MapReduce 课程设计

周华平，蒋雅楠，高翼枭

2018 年 2 月 28 日

小组信息

学号	姓名	邮箱	导师	研究领域
MG1733098	周华平	zhp@smail.nju.edu.cn	田臣	数据中心网络
MF1733026	蒋雅楠	mf1733026@smail.nju.edu.cn	田臣	数据中心网络
DZ1733004	高翼枭	dz1733004@smail.nju.edu.cn	田臣	数据中心网络

课题分工

本课题的选题和讨论由三人共同完成；项目的代码部分包含多个 MapReduce Job。具体分工如下：

姓名	主要工作
周华平	整体架构设计；各 Job 中流程设计与实现
蒋雅楠	文件输入输出；Distributed Cache 使用
高翼枭	数据连接；HDFS 读取、写入、删除文件

课程设计题目

我们研究的题目为“基于近邻成分分析的距离度量学习算法的并行化”。该题目来源于实现高级机器学习作业的过程中遇到的实际问题：在样本数据量较大或者维度较高的情况下，由于计算和存储开销较大，在单机上无法执行全梯度下降，因此该算法需要使用随机梯度下降代替常规的梯度下降来完成计算，这可能会导致算法无法正确收敛。

因此，在本课题中我们计划使用 MapReduce 对 NCA 算法进行并行化，使得该算法能够完整地运行在在较高维度和较大规模的数据集上，并且能够缩短训练时间。

摘要

本次课题，我们基于对度量学习过程中大量计算的并行化想法，来设计 MapRedue 任务。实现算法采用近邻成分分析，并使用梯度下降法求解目标函数。由于算法的数据量较

大，而且需要计算任意样本对当前分类样本的影响，求取样本被正确分类的最大概率，对计算与存储的要求都很高。于是我们将这样一个算法拆解成若干个 MapReduce Job，建立一个具有依赖关系的任务链。首先计算各个样本间的距离 x_{ij} 并写出至文件，便于下次使用；将需要更新的矩阵 A 放入 Distributed cache 中，便于各个节点读取操作；根据已得到的 x_{ij} 与矩阵 A 求解样本分类影响 p_{ij} 和 p_i 并输出；再利用中间结果进行数据连接计算 $p_{ij}x_{ij}x_{ij}^\top$。然后可以通过简单操作得到当前梯度，更新矩阵 A，并进入下一轮迭代，直至 A 的变化小于某一阈值，或达到迭代次数终止。

研究问题背景

在机器学习领域中，如何选择合适的距离度量准则一直都是一个重要而困难的问题。因为度量函数的选择非常依赖于学习任务本身，并且度量函数的好坏会直接影响到学习算法的性能。为了解决这一问题，我们可以尝试通过学习得到合适的度量函数。距离度量学习 (Distance Metric Learning, DML) 的目标是学习得到合适的度量函数，使得在该度量下更容易找出样本之间潜在的联系，进而提高那些基于相似度的学习器的性能。

在本课题中，我们采用近邻成分分析 (Neighbourhood Component Analusis, NCA) 来实现距离度量学习，

度量函数学习目标

根据马氏距离的定义

$$dist_{mah}^2(x,y) = (x-y)^\top Q(x-y) = (Ax-Ay)^\top(Ax-Ay)$$

其中 Q 称为 “度量矩阵”，而 DML 则是对 Q 的学习。为了保持距离非负且对称，Q 必须是 (半) 正定对称矩阵，即必有正交基 A 使得 Q 能写为 $Q = AA^\top$。

为了提高近邻分类器的性能，我们将 Q 直接嵌入到近邻分类器的评价指标中去，通过优化该性能目标相应地求得 Q。在本实验中我们采用近邻成分分析进行学习。

近邻分类器在进行判别时通常使用多数投票法，领域中的每个样本投 1 票，领域外的样本投 0 票。NCA 将其替换为概率投票法，对于任意样本 x_j，它对 x_i 分类结果影响的概率为

$$p_{ij} = \frac{\exp(-\|Ax_i - Ax_j\|^2)}{\sum_{k\neq i}\exp(-\|Ax_i - Ax_k\|^2)}, \qquad p_{ii} = 0$$

若以留一法正确率的最大化为目标，则可计算 x_i 的留一法正确率，即它被自身之外的所有样本正确分类的概率为

$$p_i = \sum_{j\in C_i} p_{ij}$$

其中 $C_i = \{j|c_i = c_j\}$，即与 x_i 属于相同类别的样本的下标集合。于是，整个样本集上被正确分类的点的个数的期望为

$$f(A) = \sum_i \sum_{j\in C_i} p_{ij} = \sum_i p_i$$

NCA 的优化目标是使得 $f(A)$ 最大化，即

$$\max_A \sum_i \sum_{j \in C_i} \frac{\exp(-\|Ax_i - Ax_j\|^2)}{\sum_{k \neq i} \exp(-\|Ax_i - Ax_k\|^2)}$$

优化算法

梯度下降法可以用来求解目标函数：通过求 f 对 A 的偏导，可以得到梯度公式（令 $x_{ij} = x_i - x_j$）

$$\frac{\partial f}{\partial A} = -2A \sum_i \sum_{j \in C_i} p_{ij}(x_{ij}x_{ij}^\top - \sum_k p_{ik}x_{ik}x_{ik}^\top)$$

根据该公式，使用梯度下降法即可求解 NCA 的目标函数。得到最大化近邻分类器留一法正确率的距离度量矩阵 Q。

主要技术难点和拟解决的问题

在高级机器学习课程中，我们使用了 Python 实现了该算法。在该实现中，为了尽可能利用 numpy 高效的矩阵操作，我们需要将中间结果的计算尽可能转化为矩阵的运算，从而提高并行度。然而当样本的维度较高或者样本数据量较大时，无论是存储中间结果矩阵还是计算梯度的开销都会大到单机无法承受的程度。

因此，在实际实验中我们会用随机梯度下降来代替全梯度下降，即以计算梯度中的某些项来代替梯度全体，以此降低计算和存储开销。然而随机梯度下降并不是保证迭代将沿着目标函数的最快下降方向前进，甚至不保证其沿着下降方向前进。这有可能导致算法收敛过慢、无法正确地收敛、以及在接近最优解附近时精度较差等问题。

在本课题中我们计划使用 MapReduce 来完成 NCA 算法中梯度下降的并行化，使其在能够处理维度较高、数据量较大的样本训练的同时缩短训练时间。

该算法需要使用多个 MapReduce 作业，通过迭代的方式完成计算。对 $\frac{\partial f}{\partial A}$ 的计算分为了几个阶段，我们需要将其中的计算抽象为若干个 MapReduce 过程，合理地设计每个阶段的输入和输出，并将不同阶段组织成具有依赖关系的任务链；在中间结果的计算中还需要涉及多数据源的连接。简单来说，我们首先得到样本间的距离文件 x_{ij} 和初始矩阵 A，然后可以利用 MapReduce 并行地计算 $\exp(-\|Ax_i - Ax_j\|^2)$，此处并行可以大量减少计算时间；接着将该结果作为输入，可以并行地计算 p_{ij} 和 p_i；利用上述中间结果，我们可以通过 DataJoin 来计算 $p_{ij}x_{ij}x_{ij}^\top$；最后计算出当前的 $\frac{\partial f}{\partial A}$ 并更新矩阵 A，完成一次迭代。

基本解决方法和设计思路

NCA 模型训练

对于 NCA，我们拟采用组合式 MapReduce 计算作业来实现。由于梯度下降法需要用迭代方法求得逼近结果，因此在 NCA 的主控程序中，需要使用一个循环来控制 MapReduce 作业的执行，直到第 n 次迭代后结果与第 $n-1$ 次的结果小于某个指定的阈值时结束，或者通过经验值可确定在运行一定的次数后能得到接近的最终结果，也可以控制循环固定的次数。

对于梯度下降中需要求解的变量，我们采用顺序组合式 MapReduce 作业来依次计算：首先我们可以通过 DataJoin 对集合 X 做笛卡尔积，进而可以计算出 x_{ij}；在此基础上以 x_{ij} 与矩阵 A 作为输入，我们可以进一步执行 MapReduce 任务计算出中间结果 r1: $\exp(-\|Ax_i - Ax_j\|^2)$，其中矩阵 A 作为 Distributed Cache 在 Mapper 的 setup() 阶段读入；基于以上结果，我们可以通过下一次 MapReduce 任务对所有 i 计算出中间结果 r2: $\sum_{k\neq i}\exp(-\|Ax_i - Ax_k\|^2)$，再通过 DataJoin 连接上述两个中间结果 r1 与 r2，我们可以计算出 p_{ij}。

p_i 的计算需要考虑 x_j 与 x_i 的 label 是否相同。因此我们可以通过接下来的 MapReduce 任务，在 setup() 中读入 distributed cache 中的 lable 文件，生成一个 i 与 lable 的映射表，并通过 Mapper 过滤掉 x_j 与 $x_i$label 不同的元素，在 Reducer 端只需做简单的求和即可。

基于以上数据，我们可以对梯度 $\frac{\partial f}{\partial A}$ 进行求解。我们首先使用 DataJoin 计算出中间结果 $p_{ij}x_{ij}x_{ij}^\top$，然后使用不同的 Mapper 过滤元素，分别计算 $\sum_k p_{ik}x_{ik}x_{ik}^\top$ 以及 $\sum_{j\in C_i} p_{ij}x_{ij}x_{ij}^\top$，需要注意的是由于这两者的计算没有依赖关系，所以可以通过配置 Job 使它们并行执行。在计算出以上两个求和的结果后，梯度 $\frac{\partial f}{\partial A}$ 可以通过简单的运算操作得出，这里不做赘述。在每次迭代的最后，我们利用当前位置的梯度对矩阵 A 进行更新，并启动下一次迭代。

详细设计说明

自定义数据类型

MatrixWritable

由于 NCA 中涉及了大量的矩阵运算，为了更加方便和效率地存储和表示矩阵，我们实现了 MatrixWritable 类。该类实现了 Writable 接口，能够对矩阵进行序列化和反序列化。MatrixWritable 底层使用 math3.linear.RealMatrix 来表示矩阵，该类来自于 commons-math3 包,它提供了一系列矩阵相关的操作。NCA 中大部分的中间结果都采用 MatrixWritable 来表示和存储。

TaggedEntry

在使用 DataJoin 时，DataJoinMapperBase 默认会忽略 map 函数的 key 参数，而将 GroupKey 作为输出的 Key。这样做的问题在于当处理 SequenceFile 或者其他 Key-Value 类型的输入数据时，Key 参数在 map 阶段就会被丢弃。

针对这种情况，我们实现了 TaggedEntry 基类，它继承自 TaggedMapOutput 基类，将 Key 也作为序列化数据的一部分，这样在 Reducer 端除了 GroupKey 以外，记录的 Key 依旧能够得到保留。TaggedMatrix 继承了该类，它的 Value 类型为 MatrixWritable。

中间数据格式

由于 NCA 采用了组合式 MapReduce 计算作业来实现，程序需要多趟迭代，每一趟又由多个 Job 串联而成，因此对于中间结果我们需要使用更高效的方式来存储。为此我们使用 SequenceFile 作为中间数据的文件格式，其中 Key 类型为 Text，Value 类型为 Ma-

trixWritable。Key 总共有两种形式：对于类似 p_{ij} 的中间数据，Key 为 “i,j” 这样的值对，中间用逗号分隔；对于类似 p_i 的中间数据，Key 为 “i”。

功能模块

在 NCA 中的涉及的运算主要可以分为 3 类。首先是输入和输出一一对应的情况。例如根据 x_{ij} 计算 $x_{ij}x_{ij}^\top$。对此我们只需要对每种运算分别实现一个 Mapper 即可。例如 XXtMapper 和 ExpSquaredNormMapper 分别用于计算 $x_{ij}x_{ij}^\top$ 和 $\exp(-\|Ax_i - Ax_j\|^2)$，

其次是输出对输入聚合的情况。例如根据 p_{ij} 计算 $p_i = \sum_{j\in C_i} p_{ij}$。这种是典型的 MapReduce 的应用：我们使用 Mapper 对输入进行过滤 (Filter)，然后在 Reducer 端对结果进行聚合。

由于在程序中多次使用到了类似的编程范式，因此我们对于这些算子进行了进一步的抽象。具体说来，GroupMapper 将 KEYIN“i,j” 中的 i 作为 map() 函数的 KEYOUT，而 SameLabelMapper 则会忽略掉那些 $y_i \neq y_j$ 的元组，即 Mapper 输出的结果必定具有相同的 Label。对于这类运算我们仅实现了 SumMatReducer 这一种 Reducer。通过 Mapper 和 Reducer 的不同组合，我们可以实现不同的运算。例如使用 SameLabelMapper 和 SumMatReducer 即可计算 $\sum_{j\in C_i} p_{ij}$；而使用 GroupMapper 和 SumMatReducer 即可计算 $\sum_k p_{ik}x_{ik}x_{ik}^\top$。

最后一类是输出需要连接多个输入源的情况。例如根据 p_{ij} 和 $x_{ij}x_{ij}^\top$ 来计算 $p_{ik}x_{ik}x_{ik}^\top$。为此我们使用 DataJoin 来进行多数据源的连接。由于在 NCA 中我们主要使用二元连接操作，因此我们对这种情况进行了特殊的处理，将输入源分为了 Left 和 Right，代表运算符两边的操作数。我们实现了 EntryJoinMapperBase 基类，它继承自 DataJoinMapperBase。我们通过 configure() 读取在 Configuration 中设置的 Left 和 Right 对应的文件名，在此基础上我们重载了 generateInputTag()，根据 inputFile 的路径判断该分片是属于 Left 还是 Right，并设置相应的 Tag。除此之外，我们还需要修改 map()，使得 Key 能够通过 TaggedEntry 来保留。至此我们就能够将输入源打上 Tag 并发送到 Reducer 端。并且由于 Reducer 端接收到的元组会根据它们的 Tag 进行排序，我们可以很容易地通过巧妙设置 Tag 的字面值，达到在 Array 中 Left 必定在 Right 前面的效果，所以在 Reducer 端不需要判断 Tag 的字面值而可以直接使用变量值。

在 EntryJoinMapperBase 的基础上，我们分别实现了 DefaultMapper 和 GroupMapper。其中 GroupMapper 和用于聚合的 GroupMapper 具有相似的作用，而 DefaultMapper 将元组的 Key 直接作为 GroupKey。

通过继承 EntryJoinReducerBase，我们实现了连接端不同的矩阵运算操作：例如 NumMulMatReducer 和 MatSubReducer 分别实现了数乘矩阵和矩阵减法操作。

通过组合使用 EntryJoinMapperBase 和 EntryJoinReducerBase，我们可以实现各种矩阵的二元运算，这直接体现在了代码中，这里就不过多赘述。

程序框架说明

上一节中所介绍的第一和第二类运算相关的实现位于 NCA 类中；而第三类运算的实现位于 MatJoin 类中。NCADriver 类负责对输入参数进行处理并调度 MapReduce Job 的执

行。NCADriver 在开始执行时会调用一次 init()，用于对输入数据转化为 SequenceFile 格式，然后计算 x_{ij} 和 $x_{ij}x_{ij}^{\top}$，将以上输出结果保存到 HDFS 中以供后续训练使用。这样能省去许多重复计算。

接下来 NCADriver 每调用一次 train() 都会调度一系列 MapReduce Job 来计算梯度，最后将结果累加到 A 上并保存到 HDFS 上，从而完成一趟训练。

输入输出文件格式

输入文件格式

输入文件均为文本文件，其中一个是 x_{ij}，即 $x_i - x_j$。其每行的格式为 i,j\t$x_i[0] - x_j[0] \dots x_i[n] - x_j[n]$，向量中的每个值用空格分隔；

另一个输入文件是 y_i，即每个元组的标记信息。其每行的格式为 i\t$label_i$。

输出文件格式

输出的数据是训练得到的距离度量矩阵 A,它是一个二进制文件,通过使用 Utils.serializeMatrix() 和 Utils.deserializeMatrix() 可以对其进行序列化和反序列化操作。

程序运行实验结果和说明分析

为了测试 NCA 的性能，我们使用一个包含 1000 条数据的数据集，其中每条数据的维度为 16。通过提交到集群执行，我们观察到每一趟训练产生的中间结果可以达到将近 10G，这还是在输入数据量不是很大的情况下得到的结果，我们能够很明显感受到维度爆炸带来的影响：这样的数据量很难在单机上有效处理。

运行方式说明

我们通过以下命令执行 NCADriver。其中 raw_x_ij.txt 代表文本文件 x_{ij}；label.txt 代表文本文件 y_i；dim 代表数据维度；train 代表工作目录；matA 代表矩阵 A 所在的路径；lr 代表学习率；epoch 代表迭代轮数。

```
hadoop jar nca.jar nca.NCADriver
        raw_x_ij.txt label.txt dim train matA lr epoch
```

图 1: 部分执行过程

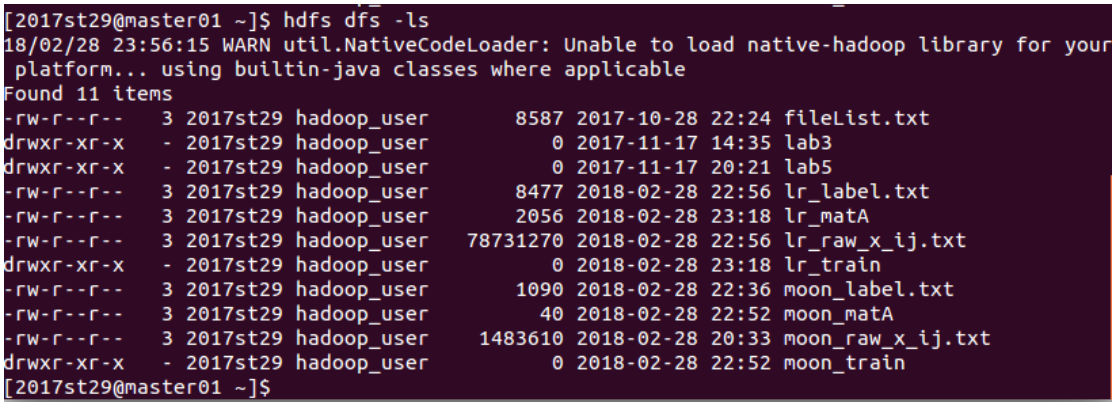
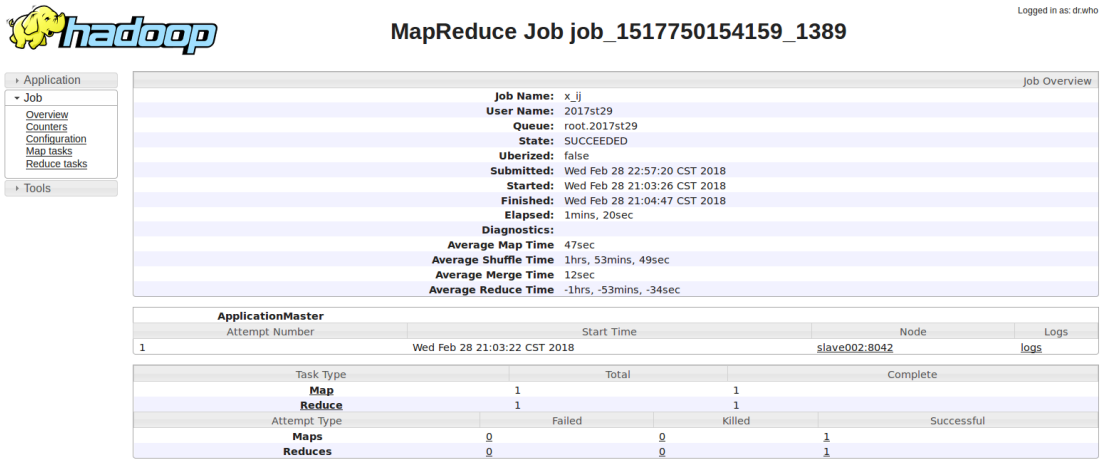


图 2: HDFS 文件列表

程序执行报告

以下为某一趟训练产生的 Job 的执行报告。



Logged in as: dr.who

MapReduce Job job_1517750154159_1389

- Application
- Job
  - Overview
  - Counters
  - Configuration
  - Map tasks
  - Reduce tasks
- Tools

Job Overview

Job Name:	x_ij
User Name:	2017st29
Queue:	root.2017st29
State:	SUCCEEDED
Uberized:	false
Submitted:	Wed Feb 28 22:57:20 CST 2018
Started:	Wed Feb 28 21:03:26 CST 2018
Finished:	Wed Feb 28 21:04:47 CST 2018
Elapsed:	1mins, 20sec
Diagnostics:	
Average Map Time	47sec
Average Shuffle Time	1hrs, 53mins, 49sec
Average Merge Time	12sec
Average Reduce Time	-1hrs, -53mins, -34sec

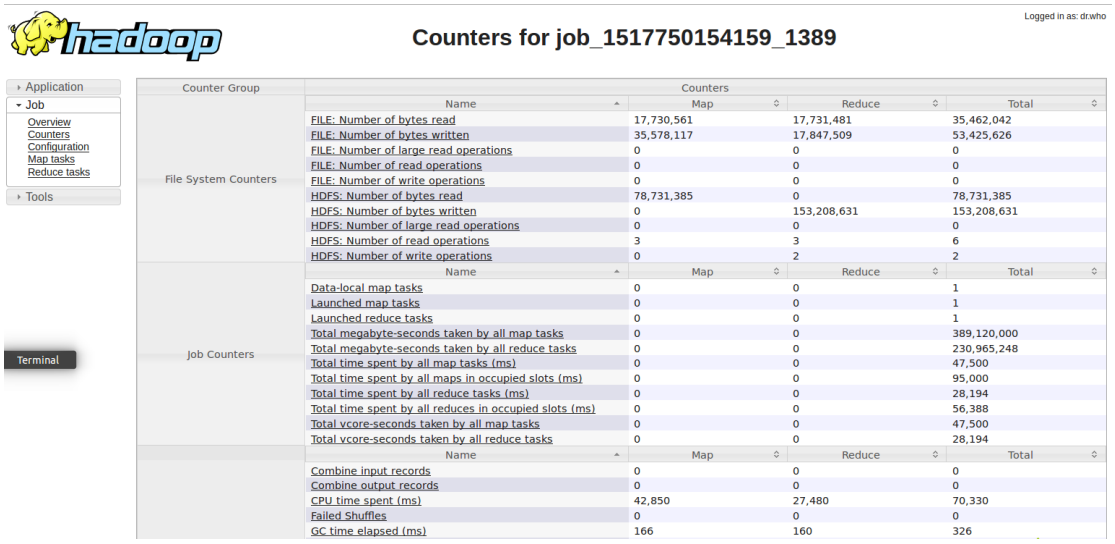
ApplicationMaster

Attempt Number	Start Time	Node	Logs
1	Wed Feb 28 21:03:22 CST 2018	slave002:8042	logs

Task Type	Total	Complete
Map	1	1
Reduce	1	1

Attempt Type	Failed	Killed	Successful
Maps	0	0	1
Reduces	0	0	1

图 3: x_ij Job



Logged in as: dr.who

Counters for job_1517750154159_1389

- Application
- Job
  - Overview
  - Counters
  - Configuration
  - Map tasks
  - Reduce tasks
- Tools

Terminal

Counter Group	Name	Map	Reduce	Total
File System Counters	FILE: Number of bytes read	17,730,561	17,731,481	35,462,042
	FILE: Number of bytes written	35,578,117	17,847,509	53,425,626
	FILE: Number of large read operations	0	0	0
	FILE: Number of read operations	0	0	0
	FILE: Number of write operations	0	0	0
	HDFS: Number of bytes read	78,731,385	0	78,731,385
	HDFS: Number of bytes written	0	153,208,631	153,208,631
	HDFS: Number of large read operations	0	0	0
	HDFS: Number of read operations	3	3	6
	HDFS: Number of write operations	0	2	2
Job Counters	Name	Map	Reduce	Total
	Data-local map tasks	0	0	1
	Launched map tasks	0	0	1
	Launched reduce tasks	0	0	1
	Total megabyte-seconds taken by all map tasks	0	0	389,120,000
	Total megabyte-seconds taken by all reduce tasks	0	0	230,965,248
	Total time spent by all map tasks (ms)	0	0	47,500
	Total time spent by all maps in occupied slots (ms)	0	0	95,000
	Total time spent by all reduce tasks (ms)	0	0	28,194
	Total time spent by all reduces in occupied slots (ms)	0	0	56,388
	Total vcore-seconds taken by all map tasks	0	0	47,500
	Total vcore-seconds taken by all reduce tasks	0	0	28,194
	Name	Map	Reduce	Total
	Combine input records	0	0	0
	Combine output records	0	0	0
	CPU time spent (ms)	42,850	27,480	70,330
	Failed Shuffles	0	0	0
	GC time elapsed (ms)	166	160	326

图 4: x_ij Counters

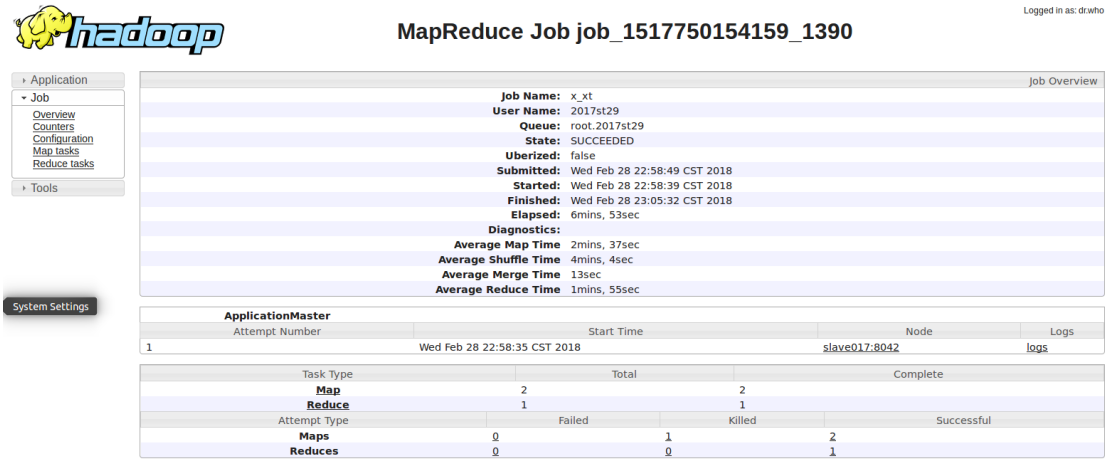


图 5: x_xt Job

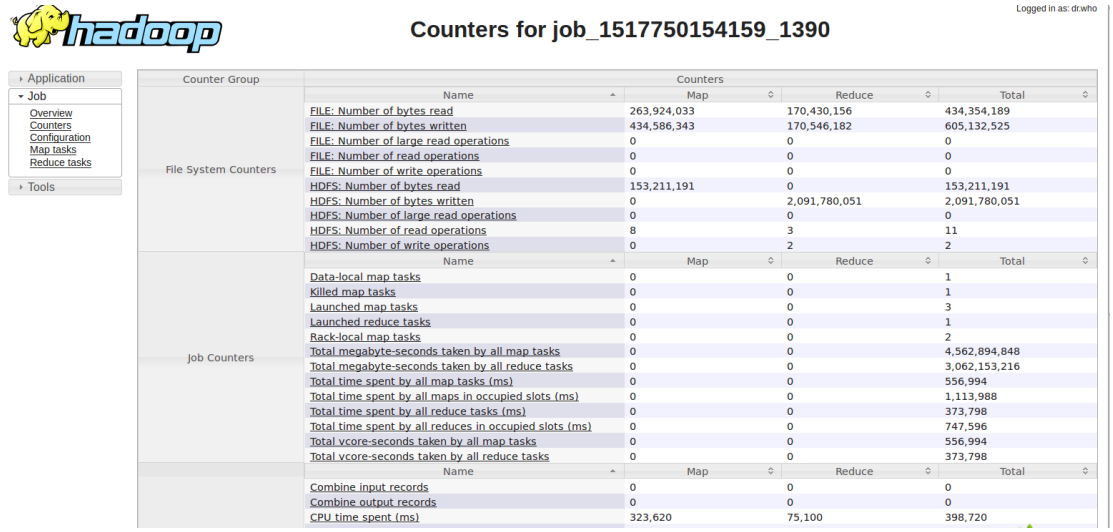


图 6: x_xt Counters

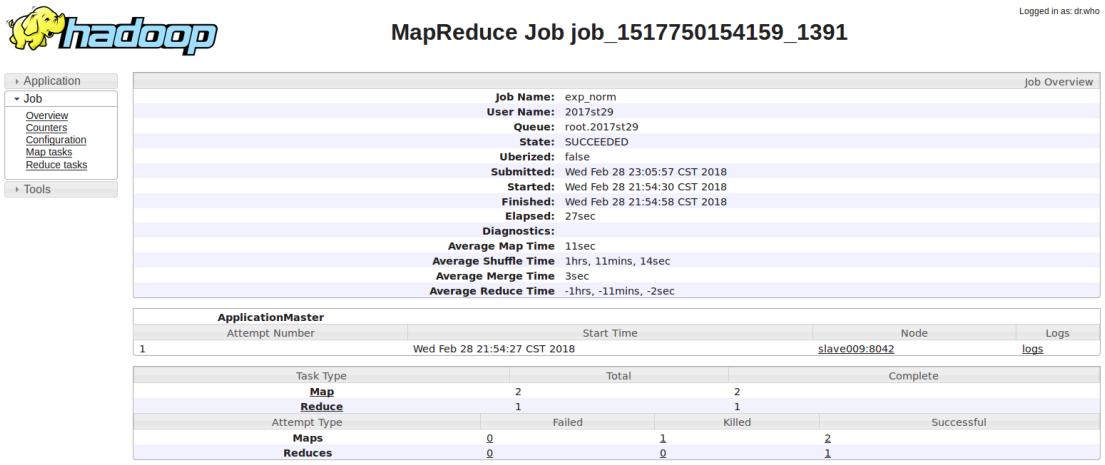


图 7: exp_norm Job

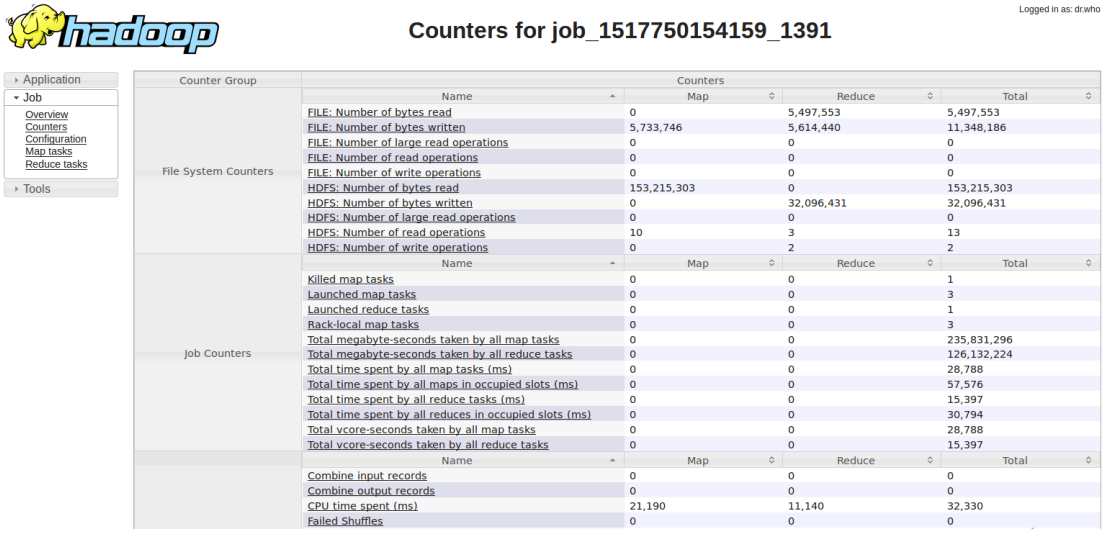


图 8: exp_norm Counters

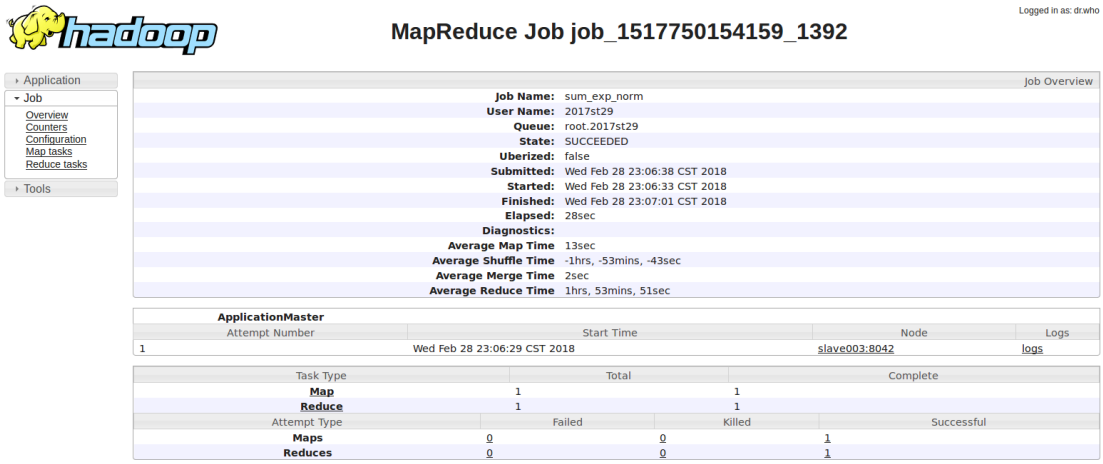


图 9: sum_exp_norm Job

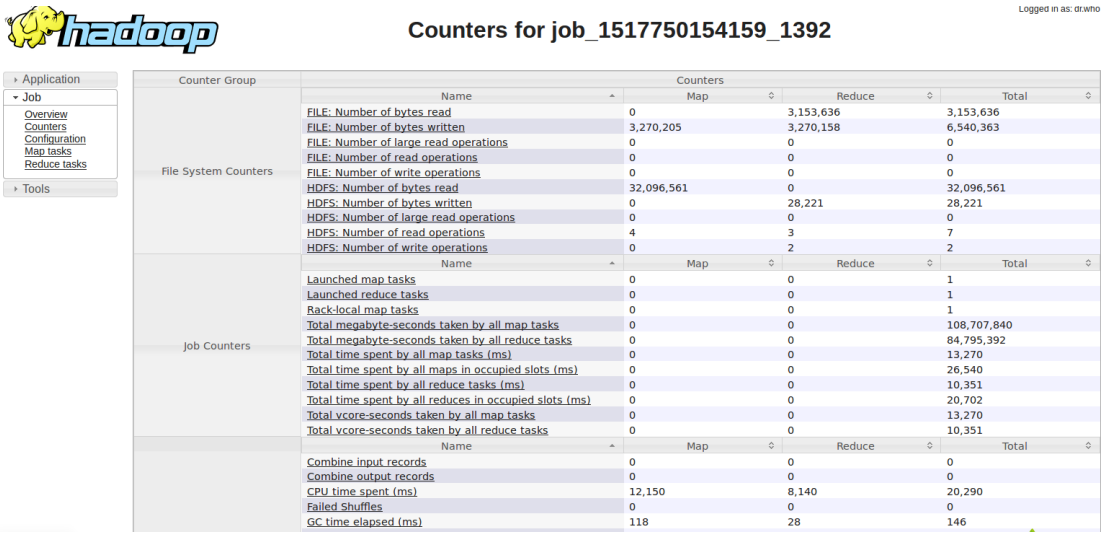
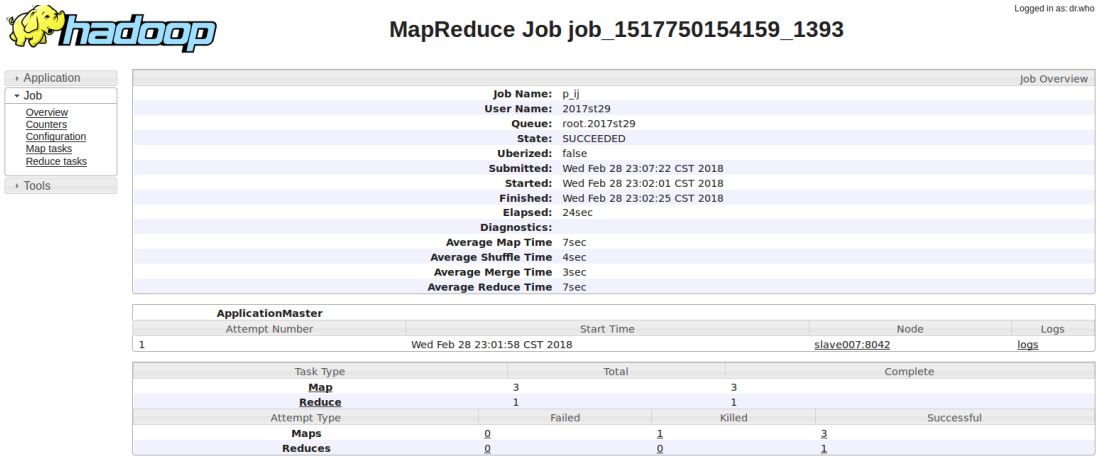


图 10: sum_exp_norm Counters



Logged in as: dr.who

MapReduce Job job_1517750154159_1393

Application
Job
Overview
Counters
Configuration
Map tasks
Reduce tasks
Tools

Job Overview

Job Name:	p_ij
User Name:	2017st29
Queue:	root.2017st29
State:	SUCCEEDED
Uberized:	false
Submitted:	Wed Feb 28 23:07:22 CST 2018
Started:	Wed Feb 28 23:02:01 CST 2018
Finished:	Wed Feb 28 23:02:25 CST 2018
Elapsed:	24sec
Diagnostics:	
Average Map Time	7sec
Average Shuffle Time	4sec
Average Merge Time	3sec
Average Reduce Time	7sec

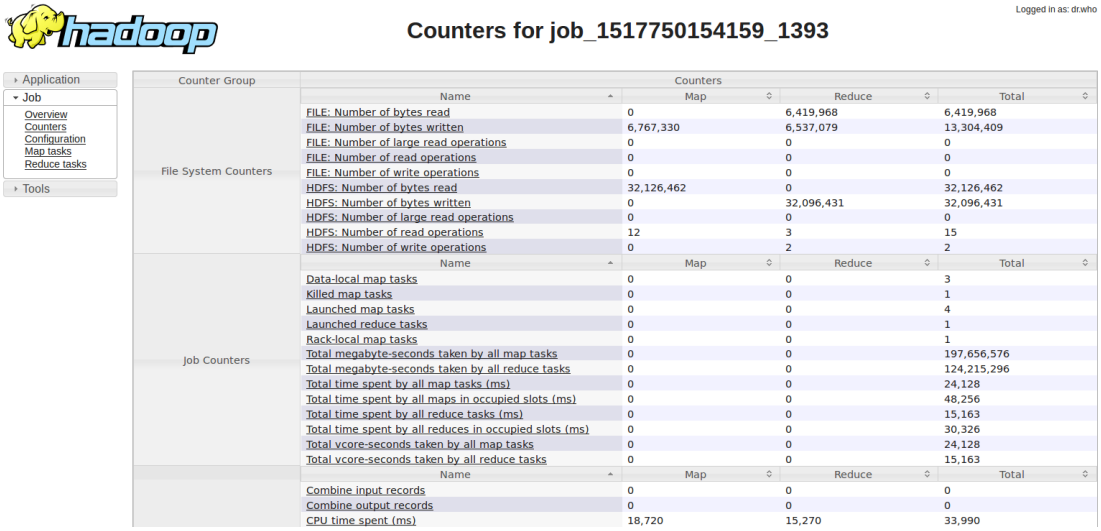
ApplicationMaster

Attempt Number	Start Time	Node	Logs
1	Wed Feb 28 23:01:58 CST 2018	slave007:8042	logs

Task Type	Total	Complete
Map	3	3
Reduce	1	1

Attempt Type	Failed	Killed	Successful
Maps	0	1	3
Reduces	0	0	1

图 11: p_ij Job



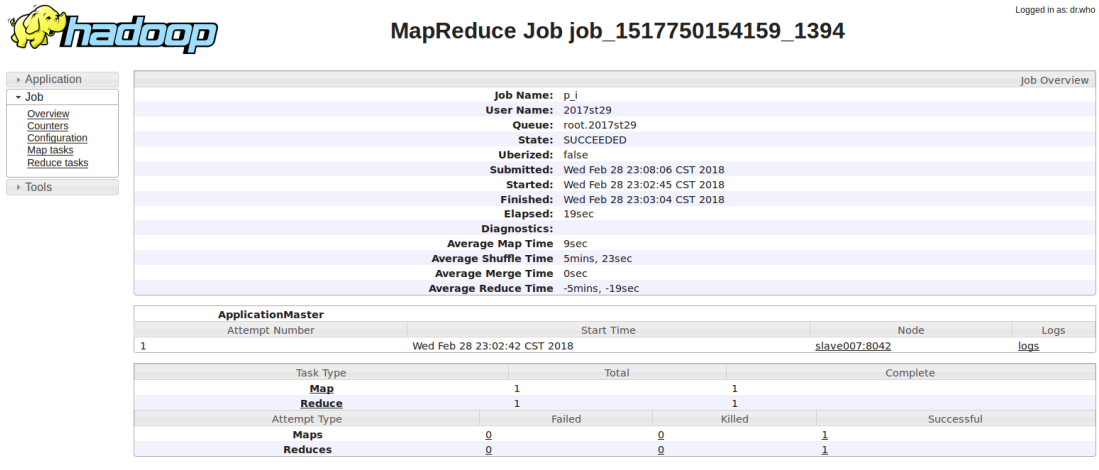
Logged in as: dr.who

Counters for job_1517750154159_1393

Application
Job
Overview
Counters
Configuration
Map tasks
Reduce tasks
Tools

Counter Group	Name	Map	Reduce	Total
File System Counters	FILE: Number of bytes read	0	6,419,968	6,419,968
	FILE: Number of bytes written	6,767,330	6,537,079	13,304,409
	FILE: Number of large read operations	0	0	0
	FILE: Number of read operations	0	0	0
	FILE: Number of write operations	0	0	0
	HDFS: Number of bytes read	32,126,462	0	32,126,462
	HDFS: Number of bytes written	0	32,096,431	32,096,431
	HDFS: Number of large read operations	0	0	0
	HDFS: Number of read operations	12	3	15
	HDFS: Number of write operations	0	2	2
Job Counters	Name	Map	Reduce	Total
	Data-local map tasks	0	0	3
	Killed map tasks	0	0	1
	Launched map tasks	0	0	4
	Launched reduce tasks	0	0	1
	Rack-local map tasks	0	0	1
	Total megabyte-seconds taken by all map tasks	0	0	197,656,576
	Total megabyte-seconds taken by all reduce tasks	0	0	124,215,296
	Total time spent by all map tasks (ms)	0	0	24,128
	Total time spent by all maps in occupied slots (ms)	0	0	48,256
	Total time spent by all reduce tasks (ms)	0	0	15,163
	Total time spent by all reduces in occupied slots (ms)	0	0	30,326
	Total vcore-seconds taken by all map tasks	0	0	24,128
	Total vcore-seconds taken by all reduce tasks	0	0	15,163
	Name	Map	Reduce	Total
	Combine input records	0	0	0
	Combine output records	0	0	0
	CPU time spent (ms)	18,720	15,270	33,990

图 12: p_ij Counters



Logged in as: dr.who

MapReduce Job job_1517750154159_1394

Application
Job
Overview
Counters
Configuration
Map tasks
Reduce tasks
Tools

Job Overview

Job Name:	p_i
User Name:	2017st29
Queue:	root.2017st29
State:	SUCCEEDED
Uberized:	false
Submitted:	Wed Feb 28 23:08:06 CST 2018
Started:	Wed Feb 28 23:02:45 CST 2018
Finished:	Wed Feb 28 23:03:04 CST 2018
Elapsed:	19sec
Diagnostics:	
Average Map Time	9sec
Average Shuffle Time	5mins, 23sec
Average Merge Time	0sec
Average Reduce Time	-5mins, -19sec

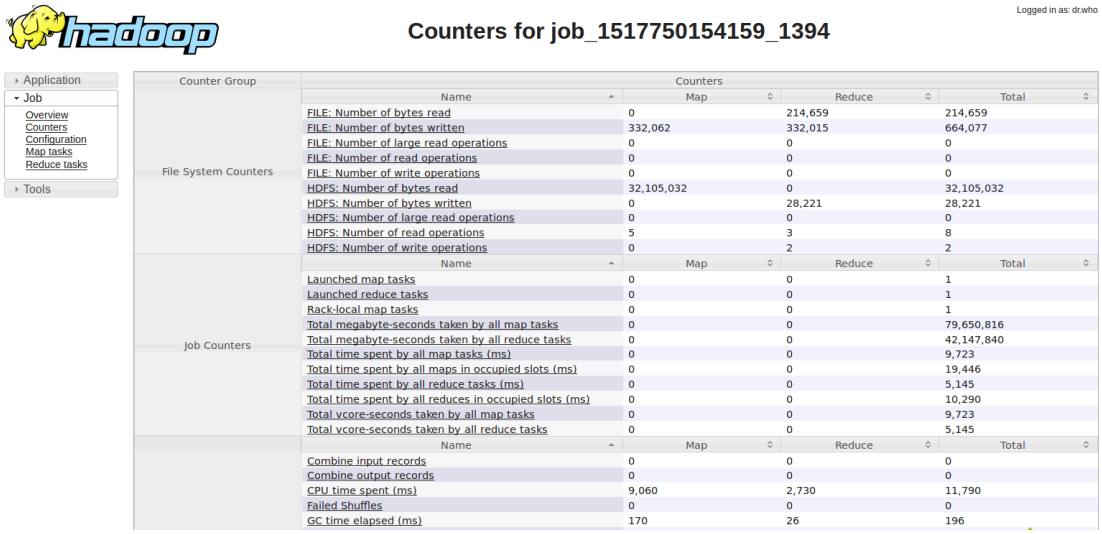
ApplicationMaster

Attempt Number	Start Time	Node	Logs
1	Wed Feb 28 23:02:42 CST 2018	slave007:8042	logs

Task Type	Total	Complete
Map	1	1
Reduce	1	1

Attempt Type	Failed	Killed	Successful
Maps	0	0	1
Reduces	0	0	1

图 13: p_i Job



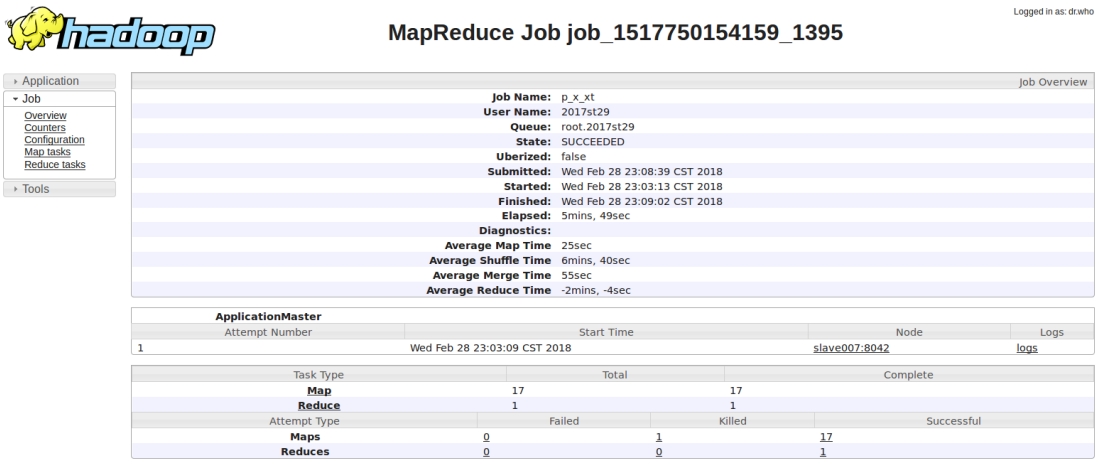
Logged in as: dr.who

Counters for job_1517750154159_1394

- Application
- Job
  - Overview
  - Counters
  - Configuration
  - Map tasks
  - Reduce tasks
- Tools

Counter Group	Name	Map	Reduce	Total
File System Counters	FILE: Number of bytes read	0	214,659	214,659
	FILE: Number of bytes written	332,062	332,015	664,077
	FILE: Number of large read operations	0	0	0
	FILE: Number of read operations	0	0	0
	FILE: Number of write operations	0	0	0
	HDFS: Number of bytes read	32,105,032	0	32,105,032
	HDFS: Number of bytes written	0	28,221	28,221
	HDFS: Number of large read operations	0	0	0
	HDFS: Number of read operations	5	3	8
	HDFS: Number of write operations	0	2	2
Job Counters	Launched map tasks	0	0	1
	Launched reduce tasks	0	0	1
	Rack-local map tasks	0	0	1
	Total megabyte-seconds taken by all map tasks	0	0	79,650,816
	Total megabyte-seconds taken by all reduce tasks	0	0	42,147,840
	Total time spent by all map tasks (ms)	0	0	9,723
	Total time spent by all maps in occupied slots (ms)	0	0	19,446
	Total time spent by all reduce tasks (ms)	0	0	5,145
	Total time spent by all reduces in occupied slots (ms)	0	0	10,290
	Total vcore-seconds taken by all map tasks	0	0	9,723
	Total vcore-seconds taken by all reduce tasks	0	0	5,145
	Combine input records	0	0	0
	Combine output records	0	0	0
	CPU time spent (ms)	9,060	2,730	11,790
	Failed Shuffles	0	0	0
	GC time elapsed (ms)	170	26	196

图 14: p_i Counters



Logged in as: dr.who

MapReduce Job job_1517750154159_1395

- Application
- Job
  - Overview
  - Counters
  - Configuration
  - Map tasks
  - Reduce tasks
- Tools

Job Overview

Job Name:	p_x_xt
User Name:	2017st29
Queue:	root.2017st29
State:	SUCCEEDED
Uberized:	false
Submitted:	Wed Feb 28 23:08:39 CST 2018
Started:	Wed Feb 28 23:03:13 CST 2018
Finished:	Wed Feb 28 23:09:02 CST 2018
Elapsed:	5mins, 49sec
Diagnostics:	
Average Map Time	25sec
Average Shuffle Time	6mins, 40sec
Average Merge Time	55sec
Average Reduce Time	-2mins, -4sec

ApplicationMaster

Attempt Number	Start Time	Node	Logs
1	Wed Feb 28 23:03:09 CST 2018	slave007:8042	logs

Task Type	Total	Complete
Map	17	17
Reduce	1	1

Attempt Type	Failed	Killed	Successful
Maps	0	1	17
Reduces	0	0	1

图 15: p_x_xt Job

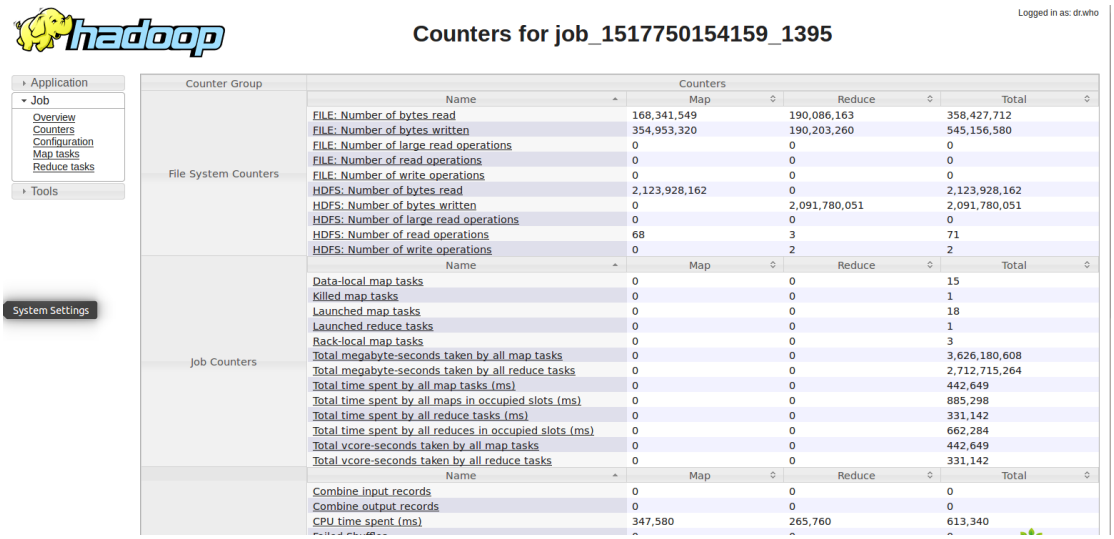


图 16: p_x_xt Counters

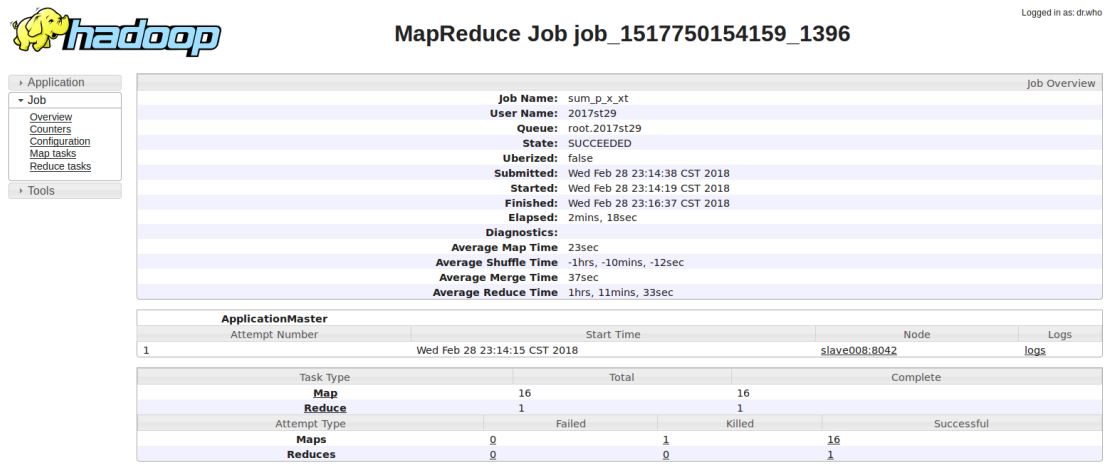
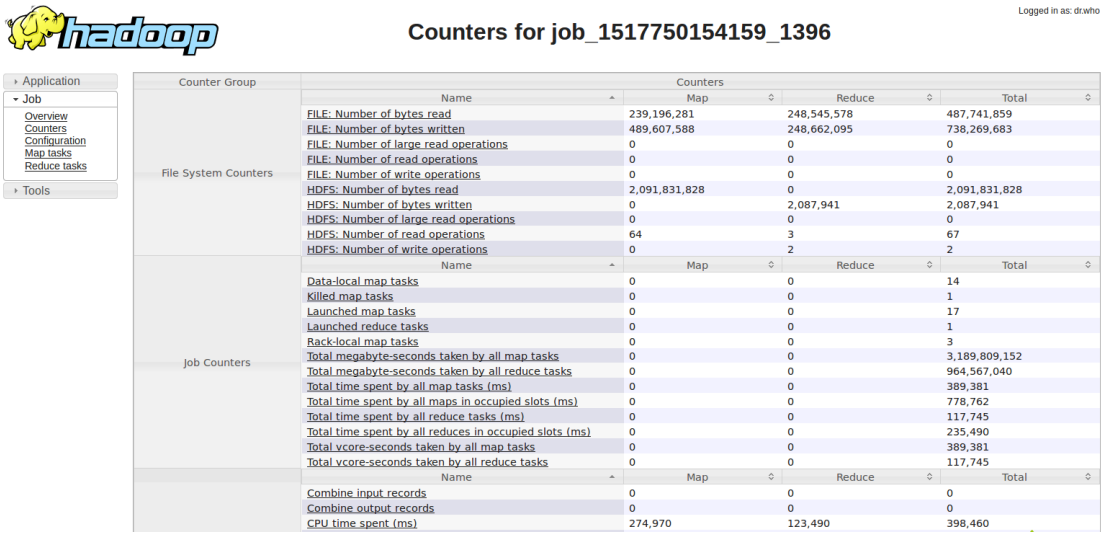


图 17: sum_p_x_xt Job



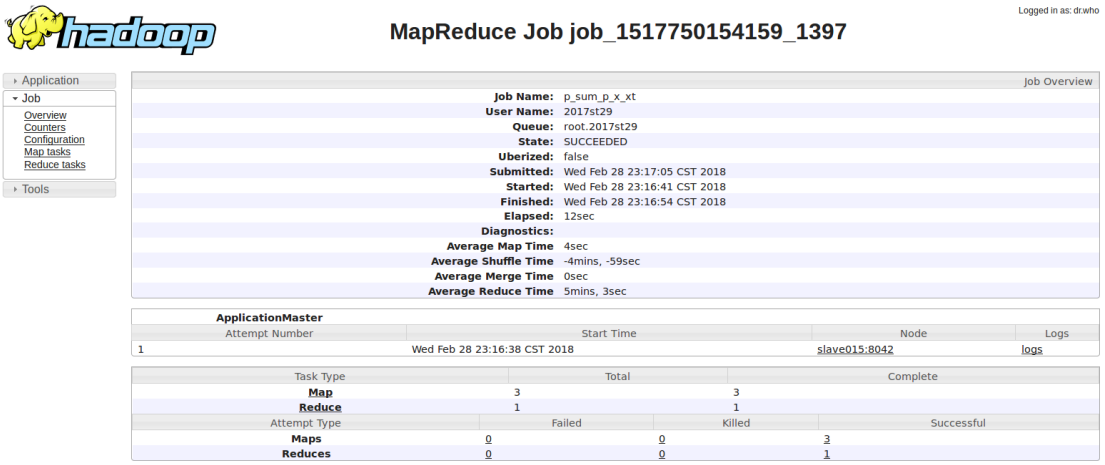
Logged in as: dr.who

Counters for job_1517750154159_1396

- Application
- Job
  - Overview
  - Counters
  - Configuration
  - Map tasks
  - Reduce tasks
- Tools

Counter Group	Counters			
File System Counters	Name	Map	Reduce	Total
	FILE: Number of bytes read	239,196,281	248,545,578	487,741,859
	FILE: Number of bytes written	489,607,588	248,662,095	738,269,683
	FILE: Number of large read operations	0	0	0
	FILE: Number of read operations	0	0	0
	FILE: Number of write operations	0	0	0
	HDFS: Number of bytes read	2,091,831,828	0	2,091,831,828
	HDFS: Number of bytes written	0	2,087,941	2,087,941
	HDFS: Number of large read operations	0	0	0
	HDFS: Number of read operations	64	3	67
	HDFS: Number of write operations	0	2	2
Job Counters	Name	Map	Reduce	Total
	Data-local map tasks	0	0	14
	Killed map tasks	0	0	1
	Launched map tasks	0	0	17
	Launched reduce tasks	0	0	1
	Rack-local map tasks	0	0	3
	Total megabyte-seconds taken by all map tasks	0	0	3,189,809,152
	Total megabyte-seconds taken by all reduce tasks	0	0	964,567,040
	Total time spent by all map tasks (ms)	0	0	389,381
	Total time spent by all maps in occupied slots (ms)	0	0	778,762
	Total time spent by all reduce tasks (ms)	0	0	117,745
	Total time spent by all reduces in occupied slots (ms)	0	0	235,490
	Total vcore-seconds taken by all map tasks	0	0	389,381
	Total vcore-seconds taken by all reduce tasks	0	0	117,745
	Name	Map	Reduce	Total
	Combine input records	0	0	0
	Combine output records	0	0	0
	CPU time spent (ms)	274,970	123,490	398,460

图 18: sum_p_x_xt Counters



Logged in as: dr.who

MapReduce Job job_1517750154159_1397

- Application
- Job
  - Overview
  - Counters
  - Configuration
  - Map tasks
  - Reduce tasks
- Tools

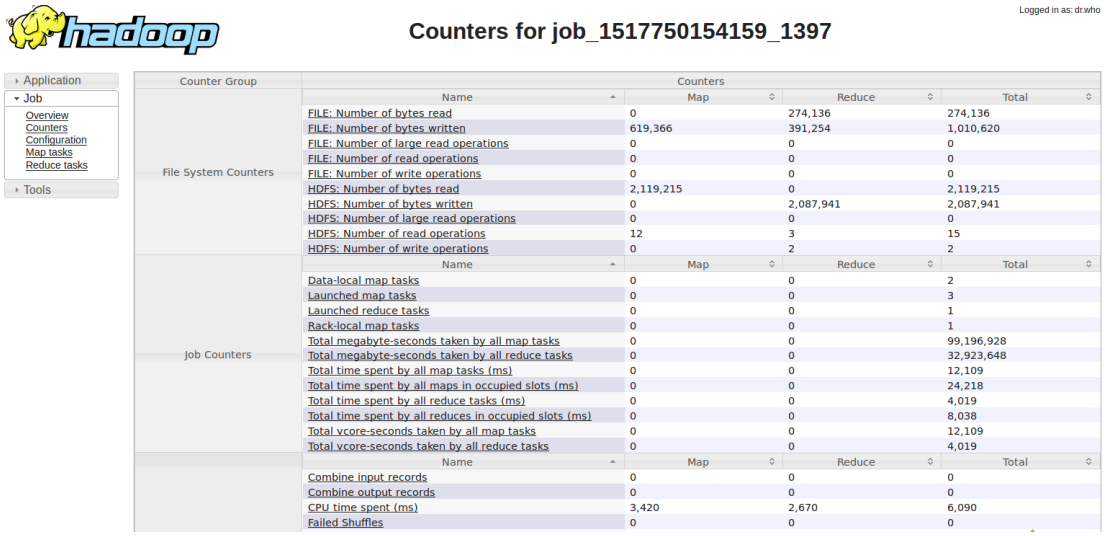
Job Overview	
Job Name:	p_sum_p_x_xt
User Name:	2017st29
Queue:	root.2017st29
State:	SUCCEEDED
Uberized:	false
Submitted:	Wed Feb 28 23:17:05 CST 2018
Started:	Wed Feb 28 23:16:41 CST 2018
Finished:	Wed Feb 28 23:16:54 CST 2018
Elapsed:	12sec
Diagnostics:	
Average Map Time	4sec
Average Shuffle Time	-4mins, -59sec
Average Merge Time	0sec
Average Reduce Time	5mins, 3sec

ApplicationMaster			
Attempt Number	Start Time	Node	Logs
1	Wed Feb 28 23:16:38 CST 2018	slave015:8042	logs

Task Type	Total	Complete
Map	3	3
Reduce	1	1

Attempt Type	Failed	Killed	Successful
Maps	0	0	3
Reduces	0	0	1

图 19: p_sum_p_x_xt Job



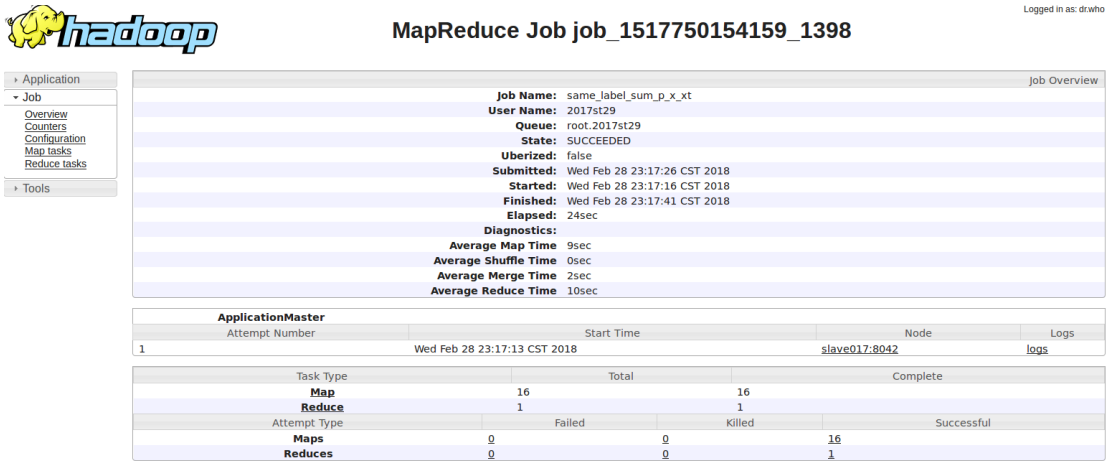
Logged in as: dr.who

Counters for job_1517750154159_1397

- Application
- Job
  - Overview
  - Counters
  - Configuration
  - Map tasks
  - Reduce tasks
- Tools

Counter Group	Name	Map	Reduce	Total
File System Counters	FILE: Number of bytes read	0	274,136	274,136
	FILE: Number of bytes written	619,366	391,254	1,010,620
	FILE: Number of large read operations	0	0	0
	FILE: Number of read operations	0	0	0
	FILE: Number of write operations	0	0	0
	HDFS: Number of bytes read	2,119,215	0	2,119,215
	HDFS: Number of bytes written	0	2,087,941	2,087,941
	HDFS: Number of large read operations	0	0	0
	HDFS: Number of read operations	12	3	15
	HDFS: Number of write operations	0	2	2
Job Counters	Data-local map tasks	0	0	2
	Launched map tasks	0	0	3
	Launched reduce tasks	0	0	1
	Rack-local map tasks	0	0	1
	Total megabyte-seconds taken by all map tasks	0	0	99,196,928
	Total megabyte-seconds taken by all reduce tasks	0	0	32,923,648
	Total time spent by all map tasks (ms)	0	0	12,109
	Total time spent by all maps in occupied slots (ms)	0	0	24,218
	Total time spent by all reduce tasks (ms)	0	0	4,019
	Total time spent by all reduces in occupied slots (ms)	0	0	8,038
	Total vcore-seconds taken by all map tasks	0	0	12,109
	Total vcore-seconds taken by all reduce tasks	0	0	4,019
	Combine input records	0	0	0
	Combine output records	0	0	0
	CPU time spent (ms)	3,420	2,670	6,090
	Failed Shuffles	0	0	0

图 20: p__sum__p__x__xt Counters



Logged in as: dr.who

MapReduce Job job_1517750154159_1398

- Application
- Job
  - Overview
  - Counters
  - Configuration
  - Map tasks
  - Reduce tasks
- Tools

Job Overview

Job Name:	same_label_sum_p_x_xt
User Name:	2017st29
Queue:	root.2017st29
State:	SUCCEEDED
Uberized:	false
Submitted:	Wed Feb 28 23:17:26 CST 2018
Started:	Wed Feb 28 23:17:16 CST 2018
Finished:	Wed Feb 28 23:17:41 CST 2018
Elapsed:	24sec
Diagnostics:	
Average Map Time	9sec
Average Shuffle Time	0sec
Average Merge Time	2sec
Average Reduce Time	10sec

ApplicationMaster

Attempt Number	Start Time	Node	Logs
1	Wed Feb 28 23:17:13 CST 2018	slave017:8042	logs

Task Type	Total	Complete
Map	16	16
Reduce	1	1

Attempt Type	Failed	Killed	Successful
Maps	0	0	16
Reduces	0	0	1

图 21: same__label__sum__p__x__xt Job

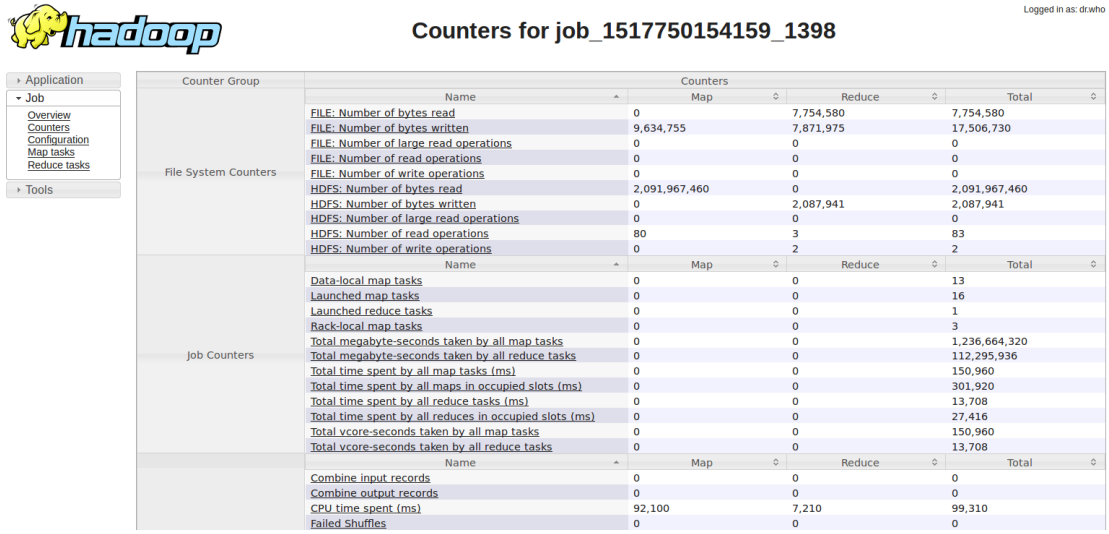


图 22: same_label_sum_p_x_xt Counters

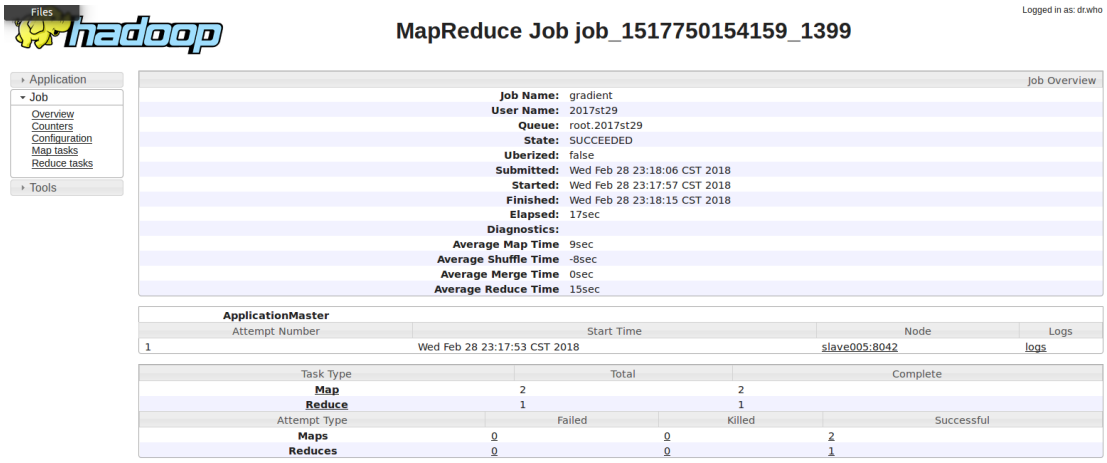
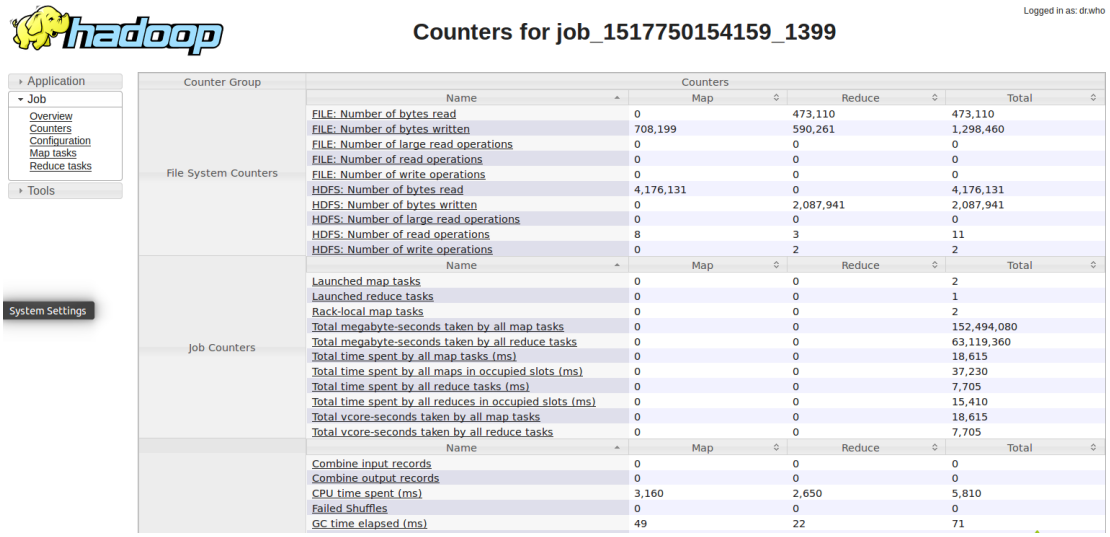


图 23: gradient Job



Logged in as: dr.who

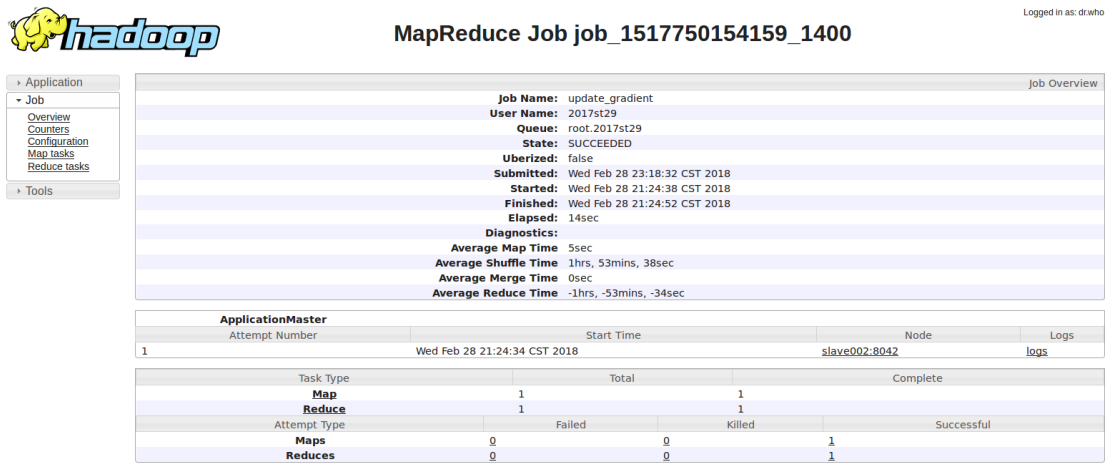
Counters for job_1517750154159_1399

- Application
- Job
  - Overview
  - Counters
  - Configuration
  - Map tasks
  - Reduce tasks
- Tools

System Settings

Counter Group	Counters			
	Name	Map	Reduce	Total
File System Counters	FILE: Number of bytes read	0	473,110	473,110
	FILE: Number of bytes written	708,199	590,261	1,298,460
	FILE: Number of large read operations	0	0	0
	FILE: Number of read operations	0	0	0
	FILE: Number of write operations	0	0	0
	HDFS: Number of bytes read	4,176,131	0	4,176,131
	HDFS: Number of bytes written	0	2,087,941	2,087,941
	HDFS: Number of large read operations	0	0	0
	HDFS: Number of read operations	8	3	11
	HDFS: Number of write operations	0	2	2
Job Counters	Name	Map	Reduce	Total
	Launched map tasks	0	0	2
	Launched reduce tasks	0	0	1
	Rack-local map tasks	0	0	2
	Total megabyte-seconds taken by all map tasks	0	0	152,494,080
	Total megabyte-seconds taken by all reduce tasks	0	0	63,119,360
	Total time spent by all map tasks (ms)	0	0	18,615
	Total time spent by all maps in occupied slots (ms)	0	0	37,230
	Total time spent by all reduce tasks (ms)	0	0	7,705
	Total time spent by all reduces in occupied slots (ms)	0	0	15,410
	Total vcore-seconds taken by all map tasks	0	0	18,615
	Total vcore-seconds taken by all reduce tasks	0	0	7,705
	Name	Map	Reduce	Total
	Combine input records	0	0	0
	Combine output records	0	0	0
	CPU time spent (ms)	3,160	2,650	5,810
	Failed Shuffles	0	0	0
	GC time elapsed (ms)	49	22	71

图 24: gradient Counters



Logged in as: dr.who

MapReduce Job job_1517750154159_1400

- Application
- Job
  - Overview
  - Counters
  - Configuration
  - Map tasks
  - Reduce tasks
- Tools

Job Overview	
Job Name:	update_gradient
User Name:	2017st29
Queue:	root.2017st29
State:	SUCCEEDED
Uberized:	false
Submitted:	Wed Feb 28 23:18:32 CST 2018
Started:	Wed Feb 28 21:24:38 CST 2018
Finished:	Wed Feb 28 21:24:52 CST 2018
Elapsed:	14sec
Diagnostics:	
Average Map Time	5sec
Average Shuffle Time	1hrs, 53mins, 38sec
Average Merge Time	0sec
Average Reduce Time	-1hrs, -53mins, -34sec

ApplicationMaster

Attempt Number	Start Time	Node	Logs
1	Wed Feb 28 21:24:34 CST 2018	slave002:8042	logs

Task Type	Total	Complete
Map	1	1
Reduce	1	1

Attempt Type	Failed	Killed	Successful
Maps	0	0	1
Reduces	0	0	1

图 25: update_gradient Job

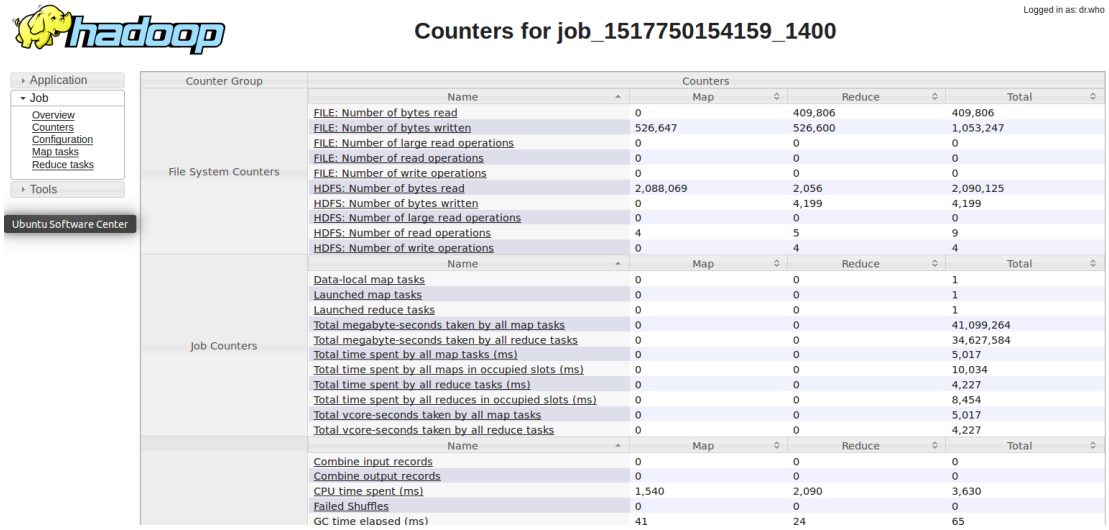


图 26: update_gradient Counters

总结

通过本次实验，我们尝试解决在机器学习课程中遇到的问题：在数据量较大和数据维度较高的情况下，如何高效地进行梯度下降计算。这一目标可进一步归结为如何将 Tensor 运算并行化。在本次实验中，我们将 Tensor 进一步拆分为矩阵的列表，并使用 MapReduce 设计了不同的矩阵运算算子来处理 NCA 中需要用到的矩阵运算。通过将数据分布到集群上处理，我们可以执行单机难以处理的计算，并利用并行化的优势加快其执行速度。

我们使用 Apache 的 commons-math3 作为矩阵运算库，但是它仅仅是针对简单的矩阵运算而设计的，并且没有对运算速度作太多的优化，可能在单个矩阵操作的运算速度上不及 numpy 之类；但是我们可以很方便地用高效的矩阵运算库来替代当前的实现，从而获得性能上的提升。

在实现 NCA 的过程中，我们其实构建了一个简单的并行矩阵处理的框架。在以后如果遇到类似的计算问题，我们可以在此基础上通过实现新的矩阵运算算子和对这些算子进行组合来解决。这确实为我们解决问题提供了新的思路。

参考文献

[1] Sam Roweis, Geoffrey Hinton, and Ruslan Salakhutdinov. Neighbourhood component analysis. *Adv. Neural Inf. Process. Syst.(NIPS)*, 17:513–520, 2004.

[2] 周志华. 机器学习. Qing hua da xue chu ban she, 2016.

[3] 黄宜华 and 苗凯翔. 深入理解大数据: 大数据处理与编程实践, 2014.