

# 大数据流式计算： 应用特征和技术挑战

孙大为

中国地质大学信息工程学院 北京 100083

## 摘要

在大数据时代,数据的时效性日益突出,数据的流式特征更加明显,越来越多的应用场景需要部署在流式计算平台中。大数据流式计算作为大数据计算的一种形态,其重要性也不断提升。针对大数据环境中流式计算应用所呈现出的诸多鲜明特征进行了系统化的分析,并从系统架构的角度,给出了大数据流式计算系统构建的原则性策略。结合当前比较典型的流式计算平台,重点研究了当前大数据流式计算在在线环境下的资源调度和节点依赖环境下的容错策略等方面的技术挑战。

## 关键词

大数据;流式计算;应用特征;在线调度;系统容错

doi: 10.11959/j.issn.2096-0271.2015032

## *Big Data Stream Computing: Features and Challenges*

Sun Dawei

School of Information Engineering, China University of Geosciences, Beijing 100083, China

## *Abstract*

In big data era, the timeliness of data has become one of the most important factors, and the streaming feature of data has become more obvious. More and more applications need to be deployed in stream computing platforms. Big data stream computing as a major form of big data computing has become more and more important. The features of big data stream computing application were systematically analyzed. The principle strategies to build a big data stream computing system were given from the perspective of system architecture. Combined with some typical big data stream computing systems, some technology challenges in big data stream computing environments were focused, such as resource scheduling in online environments, fault tolerance strategy in node-dependence environments.

## *Key words*

big data, stream computing, application feature, online scheduling, system fault tolerance

## 1 引言

云计算、物联网等新兴信息技术和应用模式的快速发展,推动人类社会迈入大数据新时代<sup>[1~3]</sup>。一般意义上,大数据是指利用现有理论、方法、技术和工具难以在可接受的时间内完成分析计算、整体呈现高价值的海量复杂数据集合。大数据蕴含大信息,大信息提炼大知识,大知识将在更高的层面、以更广的视角、在更大的范围内帮助用户提高洞察力、提升决策力,为人类社会创造前所未有的大价值。但与此同时,这些总量极大的价值往往隐藏在大数据中,表现出了价值密度极低、分布极其不规律、信息隐藏程度极深、发现有用价值极其困难等鲜明特性,这些特征必然为大数据的计算带来前所未有的挑战和机遇。

大数据的计算模式<sup>[2~5]</sup>主要分为批量计算(batch computing)、流式计算(stream computing)、交互计算(interactive computing)、图计算(graph computing)等。其中,流式计算和批量计算是两种主要的大数据计算模式,分别适用于不同的大数据应用场景。对于先存储后计算,实时性要求不高,同时数据的准确性、全面性更为重要的应用场景,批量计算更加适合;对于无需先存储,可以直接进行数据计算,实时性要求很严格,但数据的精确度往往不太苛刻的应用场景,流式计算具有明显优势。流式计算中,数据往往是最近一个时间窗口内的增量数据,因此数据时延往往较短,实时性较强,但数据的信息量往往相对较少,只限于一个时间窗口内的信息,不具有全量信息。流式计算和批量计算具有明显的优劣互补特征,在多种应用场合下可以将两者结合起来使用,通过发挥流式计算

的实时性优势和批量计算的计算精度优势,满足多种应用场景在不同阶段的数据计算要求。

在大数据时代,数据的时效性日益突出,数据的流式特征更加明显,越来越多的应用场景需要部署在流式计算平台中。大数据流式计算作为大数据计算的一种形态,其重要性也在不断提升。大数据时代的流式计算呈现出了鲜明的高带宽、低时延的应用需求,传统的流式计算平台的构建往往是以数据库为基础,且数据规模较小,数据对象较单一,无法满足大数据流式计算需求。如何构建低时延、高带宽、持续可靠、长期运行的大数据流式计算系统是当前亟待解决的问题。本文针对大数据环境中,流式计算应用所呈现出的诸多鲜明特征进行了系统化的分析,并从系统架构的角度,给出了大数据流式计算系统构建的原则性策略。结合当前比较典型的流式计算平台,重点研究了当前大数据流式计算在在线环境下的资源调度和节点依赖环境下的容错策略等方面的技术挑战。

## 2 流式应用分析

大数据流式计算有着悠久的历史,早在20世纪80年代就有部署和应用,在今天的大数据时代,其应用的领域和范围在不断地扩大,也呈现出了诸多新的特征和要求<sup>[6~9]</sup>。本节将从大数据流式计算的典型应用场景出发,系统地分析大数据环境中流式应用所呈现出来的诸多鲜明特征,并从系统架构的角度,给出大数据流式计算系统建设的原则性策略。

### 2.1 应用及特征

大数据流式计算可以广泛应用于金融

银行、互联网、物联网等诸多领域，如股市实时分析、插入式广告投放、交通流量实时预警等场景，主要是为了满足该场景下的实时应用需求。数据往往以数据流的形式持续到达数据计算系统，计算功能的实现是通过有向任务图的形式进行描述，数据流在有向任务图中流过后，会实时产生相应的计算结果。整个数据流的处理过程往往是在毫秒级的时间内完成的。

通常情况下，大数据流式计算场景具有以下鲜明特征。

- 在流式计算环境中，数据是以元组为单位，以连续数据流的形态，持续地到达大数据流式计算平台。数据并不是一次全部可用，不能够一次得到全量数据，只能在不同的时间点，以增量的方式，逐步得到相应数据。

- 数据源往往是多个，在进行数据流重放的过程中，数据流中各个元组间的相对顺序是不能控制的。也就是说，在数据流重放过程中，得到完全相同的数据流（相同的数据元组和相同的元组顺序）是很困难的，甚至是不可能的。

- 数据流的流速是高速的，且随着时间在不断动态变化。这种变化主要体现在两个方面，一个方面是数据流流速大小在不同时间点的变化，这就需要系统可以弹性、动态地适应数据流的变化，实现系统中资源、能耗的高效利用；另一方面是数据流中各个元组内容（语义）在不同时间点的变化，即概念漂移，这就需要处理数据流的有向任务图可以及时识别、动态更新和有效适应这种语义层面上的变化。

- 实时分析和处理数据流是至关重要的，在数据流中，其生命周期的时效性往往很短，数据的时间价值也更加重要。所有数据流到来后，均需要实时处理，并实时产生相应结果，进行反馈，所有的数据元组也仅会被处理一次。虽然部分数据可能以

批量的形式被存储下来，但也只是为了满足后续其他场景下的应用需求。

- 数据流是无穷无尽的，只要有数据源在不断产生数据，数据流就会持续不断地到来。这也就需要流式计算系统永远在线运行，时刻准备接收和处理到来的数据流。在线运行是流式计算系统的一个常态，一旦系统上线后，所有对该系统的调整和优化也将在在线环境中开展和完成。

- 多个不同应用会通过各自的有向任务图进行表示，并将被部署在一个大数据计算平台中，如图1所示，这就需要整个计算平台可以有效地为各个有向任务图分配合理资源，并保证满足用户服务级目标。同时各个资源间需要公平地竞争资源、合理地共享资源，特别是要满足不同时间点各应用间系统资源的公平使用。

## 2.2 流式计算系统构建原则

在大数据流式计算系统架构方面，由于大数据流式计算、批量计算、图计算

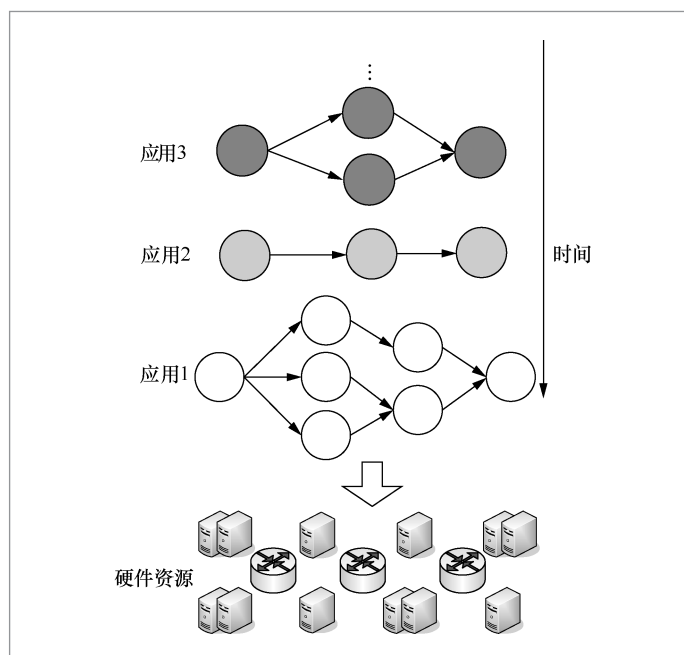


图1 大数据流式应用部署

等模式间存在显著不同,在一个计算平台中,试图将大数据批量计算、流式计算、图计算等多种不同计算形态集成起来,往往是很困难的。其中,也有些平台试图做这样的工作,如在大数据批量计算平台Spark中,加入了流式计算功能,形成了Spark Streaming子系统,其实现的关键是将Spark中任务处理步长尽可能地缩短,从而降低数据处理的时延。但Spark Streaming也只能达到亚秒级的用户响应,并不能真正满足流式计算需要。此外,YARN等平台也试图对不同的计算形态进行集成,由于这些计算形态差异很大,应用场景截然不同,需要提供的体系结构、资源管理等各方面也不完全一样,所以必须针对具体的计算场景,开展针对性的计算平台的研究,才能更好地适应实际需要。

大数据流式计算系统的体系结构需要重点考虑以下因素。

#### (1) 分布式体系结构

集中式体系结构对于大数据流式计算系统往往很难满足其可扩展性的需要,人们往往会优先采用分布式体系结构来构建流式计算系统,这样可以可扩展地适应数据流压力的动态变化,灵活地进行系统的伸展和收缩,实现对相关资源和任务在在线环境中的调整和优化。

#### (2) 内存计算是首要考虑因素

当一个有向任务图被提交到系统中后,该有向任务图将常驻内存,并将永远在线运行(除非被显示终止或系统崩溃),所有数据流到达系统后,也将直接在内存中完成相关计算,并实时产生和输出计算结果。部分数据可能会被选择性地存储在外存介质上,显然内存成为这个数据计算过程中的主要场所和重要位置。因此,在系统结构的设计过程中,需要合理、高效地利用内存资源。

#### (3) 时效性是系统设计的首要目标

大数据流式计算环境中,数据的时间价值是首要的,当数据流到来后,必须在毫秒级的时间内完成对数据流中相应知识的发现,这个过程中时效性是首要的,准确性是次要的,部分、及时、相对准确的计算结果是好于全面、延迟、精确的计算结果的。根据具体场景的需要,数据流可以被选择性地存储起来,后续可以进行批量计算,更为准确、全面地发现其中的知识,实现大数据流式计算和批量计算间的优势互补。

#### (4) 在线运行环境将是系统的常态

所有关于系统的调整和优化将会在在线环境中开展,这种情况下,需要将系统的优化性能、稳定性、波动等因素综合考虑进来,最优的方案在在线环境中进行调整未必是最合理的,当前这一时刻对于系统而言,可能是最优的,但是要达到这种最优状态需要调整过多的系统资源,可能会对系统的稳定性产生影响。而且,在下一时刻,数据流压力发生变化后,这种当前的最优方案可能又变为非最优状态了。因此,对系统性能有所改善的方案可能会达到更佳效果,要综合权衡系统性能、稳定性、动态在线变化等多个方面的因素。

## 3 技术挑战

大数据流式计算系统存在诸多挑战<sup>[9-13]</sup>,如资源调度、系统容错、动态时间窗口、高效索引策略等诸多方面。本节将从大数据流式计算系统架构的角度,针对当前大数据流式计算环境中存在的两个方面的典型问题进行系统化的分析,即在线环境下的资源调度问题和节点依赖环境下的容错策略问题,并原则性地分别给出了两类问题的解决策略。

### 3.1 在线环境下的资源调度

资源调度是分布式系统中资源管理的关键与核心,也是NP难问题,制约着整个系统的高效运行。在大数据流式计算环境中,在线环境下的资源调度又更加困难,任何一个资源或要素的调整,都会对运行着的系统产生实时影响,也会对整个系统的稳定性带来一定程度上的波动。Storm系统作为业界最具影响力的大数据流式计算系统,目前其所选用的资源调度策略为轮询方式,只是简单地将有向任务图中各个节点按照一定的拓扑序列放置到各个物理机器上去,这个调度策略没有考虑物理机器的性能以及物理机器间的拓扑结构,没有考虑有向任务图中各个节点的计算压力和节点间的通信压力。在在线调整过程中,这个调度策略没有考虑当前各个节点的资源分配情况,也就是说在实现对新的环境优化和适应的过程中,没有考虑尽可能地减少系统中节点的变动,提升系统的稳定性。这些因素的缺失,对于Storm系统的性能必然带来一定程度上的损伤。

在资源调度方面,针对大数据流式计算环境中,应用均是通过有向任务图进行描述的客观事实,需要构建一个弹性、自适应的在线调度策略,满足大数据流式应用一旦开启将永远运行下去的在线场景下的资源调度需要,即一方面要有效地适应数据流、资源等方面的动态变化,另一方面也要保持系统的稳定性,避免因调整导致的系统大幅度波动,影响系统的稳定性。具体包括以下内容:在有向任务图节点计算量和节点通信量的量化方面,节点处理时延影响因素有逻辑节点的功能、数据处理功能、数据流流速大小等;节点间传输时延影响因素有节点间传输数据流大小、网络带宽,物理距离等。整个有向任务

图在任何时刻都会存在一条关键路径,其时延也是由该图的当前关键路径决定的,关键路径将是整个有向任务图的核心和瓶颈,明确了当前的关键路径,就可以找到改善系统性能的要害。同时随着不同数据流压力的变化,各个节点的计算时延和传输时延也会发生变化,这样不同数据流压力情况下的有向任务图的关键路径也会动态变化。在有向任务图到系统资源的放置策略以及在线调整方面,对于一个经过优化和调整后的任务拓扑结构的实例图,随着数据流和系统环境的在线和实时变化,以关键路径为核心,动态地调整任务拓扑图中各节点实例在各台物理机器间的分配策略(如图2所示),可以实现对系统响应时间的显著改善。同时,当数据流压力发生变化后,只需要调整关键路径上的部分节点,就可以实现对系统性能的改善,这样就可以在尽可能地保持系统稳定性的前提下,最大程度地改善系统性能,在动态调整和优化过程中,实现对历史存量信息的最大利用。在多个有向任务图分别被提交到系统中后,需要保证在不同时间点提交的各个有向任务图可以公平地使用系统资源,这就需要明确各物理机器的计算压力和拓扑结构、各个有向任务图中节点的计算压力和传输压力,并通过一定的分配策略,实现资源的合理利用以及各有向任务图间的资源公平占有和动态调整。

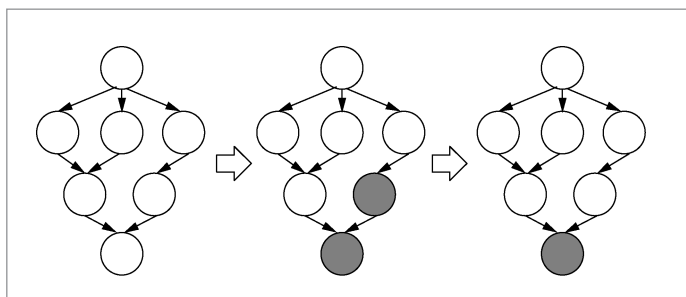


图2 资源分配和动态调整

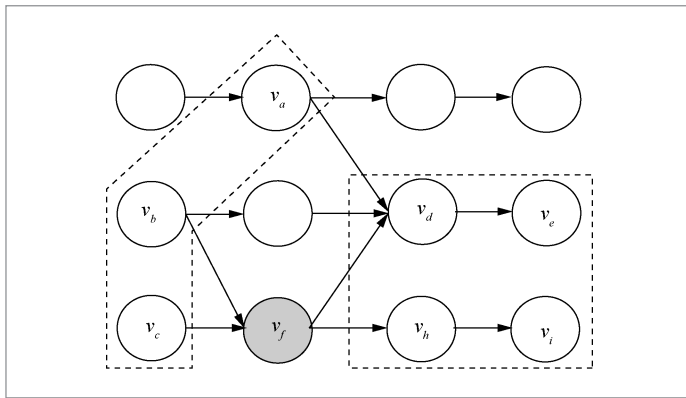


图3 故障节点最小依赖集合的容错策略

### 3.2 节点依赖环境下的容错策略

系统容错是分布式系统必不可少的一部分,特别是对于大数据流式计算系统而言,容错的价值显得尤为突出。在大数据流式计算环境中,数据流到来后往往只有一次处理的机会,重放数据流是很困难的,甚至是不可能的。这仅有的一次机会给容错带来了更高的挑战。另外,系统的实时性是大数据流式计算系统的首要目标,这也为容错策略提出了更高的要求,当高效的容错策略需要过长的时延时,会导致容错变得没有意义。Storm系统通过系统级组件Acker实现对数据流的全局计算路径的跟踪,并保证该数据流被完全执行。错误的检查是通过超时机制实现的,默认的超时时间为30 s。很显然,这么长的时延在流式应用中显得毫无意义了。

在系统容错方面,针对大数据流式计算环境中数据到来后立即进行处理,重现数据往往很困难的客观事实,需要构建一个轻量级、快速的系统容错策略,满足大数据流式计算环境中对系统容错的要求。具体研究内容包括:研究在不同应用场景下的系统容错精度方案,并对具体案例进行分析方面,主要是考虑用户的具体应用场景对容错精度的需求,进行相应容错精

度方案的设计;在建立有向任务图故障节点最小依赖集合的容错策略方面,当各个节点都进行了中间状态存储、检查点等信息的存储后,在具体的故障恢复过程中,可以选用更加高效的故障节点、最小依赖集合的容错策略,将容错的范围缩小到最小的节点集合中,如图3所示。根据有向任务图中各个节点及节点间上下游的关系以及每个节点所在的物理机器性能特征等多方面因素,为不同的节点设置有区分的、有差异的检查点频率以及不同的全量和增量式容错策略。

## 4 结束语

在大数据时代,随着越来越多的应用场景对时效性的要求不断增强,大数据流式计算作为大数据计算的一种形态,其重要性也在不断增强。本文针对大数据环境中流式计算应用所呈现出的诸多鲜明特征进行了系统化的分析,并从系统架构的角度,给出了大数据流式计算系统构建的原则性策略。结合当前比较典型的流式计算平台,重点研究了当前大数据流式计算在在线环境下的资源调度和节点依赖环境下的容错策略等方面的技术挑战。在未来的工作中,将结合应用场景的需求,开展系统架构的优化,改善整个系统的可扩展性、稳定性等多方面的特征。同时,针对应用的要求,开展上层应用同系统架构、数据流之间的动态、弹性自适应性的优化。

## 参考文献

- [1] Assunção M D, Calheiros R N, Bianchi S, *et al.* Big data computing and clouds: trends and future directions. *Journal of Parallel and Distributed Computing*, 2015(79~80): 3~15

- [2] Chen C L P, Zhang C Y. Data-intensive applications, challenges, techniques and technologies: a survey on big data. *Information Sciences*, 2014, 275(11): 314~347
- [3] Kambatla K, Kollias G, Kumar V, *et al.* Trends in big data analytics. *Journal of Parallel and Distributed Computing*, 2014, 74(7): 2561~2573
- [4] 李学龙, 龚海刚. 大数据系统综述. *中国科学: 信息科学*, 2015, 45(1): 1~44  
Li X L, Gong H G. A survey on big data systems. *Science China: Information Sciences*, 2015, 45(1): 1~44
- [5] 孟小峰, 慈祥. 大数据管理: 概念、技术与挑战. *计算机研究与发展*, 2013, 50(1), 146~169  
Meng X F, Ci X. Big data management: concepts, techniques and challenges. *Journal of Computer Research and Development*, 2013, 50(1): 146~169
- [6] Dehne F, Kong Q, Rau-Chaplin A, *et al.* Scalable real-time OLAP on cloud architectures. *Journal of Parallel and Distributed Computing*, 2015(79~80): 1920~1948
- [7] Zhang H, Chen G, Ooi B C, *et al.* In-memory big data management and processing: a survey. *IEEE Transactions on Knowledge and Data Engineering*, 2015, 27(7): 1920~1948
- [8] Zaharia M, Das T, Li H Y, *et al.* Discretized streams: fault-tolerant streaming computation at scale. *Proceedings of the 24th ACM Symposium on Operating Systems Principles, California, USA*, 2013: 423~438
- [9] Lv Y S, Duan Y J, Kang W W, *et al.* Traffic flow prediction with big data: a deep learning approach. *IEEE Transactions on Intelligent Transportation Systems*, 2015, 16(2): 865~873
- [10] Agerri R, Artola X, Beloki Z, *et al.* Big data for natural language processing: a streaming approach. *Knowledge-Based Systems*, 2015(79): 36~42
- [11] Sfrent A, Pop F. Asymptotic scheduling for many task computing in big data platforms. *Information Sciences*, 2015(319): 71~91
- [12] Yang F, Qian Z P, Chen X W, *et al.* Sonora: a platform for continuous mobile-cloud computing. <http://research.microsoft.com/apps/pubs/default.aspx?id=161446>, 2012
- [13] Andreolini M, Colajanni M, Pietri M, *et al.* Adaptive, scalable and reliable monitoring of big data on clouds. *Journal of Parallel and Distributed Computing*, 2015(79~80): 67~79

## 作者简介



孙大为, 男, 博士后, 中国地质大学(北京)信息工程学院讲师, 目前主要从事大数据计算、云计算、可信计算等方面的研究工作。

收稿日期: 2015-08-26

基金项目: 中国博士后科学基金资助项目(No. 2014M560976), 中央高校基本科研业务费专项资金(No. 2652015338)

**Foundations Items:** The China Postdoctoral Science Foundation Under Grant (No. 2014M560976), The Fundamental Research Funds for the Central Universities Under Grant (No. 2652015338)

论文引用格式: 孙大为. 大数据流式计算: 应用特征和技术挑战. *大数据*, 2015032

Sun D W. Big data stream computing: features and challenges. *Big Data Research*, 2015032