

# 1号店电商峰值与流式计算

21CTO 2015-12-04

## 导语

1号店结合自己的业务需求，在力求降低成本的前提下，最终采纳Storm计算框架来实现自己的分布式流计算平台。本文中详细阐释了这一过程中的最佳技术实践。

京东618、1号店711，还有全民购物狂欢节双11，电商促销的浪潮此起彼伏。然而，在买家和卖家欢呼雀跃的同时，电商平台正在经历着非常严峻的考验。面对一天之内犹如洪水般的网购流量，哪怕出现几分钟的闪失，都可能造成众多笔订单的损失，以及无法挽回的销售收入。时间就是金钱，在这一刻体现得淋漓尽致。如何设计电商峰值系统来更好地满足客户需求，如何在海量数据处理中实时发现有效信息并转化为商机，成为众多电商企业都在认真思考的问题。

电商峰值系统要满足三面的要求：低时延，系统对数据的处理速度要快；高可用性，故障可及时恢复，对业务影响降到最小；易扩展性，可动态添加新的计算节点，不影响在线业务，甚至可以自动做出优化调整。从这三个角度出发，Hadoop框架更适合于大数据批量处理。毕竟，它是先存储再访问，属于“一次写入多次读取”的业务类型。然而电商行业的很多业务，强调访问处理的实时性，包括电商搜索引擎、基于用户购买行为的实时商品推荐和针对网站流量的监控及反作弊等功能。这些业务要求处理行为达到秒级甚至毫秒级的时延，从而促进了以Storm为代表的流式计算系统的推广和应用。

## 流式计算解决方案

1号店结合自己的业务需求，在力求降低成本的前提下，最终采纳Storm计算框架来实现自己的分布式流计算平台。1号店实时数据处理的整体流程如图1所示。



图1 1号店分布式流计算系统

Tracker是1号店独自开发的数据记录方案，它结合Flume构成网站的数据采集模块，能在保证日志记录效率和稳定性的同时，更好地支持横向扩展，以应对日益庞大的数据量。我们采用Kafka作为前端消息数据缓冲，尽可能降低数据丢失率，同时能够灵活满足不同业务对消息处理并行性和有序性的偏好要求。Storm应用的计算结果，按照各自业务需求选择持久化存储或丢弃。

为了更进一步保证流式计算系统的稳定性，光有容错处理是不够的，我们还必须想方设法减少计算平台被过重负载压垮的风险。Linux容器技术为我们的需求提供了极大便利。它以CGroup内核功能为基础，可以支持对CPU、内存、块设备读写和网络流量等资源的隔离限制，而且此类限制可以细化到进程级别（CGroup是以进程组为基本工作单位的）。通过采用Linux容器技术，可以避免某些业务进程侵占过多系统资源，从而导致节点系统响应缓慢甚至完全瘫痪。

很遗憾，Storm自身还没有支持CGroup资源隔离功能。目前的Storm on Yarn还是个概念性实现，因为YARN对Storm的资源隔离，仅针对整个Storm集群所占用的资源，以实现和Hadoop批量计算框架在同一集群上的共存。而我们的客观需求是希望能对Storm平台上Topology一级进行资源限制，对应到每个计算节点上，就是一种进程级别的资源隔离需求。最终的解决办法是，我们独立设计自己的CGroup资源管理框架，来实现对Storm Topology的资源隔离限制。在图2所示的设计方案中，我们主要考虑以下几点。

图2 1号店Storm集群上的资源管理框架

### 更简单的用户体验

1. 用户不需掌握CGroup的复杂细节（层级、子系统、组概念及相关操作系统和文件系统知识）。
2. 仅需掌握三种操作：

创建/删除某一类型的CGroup，目前支持cpu、memory和cpuset三种类型；  
分配进程到指定的CGroup。
3. 如上操作可以通过重新设计实现的客户端指令（ycgClient）完成。
4. 用户仅需在Storm UI上对Storm Topology指定优先级（该处优先级定义为进程所在的进程组可获取资源的大小）。
5. 由守护进程（ycgManager）自动管理各计算节点的进程级优先级。

### 自动化方案降低手动管理成本

Storm Topology的优先级信息存储在ZooKeeper集群中。  
资源管理框架可以自适应异构机器结点的动态添加。  
便于集群管理（机器配置信息会自动存储在ZooKeeper中，包括逻辑CPU个数、内存和交换分区大小）。

### 能对Storm集群的Worker/Executer级别进行资源限制

这是目前Storm on Yarn做不到的。

### 模块独立于计算框架本身，方便安装部署及回滚

适当的改进可以很方便地应用于其他框架平台中。

针对Storm UI不能满足我们业务需求的问题，我们基于非Clojure语言重新设计实现了其UI模块。之所以不采用Clojure语言及其Web框架，主要是考虑降低日后的学习和维护成本。在早期的使用中，我们发现原始

的Storm UI有以下功能缺陷。

只能通过命令行提交Topology Jar包。

无法记录Topology操作日志，给运维管理带来不便。

缺少用户权限控制。

通过重新设计Storm UI，以上问题得到解决。管理员可以操作Storm UI的用户及其权限，也可以在UI上查看Topology的操作日志，对Storm应用进行有效跟踪管理。未来我们会在UI上对Topology做更细致的实时监控。

随着1号店越来越多的应用被部署在Storm平台上，实时计算的效益日趋明显。2014年上半年部分业务每天产生的日志大约有2亿多条，部署在Storm平台上的商品流量监控、个性化推荐和BI应用，都能做到个位数的秒级响应。同时，我们的日志流量在以极快的速度递增。除此之外，我们还有安全日志分析、反欺诈和订单监控等应用在有效运行。

## 总结

1号店作为一家成立时间较短的中小规模电商，业务增长十分迅速。我们意识到自己的实时计算平台在今后会有更大的压力和考验。在保证现有系统正常运行的同时，我们也在积极搜集业务部门的各种反馈，基于目前的平台和技术做进一步的调研和二次开发。如何提高系统峰值处理时的性能和可靠性，我们依然任重道远。

原文：<http://blog.csdn.net/qw1w1w/article/details/41020763>

### 关于21CTO

21CTO.com是中国互联网第一技术人脉与社交平台。我们为国内最优秀的开发者提供社交、学习等产品，帮助企业快速对接开发者，包括人才招聘，项目研发，顾问咨询服务。

看微信文章不过瘾，请移步到网站，诚挚欢迎您加入社区作者团队。

网站地址：[www.21cto.com](http://www.21cto.com)

投稿邮箱：[info@21cto.com](mailto:info@21cto.com)