基于 Spark 平台的 NetFlow 流量分析系统

丁圣勇,闵世武,樊勇兵

(中国电信股份有限公司广东研究院 广州 510630)

摘 要:目前典型的 NetFlow 分析系统多为基于私有架构或平台的第三方系统,面临扩展性较低、开放性不足、扩容代价大、分析时延长等问题。大数据技术的快速发展尤其是内存式计算平台如 Spark 的出现为集中处理大规模 NetFlow 数据提供了可能,本文提出了基于 Spark 的 NetFlow 分析系统,验证了核心算法(如流量应用构成统计)在 Spark 平台的性能。实验表明,基于 Spark 的 NetFlow 分析系统具有很高的性能和很强的扩展能力,较之 Hadoop MapReduce 有显著的性能提升。

关键词:NetFlow;Spark;流量分析

doi: 10.3969/j.issn.1000-0801.2014.10.009

A Large Scale NetFlow Analysis System Based on Spark

Ding Shengyong, Min Shiwu, Fan Yongbing

(Guangdong Research Institute of China Telecom Co., Ltd., Guangzhou 510630, China)

Abstract: The existing systems usually adopt private distributed architectures, which face scalability, openness, cost and latency problems. The development of big data technology such as Spark offers new opportunity for large scale NetFlow processing systems. A new analysis system based on Spark platform was proposed and the effectiveness of the method was verified. The experimental results show its superior performance.

Key words: NetFlow, Spark, traffic analysis

1 引言

NetFlow 是由 Cisco(思科)公司在 1996 年开发的内置于 Cisco IOS 的一种网络协议,目的是收集 IP 流量信息和监控网络的使用情况。NetFlow 被广泛应用于 Cisco 的路由器和交换机中,类似技术也得到 Juniper、华为等公司生产的路由器的支持,是网络规划、运营和优化的重要依据。从路由器中发送出来的 NetFlow 记录核心属性包括源地址、目标地址、IP 地址类型、源端口号、目标端口号、报文大小,本文中的 NetFlow 泛指能够支持类似功能的网络协议。

对于一个骨干网络,为了统计流量模式,需要在所有边缘路由器入口开启 NetFlow 采集,导致 NetFlow 系统每天需要处理数十 TB 甚至上百 TB 的原始记录,并且随着网络规模的不断扩张,系统需要具有接近线性的扩展能力。目前大型网络一般使用第三方流量分析系统,这类系统多采用私有分布式架构支持大规模处理,通过将采集器和分析器分离实现容量扩展。由于价格昂贵,这类系统在网络持续扩容时面临较大的成本压力,此外,开放性也相对较低,客户很难实现特定的分析需求。

另一方面,伴随着大数据技术的快速发展,新的分布

式平台为处理大规模 NetFlow 数据提供了契机，尤其是 Spark 技术的出现，大规模数据处理在通用服务器集群上能够达到准实时的性能。本文提出了一种基于 Spark 平台的大规模 NetFlow 处理系统，通过实验表明，该系统具有很强的扩展能力，性能显著优于 Hadoop MapReduce 计算平台。

2 Spark 技术介绍

Apache Spark 是由 UC Berkeley 开源的类 MapReduce 新一代大数据分析框架，拥有 Hadoop MapReduce 的所有优点，与 MapReduce 不同的是，Spark 将计算的中间结果数据持久地存储在内存中，通过减少磁盘 I/O，使后续的数据运算效率更高。Spark 的这种架构设计尤其适合于机器学习、交互式数据分析等应用，这些应用都需要重复地利用计算的中间数据。在 Spark 和 Hadoop 的性能基准测试对比中，运行基于内存的 logistic regression，在迭代次数相同的情况下，Spark 的性能超出 Hadoop MapReduce 100 倍以上。

Spark 在设计上参考了 Hadoop MapReduce，完全和 Hadoop 生态系统相兼容，如 Spark 底层的数据持久化部分就完全重用了 Hadoop 的文件系统，Spark 也可以运行在 Hadoop Yarn 资源管理系统上。从大数据生态系统的发展来看，Spark 的出现不是对以 Hadoop 为中心的大数据生态系统的取代，而是补充，更多的是深耕于 MapReduce 不适用的应用领域（如机器学习、流式计算、实时计算、交互式数据挖掘等）。但 Spark 又不局限于 MapReduce 简单的编程范式，Spark 立足于内存计算，同时在上层支持图计算、迭代式计算、流式计算、内存 SQL 等多种计算范式，因此相对于 MapReduce 更具有通用性。

为了支持在多次迭代计算过程中重复利用内存数据集，Spark 在借鉴传统分布式共享内存思想的基础上，提出了一种新的数据抽象模型 RDD（resilient distributed dataset），RDD 是只读、支持容错、可分区的内存分布式数据集，可以一部分或者全部缓存在集群内存中，以便在多次计算过程中重用。用户可以显式控制 RDD 的分区、物化、缓存策略等，同时 RDD 提供了一套丰富的编程接口，供用户操作。RDD 是 Spark 分布式计算的核心，Spark 的所有计算模式都必须围绕 RDD 进行。

3 基于 Spark 的 NetFlow 流量分析

NetFlow 作为一种通用的数据报文格式，是典型的结构化数据，随着运营商网络的扩容与升级，NetFlow 数据的生成速率和数据规模都出现了大规模的增长，NetFlow 的数据分析是天然的大数据处理。典型的 NetFlow 解决方案是使流量采集系统和流量分析系统相分离，本文只讨论流量分析系统。

一种大规模的处理方法是使用 MapReduce（以下简称 MR）方案。在该方案中，当需要对 NetFlow 数据进行多维度、多次数的统计时，需要编写多个 MR 任务，这些任务被分别提交到集群上，以串行或者并行方式执行，任务之间无法共享内存数据，数据需要反复在内存和磁盘之间转移，导致 MR 分析任务有性能低、分析时延长、内存占用大等缺点。

Spark 的出现有效地解决了 MR 执行过程中，中间数据不能缓存在内存中的问题。在基于 Spark 的分析系统中，NetFlow 流量数据只需要从磁盘加载到内存一次，即可在该缓存数据上进行多维度、多次数的分析和查询，通过减少磁盘 I/O，提高了分析性能，降低了分析时延，特别适合于 NetFlow 这种生成速率快、数据规模大的应用场景。

4 基于 Spark 的 NetFlow 流量分析系统架构

基于 Spark 的 NetFlow 数据分析数据流图如图 1 所示。NetFlow 以文本记录的形式保存在 HDFS 上，Spark 计算引擎调用 textFile 方法从 HDFS 上加载 NetFlow 数据到集群内存中，并将 NetFlow 数据转换为 HadoopRDD[string] 的形式。然后即可调用 RDD 上的编程接口如 map、filter、reduce、join 等，对 NetFlow 数据进行多维度统计分析。由于 RDD 的只读特性，每次对 RDD 的操作都会生成新的 RDD，整个分析流程便形成如图 1 中所示的“管道”。计算过程中，可以根据需要对任意 RDD 通过调用 persist 的方法将该 RDD 缓存在集群内存中，以便后续基于该 RDD 的分析效率更高。对于不再需要的 RDD，调用 unpersist 方法即可将该 RDD 从集群内存中清除，释放该 RDD 占用的内存空间。分析完成后，对于包含结果数据的 RDD 调用 saveAsTextFile 方法可将结果 RDD 中的数据持久化地存储到 HDFS 上。

上述 NetFlow 分析任务在 Spark 集群上运行时的图解如图 2 所示。整个分析任务由一个全局的用户 driver 程序、机器主节点上的 master 和若干集群从节点上的 worker 共同组成。提交到集群的 driver 程序和 master 进行通信，

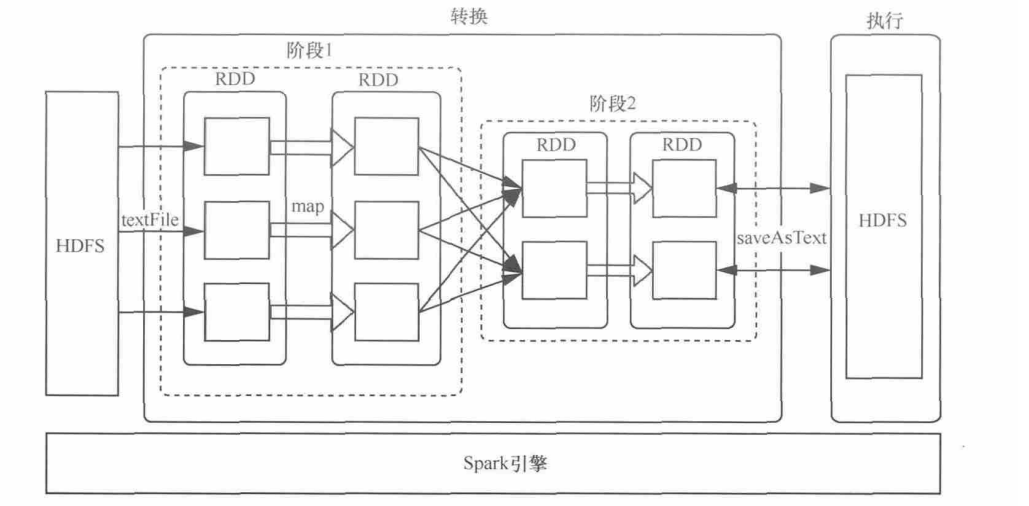

图 1 NetFlow 分析流图

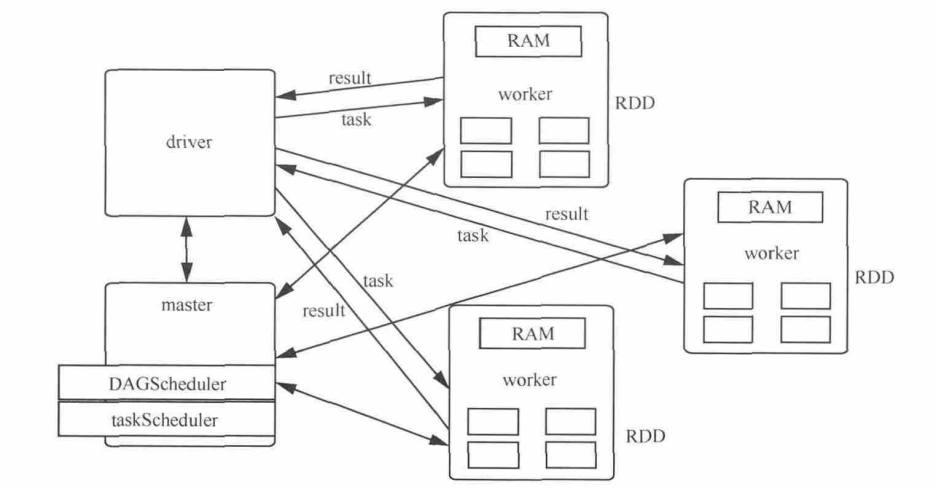

图 2 NetFlow 分析运行图

master 内的 DAGScheduler 根据分析任务的具体情况对 RDD 上的计算过程划分阶段,如图 1 中分成了阶段 1 和阶段2。划分完阶段后,master 内的 taskScheduler 将这些阶段包装成 task 的形式调度到 worker 上运行,worker 在运行任务时需要和 driver 保持通信,以进行一些数据汇总的操作。在连续的多个分析任务运行过程中,RDD 始终保持在 worker 的内存中,因此分析任务的执行速度很快、效率更高。

5 实验验证

本实验以基于 Hadoop 的应用构成统计为例,对比 Spark 和 MapReduce 在实际运行过程中的性能。应用构成统计根据应用的目标端口号对流量数据进行分类,统计不同类型应用流量的大小和占比情况。分别使用 MR 的编程接口和 Spark 的编程接口编写统计分析作业,并将作业代码打包上传到集群上执行。集群配置情况见表 1。在集群

表 1 集群配置

服务器类型	服务器数量	内存	CPU 核数	网卡速率	存储
刀片型	8 台	8×128 GB	8×16 个	1 Gbit/s	8×1 TB

上同时搭建 Hadoop 2.2 和 Spark 1.0 分布式环境，选择其中一台服务器作为 Hadoop 和 Spark 的主节点，其他 7 台服务器作为从节点。

通过调整输入数据量的规模大小，测试 MR 和 Spark 作业的完成时间，性能对比结果如图 3 所示。从图 3 中可以看到，在输入数据规模相同的情况下，Spark 作业的完成时间比 MR 更少，而且随着数据量的增大，Spark 的性能超出了 Hadoop 2 倍以上。通过分析 Spark 和 MR 运行原理，可知 Spark 将 HDFS 上的数据抽象成 RDD，并在计算过程中缓存在计算节点的内存中，加快了计算速度。同时 Spark 由于 RDD 的引入，每个节点上的计算任务以多线程的方式执行，相对于 MapReduce 为每个任务启动单独的 Java 虚拟机，效率更高。因此可知，在同等条件下 Spark 的计算性能要明显优于 MapReduce。

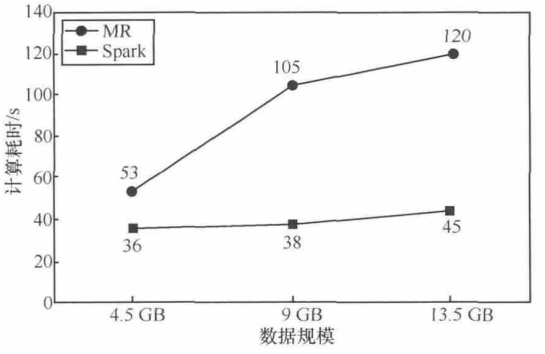

图 3　Spark 和 MapReduce 性能对比

6　结束语

本文在研究 Spark 大数据平台基本原理的基础上，结合 NetFlow 数据规模大、生成速度快等特征，提出了使用 Spark 大数据平台对 NetFlow 数据进行统计分析的基本方法，并在具体的实验中对比了 Spark 和 MapReduce 在 NetFlow 数据处理上的性能差异。验证了 Spark 在 NetFlow 数据分析方面相对 MapReduce，处理速度更快，效率更高。

参考文献

1　White T. Hadoop: the Definitive Guide. O'Reilly Media Inc, 2012

2　Zaharia M, Chowdhury M, Das T, *et al*. Resilient distributed datasets: a fault-tolerant abstraction for in-memory cluster computing. Proceedings of the 9th USENIX Conference on Networked Systems Design and Implementation, San Jose, CA, USA, 2012

3　Rossi D, Silvio V. Fine-grained traffic classification with NetFlow data. Proceedings of the 6th International Wireless Communications and Mobile Computing Conference, Shenzhen, China, 2010

[作者简介]

丁圣勇，男，现就职于中国电信股份有限公司广东研究院，长期从事网络运营数据分析与挖掘工作，主要研究方向为网络大数据。

闵世武，男，现就职于中国电信股份有限公司广东研究院，主要从事 Hadoop、Spark 相关的大数据开发工作。

樊勇兵，男，博士，现就职于中国电信股份有限公司广东研究院，主要从事云计算及大数据相关研究工作。

（收稿日期：2014-09-15）

·简讯·

2014 年 8 月电话用户分省情况

2014 年 8 月份统计数据显示，全国固定电话用户数为 25 511.2 万户，其中城市电话用户数为 17 873.5 万户，农村电话用户数为 7 637.7 万户；全国移动电话用户数为 126 698.5 万户。

全国的电话用户主要分布在东部、中部和西部地区。东部地区固定电话用户数为 13 874.7 万户，其中城市电话用户数为 9 594.5 万户，农村电话用户数为 4 280.2万户；东部地区移动电话用户数为 62 585.8 万户。中部地区固定电话用户数为 6 206.2 万户，其中城市电话用户数为 4 241.7 万户，农村电话用户数为 1 964.5 万户；中部地区移动电话用户数为 33 063.0 万户。西部地区固定电话用户数为 5 430.4 万户，其中城市电话用户数为 4 037.3 万户，农村电话用户数为 1 393.1 万户；西部地区移动电话用户数为 31 049.0 万户。

（来源：工业和信息化部运行监测协调局）