RAID4S

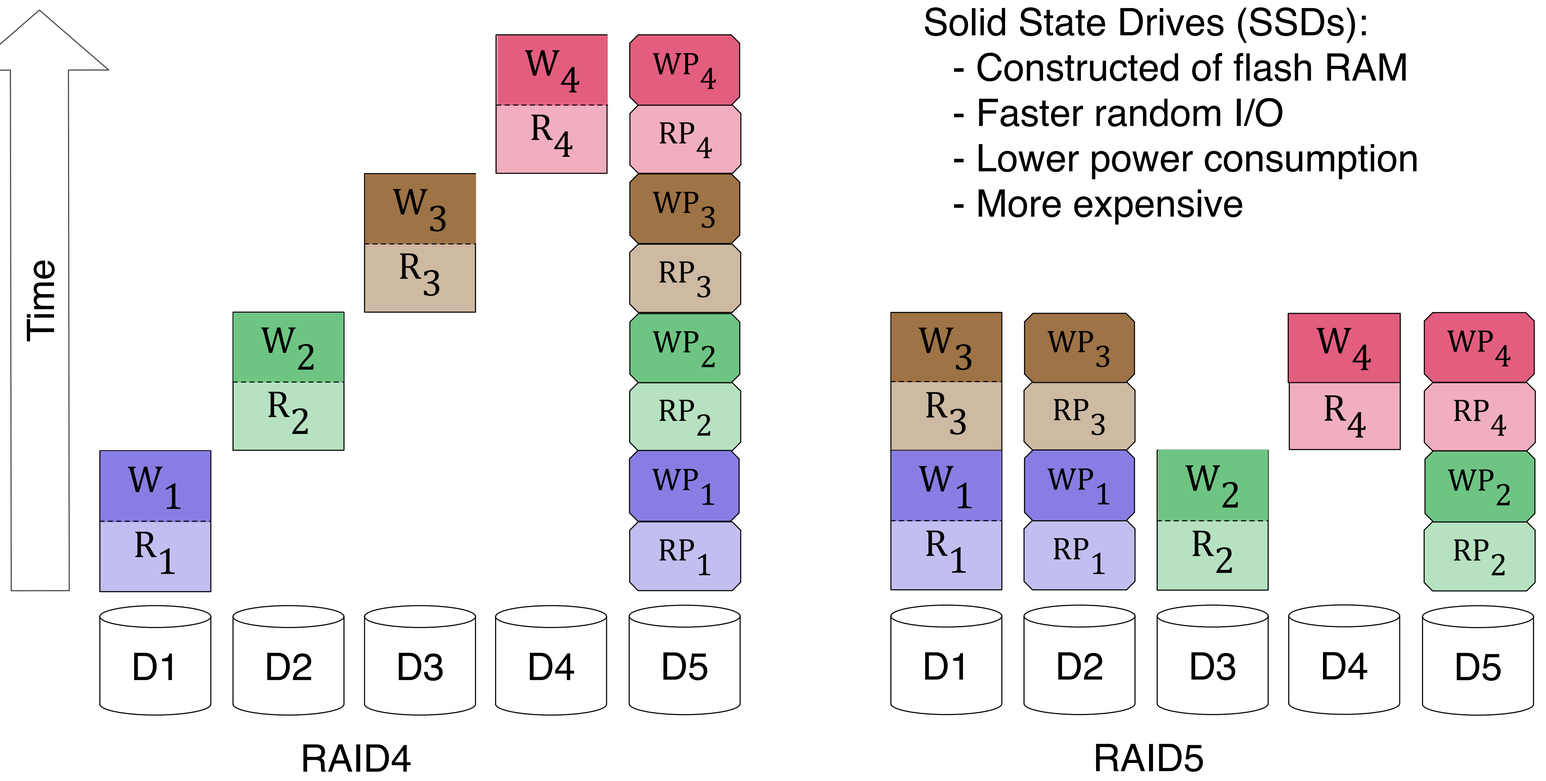
Adding SSDs to RAID Arrays

RAID4S stores parity data on solid state drives in a disk drive-based RAID4 layout. Initial results show up to 2x performance improvement.

Rosie Wacha[1], Scott Brandt[1], John Bent[2], and Carlos Maltzahn[1]

[1]University of California, Santa Cruz, [2]Los Alamos National Laboratory

Solving the Small Write Problem with RAID4S



Solid State Drives (SSDs):
- Constructed of flash RAM
- Faster random I/O
- Lower power consumption
- More expensive

Small writes limit RAID performance
- 4 I/Os per data block written

RAID4: parity drive is bottleneck
RAID5: data and parity accessed in parallel
RAID4S: parity offloaded to SSD and all data accesses are parallelized

Figure 1. The time to complete four small writes is compared. RAID4S with four disk drives allows the complete parallelization of M=4 accesses to different disks.

Calculated Throughput of RAID4S vs. Disk-Based Schemes

M outstanding I/Os at the controller
N = 4 disk drives for RAID4S + 1 SSD
N+1 = 5 disk drives for RAID 4 and RAID5

Assumptions:
Each small write incurs a seek and transfer time. Small writes are to different stripes. Thus, the I/O is parallelized as much as possible.

Throughput calculation:
- 64KB small writes
- Each I/O incurs a seek time and a transfer time
- Outstanding I/Os are queued at the controller and completed in parallel, if possible.

Hardware:
- Western Digital WD20EADS (low power disk drive)
- Intel X25-E (SSD)

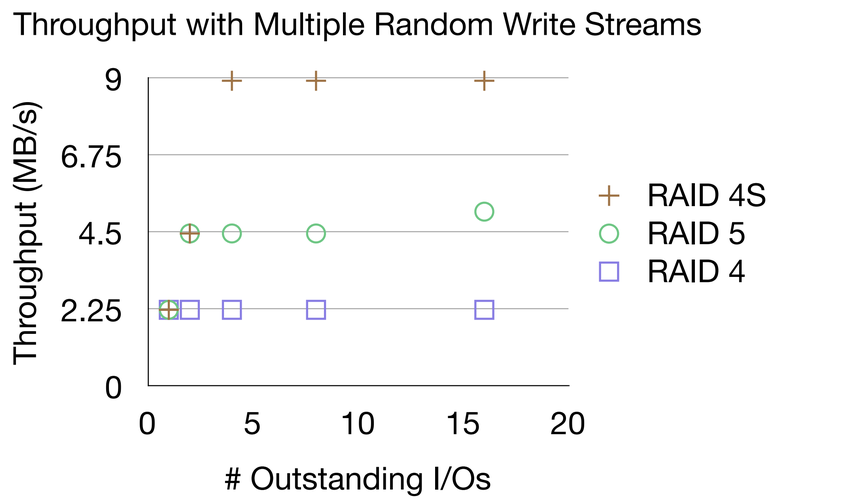


Figure 2. Throughput of small writes for a fully random workload (all small writes). RAID4S completes N small writes in the time that RAID4 completes one small write, as long as the SSD is N times faster than hard disks in the array.

RAID4S is up to 2X faster than RAID5 and 4X faster than RAID4

Conclusions

Flash SSDs provide faster random I/O performance and new opportunities for data storage in distributed systems. This work replaces the parity disk drive in a RAID4 system with SSD. Initial results show improvements of up to (N/2)X speedup over RAID5 and NX speedup over RAID4.

Future Work

Cost analysis with power usage
More general workloads
- Mix small and large writes
- Workload traces

Build RAID4S with hard drives and flash and measure:
- Throughput
- Latency
- Power

Investigate reliability impact of RAID4S

Acknowledgements

This work was funded by the Institute for Scalable Scientific Data Management (ISSDM) and completed at UC Santa Cruz.

SRL UCSC Systems Research Lab
Baskin Engineering UC SANTA CRUZ
ISSDM