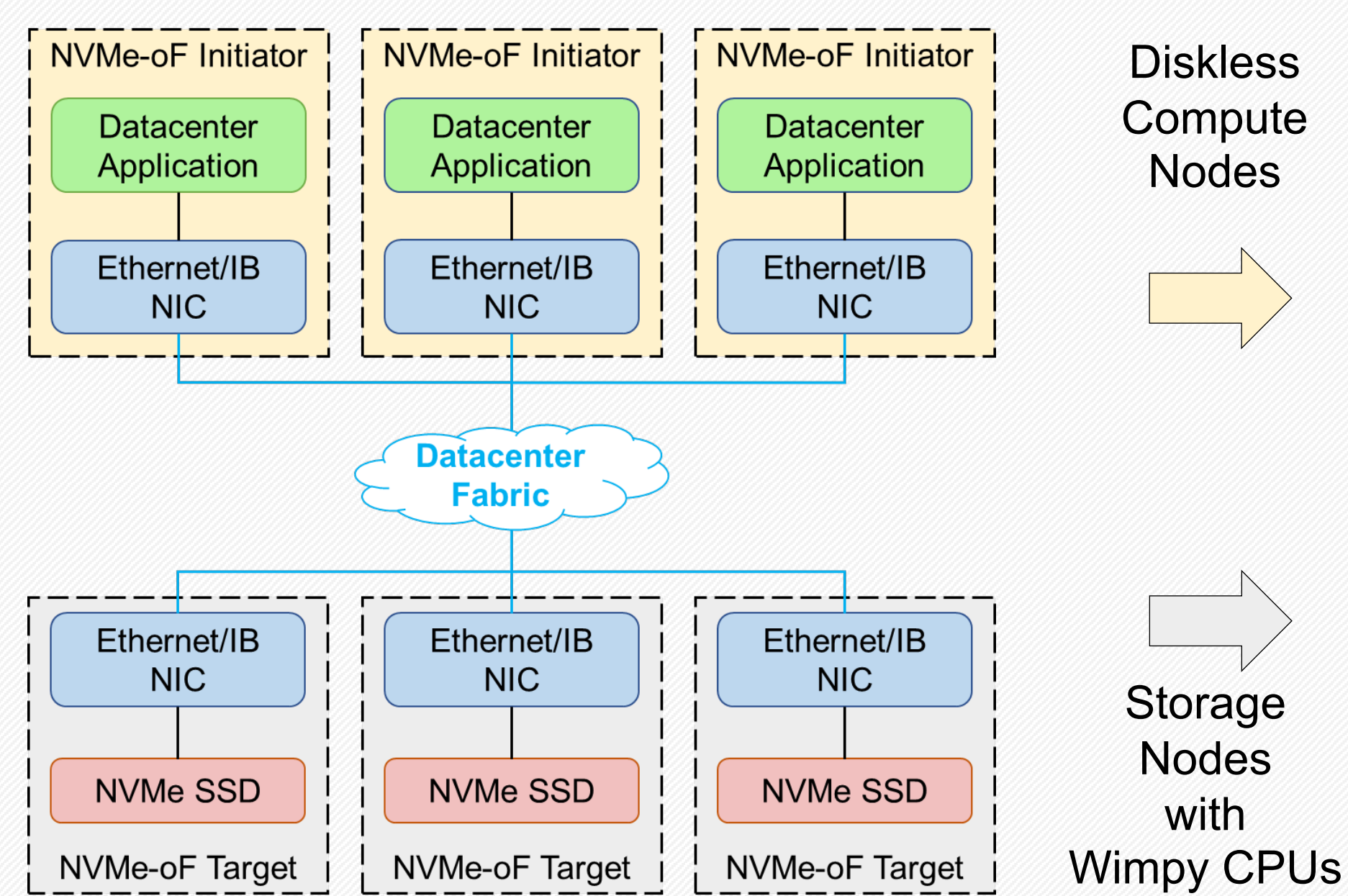


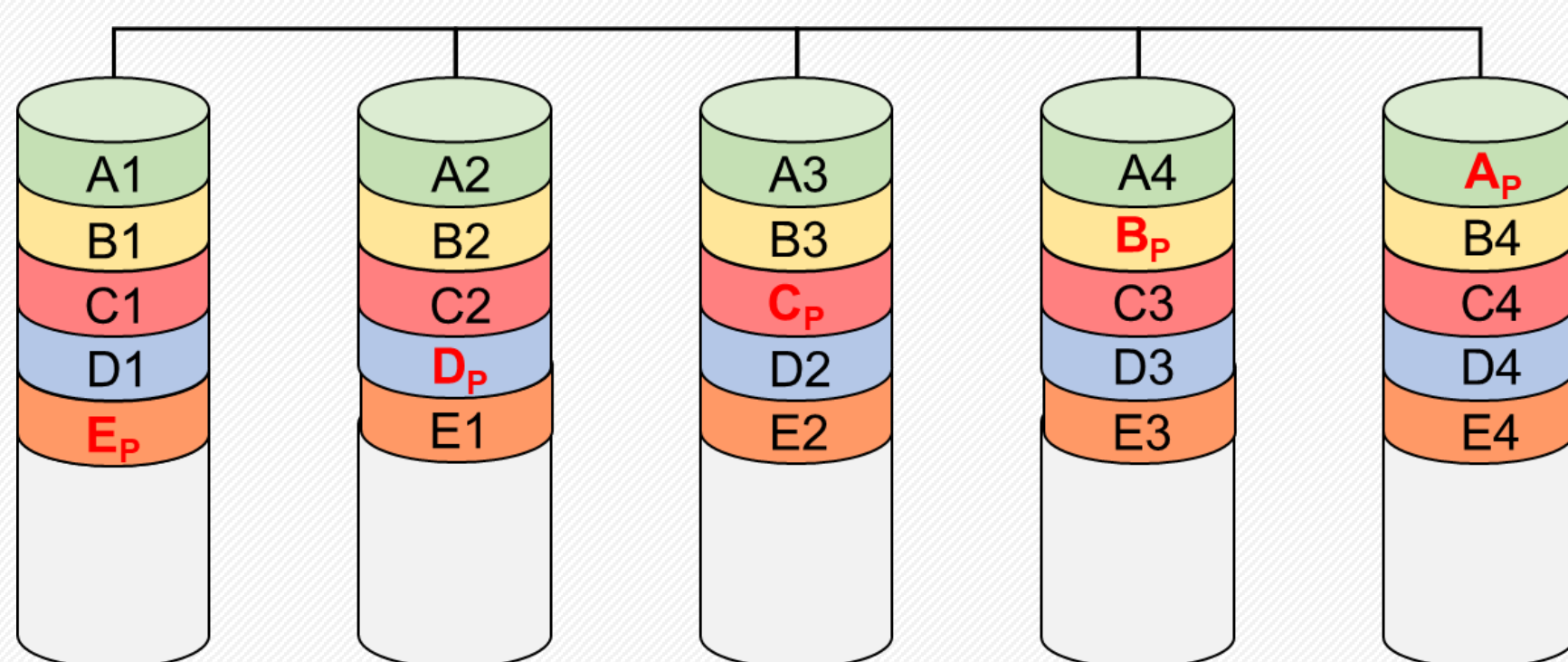


Background

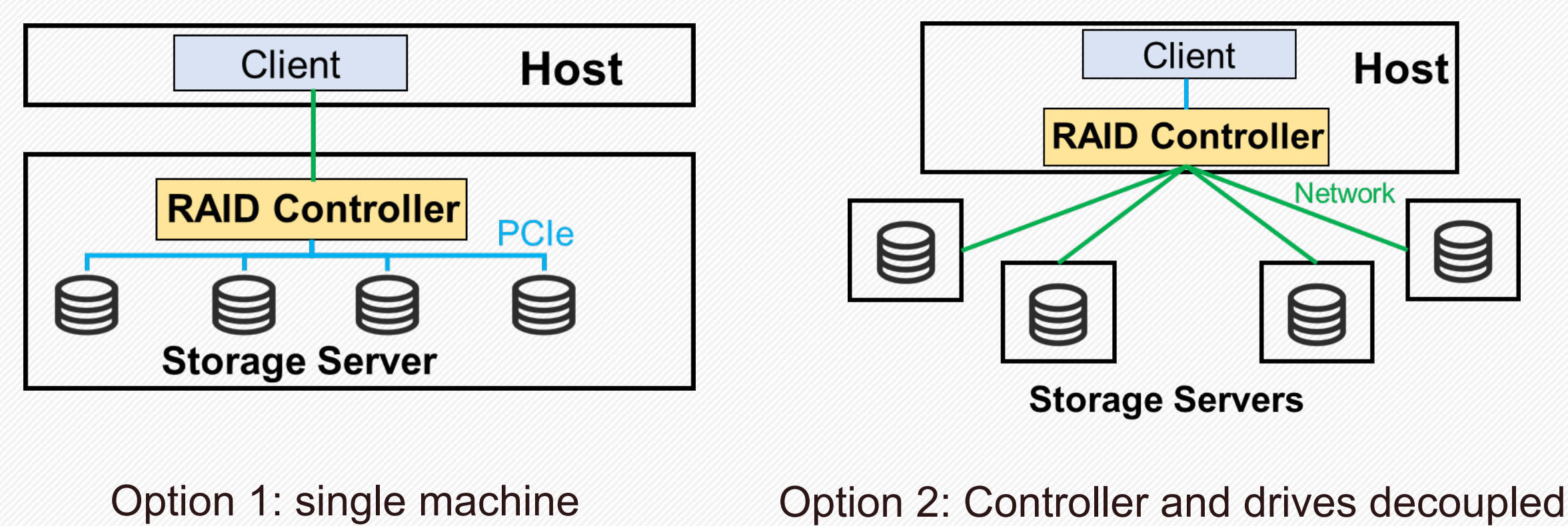
- Disaggregated storage becomes popular in datacenters



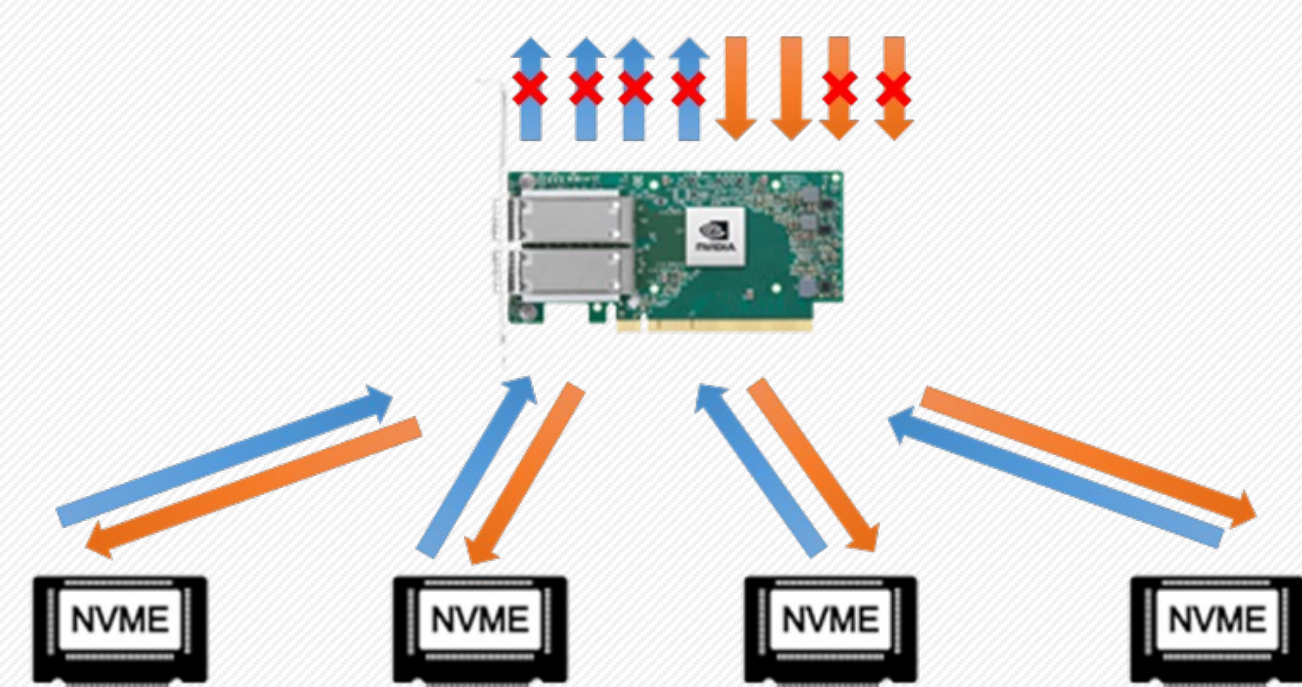
- While RAID is useful for higher BW/capacity/reliability



- Users can benefit from combining them



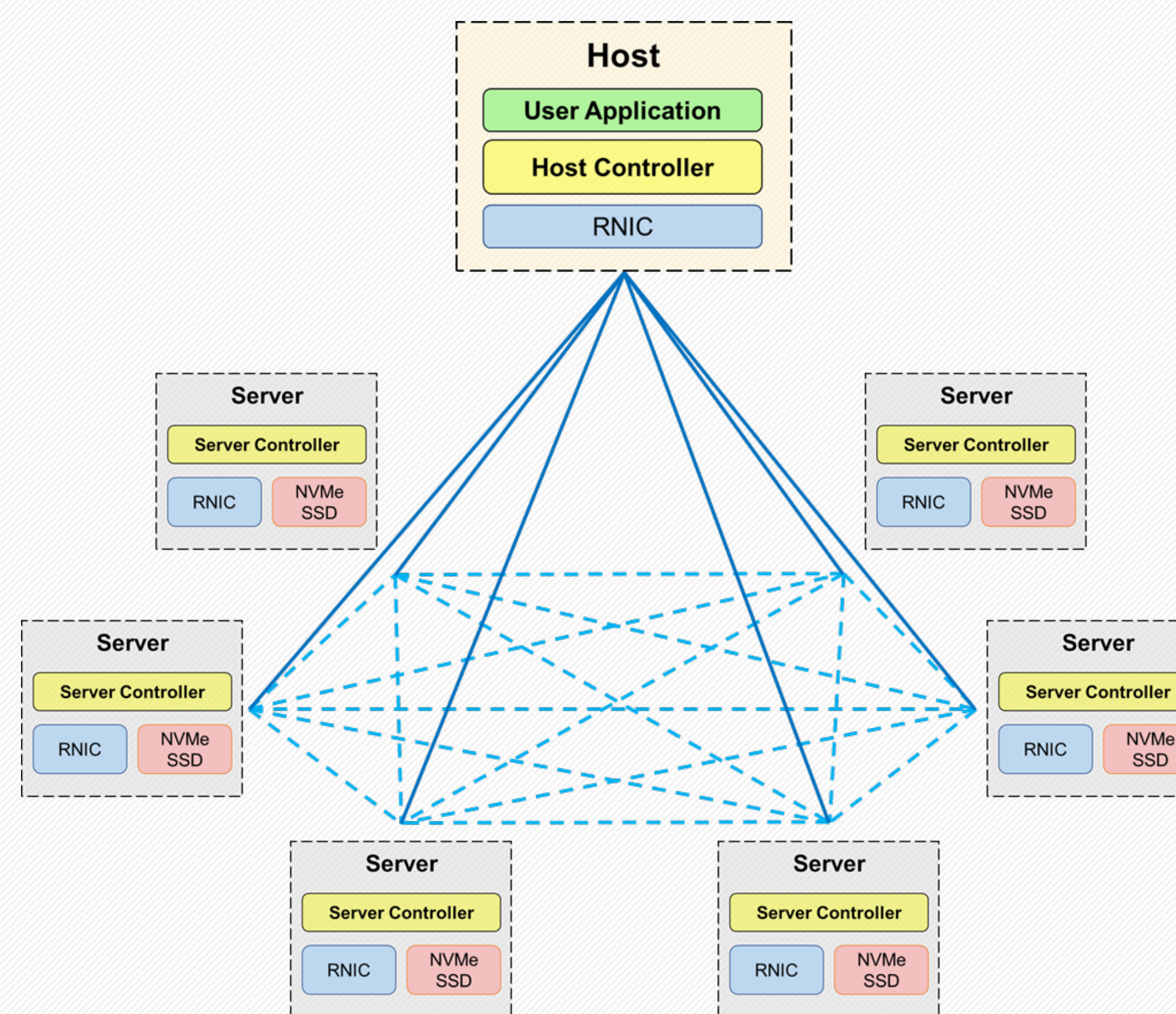
- I/O amplification on partial stripe write and reconstruction



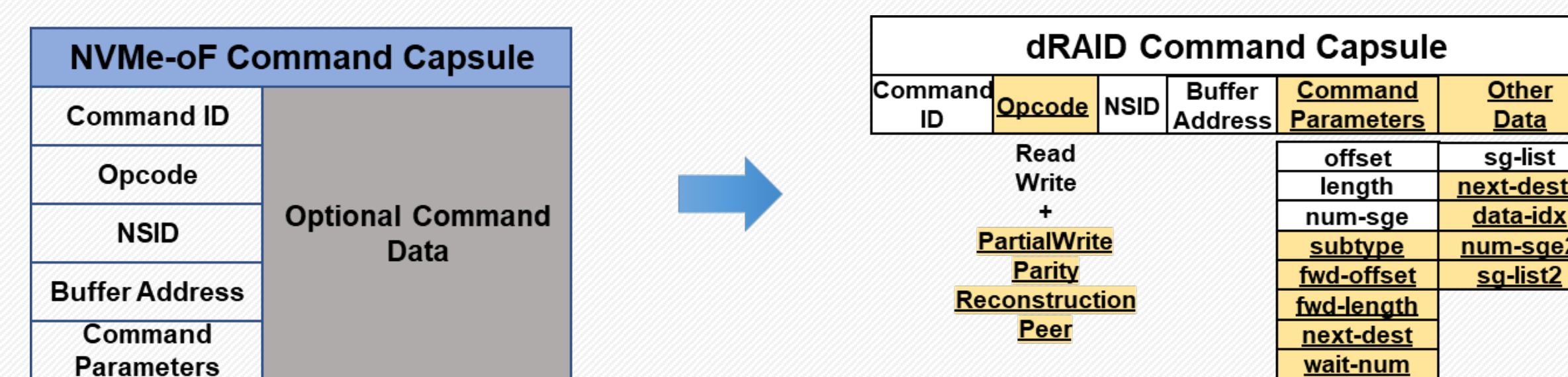
- 100G NIC can saturate read/write of 4 enterprise-grade SSDs
- A partial stripe write triggers 2 reads and 2 writes

System Architecture & Design

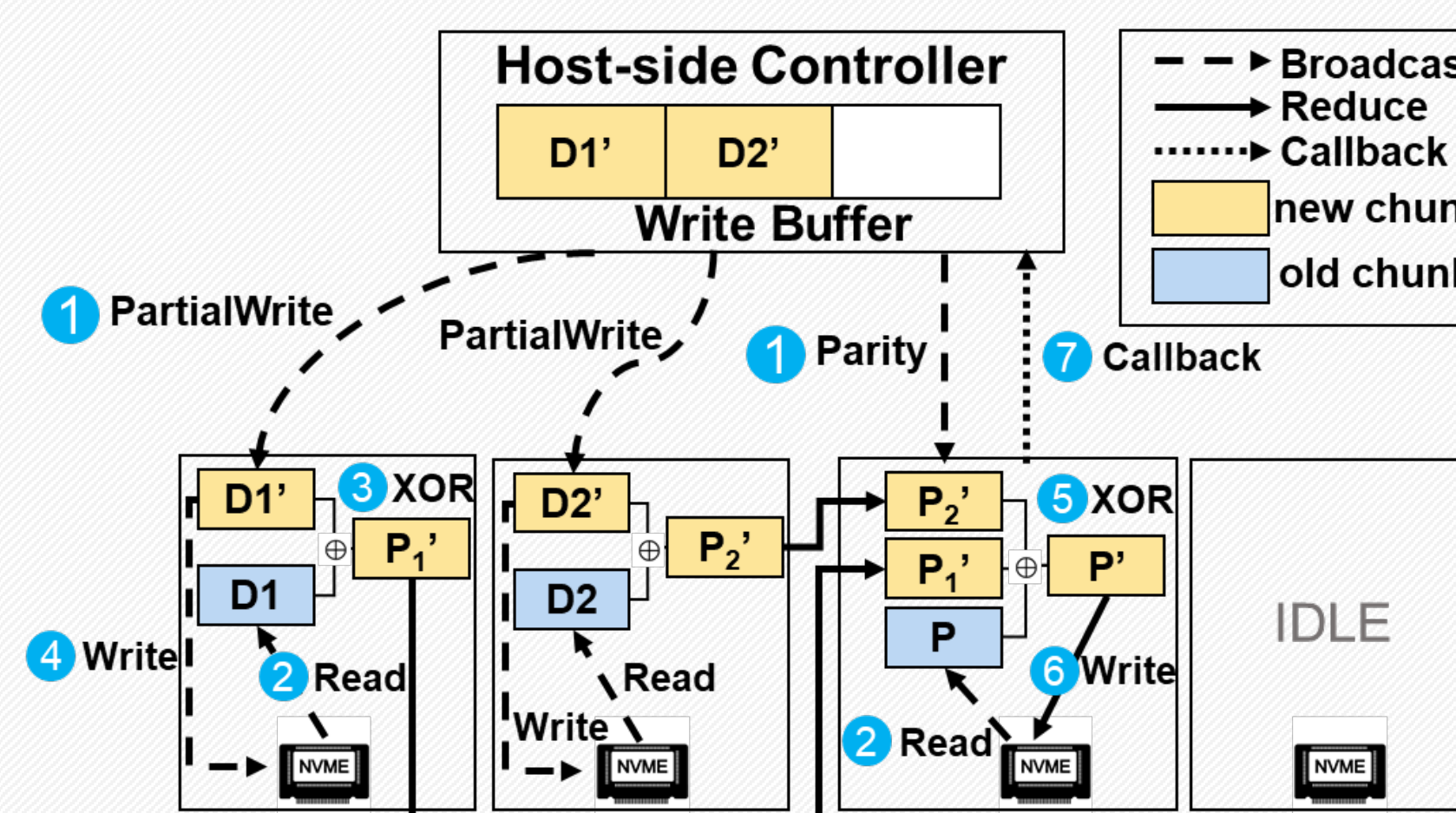
- Our key idea: **decompose** the controller to host and servers



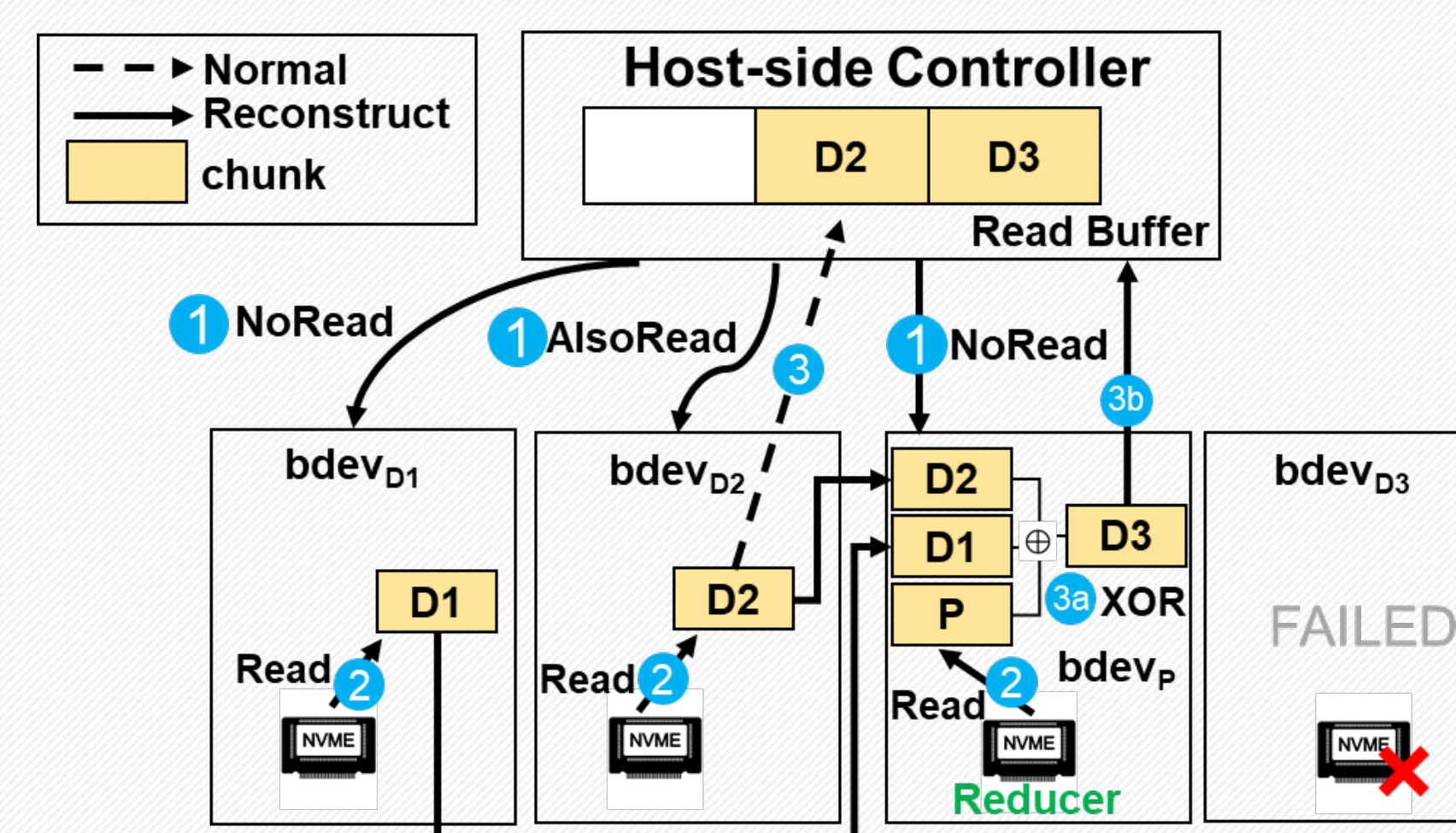
- dRAID protocol is an **extension** of the NVMe-oF protocol



- Partial stripe write workflow



- Data reconstruction workflow



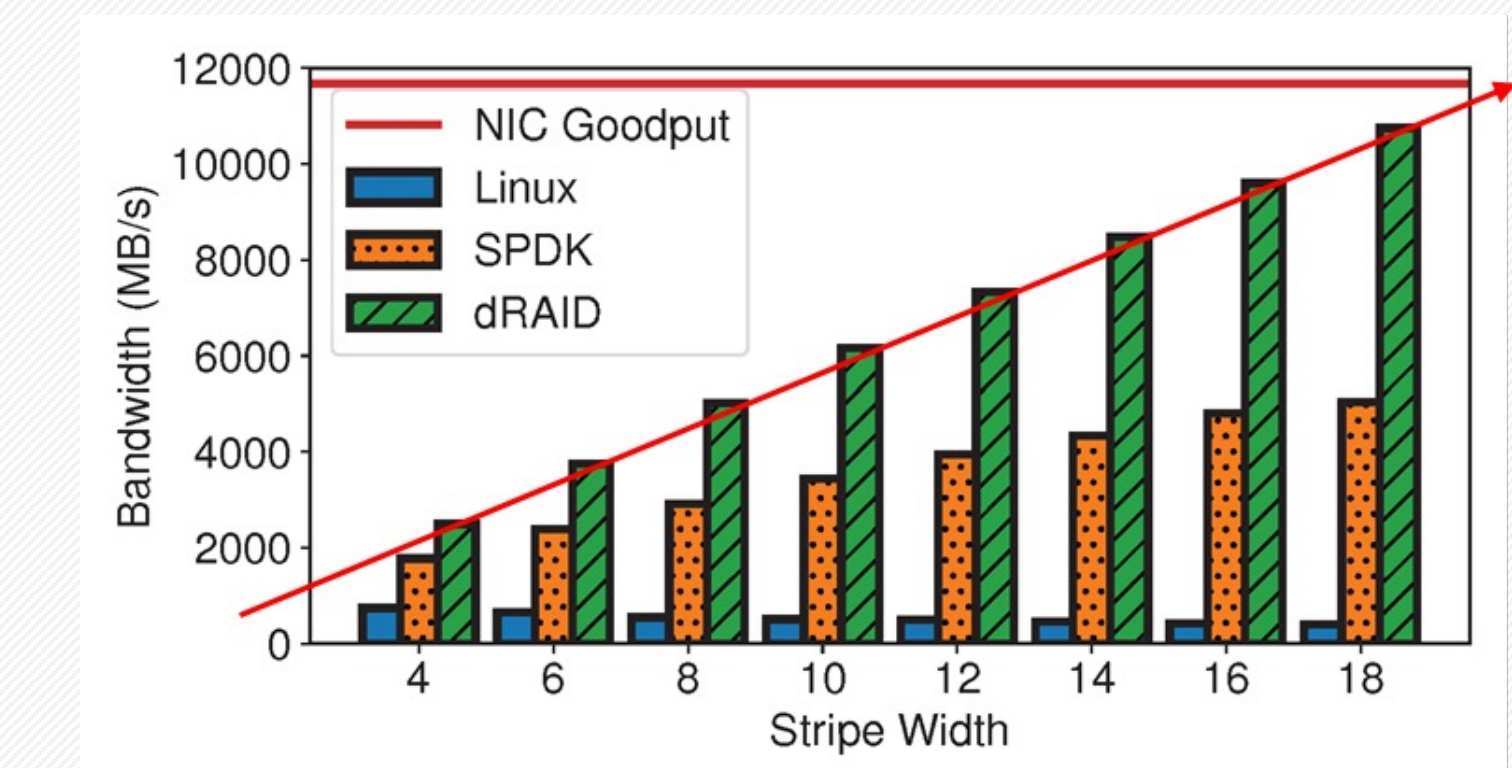
Implementation & Evaluation

- Implementation

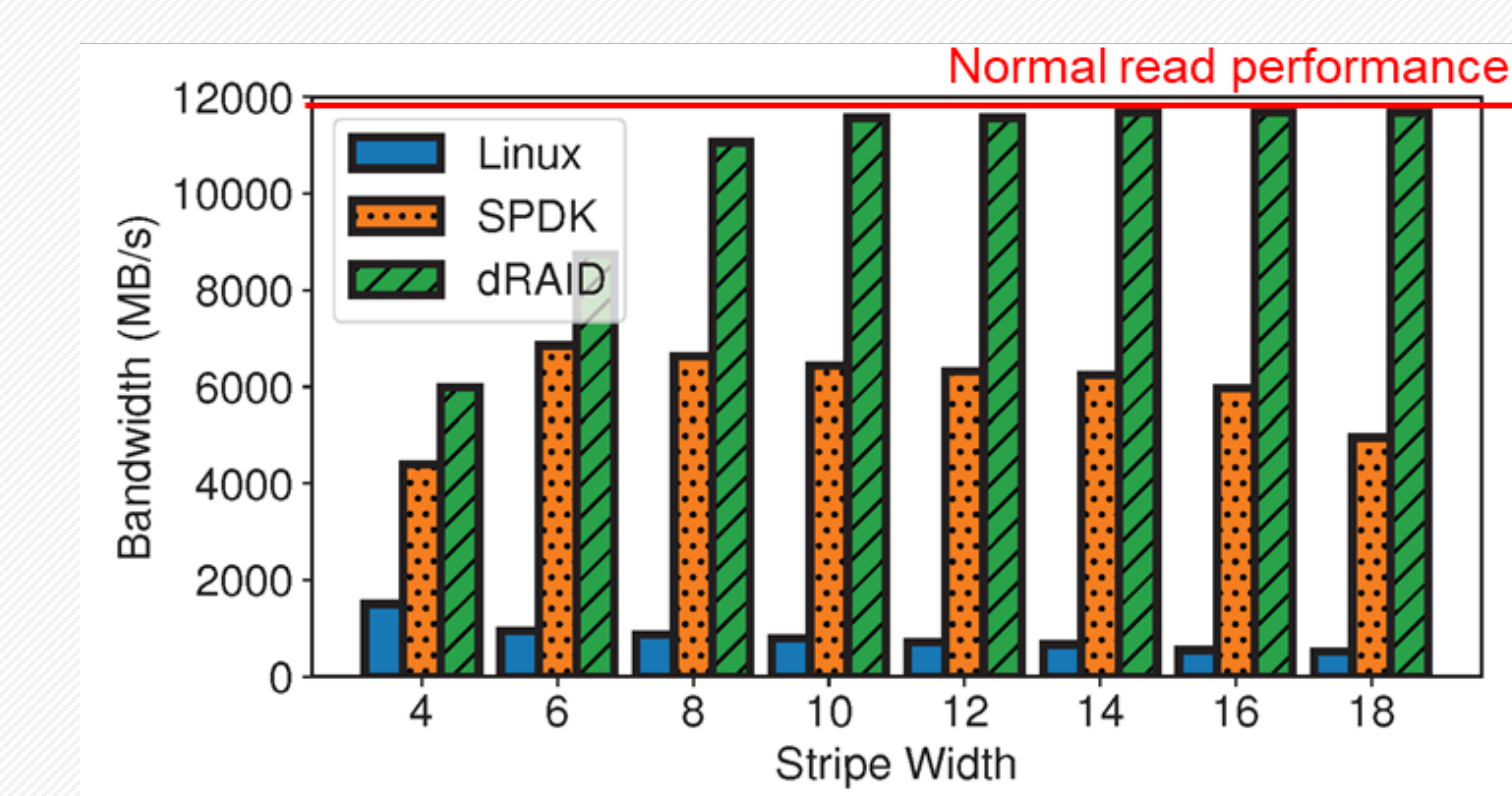
- ~9,800 lines of C++ code based on SPDK 21.10
- Open-sourced at <https://github.com/pkusys/dRAID>



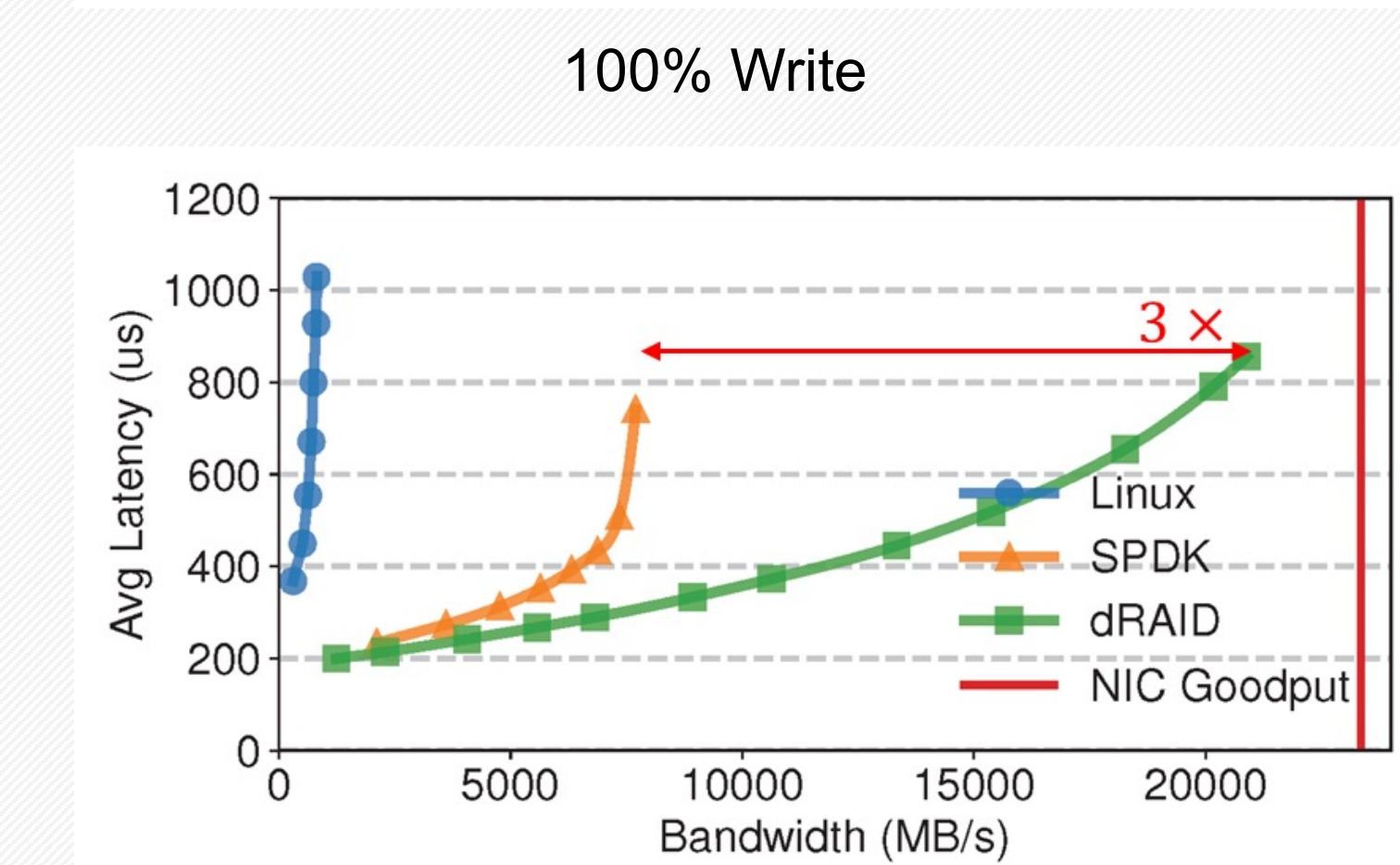
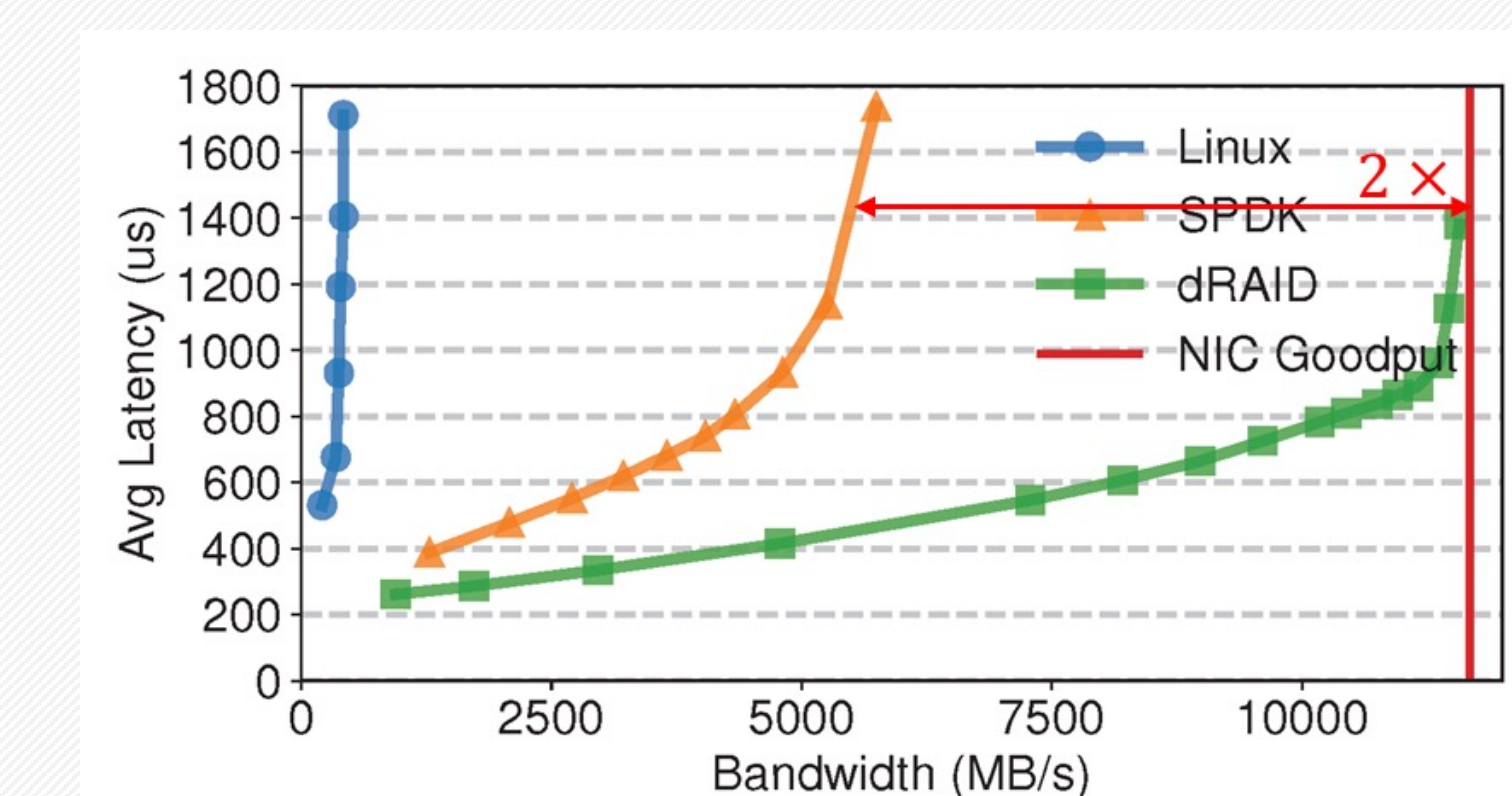
- Node scalability on write-only workload



- Node scalability on read-only workload under degraded state



- I/O scalability



50% Read + 50% Write