



应对告警风暴告警的系统优化策略

甘 雯¹, 文 锋¹, 宫大鹏¹, 徐 钊¹, 黄甫光¹, 张 健¹, 苏 雷²

(1. 中国移动通信集团广西有限公司 南宁 530022;

2. 亿阳信通股份有限公司 南宁 530022)

摘 要: 长期以来, 随着网络规模的扩大, 告警风暴一直是电信网络管理工作的重点与难点。通过对现有告警堆积日常处理经验的研究, 包括融合对硬件架构调整、软件调整、中间件配置优化等方法, 得到一套解决告警风暴的处理思路, 为电信运营商解决此类问题提供借鉴经验。

关键词: 告警风暴处理; 综合告警架构优化; 应对告警量突增

doi: 10.11959/j.issn.1000-0801.2015132

System Optimization Strategy of Alarm Storm

Gan Wen¹, Wen Feng¹, Gong Dapeng¹, Xu Tan¹, Huang Fuguang¹, Zhang Jian¹, Su Lei²

(1. China Mobile Guangxi Branch, Nanning 530022, China; 2. BOCO Inter-Telecom Inc., Nanning 530022, China)

Abstract: For a long time, with the expansion of the network, the alarm storm has been the focus and difficulty of telecommunication network management. Through the study of daily experience of the existing alarm accumulation, including fusion of hardware structure adjustment, software adjustment, middleware configuration optimization, and optimize the allocation of method to obtain a set of solving ways of the alarm storm was obtained, providing the experience for telecom operators to solve such problems.

Key words: alarm storm processing, integrated alarm framework optimization, alarm overload processing

1 引言

综合告警系统是 CMOSS2.0 规划的综合网管系统, 系统主要从 OMC(operation and maintenance center, 操作维护中心)等厂商网管系统获取资源、告警、性能数据, 包括从资源管理系统获取资源数据、从各专业网管获取告警和性能数据、从电子运维系统获取工单数据等, 然后通过列表、图表、拓扑、GIS(geographic information system)等方式进行数据的汇总呈现, 帮助监控人员了解全网的运行状态。

在通信网络运行过程中, 告警是网络管理员最为关注的。当系统出现影响正常业务的故障时, 这些重要的故障信息会以告警的方式在第一时间通知管理人员并应该立即得到解决, 否则, 可能会导致提供服务失败。为了方便运维人员处理告警和迅速定位告警源, 综合告警系统采取了各种方式处理这些告警信息, 比如告警展现、长时间未处理的告警提示、告警转发、告警过滤、告警相关性分析等。

在一种极端的情况下, 众多网元(BTS(base transceiver station, 基站收发信台)、RNC(radio network controller, 无线

收稿日期: 2015-03-05; 修回日期: 2015-05-14

论文引用格式: 甘雯, 文锋, 宫大鹏等. 应对告警风暴告警的系统优化策略. 电信科学, 2015132

Gan W, Wen F, Gong D P, et al. System optimization strategy of alarm storm. Telecommunications Science, 2015132



网络控制器)、Node B 等)由于不特定原因,同时并且长时间地向网管系统上报大量的告警,导致告警风暴的发生。如果告警系统没有及时处理,容易造成海量告警的堆积,导致网管系统瘫痪,失去管理和监控网络的能力,更不能有效遏制网络故障的进一步扩大。所以告警风暴的危害是巨大的,一方面应该尽量避免告警风暴的产生;另一方面,当告警风暴到来时,系统应有能力及时应对,将告警风暴的危害降到最低。

国外的研发机构一般通过研究告警关联性和告警挖掘技术解决此项难题,例如,惠普公司的 Event Correlation Services,是基于规则的方法研究出的告警相关性分析系统;IBM 采用基于事例的告警相关性分析方法,研制了 NetFACT 系统,利用告警相关性的研究对告警发生进行业务关联,达到突出主用告警、抑制无用告警的目的,从而大大减少告警量。

近年来,国内进行应对告警风暴研究的主要以设备厂商为主,例如中兴通讯、华为技术、大唐电信等,都开始着力研制开发智能高效的移动通信网络管理系统,包括告警相关性分析系统等。通过将一些研究成果融合到 OMC、传输 EMS(element management system,网元管理系统)等网管系统中,实现部分告警风暴抑制。

本文旨在探讨对于架构在设备厂商网管之上的综合网管系统,如何通过升级现有架构来应对告警风暴的发生,寻找合理的抑制告警风暴手段。

2 业务的高速发展和告警风暴的产生

所谓告警风暴往往被定义为在短时间内,产生了大量的告警事件,在这些事件中,有的互相存在一定关联,是由

于某种共用因素导致的。这样大量的告警在短时间内挤压告警上报通道,导致通道的堵塞甚至上层监控业务的崩溃。近年来由于告警风暴导致的数据库锁死、队列溢出、告警监控服务挂死等现象时有发生,严重影响了运营商运维管理人员的日常监控作业。

随着网络规模的不断扩大、新类型网元接入以及跨专业监控管理的现实需求,网管系统的数据接入量在不断增加,但任何一个系统的运行均有其极限值,当数据处理效率不能满足系统正常运行要求时,系统的可用性就会随之下降,对于告警监控这类实时消息系统表现得就更为明显。当在一段时间内的数据突增,超出系统处理能力时,会导致数据处理延时或者数据丢失,从而引发监控系统的不可用。

如图 1 所示,在具体的生产实践中,工程割接、自然灾害、OMC 重启、传输链路异常中断等引起的告警风暴是数据风暴的典型场景。通常反应为:在超出现有告警监控系统数据处理能力的情况下,监控系统无法及时准确地将某个网元或者相关网元的告警解析呈现,从而无法起到监控作用。

为了适应网络规模和业务持续高速增长给网络运维工作带来的巨大压力和挑战,有力支撑公司未来发展,中国移动通信集团公司于 2013 年率先提出了“集中化”的网络运维体制改革思路,经过不断建设和深化,已经形成了全网跨专业的“集中监控、集中维护、集中管理”格局。集中监控的工作内容之一是告警的集中监控。随着集中监控规模的不断发展,告警监控工作面临着前所未有的挑战:由于集中化监控的设备数量逐年快速增长,集中监控对故障的处理质量提出了更高的要求,特别是当告警风暴发生

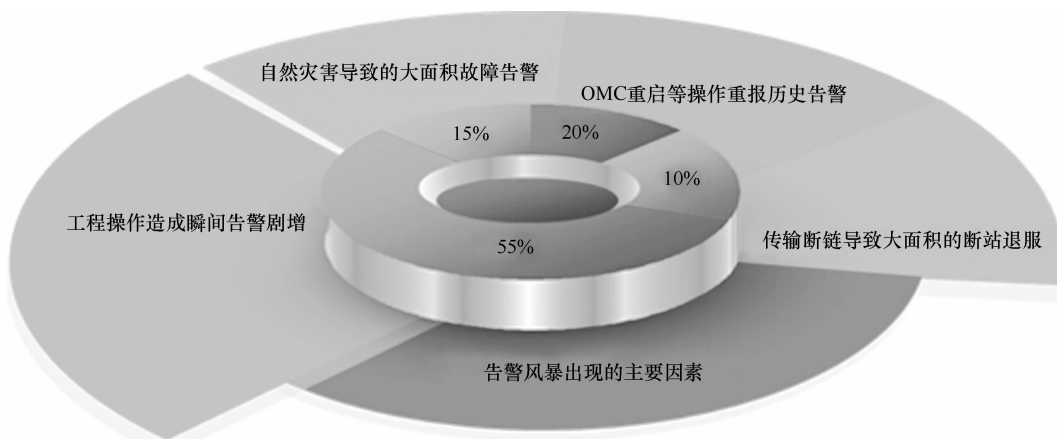


图 1 告警数据风暴出现的主要场景分析

时,系统崩溃导致值班人员无法正常开展监控工作,进而影响通信保障抢修。为此,迫切需提高告警的准确性、有效性,并通过升级基础架构提高告警处理的效率,应对极端情况下的告警风暴发生。

3 综合告警的网管技术架构升级

3.1 应对告警风暴的性能提升思路

在实际优化过程中,影响系统运行效率的原因是多方面的。网络传输效率、硬件性能、MQ 配置、应用软件部署方式、采集软件配置、数据库效率、产品设计都会对系统性造成影响。

如图 2 所示,告警风暴主要通过 OMC 产生,一旦产生,需要监控人员引起重视。告警风暴处理包含风暴抑制和风暴后处理两部分内容:风暴抑制是系统判断告警风暴产生后,将告警经过特殊通道直接传递到消息平台,并将风暴告警保存到文件中,或者停止对某个网元某个告警标题的采集;风暴后处理是将风暴期间的告警入库,以便于统计分析,并且将风暴期间漏掉的告警通过正常流程发送上去。告警风暴产生或已经停止时,系统产生预警消息,包括活动告警、清除告警。

从告警来源和架构上分析,为有效应对告警风暴,考虑从硬件、软件、中间件等方面入手进行针对性的调整。同时通过配置合理的日常维护管理制度,可大大减少告警风暴导致的业务中断情况发生的几率。

3.2 网络及系统硬件调整手段

2014 年以前,中国移动综合告警系统主要实现了全专业告警接入、告警标准化、告警关联、集客业务监控、性能场景监控等功能模块。以广西移动综合告警项目为例,在优化前,网管硬件平台的网络拓扑如图 3 所示,系统部署在网管硬件平台的 2 台 M9000 服务器上,存储使用网管硬件平台的 DX8700 磁盘阵列。

随着集中故障管理业务上线,经过对告警业务的处理量、并发事务数、峰值进行测算,得到硬件 TPC-C 吞吐率(TPC-C 使用 3 种性能和价格度量,其中性能由 TPC-C 吞吐率衡量,单位是 tpmC,其含义为每分钟内系统处理的新订单个数)的需求,进而推算出扩容方案,具体如下。

- 系统数据库服务器需要约 149 万 tpmC 处理能力,当数据库与 MQ 服务器同时部署在一台服务器的时候,建议把综合告警数据库单独拆分出来,利用已有 M9000-2 分区的计算能力,扩容 4 颗 4 核 CPU,以满足需求。
- 采集服务器需要约 127 万 tpmC 处理能力,将应用服务器的 M9000 分区的计算能力由 6 颗 4 核 CPU 内存扩容至 8 颗 4 核 CPU,扩容后可满足需求。
- 消息服务器需要约 128 万 tpmC 处理能力。将目前应用服务器的 M9000 分区由 6 颗 4 核 CPU 扩容至 8 颗 4 核 CPU,扩容后分区能满足需求。
- 拓扑服务器的能力需求为 34.86 万 tpmC,从其他机房的 x86 资源池划分 1 台总能力大于 35 万 tpmC 的虚拟机以满足需求。另外,为了保证必要的数据缓存机制,还增加了 2 TB 的存储空间。

最终调整结束后,硬件网络拓扑如图 4 所示。

此外,对于上层应用服务器也需要做负载均衡。根据广西 14 个地市的网络及业务规模,将其分成大、中、小 3 类地市,制定了动态分配应用服务器的负载均衡策略。例如,当沿海城市发生台风等自然灾害,告警突增时,策略器就会自动分配登录服务在当前空闲的机器上,保证性能使用的最大化。

通过硬件调整,可处理 200 万条/天的告警并发负载量,参照历史最严重的告警风暴发生,评估得出该硬件架构能力可并发处理 90%的告警。

告警风暴消息详情 2013-03-19 00:43:11						
序号	发生时间	标题	状态	级别	采集OMC	操作
1	2013-03-05 16:03:29	告警量超门限800	活动	二级	OMC01	
2	2013-03-05 16:03:29	OMC1风暴产生告警	活动	一级	OMC02	手工同步 查看详情 手工入库 处理设置
3	2013-03-04 12:03:29	OMC2风暴产生告警	已清除	一级	OMC03	手工同步 查看详情 手工入库 处理设置

图 2 告警风暴实例

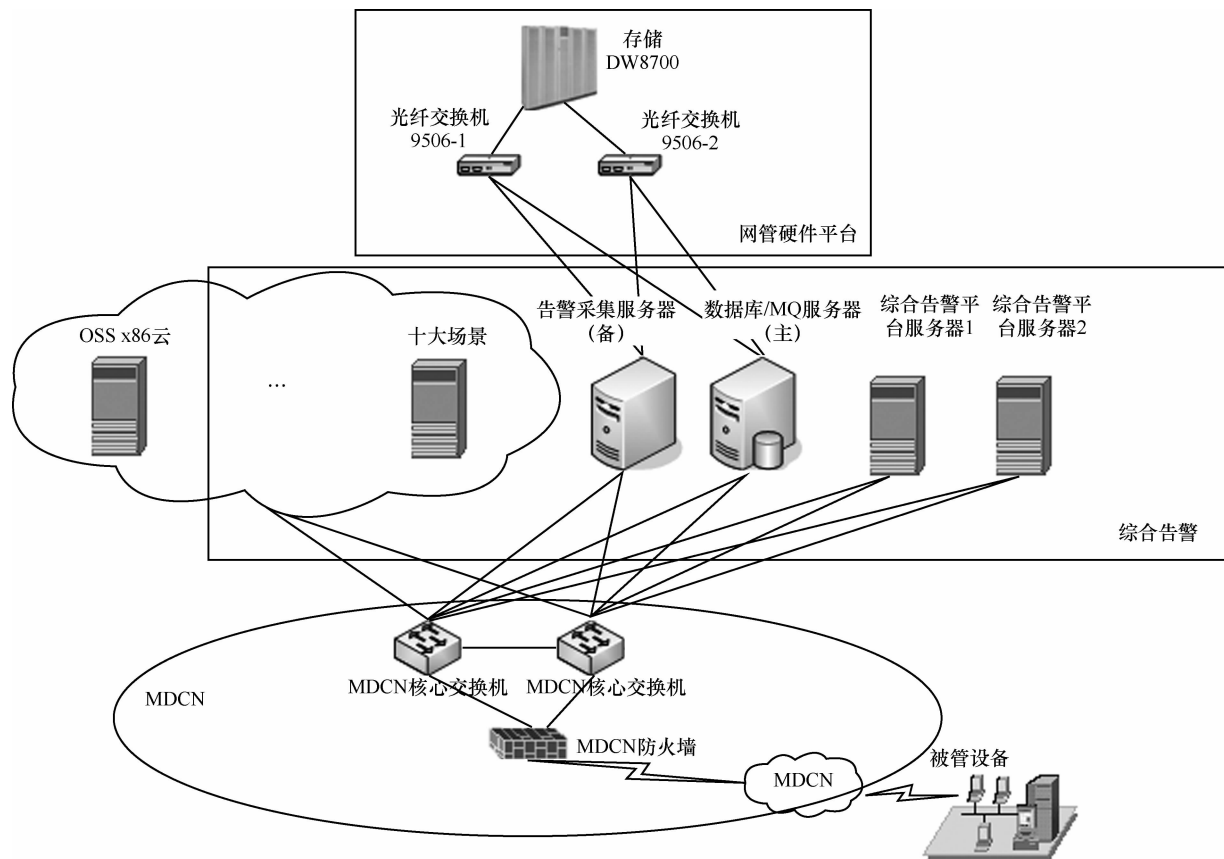


图3 综合告警系统部署调整前拓扑

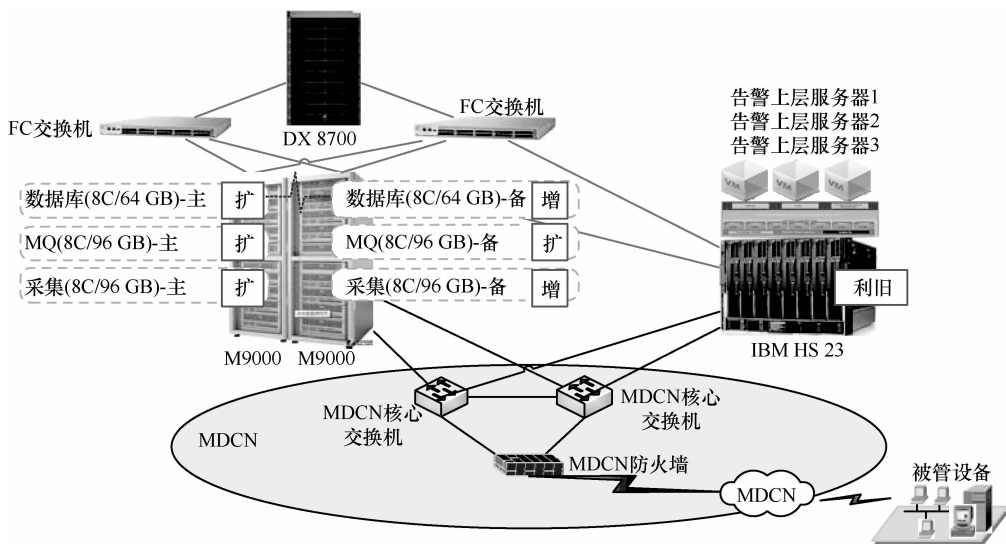


图4 综合告警系统部署调整后拓扑

3.3 软件结构升级手段

除了对硬件进行改造外,还通过对现有告警消息处理机制进行优化来应对告警风暴。如图5所示,告警在设备北向采集后会经过清洗阶段、标准化阶段、分发入库阶段和上层展现阶段,每个阶段都可能会成为告警上报的性能瓶颈。特别是在大量告警队列堆积时,有针对性地调整脚

本的处理效率,可大大提升告警处理效率。

优化手段可以从以下几个方面考虑。

(1) 智能并发告警采集

针对采集通道进行并行处理,设计软负载均衡队列管理语法,根据告警流量智能控制并行处理单元数,假设当A通道告警是B通道的两倍时,会在给A通道多分配50%

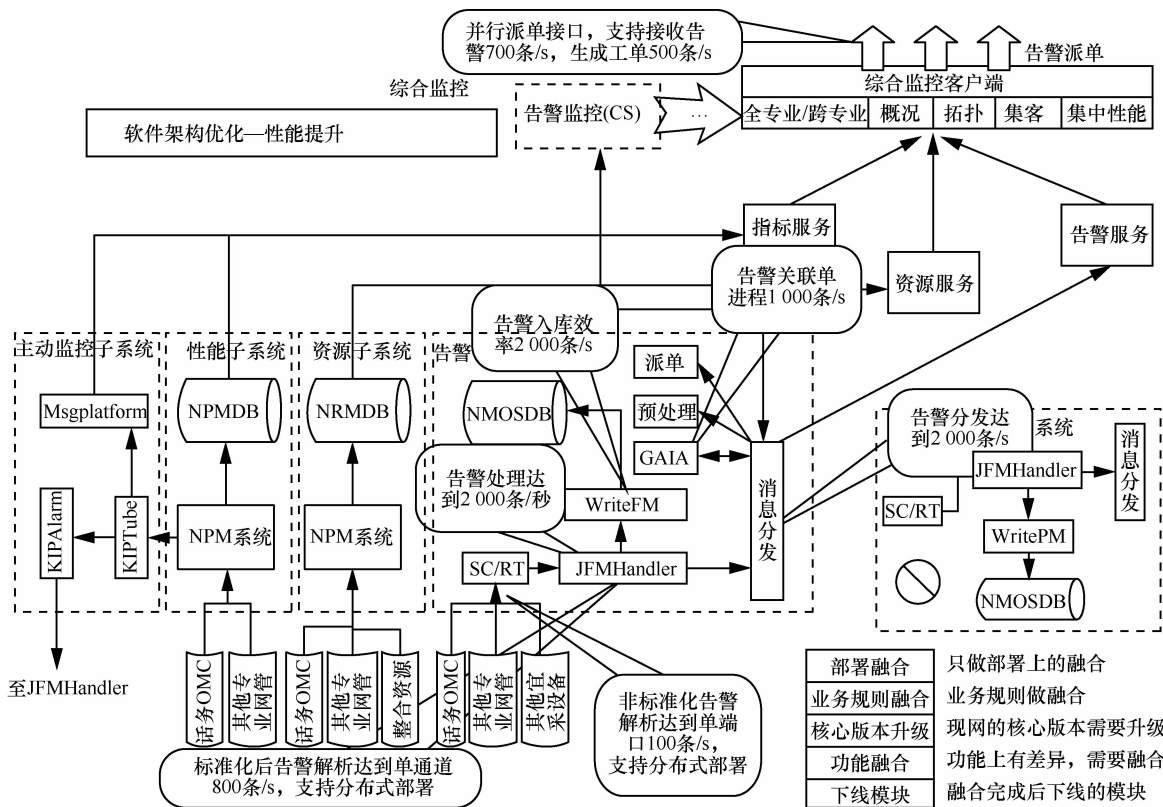


图 5 采集解析模块优化架构

的处理进程,按照该算法,通过调整可提升单通道采集能力达到 60 条/s。

(2) 基于高速缓存的告警处理系统

由于一般在告警标准化过程中,需要反复读取告警进行标准化字段的翻译。针对这种特点,可以利用高速缓存技术,如图 6 所示,将待处理告警所需要的信息存储在内存中,加快了 CPU 的处理效率,迅速提升告警标准化、工程预约等标准化处理能力,提升整体处理能力达到 400 条/s。

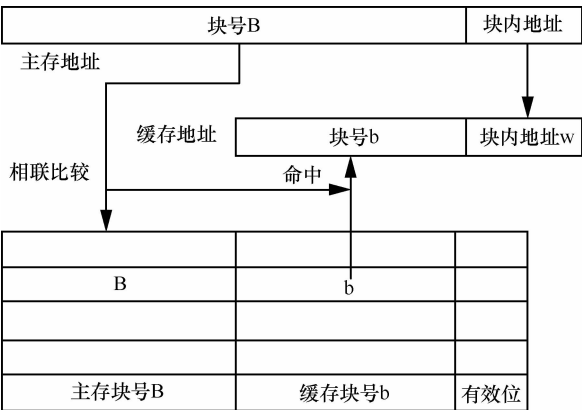


图 6 告警缓存建立示意

(3) 基于高效规则处理和并发的告警订阅分发系统

通过建立高效规则处理实现告警过滤器能力,识别多种闪现告警和关联告警并予以合并或抛弃,并结合多并发分发处理,在并发多个过滤器(50 个)情况下,提升分发效率,可以实现 400 条/s 的处理速度。

基于流计算的告警关联系统摒弃使用数据库计算的方式,改为在接收告警流数据时就开始对其进行分析,通过实时的对象流计算,可快速处理告警之间的关联关系,提升单规则处理能力,实验表明可提升到 200 条/s。

(4) 基于智能调度的告警派单处理能力提升

针对上层应用,告警的落地除了展现,更重要的是派单,提升 EOMS 派单系统自身处理效率可在业务上解决告警风暴的问题。针对 EOMS 单通道处理能力不够的情况,通过建立智能的调度机制,动态并发调用,大幅提升派单能力,派单系统自身处理能力达到 150 条/s。

(5) 数据库调优

通过对数据库 Oracle/Informix 进行分片机制,增加索引,并对大量历史数据进行拆分,也可以使得告警实时查询速率提升。广西移动采集模块优化测试结果见表 1,以广西移动实际项目测试结果为例,经过优化,可获得显著的提升效果。



表 1 广西移动采集模块优化测试结果

厂商	节点(单通道)	优化前效率(每秒处理告警数)/条	优化后效率(每秒处理告警数)/条
爱立信	101111	100	200
	1602111	70	120
	1603111	350	500
中兴通讯	702111	150	250
	7601111	6	300
上海贝尔	12601111	90	300
	12602111	40	400
	824111	131	243
华为技术	8601111	25	112
	813111	70	110
	814111	71	133
	823111	103	239
	8606111	14	155

3.4 中间件参数配置手段

如图 7 所示,消息队列(MQ)是一种应用程序对应用程序的通信方法。应用程序通过“写”和“检索”出入列队的针对应用程序的数据(消息)来通信,而无需专用连接器来链接它们。消息传递指的是程序之间通过在消息中发送数据进行通信,而不是通过直接调用彼此来通信,直接调用通常是用于诸如远程过程调用的技术。排队指的是应用程序通过队列来通信。队列的使用除去了接收和发送应用程序同时执行的要求。

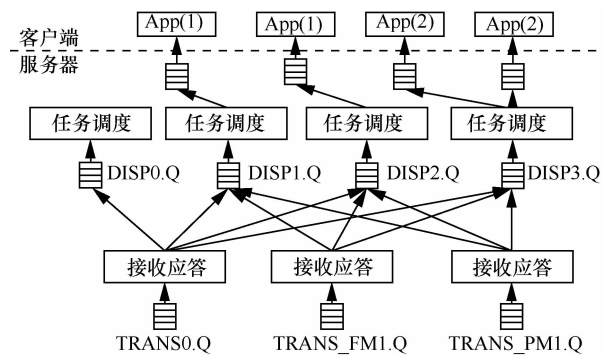


图 7 告警中间件 MQ 流程

IBM WebSphere MQ 是综合告警平台使用的队列中间件,对该中间件的参数配置进行优化,可以起到优化队列长度,提高告警处理效率的效果。

对 MQ 进行了优化调校,主要从服务器参数、MQ 参数、MQ 日志配置、队列参数、缓冲区参数、程序 API 调用等多个方面,增强了服务器处理性能,提高了 MQ 服务的处理速度,减少了核心程序对 MQ 的压力。具体为以下几

个方面的调优。

- 调整/etc/system 系统参数,优化 MQ 服务的数据处理性能。
- 优化 MQ 的断网续传参数,在 Sun 平台下调整为:/usr/sbin/ndd-set/dev/tcpkeepalive_interval 15000。
- 优化 MQ 的日志配置,修改/var/mqm/qmgrs/队列管理器名称为/qm.ini,调整日志文件的个数、每个日志文件的大小、日志缓冲区大小。
- 修改侦听的启动方式,采用 runmqslr 方式提高通道相关的性能。
- 设置侦听程序采用 trusted 方式运行,降低 CPU 和内存消耗。
- 增加通道的 PipeLineLength 属性,设置 MCA 参数采用多个线程的方式传输消息,从而提高通道性能。
- 修改/var/mqm/qmgrs/队列管理器名称/qm.ini 的 TCP 选项 KeepAlive=Yes,使操作系统的 TCP/IP 参数设置对 WebSphere MQ 生效。
- 修改核心程序对 MQ 的操作方式:MQCONN 和 MQDISC 是最耗 CPU 的两个函数,减少 MQOPEN 和 MQCLOSE 函数的调用,使用 MQCONNX 函数建立与队列管理器的连接,使应用程序和本地队列管理器代理组成同一个进程,从而提高性能。

以 JFMHandler 处理能力为例,经过测试,效果如图 8 所示。

优化后,JFM 处理消息的性能大约是 544 条/s,相比优化前的 400 条/s 处理能力提升了 36%。

队列 (发送/接收)	消息数量	开始时间	结束时间	时长/s	处理性能
P_FMH01.Q/P_WFM01.Q	38 801	17:58:13	18:04:19	354	110
P_FMH02.Q/P_WFM02.Q	38 801	17:58:13	18:04:16	357	109
P_FMH03.Q/P_WFM03.Q	38 801	17:58:13	18:04:17	356	109
P_FMH04.Q/P_WFM04.Q	38 801	17:58:13	18:04:15	358	108
P_FMH05.Q/P_WFM05.Q	38 801	17:58:13	18:04:15	358	108

图8 测试效果

3.5 服务异常检测手段

为保证在告警风暴出现时,系统有能力对告警管理人员进行支撑,需要对告警系统新增服务负荷监控模块,对综合告警客户端发起连接到服务端请求的过滤器服务进行监控,包括连接状态、连接所占资源信息,方便监控人员及时定位服务异常情况。一旦发现服务内存溢出系统无法自行恢复时,可在最短时间内通过手工进行恢复,保证业务连续性。

4 升级改造后取得的成效

在2013年,广西移动发生过类似台风灾害场景产生180万条/天的告警量,平均处理能力只有12.66条/s。在经过了硬件扩容、软件结构优化等调优工作,并进行测试和部署调整后,经过6个月的现网环境观察,告警风暴处理能力有明显提升。2014年7月,台风威马逊袭击中国东南沿海,广西受灾。每天全网产生280万条告警,形成特大告警风暴。平均每小时产生12万条告警,峰值告警为15万条,远远超过了系统的负载能力。经过事后日志分析,告警风暴最严重的7月19日,平均处理能力可达到120.4条/s,入库60.3条/s。峰值挤压时可在1h内处理完毕。处理效率提高将近10倍,大大缓解了告警数据量的压力。告警处理提升速度对照如图9所示。

告警风暴优化后,处理速度比优化前翻一番,细化了综合监控的支撑粒度,解决了遇到重大故障不可用的问

题,实现了由面向网络到面向客户、由被动运维向主动运维的转变。以业务主体为核心,实现故障的集中监控、集中管理、工单直派一线,提高故障定位速度、分析和处理的效率,最终达到提升网络运维质量的目的。

5 结束语

随着移动通信4G网络的建设以及集中化运维体制改革的深入推进,一个能有效管理网络的运行支撑系统就成为必不可少的工具,成为影响运营质量的重要因素之一。通过升级综合告警系统的软硬件架构,使之具备应对告警风暴的能力,能够使得综合告警平台为网络运维创造更多价值。

在研究过程中,参考了部分省级运营商的先进经验,结合自身“4+1”网管建设推进部署,每年都都对系统进行深入研究,从业务(告警关联)到技术(技术架构)进行深度优化,持续改进提升,以应对不断变换的网络环境,发挥网管系统的最大价值。

参考文献

1 ITU-T Rec M 3200. TMN Management Services and Communications Managed Areas: Overview, 2001

2 郑庆国,吕卫峰.通信网络中的告警相关性研究. 计算机工程与应用,2002(2):11~15

Zeng Q G, Lv W F. Study on alarm correlation in communication network. Computer Engineering and Applications, 2002(2):11~15

3 石永革,梅玉洁,石峰.通信网管告警过滤机制的研究与应用. 计算机工程与设计,2008,29(9):2169~2171

Shi Y G, Mei Y J, Shi F. Research and application of communication network management alarm filtering mechanism. Computer Engineering and Design, 2008, 29(9): 2169~2171

4 klemettinen M, Mannila H, Toivonen H. Rule discovery in telecommunication alarm data. Journal of Network and Systems Management, 1999(4): 395~423

5 潘沛. 电信网络综合告警系统需求分析与设计. 大众科技, 2009(5)

Pan P. Analysis and design of telecom network integrated alarm system requirements. Public Science and Technology, 2009(5)

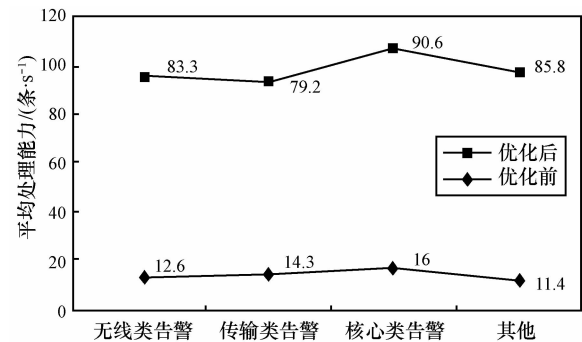


图9 告警处理提升速度对照



- 6 闻海舟. 浅谈电信网络综合告警系统建设方案. 广西通信技术, 2011(2)

Wen H Z. The construction scheme of telecom network integrated alarm system. Guangxi Communication Technology, 2011(2)

- 7 冯婧姝, 李兴明. 基于加权关联模式的通信网告警相关性分析. 电信科学, 2007, 23(11)

Feng J Y, Li X M. Telecommunication alarm correlation analysis model based on weighted association. Telecommunications Science, 2007, 23(11)

- 8 李宝山, 王苏东. 告警管理系统中的告警同步模块的设计. 通信技术, 2013(4)

Li B S, Wang S D. Design of alarm synchronization module in the alarm management system. Communications Technology, 2013(4)

- 9 Kettischau H J, Bruck S, Schefezik P L. An expert system for intelligent fault management and alarm correlation. Proceedings of Network Operation and Management Symposium (NOMS), Florence, Italy, April 2002

[作者简介]



甘雯, 男, 中国移动通信集团广西有限公司信息系统部高级工程师, 主要研究方向为 BSS/OSS 技术架构、云计算技术应用、大数据系统架构。



文锋, 男, 中国移动通信集团广西有限公司信息系统部室经理, 主要研究方向为网管支撑系统、信息安全系统、互联网。

宫大鹏, 男, 中国移动通信集团广西有限公司信息系统部项目经理, 主要研究方向为网管支撑系统、系统规划。

徐钊, 男, 中国移动通信集团广西有限公司网络运营中心网管支撑室室经理, 主要研究方向为网管支撑系统、网络信息安全。

黄甫光, 男, 中国移动通信集团广西有限公司网络运营中心工程师, 主要研究方向为网管支撑系统、IT 系统维护与管理。

张健, 男, 中国移动通信集团广西有限公司区网络运营中心网管支撑, 主要研究方向为网管支撑系统。

苏雷, 男, 亿阳信通股份有限公司规划咨询工程师, 主要研究方向为行业大数据、运营商 OSS 规划。

2015132-8

2015 年 3 月通信业主要指标完成情况(一)

指标名称	单位	本年本月止累计到达	比上年同期累计(±%)	本月
营业收入	亿元	3 301.4	-2	1 171.5
其中:电信业务收入	亿元	2 769.3	2.9	1 018.6
固定资产投资完成额	亿元	489	-4.2	267.1
固定本地电话通话时长合计	万分钟	5 587 886.5	-12.7	2 008 070.7
固定长途电话通话时长合计	万分钟	1 168 138.8	-8.2	426 893.3
移动电话去通话时长合计	万分钟	69 418 020.5	-2	24 889 720.4
其中:国内长途通话时长	万分钟	15 865 247.7	-3.6	5 815 245.7
国际长途通话时长	万分钟	28 459	-9.7	9 915
移动短信业务量	万条	18 358 664.9	-2.7	6 031 169.8
移动互联网接入流量	万吉字节	76 398.2	87.9	27 992.4

注:(1)电信业务收入增长率按可比口径计算。(2)固定长途电话通话时长和移动电话通话时长均包含相应的 IP 电话通话时长。(3)通话时长各项指标均为去话通话时长。

(来源:工业和信息化部运行监测协调局)