

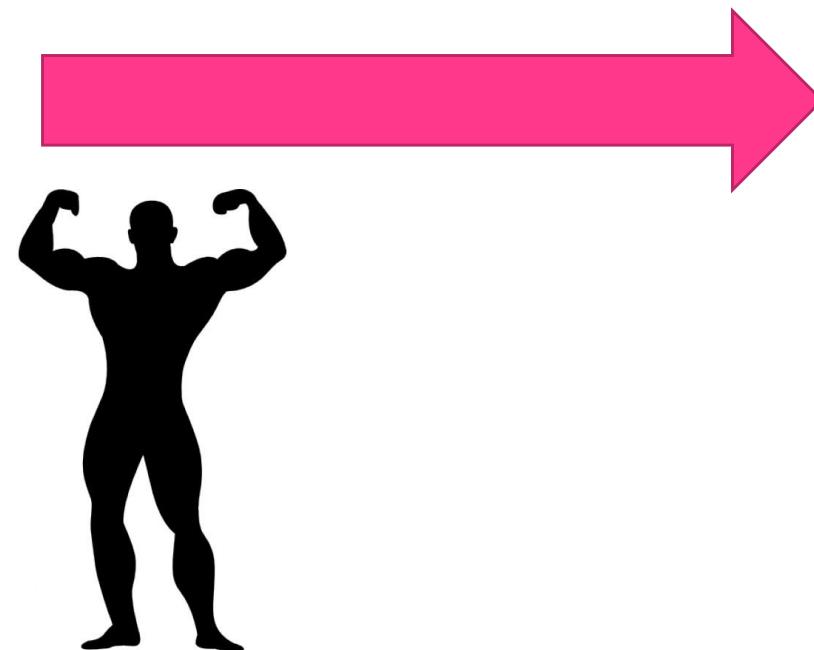
Tarefa de Carregar Pedras



Opções

- Contratar uma pessoa mais forte
- Contratar várias pessoas
- Dividir a pedra e contratar pessoas comuns para levar as partes

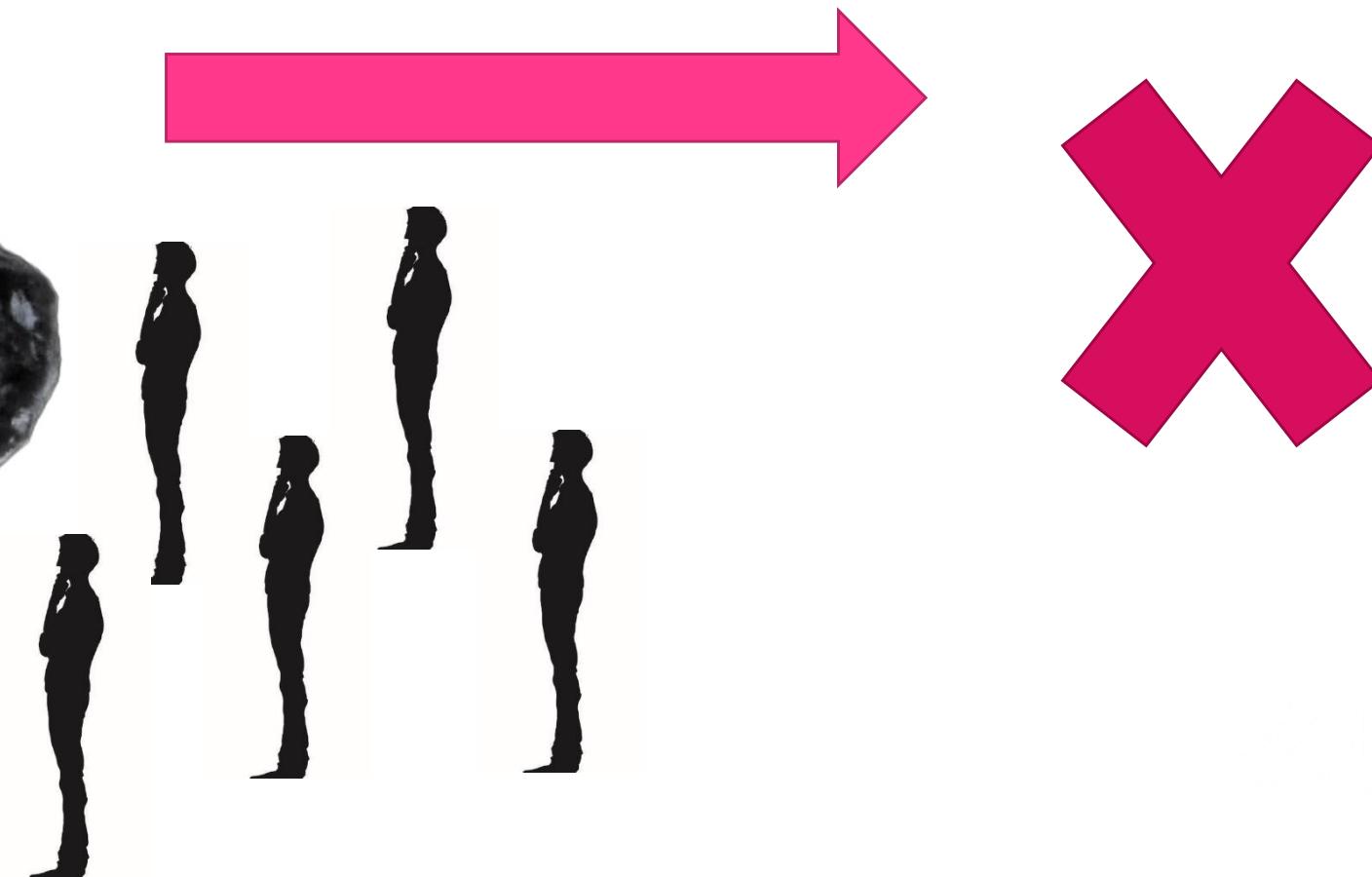
Contratar uma pessoa mais forte



Contratar uma pessoa mais forte

- E se as pedras aumentarem mais ainda o tamanho?
- E se forem pedras pequenas?

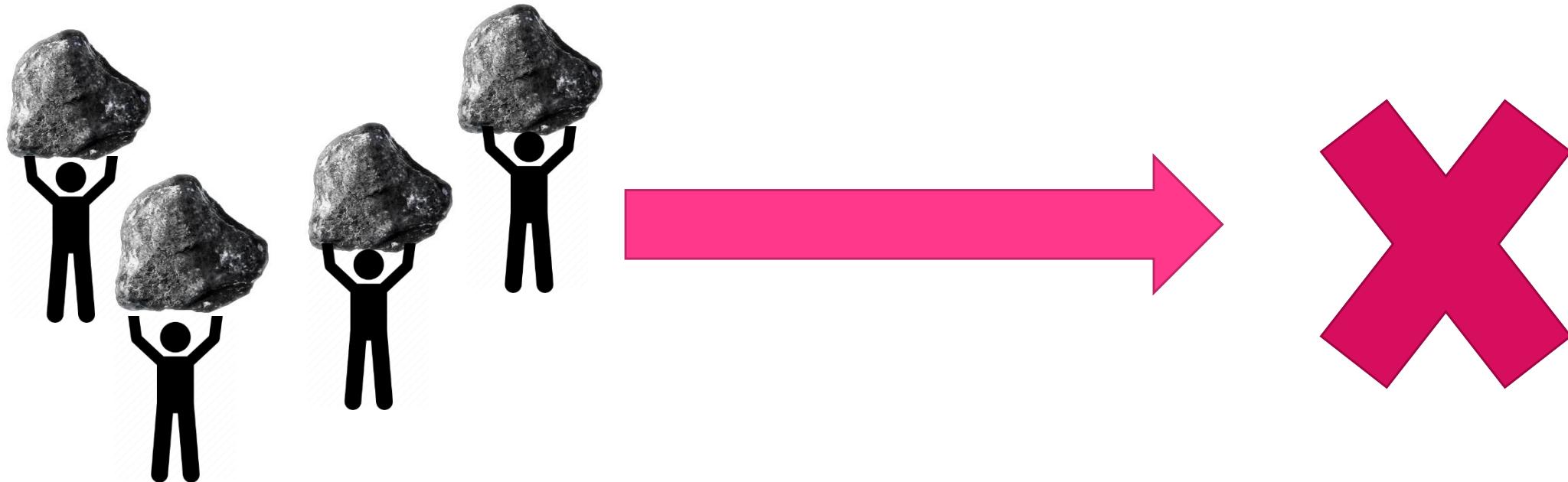
Contratar várias pessoas



Contratar várias pessoas

- E se as pedras aumentarem mais ainda o tamanho?
- E se forem pedras pequenas?
- Problemas de gerenciamento

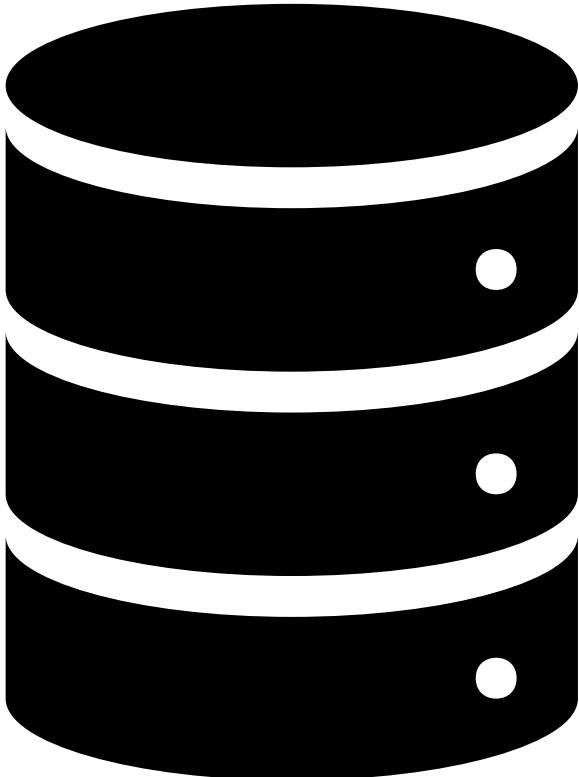
Dividir a pedra e contratar pessoas comuns para levar as partes



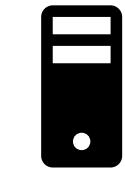
Dividir a pedra e contratar pessoas comuns para levar as partes

- Melhor gerenciamento
- Você contrata as pessoas de acordo com a necessidade
- Fácil achar pessoas comuns

Big Data



Big Data



Map Reduce

- Dividir tarefas de processamento de dados em vários nós
 - Dados são divididos em blocos
 - Divisão de problemas grandes e/ou complexos em pequenas tarefas
- Fundamento para Map Reduce e Hadoop:
 - MapReduce: Simplified Data Processing on Large Clusters
 - 2004: Google: <https://ai.google/research/pubs/pub62>

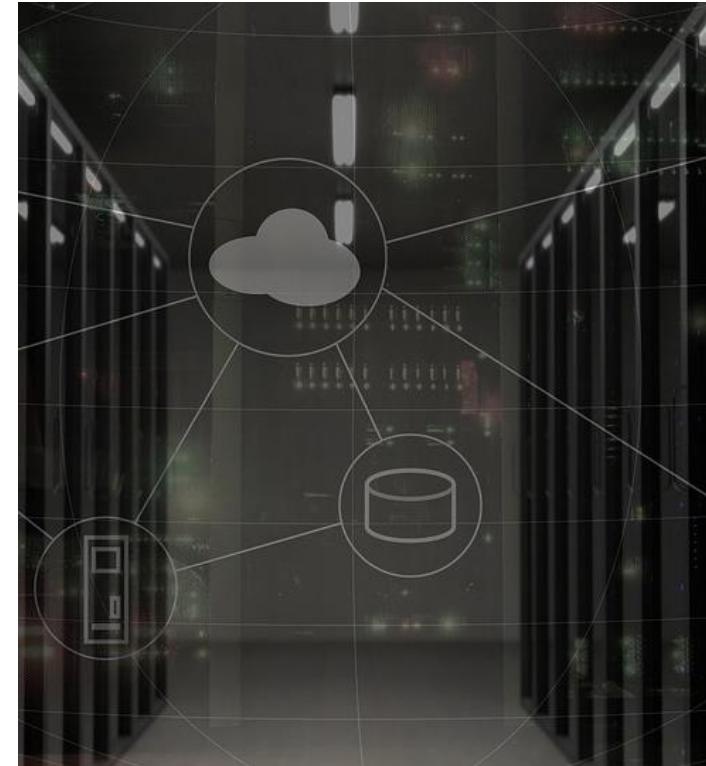
Map e Reduce

- Escalável
- Tolerante a falhas
- Disponibilidade
- Confiável
- Usa conceito de chave/valor
- Não cria gargalos na rede, pois dados não trafegam (processamento no nó)



Importante

- Processamento em Batch
- Grande volumes de dados
- Processamento distribuído
- Linguagem Imperativa (JAVA)



MapReduce

- Mapeamento é executado em paralelo nos nós
- Apenas quando Mapeamento é encerrado, redução inicia, também em paralelo
- Fase intermediária: Shuffle
- Existem tarefas que requerem apenas a etapa de Mapeamento

