

# Informe Laboratorio 2 - Sistemas Distribuidos

Mauricio Cortés - 202004529-1

Gianfranco Dissi - 202004514-3

Claudio Espinosa - 202004539-9

## 1. Preguntas planteadas

1. ¿En este sistema, los datos se encuentran distribuidos equitativamente? Justifique.

La pregunta planteada es posible analizarla tanto desde el punto de vista de los Continentes, para lo nombres que deben enviar a la OMS, como para los DataNodes, en cuanto a los nombres que reciben desde esta misma.

Viendolo por el lado de los Continentes, se desarrolló el laboratorio de tal manera que manteniendo una misma copia de un archivo names.txt para cada Continente, estos trabajan distintas líneas del mismo. Para ello cada continente contiene un parámetro en su archivo main.go denominado "lineToRead" el cual indicará que línea debe trabajar, realizando saltos de a 4. Por tanto, a nivel de Continentes, los datos se encuentran distribuidos equitativamente.

Por otro lado, el comportamiento a nivel de DataNodes es distinto. Por un lado, se depende mucho de qué Continente se esté ejecutando, dado que lee siempre las mismas líneas. Esto puede conllevar a que un Continente lea solamente apellidos de inicial entre A y M, los cuales irán todos al DataNode1, mientras el DataNode2 recibirá pocos o nulos datos, tal como sucedió en la prueba del laboratorio, donde se debió levantar un segundo continente para enviar más datos al DataNode2. Otro factor que puede desbalancear la distribución de los datos en los DataNodes, es los mismos nombre que se encuentre en el archivo names.txt, dado que si se encuentran en su mayoría nombres con la inicial de su apellido entre A y M, se enviarán constantemente al DataNode1, dejando en su mayoría ocioso al DataNode2.

2. ¿Cuanto se demora su programa en realizar cada consulta desde la ONU?

Para evaluar este punto, se realizaron tres pruebas con aproximadamente 50, 100 y 200 datos, solicitando para ambas cantidades Infectados y Muertos, obteniendo así los siguientes resultados:

	Infectados	Muertos
50 datos	10.40[s]	10.57[s]
100 datos	7.38[s]	10.44[s]
200 datos	9.27[s]	9.82[s]

3. Si se agrega otro nodo que se encargue solo de recibir desde los continentes y subir la información a los DataNodes. ¿Aumenta o disminuye en el tiempo para realizar cada consulta? Justifique.

Actualmente la OMS es la entidad encargada de realizar las dos tareas relevantes en el sistema, recibir las personas desde los continentes y enviárselos a los DataNodes y atender las consultas de la ONU. Si se añadiese otro nodo que se encargase de la tarea de recibir y subir la información de los continentes, implicaría realizar coordinación extra con este nodo, con tal de obtener correctamente los nombres desde los datanodes, sin mencionar que además se requeriría coordinación extra por parte de los DataNodes respecto a las consultas que le llegan, las cuales ahora serían desde entidades diferentes. Por tanto, no podemos asegurar con certeza si el tiempo de la consulta disminuiría o aumentaría, dado que el incluir un nodo extra requiere realizar una mayor coordinación, la cual ya se realiza en el planteamiento inicial con la OMS.

4. A partir del sistema realizado, ¿Cual seria una forma mas eficiente de balancear la carga del sistema?

Tal como se mencionó anteriormente, el desbalance de la carga se produce respecto a la cantidad de datos que reciben los DataNodes. Una medida de solución que planteamos para aligerar la carga es ir intercalando el DataNode, al cual se le envía la información. De esta manera, no nos preocupamos de la letra del apellido recibido, criterio que se utiliza en el planteamiento inicial, sino que ahora simplemente mandamos al DataNode que sea su turno de recibir la información, equiparando así la carga.