

高性能计算调度器的优化

新华三技术有限公司

刘幸 接鹏

HPC中的调度器

在高性能计算（HPC）环境中，调度器是负责管理和分配计算资源（如计算节点、处理器核心、内存等）给待执行任务的重要组件。根据系统架构和用户需求的差异，不同的HPC系统可能使用不同的调度器。Slurm是GPLv2协议的开源项目，目前是HPC领域中使用最为广泛的集群管理和作业调度系统之一。由SchedMD公司主导维护，多家大学、企业、研究机构等组织参与合作。

优点

- **灵活性：** Slurm设计灵活，可以管理各种规模和类型的集群，从小型部门集群到大型超级计算机系统。
- **高可靠性：** Slurm已经在许多大型科学计算中心和企业中得到了广泛的应用，它经过了充分的测试和性能优化。
- **可扩展性：** Slurm支持插件机制，可以轻松地集成其他工具和组件，从而增强其功能。

缺点

- **学习曲线：** 对于新用户来说，Slurm需要一些时间来学习和理解其复杂的配置和特性。
- **配置复杂性：** 尤其是在大规模集群中，Slurm的配置可能会变得复杂，需要仔细的规划和管理。

01 拓展数据库适配性

02 易用性提升

03 云原生改造

1 拓展数据库适配性

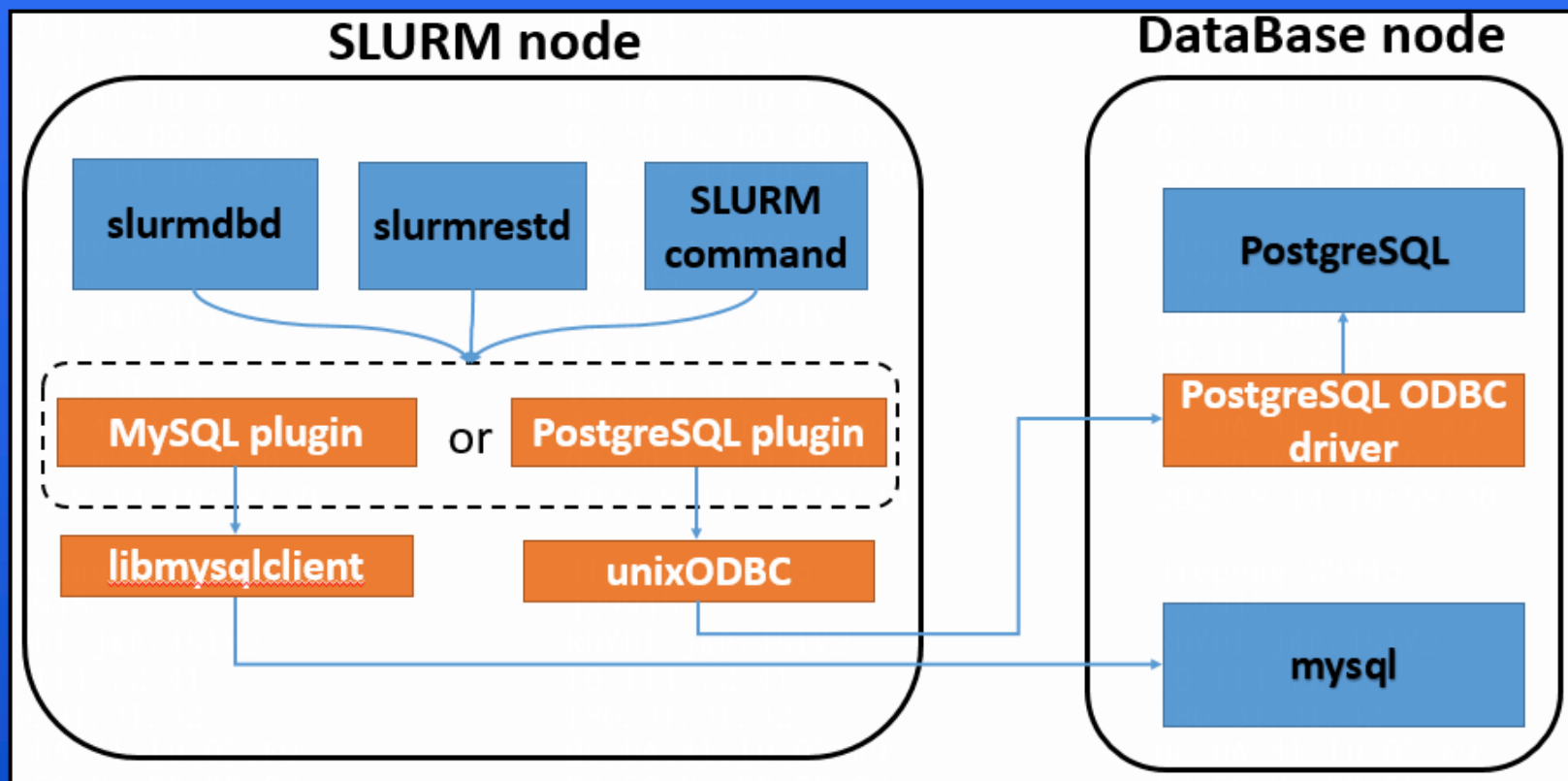
当前Slurm可对接的数据库仅有MySQL和MariaDB，而且MariaDB是在MySQL被Oracle接管后，由MySQL创始人带领团队基于MySQL的主要技术继续开发的，两者近似，所以Slurm对数据库的支持是比较有限的。另外，Slurm历史上曾支持过PostgreSQL，遗憾的是由于当时的PostgreSQL需求份额较低，SchedMD出于投入回报比不足中止了开发。因此，我们的第一步是尝试通过ODBC让Slurm支持PostgreSQL数据库。

PostgreSQL插件

我们参考Slurm的MySQL插件，新增了一个Slurm的PostgreSQL插件。该插件已经开发完成，在内部通过了测试。该插件在功能上和原生的MySQL插件一致，在性能上表现上和MySQL插件相当，对Slurm用户来说是无感知的。

该插件通过UnixODBC实现，可以轻松地扩展到其他数据库，能极大地丰富Slurm以往单一的数据库环境。

1 拓展数据库适配性



该插件通过unixODBC及PostgreSQL提供的ODBC驱动访问PostgreSQL，并向上层各个Slurm数据库操作程序提供一组接口，各Slurm进程加载该插件调用接口访问PostgreSQL。

2 易用性提升

传统Slurm的使用方式都是通过命令行，用户需要记忆繁杂的Slurm命令。近年来，随着计算管理平台的兴起，用户可以通过web方式更方便地管理集群、提交作业，无需进入后台执行复杂的命令行，极大地降低了Slurm的学习成本。

但是我们也注意到，Slurm自身提供的RESTAPI还处在尚不完善的阶段，缺失了很多命令的功能。我们计划让RESTAPI支持更多的功能，为计算管理平台提供更全面的操作选项，以此提升Slurm的易用性。

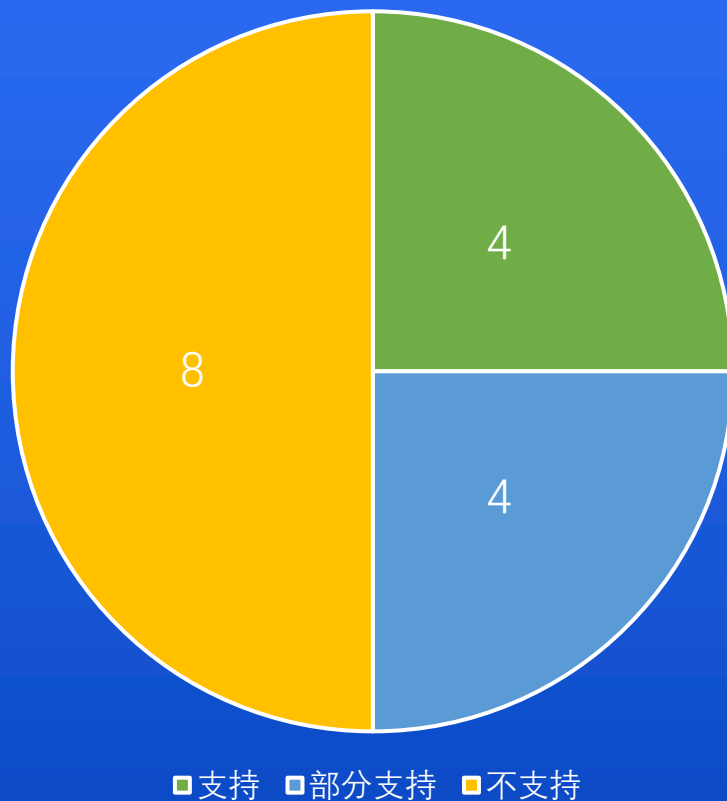
优化方案

针对scontrol命令，新增了reconfigure等RESTAPI，通过web方式就可以发送reconfigure指令，无需后台登录节点、输入Slurm命令行。

对用户来说，以往需要熟记的Slurm命令如今只需在网页上点击一个按钮即可完成。

2 易用性提升

RESTAPI对Slurm常用命令的支持情况



初步统计，在Slurm的常用命令中，当前的RESTAPI(v0039)能够完全支持的命令数量非常有限。拓展更多的RESTAPI后，将极大地提升Slurm的易用性。

3 云原生改造

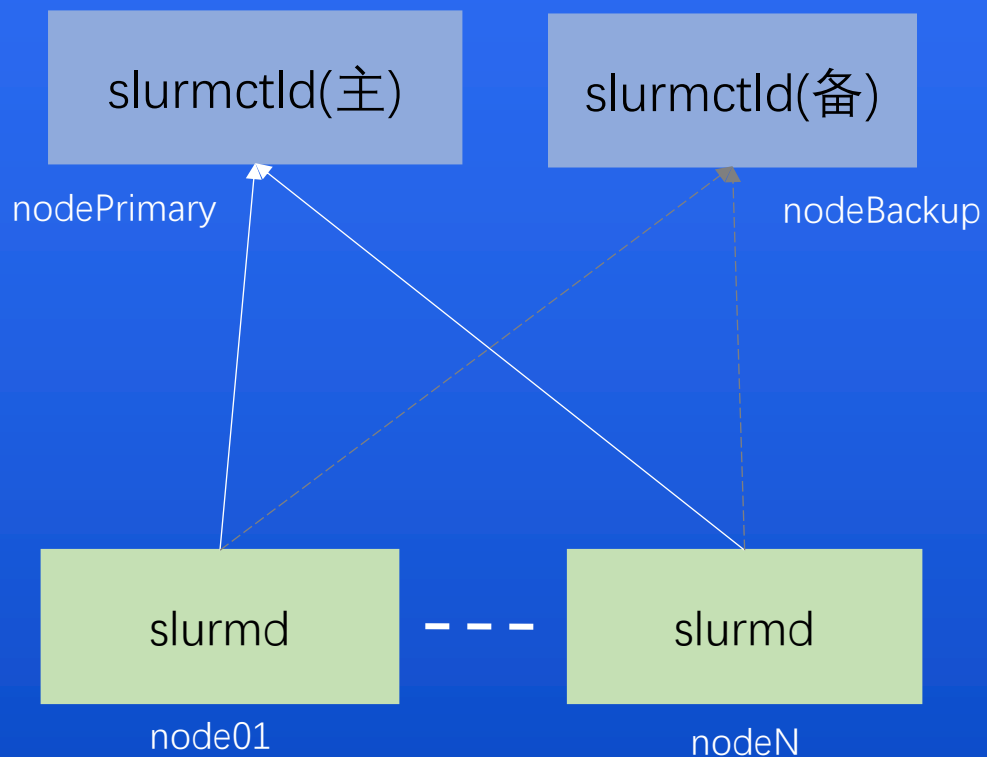
传统的Slurm部署方式是直接部署在物理机上，由于Slurm的依赖项众多，导致安装部署的过程较为繁琐，时间成本较高。而且主备服务运行在不同的物理机上，需要重复部署过程。直接运行在物理机上的Slurm服务，其高可用性也不如k8s。利用容器化技术和云原生技术，可以解决以上问题。Slurm容器化主要将Slurm的slurmctld，slurmdbd以及slurmrestd等服务部署在k8s容器中。优势在于：

- 容器隔离运行环境，可以适配多种操作系统，可移植性强
- 一键打包，简化部署
- 利用k8s高可用性保障Slurm的高可用

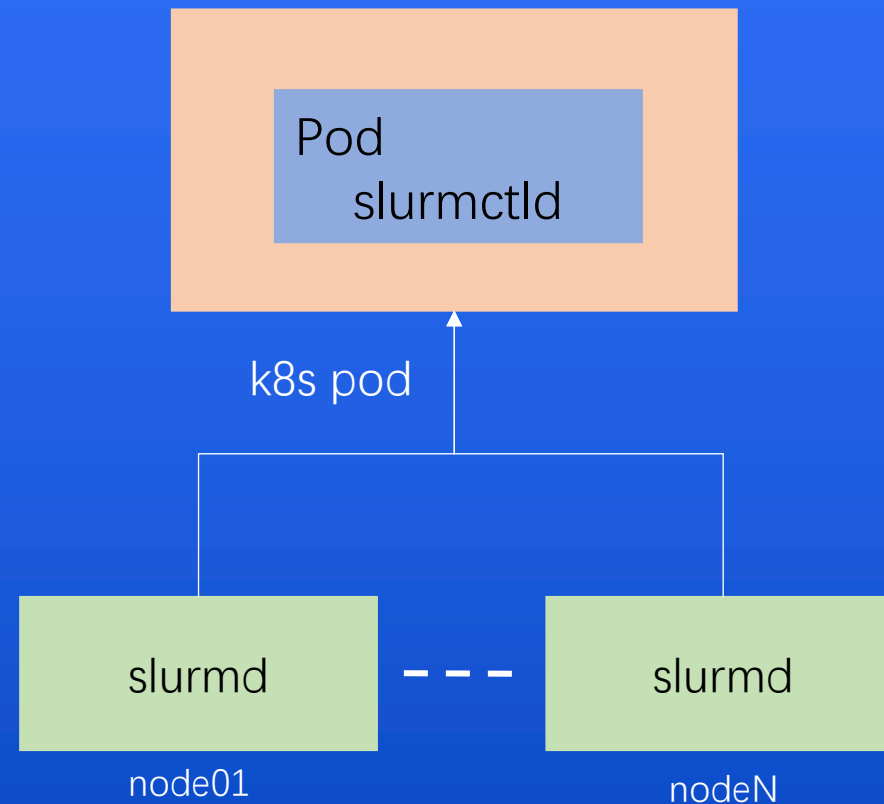
优化方案

Slurm的容器化部署，将Slurm服务运行在容器中，极大地简化了Slurm的部署流程，同时也能凭借k8s保证更高的可用性。

3 云原生改造



传统物理机方案



容器化改造方案

社区Slurm项目

欧拉社区HPC SIG Slurm项目：

<https://gitee.com/src-openeuler/slurm>

诚挚邀请所有对高性能计算感兴趣的开发者一起加入，共同贡献代码，改进功能，解决挑战！

THANKS

新华三技术有限公司

THANKS

新华三技术有限公司

THANKS

新华三技术有限公司