

一种基于 SNMP 协议的集群设备实时监控系統

张程焯

(中国人民解放军 91872 部队 北京 102442)

摘要: 针对大规模集群系统中通用性管理监控手段不足的现状,本文设计实现了一种基于 SNMP 的集群设备实时状态监控系统。该系统主要功能包括:设备状态查询、异常报警和报警信息接收等,有效避免了集群系统中硬件接口不一致造成的管理困难等问题,为集群系统提供了通用的管理机制。提出的基于 MMC 的查询与中断结合的实时监控策略,在实现对设备有效监控及维护的同时,降低了监控软件设计实现的复杂度。测试结果表明,在 eCos 环境下运用 SNMP 协议对集群设备进行监控,具有较高的时效性和可靠性。

关键词: 简单网络管理协议;集群设备;状态监控;异常报警

中图分类号: TP393.1 **文献标识码:** A **国家标准学科分类代码:** 520.3040

SNMP protocol based real-time monitoring system of cluster equipment

Zhang Chengye

(PLA No. 91872, Beijing 102442, China)

Abstract: Aiming at the lack of versatility method on management and monitoring in cluster system, a SNMP protocol based real-time monitoring system of cluster equipment is proposed. Functions such as equipment status checking, exception alarming and alarm information receiving are integrated into a general management mechanism which can effectively avoid the problem on inconsistency of the hardware interface in the cluster system. A real-time monitoring strategy combined of checking and interruption is proposed to reduce the complexity of the monitoring system design. The test results show that the timeliness and reliability of the system is desirable.

Keywords: SNMP; cluster equipment; status monitor; exception alarm

1 引言

随着社会发展和信息技术的不断进步,计算系统已经渗透到人类工作和日常生活的每个角落。无论是科学计算还是商业服务,应用程序对于计算性能的要求越来越高,提高单一设备性能已无法满足日益严苛的应用需求,因此,基于网络横向扩展的资源整合方式成为提高信息系统性能和存储容量的一种新的技术发展趋势。集群即对这一种互联计算系统的统称。集群由若干数量的计算机或服务器构成,节点之间相互独立,通过网络互联,对用户形成统一的应用视图。集群技术的实现使得计算机系统不再受制于地理和数量规模的限制,获取高性能和高扩展性。

近年来,集群技术在云计算及物联网等热门领域得到广泛应用,大规模计算和存储资源对集群系统的可靠性提出了新的迫切需求。集群中的设备通常由商用硬件组成,故障常态发生。为保证集群所支撑的上层应用安全、可靠、高效运作,不仅需要软硬件设备本身具备良好的稳定性,也

要求系统和管理员能够拥有灵活快速的故障排查与应变能力^[1]。

为提高集群的可靠性,目前已有许多通用集群监控管理软件能够为商用集群提供基础的资源管理。典型项目包括 Ganglia、Cacti 和 IPMI^[2],其中,Ganglia 是 UC Berkeley 发起的一个开源集群监控项目,针对 Linux 集群提供 CPU、内存、磁盘和网络流量等系统性能监控功能。IPMI 协议定义了开放标准的硬件管理接口规格,在集群中运用嵌入式硬件进行硬件管理,减轻了集群操作系统的工作负载。但 Ganglia 等工作的通用性较差,无法针对集群硬件设备的差异性实施不同的管理与监控手段,IPMI 协议具有较强的平台相关性,要求硬件设备兼容,无法成为通用的集群系统管理监控方法。

SNMP(simple network management protocol,简单网络管理协议)能够有效避免硬件接口不一致造成的管理困难,目前已被广泛应用于监测网络状态、网络设备运行情

况、各种 PC 设备以及一些辅助的外围设备^[3]。在网络远程监视和设备功能配置方面,SNMP 已经成为 Internet 事实上的网络管理协议标准。

为解决对集群网络中的大量硬件设备的监控管理问题,本文在 eCos 环境下设计实现了以 SNMP 协议为基础的实时监控策略。主要功能组件包含设备状态查询、异常报警和全系统报警信息接收模块。设备状态查询模块实现对设备的实时状态信息收集和分析,异常报警模块实现对日志记录处理和异常报警信息的发送,当状态查询模块检测到异常时,调用异常报警模块将 SNMP 报警信息发送至系统级报警信息接收模块,进行异常信息的接收和管理。通过模拟上述监控方案在本地集群网络中的运行,对其时效性与可靠性进行了验证。

2 SNMP 简单网络管理协议

2.1 SNMP 概述

随着互联网应用的拓展,提高网络运行效率、增强网络可靠性日趋重要。在网络远程监视和设备功能配置方面,SNMP 已经成为 Internet 事实上的网络管理协议标准。

SNMP 由一组网络管理的标准组成,架构在管理基站(NMS, network management station)和代理进程(agent)之间,能够支持网络管理系统,用以检测网络设备是否有管理异常情况。管理员通过 SNMP 能够查询设备信息、监控设备状态、自动发现故障并生成报告。

SNMP 简单网络管理协议主要由三部分组成:管理信

息结构 SMI(structure of management information)、管理信息库 MIB(management information base)与管理协议。其中 SMI 是描述管理信息标准符号,定义了 MIB 中的数据类型及网络资源的名称或表示。MIB 是网络管理协议访问的管理对象数据库,包含待管理的各类设备的信息变量。MIB 为每个管理对象说明具体对象实例,并为每个实例绑定一个值,具体实现采用树形结构^[4]。SNMP 管理协议,是管理基站和代理之间的异步请求和响应协议。其基本功能是获取、设置和接收意外报警信息。

2.2 SNMP 工作原理

SNMP 采用 C/S 模型,主要使用 UDP 协议作为传输层协议^[5],对网络的管理和维护通过管理基站和 SNMP 代理之间的交互完成^[6],该交互可采用两种方式:

- 1) 管理基站对每一个代理进行查询,每个代理回答管理信息库定义信息的各种查询;
- 2) 由代理产生 trap 自陷信息,向管理基站通报。

如图 1,在 SNMP 管理框架中,管理进程运行在管理基站 NMS 上,向代理发送指令并接收代理返回的响应,收集所有被管实体设备的状态和性能信息,并确定设备及网络状态。被管网络实体上运行着管理代理,相对应地,其功能为倾听和接收管理进程指令和请求,主动向 NMS 报告设备异常情况,以及向 NMS 发送代理的参数配置信息和运行状态^[7]。管理进程与代理进程中间利用 SNMP 报文实现通信,在互连网络中,用于传送 SNMP 报文的网络协议则是 TCP/IP 协议^[8]。

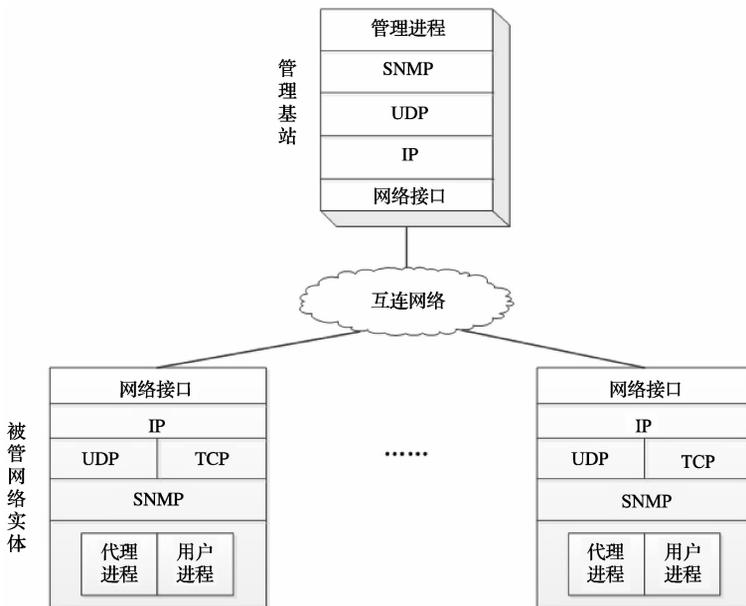


图 1 SNMP 网络管理模型

3 集群设备实时监控软件的设计实现

3.1 系统结构分析

集群架构中包含:网络互联结构、计算节点及系统层高性能客户机、监控管理子系统、应用层等^[9]。其中,监控管理子系统通常采用分布式集中管理结构,监控对象包括主机系统、存储设备、电源供电系统、制冷系统等,能够实现全系统的实时安全监测、系统控制和调试诊断等功能^[10]。

状态监控管理硬件平台主要包括系统监控控制单元 SMU(system manage unit)、系统监控管理控制端 MMC (monitor manage controller)和监控网络三大部分,其硬件架构如图 2 所示。SMU 通常采用商用服务器,实现全系统所有设备的信息采集和状态监测以及系统控制等功能。MMC 上的监控软件实现了对机框/服务器内所有设备的管理与监测。本文所提出的实时状态监控软件主要是基于 MMC 来实现的。

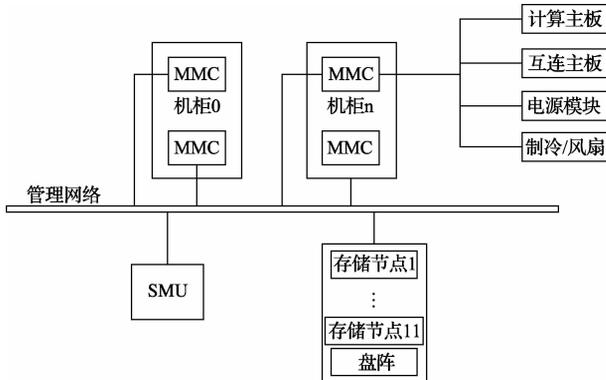


图 2 监控系统硬件平台

集群由若干个功能不同的机框组成,每个机框包含多个节点,每个节点均配置 MMC。MMC 由嵌入式 CPU、内存和 FlashROM 等组成,运行 eCos 嵌入式操作系统。MMC 与节点主板上的各类设备进行互连,对集群的全系统实施实时状态监测和控制,虽然监测的设备对象略有不同,但原理和功能类似。

MMC 和 SMC 分别对应 SNMP 管理模型中的代理和管理基站。在监控软件运行过程中,MMC 与 SMU 进行通信,上报机框内各种资源的工作状态,同时接收来自 SMU 的各种系统控制命令。

3.2 MMC 状态监控策略

从系统资源管理的角度看,集群由多个功能不同的子系统组成。每个子系统包含多个机框,机框类型不同,构成部件不同,MMC 所监控的主板类型和数量也不同。虽然每个子系统的 MMC 完成了不同的工作,但 MMC 的基本结构和功能是相同的,因此本文设计实现的状态监控软件采用相同的策略。# MMC 采用查询与中断相结合的实

时监控策略,即在每一轮监控过程中,MMC 依次查询每一块主板上所有传感器的当前状态,而在查询两块主板状态之间以及整个机框查询结束后监控进程被挂起的时间段内检测是否有与主板状态相关的中断出现。查询一块主板状态的操作是原子的,它保证状态区中同一块主板上所有传感器的值是在同一个采样时间段上获取的。MMC 收到的中断包括各主板上报的电压、温度等异常,以及硬件和命令行引起的系统(和主板)加/切电、复位操作。

图 3 为 MMC 采用的异常处理流程示意图。在对主板传感器值的查询过程中,MMC 状态监控软件将所有查询所得的值记录在一个公共状态区中,如果所得的数值正常,则将继续执行监控过程,但若数值异常,MMC 将对事件进行记录,并向 SMU 发送报警信号,报警消息的发送采用 SNMP 协议。

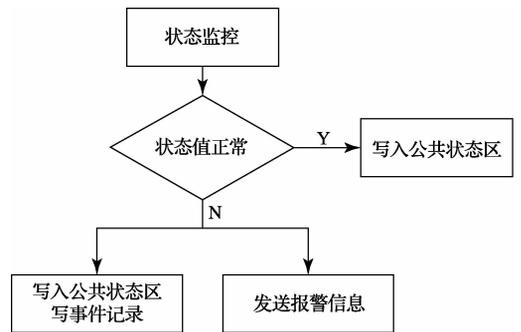


图 3 监控状态值判定处理

3.3 MMC 报警机制

报警信息的发送依赖于 SNMP trap 实现。trap 报文中可以定义报警信息的来源及其他基本特征。trap PDU 的报文构成^[11]示意图如图 4。

PDU	enterprise	agent-addr	generic-trap	specific-trap	time-stamp	var_list	...
-----	------------	------------	--------------	---------------	------------	----------	-----

图 4 trap PDU 格式

PDU 报文各字段分别标示了产生报文的系统设备 OID、设备地址、trap 类型、时间戳等。其中, var-list 为 PDU 绑定的变量^[12],在 MMC 报警机制的具体实现中,可以利用 snmp_add_var() 接口函数^[14] 自定义 var-list 所指向的 trap 报警信息^[13]。

本文基于 SNMP API 实现了 MMC 上的异常报警功能,并利用 snmp trap 接口实现了一个发送 trap 报文的接口函数 send_trap(),其具体执行步骤为:

- 1) 构建并初始化一个 snmp 会话;
- 2) 在会话中创建 trap PDU,并向该 PDU 内填充相应的信息,包括传感器 id,传感器值等信息;
- 3) 发送 trap 报文;
- 4) 关闭会话,进行清理工作。

打开 SNMP session 会话后,运行在 MMC 上的

SNMP agent 可以通过该会话向 SMU 发送 trap 信息,也可以执行相应 get/set 等操作。发送 trap 信息需对相应的 PDU 进行填充。具体实现过程如下:

上述即为 MMC 发送报警信息的主要过程。本文实现了 trap 触发条件和 trap 变量信息的自定义来构成系统的报警机制。在具体应用中,报警信息通过 162 号端口发送。

3.4 报警信息接收

MMC 在发现异常状态后,将根据预先设定的配置信息,向 SMU 发送 trap 报文。SMU 利用 snmptrapd 服务程序接收 trap 信息,并对接收到的报文进行解析。SMU 接收报警信息后,需要进行异常事件处理,本文通过向 snmptrapd 中添加 traphandle 以实现 trap 信息的处理。示例如下:

```
traphandle SNMPv2-MIB::coldStart
/usr/nba/bin/traps cold
```

traphandle 的第一个参数为 trap 报警信息中包含的设备 OID,第二个参数为调用的处理程序。

除 MMC 外,SMU 所接收到的报警信息来源还包括:存储节点的服务器异常、文件系统故障、盘阵异常;或空调制冷模块中出/入风口温度异常、进/回水温度异常、风扇转速异常等^[15]。在集群系统中发送 trap 报文的设备有多种,每个设备产生 trap 报文的传感器也有多个,因此在 snmptrap.conf 中对每个 OID 进行配置会比较繁琐,本文利用 default 功能来处理所有设备的所有 snmp trap。配置如下:

```
traphandle default
/usr/local/sbin/trapreceiver
```

trap receiver 能够处理所有的异常信息,具体流程如下。

- 1) 对 trap PDU 进行解析,解析包括 OID、设备 IP 地址、传感器编号,标志位等报文内容;
- 2) 进行事件记录,将事件信息写入数据库;
- 3) 调用相关的处理程序或命令行进行异常处理,如切电;
- 4) 将报警信息广播到客户端。

4 实验与分析

基于 SNMP 的集群设备实时状态监控策略通过对集群的监控管理子系统进行 SNMP 协议部署及自定义,实现了集群监控芯片 SMU 对管理代理 MMC 的周期性轮询监控,以及 MMC 主动报警机制。

如图 5 所示,通过图形化的 trap 接收程序检验 SNMP trap 报警信息的收发情况^[16],证实报警信息收发有效。

Trap 序号	类型	Trap 名称	时间戳	值 1	值 2	值 3	Trap 版本
202:197.30.23	Trap 消息		0:00:12.87	2.2.2.2	trap_send_success!		snmpV2
202:197.30.23	Trap 消息		0:00:04.30	2.2.2.2	trap_send_success!		snmpV2
202:197.30.23	Trap 消息		0:00:21.81	2.2.2.2	trap_send_success!		snmpV2
202:197.30.23	Trap 消息		0:00:13.29	2.2.2.2			snmpV2
202:197.30.23	Trap 消息		0:00:14.16				snmpV2
202:197.30.23	Trap 消息		0:00:11.19				snmpV2
202:197.30.23	Trap 消息		0:00:04.32				snmpV2
202:197.30.23	Trap 消息		0:01:53.76				snmpV2
202:197.30.23	Trap 消息		0:01:50.26				snmpV2

图 5 监控报警信息接收

实验集群中实现了对集群全系统设备的异常监控和管理功能,主要的监控对象如表 1 所示。其中,适配多设备的报警信息收发模块主要通过对 trap PDU 的扩展以及 snmptrapd 的自定义配置。相比于目前通用的集群管理软件,本文通过自定义设备报警信息,使监控管理策略更具通用性。

表 1 主要监控功能

编号	监控值	异常处理方式
1	主板电源电压	Power off
2	主板 CPU 温度	Power off
3	主板风扇转速异常	Alarm
4	空调出入风口温度	Alarm
5	盘阵异常	Power off
6	网络链路异常	Power off

为进一步论证本文所提出的状态监控策略的实时性和可靠性,本文通过手动模拟设备异常,记录 trap 报警信息的生成时间,并计算异常事件的误报和漏报率。

误报,即设备本身并未发生异常时 MMC 发送了 trap 报警信息,误报率由误报事件数/实际异常事件数计算得到。漏报,即设备本身出现异常而 MMC 并未报警,漏报率由漏报事件数/实际异常事件总数计算到。

如图 7,在实验集群运行的 24 小时中每隔半小时随机生成了共计 48 个异常事件,SMU 收到 50 条来自 MMC 的 trap 报警信息。通过对系统操作日志中传感器实时数值记录的排查,发现实际异常事件总数为 52 个,其中 50 个与 SMU 收到的报警信息吻合,且 50 条报警信息中包含了全部的 48 个手动设置异常事件。实验中额外产生的 4 个异常事件,可能是由于某些异常事件的联动性造成的,比如异常处理时造成的节点 CPU 过载引起温度过高、或者设置空调入风口温度异常时导致机柜内风扇转速异常等等,其中 2 个被捕获,2 个遗漏。因此,本测试中 SNMP 集群实时状态监控系统的误报率为 0%,漏报率为 3.85%。

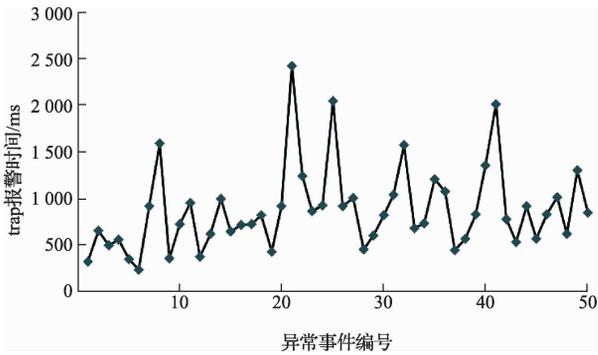


图 7 监控报警信息接收

对于每一次 trap 报警信息,其平均报警时间为约为 861 ms,实时性较高。在图 7 中出现陡峰处所指示的异常事件日志中,其报警信息的长延时主要由于网络链路状况不佳,导致报警信息的传输链路拥塞。

5 结 论

SNMP 网络管理协议在众多网络设备中都有应用,能够在集群系统中实现对设备的报警监控。本文以 eCos 环境下基于 SNMP 协议的报警和监控为主线,围绕 SNMP trap 报警信息的发送与接收处理进行了讨论,基于集群系统 MMC 和 SMU 硬件架构特点进行了相关实验。主要概述了 SNMP 简单管理协议,实现 eCos 对 SNMP 协议的支持,设计实现了系统监控管理控制端 MMC 上查询与中断相结合的实时监控策略、MMC 报警信息的发送及 SMU 上报警信息接收。

测试表明,本文所提出的基于 SNMP 的集群设备实时监控系统的能够适应多设备的监控管理需求,有效运行报警机制接收处理,具备通用性、实时性和可靠性优势。

参考文献

- [1] 赵毅. Linux 集群体系结构[D]. 青岛:青岛大学, 2009.
- [2] 于之平. 服务器集群故障预警技术的研究与实现[D]. 西安:西安电子科技大学, 2014.

- [3] 徐晶晶. 基于嵌入式 SNMP 网管代理系统的设计及实现[D]. 武汉:武汉理工大学, 2008.
- [4] 吴怡风, 归强, 罗明宇, 等. 集群计算机监控技术研究[J]. 计算机与现代化, 2013, 1(11): 218-222.
- [5] 刘荷花. 局域网络监视系统的开发与实现[J]. 电子测量技术, 2007, 10: 115-118.
- [6] 常文娜, 郭宇春. SNMP 安全机制的实现与发展[J]. 网络安全技术与应用, 2007, 5(3): 14-16.
- [7] 哈内德. 简单网络管理协议教程[M]. 北京:电子工业出版社, 1999.
- [8] 张飏, 陆继剑. SNMP 协议解析及综述[J]. 曲靖师范学院学报, 2012, 31(6): 86-89.
- [9] 潘虎. 网络集群管理技术应用研究[J]. 福建电脑, 2010, 26(1): 33-34.
- [10] 张建军, 田晓, 张本宏, 等. OSEKNM 机制在 FlexRay 网络中的研究与实现[J]. 电子测量与仪器学报, 2012, 12: 1037-1042.
- [11] 蒋句平. 嵌入式可配置实时操作系统 eCos 开发与应用[M]. 机械工业出版社, 2004.
- [12] 戴尔哈, 孙海安, 王汝传. SNMPOverFieldbus 技术研究[J]. 国外电子测量技术, 2005, 11: 15-18.
- [13] 胡方炜, 张丹, 杨扬. Linux 环境下基于 ucd-snmp 的 SNMP 代理开发[J]. 通信技术, 2003, 26(6): 66-68.
- [14] 付强, 王玥. 基于 NET-SNMP 的代理扩展开发[J]. 工业控制计算机, 2011, 24(7): 56-57.
- [15] 熊齐, 唐佳明. Linux 集群实时监控系统的实现方法[J]. 计算机系统应用, 2013, 22(9): 50-53.
- [16] 肖鹏, 李媛媛, 王云鹏. 基于 Linux 集群的分布式进程通信系统的实现[J]. 仪器仪表学报, 2006, S1: 678-680.

作者简介

张程焯, 1987 年出生, 工程师, 硕士, 主要研究领域计算机可靠性。