

Recogida y análisis de métricas de monitorización de clusters de computación Y Desarrollo de aplicaciones web con Python/Flask

Daniel Rosique Egea - Graduado en Ingeniería Informática
Contratado Prácticas Empleo Joven, CSIC 2025



MINISTERIO
DE CIENCIA, INNOVACIÓN
Y UNIVERSIDADES



CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS



VNIVERSITAT
ID VALÈNCIA



INSTITUT DE
FÍSICA
CORPUSCULAR



EXCELENCIA
SEVERO
OCHOA 1



Sobre Mí:

Contratado Prácticas Empleo Joven, CSIC

Inicio del contrato: 17/7/2025

Finalización: 16/7/2026

Tareas a realizar:

- Desarrollo de pequeñas aplicaciones WEB con Python/Flask/SQLAlchemy para apoyo a administración e infraestructuras de cálculo científico.
- Recogida y análisis y presentación de métricas de monitorización con InfluxDB, Elasticsearch y Grafana y otras herramientas de clusters de computación y redes.

Estudios: Graduado de Ingeniería Informática (UCAM) 2019-2023 y Estudiante de Máster en Robótica (UMH) 2024-202x

Recogida, análisis y presentación de métricas de monitorización con InfluxDB, ElasticSearch, Grafana y otras herramientas de clusters de computación y redes.



MINISTERIO
DE CIENCIA, INNOVACIÓN
Y UNIVERSIDADES



CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS



VNIVERSITAT
DE VALÈNCIA



INSTITUT DE
FÍSICA
CORPUSCULAR



EXCELENCIA
SEVERO
OCHOA 3



Recogida, análisis y presentación de métricas (monitorización)

Recogida:

- Obtención de datos y métricas de los servidores y servicios.

Análisis:

- Identificación de relaciones, patrones o posibles problemas a partir de las métricas.

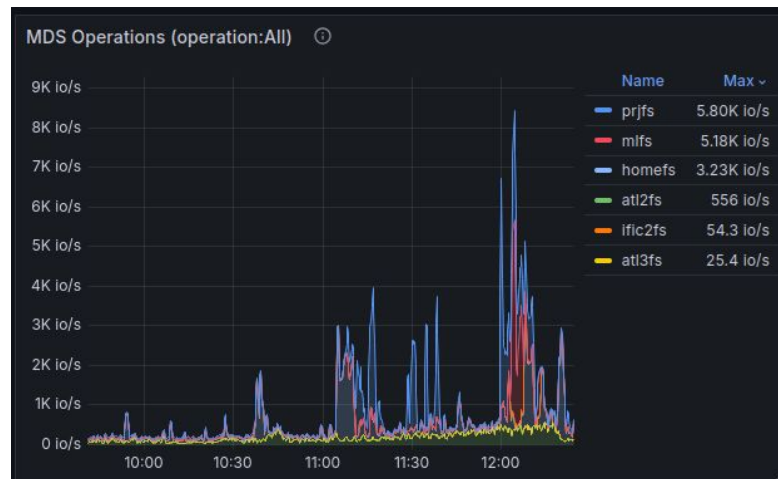
Presentación:

- Visualización clara y accesible de los datos, necesaria para comprender y analizar fácilmente el sistema.

Importancia de la Presentación



Mala presentación



Buena presentación



Necesidad de monitorización

Dependemos de los servidores ubicados en la sala de cómputo para el desarrollo de los proyectos del Instituto:

- Gluon
- Artemisa
- Atlas
- Graviton
- ...

La caída o ralentización de estos servidores puede provocar:

- Pérdida de datos
- Filtración de información confidencial
- Interrupción de proyectos críticos

Necesidad de monitorización II

La monitorización nos permite obtener métricas sobre:

- **Uso de CPU / GPU / Memoria**

Estas métricas nos sirven para:

→ detectar procesos atascados, memory leaks, ajustar supervisores y cuotas.

- **Uso de almacenamiento**

→ anticipar saturación y planificar ampliaciones.

- **Uptime**

→ detectar caídas de servicio y medir su duración.

- **Número de peticiones y respuestas** → identificar pérdidas de conexión o ataques DoS.



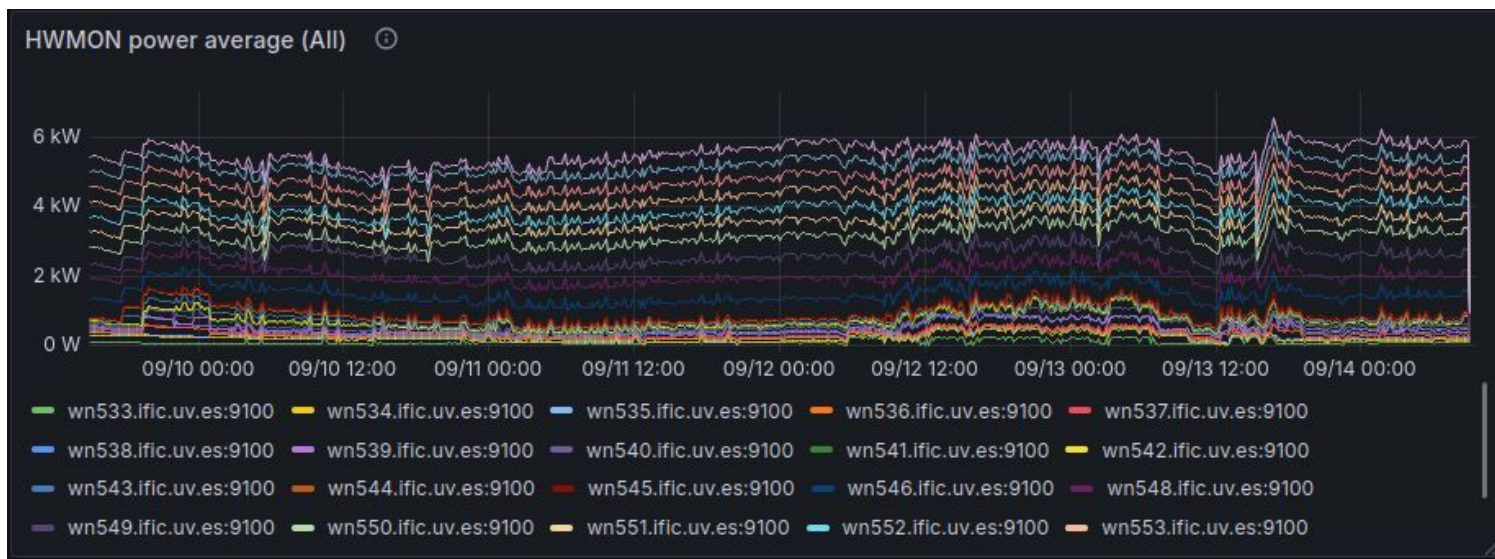
Beneficios de la monitorización

- **Optimización del rendimiento** → detectar cuellos de botella.
- ***Uso eficiente de recursos** → reducir costes y sobrecargas.
- **Detección proactiva de problemas** → alertas tempranas antes de caídas.
- **Resolución rápida de incidencias** → menos tiempo de inactividad.
- **Planificación de capacidad** → prever necesidades futuras.
- **Mejora en seguridad y cumplimiento** → detectar accesos anómalos.

Todos estos beneficios ayudan a minimizar el impacto de los problemas sobre los usuarios y detectar la causa de estos para evitar que vuelvan a ocurrir.

<https://www.ibm.com/think/topics/kubernetes-monitoring>

Metricas de consumo energetico





Herramientas de Monitorización - Grafana



¿Qué es?

- Plataforma de visualización de métricas y logs.
- Permite integrar datos de múltiples fuentes (**Prometheus, InfluxDB, Elasticsearch, Graphite**, etc.).

¿Qué ofrece?

- Dashboards personalizables para representar métricas en tiempo real.
- Alertas configurables (email, Slack, etc.).
- Exploración interactiva de datos históricos y actuales.
- Compatible con clusters, redes y aplicaciones.

Grafana se encuentra en un servidor “antiguo” y es necesario actualizar la instancia y migrarla*

<https://grafana.com/>

Herramientas de Monitorización - Grafana





Herramientas de Monitorización - InfluxDB



¿Qué es?

- Base de datos orientada a series temporales (time-series DB).
- Actualmente utilizada como recolector de métricas (en proceso de migración a Prometheus).

¿Qué ofrece?

- Almacenamiento eficiente de datos en tiempo real (CPU, memoria, red, logs).
- Lenguaje de consulta propio (InfluxQL / Flux).
- Integración con herramientas de visualización como Grafana.
- Buen rendimiento para consultas históricas y análisis de tendencias.

<https://www.influxdata.com/>



Herramientas de Monitorización - Prometheus



¿Qué es?

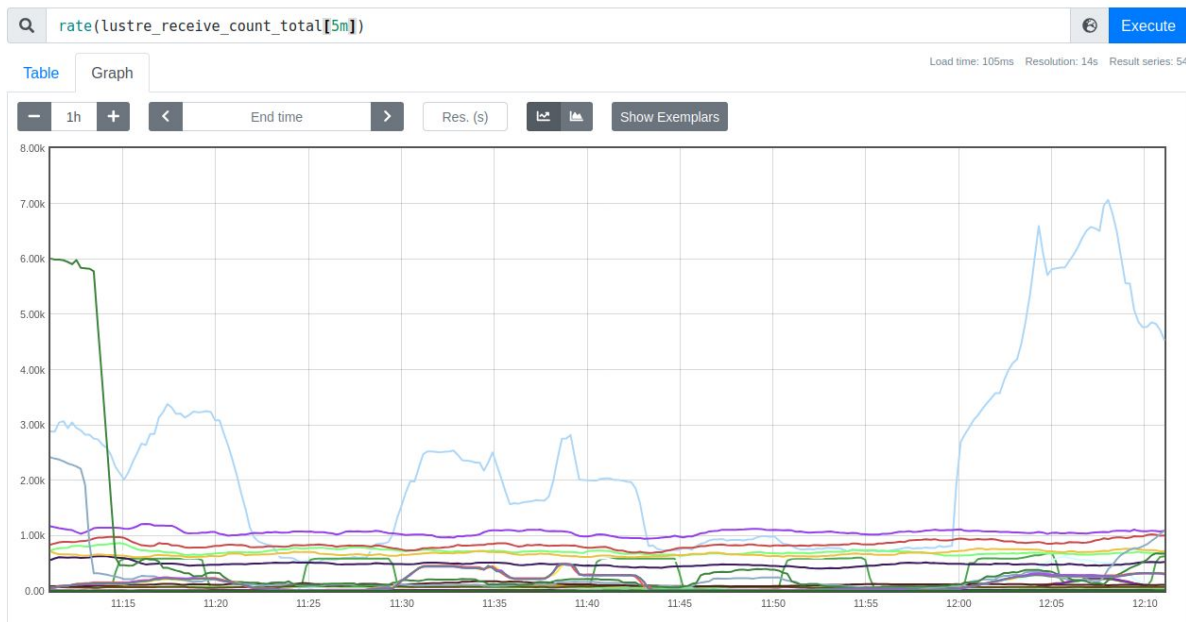
- Sistema de monitorización y alerta de código abierto.
- Diseñado para recopilar métricas de servicios y aplicaciones en tiempo real.
- Se está realizando una migración para adaptar las dashboards de grafana a Prometheus

¿Qué ofrece?

- Recolección de métricas por “scraping” mediante endpoints HTTP.
- Lenguaje de consultas PromQL para explorar y filtrar métricas.
- Alerta temprana a través de Alertmanager (email, Slack, etc.).
- Integración nativa con Grafana para la visualización.
- Gran cantidad de exporters disponibles para los distintos servicios mencionados.

<https://prometheus.io/>

Herramientas de Monitorización - Prometheus



<https://prometheus.io/>

Servicios esenciales en el Centro de Computación

- ★ important to be done soon
- ⚠ alert
- 🚧 work in progress
- 👉 low priority
- 👍 all right
- 💤 sometimes in the future

- HTCondor* → gestión de trabajos y cargas de cómputo.
- 🚧★ Lustre → almacenamiento distribuido de alto rendimiento (servicio transversal).
- Slurm → planificador de recursos y colas de ejecución.
- MySQL Server/Router → base de datos y balanceo de conexiones.
- PostgreSQL → servicio de base de datos relacional.
- Nginx → servidor web y proxy inverso.
- Apache → servidor web y de aplicaciones.
- Squid → proxy caché para control de tráfico web.
- Varnish → acelerador de contenido web mediante caché.
- 🚧 Nodo → métricas de CPU, memoria y recursos de cada máquina.
- 🚧★ IPMI → **consumo energético**, métricas de eficiencia.
- Red/Tráfico de red → antes monitorizado con Cacti, aspecto transversal.

Ejemplo de gráfica obtenida usando Grafana

Trabajo realizado hasta ahora:

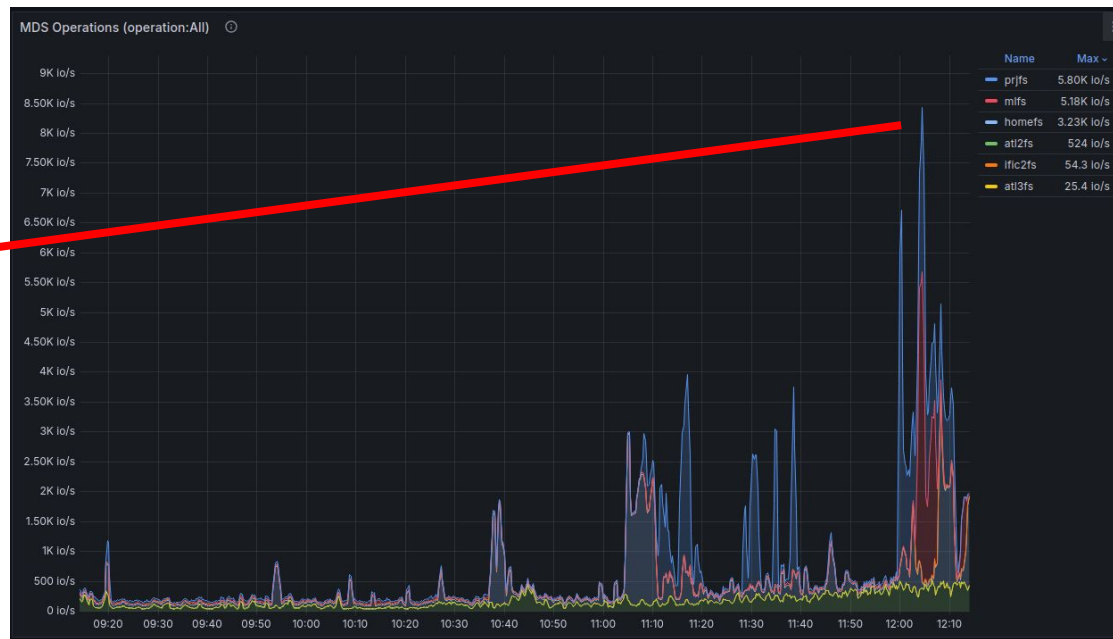
- Con **Prometheus** se monitoriza el estado de cada nodo del clúster y las métricas de **Lustre** (90%).

Ejemplo (gráfica derecha):

- Se observa un **pico de peticiones en el sistema de ficheros mlfs**.
- Puede deberse a un **error...** o a un **uso puntual legítimo**.

Valor de la monitorización histórica:

- La recolección y visualización de datos históricos permite **identificar anomalías**.
- Con métricas instantáneas (o sin métricas) esto sería **imposible de detectar**.



Desarrollo de pequeñas aplicaciones WEB con Python/Flask/SQLAlchemy para apoyo a administración e infraestructuras de cálculo científico.



MINISTERIO
DE CIENCIA, INNOVACIÓN
Y UNIVERSIDADES



CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS



VNIVERSITAT
DE VALÈNCIA

IFIC INSTITUT DE
FÍSICA
CORPUSCULAR

EXCELENCIA
SEVERO
OCHOA¹⁷



Migración de las aplicaciones web en uso

¿Por qué migrar?

- Necesidad de **upgrade de servidores/máquinas**.
- El servidor “**hal**” concentra varias aplicaciones críticas y es el principal servidor al que hacer un upgrade.
- Muchas de ellas fueron desarrolladas en **PHP3 (años 1999–2003)** → tecnología **obsoleta, insegura y difícil de mantener**.
- Riesgos actuales:
 - Fallos de compatibilidad en nuevas máquinas.
 - Vulnerabilidades de seguridad.
 - Dificultad para incorporar mejoras o integraciones.

Estado actual de las aplicaciones en producción

- ★ important to be done soon
- ⚠ alert
- 🚧 work in progress
- 🐌 low priority
- 👍 all right
- 💤 sometimes in the future

Prioridad	Nombre	Servidor	Año	Tecnología	Descripción
★	personal	hal	1999	PHP3	Gestión de personal
★	biblio	hal	2000	PHP3	Gestión de biblioteca
★	wlaccess	hal	2003	PHP3	Gestión IFIC_0
🐌	ivote	ivote	2011	TurboGears2	Voto electrónico
🐌	idoors	ifictv	2018	Flask	Control de accesos y fichaje
💤	art	artemisa	2019	Flask	Gestión de proyectos
💤	clockin	bit	2022	Flask	Fichaje remoto



Estrategia de Migración

Herramientas propuestas:

- **Base:** Python + Flask (backend ligero y flexible) + SQLAlchemy.
- **Frontend:** HTML, JS y CSS.
- **Frameworks opcionales:** Bootstrap, React, Vue, Foundation, SPA.

Importancia de la elección tecnológica:

- **Mantenibilidad** → facilitar futuras actualizaciones.
- **Escalabilidad** → crecer con nuevas funcionalidades.
- **Experiencia de usuario** → interfaces modernas y usables.

Estado actual del proyecto:

- 🚧 **No iniciado aún.**
- Hasta el momento el esfuerzo se ha centrado en la **monitorización de sistemas y servicios.**
- Próximo paso: definir arquitectura y priorizar migración de apps críticas (personal, biblio, waccess).

Fin.



MINISTERIO
DE CIENCIA, INNOVACIÓN
Y UNIVERSIDADES



CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS



VNIVERSITAT
DE VALÈNCIA

IFIC INSTITUT DE
FÍSICA
CORPUSCULAR



EXCELENCIA
SEVERO
OCHOA²¹