Creating Flash-Aware Applications

Nisha Talagala

NVM (Flash, other) is different from Disk

Area	Hard Disk Drives	Flash Devices
Logical to Physical Blocks	Nearly 1:1 Mapping	Remapped at every write
Read/Write Performance	Largely symmetrical	Heavily asymmetrical. Additional operation (erase)
Sequential vs Random Performance	100x difference. Elevator scheduling for disk arm	<10x difference. No disk arm – NAND die
Background operations	Rarely impact foreground	Regular occurrence. If unmanaged - can impact foreground
Wear out	Largely unlimited writes	Limited writes
IOPS	100s to 1000s	100Ks to Millions
Latency	10s ms	10s-100s us

https://opennvm.github.io

http://www.opencompute.org/projects/storage/

Contributions

FUSION-IO

Current OpenNVM Repositories

Flash-aware Linux swap

When working set size exceeds the capacity of DRAM, demand page from a flash-aware virtual memory subsystem.

Repository | Learn More

Key-value interface to flash

Create NoSQL databases faster. Automate garbage collection of expired data.

Repository | Learn More

Flash programming primitives

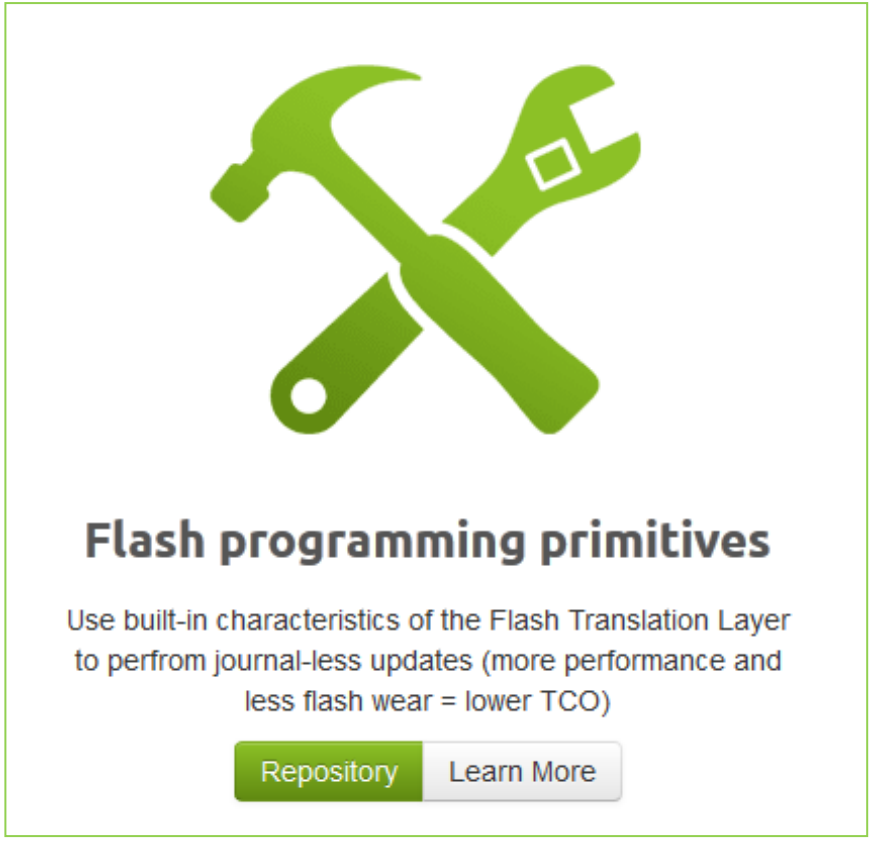
Use built-in characteristics of the Flash Translation Layer to perfrom journal-less updates (more performance and less flash wear = lower TCO)

Repository | Learn More

https://opennvm.github.io

1st Contribution: Flash Primitives

FUSION-IO



https://opennvm.github.io

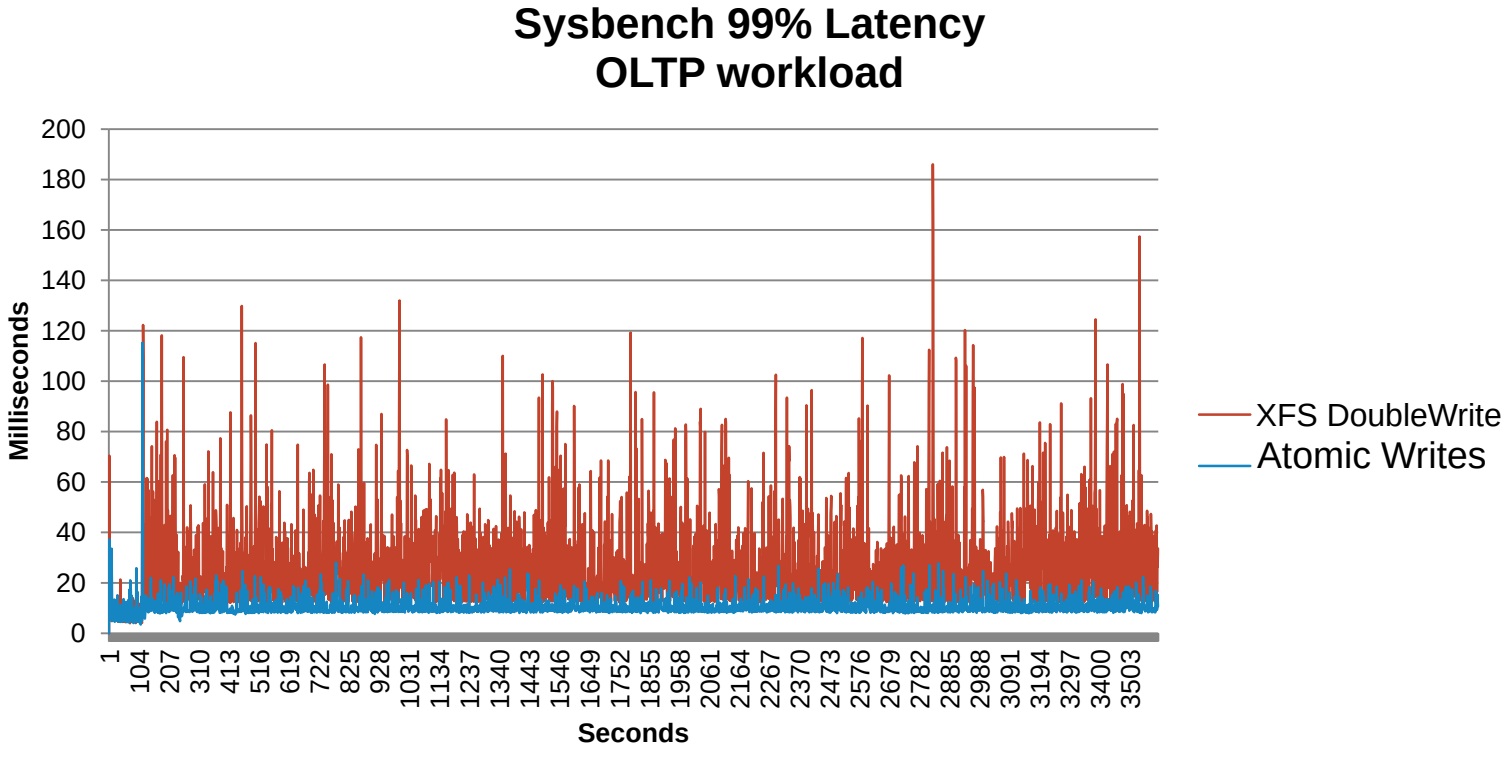
On GitHub:

- API specifications, such as:
  - *nvm_atomic_write()*
  - *nvm_batch_atomic_operations()*
  - *nvm_atomic_trim()*
- Sample program code

MySQL Example: Latency Improvement

FUSION-IO

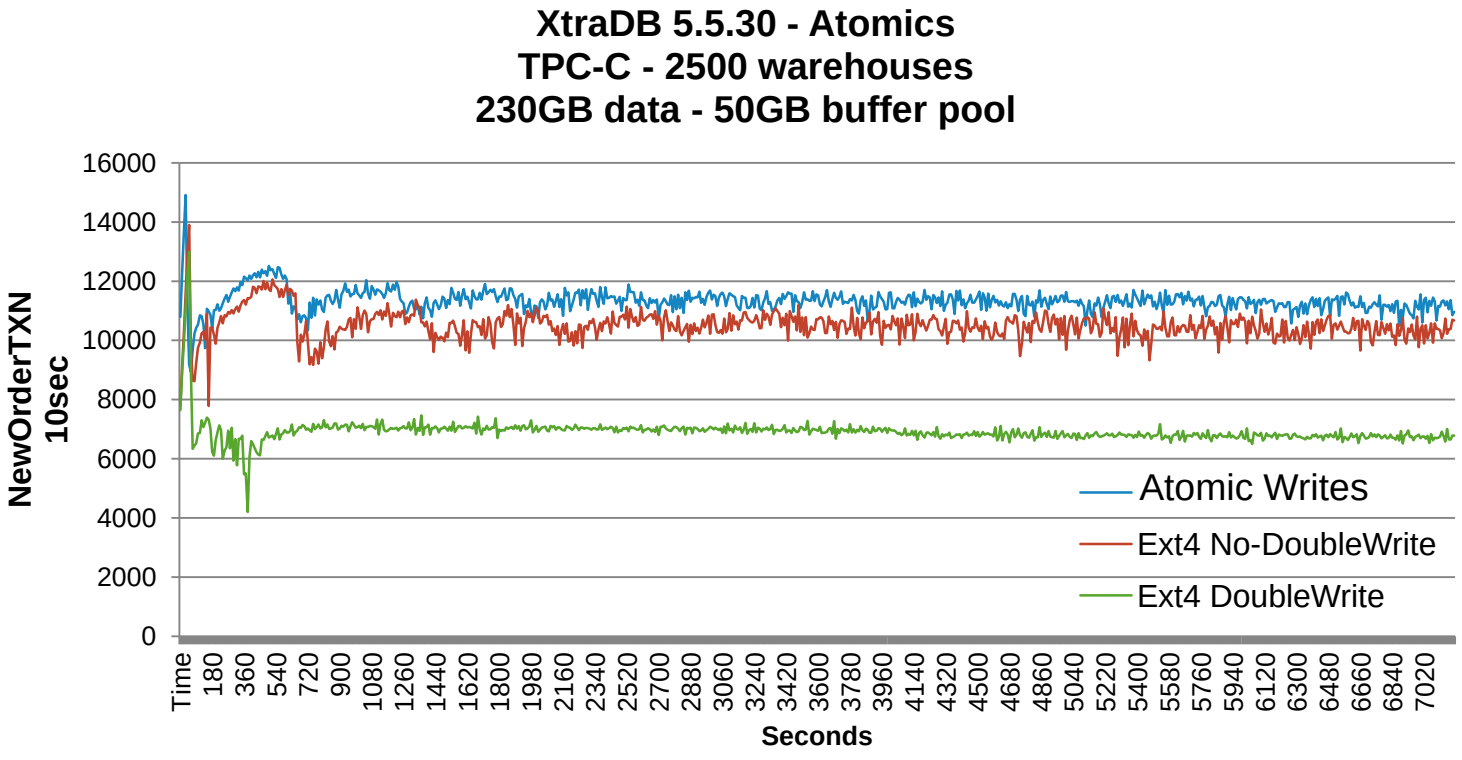
2-4x Latency Improvement on Percona Server



MySQL Example: Throughput Improvement

FUSION-IO

70% Transactions/sec Improvement on MariaDB Server



2nd Contribution: Linux Fast-Swap

FUSION-IO

On GitHub:

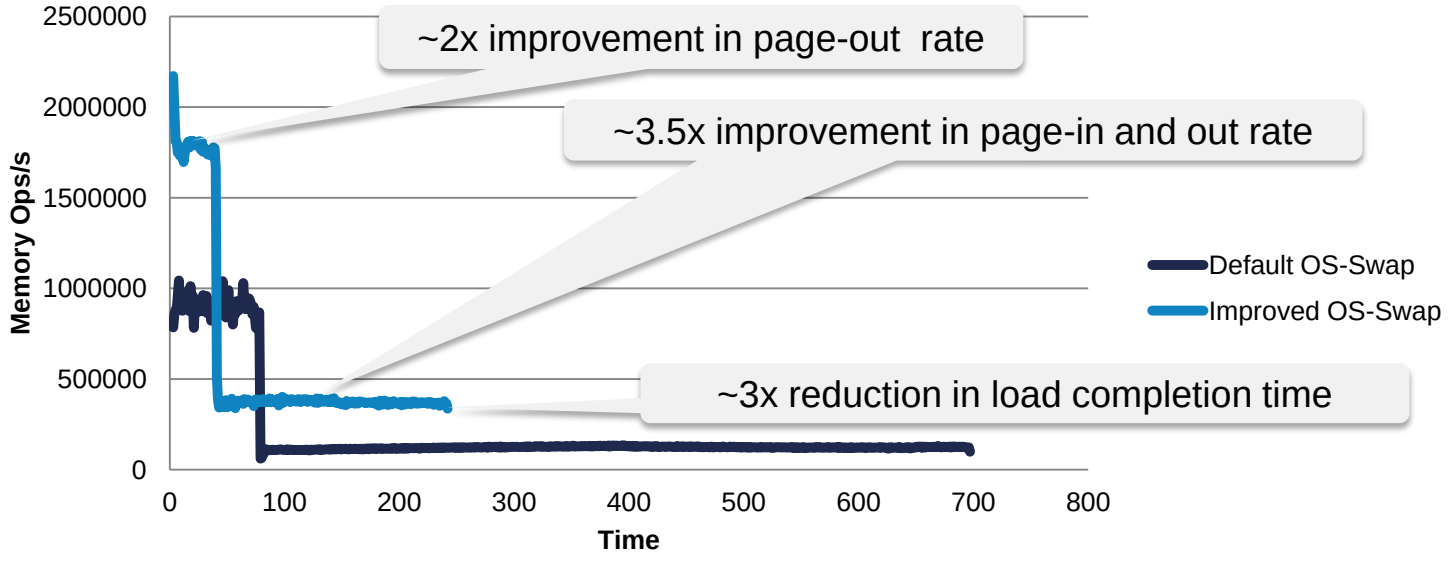
- Documentation
- Experimental Linux kernel with virtual memory swap patch (3.6 kernel)
- Benchmarking utility

https://opennvm.github.io

Fast Swap - Performance

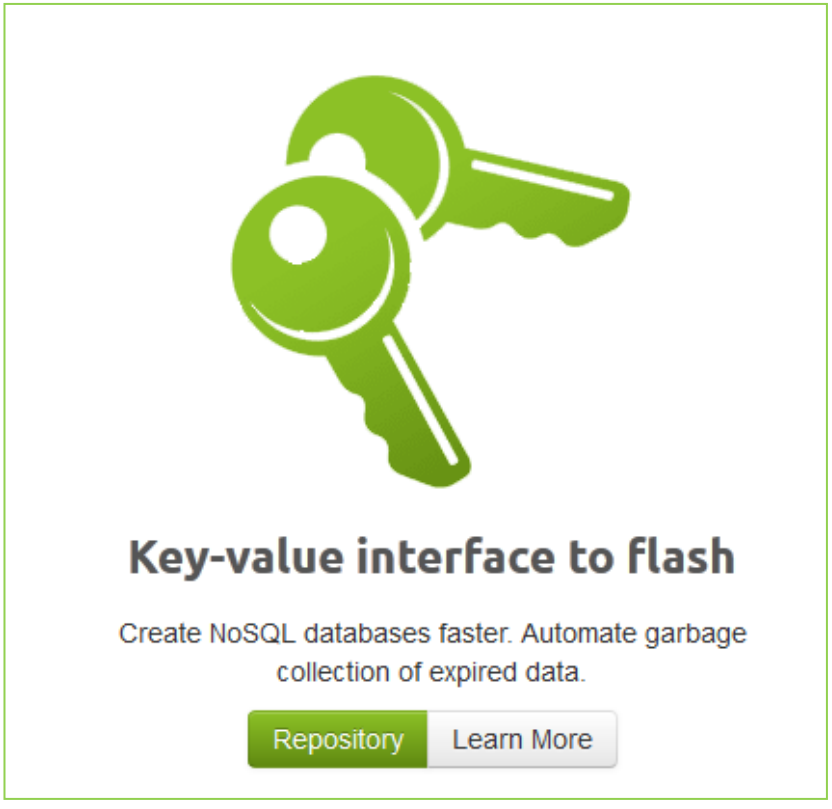
FUSION-IO

3x reduction in load completion time with fast swap



3rd Contribution: Key-Value Interface

FUSION-IO



https://opennvm.github.io

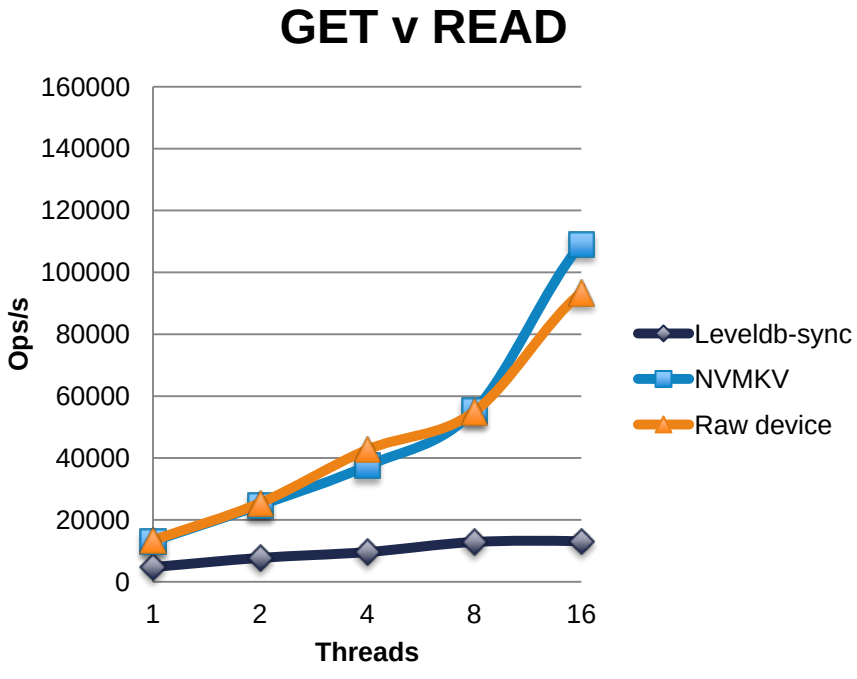
On GitHub:

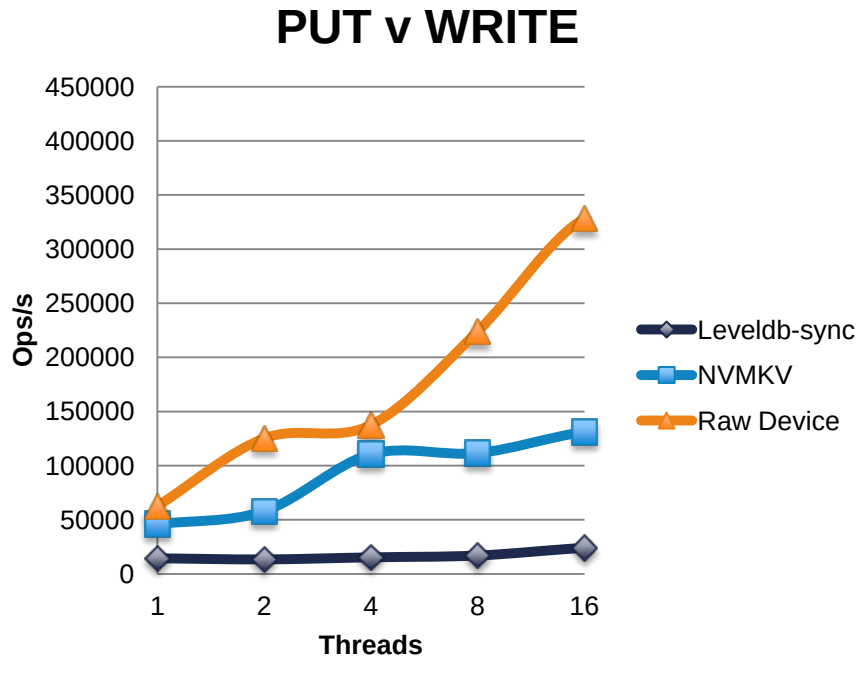
- API specifications, such as:
  nvm_kv_put()
  - *nvm_kv_get()*
  - *nvm_kev_batch_put()*
  - *nvm_kv_set_global_expiry()*
- Sample program code
- Benchmarking utility
- Source code (30 Aug)

Key-Value Interface - Performance

FUSION-IO

Key-Value get/put vs. Raw read/write vs. levelDB read/write





OpenNVM, Standards, and Consortiums

FUSION-IO

- opennvm.github.io
  - Primitives API specifications, sample code
  - Linux swap kernel patch and benchmarking tools
  - key-value interface API library, sample code, benchmark tools
- INCITS SCSI (T10) active standards proposals:
  - SBC-4 SPC-5 Atomic-Write
    http://www.t10.org/cgi-bin/ac.pl?t=d&f=11-229r6.pdf
  - SBC-4 SPC-5 Scattered writes, optionally atomic
    http://www.t10.org/cgi-bin/ac.pl?t=d&f=12-086r3.pdf
  - SBC-4 SPC-5 Gathered reads, optionally atomic
    http://www.t10.org/cgi-bin/ac.pl?t=d&f=12-087r3.pdf
- SNIA NVM-Programming TWG draft guide:
  http://snia.org/forums/sssi/nvmp

Apps Using OpenNVM technology

FUSION-IO

Learn More »

Learn More »

https://opennvm.github.io

https://opennvm.github.io

http://www.opencompute.org/projects/storage/


THANK YOU
fusionio.com | REDEFINE WHAT'S POSSIBLE