

具有自恢复能力的机群监控系统结构

范中磊

(长安大学 信息工程学院, 陕西 西安 710064)

摘 要: 提出了一种具有自恢复能力的机群监控系统结构, 采用该结构的监控系统可以自动、及时地完成机群系统中故障的发现和排除, 从而可极大地增强机群监控系统的智能化和主动控制能力, 有效提高整个机群系统的可靠性. 采用该结构开发的机群监控系统已经在 Linux 机群存储系统中得以应用, 取得了良好的效果.

关键词: 机群系统; 机群监控系统; 可靠性; 自恢复

中图分类号: TP302

文献标识码: A

文章编号: 1000-7180(2011)05-0079-03

Architecture of Cluster Monitoring System with Self-Healing Ability

FAN Zhong-lei

(School of Information Engineering, Chang'an University, Xi'an 710064, China)

Abstract: This paper presents a kind of architecture of cluster monitoring system with self-healing ability, which enables cluster monitoring system to find and eliminate faults of a cluster system automatically and timely, and strengthens the intelligence and active control ability of a cluster monitoring system itself greatly, and then improves the reliability of a cluster system. A cluster monitoring system developed by employing this proposed architecture has been applied to Linux-based cluster storage system and achieved good effectiveness.

Key words: cluster system; cluster monitoring system; reliability; self-healing

1 引言

随着网络技术的迅速发展, 基于网络互联的各种分布式应用系统也逐渐增多. 如提供 WWW 服务的 Web 服务器机群、提供数据库服务的数据库服务器机群、提供计算服务的计算服务器机群等. 在很多应用场合, 如数据中心、电子商务、科学计算等, 机群中包含大量的服务器节点和网络设备, 而且这些服务器节点和设备的物理位置分布范围也可能是很广的. 如何对上述这样的机群系统进行有效地监测和控制, 以及及时发现并排除系统中已存在的和潜在的故障, 保证整个系统高效可靠的运行, 便成为机群监控系统需要解决的一个重要问题^[1-3].

传统的机群监控系统大多集中在信息获取方面, 也就是监测方面, 而在对系统的控制能力上相对是比较弱的, 特别是在自动控制能力上就更加薄弱了. 比较有代表性的就是基于简单网络管理协议

(SNMP) 的监控系统^[4-5], 其大都是利用 SNMP 来获取所需的信息, 加以显示和分析, 当发现一些可能的问题时便通知管理员, 通过人工干预进行故障发现和排除. 这种人工控制方式会导致系统的故障不能及时得到处理, 影响系统的工作. 在有大量节点, 而且节点物理位置分布广泛的机群应用中, 这个问题会更加突出^[6]. 文中提出了一种具有自恢复能力的机群监控系统结构, 采用该结构的系统可以自动、及时的完成机群系统中故障的发现和排除, 无须人工干预, 可以极大提高整个机群系统的可靠性, 降低管理成本.

2 系统结构

图 1 为传统的机群监控系统结构示意图. 被监控方 1 通过其上的监控代理 2 传送监测信息 3 到监控方 4, 监控方 4 上的监测、显示和报警 5 接受其所需的监测信息 3, 进行显示或报警等; 监控方 4 上的

监测、显示和报警 5 也可以通过被监控方 1 上的监控代理 2 向其传送字符型或整型数据信息 6, 进行一些阈值的设定等。

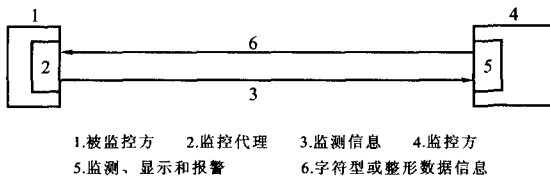


图1 传统的机群监控系统结构

图1所示的传统机群监控系统结构,其信息大多是从被监控方流向监控方,缺乏一个有效的从监控方到被监控方的控制信息流。尽管像 SNMP 协议也提供了从监控方到被监控方的信息流功能,但其功能较弱,一般只能通过其进行字符型或整型数据的传送,很难满足复杂控制的需求。

基于上述问题,文中提出一种具有自恢复能力的机群监控系统结构,如图2所示。与图1所示的传统机群监控系统结构相比,其不同在于,监控方4上的监测、显示和报警5接受到监测信息3之后,将其传送到中心分析和控制7,中心控制与分析7对接受的监测信息3进行分析决策,通过其上的控制信息通道接口8和被监控方1上的控制信息通道接口9传送文件数据或控制命令数据10到被监控方1,并控制被监控方1执行所需的处理程序或控制命令。这里10中所指的文件数据包括各种应用程序(如故障检测程序、故障处理程序、其它可执行程序等)和配置文件,控制命令数据包括下载命令、运行程序命令和各种系统命令(如启动、关机、重启、资源查询、状态查询等)。

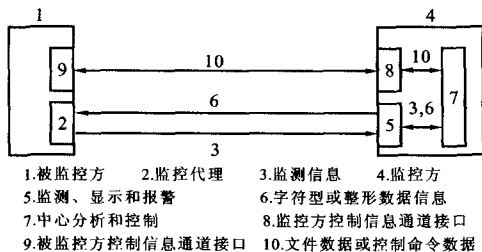


图2 具有自恢复能力的机群监控系统结构

采用图2所示的机群监控系统,当监控方的分析控制软件发现被监控方出现故障或存在潜在的故障时,便可以根据需要,主动地、有针对性地、动态地下载一些故障检测程序到被监控方,改变被监控方的监测信息(如增加或减少监视信息的节点种类和上报频率、提高监视信息密度等),为其提供进一步分析控制所需的监测信息;在获取充分的监测信息之后,通过故障原因分析,最终定位故障原因;针对

其故障,确定相应的处理方案,并通过控制信息通道下载相应的处理程序到被监控方并执行,排除故障。这样,就提高了监控方对被监控方的控制能力,实现了自动的故障分析和排除。不仅使机群监控系统具有了自恢复能力,提高了整个机群系统的可靠性,而且为系统的升级和维护带来了很大的方便。

3 实现方法

图3给出了具有自恢复能力的机群监控系统结构的一种实现方法,它已在我们的系统中得以实现和应用。其工作流程如下:

(1) 监控方4上的中心控制与分析7通过其上SNMP程序5定购其所需的监测信息,SNMP程序5将定购信息转化为SNMP SET信息6传送到被监控方1上的SNMP程序2;

(2) 被监控方1上的SNMP程序2根据收到的定购信息进行处理,将监控方4所需的监测信息转化为SNMP GET信息3传送到监控方4上的SNMP程序5,监控方4上的SNMP程序5对收到的监测信息整理和预处理之后,将其传给监控方4上的中心分析和控制7;

(3) 中心控制与分析7对接受到的监测信息进行分析、处理和决策,如有必要,决定动态增加监视信息节点,进一步收集系统信息;(这一步骤有可能被多次重复)

(4) 中心控制与分析7对接受到的监测信息进行分析、处理和决策,从文件和控制命令库13中获取所需的文件和控制命令属性信息14,转化为相应的控制命令,通过其上的控制命令发送接口8传送控制命令数据10到被监控方1上的控制命令接受和执行程序9;

(5) 被监控方1上的控制命令接受和执行程序9根据所接受的控制命令,采取相应的处理。如果是运行已经下载的程序或系统命令,则直接控制执行;如果是需要下载文件,则调用被监控方1上的FTP客户端程序11下载文件;

(6) 被监控方1上的FTP客户端程序11接受来自控制命令接受和执行程序9的下载命令17后,发送文件访问请求15到监控方4上的FTP服务器程序12,FTP服务器程序12从文件和控制命令库13中取得文件数据16,将其传递给FTP客户端程序11,FTP客户端程序11在完成文件传输后,返回结果给控制命令接受和执行程序9;

(7) 控制命令接受和执行程序9接受FTP客

户端程序 11 的返回结果以及其直接控制运行程序的返回结果,通过监控方 4 上的控制命令发送接口 8 返回结果给中心分析和控制 7。

这样,就实现了从信息的订购、采集、分析处理、故障发现到故障排除的自动化。当然,这只是一种实现方式,根据具体的应用环境还会有其它的实现方式。如不采用 FTP 协议传递文件等。

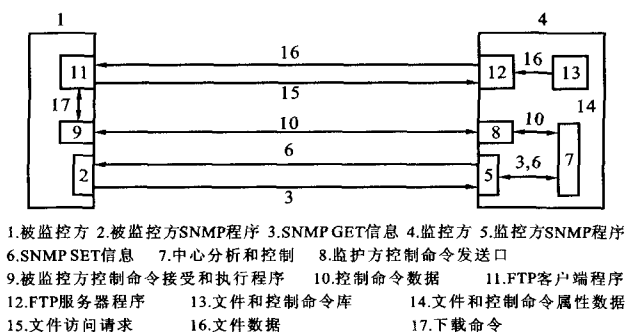


图3 具有自恢复能力的机群监控系统实现

4 结束语

文中提出的具有自恢复能力的机群监控系统结构是基于动态、主动监视和控制的设计思路,不同于传统的静态配置和管理的设计思想,可以极大提高机群监控系统的主动控制、精确控制和升级维护能力以及整个机群系统的可靠性。采用该结构的机群监控系统已在实际的 Linux 机群存储系统中得到应用,所开发的系统提供了良好的接口,允许用户动态

添加新的信息处理程序和故障处理程序,具有良好的可扩展性。

参考文献:

- [1] 武林平,孟丹,詹剑锋,等. 机群系统可靠性研究的新思路—主动故障检测机制研究[EB/OL]. (2007)[2010-08-13]. <http://www.docin.com/p-16397101.html>.
- [2] Treaster M. A survey of fault-tolerance and fault-recovery techniques in parallel systems. cs DC/0501002 [R]. CORR; National Center for Supercomputing Applications of University of Illinois, 2005.
- [3] 张洋,陈文波,李廉. 基于 SCMSWEB 的网络资源监控研究与实现[J]. 微电子学与计算机, 2008, 25(10): 166—168.
- [4] 李明江. SNMP 简单网络管理协议[M]. 北京:电子工业出版社, 2007.
- [5] 李娟,毛宏燕,马光思. SNMP v3 数据采集器 DCP 的研究与实现[J]. 微电子学与计算机, 2008, 25(12): 137—140.
- [6] Sun Jimeng, Hoke Evan, Strunk John D, et al. Intelligent system monitoring on large clusters[C]// Proceedings of the 3rd Workshop on Data Management for Sensor Networks; in Conjunction with VLDB 2006. NY, USA: ACM, 2006: 47—52.

作者简介:

范中磊 男, (1970—), 博士, 教授. 研究方向为分布式计算与存储。

(上接第 78 页)

- [5] Tan L, Gong Y, Chen G. A balance parallel clustering protocol for wireless sensor networks using K-means techniques[C]// Proceedings of the 2008 Second International Conference on Sensor Technologies and Application. Washington, DC, USA: IEEE, 2008: 300—305.
- [6] Manjeshwar A, Agrawal D P. Teen: A routing protocol for enhanced efficiency in wireless sensor networks [C]// Proc 1st International Workshop on Parallel and Distributed Computing Issues in Wireless Networks and Mobile computing. San Francisco, California, USA: IEEE, 2001: 189—196.
- [7] Zhao X, Makki S K, Pissinou N. Determining the routing path for in-networking aggregation[C]// Proc of the IEEE ICHIT. Cheju Island: IEEE, 2006: 318—327.
- [8] Yang S, Dai F, Cardei M, etc. On Multiple Point Coverage in Wireless Sensor[C]. IEEE Conf on Mobile Adhoc and Sensor Systems. Washington, DC, USA: IEEE, 2005: 757—764.

- [9] Liu Y, Liang W. Approximate coverage in wireless sensor[C]// IEEE Conf on Local Computer Networks 30th Anniversary. Washington, DC, USA: IEEE Computer Society, 2005: 68—75.
- [10] 刘明, 龚海刚, 毛莺池, 等. 高效节能的传感器网络数据收集和聚合协议[J]. 软件学报, 2005, 16(12): 2106—2116.
- [11] Yang J, Xu M, Xu B. A multipath routing algorithm based on parametric probability for wireless sensor networks[C]// International Conf on Intelligent Computation Technology and Automation. Washington, DC, USA: IEEE Computer Society, 2009: 417—420.

作者简介:

常波 女, (1970—), 讲师, 硕士. 研究方向为无线通信、自组织网络。

徐保国 男, (1950—), 教授, 博导. 研究方向为无线传感器网络及其控制、工业过程控制与优化、智能仪器及现场总线技术。