**轻量级企业分布式计算系统构建方案**

dungeonsnd@gmail.com

v0.1

2015/7/3

一 背景

本方案的目的是把很多企业的空闲机器资源利用起来，构建一个轻量级的企业分布式计算系统，使企业在有限资源的情况下也能拥有较强的计算能力。

很多企业有很多PC机(实体的或虚拟的)，但是大部分机器的绝大部分时间资源(如CPU)是处于闲置状态。 比如很多企业有较多的产品线，而不同的产品线有各种机器环境，如生产环境、准生产环境、测试环境、开发环境、管理环境等。 如何把这个机器的计算能力收集并整合起来提供统一的高效计算服务，是很多企业面临的问题。显然，要实现这样的整合就需要一个简洁、轻量、可扩展和易维护的基础框架。本文即是解决此问题的一种探索。

二 方案

作者对市面上已有的分布式计算框架(如Hadoop、Storm、Spark)暂时没有深入研究学习，所以不了解能否或已经实现了本文的目标。不过作者根据实际需要，觉得要实现上述的资源整合功能，下图所描述系统架构应该是一种有效探索研究。

用户请求

**服务入口(LVS集群)**

**负责任务调度中心的负载均衡**

**服务入口(LVS集群)**

请求转入任务调度中心

**任务调度中心(JobTracker)**

**负责任务接收、切分、合并、回复等工作。 切分和合并涉及到具体任务类别，所以可以 暂不考虑。**

**任务调度中心(JobTracker)**

**结点管理中心**

**根据NodeService提供的信息对各结点的计算能力权值(固有能力权值及实时能力权值)进行统计。**

根据任务类别查找结点信息，并确定合适的处理结点来发送任务。另外，可能存储任务详情， 以便恢复任务的执行。

**结点管理中心**

存储结点信息

**缓存中心(Redis集群)**

**缓存中心(Redis集群)**

**负责服务信息和结点计算能力等临时存储及任务临时存储等高速缓存场合。**

任务转入任务结点

可能会取任务队列，或是重新执行被回滚的任务

定时汇报结点信息

**任务结点集合 (TaskTracker)**

**负责任务执行。每个结点都运行NodeService应用，负责定时上报结点相关状态和能力数据。同时注册/解注册结点的服务类别。**

**真正提供服务的模块会向缓存中心注册所能提供的服务类别(如"音频转换:amr->mp3")，并供JobTracker调用。**

**服务模块可以用任何语言来编写，只要提供服务接口供调用即可。服务模块仅需要非常简单的调用一下SDK向外注册自己的服务类型以及提供调用接口，余下的工作就是真正的业务了。**

**上线后，只要写好服务模块并上传到后台管理系统就行了。 管理系统自动传到设定的结点集合中，NodeService会主动发现并注册新加入本结点的服务模块。 卸载某个服务模块也是类似，从后台管理系统上点击卸载，NodeService会解注册并删除对应的服务模块。**

**任务结点集合 (TaskTracker)**

三