

BIG DATA

Lucas Azevedo Dias

- 1) Significa dizer que as máquinas podem ser diferentes em SO e hardware, o que é permitido no *Hadoop* pela sua arquitetura mestre-escravo.
- 2) É uma arquitetura onde há o mestre que apenas gerencia e coordena as atividades dos escravos que, por sua vez, são os responsáveis pelo trabalho duro (armazenamento e processamento). Nela, o mestre é chamado *NameNode*, enquanto os escravos de *DataNode*.
- 3) O *JobTracker* e o *TaskTracker* acabam sendo responsáveis pela coordenação dos processamentos sobre os dados armazenados nos *DataNodes*.
- 4) Ela permite que os processamentos sejam feitos de forma descentralizada e paralela, o que reduz o peso sobre o nodo central e, além de aumentar a velocidade de processamento, leva a uma economia pela falta de necessidade de um nodo central ser potente.
- 5) Ele distribui várias réplicas dos blocos do arquivo em vários *DataNodes* distintos, o que gera a persistência do arquivo mesmo na perda de um nodo.
- 6) Resposta:
 - Serão armazenados 15 blocos (3 réplicas de 5 blocos);
 - Serão 4 blocos de 128MB e 1 bloco de 2MB;
 - Ao total, o arquivo ocupará 1542MB (3 réplicas de 514MB).