作者：严林  
链接：https://www.zhihu.com/question/30026369/answer/46476717  
来源：知乎  
著作权归作者所有。商业转载请联系作者获得授权，非商业转载请注明出处。

分布式系统在互联网时代，尤其是大数据时代到来之后，成为了每个程序员的必备技能之一。分布式系统从上个世纪80年代就开始有了不少出色的研究和论文，我在这里只列举最近15年范围以内我觉得有重大影响意义的15篇论文（15 within 15）。  
1. The Google File System: 这是分布式文件系统领域划时代意义的论文，文中的多副本机制、控制流与数据流隔离和追加写模式等概念几乎成为了分布式文件系统领域的标准，其影响之深远通过其5000+的引用就可见一斑了，Apache Hadoop鼎鼎大名的HDFS就是GFS的模仿之作；  
2. MapReduce: Simplified Data Processing on Large Clusters：这篇也是Google的大作，通过Map和Reduce两个操作，大大简化了分布式计算的复杂度，使得任何需要的程序员都可以编写分布式计算程序，其中使用到的技术值得我们好好学习：简约而不简单！Hadoop也根据这篇论文做了一个开源的MapReduce；  
3. Bigtable: A Distributed Storage System for Structured Data：Google在NoSQL领域的分布式表格系统，LSM树的最好使用范例，广泛使用到了网页索引存储、YouTube数据管理等业务，Hadoop对应的开源系统叫HBase（我在前公司任职时也开发过一个相应的系统叫BladeCube，性能较HBase有数倍提升）；  
4. The Chubby lock service for loosely-coupled distributed systems：Google的分布式锁服务，基于Paxos协议，这篇文章相比于前三篇可能知道的人就少了，但是其对应的开源系统zookeeper几乎是每个后端同学都接触过，其影响力其实不亚于前三篇；  
5. Finding a Needle in Haystack: Facebook's Photo Storage：facebook的在线图片存储系统，目前来看是对小文件存储的最好解决方案之一，facebook目前通过该系统存储了超过300PB的数据，一个师兄就在这个团队工作，听过很多有意思的事情（我在前公司的时候开发过一个类似的系统pallas，不仅支持副本，还支持Reed Solomon-LRC，性能也有较多优化）；  
6. Windows Azure Storage: a highly available cloud storage service with strong consistency：windows azure的总体介绍文章，是一篇很好的描述云存储架构的论文，其中通过分层来同时保证可用性和一致性的思路在现实工作中也给了我很多启发；  
7. GraphLab: A New Framework for Parallel Machine Learning：CMU基于图计算的分布式机器学习框架，目前已经成立了专门的商业公司，在分布式机器学习上很有两把刷子，其单机版的GraphChi在百万维度的矩阵分解都只需要2~3分钟；  
8. Resilient Distributed Datasets: A Fault-Tolerant Abstraction for In-Memory Cluster Computing：其实就是 Spark，目前这两年最流行的内存计算模式，通过RDD和lineage大大简化了分布式计算框架，通常几行scala代码就可以搞定原来上千行MapReduce代码才能搞定的问题，大有取代MapReduce的趋势；  
9. Scaling Distributed Machine Learning with the Parameter Server：百度少帅李沐大作，目前大规模分布式学习各家公司主要都是使用ps，ps具备良好的可扩展性，使得大数据时代的大规模分布式学习成为可能，包括Google的深度学习模型也是通过ps训练实现，是目前最流行的分布式学习框架，豆瓣的开源系统paracell也是ps的一个实现；  
10. Dremel: Interactive Analysis of Web-Scale Datasets：Google的大规模（近）实时数据分析系统，号称可以在3秒相应1PB数据的分析请求，内部使用到了查询树来优化分析速度，其开源实现为Drill，在工业界对实时数据分析也是比价有影响力；  
11. Pregel: a system for large-scale graph processing: Google的大规模图计算系统，相当长一段时间是Google PageRank的主要计算系统，对开源的影响也很大（包括GraphLab和GraphChi）；  
12. Spanner: Google's Globally-Distributed Database：这是第一个全球意义上的分布式数据库，Google的出品。其中介绍了很多一致性方面的设计考虑，简单起见，还采用了GPS和原子钟确保时间最大误差在20ns以内，保证了事务的时间序，同样在分布式系统方面具有很强的借鉴意义；  
13. Dynamo: Amazon’s Highly Available Key-value Store：Amazon的分布式NoSQL数据库，意义相当于BigTable对于Google，于BigTable不同的是，Dynamo保证CAP中的AP，C通过vector clock做弱保证，对应的开源系统为Cassandra；  
14. S4: Distributed Stream Computing Platform：Yahoo出品的流式计算系统，目前最流行的两大流式计算系统之一（另一个是storm），Yahoo的主要广告计算平台；  
15. Storm @Twitter：这个系统不多说，开启了流式计算的新纪元，几乎是所有公司流式计算的首选，绝对值得关注；  
最近一两年时间主要精力放到了机器学习上，分布式系统的研究不太多了，现阶段就列这15篇文章吧，覆盖了分布式系统的主要领域。如果想起来有遗漏再来补充。Good luck！