目录

[Hadoop 基础学习 1](#_Toc476486417)

[基础知识 1](#_Toc476486418)

[数据量进制 1](#_Toc476486419)

[存储 1](#_Toc476486420)

[4V: 2](#_Toc476486421)

[模块： 2](#_Toc476486422)

# Hadoop 基础学习

## 基础知识

### 数据量进制

1. bit = 1位
2. 1byte = 8bit
3. 1kb = 1024b
4. 1m = 1024k
5. 1g = 1024m
6. 1t = 1024g
7. 1p = 1024t
8. 1e = 1024p
9. 1z = 1024e
10. 1y = 1024z

### 存储

1. 分割

分布式 dfs: distributed file system

hdfs

1. 运算

mapreduce:map --> reduce 映射 –》化简

### 4V:

1. volumn 体量大
2. velocity 速度快
3. variaty 样式多
4. value 价值密度低

### 模块：

1. hadoop common
2. HDFS // hadoop distibuted file system
3. Hadoop YARN //yeah another resource ,作业调度和管理框架
4. Hadoop MapReduce //基于YARN 系统的大数据集并行处理

安装

1. jdk
2. tar hadoop.tar.gz
3. 配置环境变量

HADOOP\_HOME

PATH

1. 配置hadoop