

# Virtualización y Contenedores en Cómputo Científico

Data Pub  
ITESM

15 de mayo de 2024



# Centros de Super Cómputo

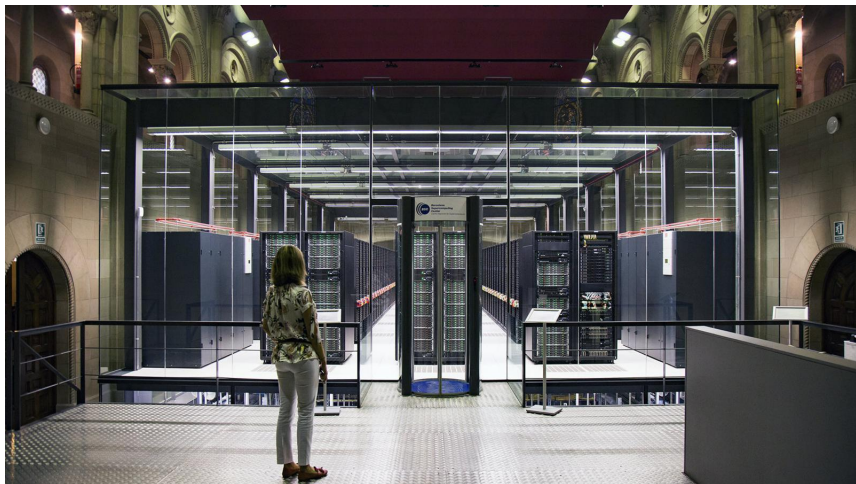
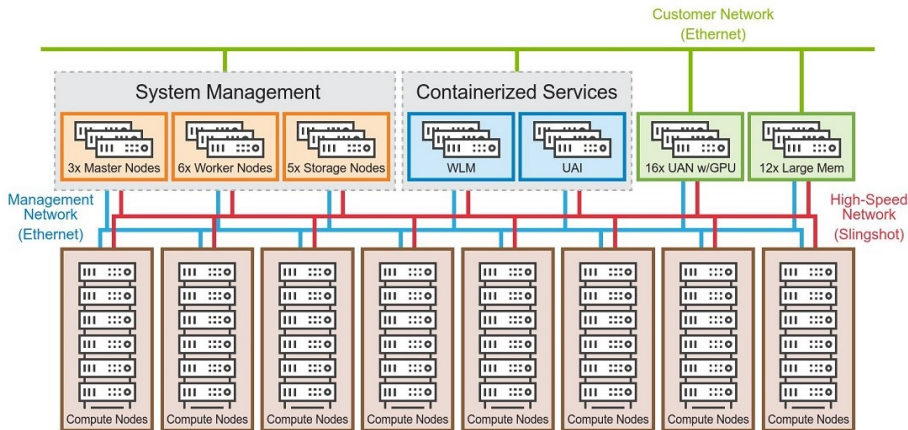


Figura: Barcelona Supercomputing Center

## Centros de Super Cómputo : Recurso compartido

- **Usuaría-Clienta.** Relacionadas con un grupo o institución.
- **Administradoras.** Satisfacen distintas necesidades de las clientas:
  - ▶ Versiones de software.
  - ▶ Compiladores.
  - ▶ Módulos.
  - ▶ Sistemas de archivos.
  - ▶ Permisos.

# Arquitectura de un Cluster de HPC



**HPC Architecture Connection Diagram**

Figura: Tomado de <https://www.gigabyte.com/Article/high-performance-computing-cluster>

# ¿Cómo se administra un Cluster con miles de nodos?

Herramientas de configuración automatizada



Job Scheduler o Resource Management System(RMS)



# Dos visiones

## Administradoras

Se aseguran que la clienta-usuario tenga las herramientas y soporte necesario para el uso eficiente de los recursos.



## Usuarios

Consume los recursos.



Todo parece estar bien ...



# La realidad es otra





# Problema

- Como en todo aquello donde existe el uso de un recurso compartido, las relaciones Sysadmin-Usuaría, Usuaría-Usuaría no son siempre cordiales.
- Hay conflicto : ¿Quién y cuándo tiene acceso al recurso? ¿Hay asignación justa de recursos? ¿Hay usuarios con mayores privilegios?
- **Sysadmin** : Definen **Reglas-Políticas de uso** con el objetivo de satisfacer las necesidades de las usuarias pero también acorde a las necesidades de mantener un recurso usable y confiable.
- **Usuaría** : Estas reglas y políticas se traduce en sistemas y software limitantes e inmóviles (Kurtzer et al., 2017).

# Problema

*This static nature coupled with distribution-specific software builds meant that service providers would ultimately end up limiting the scope of computational science that their systems could support (Kurtzer et al., 2017)*

# Propuesta de solución(Kurtzer et al., 2017)

## **Ambientes portables**

- Máquinas virtuales.
- Contenedores.

# Máquinas virtuales

# KVM



Proporcionan un ambiente completo, incluyendo dependencias de software, bibliotecas, datos que pueden ser encapsulados y ejecutados donde sea.

## Problema

Introduce una sobrecarga computacional considerable debido al nivel requerido de virtualización para emular el sistema operativo y el kernel.

# Contenedores



Características de virtualización ligera en el kernel de Linux hace posible construir virtualización más ligera.

Las características específicas del kernel son las que se ocupan de aislar procesos, en particular (Nemeth et al., 2018):

- **Namespaces** : aíslan los procesos del contenedor desde la perspectiva de las características del sistema operativo.
- **Control groups (cgroups)** limita el uso de los recursos del sistema y prioriza ciertos procesos sobre otros.

# References I

- Kurtzer, G. M., Sochat, V., and Bauer, M. W. (2017). Singularity: Scientific containers for mobility of compute. *PloS one*, 12(5):e0177459.
- Nemeth, E., Snyder, G., Hein, T. R., Whaley, B., and Mackin, D. (2018). Unix and linux system administration handbook. *USENIX Open Access Policy*, 59.