

百万亿次高性能计算系统管理和应用的实践

杨 敏, 李惠欢

(中山大学网络与信息技术中心, 广东 广州 510275)

摘 要: 近几年, 各高校级、国家级的高性能计算中心陆续建成, 超级计算平台的运维模式也从过去强调计算力向注重降低运营管理成本和提升应用服务水平转变。中山大学经历了超过10年的行业实践, 在高性能计算平台的建设管理、平台运维服务支撑、应用研究和人才培养等诸多方面具备坚实的基础积累, 建立了随需而变的运维管理机制, 充分满足多学科应用需求的个性化应用服务模式, 具备可持续发展的行业竞争实力。

关键词: 百万亿次; 高性能计算; 国家超级计算广州中心

中图分类号: TP311 **文献标识码:** A

Practice and Application of High-Performance Computing Systems

YANG Min, LI Huihuan

(School of Network and Information Technology Center, Sun Yat-sen University, Guangzhou 510275, China)

Abstract: In recent years, with the continual establishments of national and university HPC centers, the operation and maintenance pattern of the supercomputing platform has been changed from the emphasis on computing capacity to the focus on reducing the cost of operation management and improving the level of application service. After over 10 years of practice in the captive industry, Sun Yat-sen University has solid basic accumulations in respect of construction management of the high-performance computing platform, support of operation and maintenance services as well as application research and personnel training, etc. The operation and maintenance management can be adjusted in response to demands and the personalized service mode can fully meet the needs of multidisciplinary application, with the sustainable competition capacities in the HPC field.

Keywords: MegaFLOPS; HPC; National Supercomputer Center in Guangzhou

1 引言(Introduction)

中山大学高性能计算平台的建设始于2004年底, 采取分步走、分批构建、按需建设、逐步升级的策略, 现在校内共享平台已经建成聚合计算力接近200万亿次, 存储容量达到500T, 其中GPU计算力130万亿次, 使用团队已超过170多个。高性能计算系统支撑的科学计算多数都是非计算机学科的, 对于用户来说, 系统上提供的软件资源和硬件资源服务是一体化的。高性能计算系统的主要职责即是为各领域用户提供科学计算相关的资源, 包括计算资源和应用资源。我校百万亿次高性能计算集群系统分批多次建成, 硬件数量多达四百多台, 异构复杂, 用户应用领域广泛, 用户需求各异。如何进行统一的管理, 提高运维的效率, 降低用户的使用门

槛, 为用户提供易用性强的技术支持是系统应用的难点。经过十几年的经验积累, 我校在高性能计算平台的建设管理、平台运维服务支撑、应用研究和人才培养等诸方面具备坚实的基础积累, 并于2012年与广东省政府、广州市政府、国防科学技术大学四方共同建立国家超级计算广州中心, 中心的大楼坐落于中山大学广州大学城校区内, 峰值计算性能11亿亿次的“天河二号”超级计算机系统已于2015年投入运行^[1-3]。

2 运维与管理(Maintenance and management)

我校高性能计算共享服务平台有先后建于不同时期的多套异构系统, 包括公共计算平台、GPU平台、省网格、国家网格、科研团队的专业计算平台(例如空气质量预测平台),

各类平台都集成在一起进行统一建设和管理。对于不同规模的专业应用计算平台,我们采用不同的协同服务机制。系统总体架构图如图1所示。

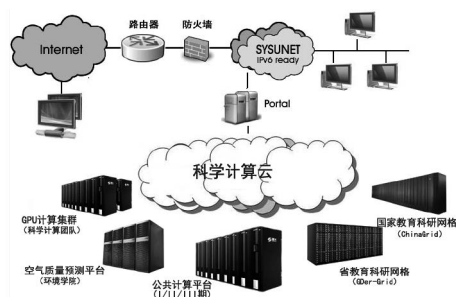


图1 系统总体架构图

Fig.1 System framework

2.1 网络监控

大规模高性能计算集群系统在管理上存在着很多的难点:(1)服务器数量多,且异构多样,专门技术人才配比少。系统一般都包括管理节点、登录节点、计算节点、I/O存储节点、光纤存储网络、高速计算网络、管理网络等。设备类型复杂,无法完全依赖人工方式完成系统运维;(2)故障类型众多,迅速故障定位要求技术专业程度高。不同类型的设备配置不一样,故障的原因和影响的范围有很大差异。单个计算节点故障只影响到提交到该节点的作业运行异常,管理节点、共享存储节点或共享计算网络故障会影响整个系统服务,导致所有用户的作业运行受到影响;(3)故障发生不可预测,要求响应处理及时。高性能计算系统需7*24小时对外提供服务,满足用户网络访问的需求,一旦发生故障必须及时通知管理员进行处理^[4,5]。

为解决以上问题,我们在开源软件ganglia和Icinga的基础上搭建了自动化的集群监控、告警运维系统。采用ganglia构建SYSU_HPC Grid网络,能够跨操作系统平台和异构硬件平台采集系统信息,持续地监测每个平台各个组件的性能,观察系统运行的情况,检测资源利用的分布趋势,并将这些信息有效地汇总和呈现,这些信息可以作为资源管理策略制定、系统架构优化和新设备的选型的依据。同时集成Icinga的告警功能,能够在问题发生之前,性能指标达到一定的阈值就发出警告,而不是在问题发生之后。在性能超出预定范围时及时通知管理员进行处理,比如温湿度过高、内存不足、CPU占用太高、磁盘剩余空间不够、ssh服务异常、网络异常等等。发出警告的途径可以由用户设置,包括MAIL、SMS、微信等方式。不同的告警级别使用不同的告警方式,一般warning警告可使用MAIL、微信告警,在网络不可用的情况可以通过SMS发短信告警。使用自动化的监控告警,使集群系统保持稳定、高效的运行状态,提高平台服务质量^[6,7]。运维管理结构图如图2所示。

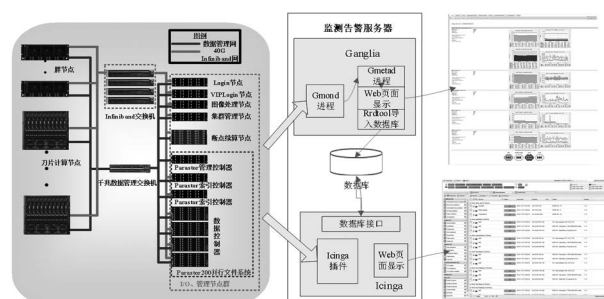


图2 运维管理结构图

Fig.2 Operation and maintenance management structure

根据监测到系统负载的分布情况,在整套平台中建立多种适合不同资源需求的作业调度管理策略,以作业申请的CPU核数、作业运行时间和用户优先级为维度划分多种队列,不同类型的作业根据资源需求情况自动进入相应队列,从而显著提高平台资源的利用率。

2.2 数据分析统计

为了提高平台使用的透明度,我们在开源软件作业管理调度系统TORQUE基础上进行了运维报表等统计分析软件的开发^[8]。主要对系统使用数据和用户使用数据进行分析统计,实现对多校区集群系统的用户作业运行情况的实时监测,并对作业调度及运行情况、集群平台使用情况及利用率等进行统计分析,形成了每周、每月、每年的运行情况报表,并进行及时发布。记录详细的CPU、内存、网络、文件系统、IO等信息的历史使用数据,按照用户指定的时间段以曲线图或报表的形式提供给用户,进行系统使用情况的统计;形成与作业系统整合的用户资源统计报表,可以按照管理员指定的时间段,统计指定用户的系统资源使用情况,或者以图表形式统计各个用户占用资源的比例。统计分析系统采用了自上而下的设计方法,技术规范、信息标准和系统安全保障横跨各个层面。系统架构分为五层,包括用户界面层、业务层、应用组件层、应用接口层和基础平台层,每一层都为上层提供服务,同时也从下层接口中获得服务^[9]。

最顶层为用户界面层直接面向最终用户,用户可以使用普通浏览器如IE或Chrome等作为客户端来访问。用户界面层为所有用户提供了统一登录入口,用户只需输入账号/密码则可登录。用户界面层作为单独一层,可以和高性能计算服务门户的用户界面层集成在一起,把系统的整个UI设计独立出来,使系统设计得更具人性化。

第二层为业务层,也是统计分析系统的主要功能模块集合。该层从用户和系统等多个纬度提供年度、季度、月度的使用情况统计查询和报表生成打印等功能,还可从分析结果中产生决策支持等功能。业务层以服务接口的方式为用户界面层提供服务,同时从应用组件层的相关引擎中获取服务。

第三层为应用组件层,将统计分析系统的业务进行逻辑

抽象出来的引擎组件。这些组件包括实现对TORQUE原始日志记录进行分析的数据分析引擎；实现日志文件访问处理的文件处理引擎；实现统计报表管理的报表引擎；实现用户操作日志记录的日志引擎等。应用组件层同样为上层提供通用的应用组件以构成业务功能模块。

第四层为应用接口层，负责与基础平台层进行对接和通信，为上层应用组件层提供服务，该层由各类适配器构成，使上层应用不需依赖于基础平台。

最底层为基础平台层，包括高性能计算平台上的各种基础运行环境，是整个系统的基础平台，涵盖集群的共享文件系统、作业管理调度系统TORQUE、各类并行程序运行环境等^[10]。

3 应用与服务(Application and service)

3.1 高性能计算门户

高性能计算系统的使用对一般用户，特别是非计算机专业的用户来说应用门槛仍然较高。我校通过建设高性能计算门户，将建于不同时期、分布在不同校区的多套计算资源进行有机地整合，为用户使用所有校内高性能计算资源提供一个统一的Web接口，有效地提高平台的易用性^[11,12]。

高性能计算门户框架主要由单点登录、远程应用管理、个性化用户界面定制等。

(1)单点登录

门户将建立单一登录功能，用户通过Web浏览器使用用户名/密码登录高性能计算门户后，门户将在用户的整个会话期内为用户管理证书，基于证书映射的角色管理为用户提供更加细致粒度的高性能计算资源和高性能计算服务的访问控制。

(2)远程应用管理

用户可以基于Web对高性能计算作业进行远程的在线应用管理，方便用户监控自己的作业。用户使用发布出来的在线高性能应用服务提交高性能计算作业。对于在线的应用管理分为三个部分：作业提交管理、作业详细信息查询、作业列表查询。

(3)个性化用户界面定制

给不同角色的用户提供不同的资源和服务视图，保证资源和服务的安全运行，同时也给不同用户赋予特定的应用体验。

高性能计算门户系统采用二级调度机制。用户从Web客户端提交的所有作业，都将被提交到系统服务，由系统服务根据整个环境的使用情况以前端服务为单位进行调度；而前端服务收到作业请求之后，则是根据其接入机群的情况，在局部范围内进行二次调度。网格上的调度不干涉高性能计算机群自身的调度系统策略，因此网格上只需要提供简单的FIFO和人工调度即可，复杂的调度策略由高性能计算机群

调度系统完成。各种高性能计算机群调度系统都有各自的特点，不过大多支持批量作业提交，支持回填、抢占、预留等高级调度算法。二级调度机制既保证了全局的调度机制，又给予局部管理范围一定的调度灵活性，可以很好地适应层次化管理需求。

高性能计算门户基于portlet技术构建。Web Portal是基于符合JSR-168 规范的Portlet 并融合了Ajax技术的Portal。Web Portal 由基础层、服务层和Portlets组件层构成。Web Portal体系结构如图3所示。

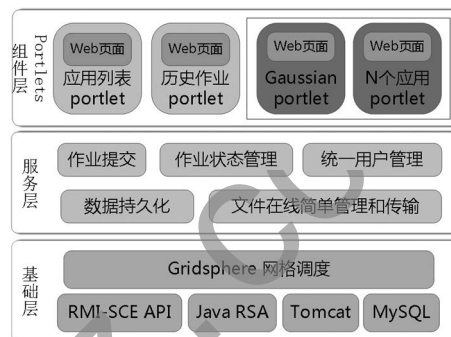


图3 计算Portal体系结构图

Fig.3 Computing Portal structure

3.2 应用服务封装

中大的高性能计算系统中除了部署高性能计算必需的编译器、数学库、并行库外，还把各科学领域用户常用的开源、商用软件分类部署到集群系统中，涉及计算化学、理论化学、理论物理、材料科学、流体力学、环境工程、气象、地理信息、生物信息等多个领域。我们通过分析主流的开源和商用大型科学与工程计算软件的应用模式，提取出科学计算软件的应用服务封装的一般化方法，实现了部分软件的基于Web技术的应用服务封装，为用户提供更加直观、易用的基于Web的科学计算软件作业提交方式^[13]。

在计算环境应用封装的基础上，在Web页面上提供应用的使用接口。Portal适合封装交互次数少、作业输入和输出数据较小的应用。在Portal上的应用封装可以分成两个部分，Web页面部分和服务部分。

(1)Web页面部分

该部分提供用户操作的界面及一些辅助功能，包括检查用户参数的完整性和合理性等功能，根据本地的cookie信息提供信息补全等功能。收集用户输入的数据并按一定的格式整理后传输到服务器。

(2)服务部分

该部分处理提交页面的请求。主要功能是接受页面的请求，通过与计算环境交互，完成相关请求并返回提示信息，如可能需要的资源列表、作业数据上传和作业提交及返回信息。

(下转第17页)