

Nuevas Tendencias en Sistemas Distribuidos

Marta Patiño

mpatino@fi.upm.es

Syllabus

- Data consistency. Scalable database replication.
- Processing and Generating Large Data Sets: MapReduce.
- Distributed Storage: Bigtable
- Key-value stores: Dynamo.

Bibliography

- “Distributed Versioning: Consistent Replication for Scaling Back—end Databases of Dynamic Content Web Sites”. Cristiana Amza, Alan L. Cox and Willy Zwaenepoel. *Proceedings of the ACM/IFIP/Usenix Middleware Conference, June 2003*
- MapReduce: Simplified Data Processing on Large Clusters Jeffrey Dean and Sanjay Ghemawat. OSDI'04: Sixth Symposium on Operating System Design and Implementation, San Francisco, CA, DeCEMber, 2004.
- "The Google file system", Sanjay Ghemawat, Howard Gobioff, Shun-Tak Leung. Proceedings of the ACM SOSP '03. December 2003.
- Large-scale Incremental Processing Using Distributed Transactions and Notifications. Daniel Peng and Frank Dabek, OSDI 2010.
- Bigtable: A Distributed Storage System for Structured Data Fay Chang, Jeffrey Dean, Sanjay Ghemawat, Wilson C. Hsieh, Deborah A. Wallach, Mike Burrows, Tushar Chandra, Andrew Fikes, and Robert E. Gruber. OSDI, Seventh Symposium on Operating System Design and Implementation, November, 2006.
- Database Replication. Bettina Kemme, Ricardo Jiménez Peris, Marta Patiño-Martínez. Morgan and Claypool Publishers, 2010.
- Dynamo: Amazon’s Highly Available Key-value Store. SOSP 07.
- Chord: A Scalable Peer-to-peer Lookup Service for Internet Applications. Proceedings of SIGCOMM'01.

Evaluation criteria

- Class attendance and participation.
- Students must read and present a two page summary of all the papers presented during the lectures.
- Programming assignment (not mandatory).
- Paper presentation.