



OPEN
Compute Project

Proposed Charter For Open Network Project

Revision History

Date	Name	Description
06-25-2013 - 07/26/2013	Peter Krey	V#1.07 initial Charter draft resulting from May 16 th 2013 First Open Compute Engineering Summit Working Group of MIT Stata Center, Cambridge MA, June 18 th 2013 Kick-Off Session, and Series of Calls w/Najam Ahmad + Jon Hoover to Create Initial Project Charter Draft Doc.
07/28/2013	Najam Ahmad	Minor edits to remove some of the place holders and correct spellings
08/26/2013	Jon Hoover	Updating with changes from OCP Engg Workshop
08/27/2013	Jon Hoover	Incorporated remaining feedback from OCP Engg Workshop
09/04/2013	Jon Hoover	V1.12 – final edits

Contents

License	4
Mission Statement	5
Project Scope	5
Organization	8
APPENDICES	9
Appendix 1 - Working Group Marker Board Notes OCP Eng Summit at MIT May 16 2013	10
Appendix 2 - OCP Eng Summit at MIT May-16-2013 – Kick-Off Team Re-Organized Etherpad Notes	11
Appendix 3 – Meeting Cadence	16
Appendix 4 - Focus Areas	17

License

As of April 7, 2011, the following persons or entities have made this Specification available under the Open Web Foundation Final Specification Agreement (OWFa 1.0), which is available at <http://www.openwebfoundation.org/legal/the-owf-1-0-agreements/owfa-1-0>:

Facebook, Inc.

You can review the signed copies of the Open Web Foundation Agreement Version 1.0 for this Specification at <http://opencompute.org/licensing/>, which may also include additional parties to those listed above.

Your use of this Specification may be subject to other third party rights. THIS SPECIFICATION IS PROVIDED "AS IS." The contributors expressly disclaim any warranties (express, implied, or otherwise), including implied warranties of merchantability, non-infringement, fitness for a particular purpose, or title, related to the Specification. The entire risk as to implementing or otherwise using the Specification is assumed by the Specification implementer and user. IN NO EVENT WILL ANY PARTY BE LIABLE TO ANY OTHER PARTY FOR LOST PROFITS OR ANY FORM OF INDIRECT, SPECIAL, INCIDENTAL, OR CONSEQUENTIAL DAMAGES OF ANY CHARACTER FROM ANY CAUSES OF ACTION OF ANY KIND WITH RESPECT TO THIS SPECIFICATION OR ITS GOVERNING AGREEMENT, WHETHER BASED ON BREACH OF CONTRACT, TORT (INCLUDING NEGLIGENCE), OR OTHERWISE, AND WHETHER OR NOT THE OTHER PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Mission Statement

The mission of Open Compute Networking Project is to create a set of networking technologies that are dis-aggregated and fully open allowing for rapid innovation in the network space.

Project Scope

High Level Descriptive Scope

The Open Networking Project is to facilitate & enable new and innovative open networking hardware & software standards, design creations & collaborations, project validation & testing, and OCP Community contributions.

The Open Networking Project is also to bring to networking technologies what has already enabled OCP open servers & storage including:

- Circumnavigate traditional closed & proprietary network switch H/W & S/W via fully open, non-lock-in networking technology stack.
- Fully disaggregated and open networking hardware & software
- Operating System - Linux based operating systems & developer tools, and ReST API's
- Fully automated configuration management & bare metal provisioning
- Universal & Multi-Form Factor Switch motherboard hardware (a la Roadrunner)
- Fully open integration & connectivity
- Energy efficient power & cooling designs
- Software Defined Networking (SDN)

Initial open switch projects are to target top-of-rack form factors, port counts, etc., while later projects are to also target spine switches & potentially other switch types.

In-Scope Technology Categories

The initial "in-scope" coverage of the Open Networking Project is described by the following layers / categories of networking technologies:

Category Level	Description
Level "3" (initially out of scope)	Developer Tools, Management Tools, ReST API's, SDN
Level "2"	Device Drivers, Boot Loader, Bare Metal Provisioning, Firmware, BIOS, Open Linux OS's
Level "1"	Universal Form Factor "Common Motherboard" H/W Standards Based H/W, Frame Processor (optional) Interconnects & Integration, Standards Based Cabling OpenRack + 19" Form Factors Energy Efficient Power Supplies & Cooling

Later phases of the project may also include exploration of a Network ASIC Chip Open Interface Layer, as a conceptual analog of how OpenCL enables a multi-language, common developers API for both multiple GPU device types and FPGAs, LibVirt & Vagrant enable a multi-language, common developers API's for multiple hypervisors, ..., etc.

Out-of-Scope Technology Categories

This project does not intend to develop areas including protocol stacks & virtualization, nor dictate network architectures & topologies. Other areas that are initially out of scope include hardware abstraction layer, deep packet inspection, hardware based security, firewall feature sets and load balancing. The project does, however, expect the results of its initiatives and contributions to support a wide variety of technologies and protocols.

Key Project Focus Areas

Per the collaborative team efforts at the First Open Compute Engineering Summit at MIT Stata Center on May 16th 2013, the following key focus areas were captured and summarized by PK per the working groups collaborative discussions & debates. (Original event content marker board photo in Appendix 1).

Open Switch Hardware

- Silicon Examples: Broadcom, Intel / Fulcrum, Marvell, Netronome, Mellanox
- Speeds, Feeds, and Environmentals
- Frame Processor and Add-In Board Interface
- Switch Control Processor / CPU: X86, ARM
- Operating System Agnostic
- Power Supplies: Input voltage, Watts, Efficiency ... OpenRack and 19" standard enterprise
- Universal Motherboard Form Factor: OpenRack and 19" standard enterprise

Form Factors

- Leaf / Top of Rack
- Spine
- Non-bladed
- N 'U' Universal Form Factor – OpenRack & 19" Traditional Enterprise
- Cooling

Power Supplies

- Input
- Watts
- Efficiency
- N 'U' Universal Form Factor – OpenRack & 19" Traditional Enterprise

Project Mission

- Open Switch Hardware Definition
 - Leaf / Top of Rack
 - Spine
- Open Switch Software Definition
 - Operating System Independent (Linux Based)
 - Open Boot Loader
 - Bare Metal Provisioning (PXE, Uboot, ONIE, ..., others)
 - Network ASIC Chip Open Interface “Driver” - “OpenCL Analog”
 - OCP Open Hardware Management Analog

Key Drivers

- Common, Standard form factors
- Capex & Opex reduction
- Tier1 & Tier2 Multi-Vendor Implementation, Standards Based
- Interop C&I Tests and metrics ... similar to OCP Server C&I Panel
- 100% open source core, low-layer
- 100% standards based hardware implementation
- Analogous to BigData, NoSQL, ..., etc . (Independent) Software & Hardware layers

Potential Futures

- Switch Hypervisor versus Linux Containers / LXC ?

Organization

The following is modeled after creation & publication of the OCP AMD 3.0 Roadrunner Server Project, though may be modified after a few iterations to best move the project forward.

1. **Project Co-Chairs** – facilitate the flow of information, determine consensus, define scope, commit documents, ..., etc.

Najam Ahmad / Facebook

Peter Krey / Krey Associates, Inc. *(PK pending Fidelity approval)*

2. **Program Management & Communications** – Jon Hoover / Facebook
3. **Core Working Group** – initial founding members committed moving the project forward between meetings. Contribution examples include guidance & advice, specification document feedback & contributions, code contributions, etc.

Najam Ahmad, Jimmy Williams, & Jon Hoover / *Facebook*

Peter Krey, Keith Shinn, and Ed Finn / *Fidelity*

Matthew Liste, Michael Valentine, and Marc Woolward / *Goldman Sachs*

Dan Pitt & TBD / *Open Networking Foundation*

Others TBD

4. **Expanded Working Group** – additional cross-industry community contributing members committed to moving the project forward between meetings. Contribution examples include guidance & advice, specification document feedback & contributions, code contributions, etc.

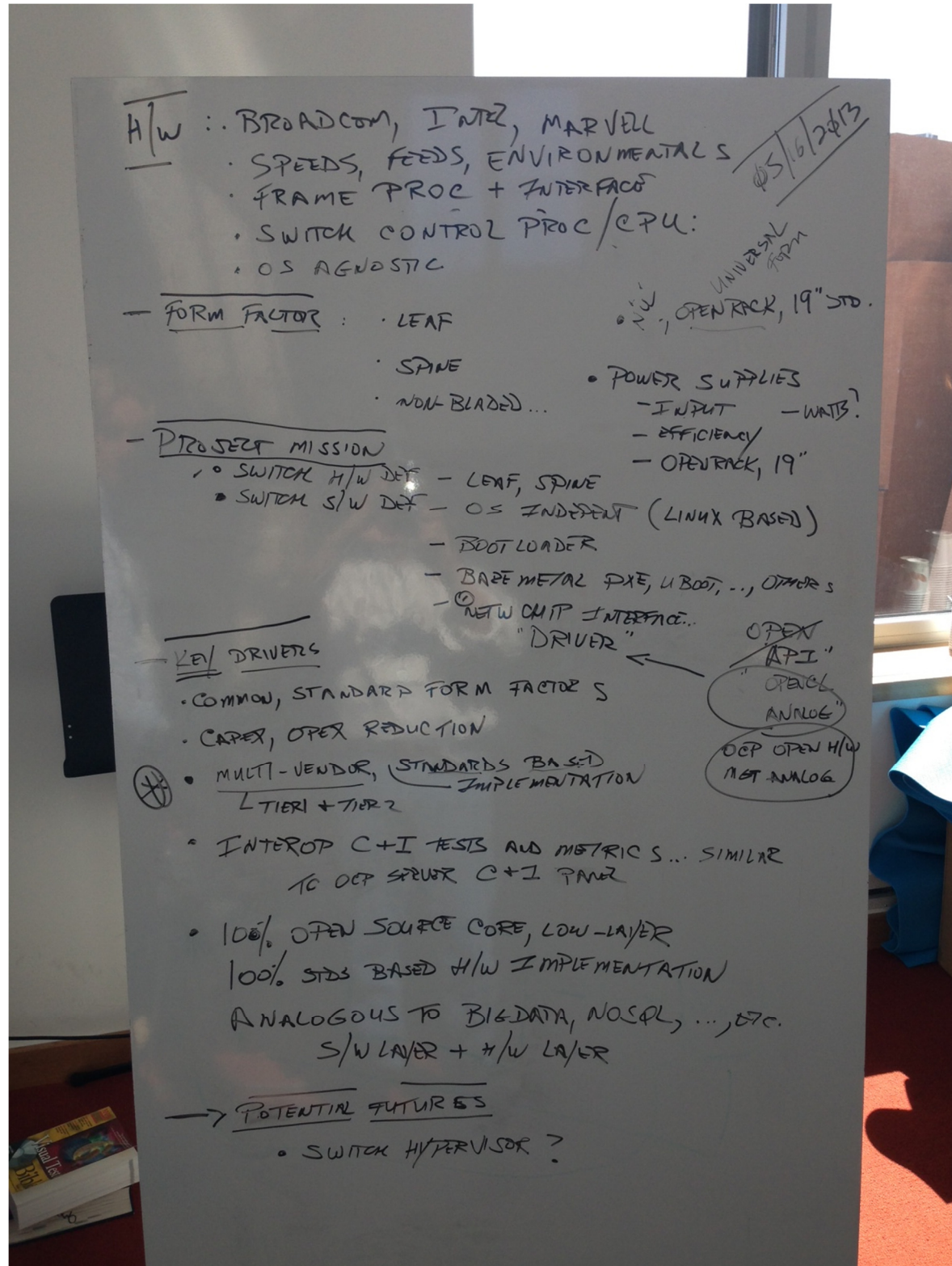
The Expanded Working Group will be organized & modeled after the AMD 3.0 “Roadrunner” and Intel Decathlete working groups expanded Enterprise Infrastructure colleagues.

Also to include additional Web 2.0 and Hyperscale colleagues.

5. **Advisory** – Monthly Advisory meetings, key topics discussions, design & development topics
6. **Open Compute Formal Events** – As per OCP schedule.

APPENDICES

Appendix 1 - Working Group Marker Board Notes
 OCP Eng Summit at MIT May 16 2013



Appendix 2 - OCP Eng Summit at MIT May-16-2013 – Kick-Off Team Re-Organized Etherpad Notes

The following is a re-organized version of the Appendix 2 reformatted & typo corrected version of the “raw” Etherpad content contributions created by the attendees of the First Open Compute Engineering Summit at MIT Stata Center on May 16th 2013.

While some content as been re-categorized & re-arranged, no contributed content has been changed and/or altered.

This appendix reflects changes agreed by Najam Ahmad, Jon Hoover, and Peter Krey resulting from team call on 07/11/2103/

OPENPACKET CHARTER :

Summary:

The Open Compute Project's focus is on server, storage, network, interoperability?? and data center design. The Open Packet project is an OCP foundation technology vertical allowing for continuous network innovation that will lower the TCO and raise the ROI of data center networking technologies through a combination of hardware and software. By leveraging industry leading technologies, Open Packet will address technology requirements for the entire range of data center networking software and hardware solutions. (this last sentence will require a statement on the relationship of this forum vs ODL, ONF, etc)

Mission Statement:

When open design of networking hardware, and software move in concert they can improve efficiency, reduce power consumption, and allow for flexible and modular specifications. define open h/w spec at network switch level along with baseline software feature set

Project Naming:

We might need a new project name: OpenPacket already taken....
'The mission of OpenPacket.org is to provide quality network traffic traces to researchers, analysts, and other members of the digital security community.'

There is no end here, so lets vote: please add a + sign to the one you like best

OpenFabric
OpenSesame (cute but many google hits)
OpenArtery
OpenStructure
OpenNexus
OpenRoute
openroad
opengate
opentube ...
OpenForward(ing)
OpenCommunicate

... continued ...

OpenConductor (our friends at Cisco have a Conductor product for TelePresence)
OpenDirector
OpenGress (ingress/egress)
OpenPacketSwitch
OpenPSwitch ++
OpenPulse
NextNet
ONN-OpenNextNet

Themes:

Per 07/11/2013 Kick-Off Team Call (Najam Ahmad, Jon Hoover, and Peter Krey), the original contributor Themes were re-organized into the following 2 categories:

R = "Requirements / "Must Do" **D** = "Design Discussion Topic / Q+A"

'SOLVE FOR TODAY, BUT DON'T LIMIT YOURSELF TO TODAY'S TECHNOLOGY' - Cole Crawford

D - How is an open source reference design for a switch different from a reference design from my favorite chip vendor?

D - Switch reference designs are usually a total package (Broadcom is all Broadcom, Intel is all Intel: What about Intel with Broadcom in the same box? (Management SoC versus switch ASIC?)

R - The focus should be disaggregation of the network OS from the TOR switch - similar to the server

R - The Open Switch spec for HW should accelerate the disaggregation of the network OS

R - Enable commoditization of the Leaf and Spine Switches closest to the servers

D - Is there an underlying assumption that there is a standard switching chip SDK that agnostic of the SDK underneath?

D - And do we define all the capabilities of this SDK? Where does our definition end, ie how far up the network control stack does OpenPacket go? (leads to the comment below)

D + R - Intersection of this effort with ONF/ODP/ONRC and other open source projects

D - Processor vendors are moving to proprietary interconnects. High performance/low latency/low power networking within the box and between boxes will need to leverage these interconnects. How do we do this as an open source hardware/software community?

D - Increase network capabilities, flexibility and agility
Application awareness/consideration
Allow migration of "network provisioning definition" to move not only between vendor equipment but between different HW at top of rack, endpoints, and or even full SW implementations.

D - Not just open hardware ... open Linux distros ... must enable access to maximum scope of "up to date" developer tools ... enable enterprise cloud automation ... and (control) processor independence.

R - Pick computationally appropriate Linux supporting processors ... allow applications to be written & hosted on the switch ...

R - Create a "universal" form factor ... switch "motherboards" must fit in both OpenRack AND traditional 19" enterprise & co-lo racks ...

D - Can we see some examples of datacenter networks being built by Facebook, etc so that we can study concrete examples that extend past the top of rack ?

D - Power architecture spec (Freescale SOC)

D - means of mapping a virtual network topology to a physical network topology for network security purposes

D - How do we stimulate creative software applications that drive the use of SDN ?

D - What level of virtualization do we want in HW/SW for the Network Switch?

D (need to clarify) - Do we consider satellite architectures, or do we just punt that problem to SDN? (clarification?)

D - Standardization of the Hardware/Device interface/SDK for Network Processors/ASICs?

D - What modularity do we want in the HW? Be able to have multiple tiers to run more or less apps in the switch? Do we want modularity to provide flexibility to easily configure HW for different port and cable types?

D (need to clarify) - Multiple layers of modular uniform interconnect switching between top of rack and outside network.

D - Should we (allow/require?) have sniffing capabilities in switch design such that we can dump packets to some form of storage either internal or external for playback with tools like Wireshark?

Scope

Data-Center - not so much defining the backbone

Initially allow 'existing' top of rack with SW vswitch, and then later architectures

But this could also have applications on backbone edge if executed properly

Top of Rack - Aggregator, evolve to Backbone switch?

Top of Rack down to node level interconnect (all elements inside data center)

Priorities: Top of Rack to begin with? Or parallel efforts on other aspects of networking

Disaggregation assumed?

Top of Rack down to virtual machines on compute/storage nodes

Uniform, reliable, low latency, low power, hardware terminated fabric

Network OS- across all devices-

Layer? 1/2/3/4?

What is the typical cluster size (how many physical node, CPU, storage node)?

Assume typical data center inter-compute/storage node distance (i.e., long reach, short reach)?

Scope should not be limited to speeds and feeds, because it depends on the protocol riding on these speeds and feeds and may limit the lifespan of this spec.

Need to divide into multiple phases: short term based on existing technology, medium term based on product in development, longer term based on new product development that leverages the output of this group's work.

Goals

TIMEFRAME?

Required Speeds and Feeds

Reliability // High Availability ??

Fault tolerance

Port count

Node count

Predictable/engineerable flow control/quality of service

Building block for of a multicomponent system to provide reliability

ie: Smaller, simpler fault units like 1RU, 2RU that create a very large network

Low, deterministic transit latency

Hardware Requirements

Decouple Hardware forwarding plane from control plane in hardware:

ie: x86 or Freescale in combination with any forwarding hardware (Fulcrum, Broadcom, Cavium, etc)

ie: ARM or x86 in combination with any forwarding hardware (Fulcrum, IDT, Broadcom, Cavium, etc.) and any open networking standards (e.g., Serial RapidIO, PCIe, Ethernet, ...)

Following 2 FPGA section were combined ...

Allow Non ASIC-based (i.e. *Xilinx* Virtex / *Altera* Stratix) forwarding / user plane?

Low latency interconnect with minimal processing overhead

We need to define a forwarding lookup scheme that is not tied to current packet formats to ensure flexibility for future technologies. Openflow tried this, but HW didn't support. New openflow is more closely tied to current H/W switching capabilities. (Hence FPGA assist :)

H/W - Direct high speed fabric interconnect on processors for direct connect to fabric infrastructure

H/W - Low latency between nodes to support high density east west

H/W vs S/W ? Graceful recovery from faults in the network at microsecond scale

Identify the reach targets and rates of the required interfaces (before picking the implementation technology) +1 performance targets

PoE? (Related: are electrical interfaces expected?)

PK – is 1588 PTP “Precision Time Protocol ? http://en.wikipedia.org/wiki/Precision_Time_Protocol
1588 ptp support for industrial/AV applications and to measure the performance of the network

Factor in cabling and transceiver cost into analysis of form factor/packaging decisions since it is becoming a dominant portion of the cost of interconnecting servers in data centers. This may, along with shrinking server form factors, new kinds of chassis-based systems

Forwarding Port Density Options (This should address both):

Low port density (rack, disaggregated, etc.)

High port density (spine switch option)

Internal Topology and how we create that topology (fiber connections, etc.)

how to accommodate different port types (RJ45, SFP+, QSFP+, etc.) and cable types (transceiver/fiber, passive/active copper, AOC)?
for top-of-rack switch prob different port requirements for in-rack and out-of-rack port types

Separate Spec:

1. Physical (power...)
2. Port configurations (Disaggregated + out of band ports for controllers/ management)
3. Device OS / BMC API ...

flexible interconnects- fiber/copper/ whatever is next

Hardware Requirements for Virtualization Support, Assume Multiple OS on Switch?

Divide hardware requirements:

Software Requirements

Added & created to separate Raw Etherpad software requirements contributions from above "Hardware Requirements" section contributions:

S/W - Vendor neutral APIs for L2/L3 etc forwarding, OpenFlow is high-level example, desire lower level fast interface.

S/W - Open flow can also be defined as low-level on the functions that are taking place

Software vendor-agnostic OS installer/loader:

Cumulus ONIE (Open Network Install Environment?)

will the switch form factor be sized to fit in an OpenU rack? will power be supplied from the power shelf and be distributed to the switch via the 3 bus bars at the back?

Appendix 3 – Meeting Cadence

“Meeting Cadence” contents will be discussed & expanded during the upcoming Aug 13th 2013 OCP Engineering Workshop. The following is a place-holder for a future document version.

The formal meetings will have the following meeting schedules:

- Working Group*** Will meet as needed between other project formal meetings. A notice of any meeting /conference call will be sent to the general list for anyone interested. In addition, the Working Group is responsible for coordinating with other OCP tracks/projects to insure uniform implementation and clarity when there is overlap.
- General Assemblies*** Will be co-terminus with the Open Compute Summits. These meeting will be for a wider audience with update on the past efforts and anticipated progress.
- Advisory*** Will be take place approximately every month and will discuss the progress made, open issues and anticipated progress. These calls are intended provide direction/focus of efforts and approve any new projects.

It is anticipated the Sub-Project meeting cadences will follow this pattern although Sub-Projects may decide on different cadences based on their requirements.

Appendix 4 - Focus Areas

“Focus Areas” contents will be discussed during the upcoming Aug 13th 2013 OCP Engineering Workshop. The following is a place-holder for a future document version.

<i>Sub-Project</i>	<i>Description</i>
Environmental Design	Mechanicals, Cooling, Power, Switch Management, Management Processor, Boot to Operating System
Switch Design	Switch Motherboard, Frame Processor, Optical Interconnects