



# By the Numbers: ZFS Performance Results from Six Operating Systems and Their Derivatives

@MichaelDexter  
editor@callfortesting.org

vBSDcon 2019

STOP  
LOSING  
DATA

I ❤️ ZFS

*But it wasn't love at first sight*

# We met at v15...

RaidZ planning was complex

Lose your SLOG and lose your pool

*“You want 16GB more RAM than you currently have”*

# But we grew...

lz4 compression simplified RaidZ planning

By v28 a SLOG would fail gracefully

Commodity hardware caught up with ZFS

# *Still* Logical Conclusions

Rely on the CPU for data checksumming,  
compression, encryption...

Stripe vdevs to increase speed with capacity

CoW for snapshots, rollback, and replication

# *Continuous Data and Backup Validation*

Validated on Read

Validated at Rest

Atomically-Identical Backups

*You have a job to do*



*Fewer and fewer  
single-platform organizations*

*Ecosystems, not monocultures*

*Diversity by choice or even mandate*

*Putting the “v” in vBSDcon...*

# OpenZFS Platforms

illumos Distributions

FreeBSD and Derivatives

GNU/Linux Distributions

macOS

Windows/Server/Hyper-V

NetBSD

ZFS vs. OpenZFS

Oracle Solaris

*OpenZFS is becoming...*

The TCP/IP | POSIX | OpenSSH  
of File Systems

# \$DAYJOB

OpenZFS Crisis Line

FreeNAS/FreeBSD/Others

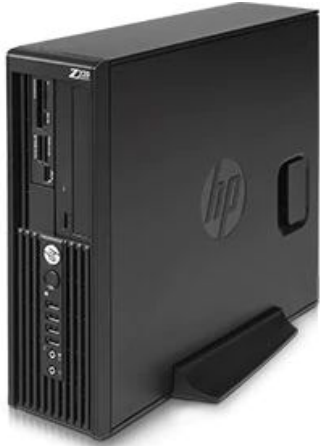
Zevo ZFS/OpenZFS on Mac since 2013

First ZFS on Windows on Hardware



We cannot have second-class  
OpenZFS Platforms

# The Lab



**HP Z220 X 10**  
Xeon E3-1225v2  
8GB RAM



**Overdrive 1000**  
AMD A1120 (A57 64-bit)  
8GB RAM



# The Lab 1.0

500GB 7200RPM Drives

Cygwin rather than PowerShell

Tribblix rather than OmniOSCE

# Methodology 1.0

Equidistant OpenSSH

touch, dd, mkdir, tree, realpath, zfs  
create dataset, native/Cygwin /bin/sh

*How's that portable BSD  
userland coming?*

# The Lab 2.0

Crucial BX500 120GB 3D NAND



*Come for the silence, stay for the reporting*

247 Host\_Program\_Page\_Count 0x0032 ... 6846185

# Methodology 2.0

Equidistant OpenSSH

FIO, the Flexible I/O Tester

(Where supported)

SmartMonTools/smart(8)

/bin/sh and PowerShell

# The Lab 2.0

macOS Mojave

OmniOSCE

Windows 10

CentOS 7

NetBSD 9.99

Debian 10

ZoL on FreeBSD 12

Oracle Solaris

Hyper-V Core 2019

FreeNAS 11.2

FreeBSD 12.0 ARM64

# The Lab 3.0 Hypervisors

HP Z420 Systems

FreeBSD/bhyve with OpenZFS

Hyper-V with OpenZFS

ESXi without OpenZFS

XCP-ng with OpenZFS

Proxmox with *root on ZFS!*



# The Lab 3.0

SmartOS

FreeBSD 11-stable

FreeBSD 12-stable

FreeBSD Current

FreeBSD/Xen 11-stable...

Ubuntu

NetBSD/Xen

GNU/Linux ARM64

Windows Server

# Benchmarks vs. Stress Testing

Mature platforms warrant  
benchmarking

New platforms warrant stress  
testing/fuzz testing



# Uncharted Territory

“There is no way touch (1) could fail.”

“That RAM upgrade couldn’t possibly  
brick this motherboard.”

“Ported, portable software is portable.”

Be Consistent • Assume Nothing

# Unchartered Territory

Always the tip of the iceberg

# zpool Strategy

```
zpool create -f \  
-O casesensitivity=insensitive \  
-O mountpoint=/tank \  
-O compression=off \  
-O atime=off \  
-O sync=standard|always
```

# FIO Strategy

```
fio --thread --randrepeat=1 \  
--direct=1 --gtod_reduce=1 \  
--name=test --bs=128k \  
--filename=/tank/fio.tmp \  
--iodepth=64 --size=8G \  
--readwrite=randrw (--sync=1)
```

# FIO Strategy

`thread pthread_create, not fork`

`randrepeat predictable across runs`

`direct=1 non-buffered I/O`

`gtod_reduce=1 reduce gettimeofday() calls`

`name, bs=128k, filename, size`

`iodepth=64 number of I/O units in flight`

`readwrite=randrw (--sync=1)`

# 8G Results sync/sync

Hyper-V ZFS	Read: 24.5MiB/s	Write: 24.6MiB/s
Hyper-V NTFS	Read: 127MiB/s	Write: 127MiB/s
CentOS ZFS	Read: 33.1 MiB/s	Write: 33.2MiB/s
FreeNAS ZFS	Read: 33.1 MiB/s	Write: 33.2MiB/s
CentOS Native	Read: 66.6MiB/s	Write: 66.7MiB/s
macOS ZFS	Read: 37.2MiB/s	Write: 37.3MiB/s
macOS HFS	Read: 83.6MiB/s	Write: 83.7MiB/s

# 8G Results sync/sync

Windows 10 ZFS	Read: 44.9MiB/s	Write: 44.9MiB/s
Windows 10 NTFS	Read: 128MiB/s	Write: 128MiB/s
NetBSD ZFS	Read: 46.9MiB/s	Write: 47.9MiB/s
NetBSD UFS	Read: 63.3MiB/s	Write: 63.4MiB/s
FreeBSD ARM64	Read: 69.2MiB/s	Write: 69.2MiB/s
Debian 10 Native	Read: 71.6MiB/s	Write: 71.7MiB/s
ZoLoF ZFS	Read: 83.7MiB/s	Write: 83.9MiB/s

OmniOSCE ZFS, Solaris ZFS, Debian ZFS TBD

# 8G Results async/async

NetBSD ZFS	Read: 69.3MiB/s	Write: 69.6MiB/s
macOS ZFS	Read: 79.5MiB/s	Write: 79.6MiB/s
macOS HFS	Read: 88.9MiB/s	Write: 89.1MiB/s
Windows 10 ZFS	Read: 87.1MiB/s	Write: 87.3MiB/s
Windows 10 NTFS	Read: 127MiB/s	Write: 127MiB/s
Hyper-V ZFS	Read: 90.9MiB/s	Write: 91.0MiB/s
Hyper-V NTFS	Read: 127MiB/s	Write: 127MiB/s



# 8G Results async/async

FreeNAS ZFS	Read: 100MiB/s	Write: 100MiB/s
CentOS ZFS	Read: 106MiB/s	Write: 106MiB/s
CentOS Native	Read: 116MiB/s	Write: 116MiB/s
FreeBSD ARM64	Read: 113MiB/s	Write: 113MiB/s
ZoLoF ZFS	Read: 143MiB/s	Write: 143MiB/s

NetBSD UFS, OmniOSCE ZFS, Solaris ZFS, Debian ZFS TBD

# 512M Results async/async

FreeBSD ARM64	Read: 473MiB/s	Write: 500MiB/s
macOS ZFS	Read: 1016MiB/s	Write: 1073MiB/s
CentOS ZFS	Read: 1456MiB/s	Write: 1538MiB/s
NetBSD ZFS	Read: 1504MiB/s	Write: 1490MiB/s
FreeNAS ZFS	Read: 1976MiB/s	Write: 2087MiB/s
ZoLoF ZFS	Read: 1915MiB/s	Write: 2023MiB/s

# Known Issues

ZFS on Windows is still Alpha

ZFS on Windows 'zfs send' needs help

macOS long-term performance

*Are we there yet?*

Universal root on ZFS

Solaris • Illumos • FreeBSD • Proxmox

Ubuntu is Promised

*Are we there yet?*

Add-on ZFS Platforms

Windows, CentOS, Debian, Ubuntu,  
XCP-ng, Zential, and more!

*Are we there yet?*

Boot environments across OS's  
Hypervisor 9pfs support

*Are we there yet?*

Baton Pass!

*Actual Output*

## **Sending from FreeNAS to ZoLoF**

Listing snapshots on ZoLoF

NAME	USED	AVAIL	REFER	
tank/baton@travel	0B	-	1.00G	-



## **Sending from FreeNAS to ZoLoF**

Listing snapshots on ZoLoF

NAME	USED	AVAIL	REFER	
tank/baton@travel	0B	-	1.00G	-

## **Sending from ZoLoF to FreeBSD ARM64**

Listing snapshots on FreeBSD ARM64

NAME	USED	AVAIL	REFER	
tank/baton@travel	0	-	1.00G	-

## **Sending from FreeBSD ARM64 to OmniOSCE**

Listing snapshots on OmniOSCE

NAME	USED	AVAIL	REFER	
rpool/ROOT/omnios...	82.0M	-	622M	-
tank/baton@travel	0	-	1.00G	-

## **Sending from FreeBSD ARM64 to OmniOSCE**

Listing snapshots on OmniOSCE

NAME	USED	AVAIL	REFER	
rpool/ROOT/omnios...	82.0M	-	622M	-
tank/baton@travel	0	-	1.00G	-

## **Sending from OmniOSCE to CentOS**

Listing snapshots on CentOS

NAME	USED	AVAIL	REFER	
tank/baton@travel	0B	-	1.00G	-

## **Sending from CentOS to Debian**

Listing snapshots on Debian

NAME	USED	AVAIL	REFER	
tank/baton@travel	0B	-	1.00G	-

## **Sending from CentOS to Debian**

Listing snapshots on Debian

NAME	USED	AVAIL	REFER
tank/baton@travel	0B	-	1.00G -

## **Sending from Debian to NetBSD**

Listing snapshots on NetBSD

```
ssh 10.13 /sbin/zfs list -t snapshot
```

## **Sending from NetBSD to macOS**

Listing snapshots on macOS

NAME	USED	AVAIL	REFER
tank/baton@travel	23K	-	1.00G
/tank/baton/.zfs/snapshot/travel			

```
macos:baton root# shasum -a 512 baton.img  
ce6210e6a71eb408517b994853af409... baton.img
```

```
macos:baton root# cat baton.img.sha512
```

```
SHA512 (/tank/baton/baton.img) =
```

```
ce6210e6a71eb408517b994853af409
```

# Wish List

OpenZFS Lite/Embedded

Permissively-licensed

Single-disk/NVMe only, no RaidZ

Minimal caching

*You have a job to do*



# Wish List cont.

About that Tweet...



A screenshot of a tweet from Michael Dexter (@michaeldexter) posted on August 9, 2019, at 11:04 PM. The tweet text is: "Looking for a few 100K to port #OpenZFS to #VMware and #OpenBSD. Has to happen, even if neither will be official. #SeenThings". The tweet has 8 retweets and 11 likes. The interface includes a profile picture, name, handle, a 'Follow' button, and a list of user avatars for interactions.

 **Michael Dexter**  
@michaeldexter Follow ⌵

Looking for a few 100K to port [#OpenZFS](#) to [#VMware](#) and [#OpenBSD](#). Has to happen, even if neither will be official. [#SeenThings](#)

11:04 PM - 9 Aug 2019

8 Retweets 11 Likes 

About that...



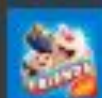
Calculator



Calendar



Camera



Candy Crush Friends



Connect

F



Feedback Hub



Firefox



Type here to search

Hyper-V Core

+

OpenZFS

+

Windows Admin Center

+

Chrome...

## hyperlab

## Tools



Overview

Azure hybrid services

Azure File Sync

Backup

Certificates

Devices

Events

Files

Firewall

## Overview

[Restart](#) [Shutdown](#) [Enable Disk Metrics](#) [Edit computer ID](#) [Manage alerts](#) [Refresh](#)

Computer name

hyperlab

Domain

-

Operating system

Microsoft Hyper-V Server

Version

10.0.17763

Installed memory (RAM)

8 GB

Disk space (Free / Total)

95.12 GB / 111.19 GB

Processors

Intel(R) Xeon(R) CPU E3-1225 V2 @  
3.20GHz

Manufacturer

Hewlett-Packard

Model

HP Z220 SFF Workstation

Logical processors

4

Windows Defender

Real-time protection: On

NIC(s)

1

Azure Backup status

[Not protected](#)

Up time

0:0:6:51

Logged in users

1

Hyper-V Tab

IPMI Tab

pfSense Tab

FreeNAS Tab

ESXi Tab

Proxmox Tab

Zential Tab...

Increasingly on OpenZFS

*See you there!*

November 4th - 5th  
San Francisco, CA



Jörgen Lundman and I will talk about  
OpenZFS Portability



# *Special Thanks*

Jörgen Lundman for  
ZFS on macOS/Windows

Conor Beh and Jason Barbier  
for Windows assistance

Verisign, Dan, and Team for vBSDcon

*Thank you!*

Questions?

@MichaelDexter  
editor@callfortesting.org