PowerLayer项目计划

**前言**

在一个Power环境里，HMC为用户提供了主机和物理资源的管理功能。但在如下两个用户场景中，我们发现了一些问题：

场景一：对一个小规模的PowerLinux主机环境 (比如5台以内的7R1)，为了满足I/O和网络性能的要求，用户希望通过直接使用物理分区 (配置物理I/O和网卡的逻辑分区) 的方式来运行他们的应用。

在这种小量低端设备的场景下，购置HMC的额外费用，将变成一个需要认真考虑的因素。如果不购置HMC，又会产生主机和物理分区的管理问题。

场景二：对一个达到一定规模的Power环境 (比如100台Power主机)，用户希望可以进行统一的管理调度。

在这个大规模的场景下，由于HMC对主机管理数量的限制，用户需要配置多台HMC来进行分组管理。这些被HMC分隔的主机组之间，无法进行资源调度。

以上两种场景，在x86的低端量化市场都比较常见，而PowerLinux要进入这个市场，就必须考虑如何解决这些场景带来的资源管理问题。

**1. 项目目标**

本项目旨在开发一个名为PowerLayer的Power物理资源综合管理工具，使得用户可以象使用虚拟资源一样方便地使用物理资源。PowerLayer将会提供类似PowerVC/OpenStack标准的REST API接口。

本项目计划将在6个月内完成。PowerLayer虽然将支持管理全线的Power服务器，但主要的应用场景是管理低端的PowerLinux主机。

项目完成后，PowerLayer的源代码，将向xCAT开发团队建议并入xCAT作为一个扩展的开源子项目。

**2. 项目背景**

最近，华胜天成发布了一个名为TEBE的大数据平台快速部署管理工具。这个TEBE工具可以通过HMC端口(FSP)管理PowerLinux的物理主机，为主机创建物理分区，并且在物理分区上安装操作系统，以及Symphony和GPFS组件。我们了解到，TEBE是基于IBM免费开源项目xCAT的最新版本进行开发的，该版本已经具备管理Power系统的初步功能。

xCAT本身是为了对大规模集群节点进行操作系统和软件的安装，配置，和维护而开发的，当它新增了对Power系统的管理能力之后，就为我们提供了除HMC之外的另一个Power物理资源管理方案。

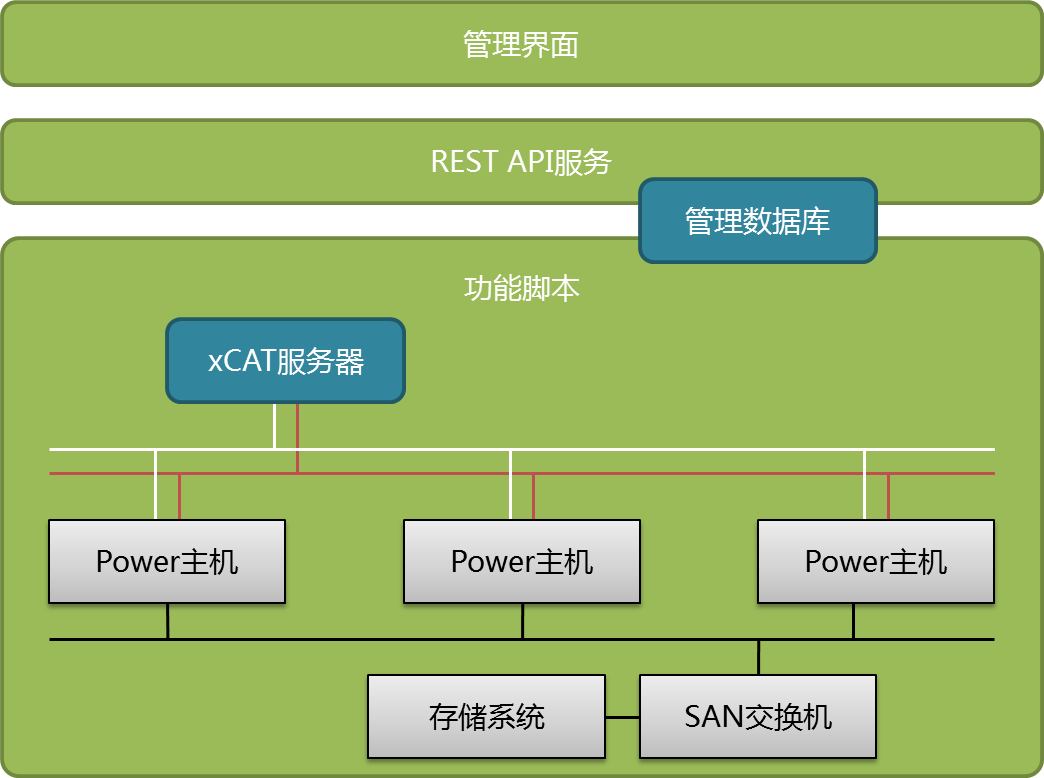
相对于HMC的管理方案，在小规模的PowerLinux环境中，xCAT可以为用户节省购置HMC所需的额外成本，在大规模的PowerLinux集群环境中，xCAT远远超过HMC在主机管理数量上的限制。简单来说，xCAT是一个十分适合低端和量化环境的管理工具。同时，华胜天成TEBE工具的成功，已经为我们验证了xCAT用于Power物理资源管理的基本可行性。

**3. 功能需求**

PowerLayer需要提供对Power物理主机和物理分区的管理功能，以及物理分区和外置存储系统的对接功能。如果条件允许，PowerLayer还将提供网络管理功能。下面的表格列出了PowerLayer管理功能的概要说明：

|  |  |
| --- | --- |
| 功能分类 | 说明 |
| 物理主机管理 | 主要功能包括：   * 发现物理主机 * 获取主机配置信息 * 发现主机上的物理分区 * 获取主机上的物理资源分配信息 * 对主机资源进行虚拟化重置 (安装VIOS) * 对主机资源进行物理化重置 (划分物理分区) * 启停主机 * 设置主机随物理分区下电 * 主机状态监控 |
| 物理分区管理 | 主要功能包括：   * 创建物理分区 * 删除物理分区 * 获取物理分区配置信息 * 调整物理分区CPU和内存配置 * 调整物理分区I/O板卡配置 * 调整物理分区网卡配置 * 配置物理分区网卡端口聚合 * 物理分区冷迁移 (关机后迁移) * 启停物理分区 * 设置物理分区随主机启动 * 捕获物理分区镜像 * 用镜像部署物理分区 * 物理分区状态监控 * 物理分区性能监控 |
| 外置存储管理 | 主要功能包括：   * 在存储设备上创建存储池 * 调整存储池的配置 * 删除存储池 * 给物理分区添加存储卷 * 从物理分区删除存储卷 * 调整存储卷容量 * 存储卷备份和恢复 * 存储系统状态监控 * 存储系统性能监控 |
| 网络管理 | 主要功能包括：   * 创建VLAN * 删除VLAN * 添加VLAN子网 * 删除VLAN子网 * 调整VLAN子网网段 * 配置物理分区VLAN IP (内网IP) * 配置物理分区公共IP (外网IP) * 网络状态监控 * 网络性能监控 |

**4. 方案设计**



PowerLayer由五个主要部分组成：

1. xCAT服务器：提供主机和物理分区管理，以及操作系统和软件的分发安装等功能。
2. 功能脚本：实现和xCAT，存储系统，网络交换机的功能对接，以及一些操作系统的shell命令调用。
3. REST API服务：调用功能脚本实现各种资源管理功能，并提供相应功能的REST调用。
4. 管理数据库：为功能脚本和REST API服务部分提供数据库功能。
5. 管理界面：调用REST API服务，实现用户操作互动。

虽然PowerLayer可以单独使用，但通常会定位在对虚拟化资源管理的补充功能，因此方案设计上主要考虑两个方面：高可用性，和安装配置的灵活性。

在高可用性方面，数据库的选择可以考虑MySQL，DB2和Oracle等流行的技术，这些数据库都提供高可用方案。xCAT本身没有正式的高可用配置，但可以通过数据备份/恢复的手段来提供一定的可用性。其他的部分主要和Web服务相关，高可用的方案也很容易实现。

在安装配置方面，数据库服务和Web服务都可以使用用户现有的环境，或者新建的实例。PowerLayer的五个主要部分可以分别单独安装配置，相互之间通过IP进行通信访问，用户可以灵活地选择各种安装组合。

**5. 设备要求**

本项目所需的开发测试环境包括如下设备：

* PowerLinux主机
* SAN交换机及外置存储系统
* 网络交换机
* xCAT管理工具

**6. 人员要求**

根据PowerLayer的5个主要部分，项目的开发将需要下列人员角色：

* 开发环境调试维护：1
* 脚本开发 (xCAT和存储功能对接)：2
* REST API封装 (Java和数据库)：2
* 管理界面 (JavaScript/CSS/HTML)：2
* 测试：2

**7. 交付内容**

本项目将交付如下条目：

* PowerLayer安装包：可以选择安装配置PowerLayer的各个功能部件
* 脚本测试工具
* REST API测试工具
* 安装维护文档
* REST API使用说明文档
* 脚本使用说明文档

**8. 项目进度**

本项目计划将在6个月 (24周) 内完成，下面是项目进度初步估算：

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 条目 | 1 | | | | 2 | | | | 3 | | | | 4 | | | | 5 | | | | 6 | | | |
| 开发测试环境搭建 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 功能设计 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| xCAT功能验证测试 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| xCAT功能对接设计 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 存储系统功能验证测试 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 存储系统功能对接设计 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| REST API设计 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 数据库设计 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 安装设计 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 管理界面设计 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 功能脚本开发 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| REST API开发 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 管理界面开发 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 安装工具开发 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 测试用例设计 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 功能脚本测试 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| REST API测试 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 管理界面测试 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 安装测试 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 综合测试 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 文档建设 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

**9. 风险评估**

本项目的风险主要有下列几个方面：

1. xCAT功能不能满足主机和物理分区的管理要求  
   这个在项目初期通过对各种用例的测试验证，可以尽早确定项目是否有效，或者要求xCAT团队实现缺失的功能。
2. 设计开发人员对xCAT的技术不熟悉  
   这个可以通过和xCAT团队沟通，寻求技术协助，比如快速培训，定期代码审核，共同开发。
3. 设计开发人员对存储系统的技术不熟悉  
   PowerLayer的存储系统基本定位在Storize接口类型的，这个可以借助ATS的DBM工具来简化开发难度。
4. 开发环境设备不足  
   目前CSC有三台7R1和一台V7000 (带SAN交换机)，是用于CAMP第二期VAD开发测试用的，目前可用于PowerLayer的项目开发。
5. PowerLayer没有足够的市场价值  
   首先，Power的虚拟化在Power用户中也并不十分普及，很多负载还是直接运行在物理资源上的，用户对通过HMC进行主机和物理分区的管理还是感到困难，PowerLayer可以在这方面提供简便操作。其次，软件开发商普遍来说对Power的体系结构没有足够的理解，在Power环境综合管理集成方面面临技术门槛过高的问题，PowerLayer可以作为一个中间件平台提供简易的集成。最后，在CAMP的新型运作模式下，华胜天成的技术力量还没到位，PowerLayer可以作为一个管理工具协助华胜天成搭建软件工厂，甚至于如果IBM选择其他代理商，PowerLayer也可以让我们掌握更多的主动性。
6. PowerLayer的开发周期过长  
   这个可以通过对人力投入的规模进行调整。目前人员估算仅限于Solution Team，可以考虑加入xCAT团队，ATS和LBS的开发力量来缩短项目周期。