

Análise e Visualização de Dados de Localização de Dispositivos Móveis Conectados à Internet

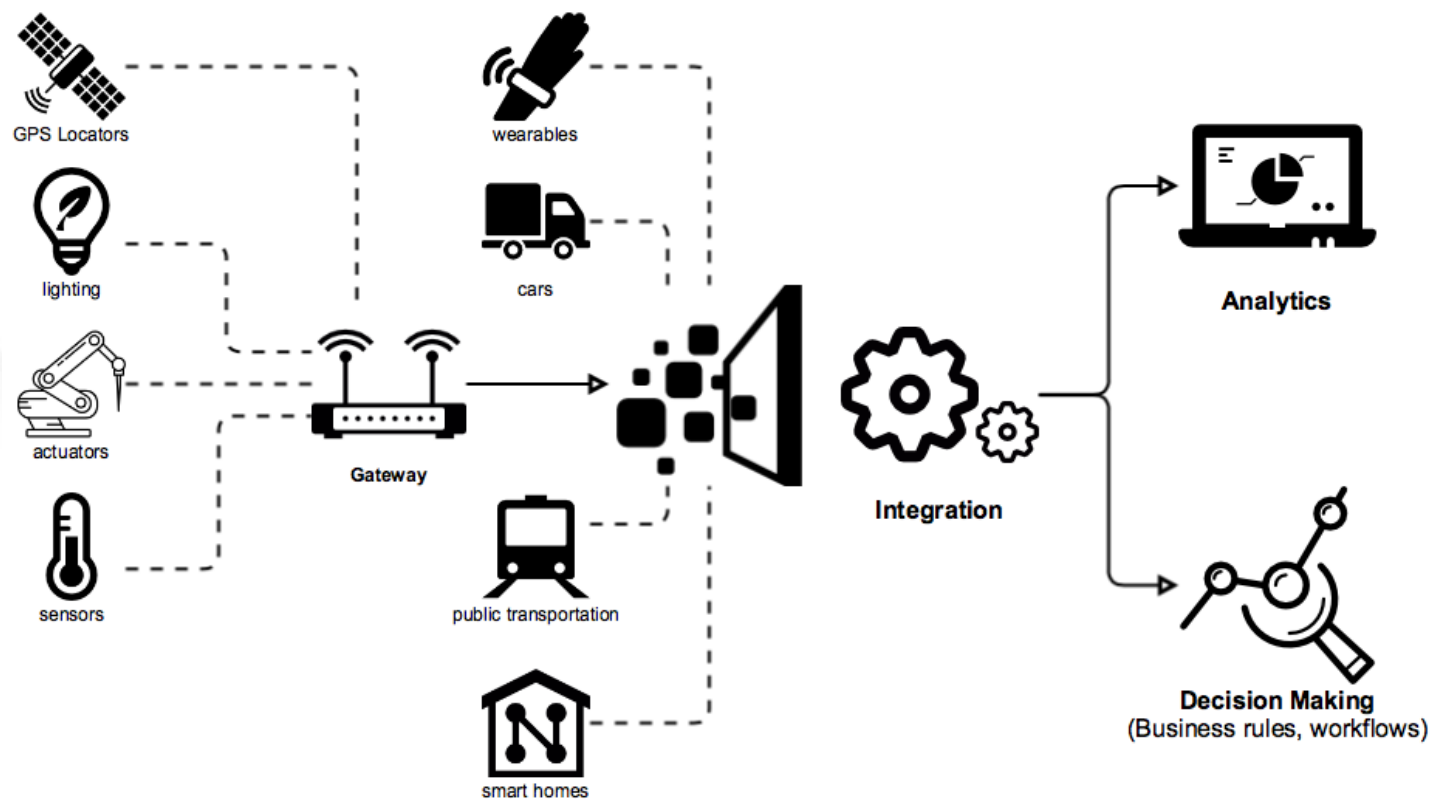
Marcelo Augusto Cordeiro

marcelo.augusto.cordeiro@gmail.com

Orientador: Prof. Dr. Eduardo Martins Morgado



Internet das Coisas



Fonte: DZone - Architecting the Internet of Things
Disponível em: <https://dzone.com/articles/architecting-the-internet-of-things>



Big Data

Tecnologias de big data são utilizadas quando a quantidade de dados é grande, complexa e variada e difícil de ser adequadamente armazenada, administrada e processada através de bases de dados e programas tradicionais.



5Vs da Big Data

- Volume
- Variedade
- Velocidade
- Valor
- Validade



Hadoop

- Criado pela Apache em 2006
- Se tornou open-source em 2009
- Definiu o modelo de programação MapReduce

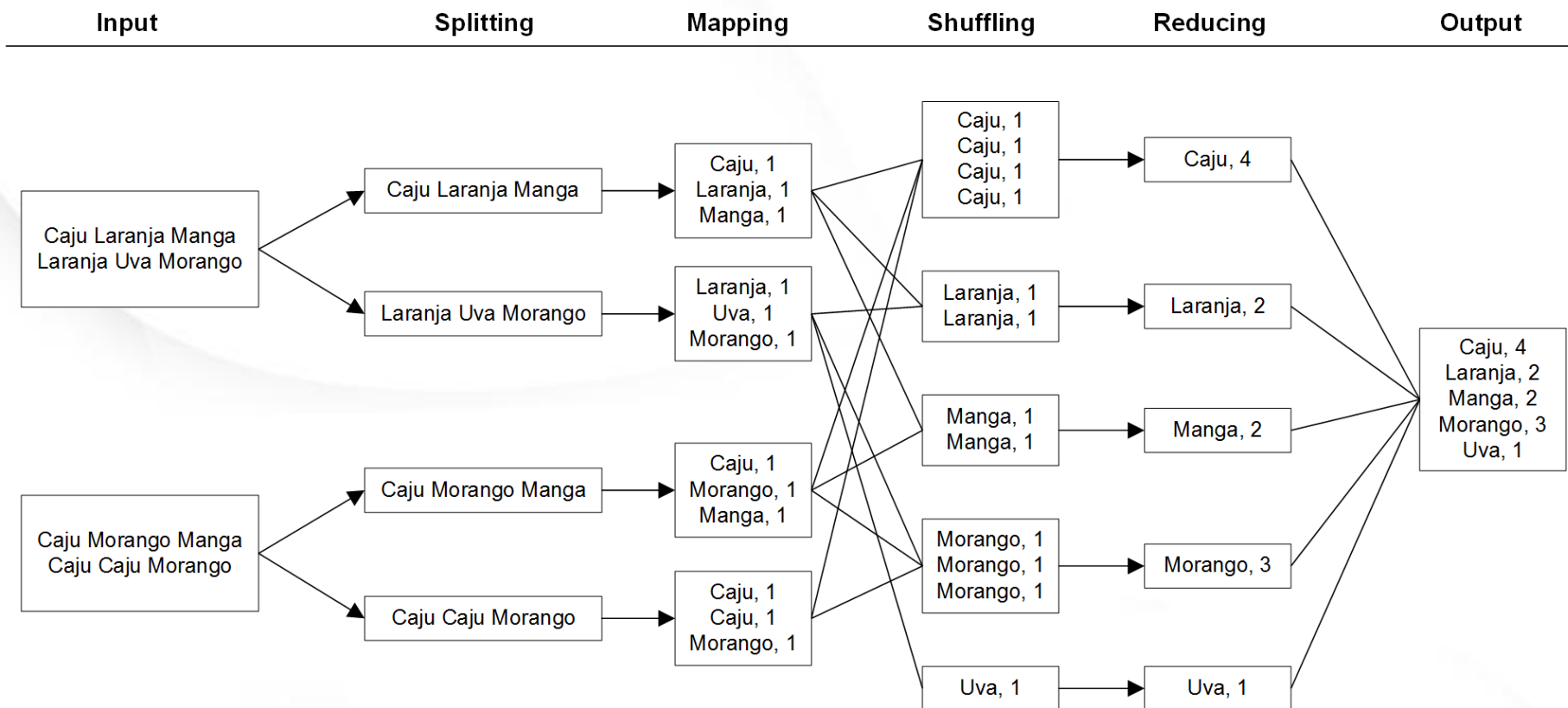


MapReduce

- Modelo de programação criado para ser utilizado no Hadoop
- Duas funções principais: *mapping* e *reducing*
- Trabalha com pares do tipo *<chave, valor>*

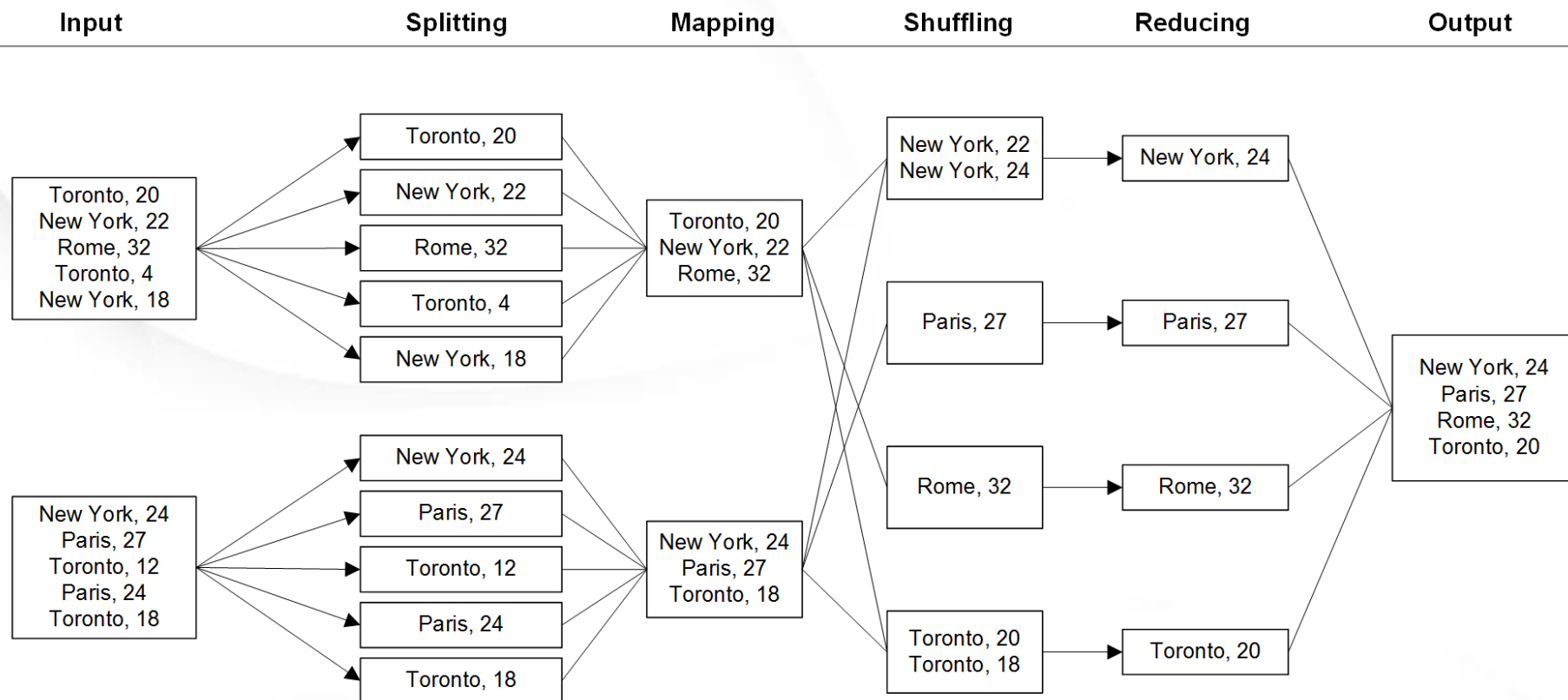


MapReduce - Exemplo



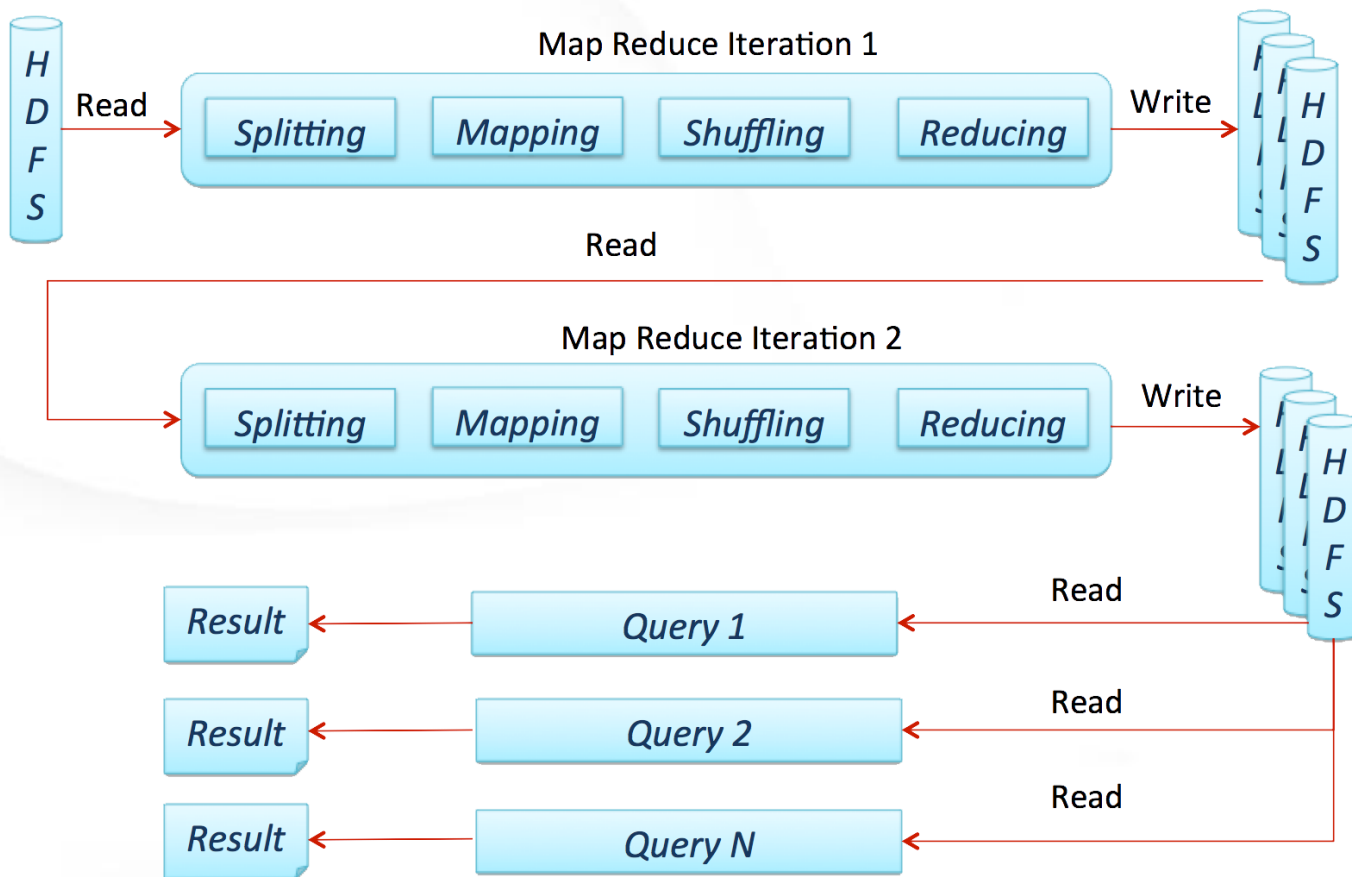


MapReduce - Exemplo





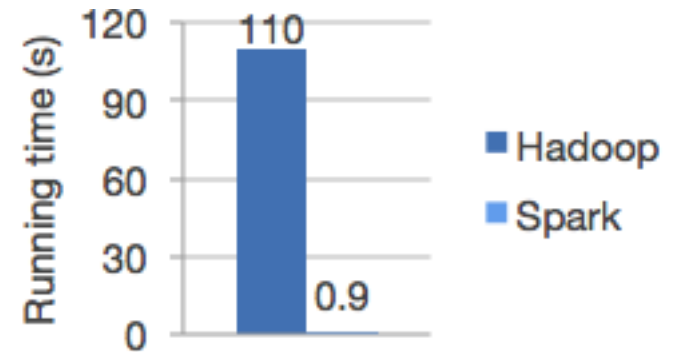
MapReduce





Spark

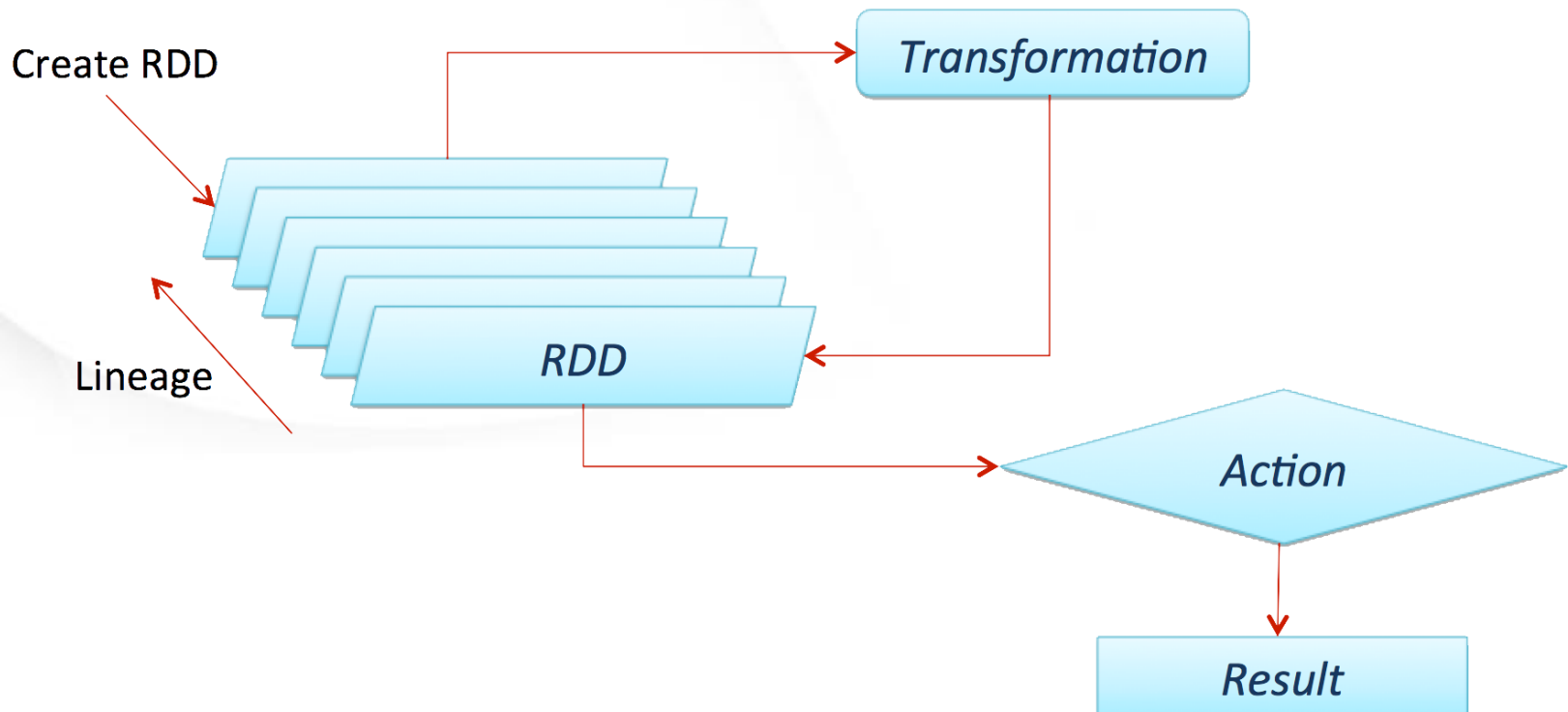
- Processamento in-memory de big data em clusters
- Até 100x mais rápido que Hadoop
- Utiliza avaliação preguiçosa



Fonte: Apache Spark
Disponível em: <http://spark.apache.org/>

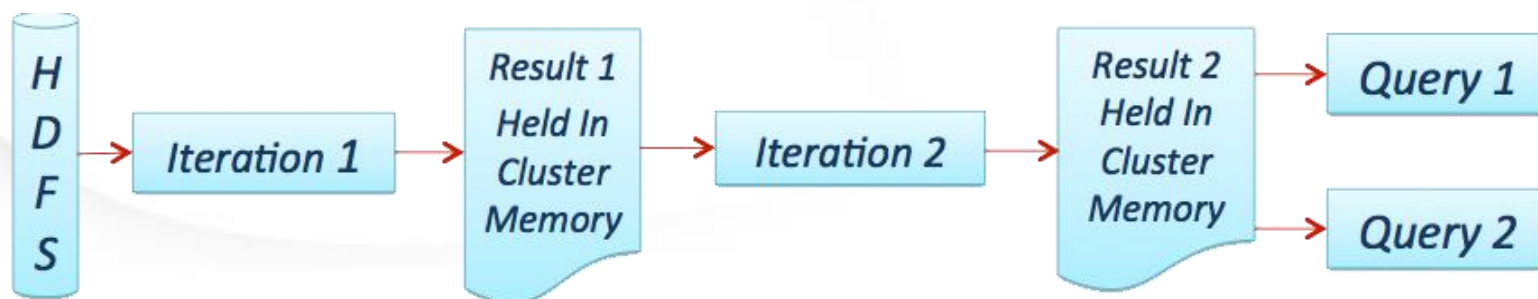


Resilient Distributed Dataset



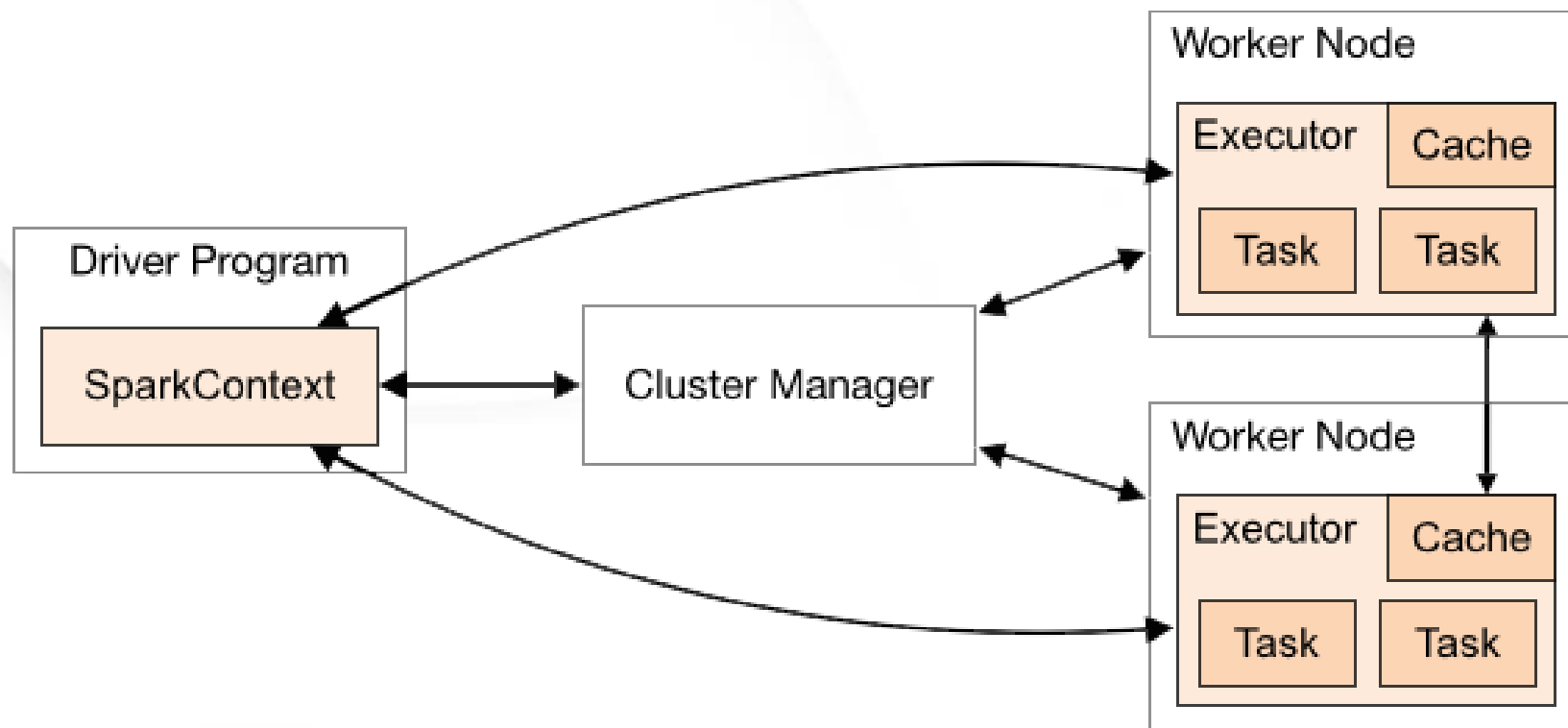


Spark





Spark

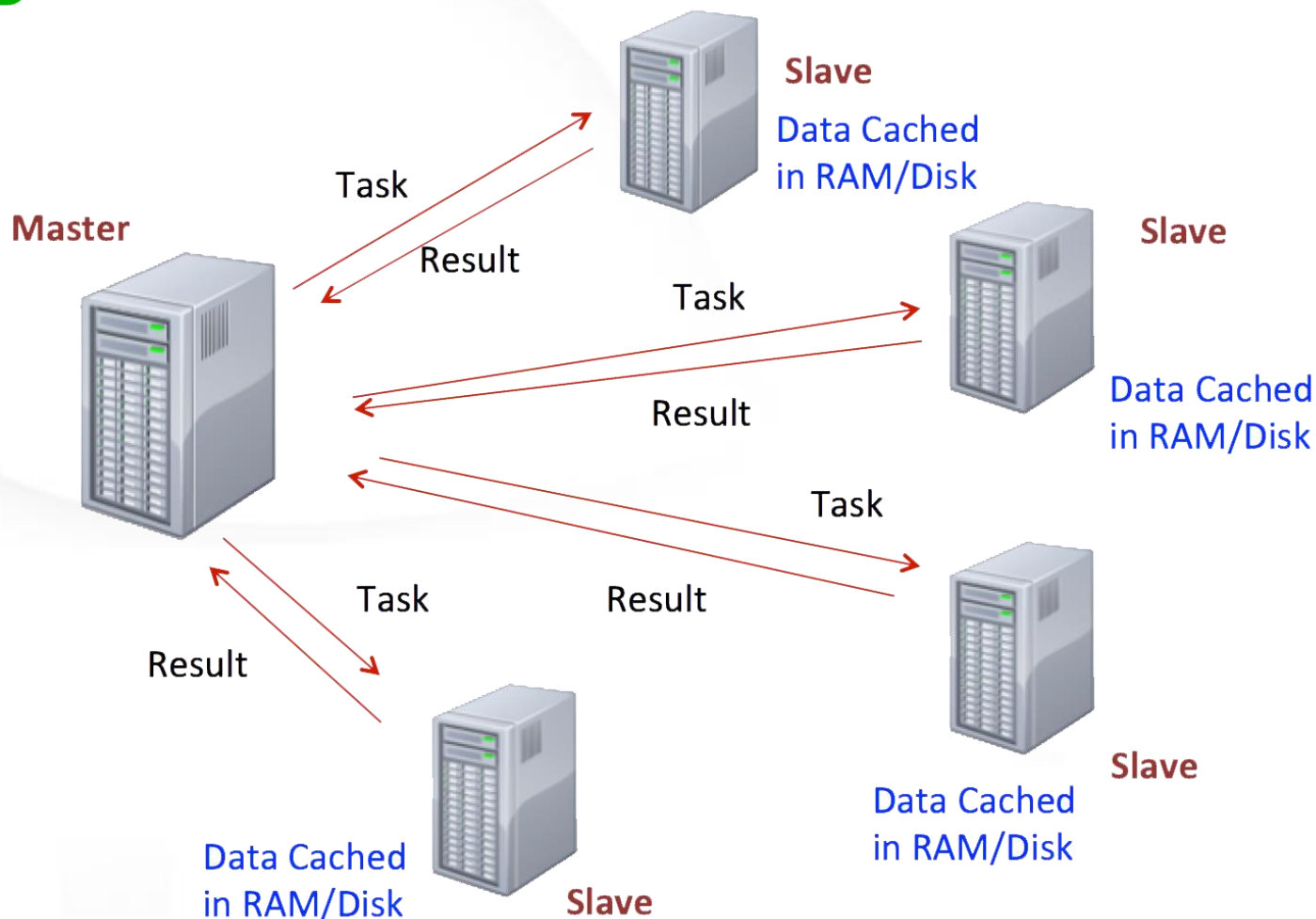


Fonte: Apache Spark - Cluster Mode Overview

Disponível em: <http://spark.apache.org/docs/latest/cluster-overview.html>



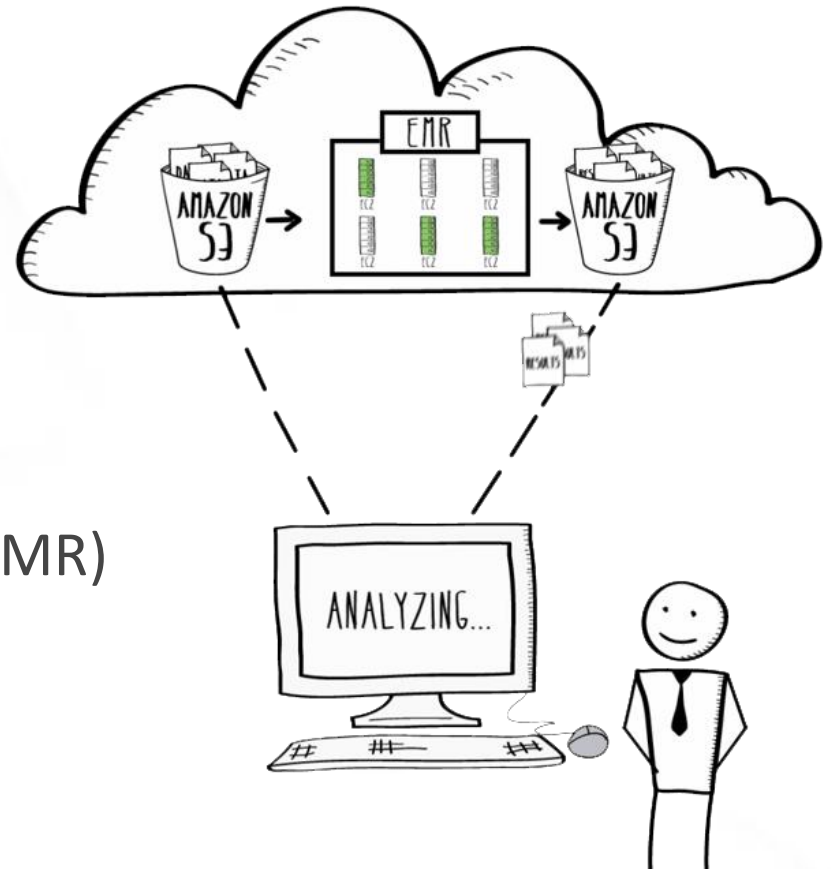
Spark





Amazon Web Services

- Suíte de serviços para computação em nuvem
- Amazon Elastic MapReduce (EMR)



Obrigado!

Marcelo Augusto Cordeiro

marcelo.augusto.cordeiro@gmail.com