

基于 Web 的集群管理系统结构^①

刘芳¹, 吴琼², 詹丽丽³

¹(东北石油大学 计算机与信息技术学院, 大庆 163318)

²(大庆油田有限责任公司勘探开发研究院, 大庆 163318)

³(哈尔滨学院, 哈尔滨 150086)

摘要: 介绍高性能计算集群管理现状, 并根据目前管理缺点与不足, 对集群管理系统进行了总体框架设计, 利用 Linux 中 /proc 文件系统、MySQL 数据库及 Web 技术, 对管理数据采集、数据存储及应用接口进行了功能设计, 最终实现了在统一 Web 界面中对集群系统运行实时情况、历史数据统计与分析进行管理。

关键词: 集群管理系统; 高性能计算; 管理数据采集; 数据统计; 集群 web 管理

Cluster Management System Based on Web

LIU Fang¹, WU Qiong², ZHAN Li-Li³

¹(Computer and Information Technology, Northeast Petroleum University, Daqing 163318, China)

²(Exploration and Development Research Institute, Daqing Oilfield Co. Ltd, Daqing 163318, China)

³(Harbin University, Harbin 150086, China)

Abstract: The paper introduced the high performance computer cluster management situation, according to the shortages of current management, the cluster management system overall framework was designed. Using the Proc file system in Linux, MySQL database and Web technology, the function of managing data acquisition, data storage and application interface was designed. Finally the paper achieved the management in the unified Web interface of the cluster system real-time running, historical data statistics and analysis.

Key words: cluster management system; high-performance computing; management data collection; data statistics; cluster web management

随着高性能计算产业的不断发展, 高性能计算集群的应用越来越普及。高性能计算系统采用标准部件构成集群系统, 极大的提高了高性能计算系统的性价比, 使集群系统的应用逐渐从科学研究发展到社会的各个领域^[1-3]。集群系统应用是集成的, 但管理却是分散的, 随着集群应用规模扩大, 集群系统中个节点运行状态管理、历史数据统计分析等工作的难度和工作量也在逐渐增加, 使得集群管理成本大大提高。此外, 在很多应用中, 不同厂商的集群系统共同为一个应用提供服务, 涉及大多套集群系统的管理与维护。因此为提高集群管理效率, 减低管理成本, 需要为集群系统提供一个灵活高效统一的管理平台。

1 集群管理现状

在集群管理方面, 国内外很多厂商及研究人员提供了一些集群系统管理解决方案。如惠普的 CMU 提供了集群管理平台, 在该平台中可以进行集群节点状态管理、系统克隆、部署等工作; IBM 的 CSM 提供集群系统远程运行命令操作, 以命令或脚本方式运行在集群中所有节点上, 动态监控系统资源使用情况等工作^[4]; 国内^[5,6]都在集群管理方面提出了基于 Web 的集群管理解决方法, 在一定程度上改变了集群的管理模式, 提高集群管理效率。但这些方案或方法也存在一定的不足之处, 例如, 厂商提供的软件通用性差, 系统功能单一, 有些需要基于命令行的管理方式进行辅

① 基金项目: 哈尔滨青年科技创新人才专项资金(RC2013QN010001)

收稿时间: 2014-12-29; 收到修改稿时间: 2015-01-29

助, 服务器主机的运转情况影响到集群的监测管理; 可移植性差, 某些管理工具采用客户端/服务器方式设计^[7], 用户需要下载安装客户端程序, 为集群实时监控和管理带来了诸多不便; 基于 Web 的集群管理对于集群实时运行情况还缺少一种更加直观的管理方式.

2 系统设计

2.1 总体框架设计

本文将系统总体框架设计分为三层: 数据采集层、服务层和应用层, 如图 1 所示.

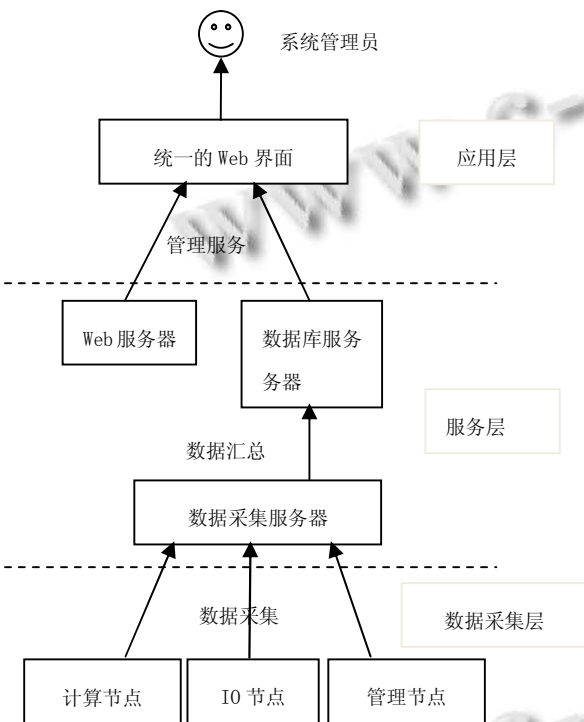


图 1 系统总体框架图

数据采集层主要目标是对集群数据进行采集, 并提供给上层数据采集服务器汇总, 由数据采集服务器进行解析. 该层的结构设计遵循客户端/服务器架构, 各个节点为客户端(计算节点、IO 节点和管理节点), 服务层的数据采集服务器为服务器端.

服务层在整体框架中承担重要角色, 以数据采集服务为基础, 将相关数据加工处理后进行汇总, 数据采集服务器负责对进行管理的节点进行数据采集, 同时对采集数据进行汇总分析, 最后汇总到数据库服务器. 该层的 Web 服务器为用户提供 Web 服务, 与数据库服务器一起为用户提供应用服务.

通过 Web 界面为用户提供统一的接口, 收集用户请求, 动态生成页面, 显示请求结果. 在该层中提供用户所需求的各种管理功能, 用户无需与被管理设备直接交互, 即可获取相关信息, 同时提供多种表现方式, 通过图形、表格、报表等形式展示用户请求.

2.2 数据采集设计

数据采集层担负着基本数据采集工作, 涉及节点多, 节点结构不统一, 操作系统多样化, 节点应用软件多样化等诸多问题. 本层设计采用多级数据采集结构, 分担数据采集和汇总压力. 如图 2 所示, 为数据采集流程图. 节点分为计算节点、IO 节点和管理节点, 包涵在节点集中. 服务器集包括多台采集服务器, 这些采集服务器根据节点种类提供采集服务. 节点不同应用的采集服务器数量和结构也是不相同的, 例如计算节点数据量多, 应用情况复杂, 其采集服务器就会根据节点应用按不同采集组进行采集汇总, 结构也如图 2 所示. 采集服务器将最终的采集结果汇总到采集汇总服务器.

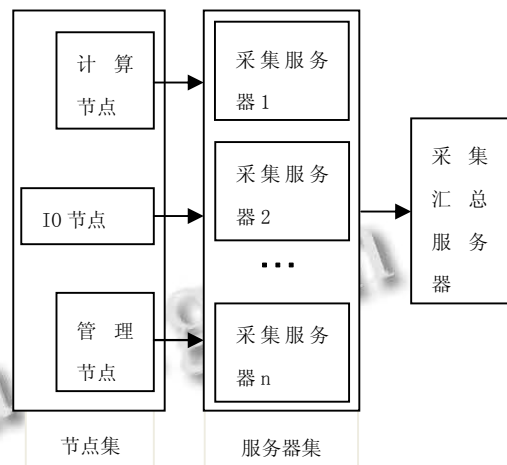


图 2 数据采集流程图

2.3 数据库设计

该层提供数据库及 Web 应用服务功能, 数据库采用 MySQL 数据库, Web 服务器采用 tomcat. 数据库表设计用于存放汇总的数据采集信息, 如表 1 所示, 系统设计的部分表名及其功能说明.

表 1 数据库表设计

序号	表名	功能
1	主机性能表	存放主机重要性能参数
2	主机 CPU 表	存放主机 CPU 信息
3	主机应用表	存放主机应用信息
4	主机状态表	存放主机当前状态

5	作业表	存放主机作业信息
6	用户表	存放用户信息
7	资产表	存放资产信息

2.4 功能设计

如图 3 所示,为集群管理系统功能模块设计图,系统分为四个主要功能模块:实时显示、系统分析、系统管理和资产管理,各功能模块有分为若干子模块。

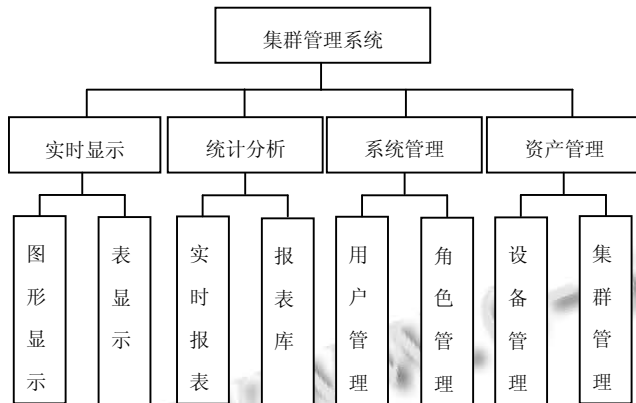


图 3 系统功能模块图

3 关键技术

3.1 采集节点数据^[8]

/proc 文件系统是 Linux 操作系统中一个虚拟文件系统,通过该文件系统可以实现 Linux 内核空间和用户空间之间进行通信,它是一种内核和内核模块用来向进程(process)发送信息的机制,管理员可以通过该文件系统获取 CPU 类、内存类、磁盘类和网络类等资源的相关数据,利用 shell 脚本技术对主机上需要采集的相关信息进行收集。

3.2 数据存储机制^[9]

XML 可扩展标记语言用于标记电子文件使其具有结构性的标记语言,是一种允许用户对自己的标记语言进行定义的源语言。与数据库不同,其数据表现形式简单,仅仅是存储数据,并且传输格式定义灵活、解析方便,对于数据采集中的不确定性有较强的适应能力。

3.3 数据传输协议^[10]

数据采集时需要对所有节点的状态进行数据收集和汇总,随着节点数量的增多,对系统的处理速度要求很高,这就要求数据包在传输过程中尽可能的减少对资源的消耗,同时加快传输的处理速度。UDP 协议属于不连接型协议,具有资源消耗小,处理速度快的优点,因此采用 UDP 协议作为采集数据传输协议。

4 系统实现

4.1 数据采集与存储

数据采集利用/proc 文件系统和 shell 脚本技术,对节点信息进行采集,而后汇聚到数据采集服务器中,在数据采集服务器上数据按着相关 XML 格式进行重新汇总,最后发送到数据库服务器。为使数据汇总信息能够与数据库对接,设计了数据库入库组件及入库流程。入库组件通过对数据采集服务器汇总的 XML 数据进行解析,使其符合数据库中对于相关数据的格式要求,如下代码为 XML 数据中存放的 CPU 相关信息。

```
<METRIC NAME="cpumetric0"
VAL="0000C04000000000DD9EBF420E397C3E00000
000000000000000000" TYPE="string" UNITS=" "
TN="0" TMAX="1200" DMAX="0" SLOPE="zero"
SOURCE="pmond">
<EXTRA_DATA>
<EXTRA_ELEMENT NAME="DESC"
VAL="cpumetric0"/>
<EXTRA_ELEMENT NAME="TITLE"
VAL="cpumetric0"/>
</EXTRA_DATA>
</METRIC>
```

将 XML 数据进行解析后成为符合数据库入库格式的数据,下面为解析器接口设计函数。

```
/* 定义 XML 文档建立与解析的接口 */
public interface XmlDocument {
/* 建立 XML 文档@param fileName 文件全路径名称*/
public void createXml(String fileName);
/* 解析 XML 文档@param fileName 文件全路径名称*/
public void parserXml(String fileName); }

```

经过解析后的数据将进入到数据库中,如图 4 所示,为数据采集入库流程。

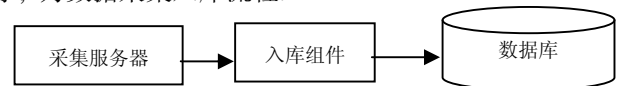


图 4 数据采集入库流程

4.2 应用显示

本系统最大特点是实现了在页面中实时显示系统运行状态。该功能在设备监控中实现,设备监控以两种方式实时显示设备的运行状态,分别是图形显示和表显示。如图 4 所示为图形显示的集群状态信息,每一个框格代表集群中的一台机器,每一台机器的管理信息设计包括 CPU 使用率、网络收发速率、磁盘读写速率、网络文件系统读写速率、内存读写速率、内存使用率、SWAP 使用率和 SWAP 实时使用率等信息。

如图 5 所示, 为表显示集群管理方式, 显示内容与图形显示相同, 但该管理方式是通过百分比(%)和数值(M/S)对每台机器的运行状态进行显示, 具体显示内容可以通过“表格配置”按钮对表格中显示内容进行修改. 该显示方式还提供高级检索方式, 管理人员可以通过过滤管理信息对集群运行状态进行显示. 以上两种方式实现在页面两秒一刷新的实时显示, 实现了在网页上动态显示集群节点运行状态的功能. 统计分析、报表库和定制报表设计为集群系统管理人员提供集群应用情况的统计分析功能. 根据统计规模大小分为按节点统计和集群统计. 在时间粒度上, 将统计设计分为半小时、一小时、今天、本周、本月、三个月、半年和一年, 这些数据全部来自于存放在数据库中获得的节点管理信息. 如图 6 所示, 为本周某集群的统计信息.

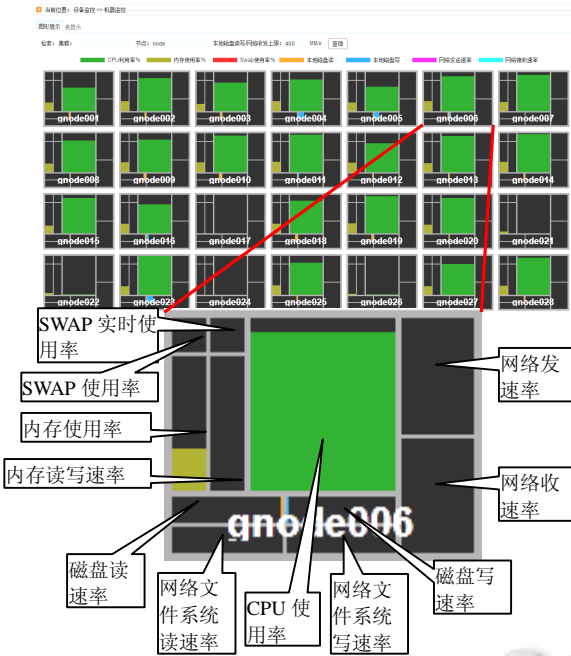


图 4 图形显示

节点	CPU使用率(%)	内存使用率(%)	内存使用量(MB)	网络收发速率(MB/s)	网络收发量(MB)
gnode001	19.20%	35.64%	5830.00	1046.00	1046.00
gnode002	34.15%	35.64%	4352.00	1120.00	1120.00
gnode004	30.88%	31.54%	4518.00	1992.00	1992.00
gnode005	38.50%	31.54%	4374.00	1398.00	1398.00
gnode007	44.67%	34.76%	4374.00	1120.00	1120.00
gnode008	35.05%	27.93%	4230.00	1190.00	1190.00
gnode010	54.41%	31.80%	8766.00	13146.00	13146.00
gnode011	37.75%	29.83%	10150.00	15884.00	15884.00
gnode012	39.85%	26.07%	6606.00	11328.00	11328.00
gnode014	38.69%	35.67%	4888.00	11680.00	11680.00
gnode018	36.23%	29.96%	4608.00	11616.00	11616.00
gnode019	36.17%	35.49%	4284.00	10682.00	10682.00
gnode021	35.25%	39.00%	3988.00	828.00	828.00
gnode022	37.58%	38.49%	8720.00	4148.00	4148.00
gnode023	34.78%	32.78%	4632.00	1884.00	1884.00
gnode024	37.44%	24.40%	3938.00	10734.00	10734.00
gnode040	37.95%	37.54%	4344.00	11460.00	11460.00
gnode041	36.10%	30.77%	4096.00	11916.00	11916.00
gnode042	37.31%	33.84%	4230.00	11958.00	11958.00
node042	38.58%	15.95%	11138000.00	11447800.00	11447800.00

图 5 表显示



图 6 某集群本周统计信息

5 结语

本文提出了基于 Web 的集群系统管理结构设计方法, 设计了系统总体框架, 实现了在页面中实时显示集群系统状态, 该方法在一个图形框格内集中显示一台主机的多个参数, 提高集群管理效率, 同时报表功能及其他功能为集群管理提供一个高效平台.

参考文献

- 1 Zima HP. Highperformance Computing. 1st edition. Springer, 2002.
- 2 Monien B. High Performance Computing-HPC 2001. 1st. edition. Springer, 2002.
- 3 Sloan JD. High Performance Linux Clusters with OSCAR, ROeks, OPenMosix, andMPI. OREILLY, 2004.
- 4 徐杰锋,舒继武,郑纬民.并行处理可视化监测环境.清华大学学报(自然科学版),2003,43(4):532~535
- 5 成 晨,刘海涛.基于 Web 的集群管理工具 CWMS.计算机辅助工程,2008,34(19):264-265,276.
- 6 冯胜鹏,郭雷.集群管理在 Web 上的设计与实现.计算机辅助工程,2006,15(1):14-17.
- 7 Liu F, Wu Q. Cluster Deployment System Based on Qt and xCAT in Linux. Advanced Materials Research.2013, 373-375: 1283-1286.
- 8 Nemeth E, Snyder G, Hein TR. Linux 系统管理员技术手册. 北京:人民邮电出版社,2007.
- 9 李刚.疯狂 XML 讲义.北京:电子工业出版社,2009.
- 10 Tanenbaum AS, Wetherall DJ.严伟.潘爱民译.计算机网络. 北京:清华大学出版社,2004.