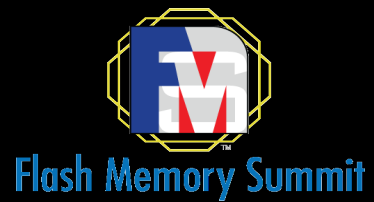


Enabling an Optimal Application Ecosystem on ZNS Storage

Session: FMAR-101-2: ZNS

Sanhita Sarkar, Ph.D.

Global Director, Software Development, Western Digital



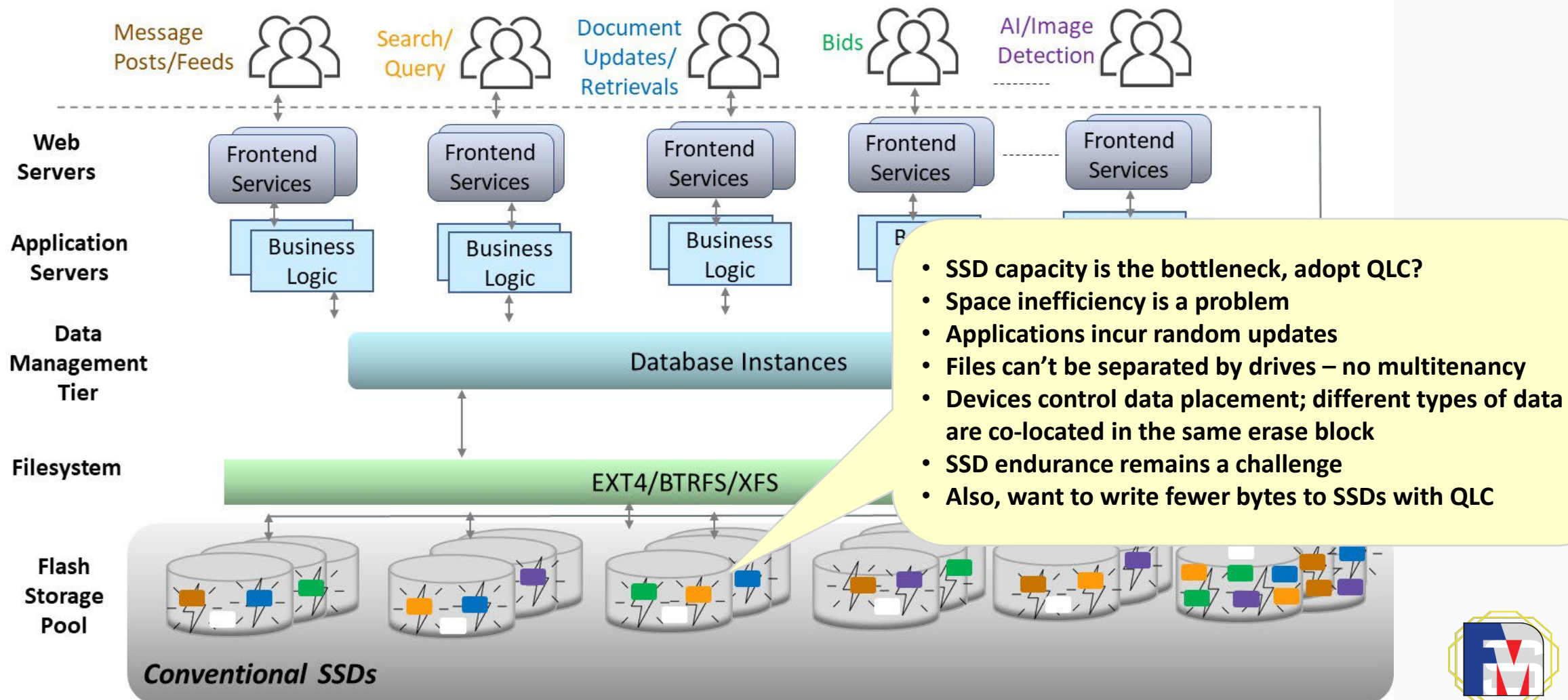
Topics

- **Application-level challenges to leverage benefits of underlying ZNS storage**
- **Deploying an optimal application ecosystem on ZNS made easy**
 - **Introducing Application-centric ZNS Reference Kits for accelerating end user deployments**
 - **Percona Server for MySQL® ZNS Reference Kit and its components**
- **Methodology for evaluating performance of the Percona MySQL solution stacks on ZNS and conventional SSDs**
 - **Performance results derived in Western Digital lab**
- **ZNS SSDs for MySQL – Value Proposition**
- **Ongoing work**



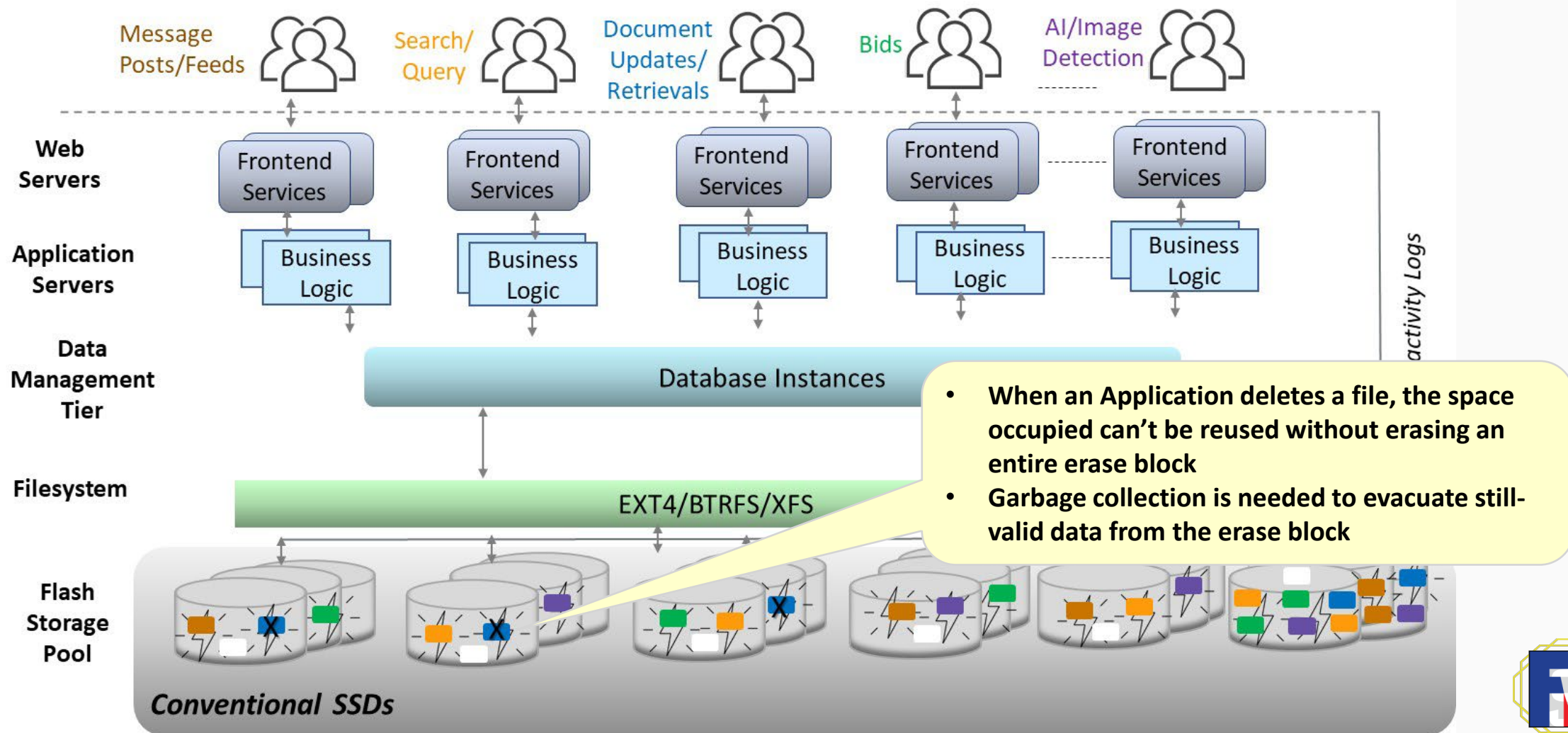
Multitenant web-scale Applications on Conventional SSDs

Challenges



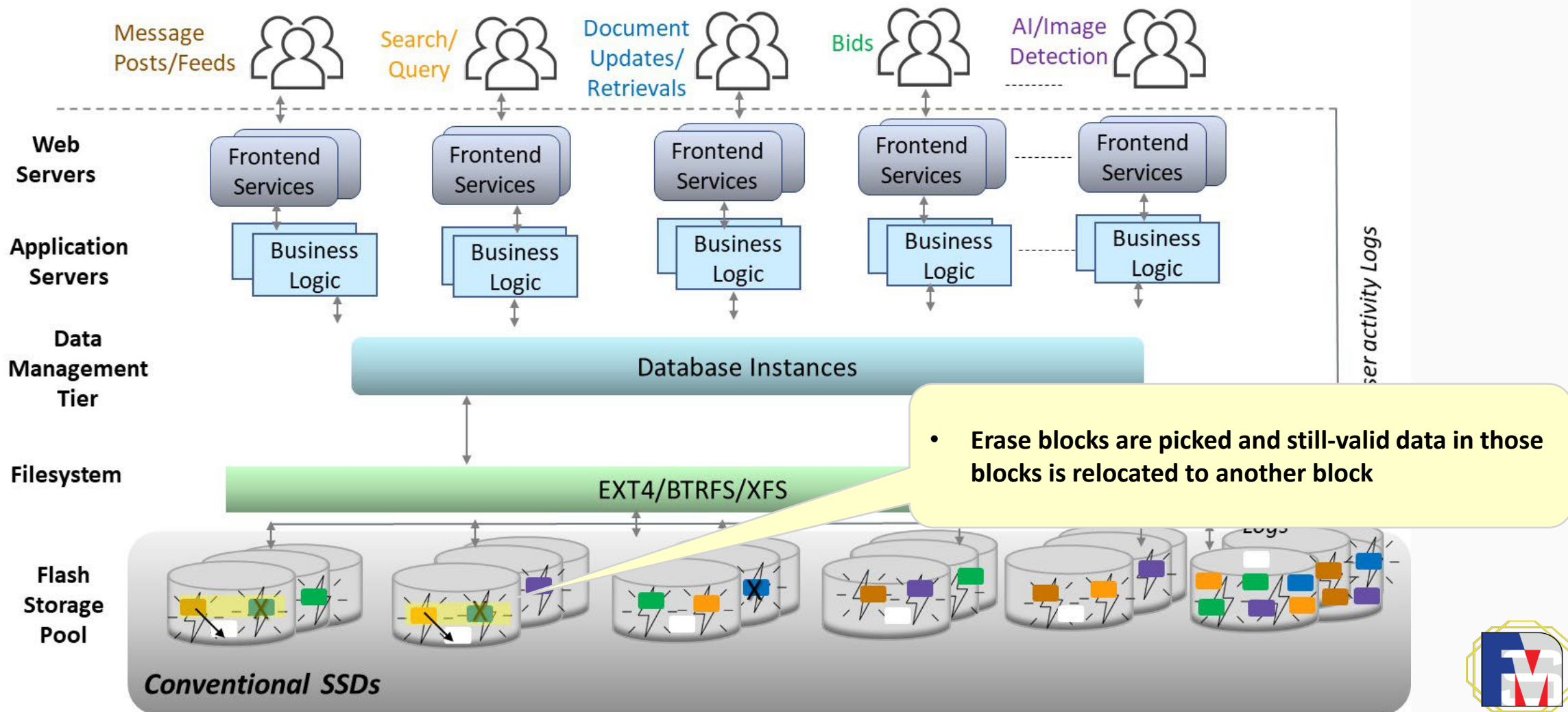
Multitenant web-scale Applications on Conventional SSDs

Challenges



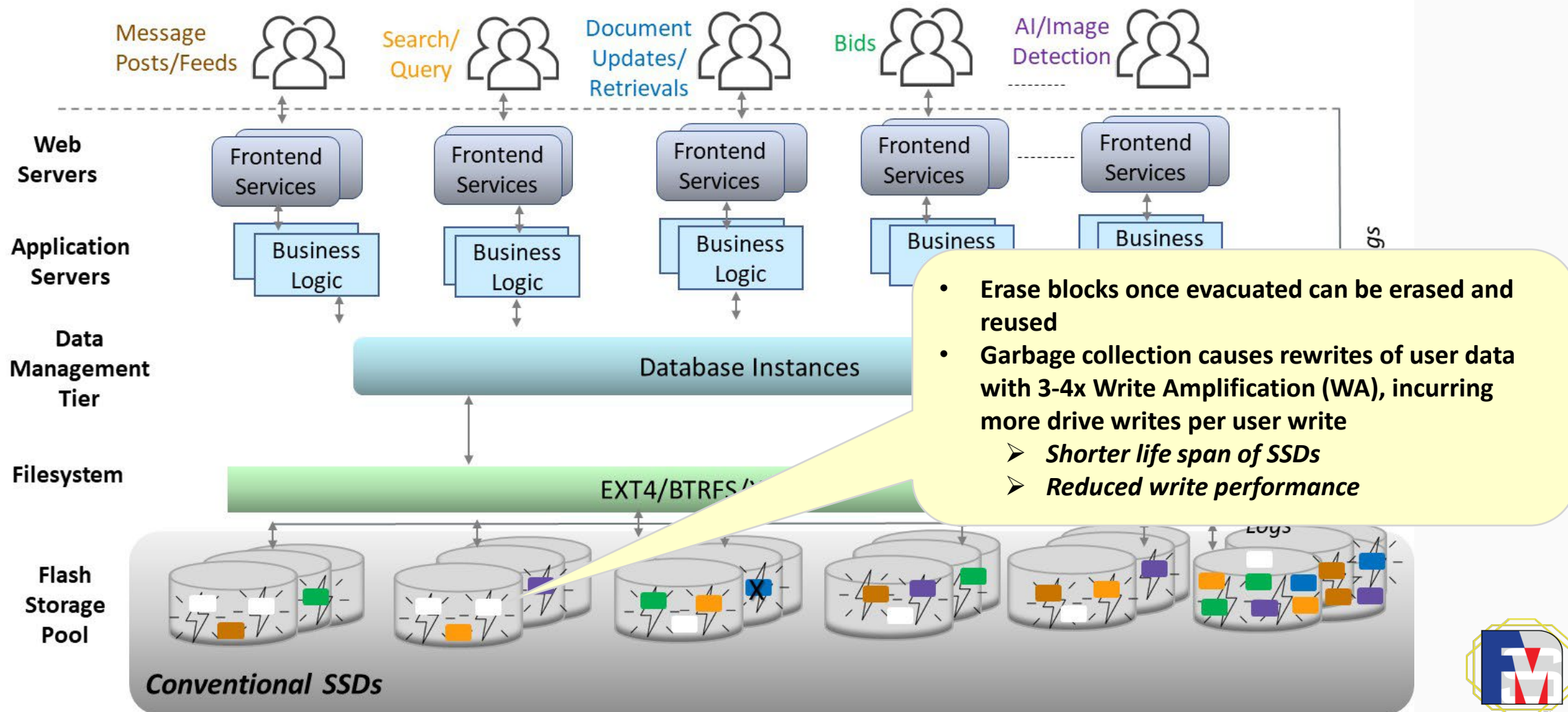
Multitenant web-scale Applications on Conventional SSDs

Challenges

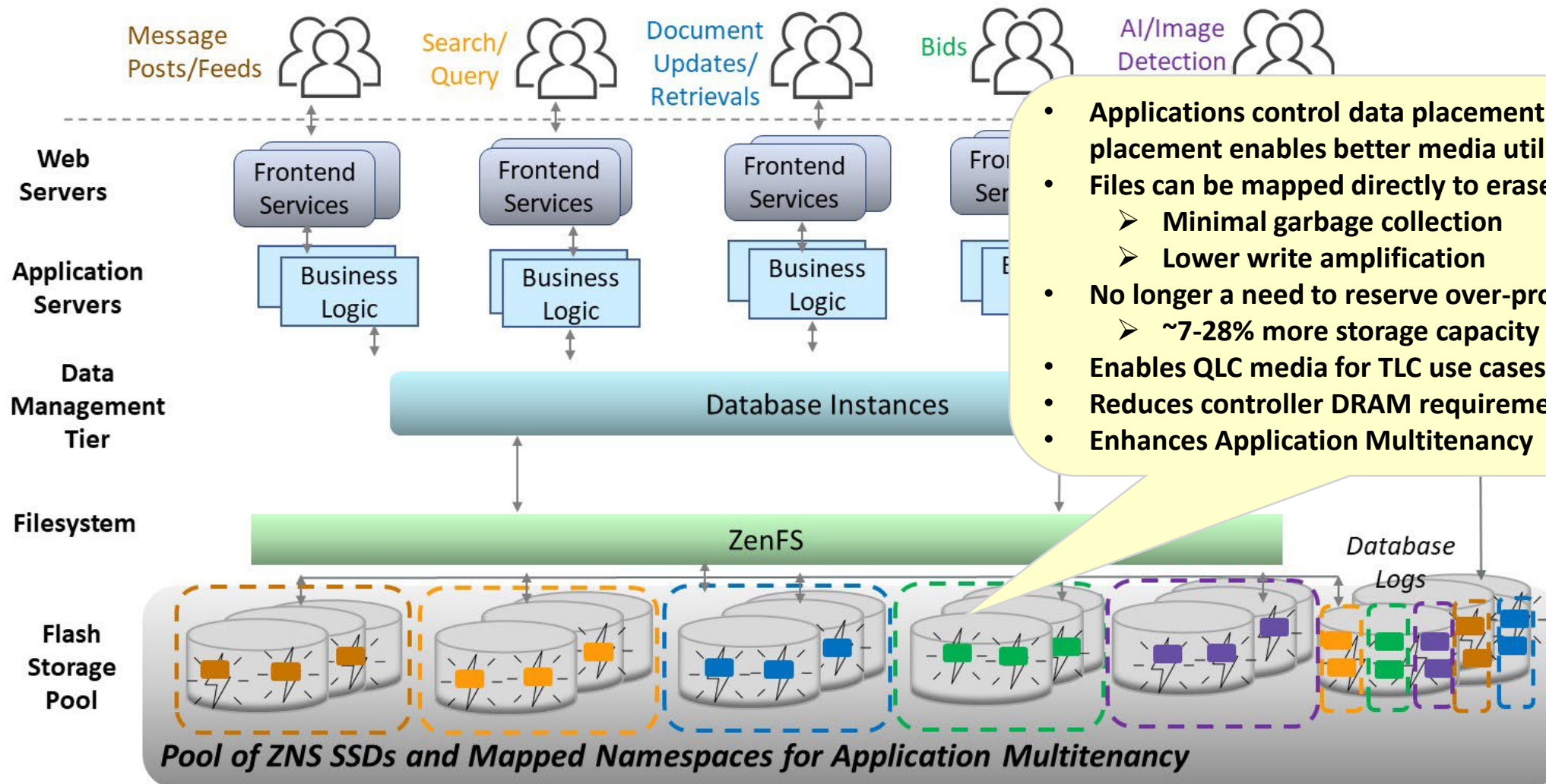


Multitenant web-scale Applications on Conventional SSDs

Challenges

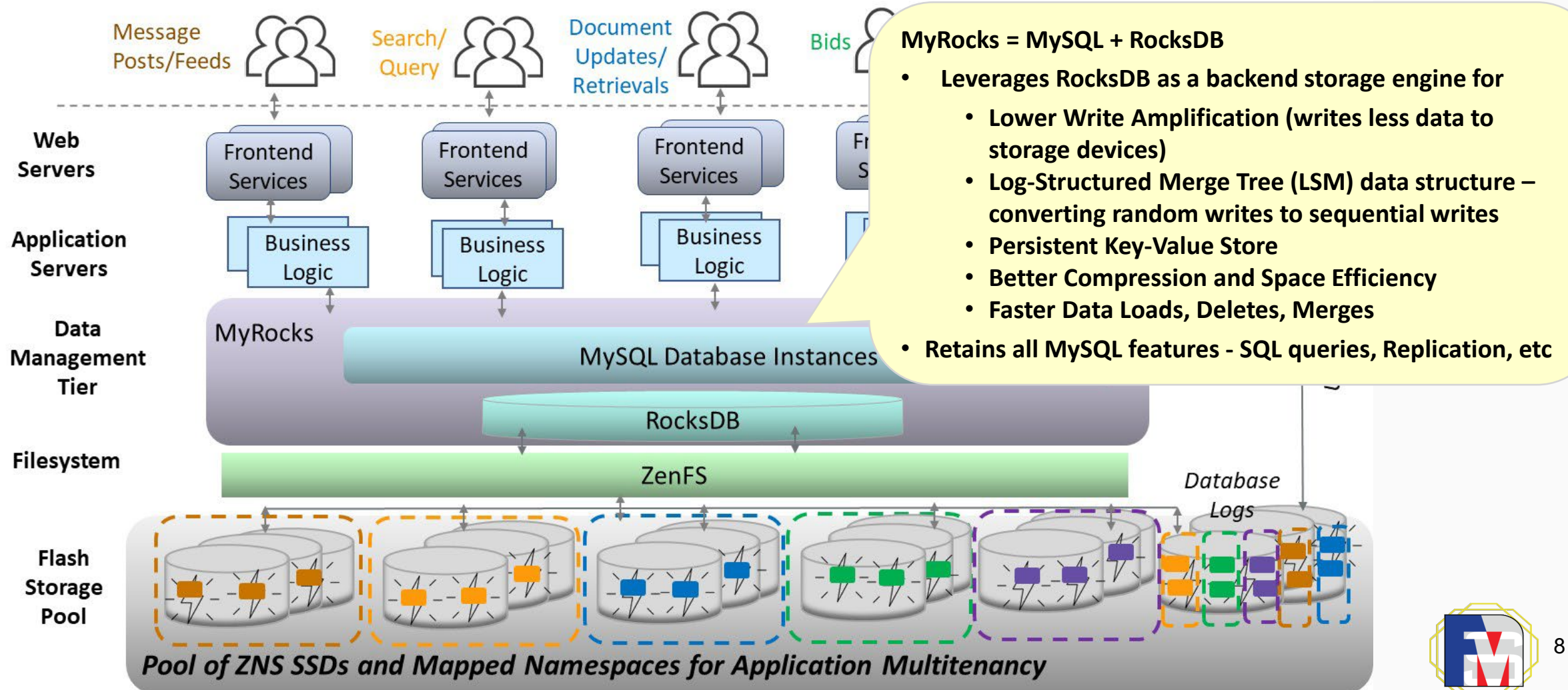


Enhanced Application Multitenancy on ZNS SSDs

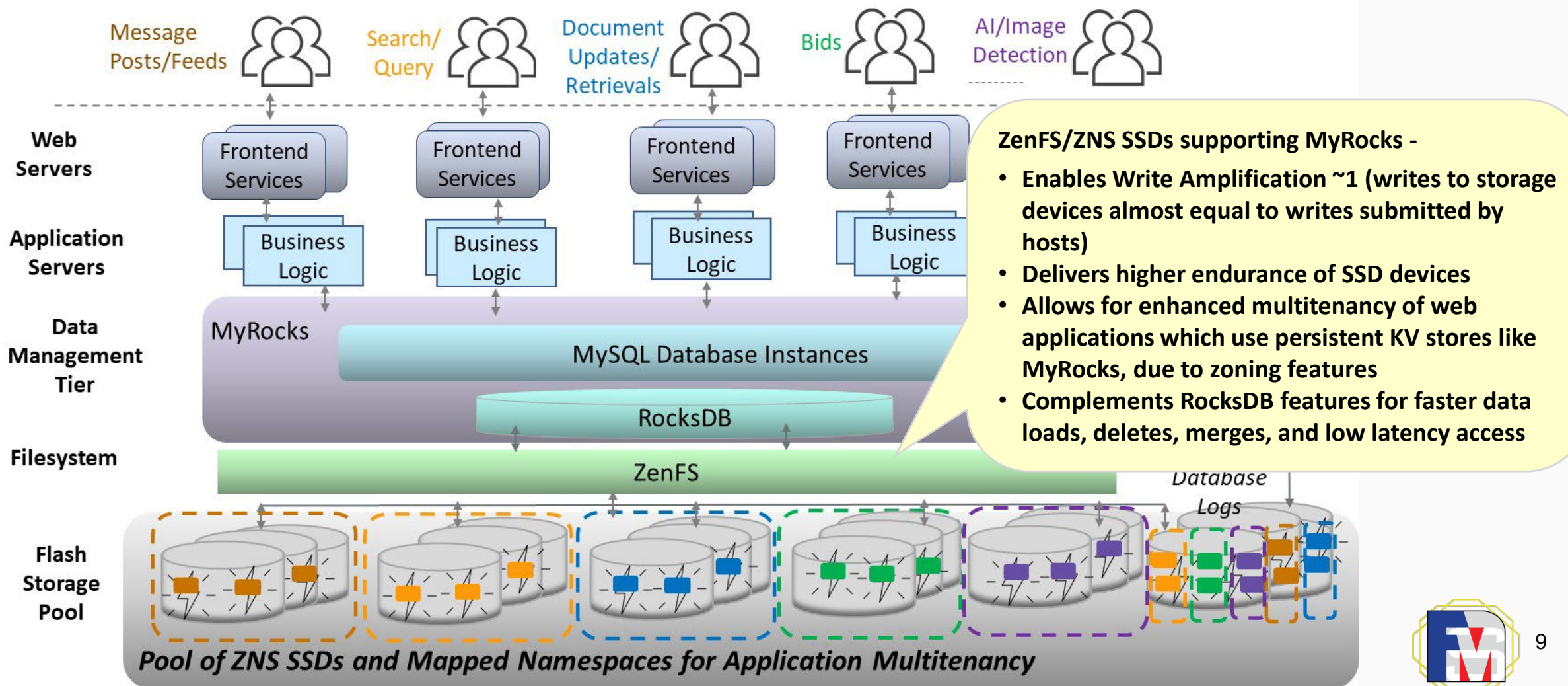


- Applications control data placement....Smart placement enables better media utilization, lower WA
- Files can be mapped directly to erase blocks (zones)
 - Minimal garbage collection
 - Lower write amplification
- No longer a need to reserve over-provisioned media
 - ~7-28% more storage capacity
- Enables QLC media for TLC use cases
- Reduces controller DRAM requirements
- Enhances Application Multitenancy

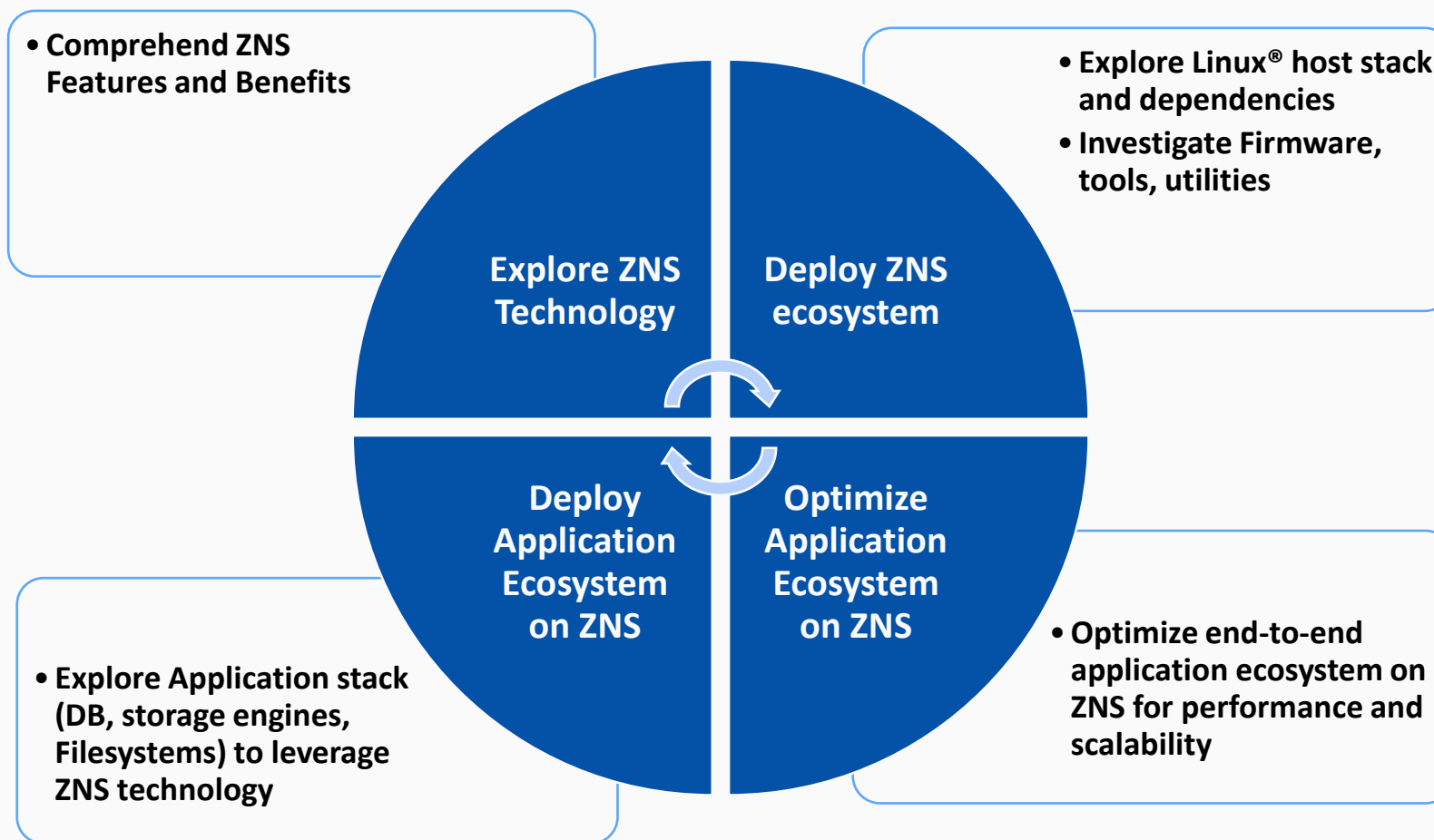
Open Storage Plugin MyRocks: MySQL + RocksDB



MyRocks leveraging ZNS SSDs



Deploying an Optimal Application Ecosystem on ZNS - Challenges



Topics

- Application-level challenges to leverage benefits of underlying ZNS storage
- **Deploying an optimal application ecosystem on ZNS made easy**
 - Introducing Application-centric ZNS Reference Kits for accelerating end user deployments
 - Percona Server for MySQL® ZNS Reference Kit and its components
- Methodology for evaluating performance of the Percona MySQL solution stacks on ZNS and conventional SSDs
 - Performance results derived in Western Digital lab
- ZNS SSDs for MySQL – Value Proposition
- Ongoing work



Deploying an Optimal Application Ecosystem on ZNS – Made Easy

Supported Linux Distro:

Kernel, utilities, and
filesystem(s) for ZNS

Application Reference Kit on ZNS

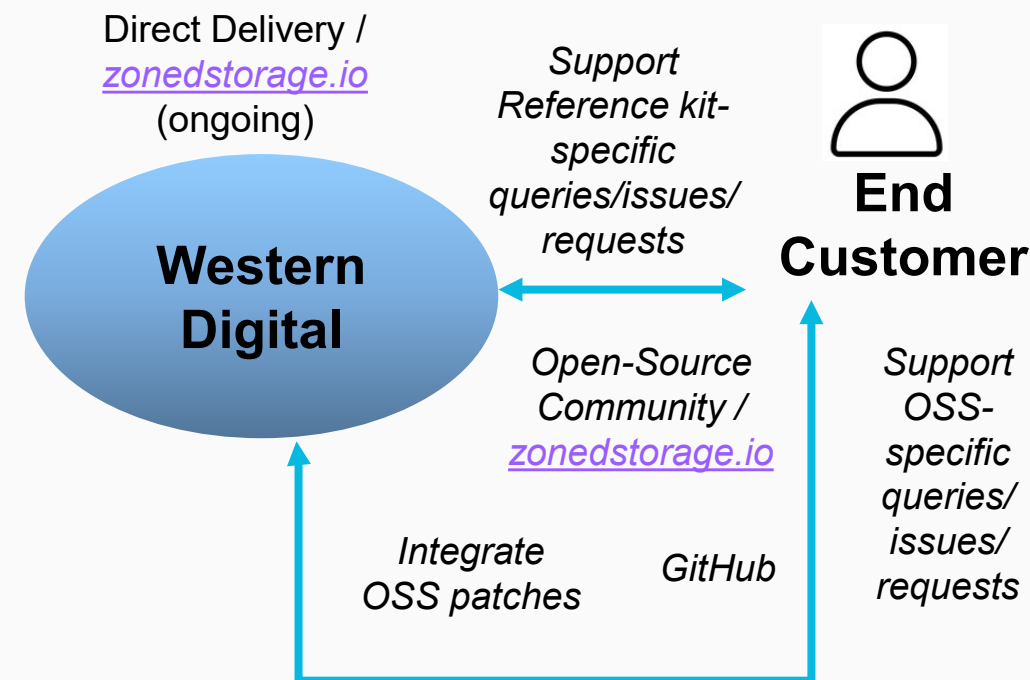


Benchmark tools, and Applications

Script to download, build and
validate the Linux host stack
along with output logs

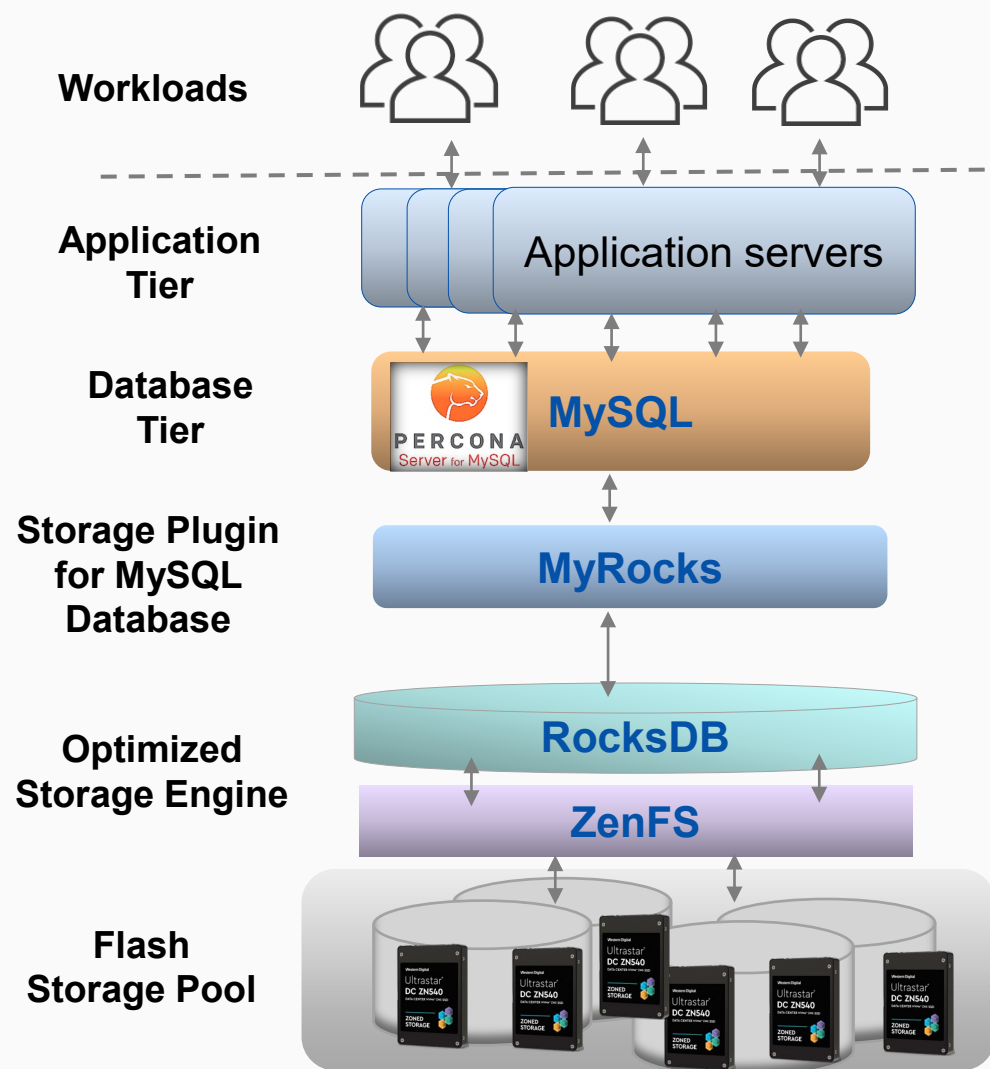
Sample scripts to configure
and benchmark an application

README with instructions for
executing the tarball, for
troubleshooting, along with a
methodology to build, configure
and benchmark an application



ZNS Storage integrated with Percona MySQL Database

Western Digital Lab Enablement



Solution



Percona Server
for MySQL

+

MyRocks

+

ZenFS

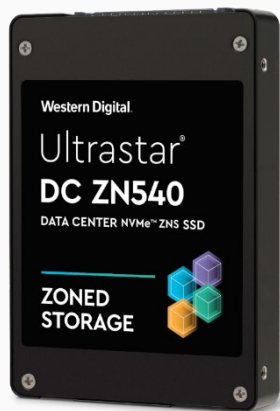
+

ZN540

Percona MySQL Reference kit from Western Digital enables ease of deployment of the solution stack for Western Digital customers

Percona MySQL Reference Kit from Western Digital

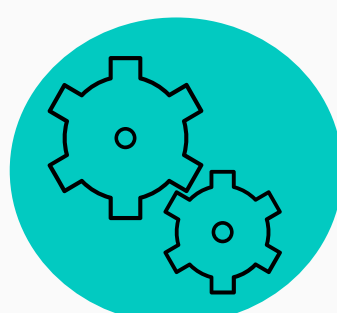
Enabling Ease of Deployment of Percona MySQL/MyRocks/ZenFS stack on Ultrastar® DC ZN540 SSDs*



ZNS SSD



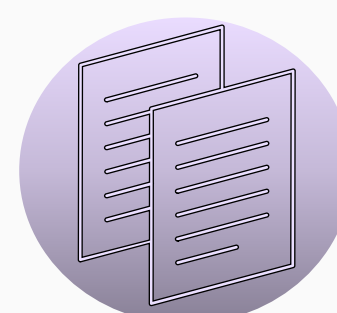
Software



Scripts



Optimization



Documentation

Western Digital Ultrastar DC ZN540* SSD:

Latest ZNS specs supported:

- TP-4053 (Zoned namespaces)
- TP-4056 (Namespace Types)
- TP-4076 (Zoned Random Write Area)

Standard Linux Distribution: Includes kernel, nvme-cli, blkzone for ZNS

Benchmark tools, Applications: FIO, Sysbench, Percona Sever for MySQL 8.0, Percona MyRocks

Tarball includes “README” instructions, script to install the Linux host stack, followed by building, configuring and benchmarking a Percona Server for MySQL 8.0 database with a transactional workload, and compare with a preconditioned conventional drive

* The Ultrastar DC ZN540 SSD has limited availability to qualified customers.



About Percona MySQL ZNS Reference Kit V3.0

What's included in the kit

1. A README document with instructions to build the Linux host software stack, followed by building and benchmarking the Percona Server for MySQL database 8.0 on ZN540* SSDs
2. A script (and a doc with manual instructions) to download, build and install the essential host stack components
3. Percona Server for MySQL database with ZenFS installation guide along with a script
4. Percona MyRocks installation guide, along with a script
5. Benchmark setup and execution and comparing performance with SN540* SSDs, and relevant scripts
6. Configuration files for executing the database on InnoDB storage engine with SN540 SSD, and on MyRocks storage engine with SN540 and ZN540 SSDs

Essential host stack & Percona Server for MySQL database components available for download using the included scripts

Linux kernel 5.10.72 LTS distro (Ubuntu 20.04.3 LTS distro)

linux-util (blkzone)

nvme-cli

FIO

Percona Server for MySQL 8.0.28-20

Percona MyRocks

Sysbench

Sample scripts and output logs included

1. fio_zns_write.sh	for FIO write
2. fio_zns_read.sh	For FIO read
3. sysbench_workload.sh	for preparing and cleaning up the database, and running sysbench
4. precondition-conv-fio.sh	for preconditioning conventional drive using fio
5. gather_debug_info	for collecting debug output



**The Ultrastar DC ZN540 and SN540 SSDs have limited availability to qualified customers*

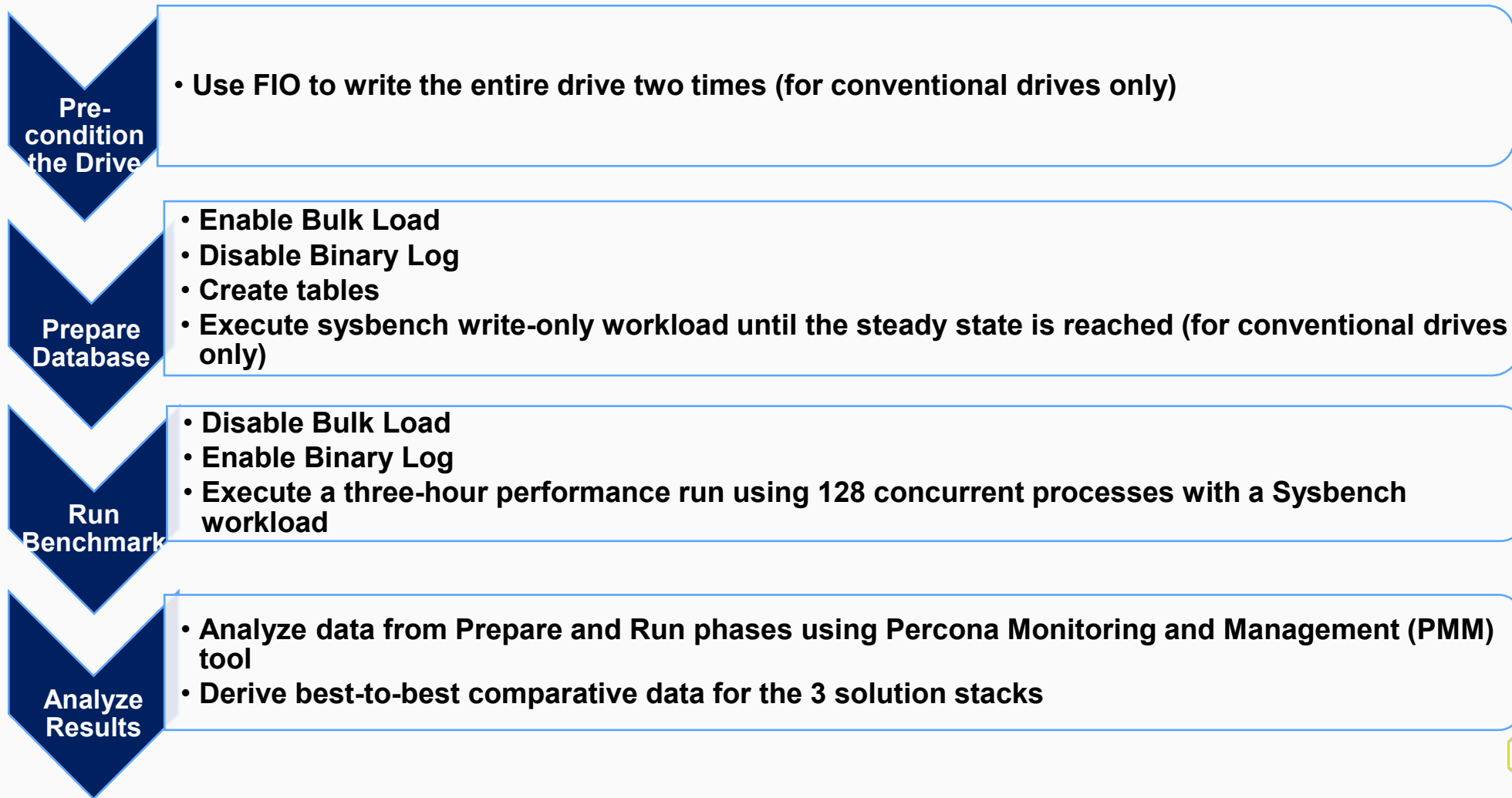
Topics

- Application-level challenges to leverage benefits of underlying ZNS storage
- Deploying an optimal application ecosystem on ZNS made easy
 - Introducing Application-centric ZNS Reference Kits for accelerating end user deployments
 - Percona Server for MySQL® ZNS Reference Kit and its components
- **Methodology for evaluating performance of the Percona MySQL solution stacks on ZNS and conventional SSDs**
 - **Performance results derived in Western Digital lab**
- ZNS SSDs for MySQL – Value Proposition
- Ongoing work



Performance Evaluation of Percona Server for MySQL

Performance Evaluation Methodology on ZNS and Conventional SSDs

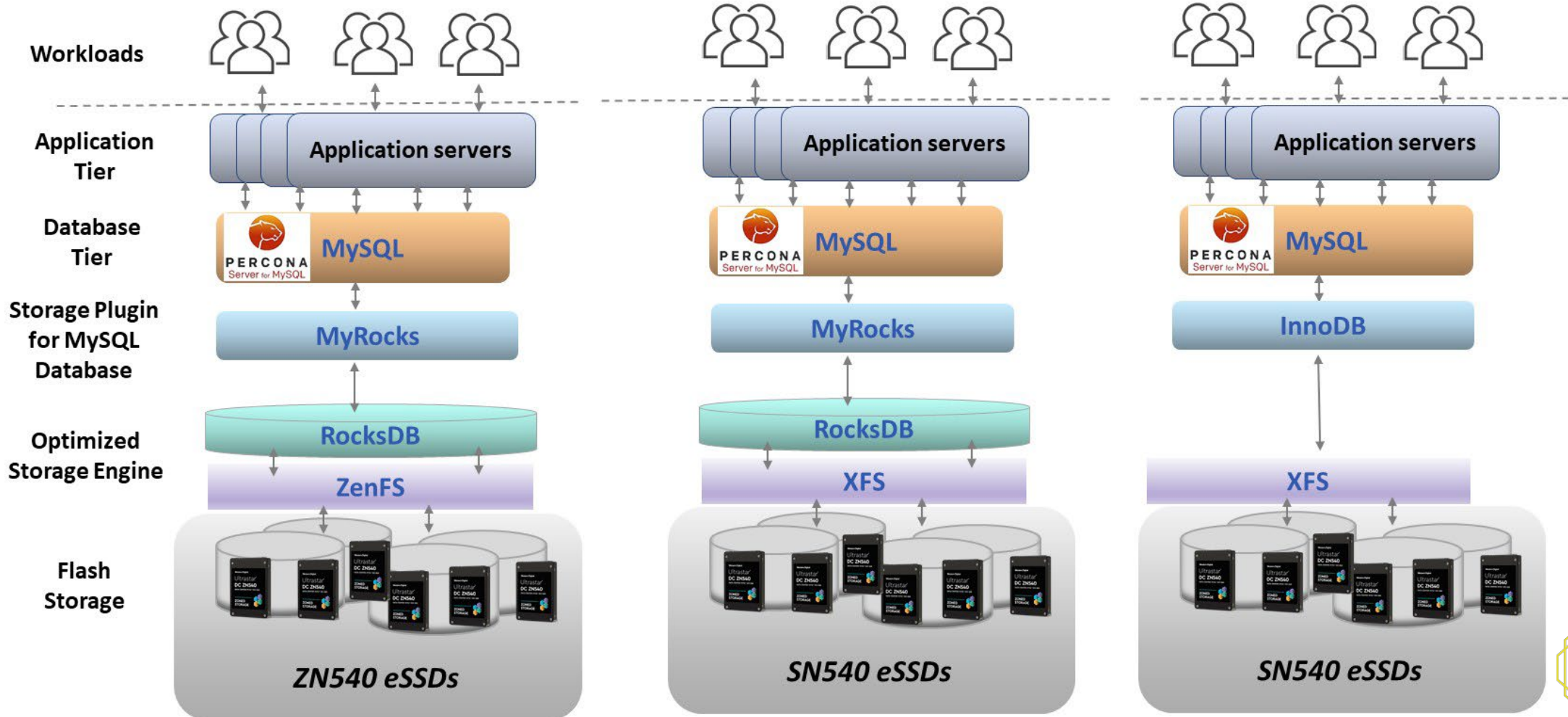


Percona Server for MySQL Solution Stacks: ZNS & Conventional

Percona MySQL/MyRocks/ZenFS/ZN540 (ZNS SSD)

Percona MySQL/MyRocks/XFS/SN540 (Conv SSD)

Percona MySQL/InnoDB/XFS/SN540 (Conv SSD)



Western Digital Example Configuration

Percona Server for MySQL 8.0 -- Lab Configuration

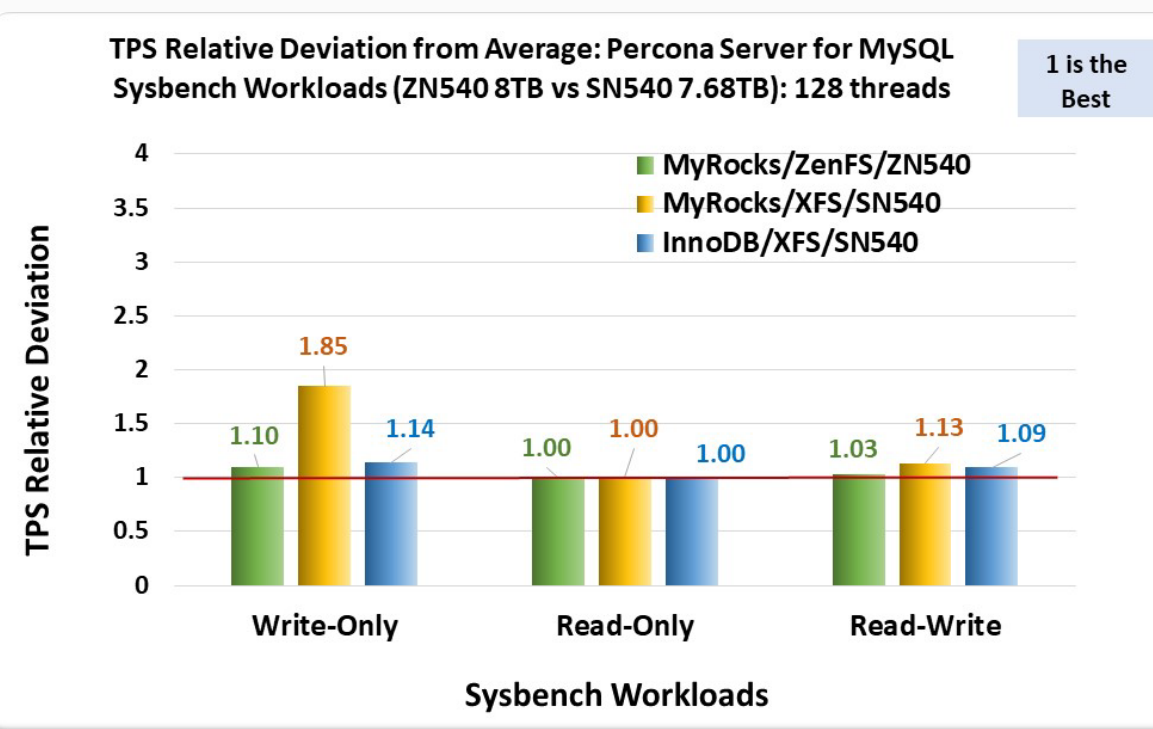
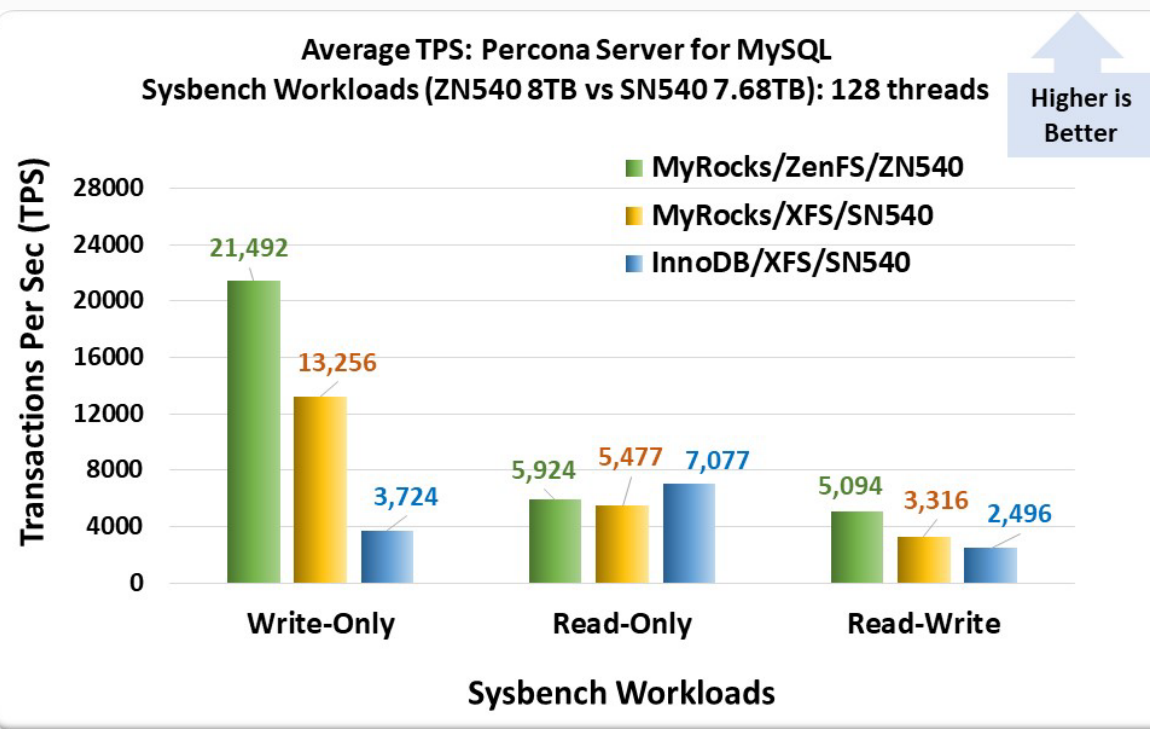
- 2U NVMe™ storage servers (Supermicro® SYS-2029U-TN24R4T)
 - 2x Intel® Xeon® 6258R 2.7 GHz 28-core
 - 384 GiB* memory
 - Two PCI-Express 3.0 x16 slots, One PCI-Express 3.0 x8 slot
 - Twenty-four NVMe hot-swap 2.5" bays, supporting U.2 form factor drives
 - populated with Ultrastar DC ZN540 eSSDs @4TB and @8TB for evaluations on ZNS drives
 - populated with Ultrastra DC SN540 eSSDs @3.84TB and @7.68TB for evaluations on conventional drives of equivalent spec
 - Kernel/OS: 5.10.72 LTS/Ubuntu 20.04.3 LTS distribution

*One terabyte (TB) equals one trillion bytes. Actual user capacity may be less due to operating environment.



Percona Server for MySQL 8.0 (ZN540 8TB vs SN540 7.68TB)

Average Transactions Per Sec (TPS) & Relative Deviation (8TB): Sysbench Workloads

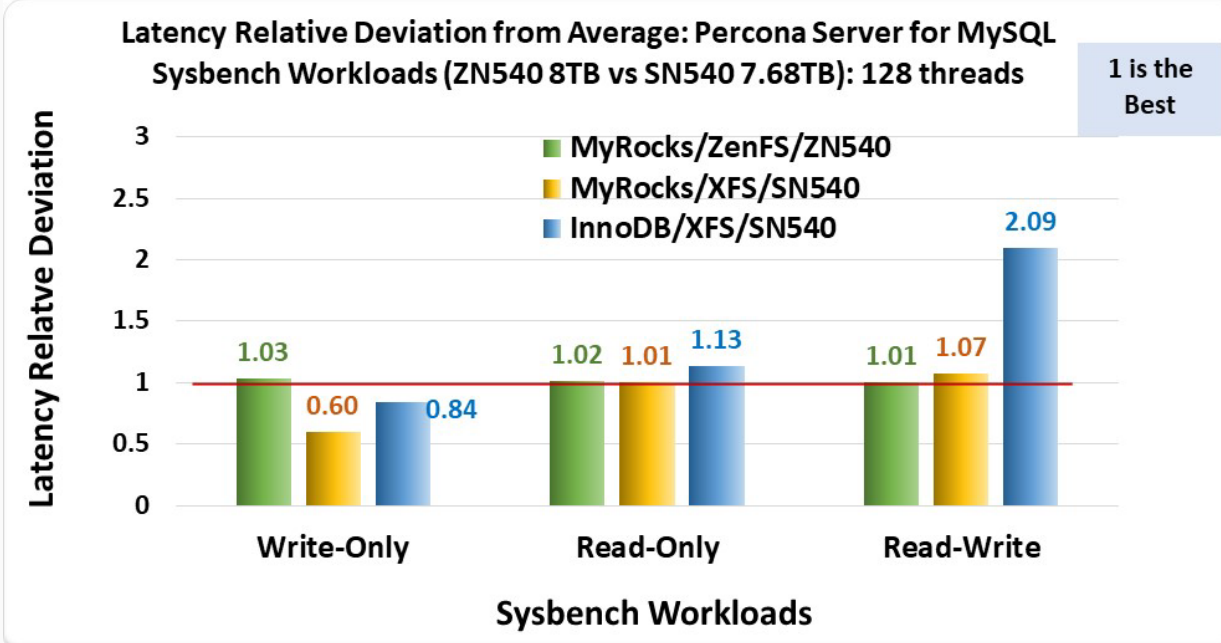
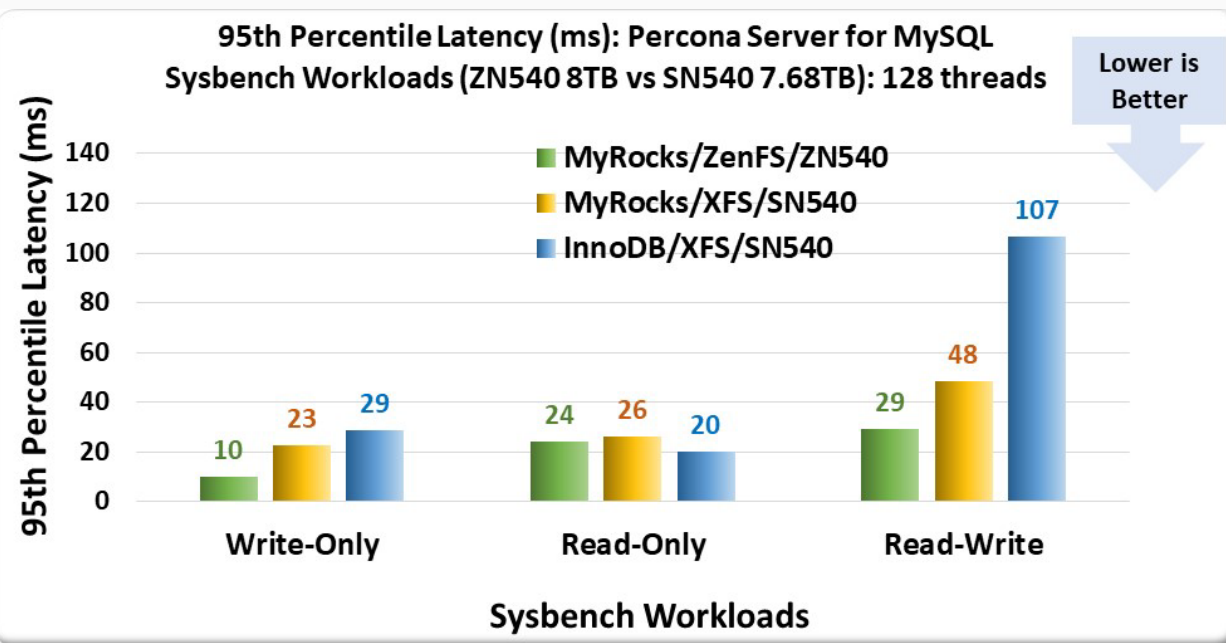


Percona Server for MySQL database 8.0 utilizing the MyRocks storage engine on ZenFS/ZN540 drives (8TB) yields an overall better TPS performance with minimal relative deviation from average, resulting in a better QoS, when compared to the MyRocks and InnoDB storage engines on preconditioned XFS/SN540 drives (7.68TB)

Disclaimer: The performance results presented are derived from iterative executions of industry-standard sysbench workloads using best known available hardware in Western Digital lab. Results may vary based on hardware specifications, database configuration, and workload goals and objectives.

Percona Server for MySQL 8.0 (ZN540 8TB vs SN540 7.68TB)

95th Percentile Latency & Relative Deviation (8TB) : Sysbench Workloads



Percona Server for MySQL database 8.0 utilizing the MyRocks storage engine on ZenFS/ZN540 drives (8TB) yields an overall lower (and better) 95th percentile latency, when compared to the MyRocks and InnoDB storage engines on preconditioned XFS/SN540 drives (7.68TB). The latency on ZenFS/ZN540 also shows a consistent minimal relative deviation across different workloads, resulting in a better QoS.

Disclaimer: The performance results presented are derived from iterative executions of industry-standard sysbench workloads using best known available hardware in Western Digital lab. Results may vary based on hardware specifications, database configuration, and workload goals and objectives.

Topics

- Application-level challenges to leverage benefits of underlying ZNS storage
- Deploying an optimal application ecosystem on ZNS made easy
 - Introducing Application-centric ZNS Reference Kits for accelerating end user deployments
 - Percona Server for MySQL® ZNS Reference Kit and its components
- Methodology for evaluating performance of the Percona MySQL solution stacks on ZNS and conventional SSDs
 - Performance results derived in Western Digital lab
- **ZNS SSDs for MySQL – Value Proposition**
- Ongoing work



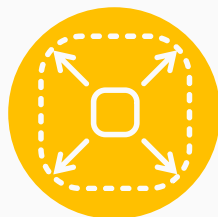
ZNS SSDs for MySQL – Value Proposition



Improves Performance and QoS

Improves responsiveness and multi-tenancy isolation for data services and apps

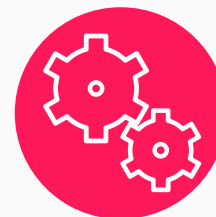
Up to **3 x** better TPS performance
Up to **4 x** reduced transaction latency *



Maximizes Capacity and Endurance

Increases SSD usable storage capacity and endurance

Up to **25%** higher usable capacity per SSD * with **3.5 DWPD** endurance



Based on Industry Standards and Open Source with Full-stack Support

Full vertical stack support for MySQL software

NVMe Zoned Namespace
ZonedStorage.io
Percona Server for MySQL



Reduces TCO

Triple the number of transactions processed on MySQL servers

Up to **62% TCO savings** for mixed read-write workloads *

* Compared to equivalent spec conventional NVMe SSD 0.8 DWPD



Topics

- Application-level challenges to leverage benefits of underlying ZNS storage
- Deploying an optimal application ecosystem on ZNS made easy
 - Introducing Application-centric ZNS Reference Kits for accelerating end user deployments
 - Percona Server for MySQL® ZNS Reference Kit and its components
- Methodology for evaluating performance of the Percona MySQL solution stacks on ZNS and conventional SSDs
 - Performance results derived in Western Digital lab
- ZNS SSDs for MySQL – Value Proposition
- **Ongoing work**



Ongoing Work

- Ongoing industry collaborations to enhance and standardize the application ecosystem around Zoned Storage with updated collateral on zonedstorage.io
- Enabling BTRFS filesystem on ZNS in Western Digital lab followed by implementing new applications like Apache Hadoop® / HDFS, Spark™, and others on BTRFS/ZNS
- Engaging with customers and partners as they continue to evaluate their workloads on ZNS technology



Western Digital®

Create What's Next

Western Digital, the Western Digital design, the Western Digital logo, and Ultrastar are registered trademarks or trademarks of Western Digital Corporation or its affiliates in the US and/or other countries. Apache®, Apache Hadoop, and Apache Spark, are either registered trademarks or trademarks of the Apache Software Foundation in the United States and/or other countries. Intel and Xeon are trademarks of Intel Corporation or its subsidiaries in the U.S. and/or other countries. Linux® is the registered trademark of Linus Torvalds in the U.S. and other countries. MySQL is a registered trademark of Oracle and/or its affiliates. The NVMe word mark is a trademark of NVM Express, Inc. SuperMicro is a trademark or registered trademark of Super Micro Computer, Inc. or its subsidiaries in the United States and other countries. All other marks are the property of their respective owners.