Related Works

Mark Madler

1 Replicated Data Stores over RDMA

1.1 Kite: efficient and available release consistency for the datacenter

This is the real entry paper. This is a replicated KVS over RDMA.[8]

1.2 Hermes: A Fast, Fault-Tolerant and Linearizable Replication Protocol

Fault tolerant protocol for replication of datastore. Linearizable. [13]

1.3 Hamband: RDMA Replicated Data Types

This is obviously a replicated system.[11]

1.4 FaRM: Fast Remote Memory

Super similar to the entry paper. A lot of other papers call it DSM so I am not totally sure about this. I think this still caches data at the page level.[3]

1.5 FaRMv2: Fast General Distributed Transactions with Opacity

Just like FaRM but with opacity. Also providing strict serializability.[15]

2 DSM systems

2.1 Scaling out NUMA-Aware Applications with RDMA-Based Distributed Shared Memory: MAGI

Page-based DSM again. [10]

2.2 Gengar: An RDMA-based Distributed Hybrid Memory Pool

This is object based dsm over rdma but with non-volatile memory as well using Intel Optane. Seems to also use this lease assignment idea like in [5] but is not page based.[4]

2.3 MENPS: A Decentralized Distributed Shared Memory Exploiting RDMA

- Page based DSM
- Special Diff merging and page sharing
- Combine write notices and logical leases (what is that?)[5]

2.4 Argo DSM (Turning Centralized Coherence and Distributed Critical-Section Execution on their Head: A New Approach for Scalable Distributed Shared Memory)

Page-based DSM again but directory coherence. This was maybe the first RDMA-based DSM paper, at least thats what the authors allude to.[14]

2.5 Evaluation of RDMA opportunities in an Object-Oriented DSM

entry paper? Interesting result is that it proves that invalidation protocols are better suited for distributed systems. [18]

3 Closely Related Works object based / not page-level granularity

3.1 Odyssey: The Impact of Modern Hardware on Strongly-Consistent Replication Protocols

This is the larger project that Kite is a part of. Not really an implemented work itself but it is sort of a survey of protocols used.[7]

3.2 Efficient Distributed Memory Management with RDMA and Caching

cache-line granularity.[2]

3.3 Distributed Shared Object Memory

object based granularity, release consistency... too old for RDMA[9]

3.4 Scalable RDMA performance in PGAS languages

This paper is for PGAS languages. Has an address hash table similar to LOCO for remote lookups.[6]

4 Loosely Related but Evaluated

4.1 TreadMarks: shared memory computing on networks of workstations

Was implemented over IP, lazy release consistency I think. Not sure of granularity yet.[1]

4.2 CoRM: Compactable Remote Memory over RDMA

page based I think (re-read this)[16]

4.3 LITE Kernel RDMA Support for Datacenter Applications

This is page based DSM using the kernel. [17]

4.4 GiantVM: A Novel Distributed Hypervisor for Resource Aggregation with DSM-aware Optimizations

Page-based DSM again but also works over TCP and RDMA[12]

5 Evaluated but not Related

5.1 Rcmp: Reconstructing RDMA-Based Memory Disaggregation via CXL

page based and uses CXL, not comparable[19]

References

[1] Amza, C., Cox, A., Dwarkadas, S., Keleher, P., Lu, H., Rajamony, R., Yu, W., and Zwaenepoel, W. Treadmarks: shared memory computing on networks of workstations. *Computer 29*, 2 (1996), 18–28.

[2] Cai, Q., Guo, W., Zhang, H., Agrawal, D., Chen, G., Ooi, B. C., Tan, K.-L., Teo, Y. M., and Wang, S. Efficient distributed memory management with rdma and caching. *Proc. VLDB Endow. 11*, 11 (July 2018), 1604–1617.

[3] Dragojević, A., Narayanan, D., Hodson, O., and Castro, M. Farm: fast remote memory. In *Proceedings of the 11th USENIX Conference on Networked Systems Design and Implementation* (USA, 2014), NSDI'14, USENIX Association, p. 401–414.

[4] Duan, Z., Liu, H., Lu, H., Liao, X., Jin, H., Zhang, Y., and He, B. Gengar: An rdma-based distributed hybrid memory pool. In *2021 IEEE 41st International Conference on Distributed Computing Systems (ICDCS)* (2021), pp. 92–103.

[5] Endo, W., Sato, S., and Taura, K. Menps: A decentralized distributed shared memory exploiting rdma. In *2020 IEEE/ACM Fourth Annual Workshop on Emerging Parallel and Distributed Runtime Systems and Middleware (IPDRM)* (2020), pp. 9–16.

[6] Farreras, M., Almasi, G., Cascaval, C., and Cortes, T. Scalable rdma performance in pgas languages. pp. 1–12.

[7] Gavrielatos, V., Katsarakis, A., and Nagarajan, V. Odyssey: the impact of modern hardware on strongly-consistent replication protocols. In *Proceedings of the Sixteenth European Conference on Computer Systems* (New York, NY, USA, 2021), EuroSys '21, Association for Computing Machinery, p. 245–260.

[8] Gavrielatos, V., Katsarakis, A., Nagarajan, V., Grot, B., and Joshi, A. Kite: efficient and available release consistency for the datacenter. In *Proceedings of the 25th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming* (New York, NY, USA, 2020), PPoPP '20, Association for Computing Machinery, p. 1–16.

[9] Guedes, P., and Castro, M. Distributed shared object memory. In *Proceedings of IEEE 4th Workshop on Workstation Operating Systems. WWOS-III* (1993), pp. 142–149.

[10] Hong, Y., Zheng, Y., Yang, F., Zang, B.-Y., Guan, H.-B., and Chen, H.-B. Scaling out numa-aware applications with rdma-based distributed shared memory. *Journal of Computer Science and Technology 34* (2019), 94–112.

[11] Houshmand, F., Saberlatibari, J., and Lesani, M. Hamband: Rdma replicated data types. In *Proceedings of the 43rd ACM SIGPLAN International Conference on Programming Language Design and Implementation* (New York, NY, USA, 2022), PLDI 2022, Association for Computing Machinery, p. 348–363.

[12] Jia, X., Zhang, J., Yu, B., Qian, X., Qi, Z., and Guan, H. Giantvm: A novel distributed hypervisor for resource aggregation with dsm-aware optimizations. *ACM Trans. Archit. Code Optim. 19*, 2 (Mar. 2022).

[13] Katsarakis, A., Gavrielatos, V., Katebzadeh, M. S., Joshi, A., Dragojevic, A., Grot, B., and Nagarajan, V. Hermes: A fast, fault-tolerant and linearizable replication protocol. In *Proceedings of the Twenty-Fifth International Conference on Architectural Support for Programming Languages and Operating Systems* (New York, NY, USA, 2020), ASPLOS '20, Association for Computing Machinery, p. 201–217.

[14] Kaxiras, S., Klaftenegger, D., Norgren, M., Ros, A., and Sagonas, K. Turning centralized coherence and distributed critical-section execution on their head: A new approach for scalable distributed shared memory. In *Proceedings of the 24th International Symposium on High-Performance Parallel and Distributed Computing* (New York, NY, USA, 2015), HPDC '15, Association for Computing Machinery, p. 3–14.

[15] Shamis, A., Renzelmann, M., Novakovic, S., Chatzopoulos, G., Dragojević, A., Narayanan, D., and Castro, M. Fast general distributed transactions with opacity. In *Proceedings of the 2019 International Conference on Management of Data* (New York, NY, USA, 2019), SIGMOD '19, Association for Computing Machinery, p. 433–448.

[16] Taranov, K., Di Girolamo, S., and Hoefler, T. Corm: Compactable remote memory over rdma. In *Proceedings of the 2021 International Conference on Management of Data* (New York, NY, USA, 2021), SIGMOD '21, Association for Computing Machinery, p. 1811–1824.

[17] Tsai, S.-Y., and Zhang, Y. Lite kernel rdma support for datacenter applications. In *Proceedings of the 26th Symposium on Operating Systems Principles* (New York, NY, USA, 2017), SOSP '17, Association for Computing Machinery, p. 306–324.

[18] Veldema, R., and Philippsen, M. Evaluation of rdma opportunities in an object-oriented dsm. pp. 217–231.

[19] Wang, Z., Guo, Y., Lu, K., Wan, J., Wang, D., Yao, T., and Wu, H. Rcmp: Reconstructing rdma-based memory disaggregation via cxl. *ACM Trans. Archit. Code Optim. 21*, 1 (Jan. 2024).