Open Compute Project

Open Rack Charter

Authors: Gio Coglitore, Pierluigi Sarti, John Kenevey, Pete Bratach

OPEN Compute Project

Contents

Contents 2
Mission Statement 3
Summary 3
Overview 4
License 4
Design Approach 5
Physical Design 5
Power Distribution (Bus Bars) 6
Battery Backup Power 8
Cable Management 8
Compute Density and Thermal Design 8
Serviceability 9
Installation Flexibility 9
Component Upgrade 9
Roadmap 9
Community Organization and Meeting Cadence 10

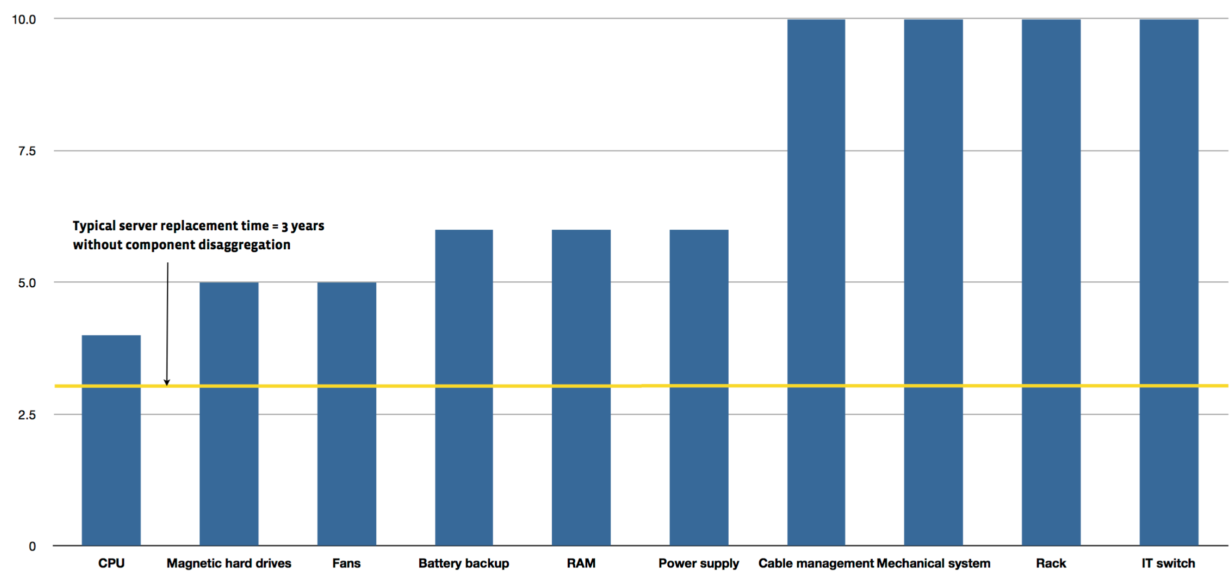
Mission Statement

Reducing TCO, improving component refresh cycles and increasing flexibility in the scale compute space through simple, innovative and modular designs; an open form factor that still allows for value-added differentiation at the component layer; a standardized rack infrastructure with common interconnects and power distribution; and optimized supply chain management.

Summary

The Open Compute Project's focus is on server, storage and data center design. The Open Rack design is a foundational technology for other Open Compute designs that aims to diminish or eliminate the traditional differentiation at the mechanical rack and power distribution layer. Open Rack is an open design that focuses on enabling standard mechanical form factors, modular power supplies, improved cable management, efficient thermal design, and flexible serviceability allowing fluid component upgrades and independent refresh cycles.

The intent is to enable what we term component disaggregation. Disaggregated compute components can be swapped and expire according to their own life cycles without replacing the entire server for an upgrade or a repair. The open form factor offers component choice without locking in to a specific vendor. This improves total cost of ownership and allows needed flexibility in the scale compute space.



Typical server component life cycles vs. the typical server replacement time.

Overview

The Open Rack can be thought of in the following areas:

- Mechanical design
- Power distribution
- Battery backup power
- Cable management
- Compute density and thermal design
- Serviceability
- Installation flexibility
- Component upgrades

The problems the Open Rack seeks to solve encompass:

- Different mechanical form factors require different implementation schemes and force vendor lock-in without associated benefit, as typified by blade servers.
- Redundant power distribution integrated into and specific to each server.
- Outdated cable management with proprietary interconnects.
- Impractical compute density, restricted by thermal design.
- Inefficient serviceability.
- Inflexible installation processes.
- Inefficient component upgrade cycles.

WHETHER OR NOT THE OTHER PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Design Approach

Using the mission statement as a guide, our approach to deliver and maintain projects within this track is:

- Any solution must scale to 100,000+ physical machines.
- Use existing standards whenever possible. If standards require adjustment, partner with standards bodies to make changes.
- Define requisite features, attributes and interface standards for OCP hardware designs.
- Encourage and work with the supplier ecosystem across the technology stacks that Open Rack relies upon.
- Maintain interoperability for equipment used in the Open Rack.
- Work with OCP hardware designers to implement required functionality.
- As part of every deliverable, a methodology to validate the functionality and maintain validity thereof must be included.
- As it is common to have some non-scale platforms in a scale environment, encourage specialty hardware manufacturers to comply with the Open Rack specification.

Physical Design

The Open Rack design should remove unnecessary differentiation at the mechanical design level, eliminating customer lock-in to a particular supplier. Customers and suppliers alike can configure their own designs for each power zone.

The Open Rack form factor features:

- 23.6" wide rack, as floor steps in most data centers are 24". This provides flexibility for the chassis form factor within a 21" wide equipment bay. Racks are traditionally 19" wide.
- "Practical density." We have designed a larger aperture (the opening in the front) to improve airflow and create a good thermal profile. This allows for higher compute density.
- 36-48" rack depth makes serviceability more practical and thermal flows more efficient. Normal equipment depth is 33", with the remainder used for bus bars and cables.
- Up to 48U rack height maximize the density allowed while also allowing deployment into most data centers. The accepted rack height is 95".
- At least 1 power zone per column. A power zone comprises the power shelf solution and the equipment bay that it powers. Power solution can be single-phase AC or three-phase AC, or HVDC. The width of the power shelf can be as wide as the equipment bay itself. The height is measured in 0.5U increments.
- The Open Rack increases compute density, demonstrates flexibility in design, improving efficiency and lowering cost.

- Maximum of three 1U rack switches, allowing iteration on the compute density reinforced by one or more power zones.
- Compatible with all form factors and provides interoperability between different suppliers.
- A modular chassis that accommodates multiple sleds, so there can be multiple motherboards in a given chassis in either horizontal or vertical layouts; the chassis can be of varying height and can be pre-integrated into the rack.
- A universally flexible combination mounting scheme for the equipment bay that provides for fixed or adjustable mounting options in a variety of configurations in 0.5U increments.
- A slot in the equipment bay is measured in 0.5OU steps. An OpenU (OU) is 48mm.

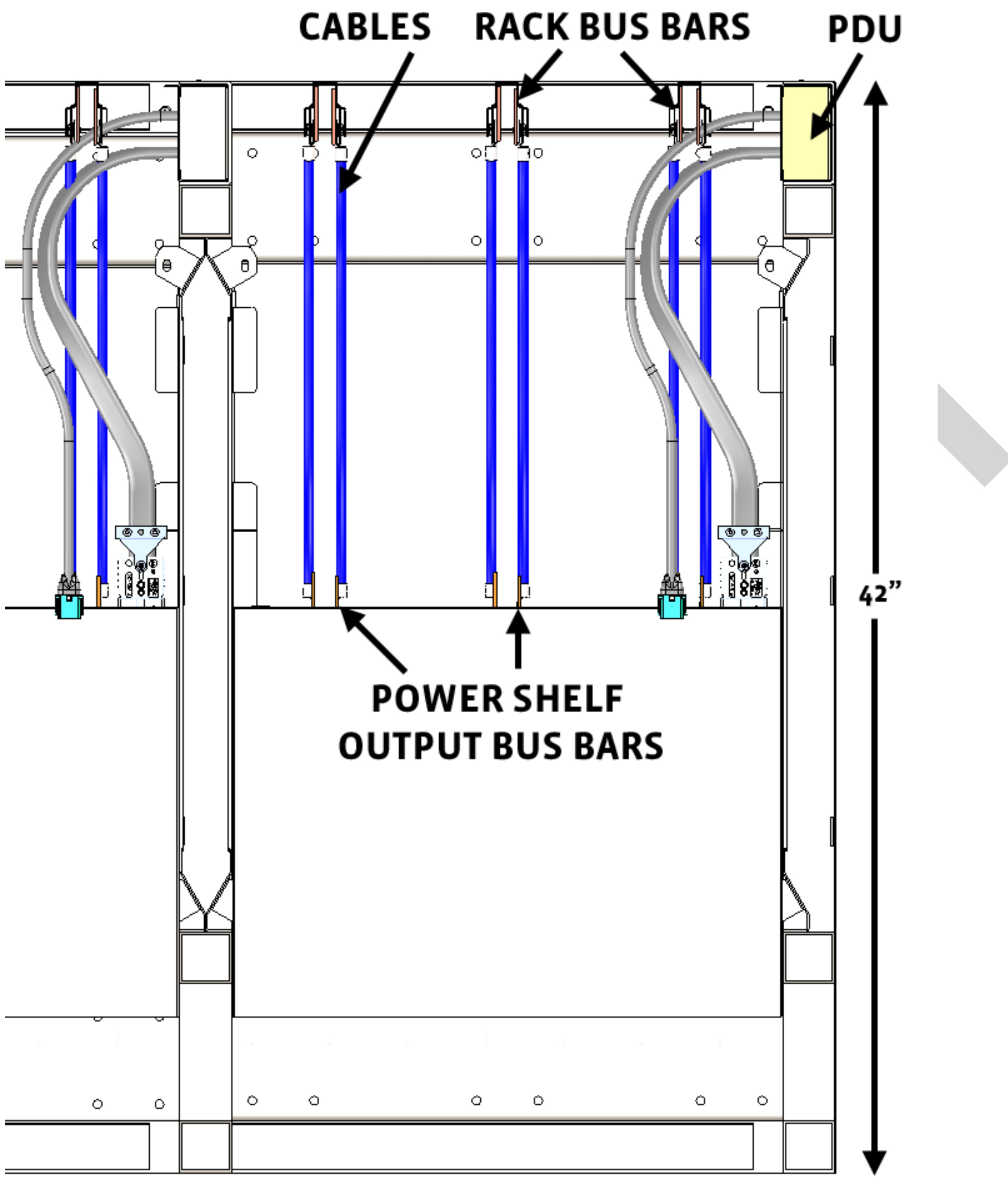
Power Distribution (Bus Bars)

Traditionally, power distribution is not modular but specific to a particular compute system. The Open Rack's power is not part of the server unit but is designed into the rack, putting the focus on efficiency.

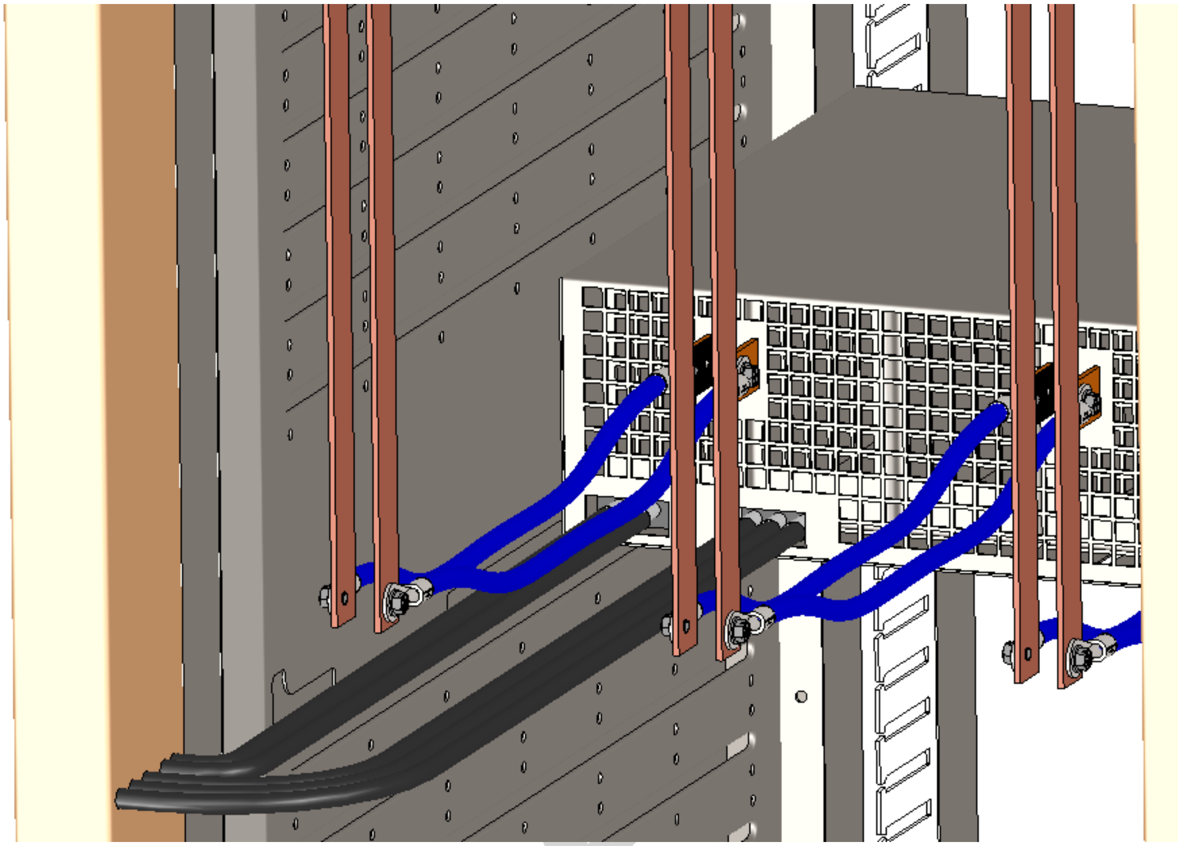
The power distribution design enables redundant power schemes and also distributes more power to a given device than a single PSU. This provides flexibility in compute designs and flexibility to reuse the PSU and adjust the density profile.

The bus bars are designed for safety, efficiency and tolerance. They incorporate these features:

- Simplicity of design
- Improved airflow
- Seamless mating of power solution and compute unit
- Improved efficiency for 12V distribution via multiple bus bars and multiple power zones
- The 12V "secondary" voltage provided by the bus bar is below the 60V SELV threshold (see http://www.electropedia.org/iev/iev.nsf/display?openform&ievref=851-15-08 and http://en.wikipedia.org/wiki/Extra-low_voltage) and allows the design to bypass regulatory oversight (which begins at 90V)
- Very tight (constrained) mechanical tolerances enable mating with the compute units
- Up to 3 bus bar pairs provide ability to install 3 motherboards per slot
- Eliminates bulky power cords that hinder airflow and are another point of failure themselves
- Enables redundancy without significant added CapEx
- Optimizes efficiency of power distribution through higher utilization



Overhead view of bus bars with vertical PDU.



Power shelf connection to vertical bus bars (can be either cables or blades). Example shows single-phase power solution.

Battery Backup Power

The Open Rack can use a flexible battery backup strategy in lieu of a centralized UPS. It may be:

- BBU within the rack or
- OCP battery cabinet as a local backup source
- 48Vdc from a central backup line or HVDC scheme

Cable Management

In consideration of cable management, there are some next generation interconnects that make it better in the data center from management, interconnect-agnostic and bandwidth points of view.

Compute Density and Thermal Design

The Open Rack enables compute density that takes into account a practical thermal design and efficiency. It is not focused on aperture reduction (that is, maximizing the airflow over each server). Its form factor allows for large number of systems in the rack without cost of thermal restrictions.

Serviceability

The Open Rack should be tool-less with easy-to-remove components that require the least amount of effort to remove.

Installation Flexibility

The rack can be delivered fully integrated and needing only to be bolted into the data center. Or the rack can be ready to receive compute components during an infrastructure build-out at the appropriate time (and cost), following a data center's build-to-deliver cycle.

Component Upgrade

The Open Rack modular design enables independent refresh cycles for all components.

V1 of this specification enables independent server and storage device upgrades without modifying the power distribution scheme. We envision v2 of the spec to enable disaggregation of other components such as network interface, DRAM, and boot devices, as well as include a BBU solution in the rack.

Roadmap

Ownership of non-recurring engineering (NRE) funds, manufacturing, validation and test plans will vary based on market demand and supplier interest.

Version	Notes	Validation
1.0	Scheduled for 4Q2012, it features a 4.2KW power shelf for an OCP server with an OCP Intel v2.0 motherboard.	Facebook will perform a full evaluation for production readiness.
1.5	This mid-point release, scheduled for late 2012, features greater than 5KW per power shelf, wide-range rectifier schemes that can take an estimated 120 to 240VDC, 1U with 1U lithium polymer battery underneath, and enables HPC and other markets to meet their business requirements (such as better cost structures).	The community's evaluation is welcome.
2	Planned for 2013. Focus is on superior design for greater efficiency in the scale compute space and other features. The form factor can accommodate a high-voltage DC feed, still with delivery to low voltage DC distribution system within the rack.	Facebook will perform a full evaluation for production readiness.

Community Organization and Meeting Cadence

The Open Rack Project Committee chair is Gio Coglitore, Facebook. The committee also includes:

- Andy Bechtolsheim, Arista Networks
- Stephen Chan, Facebook
- Shesha Krishnapura, Intel
- Pierluigi Sarti, Facebook

The committee operates under these rules:

- The project committee chair runs regular meetings and ensures minutes are recorded and shared with the community via the mailing alias. (The chair can elect a secretary from the project committee to record minutes.)
- Oral contributions fall under the OWF CLA and OWFa definition (see http://www.openwebfoundation.org/legal/the-owf-1-0-agreements/owf-contributor-license-agreement-1-0---copyright-and-patent) and can equate to actual contributions. They must be put in writing in order to be considered as a contribution to a specification.
- Minutes should be published to the project mailing alias 10 days after the meeting.