NVMe-oF Aware Filesystem Accelerates Machine Learning Workloads

Maor Ben-Dayan

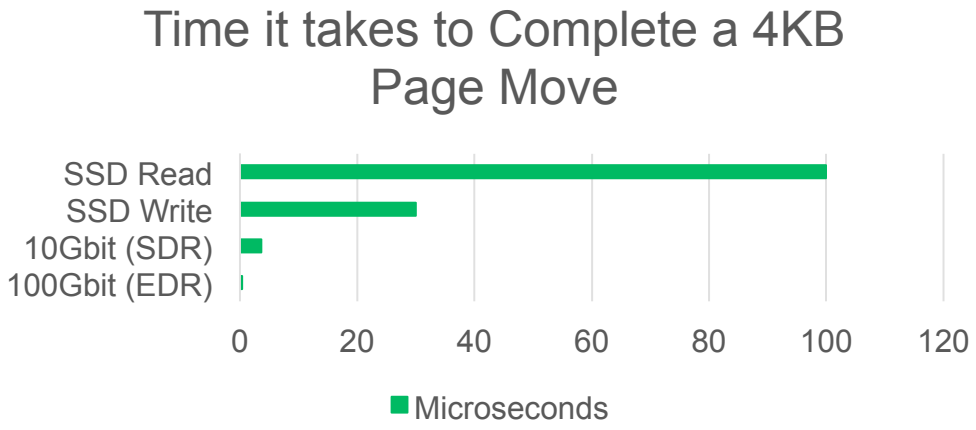
Chief Architect, Co-founder WekaIO

Flash Memory Summit 2018
Santa Clara, CA

WEKA.IO

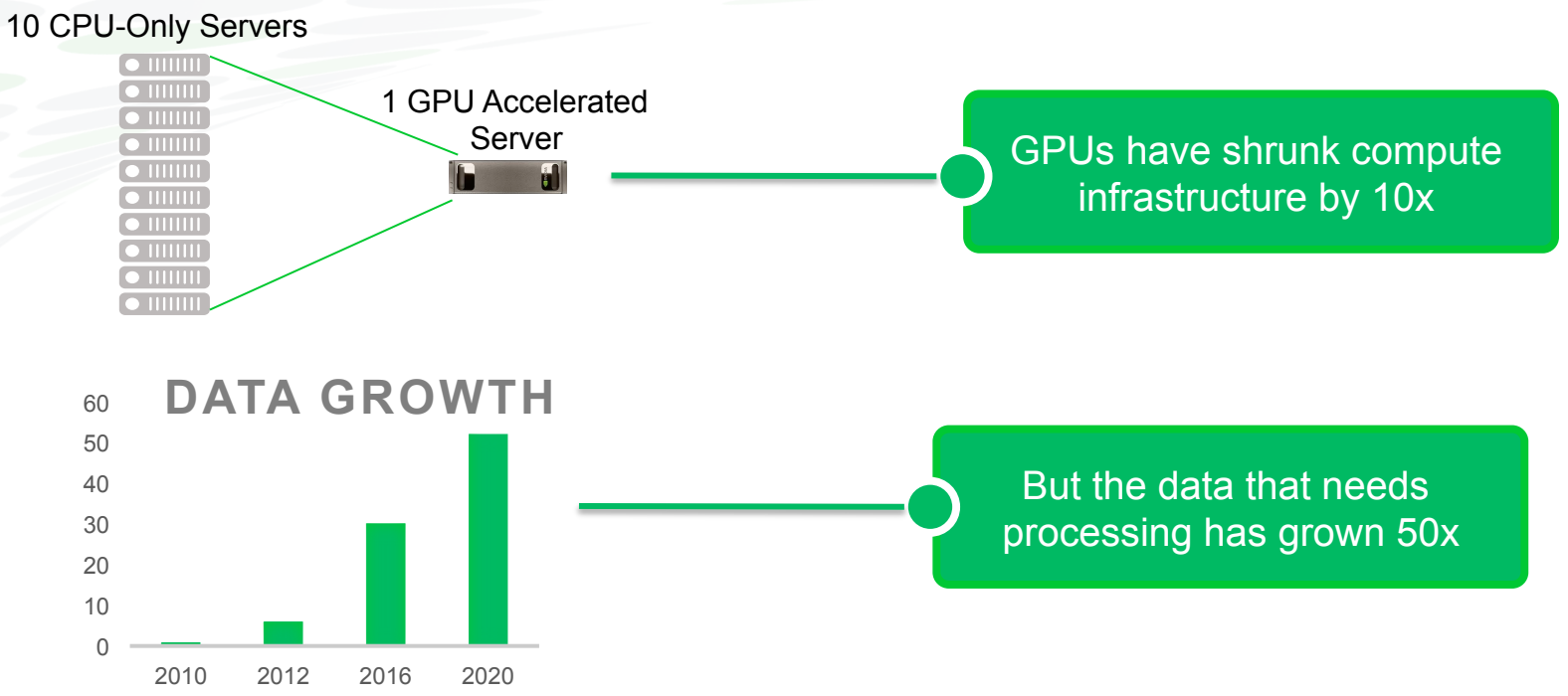
NVMe-oF

- NVMe is great!
- Still, modern networks on 100Gbit are much faster than SSD
- Locality is becoming irrelevant
- With the right stack, remote storage is more performant than local storage



WEKA.IO

The Beast Is Always Hungry



Industry is cornered into an I\O nightmare

WEKA.IO

Mission statement

Build a fast, scalable file system to serve the data hungry applications out there

WEKA.IO

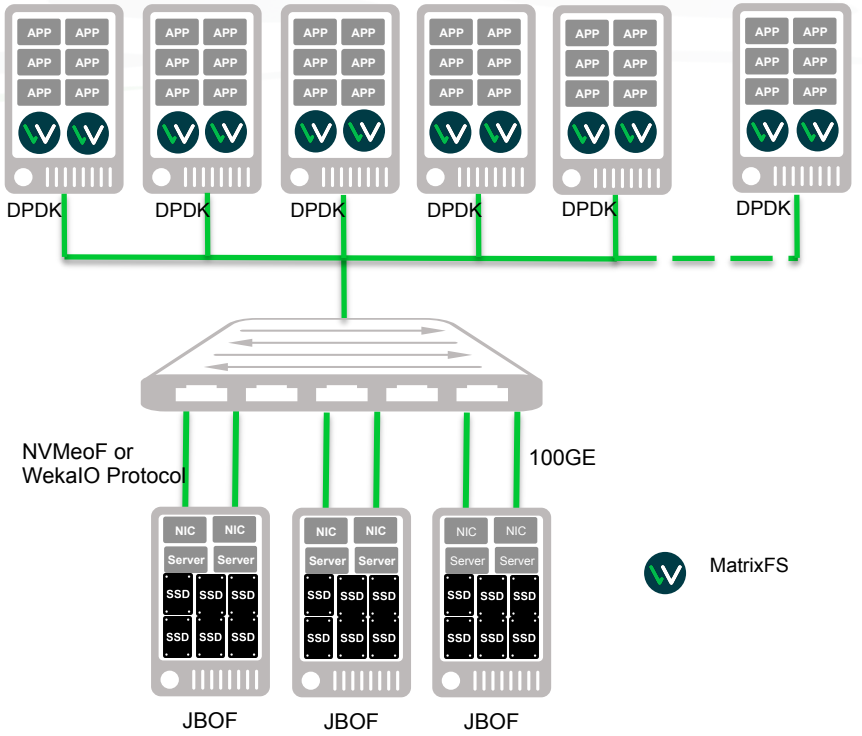
Problems to solve

- Access protocol (e.g. NFS) limitations
  - Develop our own client side agent
- Distribute data & metadata to avoid hot spots
- Work efficiently in NUMA architectures
- Avoid bottlenecks in the Linux kernel
- Keep it simple (as much as possible)

WEKA.IO

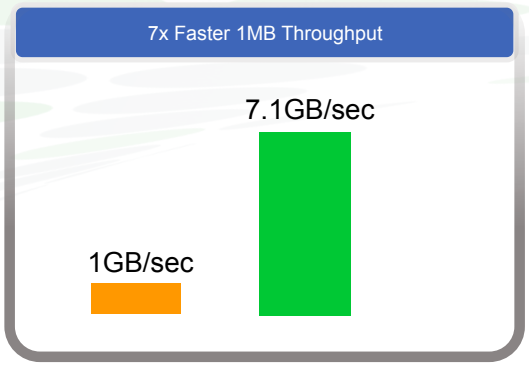
Deployment model

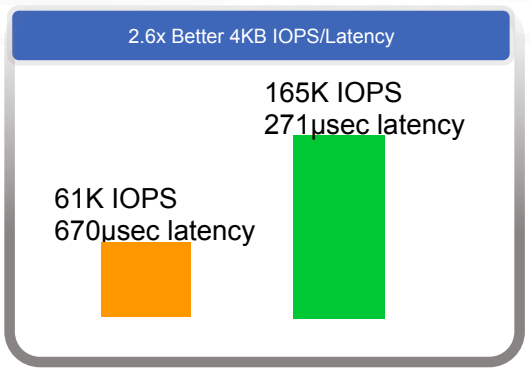
- WekaIO runs inside LXC container for isolation
- Provides POSIX VFS through lockless queues to WekaIO driver
- SR-IOV to run network stack and NVMe in user space
- I/O stack bypasses kernel
- Scheduling and memory management also bypass kernel
- Support, bare metal, container & hypervisor

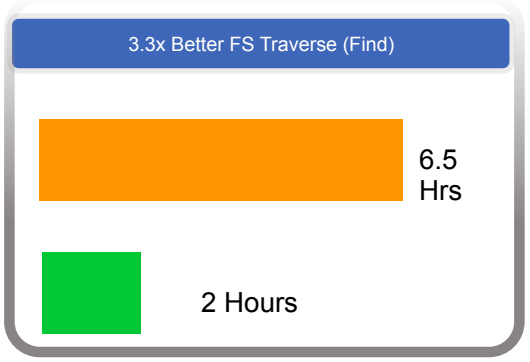


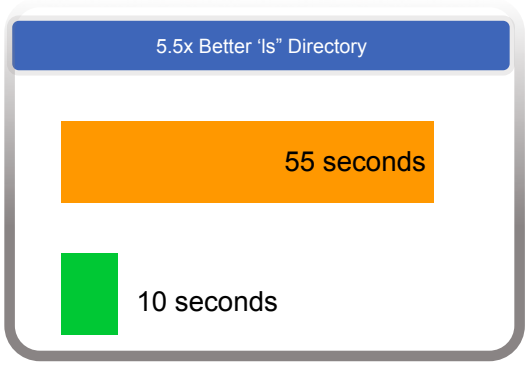
WEKA.IO

Actual Results from Bake-off vs All flash filer



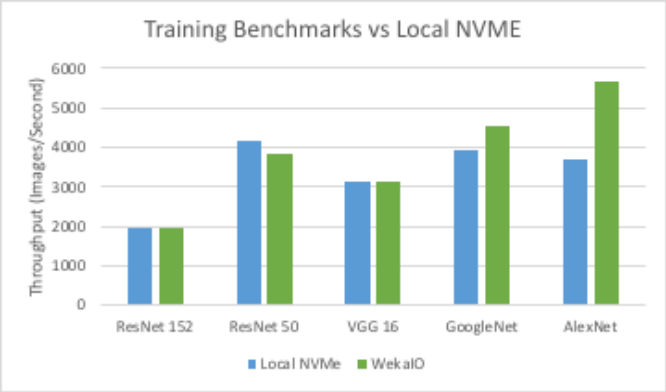


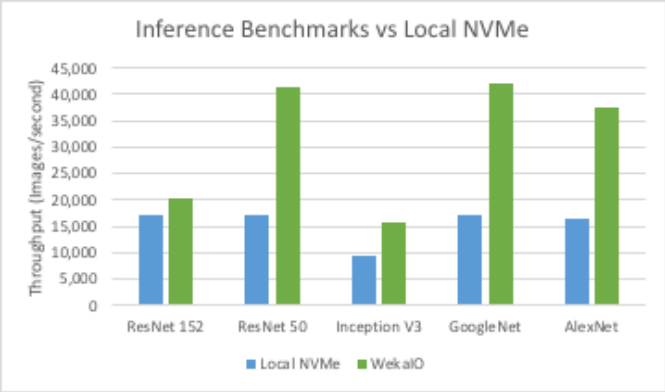


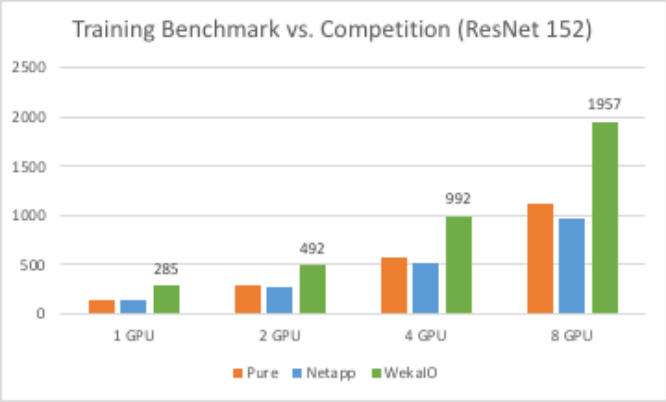


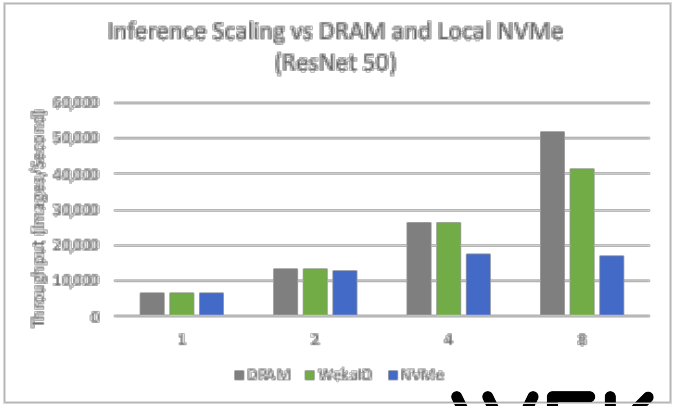
WEKA.IO

GPU Performance vs. Alternatives









https://www.theregister.co.uk/2018/06/07/pure_beats_netapp_at_ai/

http://dlpg.labs.hpe.com

WEKA.IO

Summary

- NVMe-oF enables performance breakthroughs
- Software stacks may need to be re-designed
- Call Weka.IO if you want your data to move faster

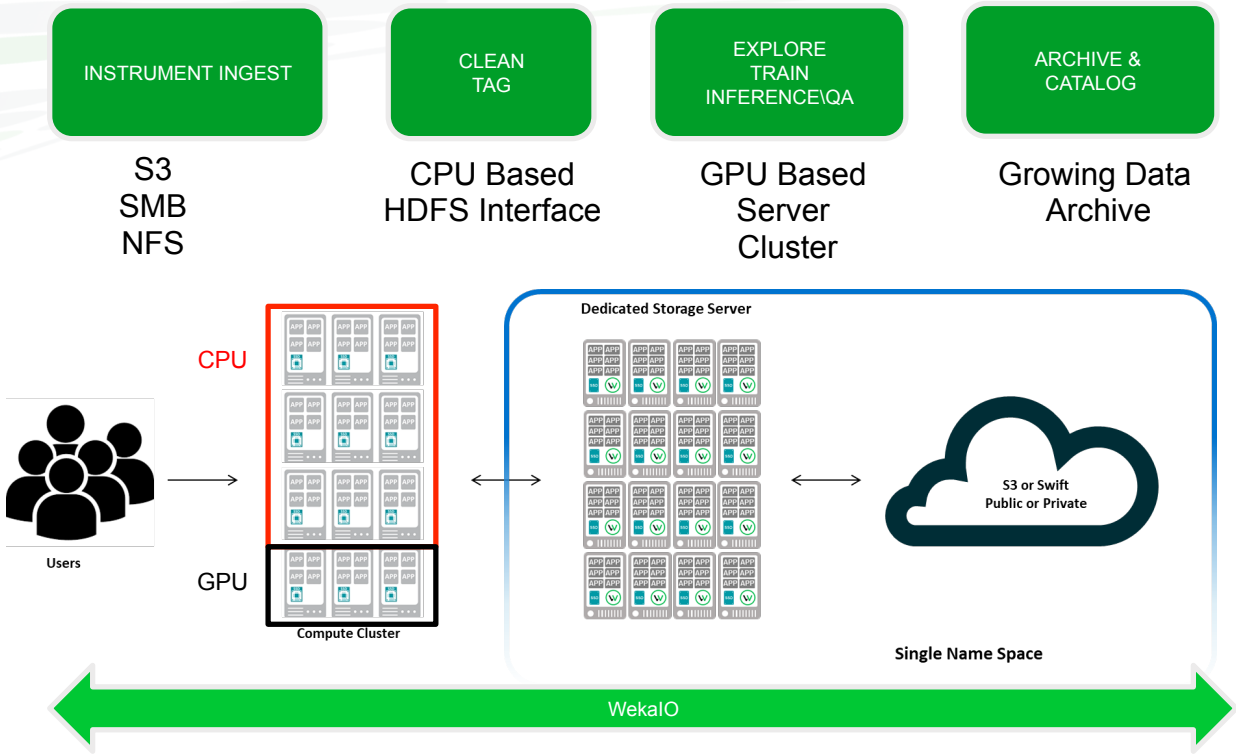
WEKA.IO

Flash Memory Summit

WEKA.IO

Flash Memory Summit 2018
Santa Clara, CA

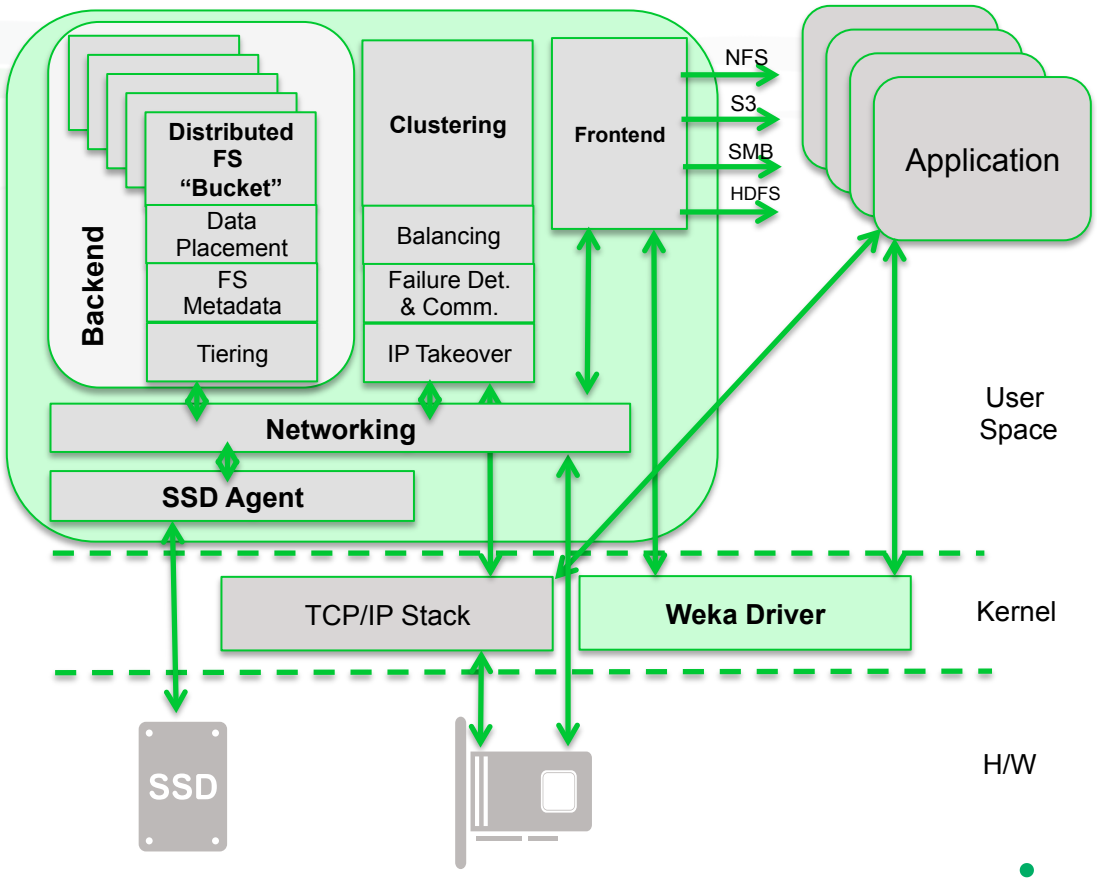
AI Requires an Appropriate IT Architecture



WEKA.IO

Software Architecture – Keep our of kernel

- Runs inside LXC container for isolation
- SR-IOV to run network stack and NVMe in user space
- Provides POSIX VFS through lockless queues to WekaIO driver
- I/O stack bypasses kernel
- Scheduling and memory management also bypass kernel
- Support, bare metal, container & hypervisor



WEKA.IO