

Parameter Server for Distributed Machine Learning

- This Paper is very similar to **Scaling Distributed Machine Learning with the Parameter Server(OSDI' 2014)**. While,I find some interesting points or the points I omit while reading the one in OSDI 2014.
- **The target clusters:** It target on the real cloud computing scenarios applicable to Google, Baidu, Amazon, Microsoft where machines are possibly unreliable, jobs may get preempted, data may be lost, and where network latency and temporary workloads lead to a much more diverse performance profile, rather than the low utilization-rate, exclusive use, high performance supercomputer clusters. 这个有两个启发吧。首先是在当下2022年的情况，这些商用的clusters还有这些缺点吗？如果没有，parameter servers是否还保持优势？其次，它说这些 profiling是diverse的，那么对于DML Profiling 来说，是否可以找到找到这些diverse中的共通点，从而简化问题。
- It provides a proof for the Asynchronous and convergence guarantee: 即其证明了，在设置了 maximum delay time (a new push call will be blocked until all previous push call t time ago have been finished) 的情况下，aggregate到的gradient和真实的gradient是很好地近似的。这也为其asynchronous communication做了很好的理论证明。