

Review: “It’s Time to Talk About HPC Storage: Perspectives on the Past and Future”

Kathryn Mohror¹

¹Affiliation not available

September 8, 2021

This is a review of the article *“It’s Time to Talk About HPC Storage: Perspectives on the Past and Future,”* for the Leadership Computing department.

Comments:

Great article! I have only a few small suggestions for clarity/readability below.

Abstract: suggest changing “especially as relates to the use of HPC” to “especially related to the use of HPC”

paragraph 2: ” ORNL Health Data Sciences Institute” -> “Oak Ridge National Laboratory (ORNL) Health Data Sciences Institute”

Fig 2 caption: “Each access must traverse multiple distinct protocol hops, with its own interrupt processing, buffering, handshaking, serialization, and access control conventions ” The “its” is ambiguous here, suggest ” Each access must traverse multiple distinct protocol hops, where each protocol hop has its own interrupt processing, buffering, handshaking, serialization, and access control conventions .”

First paragraph in section “Emerging Challenges in HPC Storage”: ” Specifically, storage servers were designed to mediate all access to hard drives. By doing so, they could shape traffic (e.g., serializing and batching), buffer data (e.g., caching based on locality), and process I/O requests on more powerful host CPUs (e.g., handling interrupts, packing and unpacking Remote Procedure Call (RPC) requests, and enforcing authorization) to make the most of hard drive capabilities.“ Suggest a couple filler words for clarity: ” Specifically, storage servers were designed to mediate all access to hard drives. By doing so, they could shape traffic (e.g., by serializing and batching requests), buffer data (e.g., through caching based on locality), and process I/O requests on more powerful host CPUs (e.g., by handling interrupts, packing and unpacking Remote Procedure Call (RPC) requests, and enforcing authorization) to make the most of hard drive capabilities. ”

Subsection headers throughout use inconsistent capitalization. I am not certain on the magazine style preference but from the department template it looks like capitalizing the first word and lowercase for the rest of the words is preferred for subsections, e.g., “Low Latency access to Storage” => “Low latency access to storage”

In the paragraph beginning ” From a device interface perspective, storage system“, suggest adding a comma to this clause between ”computation” and “reducing”: “. . . that affects device performance and cost, and as storage becomes disaggregated from computation, reducing data movement between the device and host becomes crucial. ”

Same paragraph as above comment, for sentence “. . . but recently most storage abstraction innovation has occurred elsewhere, with cloud service providers offering column stores, document stores, key-value stores,

streaming data infrastructure, and object stores just to name a few options.” Suggest rewording slightly to “... but recently most storage abstraction innovation has occurred elsewhere, with cloud service providers offering numerous options including column stores, document stores, key-value stores, streaming data infrastructure, and object stores.” Or maybe it would be clearer with a comma before “just to name a few options.”