Belief-Based Storage Systems

Dušan Ramljak
Temple University

Krishna Kant
Temple University

Motivation: The current data growth and consequent increase in complexity of its usage patterns indicate that intelligent management of data storage is becoming ever more crucial [7, 20, 23]. We rely on the claim that efficient pattern discovery and description, coupled with the observed predictability of complex patterns within many high-performance applications, offers significant potential to enable many I/O optimizations [11]. We developed a compact flexible caching and pre-fetching framework [17] that could, potentially address any imposed reliability, performance, energy efficiency requirement and have the ability to add any relevant information. Here, we discuss possible ways to extend this framework towards belief based storage systems.

Approach: In our unified caching and pre-fetching framework, that incorporates Bayesian reasoning, belief is an estimate for probability of a hit and is based on the relationships gauged from the accesses to data entities. (Hereinafter, entity could be a file, file-segment, block, object, etc.) Internal parameters of the algorithm relate to the workload characteristics [13] and are automatically adjusted offline, while the belief values are improved in an online fashion. Integrated use of belief guides both pre-fetch and eviction, as we pre-fetch entities with highest belief values, and evict ones with lowest belief values. Our framework prevents that pre-fetched data replaces more useful data from the cache and allows us to control the trade-off of expected latency and related data movement.

Results that Motivated the Idea: For the workloads we examined, the resulting framework provides the opportunity to capture complex non-sequential access patterns better than a state-of-the-art framework for optimizing cloud storage gateways [23, 8, 9]. Moreover, our framework is also able to adjust to variations in the workload faster, and does not require a static workload to be effective. However, the results of the adjustment largely depend on the internal parameters of the algorithm. For the perturbations of the same workload we want to inherit training, but not belief values. When we face a phase change (variations that basically make a different workload), we want to inherit neither belief values nor training. Looking for ways to solve this problem led us to the idea of belief based storage systems.

Idea: First extension towards belief-based storage systems is to consider to leverage provenance information. Provenance is the metadata that represents the history or lineage of a data entity. Provenance as a subject has gained high visibility, a number of provenance-related aspects have been explored [24, 19, 3, 6, 5, 4, 10, 12, 15, 18, 16, 1, 22, 21, 14, 2]. Provenance information could improve belief calculation by adding structural relationships. More importantly it could help identify which group of data we are working with, which allows us to set internal parameters correctly and improve performance of our framework.

Next, in each storage system we could identify at least three grouping levels (e.g. Block level, File level, Groups of files level). Even without having beliefs on a lower level we could bias lower level caching algorithm based on the upper level beliefs, but we would like to coordinate beliefs, form meta-belief that will guide all decisions and connect different use cases and/or multiple set of beliefs. In this hierarchical organization we would like to find ways to discover phase change.

Lastly, the idea could be applied to more than just caching. In belief based data placement, we could have an intelligent data movement mechanism across different types of devices based on device and access characteristics. Belief based redundancy removal could be seen as generalization of caching - what representation of data should be stored. Other storage subsystems and environments could benefit from belief-based framework that exposes and leverages structural and behavioral relationships.

References

[1] ABREU, R., ARCHER, D., CHAPMAN, E., CHENEY, J., ELDARDIRY, H., AND GASCÓN, A. Provenance segmentation. In *8th USENIX Workshop on the Theory and Practice of Provenance (TaPP 16)* (Washington, D.C., June 2016), USENIX Association.

[2] AINY, E., BOURHIS, P., DAVIDSON, S. B., DEUTCH, D., AND MILO, T. Approximated summarization of data provenance. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management* (2015), CIKM '15.

[3] BACHOUR, K., WETZEL, R., FLINTHAM, M., HUYNH, T. D., RODDEN, T. A., AND MOREAU, L. Provenance for the people: A user-centered look at the w3c prov standard through an online game.

[4] BATES, A., MOOD, B., VALAFAR, M., AND BUTLER, K. Towards secure provenance-based access control in cloud environments. In *Proceedings of the third ACM conference on Data and application security and privacy* (2013), ACM, pp. 277–284.

[5] BATES, A. M., TIAN, D., BUTLER, K. R., AND MOYER, T. Trustworthy whole-system provenance for the linux kernel. In *Usenix Security* (2015), pp. 319–334.

[6] BERTINO, E., GHINITA, G., KANTARCIOGLU, M., NGUYEN, D., PARK, J., SANDHU, R., SULTANA, S., THURAISINGHAM, B., AND XU, S. A roadmap for privacy-enhanced secure data provenance. *Journal of Intelligent Information Systems 43*, 3 (2014), 481–501.

[7] BREWER, E., YING, L., GREENFIELD, L., CYPHER, R., AND T'SO, T. Disks for data centers.

[8] GILL, B. S., AND BATHEN, L. A. D. Amp: Adaptive multi-stream prefetching in a shared cache. In *FAST* (2007), vol. 7, pp. 185–198.

[9] GILL, B. S., AND MODHA, D. S. Sarc: Sequential prefetching in adaptive replacement cache. In *USENIX Annual Technical Conference, General Track* (2005), pp. 293–308.

[10] GROTH, P., JIANG, S., MILES, S., MUNROE, S., TAN, V., TSASAKOU, S., AND MOREAU, L. An architecture for provenance systems.

[11] HE, J., BENT, J., TORRES, A., GRIDER, G., GIBSON, G., MALTZAHN, C., AND SUN, X.-H. I/o acceleration with pattern detection. In *Proceedings of the 22nd international symposium on High-Performance Parallel and Distributed Computing* (2013), ACM, pp. 25–36.

[12] JAYAPANDIAN, M., CHAPMAN, A., TARCEA, V. G., YU, C., ELKISS, A., IANNI, A., LIU, B., NANDI, A., SANTOS, C., ANDREWS, P., ET AL. Michigan molecular interactions (mimi): putting the jigsaw puzzle together. *Nucleic acids research 35*, suppl 1 (2007), D566–D571.

[13] KAPLAN, S. F., MCGEOCH, L. A., AND COLE, M. F. Adaptive caching for demand prepaging. In *ACM SIGPLAN Notices* (2002), vol. 38, ACM, pp. 114–126.

[14] MOREAU, L. Aggregation by provenance types: A technique for summarising provenance graphs. *arXiv preprint arXiv:1504.02616* (2015).

[15] MUNISWAMY-REDDY, K.-K., BRAUN, U., HOLLAND, D. A., MACKO, P., MACLEAN, D. L., MARGO, D. W., SELTZER, M. I., AND SMOGOR, R. Layering in provenance systems. In *USENIX Annual technical conference* (2009).

[16] MUNISWAMY-REDDY, K.-K., HOLLAND, D. A., BRAUN, U., AND SELTZER, M. I. Provenance-aware storage systems.

[17] RAMLJAK, D., ALAZZAWE, A., UVERSKY, A., AND KANT, K. Belief-based data pre-fetching and replacement in storage systems. In *The 15th USENIX Conference on File and Storage Technologies (FAST), Work in progress (WiP) session, Santa Clara, Feb 27 - Mar 2* (2017).

[18] SIMMHAN, Y. L., PLALE, B., AND GANNON, D. A framework for collecting provenance in data-centric scientific workflows. In *2006 IEEE International Conference on Web Services (ICWS'06)* (2006), IEEE, pp. 427–436.

[19] SULTANA, S., AND BERTINO, E. A distributed system for the management of fine-grained provenance. *Journal of Database Management (JDM) 26*, 2 (2015), 32–47.

[20] TURNER, V., GANTZ, J. F., REINSEL, D., AND MINTON, S. The digital universe of opportunities: rich data and the increasing value of the internet of things. *IDC Analyze the Future* (2014).

[21] XIE, Y., MUNISWAMY-REDDY, K.-K., FENG, D., LI, Y., AND LONG, D. D. Evaluation of a hybrid approach for efficient provenance storage. *ACM Transactions on Storage (TOS) 9*, 4 (2013), 14.

[22] XIE, Y., MUNISWAMY-REDDY, K.-K., LONG, D. D., AMER, A., FENG, D., AND TAN, Z. Compressing provenance graphs. In *TaPP* (2011).

[23] YANG, S., SRINIVASAN, K., UDAYASHANKAR, K., KRISHNAN, S., FENG, J., ZHANG, Y., ARPACI-DUSSEAU, A. C., AND ARPACI-DUSSEAU, R. H. Tombolo: Performance enhancements for cloud storage gateways. In *Proceedings of the 32nd International Conference on Massive Storage Systems and Technology (MSST 2016)* (2016).

[24] ZHOU, W., SHERR, M., TAO, T., LI, X., LOO, B. T., AND MAO, Y. Efficient querying and maintenance of network provenance at internet-scale. In *Proceedings of the 2010 ACM SIGMOD International Conference on Management of data* (2010), ACM, pp. 615–626.