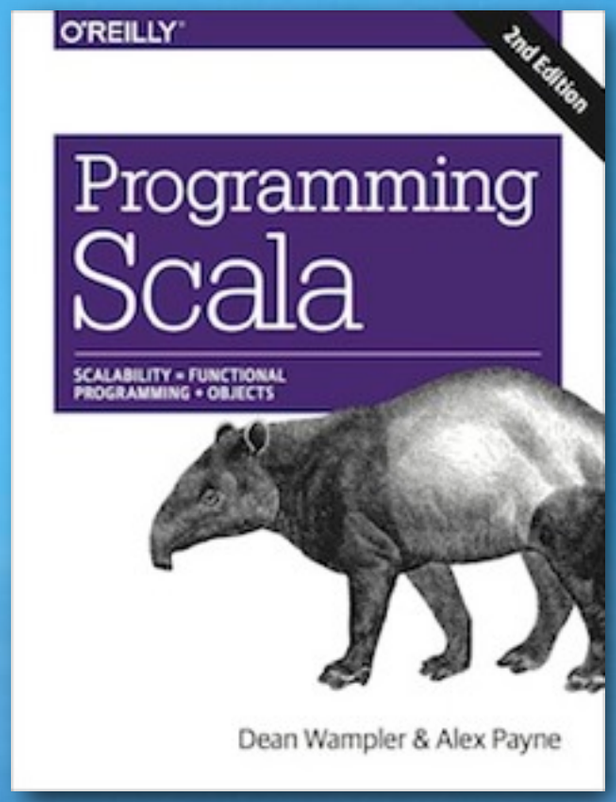
Typesafe

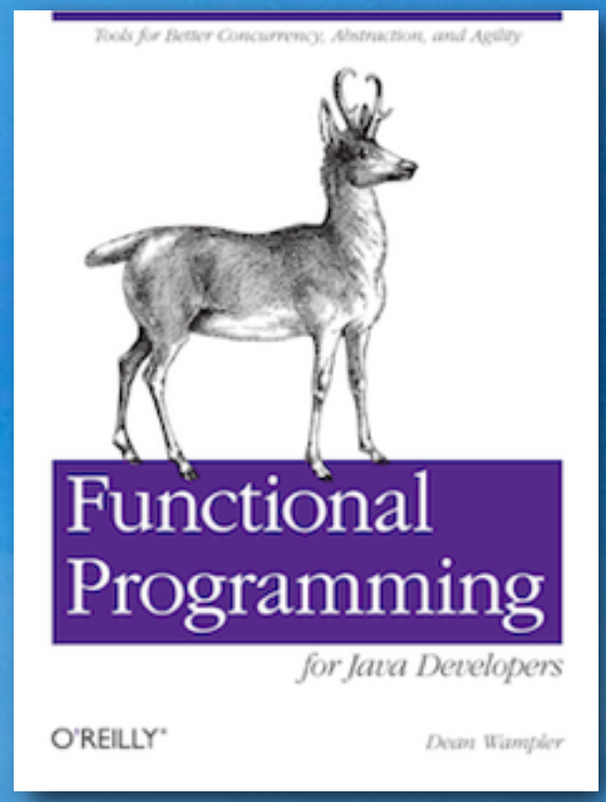
Why Spark Is the Next Top (Compute) Model

Philly ETE 2014
April 22-23, 2014
@deanwampler
polyglotprogramming.com/talks

Image: Detail of the London Eye

Dean Wampler





dean.wampler@typesafe.com
polyglotprogramming.com/talks
@deanwampler

About me. You can find this presentation and others on Big Data and Scala at polyglotprogramming.com.
Programming Scala, 2nd Edition is forthcoming.
photo: Dusk at 30,000 ft above the Central Plains of the U.S. on a Winter's Day.

Or this?

THE

Compleat Troller,

OR,

THE ART OF TROLLING.

WITH

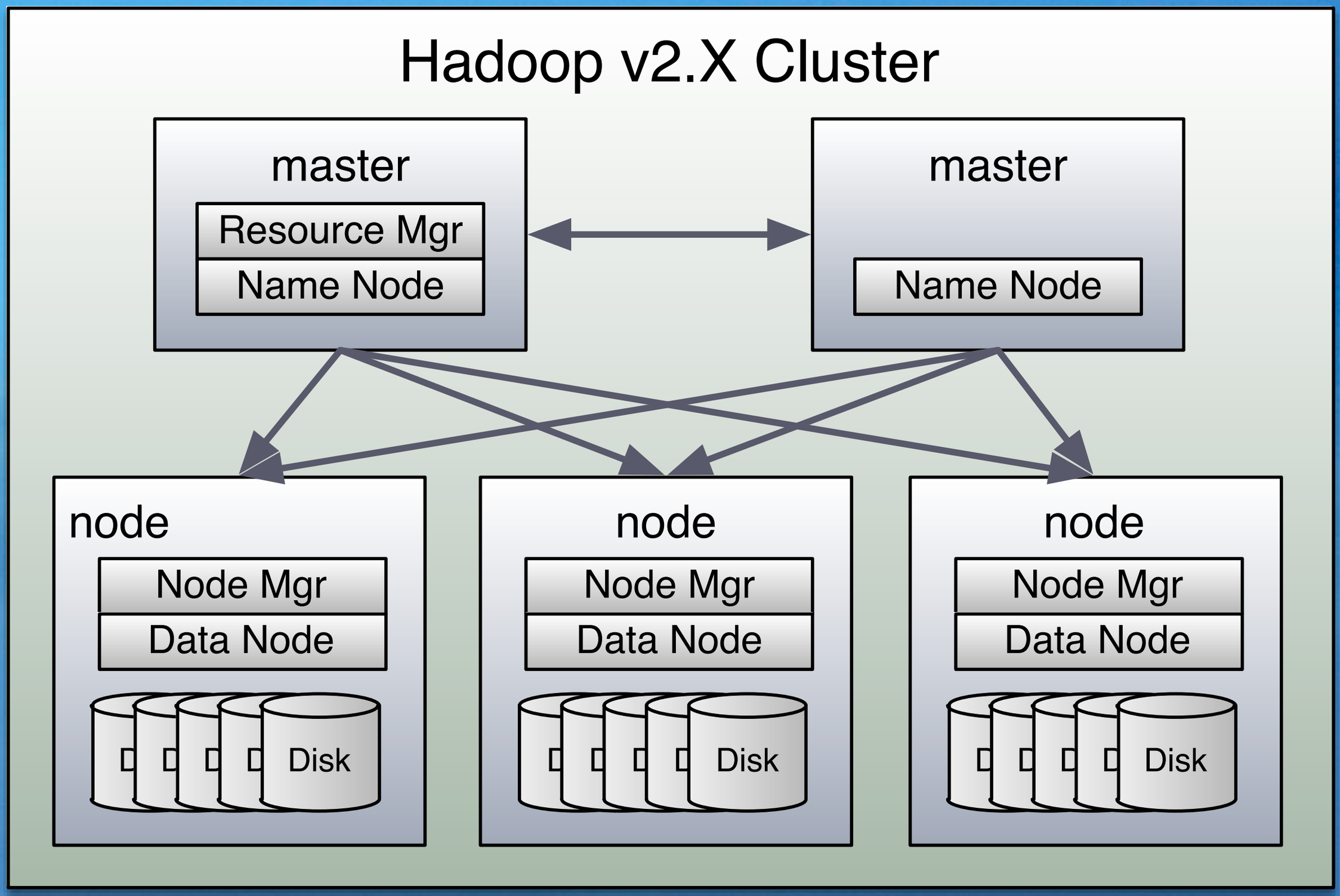
A Deſcription of all the Utenſils, Inſtruments, Tackling, and Materials requiſite thereto: With Rules and Directions how to uſe them

Saturday, April 19, 14

photo: https://twitter.com/john_overholt/status/447431985750106112/photo/1

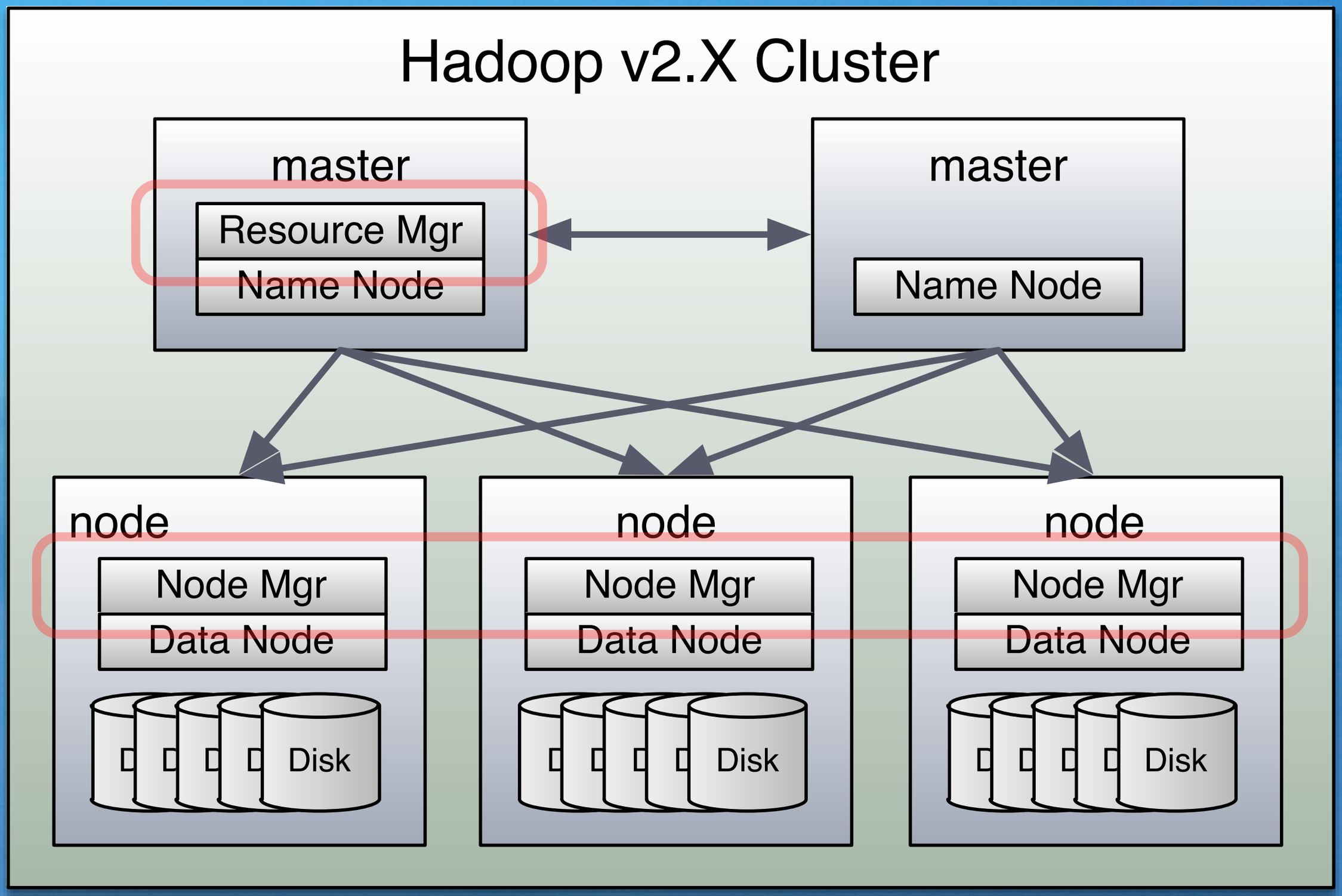
Hadoop circa 2013

The state of Hadoop as of last year.
Image: Detail of the London Eye



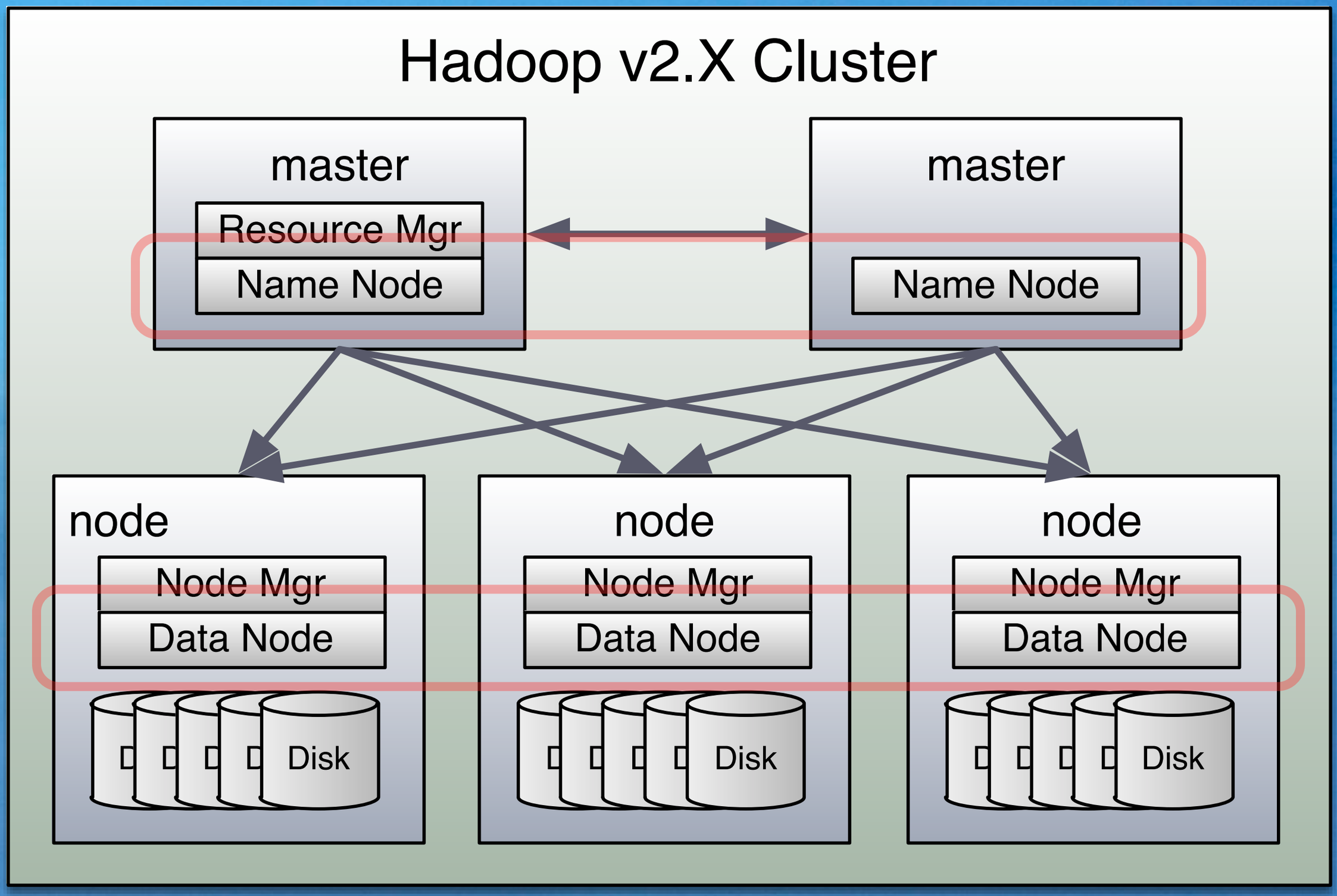
Schematic view of a Hadoop 2 cluster. For a more precise definition of the services and what they do, see e.g., http://hadoop.apache.org/docs/r2.3.0/hadoop-yarn/hadoop-yarn-site/YARN.html We aren't interested in great details at this point, but we'll call out a few useful things to know.

Resource and Node Managers



Hadoop 2 uses YARN to manage resources via the master Resource Manager, which includes a pluggable job scheduler and an Applications Manager. It coordinates with the Node Manager on each node to schedule jobs and provide resources. Other services involved, including application-specific Containers and Application Masters are not shown.

Name Node and Data Nodes



Hadoop 2 clusters federate the Name node services that manage the file system, HDFS. They provide horizontal scalability of file-system operations and resiliency when service instances fail. The data node services manage individual blocks for files.

MapReduce

The classic compute model for Hadoop

Hadoop 2 clusters federate the Name node services that manage the file system, HDFS. They provide horizontal scalability of file-system operations and resiliency when service instances fail. The data node services manage individual blocks for files.

MapReduce

*1 **map** step + 1 **reduce** step*

(wash, rinse, repeat)

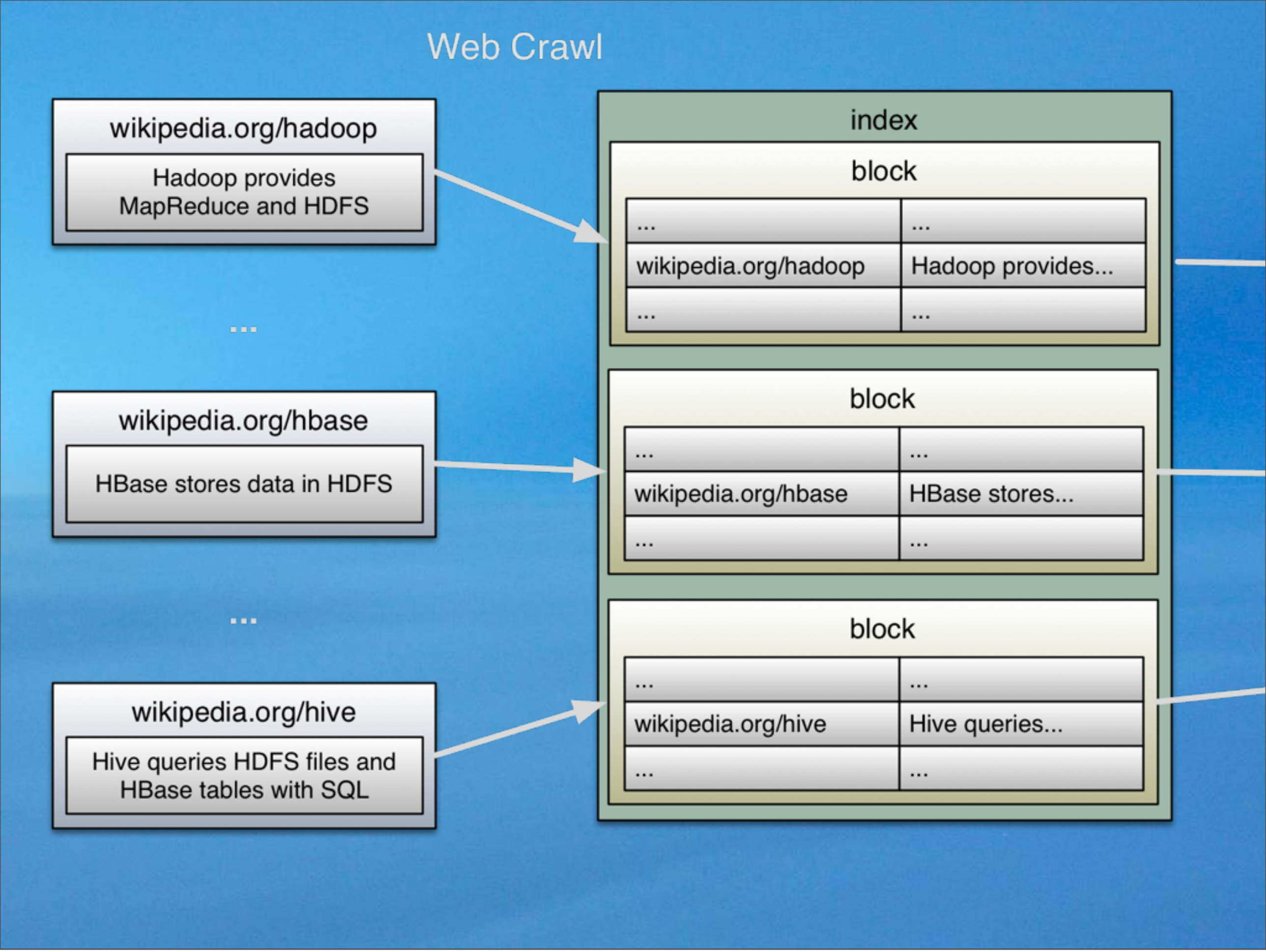
You get 1 map step (although there is limited support for chaining mappers) and 1 reduce step. If you can't implement an algorithm in these two steps, you can chain jobs together, but you'll pay a tax of flushing the entire data set to disk between these jobs.

MapReduce

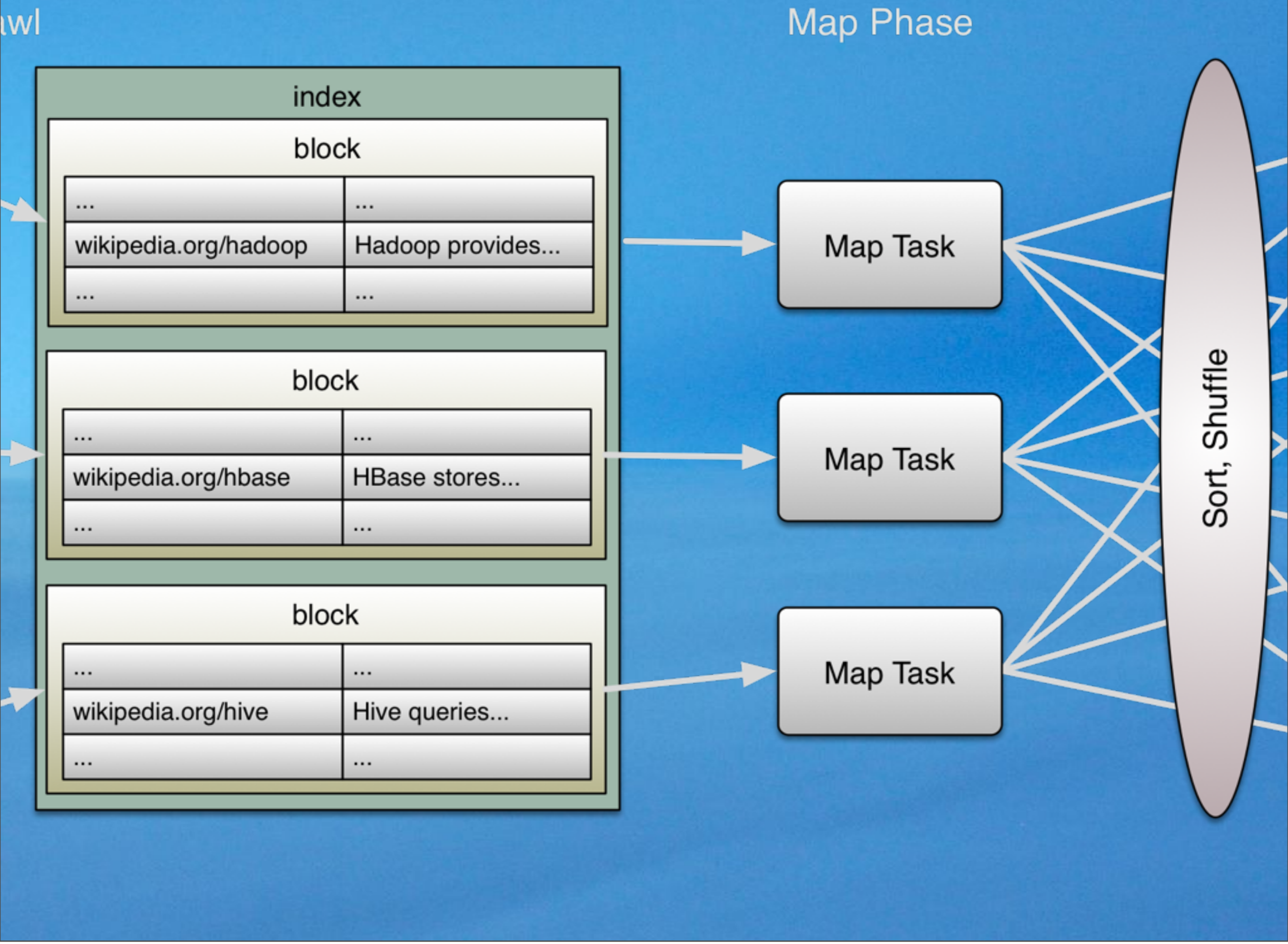
Example:

Inverted Index

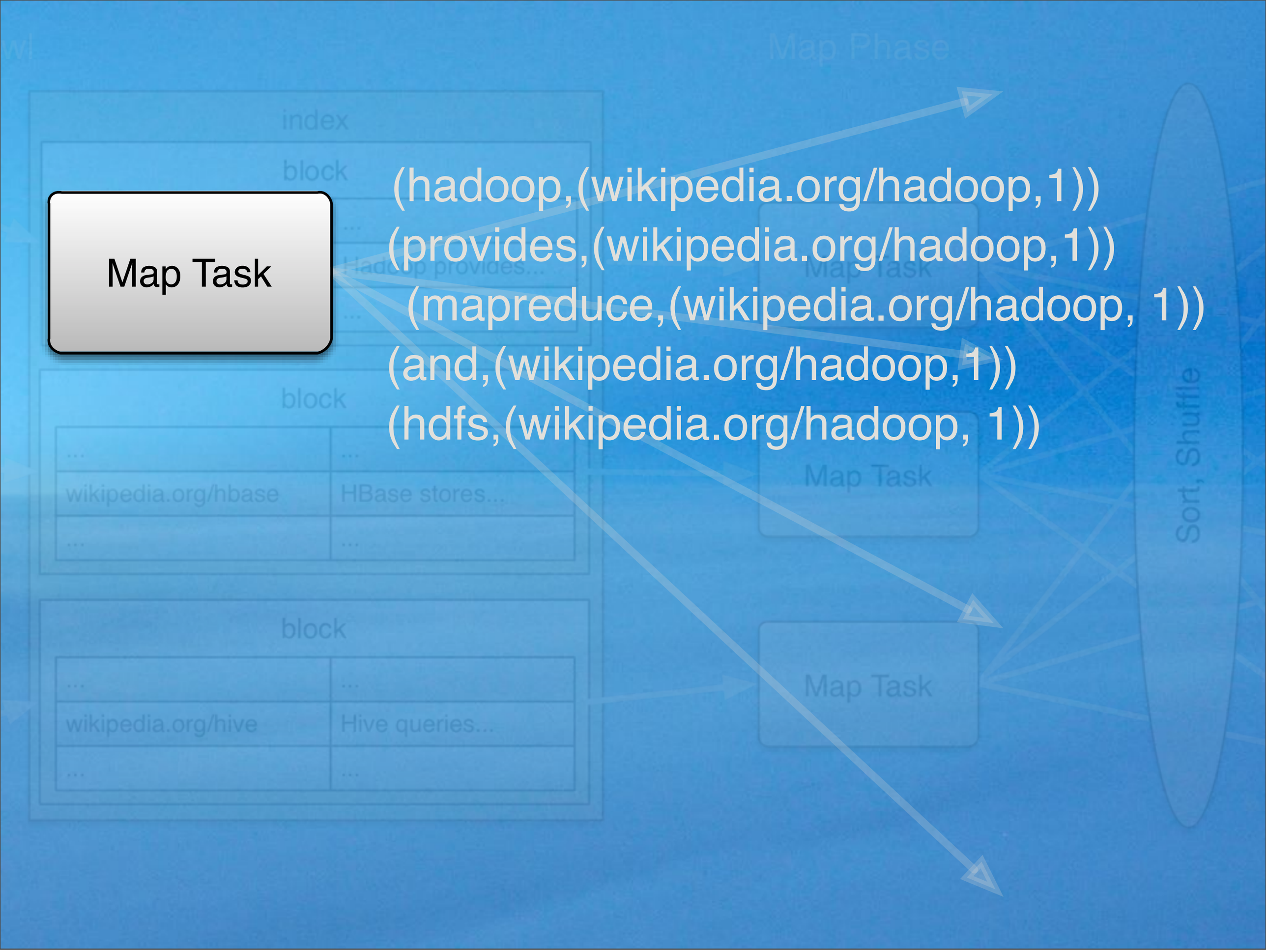
The inverted index is a classic algorithm needed for building search engines.



Before running MapReduce, crawl teh interwebs, find all the pages, and build a data set of URLs -> doc contents, written to flat files in HDFS or one of the more “sophisticated” formats.



In the map step, a task (JVM process) per file *block* (64MB or larger) reads the rows, tokenizes the text and outputs key-value pairs (“tuples”)...



Saturday, April 19, 14

... the keys are each word found and the values are themselves tuples, each URL and the count of the word. In our simplified example, there are typically only single occurrences of each work in each document. The real occurrences are interesting because if a word is mentioned a lot in a document, the chances are higher that you would want to find that document in a search for that word.

awl

Map Phase

index

block

...	...
wikipedia.org/hadoop	Hadoop provides...
...	...

block

...	...
wikipedia.org/hbase	HBase stores...
...	...

block

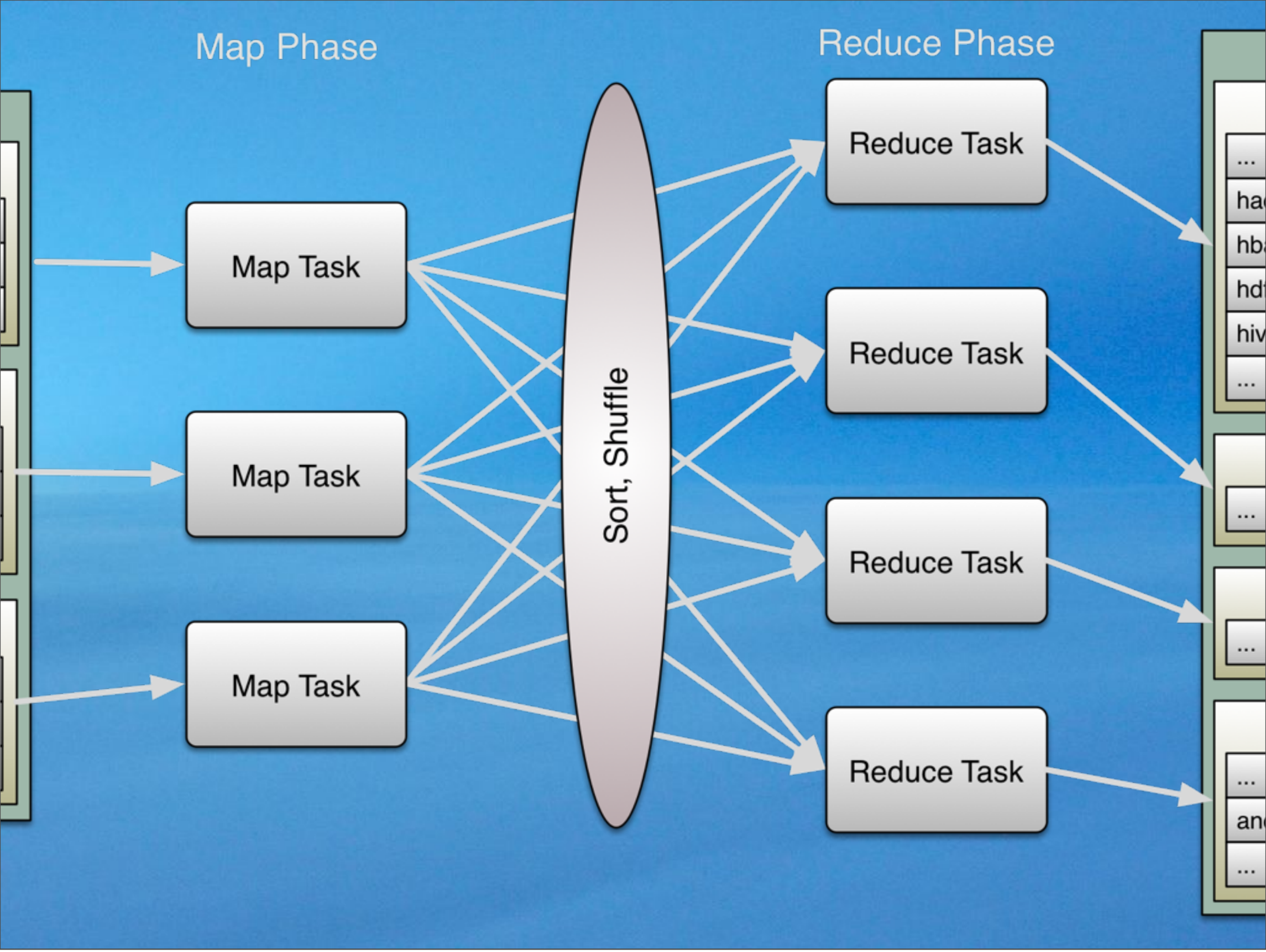
...	...
wikipedia.org/hive	Hive queries...
...	...

Map Task

Map Task

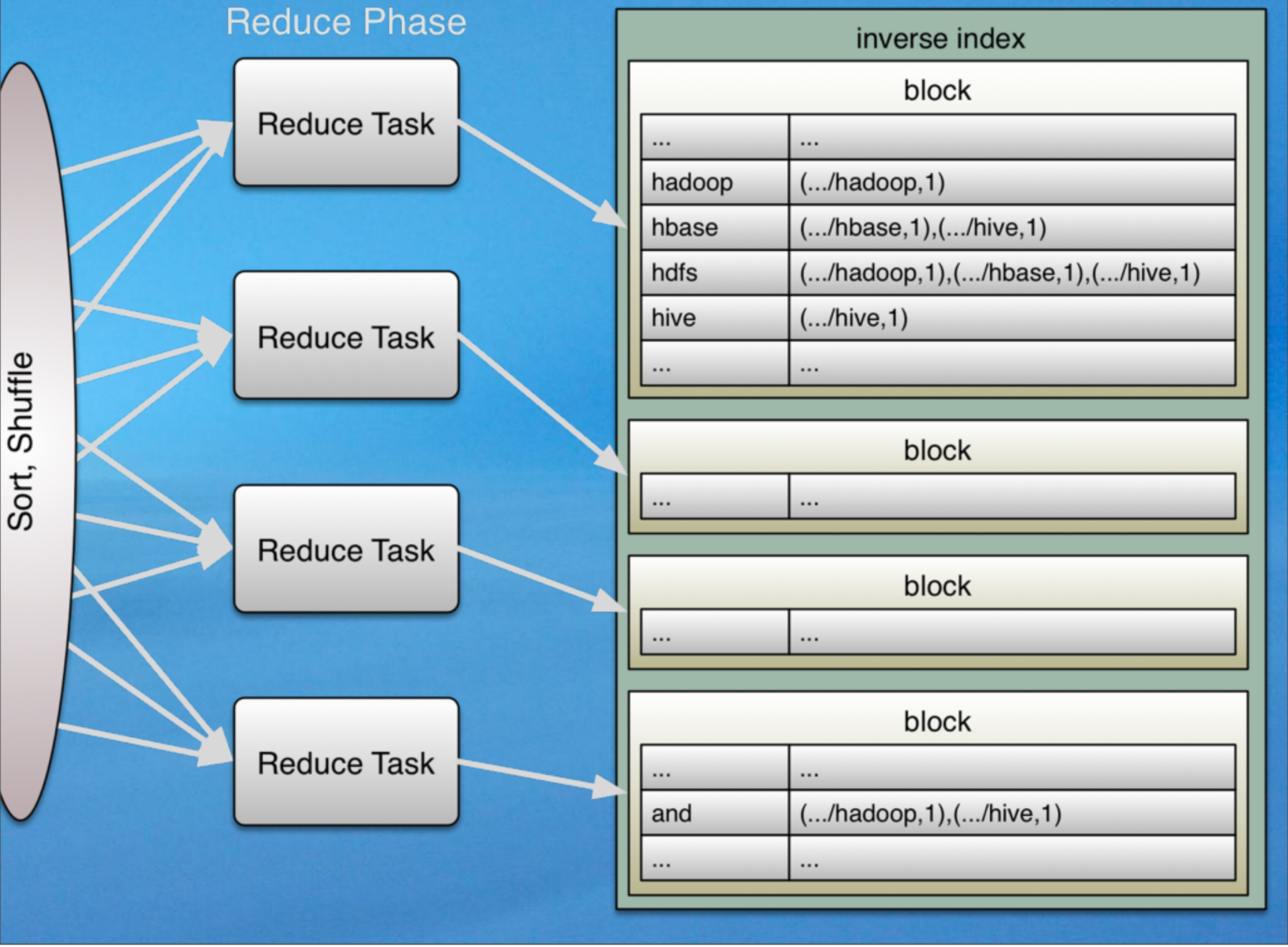
Map Task

Sort, Shuffle



Saturday, April 19, 14

The output tuples are sorted by key locally in each map task, then “shuffled” over the cluster network to reduce tasks (each a JVM process, too), where we want all occurrences of a given key to land on the same reduce task.



Finally, each reducer just aggregates all the values it receives for each key, then writes out new files to HDFS with the words and a list of (URL-count) tuples (pairs).

Typesafe

Why Spark Is the Next Top (Compute) Model

Philly ETE 2014
April 22-23, 2014
@deanwampler
polyglotprogramming.com/talks

Image: Detail of the London Eye