# 分布式追踪系统

# 项目意义及应用场景：

通过对调用链路的日志收集，实现故障的快速定位；实现各个调用环节的性能分析，快速找到性能瓶颈；通过对业务日志的可视化展示，清晰的看到系统间的依赖关系；

根据收集的业务日志进行用户行为路径分析（此功能也可参照曹庆荣的分享）。

# 系统具备的特点：

对业务系统低侵入、最小化业务系统的性能损失。

数据高效查询，秒级返回数据查询及页面绘图。

架构设计高可用。

主要由agent采集器、collector收集器、web端构成。

# 参考资料：

开源分布式追踪系统，Google Drapper、Zipkin、pinpoint、美团的CAT、Open-Falcon及其他。

阅读Google Dapper的论文，该论文是分布式追踪系统的指导思想，后期的一些开源项目都基于此。建议搭建一套开源环境，了解下依赖关系、性能分析是如何可视化的，从哪些维度展现。

# 我们的问题：

了解了这种日志收集的数据结构（也就是Dapper论文中Trace、span那一块），通过埋点，我们自己也可以实现。

也可以在开源系统上进行二次开发，只是架构设计、技术选型、页面展示，可能会有不太符合我们的地方。