CS525 UIUC SP24: Reading List

Compiled by Prof. Indranil Gupta. http://indy.cs.illinois.edu/

Link to Main Course Website: https://courses.engr.illinois.edu/cs525/sp2024
Link to SP24 Schedule: https://go.cs.illinois.edu/CS525SP24Schedule (includes slides from presentations/discussions)
This Document: https://go.cs.illinois.edu/CS525SP24ReadingList

There are a total of **22 student sessions** (early Feb to late May -- see dates below). The presentation schedule for these can be found at the main CS525 Schedule Page (https://go.cs.illinois.edu/CS525SP24Schedule).

For each session below, there are:
1) Two "Main Papers" (Paper 1, Paper 2: these are the presentation papers), and
2) Some "Additional Readings" papers.

These papers have been carefully selected, and the schedule is not negotiable. Presenters must present the two "Main Papers (Paper 1 and Paper 2)" (they can also additionally cover the other papers listed in the session). Other students must review both of the two "Main Papers (Paper 1 and Paper 2)" (unless exceptions are mentioned for a given session).

Finally, there are more papers available on topics, which we could not cover given the limited number of weeks in SP24, at the SP23 CS525 reading list, SP21 CS525 reading list, and the SP18 CS525 reading list -- *please read those for project ideas! In particular, DO NOT limit your project ideas to only papers you read in the SP24 semester during its first few weeks.* ***Read ahead in the reading lists below, well into March and April sessions! Many interesting topics are covered later in the semester, and even one or two papers that you read ahead (from March/April sessions) before your survey report deadline can give you great project ideas!***

Table of Contents (Clickable!)

***** Indy's Lectures [1/17-1/31]

1/17 Introduction

- The mathematical theory of infectious diseases and its applications, N. T. J. Bailey, 1975 (out of print)

1/19 Before, There Were Clouds

- Historical reflections: The rise, fall, and resurrection of software as a service, Martin Campbell-Kelly, CACM, May 2009.

- [MapReduce: Simplified Data Processing on Large Clusters](#), J. Dean et al, OSDI 2004 (Google)
- [Grid: a new infrastructure for 21st century science](#), I. Foster, Physics Today, 2002.
- [A view of cloud computing](#), Michael Armbrust, Michael Armbrust , Armando Fox, Armando Fox, Rean Griffith, Rean Griffith, Anthony Douglas Joseph, Anthony D. Joseph, Randy Katz, Randy Katz, Andy D Konwinski, Andy Konwinski, Gunho Lee, Gunho Lee, David Andrew Patterson, David Patterson, Ariel Rabkin, Ariel Rabkin, Ion Stoica, Matei Zaharia, vol. 53, no. 4, CACM, Apr 2010.
- [Larry Ellison's Rant on Cloud Computing](#) (Youtube video)

1/24 P2P Systems

- [The Gnutella protocol specification v 0.4](#)
- [Chord: a scalable peer-to-peer lookup service for Internet applications](#), I. Stoica et al, SIGCOMM 2001
- [Pastry: scalable, distributed object location and routing for large-scale peer-to-peer systems](#), A. Rowstron et al, Middleware 2001.
- [Kelips: Building an Efficient and Stable P2P DHT through Increased Memory and Background Overhead](#), I. Gupta. K. Birman, P. Linga, A. Demers, R. van Renesse, IPTPS 2003 (Springer).

1/26 Key-Value Stores

- [Cassandra NoSQL Presentation](#)
- [Cassandra 10 Documentation](#)
- [Cassandra 1.0 documentation at datastax.com](#)
- [HBase](#)
- Others: [MongoDB](#)

<<< 1/26 - Last day to sign up for a presentation or scriber slot (see Course Schedule link at top of this doc) >>>

1/31 Distributed Algorithms, Sensor Networks, IoT, ML

- [Time, clocks and the ordering of events in a distributed system](#), L. Lamport, Communications ACM 1978
- [Distributed snapshots: determining global states of distributed systems](#), Chandy and Lamport, ACM TOCS 1985
- [Impossibility of distributed consensus with one faulty process](#), Fischer, Lynch and Patterson, Journal ACM 1985

- [Smart Dust](#)
- [TinyOS](#)
- [Raspberry PI](#)
- [Arduino](#)

<<< Student Presentation Topics Start >>>

***** Asynchrony [2/2]

- (Base) Orientation lecture by Indy
  - L. Lamport, [Time, Clocks, and the Ordering of Events in a Distributed System](#), CACM 1978.
  - Also: Timed Asynchrony, Partial Synchrony
- **Paper 1**: Pasindu Tennage, Cristina Basescu, Lefteris Kokoris-Kogias, Ewa Syta, Philipp Jovanovic, Vero Estrada-Galiñanes, Bryan Ford: [QuePaxa: Escaping the tyranny of timeouts in consensus](#). SOSP 2023: 281-297.
- **Paper 2**: Ali Najafi, Meta; Michael Wei, VMware Research, [Graham: Synchronizing Clocks by Leveraging Local Clock Properties](#), NSDI 2022.

  Additional Readings

- C. Dwork, N. Lynch. L .Stockmeyer, [Consensus in the Presence of Partial Synchrony](#), Journal of the Association for Computing Machinery, Vol. 35, No. 2, April 1988, pp. 288-323.
- Eric Brewer, [Spanner, TrueTime and the CAP Theorem](#), 2017.
- J. C. Corbett, J. Dean, et al, [Spanner: Google's Globally-Distributed Database](#), OSDI 2012
- [Partial synchrony based on set timeliness](#), Marcos K. Aguilera, Carole Delporte-Gallet, Hugues Fauconnier, Sam Toueg, ACM PODC 2009.
- Jeremy Elson, Deborah Estrin, [Time Synchronization for Wireless Sensor Networks](#). IPDPS 2001: 186
- [Sundial: Fault-tolerant Clock Synchronization for Datacenters](#), Yuliang Li, Gautam Kumar, Hema Hariharan, Hassan Wassel, Peter Hochschild, Dave Platt, Simon Sabato, Minlan Yu, Nandita Dukkipati, Prashant Chandra, Amin Vahdat, OSDI 2020.
- [Reliable Timekeeping for Intermittent Computing](#), Jasper de Winkel, Jasper de Winkel, Carlo Delle Donne, Carlo Delle Donne, Kasım Sinan Yıldırım, Kasim Sinan Yildirim, Przemysław Pawełczak, Przemysław Pawełczak, Josiah Hester, Josiah Hester, ASPLOS 2020.
- [Time-sensitive Intermittent Computing Meets Legacy Software](#), Vito Kortbeek, Vito Kortbeek, Kasım Sinan Yıldırım, Kasim Sinan Yildirim, Abu Bakar, Abu Bakar, Jacob

Sorber, Jacob Sorber, Josiah Hester, Josiah Hester, Przemysław Pawełczak, Przemysław Pawełczak, ASPLOS 2020

- [Time Awareness in Deep Learning-Based Multimodal Fusion Across Smartphone Platforms](#), Sandeep Singh Sandha, Joseph Noor, Fatima M. Anwar Mani Srivastava, IoTDI 2020.
- [AutoSync: Learning to Synchronize for Data-Parallel Distributed Deep Learning](#), Hao Zhang, Yuan Li, Zhijie Deng, Xiaodan Liang, Lawrence Carin, Eric Xing, NeurIPS 2020.
- [No Time for Asynchrony](#), Marcos Aguilera, Michael Walfish, Usenix HotOS, 2009.

***** Parallelism [2/7]

- (Base) Indy lecture on Machine Learning Parallelism
- **Paper 1**: Mohammad Shoeybi, Mostofa Patwary, Raul Puri, Patrick LeGresley, Jared Casper, Bryan Catanzaro, [Megatron-LM: Training Multi-Billion Parameter Language Models Using Model Parallelism](), Arxiv, 2019.
- **Paper 2**: Xupeng Miao, Yining Shi, Zhi Yang, Bin Cui, Zhihao Jia: [SDPipe: A Semi-Decentralized Framework for Heterogeneity-aware Pipeline-parallel Training](). Proc. VLDB Endow. 16(9): 2354-2363 (2023).

Additional Readings

- Weiyang Wang, Moein Khazraee, Zhizhen Zhong, Manya Ghobadi, Zhihao Jia, Dheevatsa Mudigere, Ying Zhang, Anthony Kewitsch: [TopoOpt: Co-optimizing Network Topology and Parallelization Strategy for Distributed Training Jobs](). NSDI 2023: 739-767
- [Unity: Accelerating DNN Training Through Joint Optimization of Algebraic Transformations and Parallelization](), Colin Unger, Stanford University; Zhihao Jia, Carnegie Mellon University and Meta; Wei Wu, Los Alamos National Laboratory and NVIDIA; Sina Lin, Microsoft; Mandeep Baines and Carlos Efrain Quintero Narvaez, Meta; Vinay Ramakrishnaiah, Nirmal Prajapati, Pat McCormick, and Jamaludin Mohd-Yusof, Los Alamos National Laboratory; Xi Luo, SLAC National Accelerator Laboratory; Dheevatsa Mudigere, Jongsoo Park, and Misha Smelyanskiy, Meta; Alex Aiken, Stanford University, OSDI 2022.
- Peter Kraft, Fiodar Kazhamiaka, Peter Bailis, and Matei Zaharia, [Data-Parallel Actors: A Programming Model for Scalable Query Serving Systems](), NSDI 2022
- [Alpa: Automating Inter- and Intra-Operator Parallelism for Distributed Deep Learning](), Lianmin Zheng, Zhuohan Li, and Hao Zhang, UC Berkeley; Yonghao Zhuang, Shanghai Jiao Tong University; Zhifeng Chen and Yanping Huang, Google; Yida Wang, Amazon Web Services; Yuanzhong Xu, Google; Danyang Zhuo, Duke University; Eric P. Xing, MBZUAI and Carnegie Mellon University; Joseph E. Gonzalez and Ion Stoica, UC Berkeley, OSDI 2022
- [Beyond Data and Model Parallelism for Deep Neural Networks](), Zhihao Jia, Matei Zaharia, Alex Aiken, MLSys 2019.
- [Baechi: Fast Device Placement of Machine Learning Graphs](), Beomyeol Jeon, Linda Cai, Pallavi Srivastava, Jintao Jiang, Xiaolan Ke, Yitao Meng, Cong Xie, Indranil Gupta, SoCC 2020.
- Jayashree Mohan, Amar Phanishayee, and Janardhan Kulkarni, Microsoft Research; Vijay Chidambaram, University of Texas at Austin and VMware Research, [Looking Beyond GPUs for DNN Scheduling on Multi-Tenant Clusters](), OSDI 2022.
- [Unity: Accelerating DNN Training Through Joint Optimization of Algebraic Transformations and Parallelization](), Colin Unger, Stanford University; Zhihao Jia, Carnegie Mellon University and Meta; Wei Wu, Los Alamos National Laboratory and NVIDIA; Sina Lin, Microsoft; Mandeep Baines and Carlos Efrain Quintero Narvaez, Meta; Vinay Ramakrishnaiah, Nirmal Prajapati, Pat McCormick, and Jamaludin Mohd-

Yusof, Los Alamos National Laboratory; Xi Luo, SLAC National Accelerator Laboratory; Dheevatsa Mudigere, Jongsoo Park, and Misha Smelyanskiy, Meta; Alex Aiken, Stanford University, OSDI 2022

- [Orca: A Distributed Serving System for Transformer-Based Generative Models](), Gyeong-In Yu and Joo Seong Jeong, Seoul National University; Geon-Woo Kim, FriendliAI and Seoul National University; Soojeong Kim, FriendliAI; Byung-Gon Chun, FriendliAI and Seoul National University, OSDI 2022
- [Beyond Data and Model Parallelism for Deep Neural Networks](), Zhihao Jia, Matei Zaharia, Alex Aiken, MLSys 2019.
- Microsoft's [DeepSpeed]()

(Additional Readings on ML Systems below)

- [Clipper: A Low-Latency Online Prediction Serving System](), Daniel Crankshaw et al, NSDI 2017.
- [Google TensorFlow paper]()
- [Project Adam: Building an Efficient and Scalable Deep Learning Training System](), Trishul Chilimbi, Yutaka Suzue, Johnson Apacible, and Karthik Kalyanaraman, OSDI 2014
- [Scaling Distributed Machine Learning with the Parameter Server](), Mu Li, David G. Andersen, Jun Woo Park, Alexander J. Smola, Amr Ahmed, Vanja Josifovski, James Long, Eugene J. Shekita, Bor-Yiing Su, OSDI 2014
- [Exploiting iterative-ness for parallel ML computations](), Henggang Cui, Alexey Tumanov, Jinliang Wei, Lianghong Xu, Wei Dai, Jesse Haber-Kucharsky, Qirong Ho, Gregory R. Ganger, Phillip B. Gibbons*, Garth A. Gibson, Eric P. Xing, SoCC 2014
- [Matchmaker: Data Drift Mitigation in Machine Learning for Large-Scale Systems](), Ankur Mallick, Kevin Hsieh, Behnaz Arzani, Gauri Joshi, MLSys 2022

***** Fast Scheduling [2/9]

- (Base) Orientation lecture by Indy
    - Michael Mitzenmacher: The Power of Two Choices in Randomized Load Balancing. IEEE Trans. Parallel Distributed Syst. 12(10): 1094-1104 (2001)
- **Paper 1**: Midhul Vuppalapati, Giannis Fikioris, and Rachit Agarwal, Cornell University; Asaf Cidon, Columbia University; Anurag Khandelwal, Yale University; Éva Tardos: Karma: Resource Allocation for Dynamic Demands, OSDI 2023.
- **Paper 2**: Ajaykrishna Karthikeyan, Nagarajan Natarajan, Gagan Somashekar, Lei Zhao, Ranjita Bhagwan, Rodrigo Fonseca, Tatiana Racheva, Yogesh Bansal: SelfTune: Tuning Cluster Managers. NSDI 2023: 1097-1114

Additional Readings

- Romil Bhardwaj, Kirthevasan Kandasamy, Asim Biswal, Wenshuo Guo, Benjamin Hindman, Joseph Gonzalez, Michael I. Jordan, Ion Stoica: Cilantro: Performance-Aware Resource Allocation for General Objectives via Online Feedback. OSDI 2023: 623-643
- Efficient Scheduling Policies for Microsecond-Scale Tasks, Sarah McClure and Amy Ousterhout, UC Berkeley; Scott Shenker, UC Berkeley, ICSI; Sylvia Ratnasamy, UC Berkeley, NSDI 2022
- MLaaS in the Wild: Workload Analysis and Scheduling in Large-Scale Heterogeneous GPU Clusters, Qizhen Weng, Hong Kong University of Science and Technology and Alibaba Group; Wencong Xiao, Alibaba Group; Yinghao Yu, Alibaba Group and Hong Kong University of Science and Technology; Wei Wang, Hong Kong University of Science and Technology; Cheng Wang, Jian He, Yong Li, Liping Zhang, Wei Lin, and Yu Ding, Alibaba Group, NSDI 2022
- Henri Maxime Demoulin, Joshua Fried, Isaac Pedisich, Marios Kogias, Boon Thau Loo, Linh Thi Xuan Phan, Irene Zhang: When Idling is Ideal: Optimizing Tail-Latency for Heavy-Tailed Datacenter Workloads with Perséphone. SOSP 2021: 621-637
- Kay Ousterhout, Patrick Wendell, Matei Zaharia, Ion Stoica: Sparrow: distributed, low latency scheduling. SOSP 2013: 69-84
- Microsecond Consensus for Microsecond Applications, Marcos K. Aguilera, Naama Ben-David, Rachid Guerraoui, Virendra J. Marathe, Athanasios Xygkis, Igor Zablotchi, OSDI 2020.
- HovercRaft: Achieving Scalability and Fault-tolerance for microsecond-scale Datacenter Services, Marios Kogias, Edouard Bugnion, Eurosys 2020.
- RackSched: A Microsecond-Scale Scheduler for Rack-Scale Computers, Hang Zhu, Kostis Kaffes, Zixu Chen, Zhenming Liu, Christos Kozyrakis, Ion Stoica, Xin Jin, OSDI 2020.
- Delay scheduling: a simple technique for achieving locality and fairness in cluster scheduling, Matei Zaharia, Dhruba Borthakur, Joydeep Sen Sarma, Khaled Elmeleegy, Scott Shenker, Ion Stoica, Eurosys 2010.

***** Batch Processing [2/14].

- (Base) Orientation lecture by Indy
  - Jeffrey Dean Sanjay Ghemawat. [MapReduce: Simplified Data Processing on Large Clusters](). Sixth Symposium on Operating System Design and Implementation (OSDI), San Francisco, CA (2004), pp. 137-150
  - Vinod Kumar Vavilapalli, Arun C. Murthy, Chris Douglas, Sharad Agarwal, Mahadev Konar, Robert Evans, Thomas Graves, Jason Lowe, Hitesh Shah, Siddharth Seth, Bikas Saha, Carlo Curino, Owen O'Malley, Sanjay Radia, Benjamin Reed, Eric Baldeschwieler: [Apache Hadoop YARN: yet another resource negotiator](). SoCC 2013: 5:1-5:16
  - Grzegorz Malewicz, Matthew H. Austern, Aart J. C. Bik, James C. Dehnert, Ilan Horn, Naty Leiser, Grzegorz Czajkowski: [Pregel: a system for large-scale graph processing](). SIGMOD Conference 2010: 135-146
- *Two related papers by graph processing researchers*
- **Paper 1**: Amitabha Roy, Laurent Bindschaedler, Jasmina Malicevic, Willy Zwaenepoel: [Chaos: scale-out graph processing from secondary storage](). SOSP 2015: 410-424
- **Paper 2**: Frank McSherry, Michael Isard, Derek G. Murray: [Scalability! But at what COST?](). HotOS 2015. (Short paper, but presenter has to read and present next two sub-bullet articles as well)
  - (Not for review, but should be presented by Paper 2 Presenter) Epic Graph Battle on Twitter! [See this link for snippets](.)
  - (Not for review, but should be presented by Paper 2 Presenter) [Busting Graph Processing Myths](), Blog post by Amitabha Roy, 2015.

  Additional Readings

- Jingji Chen, Xuehai Qian: [Khuzdul: Efficient and Scalable Distributed Graph Pattern Mining Engine](). ASPLOS (2) 2023: 413-426.
- Two (related, short) papers by DB researchers
  - 1. David J. DeWitt, Michael Stonebraker, [MapReduce: A major step backwards](), Database Column, 2008.
  - 2. Michael Stonebraker, Daniel J. Abadi, David J. DeWitt, Samuel Madden, Erik Paulson, Andrew Pavlo, Alexander Rasin: [MapReduce and parallel DBMSs: friends or foes?]() CACM 53(1): 64-71 (2010)
- Xizhe Yin, Zhijia Zhao, Rajiv Gupta: [Glign: Taming Misaligned Graph Traversals in Concurrent Graph Processing](). ASPLOS (1) 2023: 78-92
- Matei Zaharia, Mosharaf Chowdhury, Tathagata Das, Ankur Dave, Justin Ma, Murphy McCauley, Michael J. Franklin, Scott Shenker, Ion Stoica, [Resilient Distributed Datasets: A fault-tolerant abstraction for in-memory cluster computing](), NSDI 2012
- [Map-reduce-merge: simplified relational data processing on large clusters](), H.-C. Yang et al, SIGMOD 2007
- [Naiad: a timely dataflow system](), D. G. Murray, et al, SOSP 2013

- [DryadLINQ: A System for General-Purpose Distributed Data-Parallel Computing Using a High-Level Language](), Yuan Yu et al, OSDI 2008
- [Hive - a warehousing solution over a map-reduce framework]() A. Thusoo et al, VLDB 2009
- [Pig latin: a not-so-foreign language for data processing](), C. Olston et al, SIGMOD 2008 (Yahoo!)
- [GraphX: Graph Processing in a Distributed Dataflow Framework](), Joseph E. Gonzalez, Reynold S. Xin, Ankur Dave, Daniel Crankshaw, Michael J. Franklin, Ion Stoica, OSDI 2014.
- [Spark Streaming]()

***** Stream Processing and Streaming [2/16].

- (Base) Orientation lecture by Indy
  - [Storm@Twitter]()
  - History of Evolution of Stream Processing, across different communities
- **Paper 1**: Haruna Isah, Tariq Abughofa, Sazia Mahfuz, Dharmitha Ajerla, Farhana H. Zulkernine, Shahzad Khan: [A Survey of Distributed Data Stream Processing Frameworks](). IEEE Access 7: 154300-154316. 2019.
- **Paper 2**: Zhuqi Li, Yaxiong Xie, Ravi Netravali, Kyle Jamieson: [Dashlet: Taming Swipe Uncertainty for Robust Short Video Streaming](). NSDI 2023: 1583-1599.

  Additional Reading

- Ugur Çetintemel, Daniel J. Abadi, Yanif Ahmad, Hari Balakrishnan, Magdalena Balazinska, Mitch Cherniack, Jeong-Hyon Hwang, Samuel Madden, Anurag Maskey, Alexander Rasin, Esther Ryvkina, Mike Stonebraker, Nesime Tatbul, Ying Xing, Stan Zdonik: [The Aurora and Borealis Stream Processing Engines](). Data Stream Management 2016: see pages 337-359
- Jeff Barber, Ximing Yu, Laney Kuenzel Zamore, Jerry Lin, Vahid Jazayeri, Shie Erlich, Tony Savor, Michael Stumm: [Bladerunner: Stream Processing at Scale for a Live View of Backend Data Mutations at the Edge](). SOSP 2021: 708-723
- Le Xu, Shivaram Venkataraman, Indranil Gupta, Luo Mai, Rahul Potharaju. [Move Fast and Meet Deadlines: Fine-grained Real-time Stream Processing with Cameo](). Proc. USENIX Symposium on Networked Systems Design and Implementation (NSDI). 2021.
- See Additional Papers in the Stream Processing topic of the [SP18 version of the course (2/19)]().
- [Twitter Heron: stream processing at scale](), S. Kulkarni et al, SIGMOD 2015
- [Hadoop Streaming]()
- [MapReduce Online](), T. Condie et al, NSDI 2010
- [Spark Streaming]()
- [Live Video Analytics at Scale with Approximation and Delay-Tolerance](), Haoyu Zhang et al, NSDI 2017.

- [Encoding, Fast and Slow: Low-Latency Video Processing Using Thousands of Tiny Threads](), Sadjad Fouladi et al, NSDI 2017.
- [ReStream: Accelerating Backtesting and Stream Replay with Serial-Equivalent Parallel Processing](), J. Schleier-Smith et al, SoCC 2016
- [Consistent Regions: Guaranteed Tuple Processing in IBM Streams](), G. J. da Silva et al, VLDB 2016

***** Edge and Hybrid Computing [2/21].

- (Base) Orientation lecture by Indy
  - M. Satyanarayanan, [The Emergence of Edge Computing](), IEEE Computer, Volume 50, Number 1, January 2017
  - Shanhe Yi , Cheng Li , Qun Li, [A Survey of Fog Computing: Concepts, Applications and Issues](), Mobidata '15: Proceedings of the 2015 Workshop on Mobile Big DataJune 2015 Pages 37–42
  - Brief Primer on LEO Satellites: Bradley Denby, Brandon Lucia: [Orbital Edge Computing: Nanosatellite Constellations as a New Class of Computer System](). ASPLOS 2020: 939-954
- **Paper 1**: Yuanjie Li, Hewu Li, Wei Liu, Lixin Liu, Wei Zhao, Yimei Chen, Jianping Wu, Qian Wu, Jun Liu, Zeqi Lai, Han Qiu: [A Networking Perspective on Starlink's Self-Driving LEO Mega-Constellation](). MobiCom 2023: 17:1-17:16
- **Paper 2**: Bashima Islam, Yubo Luo, Shahriar Nirjon: [Amalgamated Intermittent Computing Systems](). IoTDI 2023: 184-196.

  Additional Readings

- Bill Tao, Maleeha Masood, Indranil Gupta, Deepak Vasisht: [Transmitting, Fast and Slow: Scheduling Satellite Traffic through Space and Time](). MobiCom 2023: 24:1-24.
- [SwarmMap: Scaling Up Real-time Collaborative Visual SLAM at the Edge](), Jingao Xu, Hao Cao, and Zheng Yang, Longfei Shangguan, Jialin Zhang, Xiaowu He, and Yunhao Liu, NSDI 2022.
- [Walle: An End-to-End, General-Purpose, and Large-Scale Production System for Device-Cloud Collaborative Machine Learning](), Chengfei Lv, Chaoyue Niu, Renjie Gu, Xiaotang Jiang, Zhaode Wang, Bin Liu, Ziqi Wu, Qiulin Yao, Congyu Huang, Panos Huang, Tao Huang, Hui Shu, Jinde Song, Bin Zou, Peng Lan, and Guohuan Xu, Shaojie Tang, Fan Wu and Guihai Chen, OSDI 2022
- [Towards Federated Learning at Scale: System Design](), Keith Bonawitz, Hubert Eichner, Wolfgang Grieskamp, Dzmitry Huba, Alex Ingerman, Vladimir Ivanov, Chloé Kiddon, Jakub Konečný, Stefano Mazzocchi, Brendan McMahan, Timon Van Overveldt, David Petrou, Daniel Ramage, Jason Roselander, MLSys 2019.
- Dongzhu Xu , Anfu Zhou , Guixian Wang , Huanhuan Zhang , Xiangyu Li , Jialiang Pei , Huadong Ma, [Tutti: coupling 5G RAN and mobile edge computing for latency-critical video analytics](), Mobicom 2022.
- [Starlight: Fast Container Provisioning on the Edge and over the WAN](), Jun Lin Chen, Daniyal Liaqat, Moshe Gabel, and Eyal de Lara, University of Toronto, NSDI 2022
- [BumbleBee: Application-aware adaptation for edge-cloud orchestration](), HyunJong Lee , Shadi Noghabi, Brian Noble, Matthew Furlong , Landon Cox Symposium on Edge Computing | December 2022
- [Sl-EDGE: Network Slicing at the Edge](), Salvatore D'Oro, Leonardo Bonati Francesco Restuccia, Michele Polese, Michele Zorzi, Tommaso Melodia, Mobihoc 2020.

- [Fair Multi-resource Allocation in Mobile Edge Computing with Multiple Access Points](#), Erfan Meskar, Ben Liang, Mobihoc 2020. [[Author paper link](#)]
- [MCUNet: Tiny Deep Learning on IoT Devices](#), Ji Lin, Wei-Ming Chen, Yujun Lin, John Cohn, Chuang Gan, Song Han, NeurIPS 2020. [[Conference paper link](#)]
- [CLIO: Enabling automatic compilation of deep learning pipelines across IoT and Cloud](#), Jin Huang, Colin Samplawski, Deepak Ganesan, Benjamin Marlin, Heesung Kwon, Mobicom 2020. [[Author paper link](#)]
- [ECCO: Edge-Cloud Chaining and Orchestration Framework for Road Context Assessment](#), Vittorio Cozzolino, Jörg Ott, Aaron Yi Ding Richard Mortier, IoTDI 2020.
- Kostis Kaffes, Jack Tigar Humphries, David Mazières, Christos Kozyrakis: [Syrup: User-Defined Scheduling Across the Stack](#). SOSP 2021: 605-620

***** Visions [2/23]

- (Base) Orientation lecture by Indy
  - Michael Armbrust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy H. Katz, Andy Konwinski, Gunho Lee, David A. Patterson, Ariel Rabkin, Ion Stoica, Matei Zaharia: A view of cloud computing. Commun. ACM 53(4): 50-58 (2010) [CACM Paper] [Longer Arxiv]
- **Paper 1**: SkyPilot: An Intercloud Broker for Sky Computing, Zongheng Yang, Zhanghao Wu, Michael Luo, Wei-Lin Chiang, Romil Bhardwaj, Woosuk Kwon, Siyuan Zhuang, Frank Sifei Luan, Gautam Mittal, Scott Shenker, Ion Stoica, NSDI 2023.
- **Paper 2**: Aegean: Replication Beyond the Client-Server Model, Remzi Can Aksoy, Manos Kapritsos, SOSP 2019.

  Additional Readings

- Sarah Chasins, Alvin Cheung, Natacha Crooks, Ali Ghodsi, Ken Goldberg, Joseph E. Gonzalez, Joseph M. Hellerstein, Michael I. Jordan, Anthony D. Joseph, Michael W. Mahoney, Aditya G. Parameswaran, David A. Patterson, Raluca Ada Popa, Koushik Sen, Scott Shenker, Dawn Song, Ion Stoica: The Sky Above The Clouds. Arxiv, 2022.
- Matthew Burke, Sowmya Dharanipragada, Shannon Joyner, Adriana Szekeres, Jacob Nelson, Irene Zhang, Dan R. K. Ports: PRISM: Rethinking the RDMA Interface for Distributed Systems. SOSP 2021: 228-242
- Beyond Data and Model Parallelism for Deep Neural Networks, Zhihao Jia, Matei Zaharia, Alex Aiken, MLSys 2019.
- Toward a Generic Fault Tolerance Technique for Partial Network Partitioning, Mohammed Alfatafta, Basil Alkhatib, Ahmed Alquraan, Samer Al-Kiswany, OSDI 2020.
- Jayashree Mohan, Amar Phanishayee, and Janardhan Kulkarni, Microsoft Research; Vijay Chidambaram, University of Texas at Austin and VMware Research, Looking Beyond GPUs for DNN Scheduling on Multi-Tenant Clusters, OSDI 2022.
- Monotasks: Architecting for Performance Clarity in Data Analytics Frameworks, Kay Ousterhout et al, SOSP 2017.

***** Caching [2/28].

- (Base) Orientation lecture by Indy
  - Classical Cache Eviction (What's optimal? What're good approximations?)
- **Paper 1**: Alec Wolman, Geoffrey M. Voelker, Nitin Sharma, Neal Cardwell, Anna R. Karlin, Henry M. Levy: On the scale and performance of cooperative Web proxy caching. SOSP 1999: 16-31
- **Paper 2**: Zhenyu Song, Kevin Chen, Nuikhil Sarda, Deniz Altinbüken, Eugene Brevdo, Jimmy Coleman, Xiao Ju, Pawel Jurczyk, Richard Schooler, Ramki Gummadi: HALP: Heuristic Aided Learned Preference Eviction Policy for YouTube Content Delivery Network. NSDI 2023: 1149-1163.

  Additional Readings

- Juncheng Yang, Yazhuo Zhang, Ziyue Qiu, Yao Yue, Rashmi Vinayak: FIFO queues are all you need for cache eviction. SOSP 2023: 130-149

***** Weak Consistency [3/1].

- (Base) Orientation lecture by Indy
  - Weak Consistency Models
- **Paper 1**: João Ferreira Loff, Daniel Porto, João Garcia, Jonathan Mace, Rodrigo Rodrigues: Antipode: Enforcing Cross-Service Causal Consistency in Distributed Applications. SOSP 2023: 298-313.
- **Paper 2**: Haonan Lu, Shuai Mu, Siddhartha Sen, Wyatt Lloyd: NCC: Natural Concurrency Control for Strictly Serializable Datastores by Avoiding the Timestamp-Inversion Pitfall. OSDI 2023: 305-323.

Additional Readings

- Epidemic algorithms for replicated database maintenance, Alan Demers, Dan Greene, Carl Hauser, Wes Irish, John Larson. Scott Shenker, Howard Sturgis, Dan Swinehart, and Doug Terry, ACM PODC 1987 https://dl.acm.org/doi/pdf/10.1145/41840.41841
- Sally Floyd, Van Jacobson, Ching-Gung Liu, Steven McCanne, Lixia Zhang: A reliable multicast framework for light-weight sessions and application level framing. IEEE/ACM Trans. Netw. 5(6): 784-803 (1997) https://ieeexplore.ieee.org/abstract/document/650139
- Bimodal multicast, K Birman et al, ACM TOCS 1999
- Marc Shapiro, Nuno M. Preguiça, Carlos Baquero, Marek Zawirski: Conflict-Free Replicated Data Types. SSS 2011: Stabilization, Safety, and Security of Distributed Systems pp 386–400, 2011.
- Inho Choi, Ellis Michael, Yunfan Li, Dan R. K. Ports, Jialin Li: Hydra: Serialization-Free Network Ordering for Strongly Consistent Distributed Applications. NSDI 2023: 293-320
- Shadaj Laddad, Conor Power, Mae Milano, Alvin Cheung, Natacha Crooks, Joseph M. Hellerstein: Keep CALM and CRDT On, VLDB 2023.
- Z. Haas et al, Gossip-based ad hoc routing, IEEE Transactions on Networking 2006. (also in Infocom 2002).
- Cigdem Sengul, Indranil Gupta, Matthew J. Miller: Adaptive probability-based broadcast forwarding in energy-saving sensor networks. ACM Trans. Sens. Networks 4(2): 6:1-6:32 (2008) (Also published in ICDCS 2005).
  - *(A CS525 Project that became a conference paper and a journal paper!)*
- Marc Shapiro, Nuno M. Preguiça, Carlos Baquero, Marek Zawirski: Conflict-Free Replicated Data Types. SSS 2011: Stabilization, Safety, and Security of Distributed Systems pp 386–400, 2011.
- Randomized Rumor Spreading, Karp and Shenker, FOCS 2000
- Spatial gossip and resource location protocols, Kempe, Kleinberg and Demers, STOC 2001
- Immunology as information processing, S. Forrest et al, 2000.
- Efficient and Adaptive Epidemic-style Protocols for Reliable and Scalable Multicast, Gupta et al, IEEE TPDS, 2006.
- Gossip-Based Clock Synchronization for Large Decentralized Systems, K. Iwanicki et al, SelfMan 2006: 28-42

- [A gossip-based failure detection service](#), R. van Renesse et al, Middleware 1998
- [Peer-to-peer membership management for gossip-based protocols](#), A.J. Ganesh et al, IEEE TOC, Feb 2003.
- [T-Man: Fast Gossip-based Construction of Large-Scale Overlay Topologies](#), M. Jelasity et al, U. Bologna Tech Report.

***** Microservices [3/6].

- (Base) Orientation lecture by Indy
- **Paper 1**: Vaastav Anand, Deepak Garg, Antoine Kaufmann, Jonathan Mace: Blueprint: A Toolchain for Highly-Reconfigurable Microservice Applications. SOSP 2023: 482-497.
- **Paper 2**: Harshit Saokar, Soteris Demetriou, Nick Magerko, Max Kontorovich, Josh Kirstein, Margot Leibold, Dimitrios Skarlatos, Hitesh Khandelwal, Chunqiang Tang: ServiceRouter: Hyperscale and Minimal Cost Service Mesh at Meta. OSDI 2023: 969-985.

  Additional Readings

- Mingyu Liang, Yu Gan, Yueying Li, Carlos Torres, Abhishek Dhanotia, Mahesh Ketkar, Christina Delimitrou: Ditto: End-to-End Application Cloning for Networked Cloud Services. ASPLOS (2) 2023: 222-236
- Shutian Luo, Huanle Xu, Kejiang Ye, Guoyao Xu, Liping Zhang, Jian He, Guodong Yang, Chengzhong Xu: Erms: Efficient Resource Management for Shared Microservices with SLA Guarantees. 62-77, ASPLOS 2023.
- Yu Gan, Mingyu Liang, Sundar Dev, David Lo, Christina Delimitrou: Sage: practical and scalable ML-driven performance debugging in microservices. ASPLOS 2021: 135-151
- Yanqi Zhang, Weizhe Hua, Zhuangzhuang Zhou, G. Edward Suh, Christina Delimitrou: Sinan: ML-based and QoS-aware resource management for cloud microservices. ASPLOS 2021: 167-181

***** Serverless [3/8].

- (Base) Orientation lecture by Indy
  - [AWS Lambda](), [Azure Functions]()
- **Paper 1**: Zhuangzhuang Zhou, Yanqi Zhang, Christina Delimitrou: [AQUATOPE: QoS-and-Uncertainty-Aware Resource Management for Multi-stage Serverless Workflows](). ASPLOS (1) 2023: 1-14
- **Paper 2**: David H. Liu, Amit Levy, Shadi A. Noghabi, Sebastian Burckhardt: [Doing More with Less: Orchestrating Serverless Applications without an Orchestrator](). NSDI 2023: 1505-1519

  Additional Readings

- Sheng Qi, Xuanzhe Liu, Xin Jin: [Halfmoon: Log-Optimal Fault-Tolerant Stateful Serverless Computing](). SOSP 2023: 314-330
- Deepak Bansal, Gerald DeGrace, Rishabh Tewari, Michal Zygmunt, James Grantham, Silvano Gai, Mario Baldi, Krishna Doddapaneni, Arun Selvarajan, Arunkumar Arumugam, Balakrishnan Raman, Avijit Gupta, Sachin Jain, Deven Jagasia, Evan Langlais, Pranjal Srivastava, Rishiraj Hazarika, Neeraj Motwani, Soumya Tiwari, Stewart Grant, Ranveer Chandra, Srikanth Kandula: [Disaggregating Stateful Network Functions](). NSDI 2023: 1469-1487
- Ziming Zhao, Mingyu Wu, Jiawei Tang, Binyu Zang, Zhaoguo Wang, Haibo Chen: BeeHive: [Sub-second Elasticity for Web Services with Semi-FaaS Execution](). 74-87, ASPLOS 2023
- Alireza Sahraei, Soteris Demetriou, Amirali Sobhgol, Haoran Zhang, Abhigna Nagaraja, Neeraj Pathak, Girish Joshi, Carla Souza, Bo Huang, Wyatt Cook, Andrii Golovei, Pradeep Venkat, Andrew Mcfague, Dimitrios Skarlatos, Vipul Patel, Ravinder Thind, Ernesto Gonzalez, Yun Jin, Chunqiang Tang: [XFaaS: Hyperscale and Low Cost Serverless Functions at Meta](). SOSP 2023: 231-246
- Minchen Yu, Tingjia Cao, Wei Wang, Ruichuan Chen: Following the Data, Not the [Function: Rethinking Function Orchestration in Serverless Computing](). NSDI 2023: 1489-1504
- Zhipeng Jia, Emmett Witchel: [Boki: Stateful Serverless Computing with Shared Logs](). SOSP 2021: 691-707
- Kostis Kaffes (Stanford University); Neeraja J. Yadwadkar (UT Austin); Christos Kozyrakis (Stanford University), [Hermod: Principled and Practical Scheduling for Serverless Functions](), SoCC 2022.
- Vivek M. Bhasi (The Pennsylvania State University, University Park); Jashwant Raj Gunasekaran (Adobe Research); Aakash Sharma, Mahmut Taylan Kandemir, Chita Das (The Pennsylvania State University, University Park), [Cypress : Input size–Sensitive Container Provisioning and Request Scheduling for Serverless Platforms](), SoCC 2022
- [A Fault-Tolerance Shim for Serverless Computing](), Vikram Sreekanti, Chenggang Wu, Saurav Chhatrapati, Joseph Gonzalez, Joseph M. Hellerstein (UC Berkeley), Jose M. Faleiro, Eurosys 2020.

- [Fault-tolerant and transactional stateful serverless workflows](), Haoran Zhang, Adney Cardoza, Peter Baile Chen, Sebastian Angel, Vincent Liu, OSDI 2020.
- [FIRM: An Intelligent Fine-grained Resource Management Framework for SLO-Oriented Microservices](), Haoran Qiu, Subho S. Banerjee, Saurabh Jha, Zbigniew T. Kalbarczyk, Ravishankar K. Iyer, OSDI 2020.
- [SEUSS: Skip Redundant Paths to Make Serverless Fast](), James Cadden, Thomas Unger, Yara Awad, Han Dong, Orran Krieger, Jonathan Appavoo, Eurosys 2020.
- [Move Fast and Meet Deadlines: Fine-grained Real-time Stream Processing with Cameo](), Le Xu, Shivaram Venkataraman, Indranil Gupta, Luo Mai, Rahul Potharaju, NSDI 2021.

3/13 and 3/15 - No Class. Spring Break.

***** Elasticity [3/20]

- (Base) Orientation lecture by Indy
- **Paper 1**: Romil Bhardwaj, Kirthevasan Kandasamy, Asim Biswal, Wenshuo Guo, Benjamin Hindman, Joseph Gonzalez, Michael I. Jordan, Ion Stoica: Cilantro: Performance-Aware Resource Allocation for General Objectives via Online Feedback. OSDI 2023: 623-643
- **Paper 2**: Zibo Wang, Pinghe Li, Chieh-Jan Mike Liang, Feng Wu, Francis Y. Yan: Autothrottle: A Practical Bi-Level Approach to Resource Management for SLO-Targeted Microservices, NSDI 2024 (accepted).

    Additional Readings

- K8s' VPA and HPA (Vertical and Horizontal Pod Autoscalers, in Kubernetes)
- FIRM: An Intelligent Fine-grained Resource Management Framework for SLO-Oriented Microservices, Haoran Qiu, Subho S. Banerjee, Saurabh Jha, Zbigniew T. Kalbarczyk, Ravishankar K. Iyer, OSDI 2020.
- Autopilot: workload autoscaling at Google, Krzysztof Rzadca, Pawel Findeisen, Jacek Swiderski, Przemyslaw Zych, Przemyslaw Broniek, Jarek Kusmierek, Pawel Nowak, Beata Strack, Piotr Witusowski, Steven Hand, John Wilkes, Eurosys 2020. [direct link]
- Diandian Gu, Yihao Zhao, Yinmin Zhong, Yifan Xiong, Zhenhua Han, Peng Cheng, Fan Yang, Gang Huang, Xin Jin, Xuanzhe Liu: ElasticFlow: An Elastic Serverless Training Platform for Distributed Deep Learning. ASPLOS (2) 2023: 266-280
- Yanqi Zhang, Weizhe Hua, Zhuangzhuang Zhou, G. Edward Suh, Christina Delimitrou: Sinan: ML-based and QoS-aware resource management for cloud microservices. ASPLOS 2021: 167-181
- Nikos Armenatzoglou, Sanuj Basu, Naga Bhanoori, Mengchu Cai, Naresh Chainani, Kiran Chinta, Venkatraman Govindaraju, Todd J. Green, Monish Gupta, Sebastian Hillig, Eric Hotinger, Yan Leshinksy, Jintian Liang, Michael McCreedy, Fabian Nagel, Ippokratis Pandis, Panos Parchas, Rahul Pathak, Orestis Polychroniou, Foyzur Rahman, Gaurav Saxena, Gokul Soundararajan, Sriram Subramanian, Doug Terry: Amazon Redshift Re-invented. SIGMOD 2022: 2205-2217
- Vincenzo Gulisano, Ricardo Jiménez-Peris, Marta Patiño-Martínez, Claudio Soriente, Patrick Valduriez: StreamCloud: An Elastic and Scalable Data Streaming System. IEEE Trans. Parallel Distributed Syst. 23(12): 2351-2365 (2012)
- Resource Elasticity in Distributed Deep Learning In Efficient Model Training, Andrew Or, Haoyu Zhang, Michael Freedman, MLSys 2020.
- PLASMA: Programmable Elasticity for Stateful Cloud Computing Applications, Bo Sang, Pierre-Louis Roman, Patrick Eugster, Hui Lu, Srivatsan Ravi, Gustavo Petri, Eurosys 2020.

- Faria Kalim, Le Xu, Sharanya Bathey, Richa Meherwal, Indranil Gupta: Henge: Intent-driven Multi-Tenant Stream Processing. SoCC 2018.
- Le Xu, Shivaram Venkataraman, Indranil Gupta, Luo Mai, Rahul Potharaju. Move Fast and Meet Deadlines: Fine-grained Real-time Stream Processing with Cameo. Proc. USENIX Symposium on Networked Systems Design and Implementation (NSDI) 2021.
- Brian Cho, Muntasir Raihan Rahman, Tej Chajed, Indranil Gupta, Cristina L. Abad, Nathan Roberts, Philbert Lin: Natjam: design and evaluation of eviction policies for supporting priorities and deadlines in mapreduce clusters. SoCC 2013: 6:1-6:17
- SRIFTY: Swift and Thrifty Distributed Neural Network Training on the Cloud, Liang Luo, Peter West, Pratyush Patel, Arvind Krishnamurthy, Luis Ceze, MLSys 2022
- REX: Revisiting Budgeted Training with an Improved Schedule, John Chen, Cameron Wolfe, Tasos Kyrillidis, MLSys 2022

***** Byzantine [3/22]

- (Base) Orientation lecture by Indy
  - L. Lamport, R. Shostak, M. Pease: [The Byzantine Generals Problem](#), Concurrency: the Works of Leslie Lamport, October 2019, Pages 203–226. [[Alternate Link](#)]
  - Zibin Zheng, Shaoan Xie, Hongning Dai, Xiangping Chen, Huaimin Wang: [An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends](#). BigData Congress 2017: 557-564
  - L. Lamport, [The Weak Byzantine Generals Problem](#), JACM, Vol 30, No 3, July 1983, pp 668-676.
- **Paper 1**: Shengyun Liu, Wenbo Xu, Chen Shan, Xiaofeng Yan, Tianjing Xu, Bo Wang, Lei Fan, Fuxi Deng, Ying Yan, Hui Zhang: [Flexible Advancement in Asynchronous BFT Consensus](#). SOSP 2023: 264-280
- **Paper 2**: Marcos K. Aguilera, Naama Ben-David, Rachid Guerraoui, Antoine Murat, Athanasios Xygkis, Igor Zablotchi: [uBFT: Microsecond-Scale BFT using Disaggregated Memory](#). ASPLOS (2) 2023: 862-877

  Additional Readings

- M. Castro, B, Liskov, [Practical Byzantine Fault-Tolerance](#), OSDI 1999.
- Allen Clement, Edmund L. Wong, Lorenzo Alvisi, Michael Dahlin, Mirco Marchetti: [Making Byzantine Fault Tolerant Systems Tolerate Byzantine Faults](#). NSDI 2009: 153-168
- Ittay Eyal, Emin Gün Sirer: [Majority Is Not Enough: Bitcoin Mining Is Vulnerable](#). 436-454, 18th Financial Cryptography 2014.
- Ji Qi, Xusheng Chen, Yunpeng Jiang, Jianyu Jiang, Tianxiang Shen, Shixiong Zhao, Sen Wang, Gong Zhang, Li Chen, Man Ho Au, Heming Cui: [Bidl: A High-throughput, Low-latency Permissioned Blockchain Framework for Datacenter Networks](#). SOSP 2021: 18-34 [[Alt Paper link](#)]

***** Consensus [3/27]

- (Base) Orientation lecture by Indy
  - [Impossibility of distributed consensus with one faulty process](#), Fischer, Lynch and Patterson, Journal ACM 1985
  - Leslie Lamport: [The Part-Time Parliament](#). ACM Trans. Comput. Syst. 16(2): 133-169 (1998)
  - [Paxos Made Simple](#), L. Lamport.
- *For this day only - Only 1 paper review per session (out of 2 papers)*
- **Paper 1**: Yanhua Mao, Flavio Paiva Junqueira, Keith Marzullo: [Mencius: Building Efficient Replicated State Machine for WANs](#). OSDI 2008: 369-384.
- **Paper 2**: Ittay Eyal, Adem Efe Gencer, Emin Gün Sirer, Robbert van Renesse: [Bitcoin-NG: A Scalable Blockchain Protocol](#). NSDI 2016.

  Additional Readings

- Mahesh Balakrishnan, Chen Shen, Ahmed Jafri, Suyog Mapara, David Geraghty, Jason Flinn, Vidhya Venkat, Ivailo Nedelchev, Santosh Ghosh, Mihir Dharamshi, Jingming Liu, Filip Gruszczynski, Jun Li, Rounak Tibrewal, Ali Zaveri, Rajeev Nagar, Ahmed Yossef, Francois Richard, Yee Jiun Song: [Log-structured Protocols in Delos](#). SOSP 2021: 538-552
- Jiaping Wang, Hao Wang: [Monoxide: Scale out Blockchains with Asynchronous Consensus Zones](#). NSDI 2019: 95-112
- Mohammadreza Alimadadi, Hieu Mai, Shenghsun Cho, Michael Ferdman, Peter A. Milder, Shuai Mu: [Waverunner: An Elegant Approach to Hardware Acceleration of State Machine Replication](#). NSDI 2023: 357-374
- Haochen Pan, Jesse Tuglu, Neo Zhou, Tianshu Wang, Yicheng Shen, Xiong Zheng, Joseph Tassarotti, Lewis Tseng, Roberto Palmieri: [Rabia: Simplifying State-Machine Replication Through Randomization.](#) SOSP 2021: 472-487
- Diego Ongaro and John Ousterhout, [In Search of an Understandable Consensus Algorithm](#), Usenix ATC 2014
- [Virtual Consensus in Delos](#), Mahesh Balakrishnan, Jason Flinn, Chen Shen, Mihir Dharamshi, Ahmed Jafri, Xiao Shi, Santosh Ghosh, Hazem Hassan, Aaryaman Sagar, Rhed Shi, Jingming Liu, Filip Gruszczynski, Xianan Zhang, Huy Hoang, Ahmed Yossef, Francois Richard, Yee Jiun Song, OSDI 2020.
- [vCorfu: A Cloud-Scale Object Store on a Shared Log](#), Michael Wei et al, NSDI 2017.
- [EPaxos Revisited](#), Sarah Tollman, Seo Jin Park, John Ousterhout, NSDI 2021.
- [Fault-Tolerant Replication with Pull-Based Consensus in MongoDB](#), Siyuan Zhou, Shuai Mu, NSDI 2021.
- [Scalog: Seamless Reconfiguration and Total Order in a Scalable Shared Log](#), Cong Ding David Chu Evan Zhao Xiang Li† Lorenzo Alvisi Robbert van Renesse, NSDI 2020.
- [CRaft: An Erasure-coding-supported Version of Raft for Reducing Storage Cost and Network Cost](#), Zizhong Wang, Tongliang Li, Haixia Wang, Airan Shao, Yunren Bai, Shangming Cai, Zihan Xu, Dongsheng Wang, FAST 2020.

- [Paxos Made Transparent](), Heming Cui, et al, SOSP 2015
- [Tango: distributed data structures over a shared log](), M. Balakrishnan, et al, SOSP 2013
- [Paxos Made Moderately Complex](), R. van Renesse, D. Altinbuken, CSUR Surveys, 2015
- [Paxos made live - an engineering perspective](), T. Chandra et al, PODC 2007
- [ABCDs of Paxos](), Butler Lampson, PODC 2001.
- [Paxos Made Simple](), Leslie Lamport, 2001.
- [The Part-time Parliament](), Leslie Lamport, ACM TOCS, 1998.

***** P2P Apps [3/29]

- (Base) Orientation lecture by Indy
- *For this day only - Only 1 paper review per session (out of 2 papers)*
- **Paper 1**: Beehive : O(1) Lookup Performance for Power-Law Query Distributions in Peer-to-Peer Overlays, Venugopalan Ramasubramanian and Emin Gun Sirer, NSDI 2004.
- **Paper 2**: Xiaofei Liao, Hai Jin, Yunhao Liu, Lionel M. Ni, Dafu Deng: Anysee: p2p live streaming, Infocom 2006.

  Additional Readings
- Petar Maymounkov, David Mazières: Kademlia: A Peer-to-Peer Information System Based on the XOR Metric. IPTPS 2002: 53-65
- Corona: a high performance publish-subscribe system for the World Wide Web, V. Ramasubramanian, R. Peterson, E. G. Sirer, NSDI 2006
- Lu Fan, Philip W. Trinder, Hamish Taylor: Design issues for Peer-to-Peer Massively Multiplayer Online Games. Int. J. Adv. Media Commun. 4(2): 108-125 (2010)
- Reliable client accounting for p2p-infrastructure hybrids, P. Aditya, et al, NSDI 2012
- UsenetDHT: A Low-Overhead Design for Usenet
  Emil Sit et al, NSDI 2008
- Colyseus: A distributed architecture for interactive multiplayer games, A.R. Bharambe, Usenix NSDI 2006.
- Peer-to-peer support for massively multiplayer games, B. Knutsson et al, Infocom 2004.
- Operating system support for planetary-scale network services, A. Bavier et al, NSDI 2004.
- CoDNS: Improving DNS Performance and Reliability via Cooperative Lookups, KyoungSoo Park et al, OSDI 2004
- Wide-area cooperative storage with CFS, F. Dabek et al, SOSP 2001
- Scalability of reliable group communication using overlays, F. Baccelli et al, Infocom 2004.
- OceanStore: An Architecture for Global-Scale Persistent Storage , J. Kubiatowicz, ASPLOS 2000
- Splitstream, M. Castro et al, SOSP 2003.
- Johan A. Pouwelse, Pawel Garbacki, Dick H. J. Epema, Henk J. Sips: The Bittorrent P2P File-Sharing System: Measurements and Analysis. IPTPS 2005: 205-216

<u><<<</u> Midterm Report Due 3/31 (must contain design and preliminary results!) <u>>>></u>

<u>*****</u> Key-value/NoSQL Storage Systems [4/3]

- (Base) A history of the evolution of key-value/NoSQL stores
  - [Dynamo: Amazon's highly-available key-value store](#), DeCandia et al, SOSP 2007
- *For this day only - Only 1 paper review per session (out of 2 papers)*
- **Paper 1**: Zhuohui Duan, Jiabo Yao, Haikun Liu, Xiaofei Liao, Hai Jin, Yu Zhang: [Revisiting Log-Structured Merging for KV Stores in Hybrid Memory Systems](#). ASPLOS (2) 2023: 674-687
- **Paper 2**: [SILT: a memory-efficient, high performance key-value store](#), H. Lim, B. Fan. D. Andersen, M. Kaminsky, SOSP 2011

  Additional Readings

- Se Kwon Lee, Soujanya Ponnapalli, Sharad Singhal, Marcos K. Aguilera, Kimberly Keeton, Vijay Chidambaram: [DINOMO: An Elastic, Scalable, High-Performance Key-Value Store for Disaggregated Persistent Memory](#) VLDB 2021-2022.
- Subarna Chatterjee, Meena Jagadeesan, Wilson Qin, Stratos Idreos: [Cosine: A Cloud-Cost Optimized Self-Designing Key-Value Storage Engine](#). Proc. VLDB Endow. 15(1): 112-126, 2021.
- [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 Robert E. Gruber, OSDI 2006
- [Project Voldemort, Linkedin](#)
- [Comet: An Active Distributed Key-Value Store](#), R. Geambasu et al, OSDI 2010
- Baptiste Lepers, Oana Balmau, Karan Gupta, Willy Zwaenepoel: [KVell: the design and implementation of a fast persistent key-value store](#). SOSP 2019: 447-461
- Robert Escriva, Bernard Wong, Emin Gün Sirer: [HyperDex: a distributed, searchable key-value store](#). SIGCOMM 2012: 25-36
- [Concerto: A High Concurrency Key-Value Store with Integrity](#), Arvind Arasu et al, SIGMOD 2017.
- [SILT: a memory-efficient, high performance key-value store](#), H. Lim, B. Fan. D. Andersen, M. Kaminsky, SOSP 2011
- [Rein: Taming Tail Latency in Key-Value Stores via Multiget Scheduling](#), Waleed Reda et al, SoCC 2017.
- [C-Hint: An Effective and Reliable Cache Management for RDMA-Accelerated Key-Value Stores](#), Yandong Wang, Xiaoqiao Meng, Li Zhang, Jian Tan, SoCC 2014
- [BlueCache: A Scalable Distributed Flash-based Key-value Store](#), Shuotao Xu et al, VLDB 2017

***** Preemptibles [4/5].

- (Base) Orientation lecture by Indy
- ***For this day only - Only 1 paper review per session (out of 2 papers)***
- **Paper 1**: Fangkai Yang, Lu Wang, Zhenyu Xu, Jue Zhang, Liqun Li, Bo Qiao, Camille Couturier, Chetan Bansal, Soumya Ram, Si Qin, Zhen Ma, Íñigo Goiri, Eli Cortez, Terry Yang, Victor Rühle, Saravan Rajmohan, Qingwei Lin, Dongmei Zhang: Snape: Reliable and Low-Cost Computing with Mixture of Spot and On-Demand VMs. ASPLOS (3) 2023: 631-643.
- **Paper 2**: John Thorpe, Pengzhan Zhao, Jonathan Eyolfson, Yifan Qiao, Zhihao Jia, Minjia Zhang, Ravi Netravali, Guoqing Harry Xu: Bamboo: Making Preemptible Instances Resilient for Affordable Training of Large DNNs, NSDI 2023.

  Additional Readings

- Spotcheck: Designing a derivative iaas cloud on the spot market P Sharma, S Lee, T Guo, D Irwin, P Shenoy Proceedings of the Tenth European Conference on Computer Systems, 1-15
- Spoton: a batch computing service for the spot market S Subramanya, T Guo, P Sharma, D Irwin, P Shenoy Proceedings of the sixth ACM symposium on cloud computing, 329-341

***** Databases and Transactions [4/10]

- (Base) Orientation lecture by Indy
  - Samuel Madden, Michael J. Franklin, Joseph M. Hellerstein, Wei Hong: [TinyDB: an acquisitional query processing system for sensor networks](). ACM Trans. Database Syst. 30(1): 122-173 (2005)
- **Paper 1**: Audrey Cheng, David Chu, Terrance Li, Jason Chan, Natacha Crooks, Joseph M. Hellerstein, and Ion Stoica, Xiangyao Yu: [Take Out the TraChe: Maximizing (Tra)nsactional Ca(che) Hit Rate](). OSDI 2023.
- **Paper 2**: Kelly Kostopoulou, Pierre Tholoniat, Asaf Cidon, Roxana Geambasu, Mathias Lécuyer: [Turbo: Effective Caching in Differentially-Private Databases](). SOSP 2023: 579-594

  Additional Readings

- [Millions of Tiny Databases](), Marc Brooker, Tao Chen Fan Ping, NSDI 2020.
- [Transactuations: where transactions meet the physical world](), A. Sengupta, T. Leesatapornwongsa, M. S. Ardekani, C. A. Stuardo, Usenix ATC 2019.
- Hafiz Mohsin Bashir, Abdullah Bin Faisal, Fahad Dogar, [Network Resource Management as a Database Problem](), SoCC 2022
- Shegufta Bakht Ahsan, Rui Yang, Shadi A. Noghabi, Indranil Gupta: [Home, SafeHome: smart home reliability with visibility and atomicity](). EuroSys 2021: 590-605
- Shegufta Bakht Ahsan, Rui Yang, Shadi A. Noghabi, Indranil Gupta. [Home, SafeHome: Ensuring a Safe and Reliable Home Using the Edge](), Proc. 2nd Usenix Workshop on Hot Topics in Edge Computing (HotEdge), 2019.
- [Millions of Tiny Databases](), Marc Brooker, Tao Chen Fan Ping, NSDI 2020.
- [The SNOW Theorem and Latency-Optimal Read-Only Transactions](), H. Lu et al, OSDI 2016.
- Yongle Zhang, Junwen Yang, Zhuqi Jin, Utsav Sethi, Kirk Rodrigues, Shan Lu, Ding Yuan: [Understanding and Detecting Software Upgrade Failures in Distributed Systems](). SOSP 2021: 116-131
- [Yesquel: Scalable SQL storage for Web applications](), Marcos K. Aguilera, et al, SOSP 2015
- [Demystifying and Checking Silent Semantic Violations in Large Distributed Systems](), Chang Lou, Yuzhuo Jing, and Peng Huang, Johns Hopkins University, OSDI 2022
- Dai Qin, Angela Demke Brown, Ashvin Goel: [Caracal: Contention Management with Deterministic Concurrency Control](). SOSP 2021: 180-194
- [Cobra: Making Transactional Key-Value Stores Verifiably Serializable](), Cheng Tan, Changgeng Zhao, Shuai Mu, Michael Walfish, OSDI 2020.

- [Kvell+: Snapshot Isolation without Snapshots](), Baptiste Lepers, Oana Balmau, Karan Gupta, Willy Zwaenepoel, OSDI 2020.
- [Meerkat: Multicore-Scalable Replicated Transactions Following the Zero-Coordination Principle](), Adriana Szekeres, Michael Whittaker, Naveen Kr. Sharma, Jialin Li, Arvind Krishnamurthy, Dan R. K. Ports, Irene Zhang, Eurosys 2020.
- [Transactional storage for geo-replicated systems](), Y. Sovran, R. Power, M. K. Aguilera, J. Li, SOSP 2011
- More readings at [SP18 CS525 Website]() (see 3/26 session on "Transactions to rule them all").

***** Industry [4/12]

- (Base) Indy lecture on the interplay between industry and academia, and how it has evolved over the years
- **Paper 1**: Alireza Sahraei, Soteris Demetriou, Amirali Sobhgol, Haoran Zhang, Abhigna Nagaraja, Neeraj Pathak, Girish Joshi, Carla Souza, Bo Huang, Wyatt Cook, Andrii Golovei, Pradeep Venkat, Andrew Mcfague, Dimitrios Skarlatos, Vipul Patel, Ravinder Thind, Ernesto Gonzalez, Yun Jin, Chunqiang Tang: XFaaS: Hyperscale and Low Cost Serverless Functions at Meta. SOSP 2023: 231-246
- **Paper 2**: Marius Eriksen, Kaushik Veeraraghavan, Yusuf Abdulghani, Andrew Birchall, Po-Yen Chou, Richard Cornew, Adela Kabiljo, Ranjith Kumar S., Maroo Lieuw, Justin Meza, Scott Michelson, Thomas Rohloff, Hayley Russell, Jeff Qin, Chunqiang Tang: Global Capacity Management With Flux. OSDI 2023: 589-606

Additional Readings

- Borg: the Next Generation, Muhammad Tirmazi, Adam Barker, Nan Deng, Md Ehtesam Haque, Zhijing Gene Qin, Steven Hand, Mor Harchol-Balter, John Wilkes, Eurosys 2020.
- Autopilot: workload autoscaling at Google, Krzysztof Rzadca, Pawel Findeisen, Jacek Swiderski, Przemyslaw Zych, Przemyslaw Broniek, Jarek Kusmierek, Pawel Nowak, Beata Strack, Piotr Witusowski, Steven Hand, John Wilkes, Eurosys 2020. [direct link]
- Boris Grubic, Yang Wang, Tyler Petrochko, Ran Yaniv, Brad Jones, David Callies, Matt Clarke-Lauer, Dan Kelley, Soteris Demetriou, Kenny Yu, Chunqiang Tang: Conveyor: One-Tool-Fits-All Continuous Software Deployment at Meta. OSDI 2023: 325-342
- Elastic Cloud Services: Scaling Snowflake's Control Plane, Themis Melissaris, Kunal Nabar, Rares Radut, Samir Rehmtulla, Arthur Shi, Samartha Chandrashekar (Snowflake Inc.); Ioannis Papapanagiotou (Gemini Trust), SOCC 2022
- Andrew Newell, Dimitrios Skarlatos, Jingyuan Fan, Pavan Kumar, Maxim Khutornenko, Mayank Pundir, Yirui Zhang, Mingjun Zhang, Yuanlai Liu, Linh Le, Brendon Daugherty, Apurva Samudra, Prashasti Baid, James Kneeland, Igor Kabiljo, Dmitry Shchukin, Andre Rodrigues, Scott Michelson, Ben Christensen, Kaushik Veeraraghavan, Chunqiang Tang: RAS: Continuously Optimized Region-Wide Datacenter Resource Allocation. SOSP 2021: 505-520
- Brad Calder, Ju Wang, Aaron Ogus, Niranjan Nilakantan, Arild Skjolsvold, Sam McKelvie, Yikang Xu, Shashwat Srivastav, Jiesheng Wu, Huseyin Simitci, Jaidev Haridas, Chakravarthy Uddaraju, Hemal Khatri, Andrew Edwards, Vaman Bedekar, Shane Mainali, Rafay Abbasi, Arpit Agarwal, Mian Fahim ul Haq, Muhammad Ikram ul Haq, Deepali Bhardwaj, Sowmya Dayanand, Anitha Adusumilli, Marvin McNett, Sriram Sankaran, Kavitha Manivannan, Leonidas Rigas: Windows Azure Storage: a highly available cloud storage service with strong consistency. SOSP 2011: 143-157
- Sangmin Lee, Zhenhua Guo, Omer Sunercan, Jun Ying, Thawan Kooburat, Suryadeep Biswal, Jun Chen, Kun Huang, Yatpang Cheung, Yiding Zhou, Kaushik Veeraraghavan,

Biren Damani, Pol Mauri Ruiz, Vikas Mehta, Chunqiang Tang: Shard Manager: A Generic Shard Management Framework for Geo-distributed Applications. SOSP 2021: 553-5

- Chunqiang Tang, Kenny Yu, Kaushik Veeraraghavan, Jonathan Kaldor, Scott Michelson, Thawan Kooburat, Aravind Anbudurai, Matthew Clark, Kabir Gogia, Long Cheng, Ben Christensen, Alex Gartrell, Maxim Khutornenko, Sachin Kulkarni, Marcin Pawlowski, Tuomas Pelkonen, Andre Rodrigues, Rounak Tibrewal, Vaishnavi Venkatesan, and Peter Zhang. 2020. Twine: A Unified Cluster Management System for Shared Infrastructure. In 14th USENIX Symposium on Operating Systems Design and Implementation (OSDI 20).
- Borg: the Next Generation, Muhammad Tirmazi, Adam Barker, Nan Deng, Md Ehtesam Haque, Zhijing Gene Qin, Steven Hand, Mor Harchol-Balter, John Wilkes, Eurosys 2020.
- Gandalf: An Intelligent, End-To-End Analytics Service for Safe Deployment in Large-Scale Cloud Infrastructure, Ze Li, Qian Cheng, Ken Hsieh, Yingnong Dang, Peng Huang, Pankaj Singh, Xinsheng Yang, Qingwei Lin, Youjiang Wu, Sebastien Levy, Murali Chintalapati, NSDI 2020.
- Twine: A Unified Cluster Management System for Shared Infrastructure, Chunqiang Tang, Kenny Yu, Kaushik Veeraraghavan, Jonathan Kaldor, Scott Michelson, Thawan Kooburat, Aravind Anbudurai, Matthew Clark, Kabir Gogia, Long Cheng, Ben Christensen, Alex Gartrell, Maxim Khutornenko, Sachin Kulkarni, Marcin Pawlowski, Tuomas Pelkonen, Andre Rodrigues, Rounak Tibrewal, Vaishnavi Venkatesan, Peter Zhang, OSDI 2020.
- Akamai DNS: Providing Authoritative Answers to the World's Queries, Kyle Schomp, Onkar Bhardwaj, Eymen Kurdoglu, Mashooq Muhaimen, Ramesh K. Sitaraman, SIGCOMM 2020.
- POLARDB Meets Computational Storage: Efficiently Support Analytical Workloads in Cloud-Native Relational Database, Wei Cao,Yang Liu,Zhushi Cheng,Ning Zheng,Wei Li, Wenjie Wu,Linqiang Ouyang,Peng Wang, Yijing Wang,Ray Kuan, Zhenjun Liu, Feng Zhu,Tong Zhang, FAST 2020.
- Millions of Tiny Databases, Marc Brooker, Tao Chen Fan Ping, NSDI 2020.
- Amazon Aurora: Design Considerations for High Throughput Cloud-Native Relational Databases, Alexandre Verbitski et al, SIGMOD 2017.
- Social Hash Partitioner: A Scalable Distributed Hypergraph Partitioner, Igor Kabiljo et al, VLDB 2017.
- LittleTable: A Time-Series Database and Its Uses, Sean Rhea et al, SIGMOD 2017.

***** Grounding It (Measurement Studies) [4/17]

Main Papers

- (Base) Orientation lecture by Indy
- **Paper 1**: Borg: the Next Generation, Muhammad Tirmazi, Adam Barker, Nan Deng, Md Ehtesam Haque, Zhijing Gene Qin, Steven Hand, Mor Harchol-Balter, John Wilkes, Eurosys 2020.
- **Paper 2**: Characterizing Smart Home IoT Traffic in the Wild, M. Hammad Mazhar Zubair Shafiq, IoTDI 2020.

Additional Readings

- Adem Efe Gencer, Soumya Basu, Ittay Eyal, Robbert van Renesse, Emin Gün Sirer: Decentralization in Bitcoin and Ethereum Networks. Financial Cryptography 2018: 439-457
- Is Big Data Performance Reproducible in Modern Cloud Networks? Alexandru Uta, Alexandru Custura, Dmitry Duplyakin, Ivo Jimenez, Jan Rellermeyer, Carlos Maltzahn, Robert Ricci, Alexandru Iosup, NSDI 2020.
- Juncheng Yang, Yazhuo Zhang, Ziyue Qiu, Yao Yue, Rashmi Vinayak: FIFO queues are all you need for cache eviction. SOSP 2023: 130-149
- Characterizing, Modeling, and Benchmarking RocksDB Key-Value Workloads at Facebook, Zhichao Cao, Siying Dong, Sagar Vemuri, David H.C. Du, FAST 2020.
- On the Use of ML for Blackbox System Performance Prediction, Silvery Fu, Saurabh Gupta, Radhika Mittal, Sylvia Ratnasamy, NSDI 2021.
- Is Big Data Performance Reproducible in Modern Cloud Networks? Alexandru Uta, Alexandru Custura, Dmitry Duplyakin, Ivo Jimenez, Jan Rellermeyer, Carlos Maltzahn, Robert Ricci, Alexandru Iosup, NSDI 2020.
- Characterizing Smart Home IoT Traffic in the Wild, M. Hammad Mazhar Zubair Shafiq, IoTDI 2020.
- MLaaS in the Wild: Workload Analysis and Scheduling in Large-Scale Heterogeneous GPU Clusters, Qizhen Weng, Hong Kong University of Science and Technology and Alibaba Group; Wencong Xiao, Alibaba Group; Yinghao Yu, Alibaba Group and Hong Kong University of Science and Technology; Wei Wang, Hong Kong University of Science and Technology; Cheng Wang, Jian He, Yong Li, Liping Zhang, Wei Lin, and Yu Ding, Alibaba Group, NSDI 2022
- Werner Vogels: File system usage in Windows NT 4.0. SOSP 1999: 93-109
- Uncovering Access, Reuse, and Sharing Characteristics of I/O-Intensive Files on Large-Scale Production HPC Systems, Tirthak Patel, Suren Byna, Glenn K. Lockwood Nicholas J. Wright, Philip Carns, Robert Ross, Devesh Tiwari, FAST 2020.
- The Standby Energy of Smart Devices: Problems, Progress, & Potential, Wenpeng Wang, Jianyu Su, Zackary Hicks Bradford Campbell, IoTDI 2020.
- On the scale and performance of cooperative web proxy caching, A. Wolman et al, SOSP 1999

- [Scalability! But at what COST?](), Frank McSherry, Michael Isard, Derek G. Murray, HotOS 2015.
- Cloudera, What do Hadoop Workloads Really Look Like?
    - Webarchive link (missing images): https://web.archive.org/web/20121010095149/http://www.cloudera.com/blog/2012/09/what-do-real-life-hadoop-workloads-look-like
    - Indy's slides capturing the Cloudera study (includes images) - see slides 21-29: https://courses.engr.illinois.edu/cs425/fa2015/L23.FA15.ppt OR https://courses.engr.illinois.edu/cs425/fa2015/L23.FA15.pdf
    - (original link, broken) `http://blog.cloudera.com/blog/2012/09/what-do-real-life-hadoop-workloads-look-like/`

***** Modern RAIDs [4/19].

- (Base) Orientation lecture by Indy - [What is RAID](#)
- **Paper 1**: Thomas Kim, Jekyeom Jeon, Nikhil Arora, Huaicheng Li, Michael Kaminsky, David G. Andersen, Gregory R. Ganger, George Amvrosiadis, Matias Bjørling: [RAIZN: Redundant Array of Independent Zoned Namespaces](#). ASPLOS (2) 2023: 660-673
- **Paper 2**: Junyi Shu, Ruidong Zhu, Yun Ma, Gang Huang, Hong Mei, Xuanzhe Liu, Xin Jin: [Disaggregated RAID Storage in Modern Datacenters](#). ASPLOS (3) 2023: 147-163

Additional Readings

- [RACS: a case for cloud storage diversity](#), H. Abu-Libdeh et al, SOCC 2010

***** Potpourri: MPC and ML/Systems [4/24]

- (Base) Orientation lecture by Indy
    - [How (and how not) to write a good SOSP paper](), R. Levin and D. D. Redell, 1983
- *For this day only - Only 1 paper review per session (out of 2 papers)*
- **Paper 1**: John Liagouris, Vasiliki Kalavri, Muhammad Faisal, Mayank Varia: [SECRECY: Secure collaborative analytics in untrusted clouds](). NSDI 2023: 1031-1056
- **Paper 2**: Zhuang Wang, Zhen Jia, Shuai Zheng, Zhen Zhang, Xinwei Fu, T. S. Eugene Ng, Yida Wang: [GEMINI: Fast Failure Recovery in Distributed Training with In-Memory Checkpoints](). SOSP 2023: 364-381

Additional Readings

- Chieh-Jan Mike Liang, Zilin Fang, Yuqing Xie, Fan Yang, Zhao Lucis Li, Li Lyna Zhang, Mao Yang, Lidong Zhou: [On Modular Learning of Distributed Systems for Predicting End-to-End Latency](). NSDI 2023: 1081-1095
- Insu Jang, Zhenning Yang, Zhen Zhang, Xin Jin, Mosharaf Chowdhury: [Oobleck: Resilient Distributed Training of Large Models Using Pipeline Templates](). SOSP 2023: 382-395
- Suhas Jayaram Subramanya, Daiyaan Arfeen, Shouxu Lin, Aurick Qiao, Zhihao Jia, Gregory R. Ganger: [Sia: Heterogeneity-aware, goodput-optimized ML-cluster scheduling](). SOSP 2023: 642-657
- Kshiteej Mahajan, Ching-Hsiang Chu, Srinivas Sridharan, Aditya Akella: Better Together: [Jointly Optimizing ML Collective Scheduling and Execution Planning using SYNDICATE](). NSDI 2023: 809-824

***** Final Project Presentations - Session 1 of 2 [4/26]

***** Final Project Presentations - Session 2 of 2 [5/1]

<<< **Final Report Due 5/5** >>>

Leftover Topics (Browse for More Project Ideas!)

The following topics are extremely interesting but could not be included due to limited number of sessions. Please read them for more project ideas!

Leftover Topic: Potpourri: DSM, Far Memory, and Tails

- (Base) Lecture by Indy on DSM and its history, cross-layer research, and disaggregated distributed systems.
- Paper 1: Carbink: Fault-Tolerant Far Memory, Yang Zhou, Harvard University; Hassan M. G. Wassel, Google; Sihang Liu, University of Virginia; Jiaqi Gao and James Mickens, Harvard University; Minlan Yu, Harvard University and Google; Chris Kennelly, Paul Turner, and David E. Culler, Google; Henry M. Levy, University of Washington and Google; Amin Vahdat, Google, OSDI 2022
- Paper 2: Henri Maxime Demoulin, Joshua Fried, Isaac Pedisich, Marios Kogias, Boon Thau Loo, Linh Thi Xuan Phan, Irene Zhang: When Idling is Ideal: Optimizing Tail-Latency for Heavy-Tailed Datacenter Workloads with Perséphone. 621-637, SOSP 2021

  Additional Readings

- The Tail at Scale, Jeffrey Dean Luiz André Barroso Communications of the ACM, vol. 56 (2013), pp. 74-80, 2013.
- Hilfi Alkaff, Indranil Gupta, Luke M. Leslie: Cross-Layer Scheduling in Cloud Systems. IC2E 2015: 236-245
- Khaled Diab, Parham Yassini, and Mohamed Hefeeda, Simon Fraser University, Orca: Server-assisted Multicast for Datacenter Networks, NSDI 2022
- Can Far Memory Improve Job Throughput? Emmanuel Amaro, Christopher Branner-Augmon, Zhihong Luo, Amy Ousterhout, Marcos K. Aguilera, Aurojit Panda, Sylvia Ratnasamy, Scott Shenker, Eurosys 2020.
- Accessible Near-Storage Computing with FPGAs, Robert Schmid, Max Plauth, Lukas Wenzel, Felix Eberhardt, Andreas Polze, Eurosys 2020.
- Decibel: Isolation and Sharing in Disaggregated Rack-Scale Storage, Mihir Nanavati et al, NSDI 2017.
- Hailstorm: Disaggregated Compute and Storage for Distributed LSM-based Databases, Laurent Bindschaedler, Laurent Bindschaedler, Ashvin Goel, Ashvin Goel, Willy E Zwaenepoel, Willy Zwaenepoel March 2020, ASPLOS 2020.

Leftover Topic: Wrap Up (Cool papers to read!)

- World Brain, H. G. Wells, 1937
- The tragedy of the commons, G. Hardin, 1968
- How (and how not) to write a good SOSP paper, R. Levin and D. D. Redell, 1983

Additional Readings

- R. Hoffmann, "Why buy that theory?", 2003
- R. P. Feynman, "Metaplast Corp."

Leftover Topic: Transactions

- (Base) Indy lecture on history of transactions and key-value stores
- (**You don't need to review this paper on 4/7. can review this paper on 4/21. If you've already submitted a review on Piazza, you can keep it there but please remember to resubmit that review on 4/21.**) Paper 1: The SNOW Theorem and Latency-Optimal Read-Only Transactions, H. Lu et al, OSDI 2016.
- (No Paper 2)

  Additional Readings

- Dai Qin, Angela Demke Brown, Ashvin Goel: Caracal: Contention Management with Deterministic Concurrency Control. SOSP 2021: 180-194
- Cobra: Making Transactional Key-Value Stores Verifiably Serializable, Cheng Tan, Changgeng Zhao, Shuai Mu, Michael Walfish, OSDI 2020.
- Kvell+: Snapshot Isolation without Snapshots, Baptiste Lepers, Oana Balmau, Karan Gupta, Willy Zwaenepoel, OSDI 2020.
- Meerkat: Multicore-Scalable Replicated Transactions Following the Zero-Coordination Principle, Adriana Szekeres, Michael Whittaker, Naveen Kr. Sharma, Jialin Li, Arvind Krishnamurthy, Dan R. K. Ports, Irene Zhang, Eurosys 2020.
- Transactional storage for geo-replicated systems, Y. Sovran, R. Power, M. K. Aguilera, J. Li, SOSP 2011
- More readings at SP18 CS525 Website (see 3/26 session on "Transactions to rule them all").

Leftover Topic: Virtual/View Synchrony

- (Base) Orientation lecture by Indy
  - Kenneth P. Birman, Thomas A. Joseph: Exploiting Virtual Synchrony in Distributed Systems. SOSP 1987: 123-138

- Paper 1: Louise E. Moser, P. M. Melliar-Smith, Deborah A. Agarwal, Ravi K. Budhia, Colleen A. Lingley-Papadopoulos: Totem: A Fault-Tolerant Multicast Group Communication System. CACM 39(4): 54-63 (1996) https://dl.acm.org/doi/abs/10.1145/227210.227226
- Paper 2: M. Castro, B, Liskov, Practical Byzantine Fault-Tolerance, OSDI 1999.

  Additional Readings

- Barbara Liskov and James Cowling, Viewstamped Replication Revisited, 2017
- Indranil Gupta, Kenneth P. Birman, Robbert van Renesse, Fighting fire with fire: using randomized gossip to combat stochastic scalability limits, QREI, May 2002.
- CATOCS Controversy Papers
  - Kenneth P. Birman, Thomas A. Joseph: Exploiting Virtual Synchrony in Distributed Systems. SOSP 1987: 123-138
  - Understanding the limitations of causally and totally ordered communication, Cheriton and Skeen, SIGOPS OSR 1993
  - A response to Cheriton and Skeen's criticism of causal and totally ordered communication, Birman, SIGOPS OSR 1994
  - Why bother with CATOCS? R. van Renesse, SIGOPS OSR 1994.

Leftover Topic: Measurements

- Measurement, modeling, and analysis of a peer-to-peer file-sharing workload Krishna P. Gummadi et al, SOSP 2003
- Understanding availability, R. Bhagwan et al, IPTPS 2003
- Understanding Overlay Characteristics of a Large-scale Peer-to-Peer IPTV System, Long Vu, Indranil Gupta, Klara Nahrstedt, Jin Liang. ACM Transactions on Multimedia Computing, Communications and Applications (TOMCCAP), vol. 6, no. 4, 2009.
- What do Real-Life Hadoop Workloads Look Like, Cloudera Blog
  - Web archive link (figures missing)
- (if time permits) An Evaluation of Amazon's Grid Computing Services: EC2, S3 and SQS, Simson Garfinkel, Harvard TechRep., 2007

  Additional Readings

- Hadoop's Adolescence: An analysis of Hadoop usage in scientific workloads, K. Ren, Y. Kwon, M. Balazinska, B. Howe, VLDB 2013
- Availability and Locality Measurements of Peer-to-Peer File Systems, J. Chu et al, SPIE 2002.
- A measurement study of peer-to-peer file sharing systems, S. Saroui et al, MMCN 2002
- Free riding on Gnutella
  Adar and Huberman, First Monday, 2000
- Small-world file sharing communities, A Iamnitchi et al, Infocom 2004.

Leftover Topic: Modern OSs

- (Base) Orientation lecture by Indy
    - [How (and how not) to write a good SOSP paper](#), R. Levin and D. D. Redell, 1983
- Paper 1: Irene Zhang, Amanda Raybuck, Pratyush Patel, Kirk Olynyk, Jacob Nelson, Omar S. Navarro Leija, Ashlie Martinez, Jing Liu, Anna Kornfeld Simpson, Sujay Jayakar, Pedro Henrique Penna, Max Demoulin, Piali Choudhury, Anirudh Badam: [The Demikernel Datapath OS Architecture for Microsecond-scale Datacenter Systems](#). SOSP 2021: 195-211 [[Alt Link](#)]
- Paper 2: [Twine: A Unified Cluster Management System for Shared Infrastructure](#), Chunqiang Tang, Kenny Yu, Kaushik Veeraraghavan, Jonathan Kaldor, Scott Michelson, Thawan Kooburat, Aravind Anbudurai, Matthew Clark, Kabir Gogia, Long Cheng, Ben Christensen, Alex Gartrell, Maxim Khutornenko, Sachin Kulkarni, Marcin Pawlowski, Tuomas Pelkonen, Andre Rodrigues, Rounak Tibrewal, Vaishnavi Venkatesan, Peter Zhang, OSDI 2020.

    Additional Readings

- [SkyPilot: An Intercloud Broker for Sky Computing](#), Zongheng Yang, Zhanghao Wu, Michael Luo, Wei-Lin Chiang, Romil Bhardwaj, Woosuk Kwon, Siyuan Zhuang, Frank Sifei Luan, Gautam Mittal, Scott Shenker, Ion Stoica, NSDI 2023.
- [An Intercloud Broker for Sky Computing](#), Zongheng Yang, Zhanghao Wu, Michael Luo, Wei-Lin Chiang, Romil Bhardwaj, Woosuk Kwon, Siyuan Zhuang, Frank Sifei Luan, and Gautam Mittal, UC Berkeley; Scott Shenker, ICSI and UC Berkeley; Ion Stoica, UC Berkeley, NSDI 2023.
- [Mesos: a platform for fine-grained resource sharing in the data center](#), B. Hindman et al, NSDI 2011
- [Ray: A Distributed Framework for Emerging AI Applications](#), Philipp Moritz, Robert Nishihara, Stephanie Wang, Alexey Tumanov, Richard Liaw, Eric Liang, Melih Elibol, Zongheng Yang, William Paul, Michael I. Jordan, Ion Stoica, OSDI 2018.
- [Borg: the Next Generation](#), Muhammad Tirmazi, Adam Barker, Nan Deng, Md Ehtesam Haque, Zhijing Gene Qin, Steven Hand, Mor Harchol-Balter, John Wilkes, Eurosys 2020.
- Brendan Burns, Brian Grant, David Oppenheimer, Eric Brewer, John Wilkes, [Borg, Omega, and Kubernetes](#), ACM Queue, vol. 14 (2016), pp. 70-93, 2016.

(END OF DOCUMENT - CS525 SP24 Reading List)