Transforming an SSD into a Cost-Effective Edge Server

Neil Werdmuller
Director Storage Solutions
Arm
Who am I?

- 28 years’ experience in embedded
- Lead the storage solutions team
- Work closely with the industry’s top storage suppliers
- Previously in wireless at Texas Instruments
- BSc in Computer Science from Portsmouth University (UK)
- I enjoy brewing beer at home!
What will we cover today?

- Why compute is moving to the edge
- Design structure of servers and SSD storage
- Why run an SSD as an edge server?
- Ecosystem support available
Arm computing is everywhere

- #1 shipping processor in storage devices
- 21Bn Arm-based chips shipped in 2017
- > 5Bn people using Arm-based mobile phones
- 120Bn Arm-based chips to date
Why computation is moving to the edge
Edge server

Compute:
• Arm Cortex-A series or Intel x86

Memory:
• DRAM

Storage:
• SSD/HDD

Interfaces:
• Ethernet + PCIe/SATA...
SSD storage

Compute:
• Arm Cortex-R series or Cortex-A series

Memory:
• DRAM ~1GB for each 1TB of flash

Storage:
• Today 256GB to 64TB... flash storage

Interfaces:
• PCIe/SATA/SAS...
Compute:
- Arm-based system-on-chip (SoC)

Memory:
- Shared DRAM

Storage:
- Shared Flash

Interface
- Ethernet...

**Edge SSD**

**SSD Based Edge Server:**
- **Network Interface**
  - **DRAM**
  - **CPU(s)**
  - **Flash**

**Classic Edge Server:**
- **Network Interface**
  - **DRAM**
  - **CPU(s)**
  - **PCIe**
  - **CPU(s)**
  - **Flash**
  - **DRAM**
SSD controllers

Compute:
- Frontend: Host I/F + Flash Translation Layer
  - Typically Cortex-R or Cortex-A series
- Backend: Flash management
  - Typically Cortex-R or Cortex-M series
- Custom accelerators: encryption, LDPC,…

Memory:
- DRAM typically >1GB for each 1TB of flash
Cortex-A series apps processor running:
- High-level OS such as Linux
- Networking protocols: Ethernet, TCP/IP, RDMA...
- Open source applications and workloads
- Standard administration functions

Edge SSD server:
- Single PCB and product housing
- Apps processor(s) added in to a SSD
- Ethernet network interface
- Internal interface to SSD controller SoC
  - PCIe or could be enhanced by chip-to-chip interface
Optimizations

Separate Cortex-A series processor
- Enables any SSD (or HDD) to run Linux
- Wide performance range from Cortex-A5 to Cortex-A76

Single SoC for cost/latency reduction
- Lower latency by removing internal (PCIe) interface
- Separation of apps processor and the SSD processing
- Shared DRAM and other SoC resources

Combined into frontend/apps processor
- Hypervisor provides SSD frontend separation from Linux
- Lowest cost and tightest integration
- Lowest possible latency
- Highest internal bandwidth
Why a SSD for an Edge server? (1 of 3)

An SSD SoC has flexible compute
• And an Edge server requires storage anyway

Lowest cost
• Adding ‘additional’ compute in SoC inexpensive
  • More, or larger, cores or additional clusters
  • Custom acceleration easily integrated: ML, FPGA…
  • Single DRAM reduces requirements

Robust
• Core SSD functionality unmodified
• SSDs can include power loss protection
Why a SSD for an Edge server? (2 of 3)

Lower power
- Shared single DRAM
- Minimal data copying and movement
- Less protocol processing

High bandwidths
- Ethernet to high bandwidth internal buses
- Parallelism of wide backend processing
- Removes bottleneck of PCIe lanes

Lowest latencies
- No internal cross interface latencies
- In-storage compute or inference all in DRAM
Scalability of compute and storage
• Single low power core to multiple clusters of high-performance cores

Flexibility
• One SoC that is suitable for:
  • Edge SSD, NVMe-oF, In-storage compute,…
  • Arm NEON and Compute Library for ML…
  • Potential for additional Arm or 3rd party ML

Security
• TrustZone isolates Linux and SSD functionality
• Processing of data is all done on the drive
• Decrypted data remains on the drive
Linux ecosystem on Arm

www.linaro.org
A few ‘Works on Arm’ partners

FreeBSD®

ubuntu

debian

openstack

kubernetes

docker

Works on arm

www.worksonarm.com
Conclusions for Edge SSD

• Edge SSD is an evolution not a revolution

• Edge SSD delivers with low-cost, low-power and lowest-latency

• High-performance, high-bandwidth and highly flexible

• Edge SSD and In-storage compute opens up many possibilities
  • This presentation will be made available in FMS proceedings
  • Please come to my second talk on Thursday at 8:30am for more information...
    • COMP-301-1: “Bringing Intelligence to Enterprise Storage Drives”
I’ll be here all week!

For more information, visit storage.arm.com.

neil.werdmuller@arm.com
linkedin.com/nwerdmuller
Thank You!
Danke!
Merci!
谢谢!
ありがとう!
Gracias!
Kiitos!
감사합니다
धन्यवाद