OSF for Intel Xeon-SP OCP Servers Reaching New Height

Meta: Jonathan Zhang

Wiwynn: Johnny Lin

Quanta: Tim Chu

9E: Arthur Heymans

SysPro: Marc Jones

Connect. Collaborate. Accelerate.

OPEN Compute Project®

Agenda

- High level Status / Plan
- OCP DeltaLake server in your fingertip
- New Features / Changes
- Q&A

Status

Intel Xeon Scalable Processor	OCP Server	Intel FSP API Mode Status	OSF Status
Sky Lake SP	TiogaPass	Proof Of Concept	POC achieved (2020)
Cooper Lake SP (CPX-SP)	DeltaLake	Statement Of Work	Pre-production ready achieved, OCP acceptance achieved (2021)
Sapphire Rapids SP (SPR-SP)	Being developed	Plan Of Record	Production ready in progress

Production Ready

At least on-par with traditional firmware approach on:

- PO/EVT/DVT/PVT schedules
- Validation Scope
    - Cross functional group testings equivalent to traditional firmware
- Feature set
- Stability/Performance/Power

Eco-system

- Collaborating with multi-hyperscalers on the technology development, great progress
  - Meta targeting single socket server for production ready
  - ByteDance targeting dual socket server for production deployment
  - And several other hyper-scalers doing technical preparation
- Shared code base (private till Intel SPR-SP product launch)
- Multiple collaboration channels
- Frequent meetings

Primary technology contributors

- Intel
- Hyperscalers
  - Amazon
  - ByteDance
  - Google
  - Meta
- ODMs / OEMs
  - Inspur
  - Quanta
  - SuperMicro
  - Winwynn
- Independent Firmware Vendors
  - 9E
  - SysPro
- coreboot/LinuxBoot/kernel communities

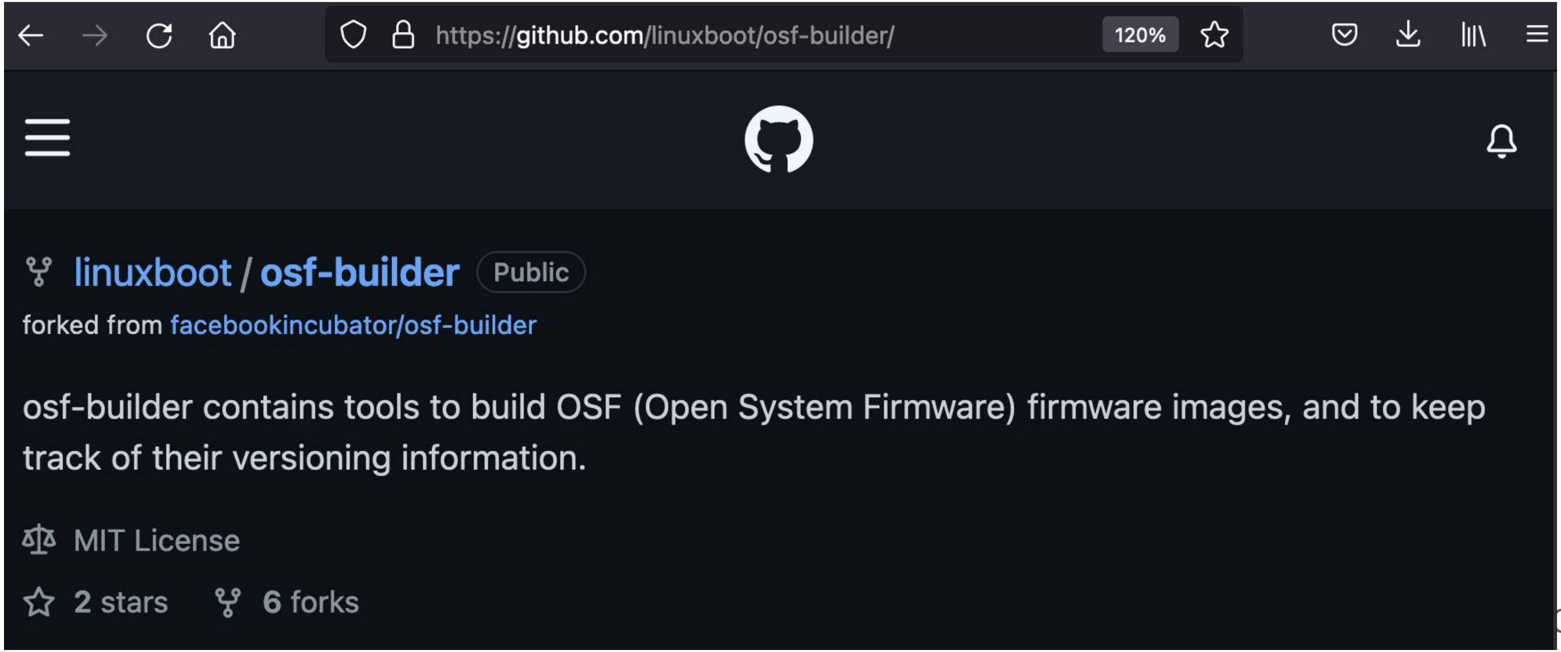
Benefits Observed

- Accelerate new technology development
- Enable vertical customization
- Improve troubleshooting velocity
- Leverage community resources

Near Term Plan (H2’ 2022)

- Along with Intel SPR-SP product launch, enable Ecosystem to benefit from OSF
    - Intel FSP binary, FSP integration guide, FSP release process
    - Intel ME ignition binary
    - coreboot support for
        - SPR-SP processor
        - Intel ArcherCity Customer Reference Board
        - Single / dual socket customer servers
    - LinuxBoot osf-builder support for Inter ArcherCity CRB, and customer servers
    - OCP OSF white papers (why/how)
        - What’s the benefits
        - How to build/run on ArcherCity CRB
        - How to enable OSF for **YOUR** SPR-SP based servers
        - How to contribute back to the ecosystem

OSF Builder



linuxboot / osf-builder Public

forked from facebookincubator/osf-builder

osf-builder contains tools to build OSF (Open System Firmware) firmware images, and to keep track of their versioning information.

MIT License

2 stars 6 forks

OCP DeltaLake Server in Your Fingertip

- OSF submission for DeltaLake https://github.com/opencomputeproject/OpenSystemFirmware/tree/master/Wiwynn/deltalake
- OCP community lab
  - OSF checklist validation, easy access for community development, testing and evaluation.
- Easy to reserve
- Easy to build/flash/boot

Wiwynn’s work through CraterLake

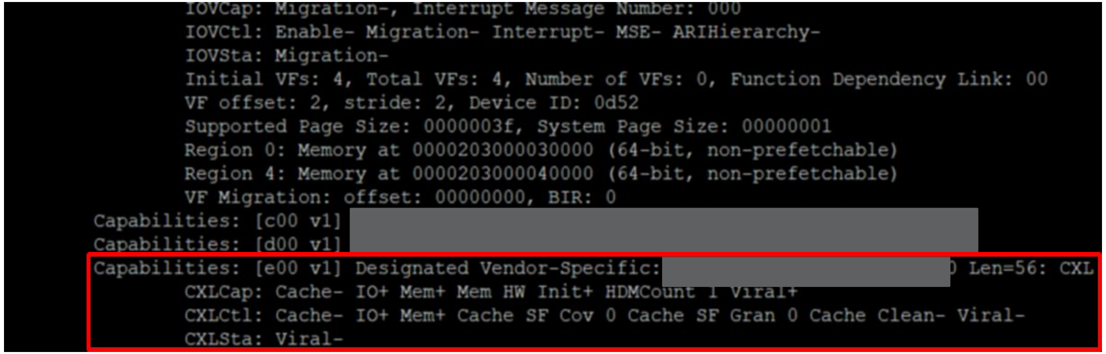
- CXL type 3 support, CXL memory available as system memory
- ME HECI-1 interface support, able to manage ME from host
- RAS: DMI, PCIe eDPC, memory, CPU firmware first error handling
- FSP EWL (Enhanced Warning Log) error handling for error DIMM reporting via BMC SEL right after MRC training
- SMM runtime serial log control via VPD firmware variable
- Improve u-root GRUB BLSCFG support in ‘boot’ command: support CentOS 8 and other newer distributions

QCI’s work through DeltaLake/S9S

- DeltaLake – Single socket server based on Intel CopperLake-SP
- S9S – Dual socket server based on Intel SapphireRapids-SP
- Early porting for SuperI/O to activate serial ports and debug port.
  - PCH Porting
    - Enable decoding of I/O address for SuperI/O (0x2E/0x2F and/or 0x4E/0x4F) and serial ports (0x3F8 - 0x3FF and 0x2F8 - 0x2FF)
  - SuperI/O Porting
    - Configure and activate above serials port address.
    - Configure respective multi-function pins for I/O 80 port to output POST codes.
  - https://review.coreboot.org/c/coreboot/+/40481 is reusable for new platform with the same PCH and SuperI/O.

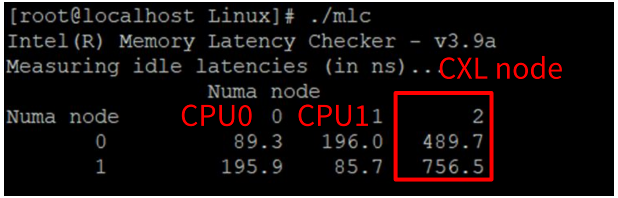
QCI's work through DeltaLake/S9S

- CXL card support on dual socket system
  - CXL capabilities can be found using lspci command



```
        IOVCap: Migration-, Interrupt Message Number: 000
        IOVCtl: Enable- Migration- Interrupt- MSE- ARIHierarchy-
        IOVSta: Migration-
        Initial VFs: 4, Total VFs: 4, Number of VFs: 0, Function Dependency Link: 00
        VF offset: 2, stride: 2, Device ID: 0d52
        Supported Page Size: 0000003f, System Page Size: 00000001
        Region 0: Memory at 0000203000030000 (64-bit, non-prefetchable)
        Region 4: Memory at 0000203000040000 (64-bit, non-prefetchable)
        VF Migration: offset: 00000000, BIR: 0
Capabilities: [c00 v1] [redacted]
Capabilities: [d00 v1]
Capabilities: [e00 v1] Designated Vendor-Specific: [redacted] Len=56: CXL
        CXLCap: Cache- IO+ Mem+ Mem HW Init+ HDMCount 1 Viral+
        CXLCtl: Cache- IO+ Mem+ Cache SF Cov 0 Cache SF Gran 0 Cache Clean- Viral-
        CXLSta: Viral-
```

  - Memory Latency Checker tool can see CXL node



```
[root@localhost Linux]# ./mlc
Intel(R) Memory Latency Checker - v3.9a
Measuring idle latencies (in ns)..CXL node
                Numa node
Numa node     CPU0  0  CPU11        2
     0          89.3    196.0    489.7
     1         195.9     85.7    756.5
```

QCI’s work through DeltaLake/S9S

- Smbios table porting on new platform, including type 8, 9, 11 and so on.
- Support NVMe Hot-plug.
- Add VPD items to control system behavior.
  - MRC warning promote
    - Control whether system to bypass memory training failure or hang in FSP.
  - Dimm frequency setting
    - Select which dimm frequency limit to be used.
- ACPI BERT table implementation (for MCA bank errors) (ongoing)

9elements work on security and coreboot core code

- Converged security suite: Open source tooling for Intel CBnT (Converged bootguard and TXT)
- CBnT setup and integration in coreboot
- CBnT + coreboot measured boot integration (early TPM)
- MP init on ‘many’ cores (spinlocks, SMM, halt before CBnT, …)

SysPro's XEON_SP contributions

- XEON_SP shared code structure
  - limited code duplication in soc/intel/xeon_sp & soc/intel/common
    - src/intel/xeon_sp: ~5500 lines code + headers
    - src/mainboard/ocp/deltalake: ~900 lines code + headers
    - src/mainboard/ocp/tiogapass: ~600 lines code + headers
- XEON_SP PORTs/STACKs coreboot and ACPI resource allocation
- ACPI - new and updated tables - DMAR, SRAT, SLIT, CEDT
  - multi socket and high count core/thread support
- LinuxBoot & u-root additional loader integration (pxeboot)

Call to Action

- **Join** open system firmware slack: https://slack.osfw.dev/
- **Use** coreboot for OCP DeltaLake server based on Intel Xeon Scalable processor:
  - https://github.com/opencomputeproject/OpenSystemFirmware/tree/master/Wiwynn/deltalake
- **Benefit and Contribute**:
  - OCP OSF for Intel SPR-SP based platform
  - OCP community lab

Open Discussion

OCP GLOBAL SUMMIT

NOVEMBER 9-10, 2021