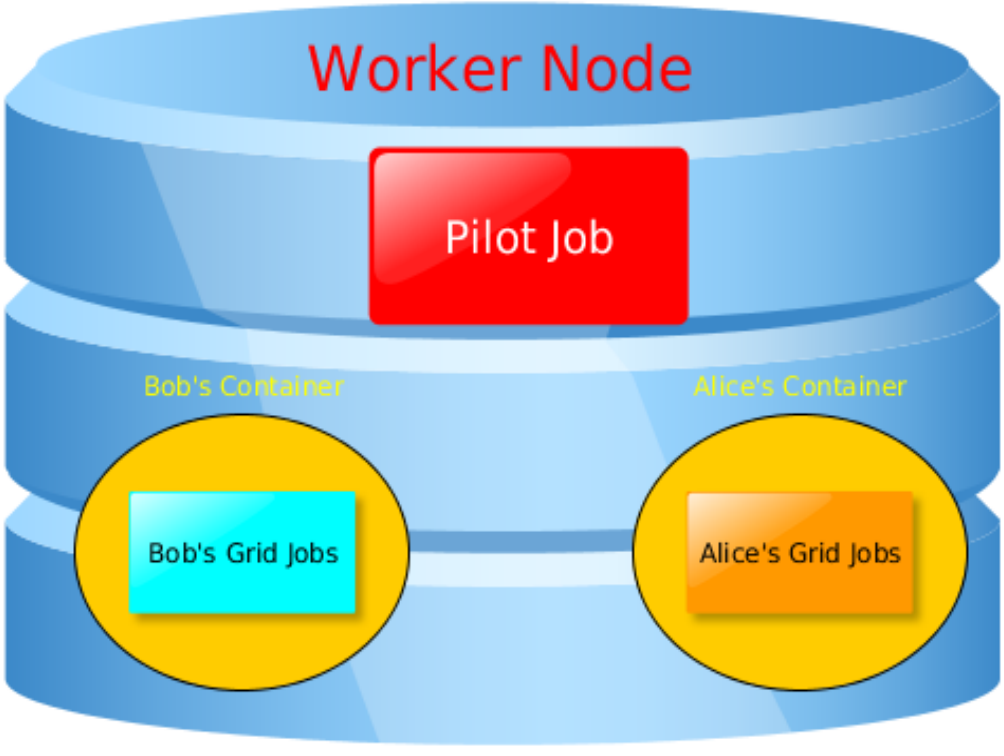
Objectives: Intrusion prevention

- We want to run the payloads in an isolated environment
- The Pilot Job would have unrestricted access to containers
- Anything running inside the container should be isolated

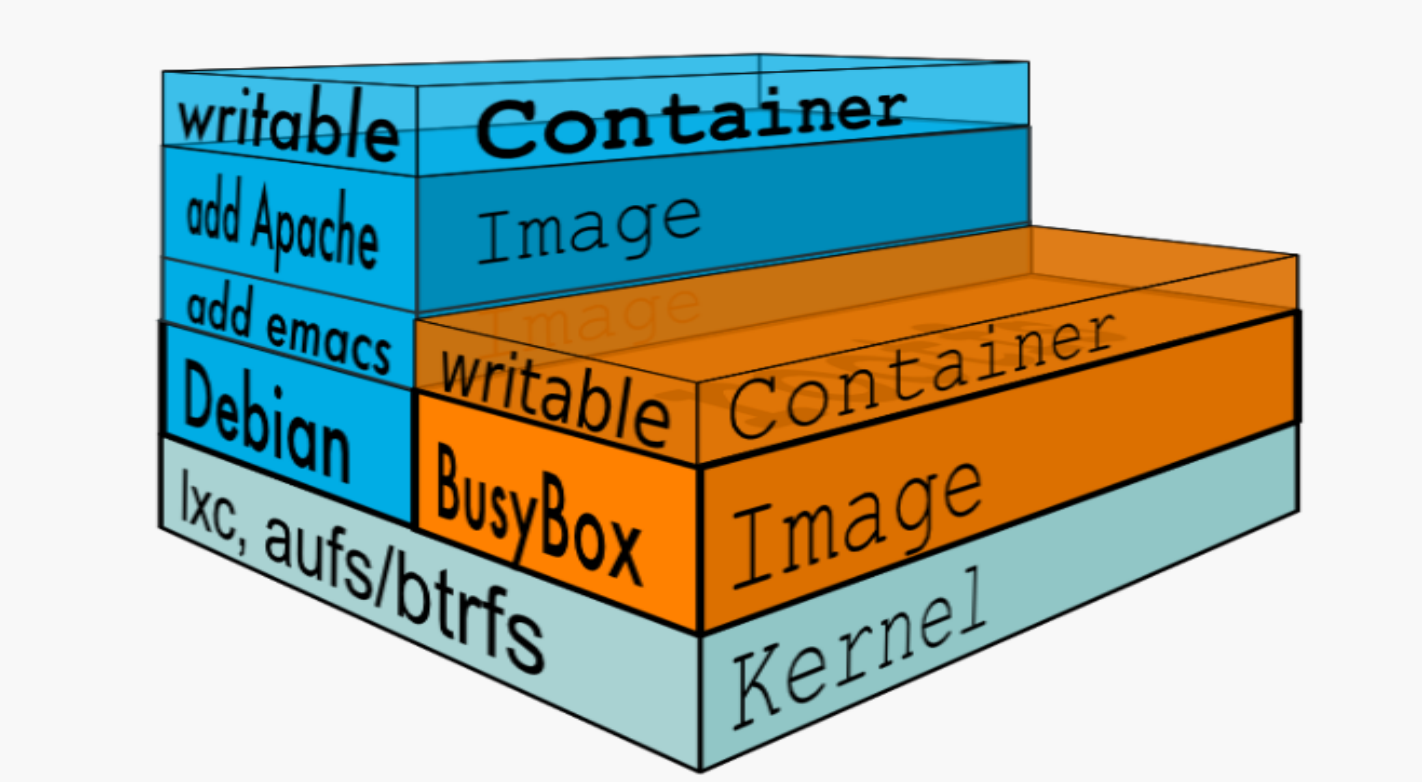


Objectives: Security by isolation

- All components run as unprivileged users
- Root emulation inside the container
- The Jobs run with less privileged Grid user
- Unprivileged Isolated Multi User Pilot Jobs
- Use containers to achieve isolation



Containers



- Lightweight, fast, disposable
- Virtual environments
- Boot in milliseconds
- Just a few MB of intrinsic disk/memory usage
- Bare metal performance is possible

Containers vs Virtual Machines: Security

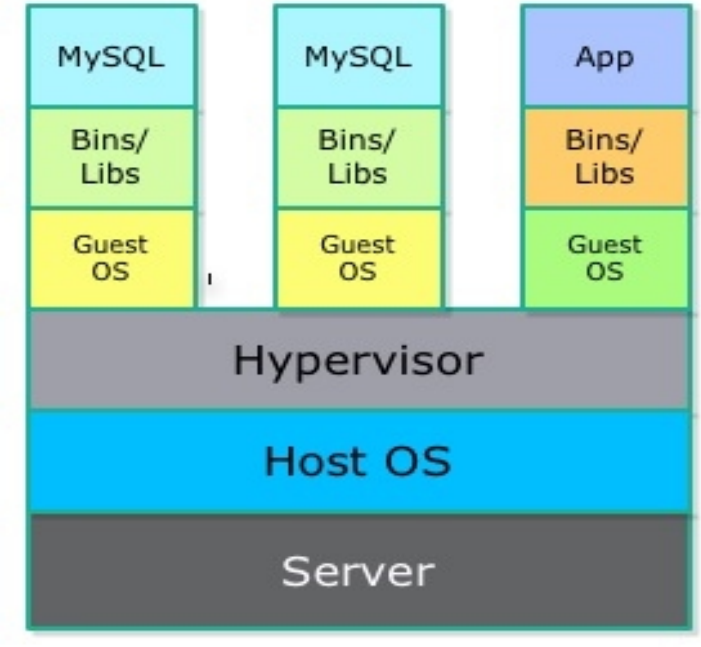
Virtual Machines:

- More layers of protection
- Huge surface of attack
- Alone, it does not solve our requirements!

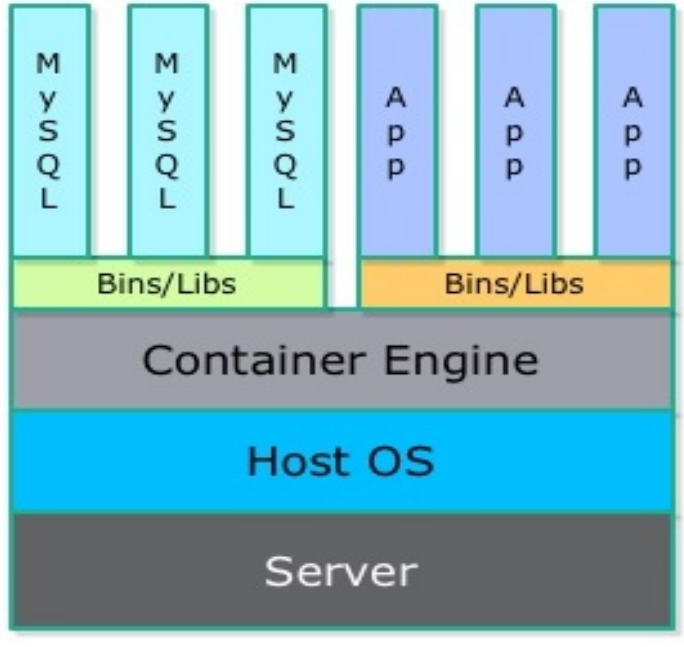
Containers:

- The kernel is directly exposed
- Less mature technology
- Reduced surface of attack
- Attenuation of kernel exposition possible
- Less time to update (kernel bugs)
- Fine-grained control

Virtual Machines



Containers



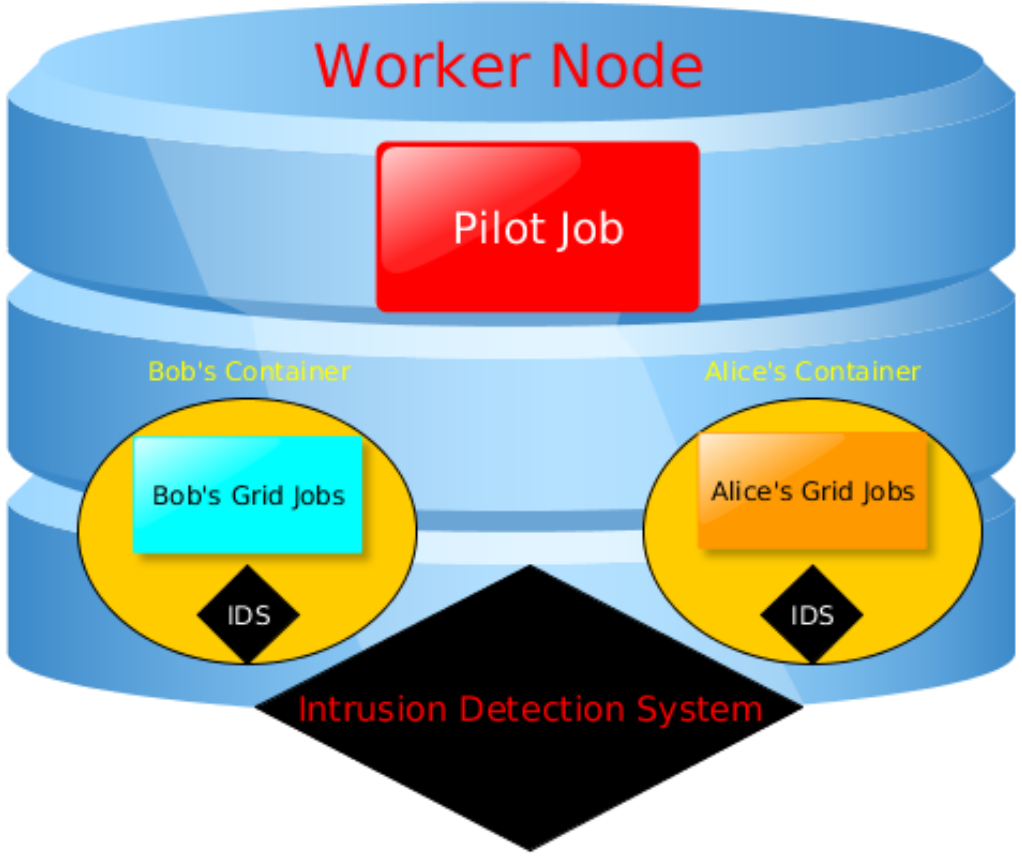
Containers: Reducing the surface of attack

- Again: Use unprivileged user and containers!
- Use **Seccomp-bpf** to filter available system calls
  - **Sandboxes**
  - Tor
  - Firefox
  - Chrome
- Use LSM technologies like Appamor
- Optionally: use Grsecurity Linux kernel patch
- Optionally: use containers over VMs

ALICE
A JOURNEY OF DISCOVERY

Objectives: Intrusion detection

- **Measure Job behavior**
- **Raise alarms** on possible attacks
- **Adapt** to dynamic environment
- Several metrics:
  - Job and system logs
  - System calls sequence
  - Common monitoring data



Intrusion detection: Machine learning

- Common IDS use **fixed** rules
- Machine learning methods can help to generalize
- Analyze **"normal"** behavior vs **"malicious"** behavior
- **Train** AI algorithm
- Specific algorithm under research

Project steps

Done

- AliEn grid running in a single machine
- Framework modified to execute Jobs inside an unprivileged container

Todo

- Create a custom site for security testing - 2015
- Modify Alien/JAlien to fully execute Jobs in containers - 2015
- Research on Machine Learning for IDS – 2015/2016
- Develop a complete prototype - 2016

ALICE
A JOURNEY OF DISCOVERY

Challenges

- Security vs performance
- What if we consider private data
- What if we consider external attacks
- How to analyze the huge amount of trace/logs data generated in a efficient way
- How to share information between several components of the Grid
- Reduce the amount of false positives and negatives

ALICE
A JOURNEY OF DISCOVERY

Summary

- Job execution environment in the Grid has to be hardened
- Containers provide security by isolation among the Grid components and the underline machine
- We have to detect intrusions coming from Jobs
- Even if a new attack method is used

ALICE
A JOURNEY OF DISCOVERY

GOETHE UNIVERSITÄT FRANKFURT AM MAIN

Thank you!

Questions?

SPONSORED BY THE

Federal Ministry of Education and Research