IMPLEMENTACIÓN Y EVALUACIÓN DE UNA HERRAMIENTA DISTRIBUIDA DE ANONIMIZACIÓN DE CAPTURAS DE RED

Emilio Rigazio **David Morocho**

El problema

- En la actualidad las empresas cuentan con redes de computadoras sobre las que circulan la hoy en día invaluable información.
- El monitoreo de esas redes es una importante herramienta de análisis.
- Sin embargo, existe un riesgo de seguridad y privacidad pues se expone información sensible de los usuarios.

Source	Destination	Protocol	Info
192.168.100.100	74.125.159.113	HTTP	GET /generate_204 HTTP/1.1
74.125.159.113	192.168.100.100	HTTP	HTTP/1.1 200 OK (text/javascript)
192.168.100.100	74.125.159.106	HTTP	GET /search?hl=en&source=hp&q=sintomas+de+VIH&btnG=Google+Search
192.168.100.100	74.125.159.113	HTTP	GET /complete/search?hl=en&q=sintomas%20de%20VIH&cp=15 HTTP/1.1
74.125.159.113	192.168.100.100	HTTP	HTTP/1.1 200 OK (text/javascript)
192.168.100.100	74.125.159.113	HTTP	GET /complete/search?hl=en&q=sintomas%20de%20VI&cp=14 HTTP/1.1
74.125.159.113	192.168.100.100	HTTP	HTTP/1.1 200 OK (text/javascript)
192.168.100.100	74.125.159.113	HTTP	GET /complete/search?hl=en&q=sintomas%20de%20V&cp=13 HTTP/1.1
74.125.159.113	192.168.100.100	HTTP	HTTP/1.1 200 OK (text/javascript)
192.168.100.100	74.125.159.113	HTTP	GET /complete/search?hl=en&q=sintomas%20de%20&cp=12 HTTP/1.1
74.125.159.113	192.168.100.100	HTTP	HTTP/1.1 200 OK (text/javascript)
192.168.100.100	74.125.159.113	HTTP	GET /complete/search?hl=en&q=sintomas%20de&cp=11 HTTP/1.1

El problema

- Las empresas restringen o prohíben el monitoreo de sus redes para minimizar los riesgos de seguridad.
- La inter-cooperación entre entidades es difícil de conseguir pues se deben formar relaciones de confianza para el intercambio de capturas de red.
- Eliminar la información sensible antes del análisis no es una solución ideal, pues lo limita al eliminar información fundamental.

Lo requerido

- Se necesita un mecanismo de “enmascaramiento” que no elimine datos sensibles, sino que los transforme.
- Además, el mecanismo debe ser capaz de procesar grandes cantidades de datos sin que esto implique tiempos de procesamiento mucho mayores.
- Se requiere entonces un mecanismo **escalable** de **anonimización**.

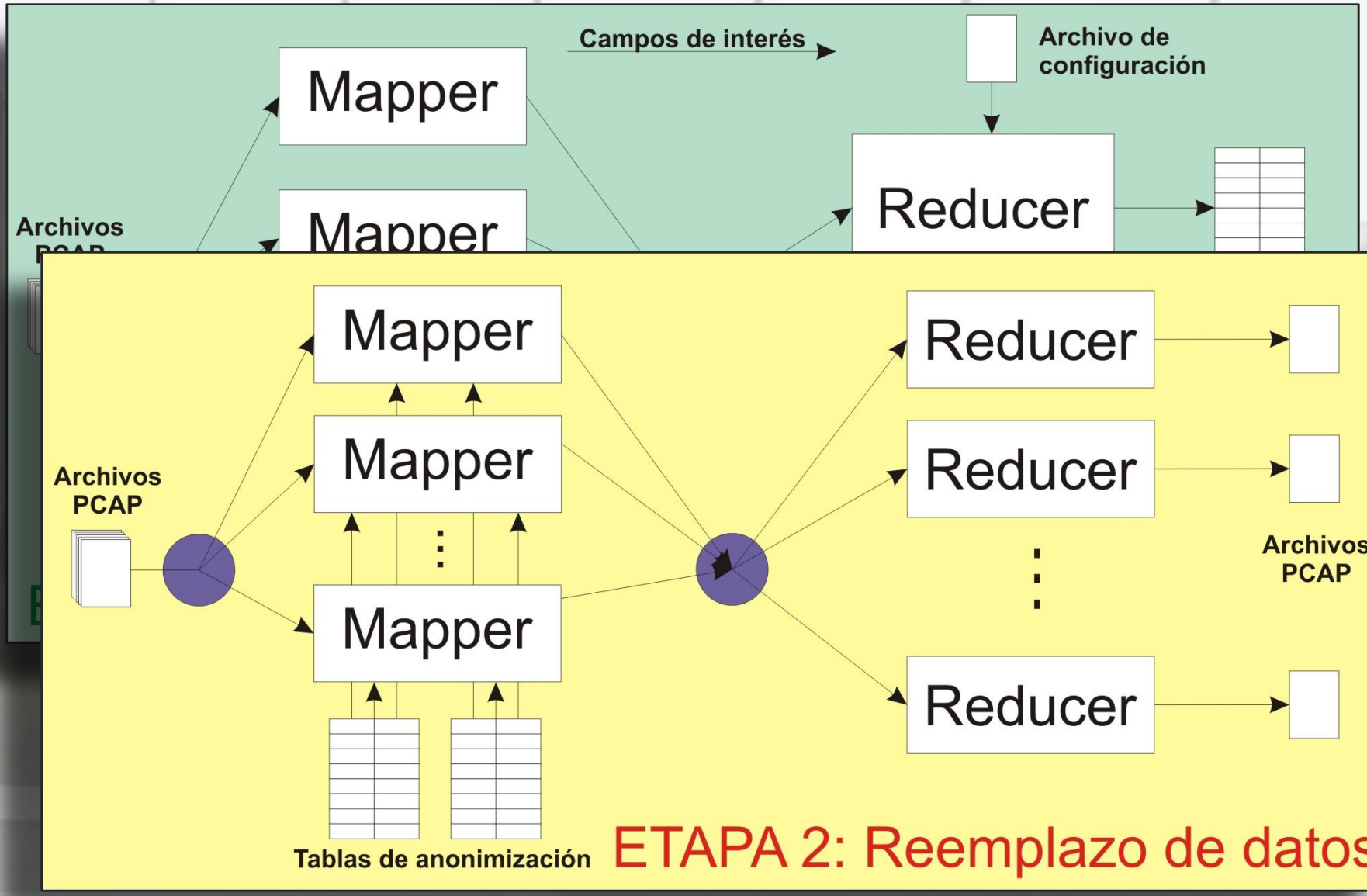
La solución propuesta

- Mecanismo **distribuido** de anonimización.
- El procesamiento de datos se realiza paralelizado en nodos que trabajan independientemente.
- La carga es distribuida en un clúster escalable, de modo que mayores cargas puedan ser procesadas en similares tiempos.

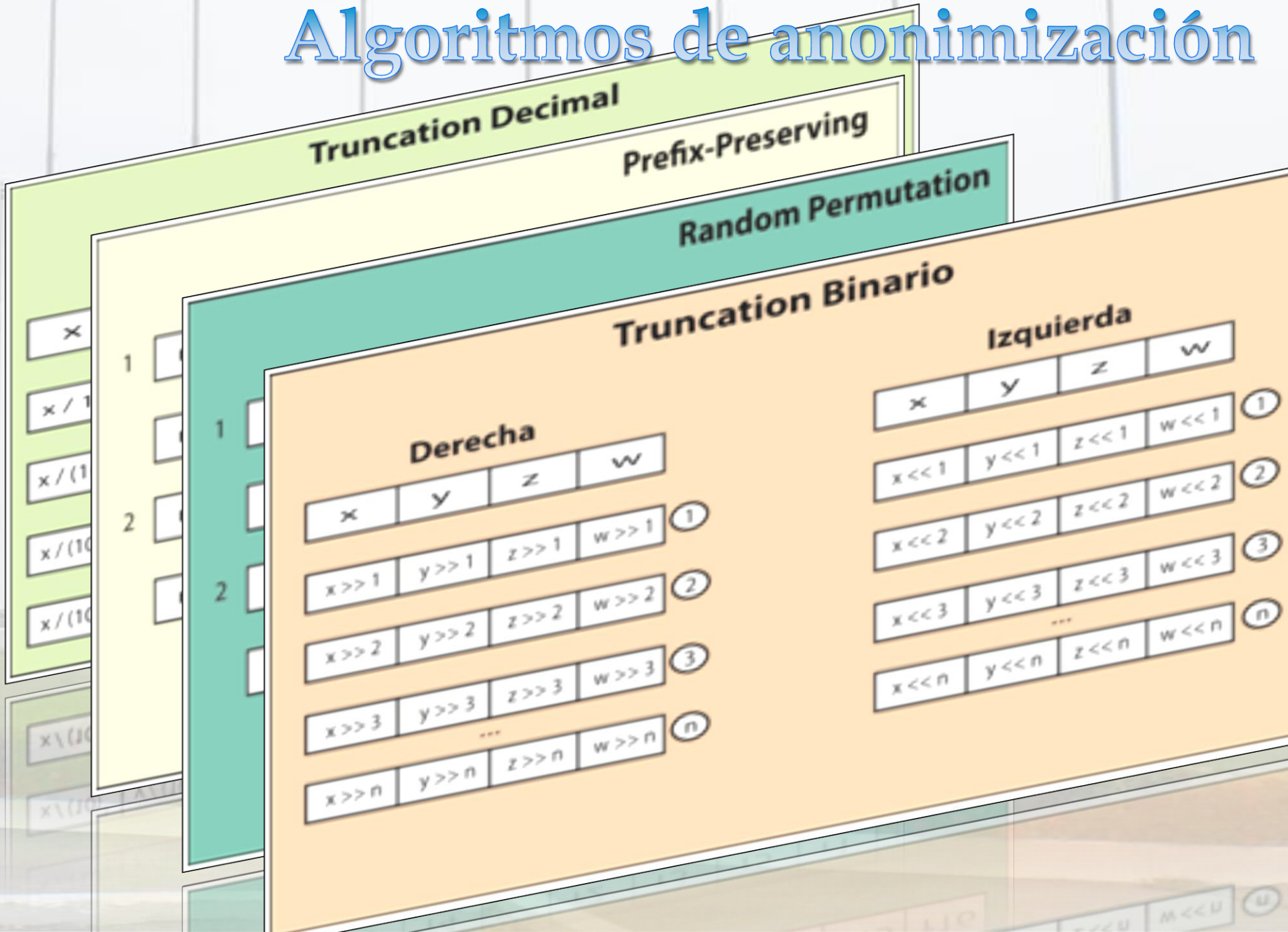
Diseño de la solución

- Uso del paradigma de programación **Map Reduce**.
- División del problema en 2 etapas generales:
  - Extracción y anonimización de datos
  - Reemplazo de datos
- Cada etapa comprende una tarea Map Reduce.
- La anonimización de datos se ejecuta utilizando algoritmos especializados con distintos niveles de seguridad.

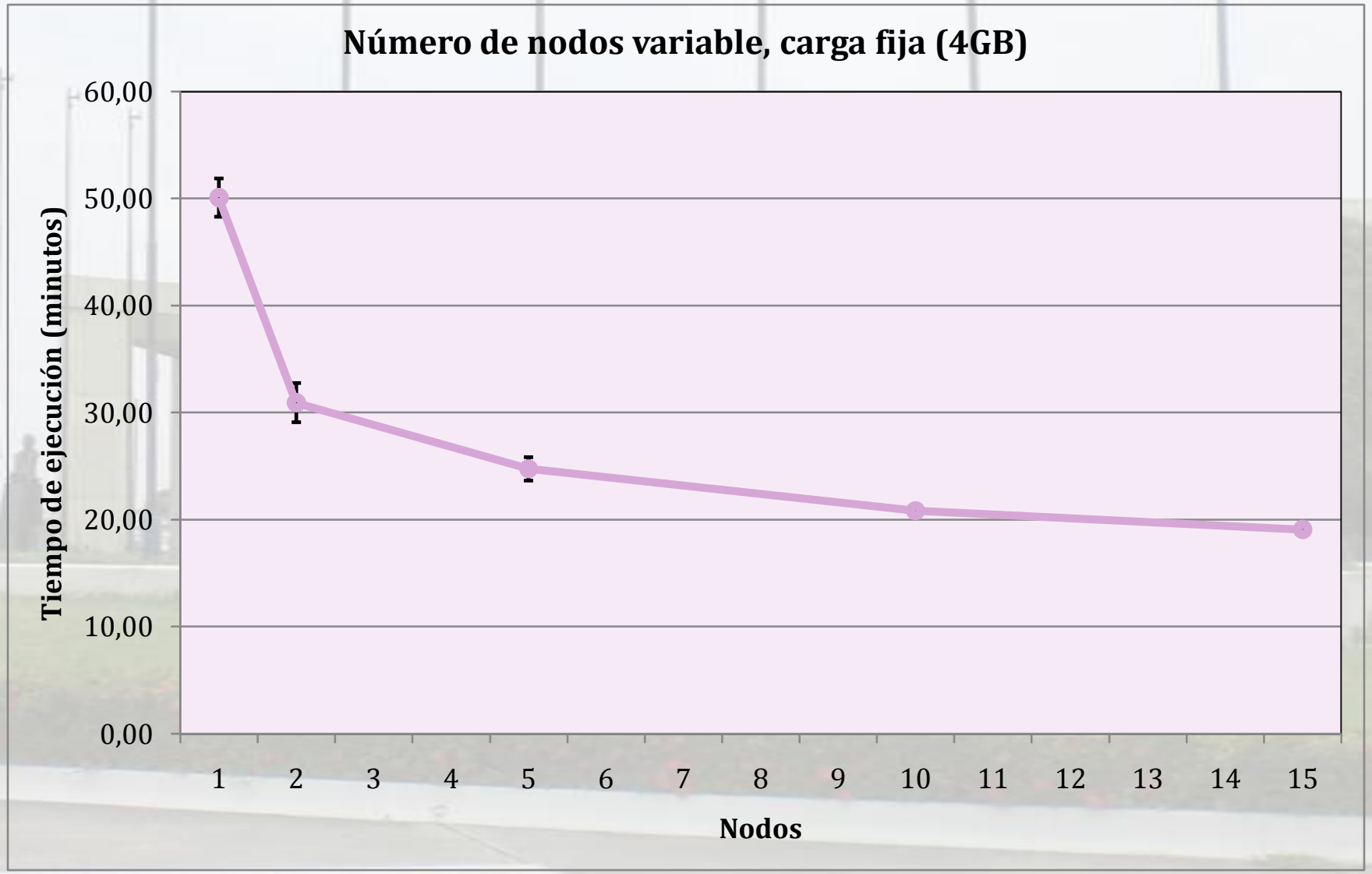
Diseño de la solución



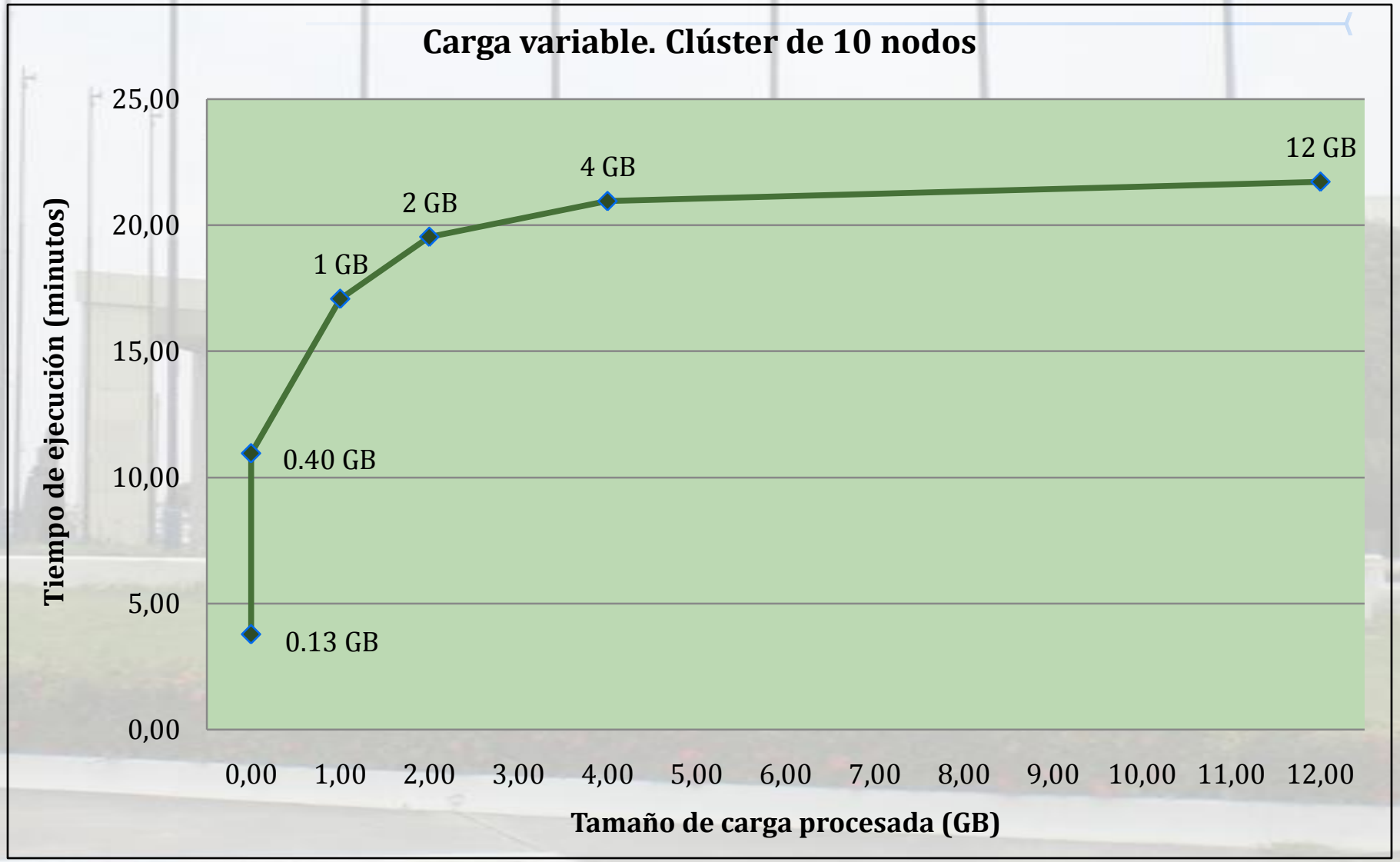
Algoritmos de anonimización



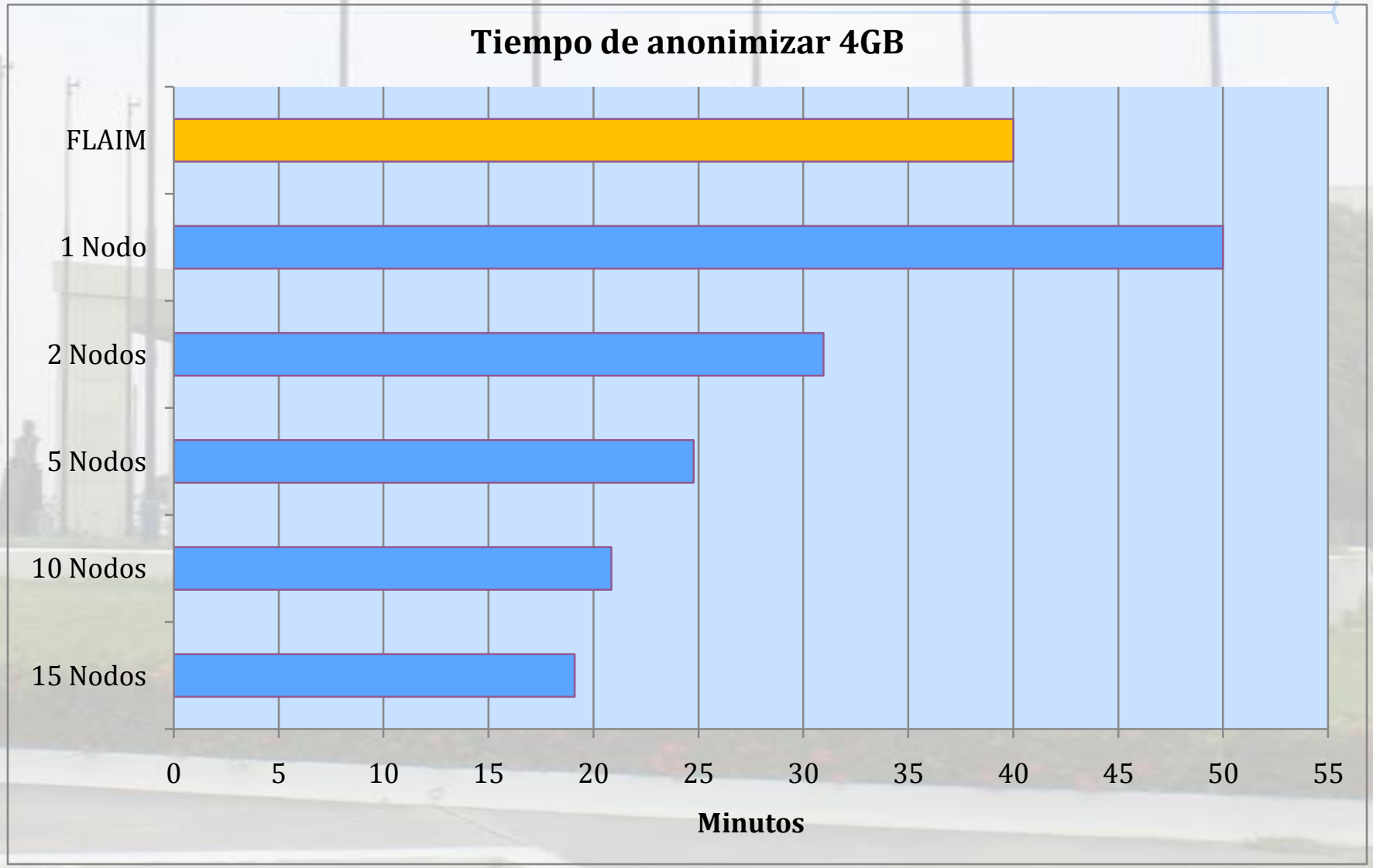
Pruebas y resultados



Pruebas y resultados



Pruebas y resultados



Conclusiones

- La anonimización de archivos es un proceso computacionalmente costoso, especialmente ejecutado secuencialmente.
- Los modelos distribuidos generan soluciones altamente escalables.
- Se generó una herramienta que aprovecha las ventajas del modelo distribuido, donde se observa un comportamiento escalable superior al de una herramienta existente.
- Existen servicios comerciales que brindan la posibilidad de ejecutar soluciones distribuidas como la planteada por costos reducidos, esto hace muy viable la implementación de las mismas.

Recomendaciones

- Incrementar la funcionalidad de la herramienta mediante la anonimización de otros campos.
- Implementar un mecanismo de escritura de archivos en formatos estandarizados.
- Desarrollar implementaciones de nuevos algoritmos de anonimización.
- Ampliar el alcance de la solución a la anonimización de archivos de log de otros tipos de aplicaciones

Preguntas

Implementación y evaluación de una herramienta distribuida de anonimización de capturas de red

Emilio Rigazio David Morocho

IMPLEMENTACIÓN Y EVALUACIÓN DE UNA HERRAMIENTA DISTRIBUIDA DE ANONIMIZACIÓN DE CAPTURAS DE RED

Emilio Rigazio **David Morocho**