

A Lightweight Data Location Service for Nondeterministic Exascale Storage Systems

[Extended Abstract]

Zhiwei Sun, Anthony Skjellum
University of Alabama at Birmingham
CH 115, 1300 University Blvd
Birmingham, AL, 35205
{zhwsun, tony}@cis.uab.edu

Lee Ward, Matthew Curry
Sandia National Laboratories
PO Box 5800
Albuquerque, NM 87185-1319
{lee, mlcurry}@sandia.gov

Abstract

We present LWDLS, a lightweight data location service designed for Exascale storage systems (storage systems with order of 10^{18} bytes) and geo-distributed storage systems (large storage systems with physically distributed locations). LWDLS provides a search-based data location solution, and enables free data placement, movement, and replication. Probe and prune protocols are introduced that reduce topology mismatch, and a heuristic flooding search algorithm (HFS) is presented with higher search efficiency than pure flooding search while delivering comparable search speed and coverage to pure flooding search.

LWDLS is lightweight and scalable in terms of incorporating low overhead, high search efficiency, no global state, and by avoiding periodic messages. LWDLS is fully distributed and can be used in nondeterministic storage systems and in deterministic storage systems.

Extensive simulations modeling large-scale High Performance Computing (HPC) storage environments provide representative performance outcomes.

1. PROBE AND PRUNE PROTOCOLS

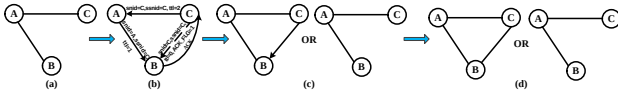


Figure 1: Example of probing operation.

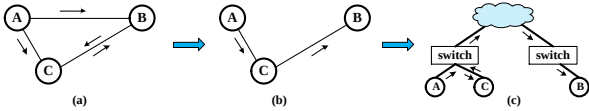


Figure 2: Pruning operation based on the ratio of the lengths of two adjacent edges of a triangle.

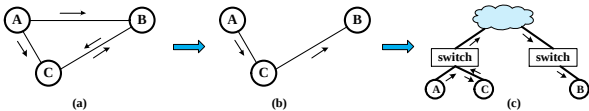


Figure 3: Examples of pruning operations.

2. STATE TRANSITION AMONG THE TYPES OF SERVER LISTS IN LWDLS

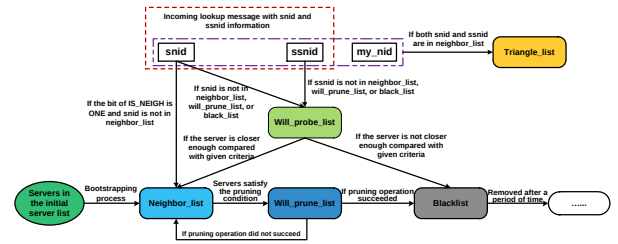


Figure 4: Diagram of state transitions for servers among the types of server lists used in LWDLS.

3. A HEURISTIC FLOODING SEARCH (HFS) ALGORITHM

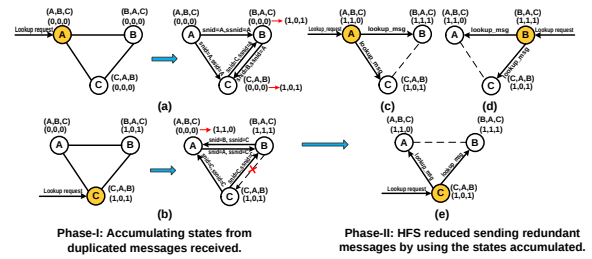


Figure 5: Accumulating states for reducing redundant messages with the support of triangles detected and knowledge bits encoded.

4. CONCLUSIONS

- LWDLS is a search-based data location service, enables free data placement, movement, and replication.
- LWDLS is able to locate data efficiently in nondeterministic Exascale storage systems.
- LWDLS is lightweight, efficient, and scalable, in terms of avoiding global state, periodic messages, or the limitations on the location of data.

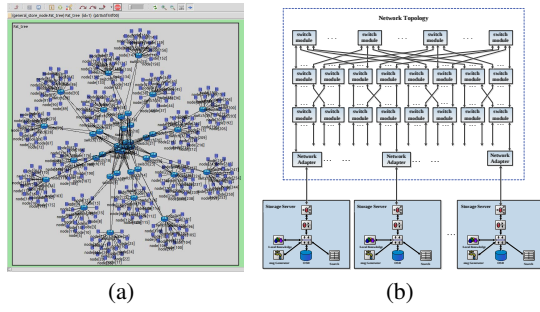


Figure 6: Simulator.

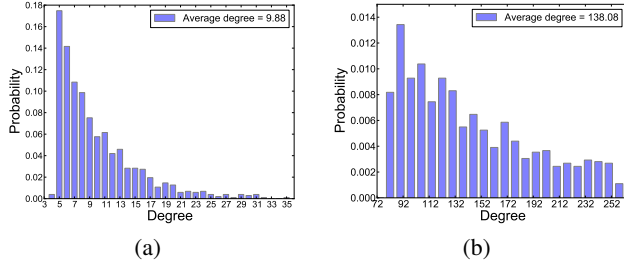


Figure 7: Two overlay network graphs used for testing LWDLs.

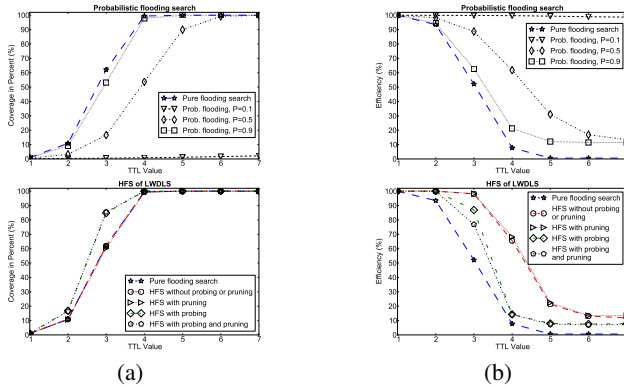


Figure 8: Comparison of search performance on the overlay network with $M=5$.

- With the probe and prune protocols and HFS algorithm, LWDLs is able to address problems of topology mismatch and inefficient search performance of pure flooding search algorithm.
- LWDLs provides higher search efficiency than pure flooding search while having comparable search speed and search coverage.
- The effectiveness of LWDLs has been tested through extensive simulations modeling large-scale HPC storage environments.

Acknowledgments

This work was supported in part by the United States Department of Energy under Contract DE-AC04-94AL85000. This work was

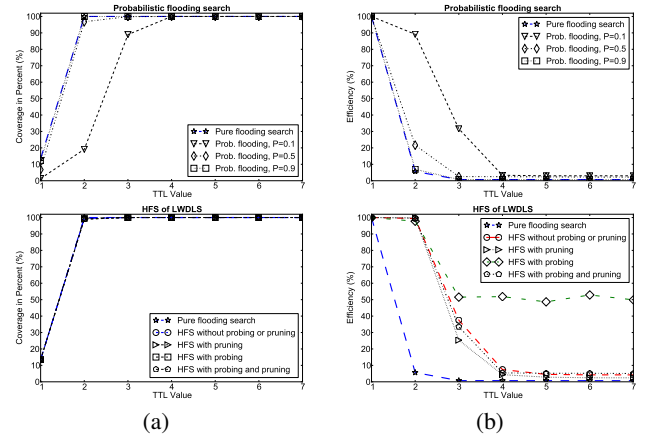


Figure 9: Comparison of search performance on the network graph with $M=80$.

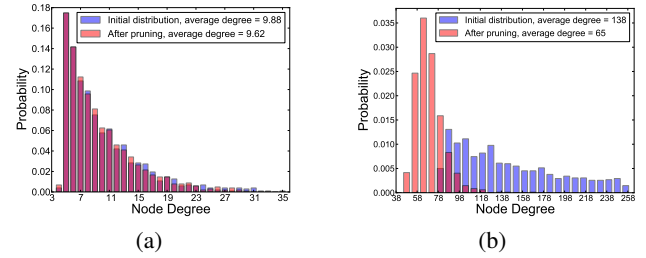


Figure 10: Effects of probing and pruning operations on the node degree distribution, tested on the two overlay network topologies.

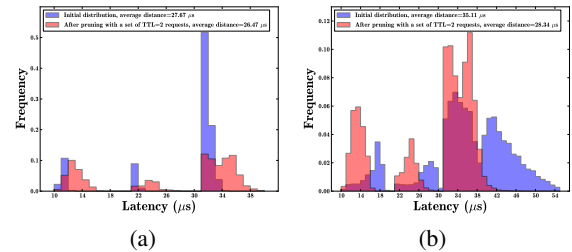


Figure 11: Distribution of neighbor distance after pruning.

supported in part by the National Science Foundation under grant OCI-1064247 and grant CCF-1239962. We want to thank UAB CIS IT staff for maintaining the Linux clusters used and providing timely supports.