

Um Servidor Escalável para Bases Massivas de Dados Geográficos

Leandro da Silva Santos

Orientador: Tiago Garcia de Senna Carneiro
Co-orientador: Ricardo Augusto Rabelo Oliveira

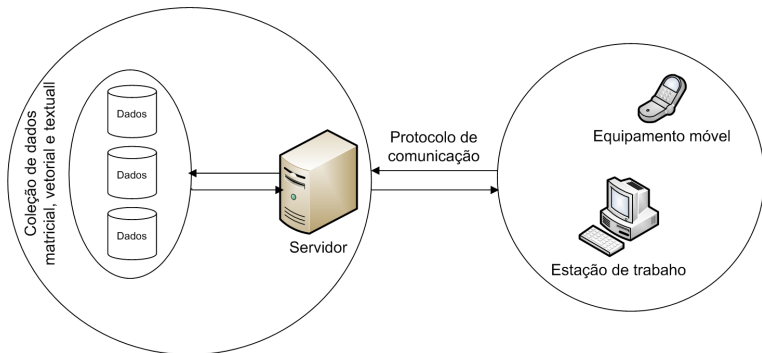
Departamento de Computação
UFOP

14 de julho de 2011

- 1 **Introdução**
 - O Problema
- 2 Trabalhos Relacionados
 - HDFS
 - MapReduce
 - HBase
- 3 Metodologia
 - Método
 - Análise de Complexidade
- 4 Conclusões

O Problema

- Características:
 - Dado na escala do terabytes.
 - Quantidade massiva de clientes.
 - Serviço em tempo real.



- 1 Introdução
 - O Problema
- 2 **Trabalhos Relacionados**
 - HDFS
 - MapReduce
 - HBase
- 3 Metodologia
 - Método
 - Análise de Complexidade
- 4 Conclusões

HDFS

- Características:
 - Alta disponibilidade.
 - Alta largura de banda.
 - Escalabilidade.
 - Gerenciamento prático.
- Entidades:
 - Namenode.
 - Datanode.

HDFS

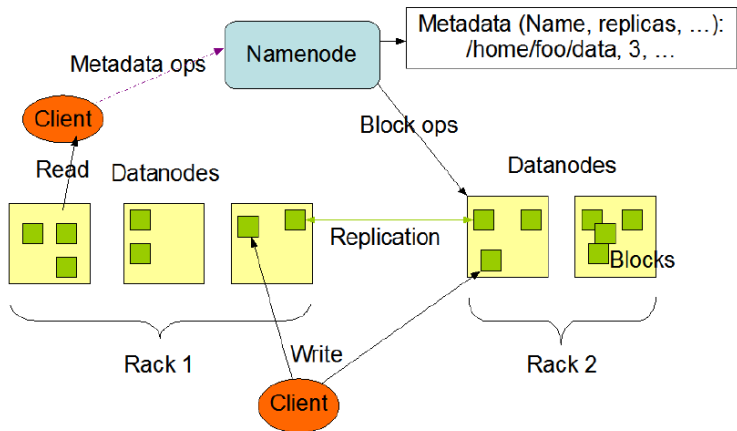


Figura: Arquitetura do sistema de arquivo distribuído HDFS [1].

MapReduce

- Características:
 - Simplicidade.
 - Ocultação do paralelismo.
- Entidades:
 - Master.
 - Worker Map.
 - Worker Reduce.

MapReduce

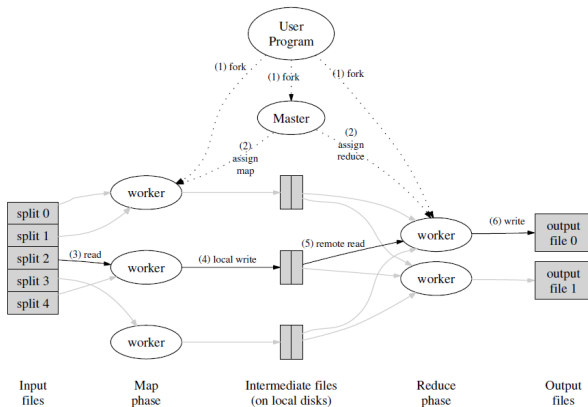


Figura: Modelo de programação MapReduce [2].

MapReduce

- Modelo de programação:
 - Implementação de duas funções.
 - $\text{map: } \text{map}(\text{chave}, \text{valor}) \rightarrow \text{list}(\text{chave}, \text{valor}; \text{intermediario})$.
 - Processa um par de entrada (*chave*, *valor*).
 - Produz um conjunto de pares intermediários.
 - $\text{reduce}(\text{chave}, \text{list}(\text{valor}; \text{intermediario})) \rightarrow \text{list}(\text{valor})$.
 - Combina todos os valores intermediários para uma chave.
 - Produz um conjunto que contém os valores de saída.

MapReduce

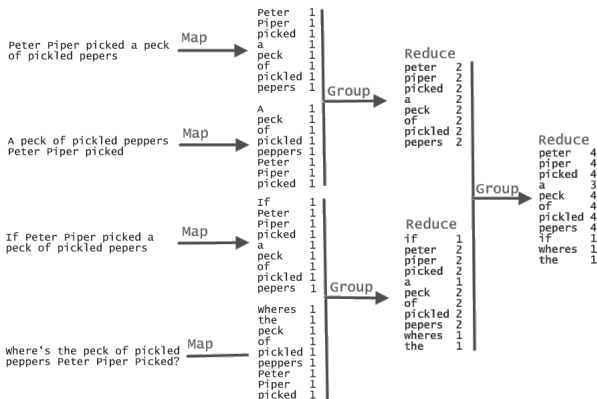


Figura: Exemplo de execução de um programa MapReduce. Contador de palavras [3].

HBase

- Características:
 - Banco de dados semi-estruturado.
 - Capacidade de indexação na escala do terabytes.
 - Recuperação da informação em tempo real.
- Entidades:
 - Colunas.
 - Linhas.
 - TimeStamp.

HBase

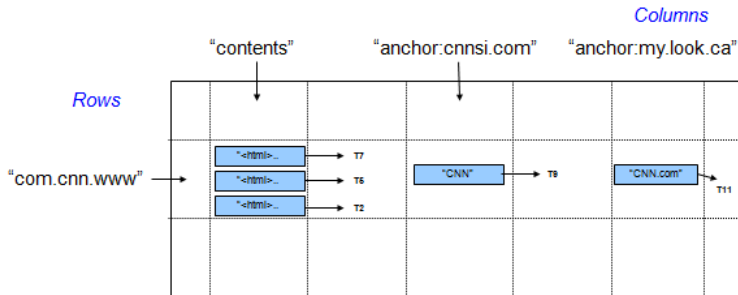


Figura: Modelo de dados básico do sistema HBase [4].

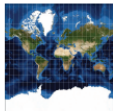
- 1 Introdução
 - O Problema
- 2 Trabalhos Relacionados
 - HDFS
 - MapReduce
 - HBase
- 3 Metodologia
 - Método
 - Análise de Complexidade
- 4 Conclusões

Método Proposto

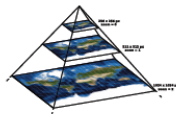
- Primeira etapa
 - Processamento do dado



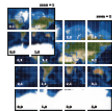
GEODETTIC COORDINATES
LATITUDE LONGITUDE
WGS 84 (EPSG:4326)



SPHERICAL MERCATOR
METERS
EPSG:3786 / EPSG:900913



PYRAMID COORDINATES
XYZ PIXELS / ZOOM
WEB VIEWERS



TILE INDEX
XYZ TILE / ZOOM
TILE MAP SERVICE

Figura: Modelo de processamento do dado. [5].

Método Proposto

- Segunda etapa
 - Indexação do dado.

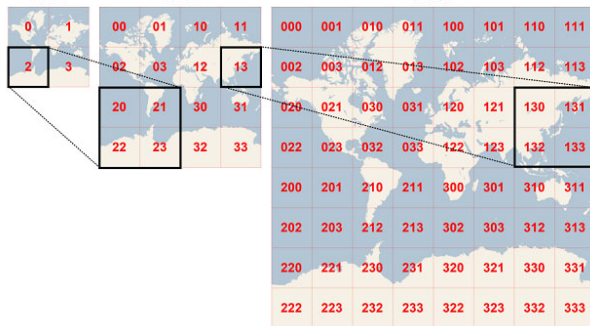


Figura: Modelo de indexação utilizando a QuadTree [6].

Análise de Complexidade

- Complexidade de tempo e espaço relacionadas ao:
 - Número de níveis do modelo de visualização - n .
 - $O(n)$.

- 1 Introdução
 - O Problema
- 2 Trabalhos Relacionados
 - HDFS
 - MapReduce
 - HBase
- 3 Metodologia
 - Método
 - Análise de Complexidade
- 4 Conclusões

Conclusão

- Necessidade de encontrar mais trabalhos relacionados as etapas propostas na metodologia.
- Experimentar os recursos existentes - HDSF, MapReduce e HBase.

-  D. Borthakur, *The Hadoop Distributed File System: Architecture and Design*, The Apache Software Foundation, 2007.
-  J. Dean and S. Ghemawat, “Mapreduce: simplified data processing on large clusters,” *Commun. ACM*, vol. 51, pp. 107–113, January 2008.
-  Tarn, “Map-reduce on mongo,”
<http://sharpthinking.com.au/author/tarn.aspx>, visitado em 04/07/2011.
-  F. Chang, J. Dean, S. Ghemawat, W. C. Hsieh, D. A. Wallach, M. Burrows, T. Chandra, A. Fikes, and R. E. Gruber, “Bigtable: A distributed storage system for structured data,” in *Proceedings of the 7th USENIX Symposium on Operating Systems Design and Implementation (OSDI'06)*, 2006.

[Online]. Available:

<http://labs.google.com/papers/bigtable.html>



K. P. Pridal, "Tiles à la google maps: Coordinates, tile bounds and projection - conversion to epsg:900913 (epsg:3785) and epsg:4326 (wgs84)," <http://www.maptiler.org/google-maps-coordinates-tile-bounds-projection/>, visitado em 11/07/2011.



"Bing maps tile system," <http://msdn.microsoft.com/en-us/library/bb259689.aspx>, visitado em 03/07/2011.