

[PrestoCon 2021 超清视频和 PPT 下载](#)

PrestoCon 2021 于2021年12月09日通过在线的形式举办完了。在 PrestoCon，来自行业领先公司的用户分享了一些用例和最佳实践，Presto 开发人员讨论项目的特性；用户和开发人员将合作推进 Presto 的使用，将其作为一种高质量、高性能和可靠的软件，用于支持全球组织的分析平台，无论是在本地还是在云端。本次会议大概有 20多个议题，干货还是挺多的。



如果想及时了解Spark、Hadoop或者HBase相关的文章，欢迎关注微信公众号：过往记忆大数据

会议资料

本次会议的 PPT 和视频请关注 过往记忆大数据 公众号，并回复 10108。

会议预览

在查询速度提升方面，英特尔的工程师 Dave Cohen 介绍了异构硬件的市场变化，并提出 PrestoDB SQL查询引擎与 Velox 和 Cachelib 库结合起来来满足快速变化的需求。Meta (Facebook) 的 Ke Wang 和 Alluxio 的 Bin Fan 致力于 RaptorX 的开发，进一步提升每个查询的 cache hit rate。RaptorX 的相关文章也可以参考 [《RaptorX: 将 Presto 性能提升十倍》](#) 另外，Presto on Velox (Prestissimo) 项目也有了比较大的提升，当前 Prestissimo 支持跑 TPCH 19条 SQL，其架构如下：

在最近比较火的数据湖方面，Iceberg 数据源社区方面在添加一些比较重要的功能，比如 row-level deletion 以及 iceberg v2。在数砖的工程师支持下，目前 Presto 支持读写 Delta Lake 的数据。

在大型 ETL 支持方面，Presto on Spark 利用了 Presto 的编译器/计算库和 Spark 的大规模处理能力，使得 Presto 支持 ETL 的任务。同时这次分享中 Intuit 的工程师介绍了其 Presto on Spark 的使用以及一些改造。关于 Presto on Spark 的更多知识也可以阅读本博客的 [《Presto on Spark：通过 Spark 来扩展 Presto》](#)、[《Presto on Spark：支持即时查询和批处理》](#) 等文章。

这次会议中，来自字节（[《Presto 在字节跳动的实践》](#)）和腾讯的国内公司也介绍了其内部对 Presto 的使用场景和优化方面的事情。

过往记忆大数据后面将挑选一些比较好的 PPT 进行分享，敬请期待。

本博客文章除特别声明，全部都是原创！
原创文章版权归过往记忆大数据（[过往记忆](#)）所有，未经许可不得转载。
本文链接: [【】（）](#)