

面向日志搜索引擎的自动化调优

Saturday, 30 November 2019 15:30 (30 minutes)

日志对大规模分布式应用的监控和性能诊断至关重要。为了适应不同的应用场景，以 Elasticsearch 为代表的日志搜索引擎在提供存储和近实时分析功能的同时，对用户暴露了大量性能相关的配置参数。由于手动配置需要大量的专业经验和时间开销，自动化调优配置参数成为近年来的研究热点。然而，日志搜索引擎的实现高度复杂，配置参数和性能之间的关系很难通过建模直接获得；大量的配置参数所组成的高维配置空间也为求配置参数调优带来了挑战。因此，本工作针对日志搜索引擎设计了一种自动化的高维配置参数调优工具 HDConfigor：引入黑盒优化方法，HDConfigor 基于随机嵌入的思想提出了 mREMBO 算法；为了实现配置参数调优过程的自动化，在 HDConfigor 中利用 Ansible 管理工具实现了配置参数的自动化部署和性能测试任务的自动化运行。在本地 Elasticsearch 集群上的实验结果表明：在三种不同类型的工作负载下，通过给定次数的配置变更，HDConfigor 可以将 Elasticsearch 的吞吐量提高到默认配置的 2.07 倍，1.46 倍和 1.07 倍。

Primary authors: Dr 窦, 晖 (中山大学); Prof. 陈, 鹏飞 (中山大学)

Presenter: Dr 窦, 晖 (中山大学)

Session Classification: 技术报告