# Architecting Flash for Scale in HPC

WEKA IO

Liran Zvibel, CTO
liran@WEKA.io
@liranzvibel

Company Overview

- Headquartered in San Jose, CA

- Engineering in Israel and USA

- Over $32.25M raised in series A and B

- Team bring deep technical and commercial expertise in storage

  - Intel, IBM, Netapp, Sandforce, XIV, Panasas, EMC

- 14 patents filed – 2 issued – 25 more identified

- Product in full production

  - Installed at leading semiconductor company and high-visibility Web 2.0 company

  - POCs in Life Sciences, EDA and Media and Entertainment

GEMINI ISRAEL VENTURES

WRV

QUALCOMM®

NVP NORWEST VENTURE PARTNERS.
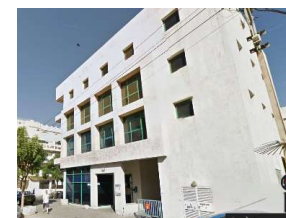
All Storage is Software Defined
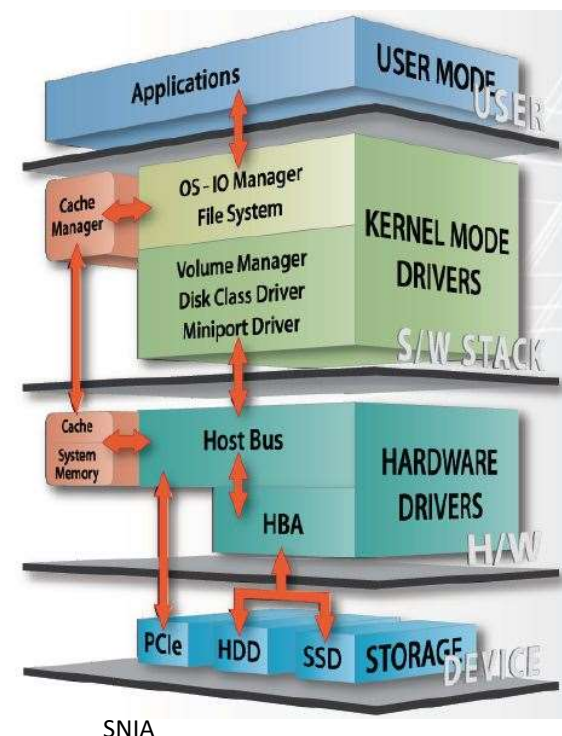
Software Defined   Vs.   Hardware Independent

Modern Storage System Design Requirements

- Architected for tomorrow's IT environments
    - High speed networks
    - Flash first
    - Hot and cold tiers for cost optimization
    - Commodity "unreliable" servers
    - Virtualized agile deployment
    - Consistent performance at scale

SNIA

Designed for Failure

Software Defined

- Homogenous
- High durability components
- Dedicated networks
- Failures are exceptional
- Tight bill-of-materials
- Design to slowest component

OLD WAY

NEW WAY

Hardware Independent

- Heterogeneous
- Cheapest components
- Poor networks
- Failures are normal
- Latest off the truck
- Design to the average of all

Leveraging the Shift to Flash

- Everything changed with SSDs – 2007 timeframe
  - o Data Structures
  - o Power failure
  - o Endurance
  - o Reliability
 - o Latency

Architecting for Scale

- Triple Replication Costs

    o 3x cost, 1/3 life

    o Impacts power, network traffic, footprint, failures

- Erasure coding reduces overhead by over 80%

    o Properly designed to minimize write amplification

    o Cannot impact performance

- Independent scaling of performance and capacity

Latency Determines Winners

- IOPS - Solved
  - o Bottleneck moved to software, controllers, and network
- Legacy stacks designed for disk latencies
  - o Requires I/O path efficiencies
- ~200 microsecond latencies over standard Ethernet

Rethinking Possibilities for HPC

- All legacy workloads run faster on a unified architecture

- Burst buffer elimination

- Downstream analysis in genomic research

- Small I/O workloads
  - EDA, financial trading, sensor data
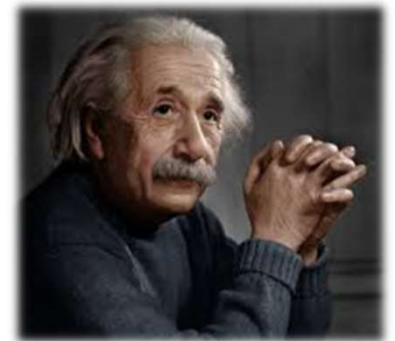
"Logic will get you from A to B. Imagination will take you everywhere."
*– Albert Einstein*

Summary

- SSD has a transformational impact on I/O
  - Data structures, I/O path, resiliency, endurance, cost
- Legacy file systems will not solve the small file challenge
- Requires a next generation file systems
  - Flash centric
  - Virtualized, tiered, scalable, modern interfaces

Questions

Characteristic	Impact
Front End	Ops and Metadata
Back End	Throughput
Archive	Scale, Cost, Capacity, Footprint, TCO

WEKA.io is the only company that addresses all three challenges
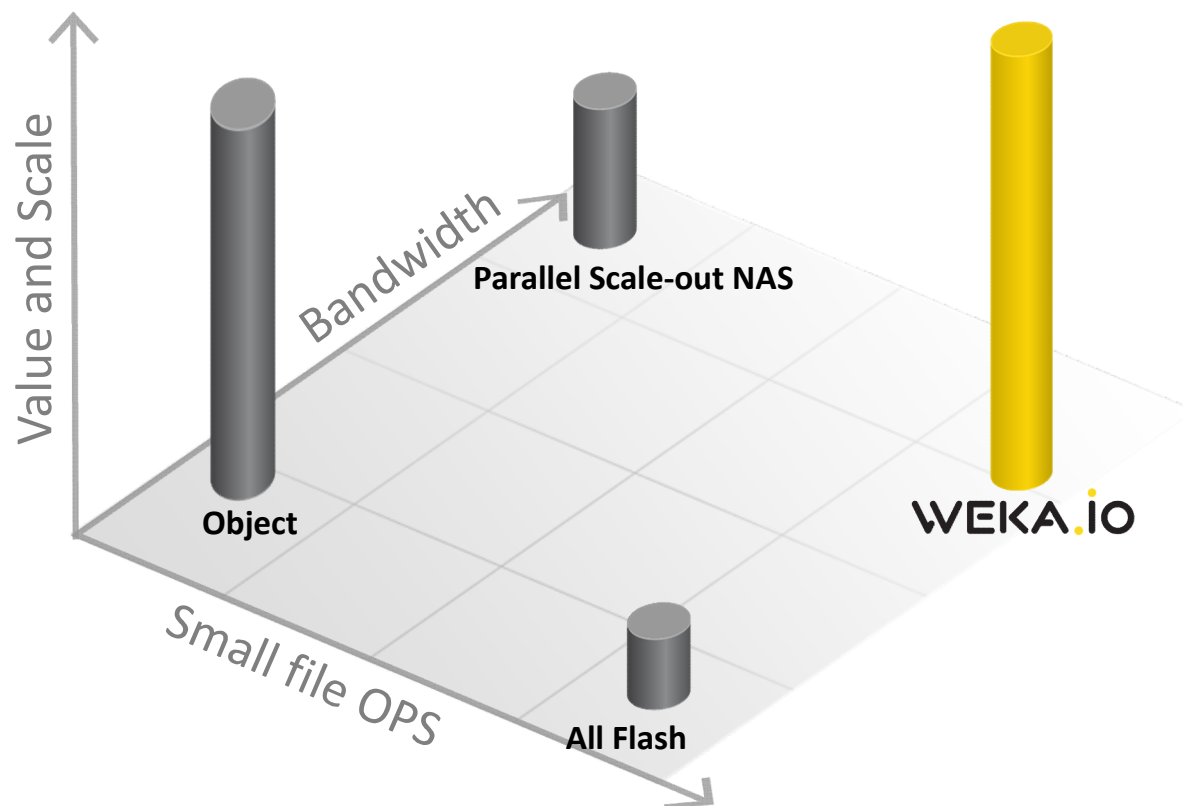
Peta (10^{15})

Exa (10^{18})

Zetta (10^{21})

Yotta (10^{23})

Xona (10^{27})

Weka (10^{30})

WEKA.io

—

Thank You