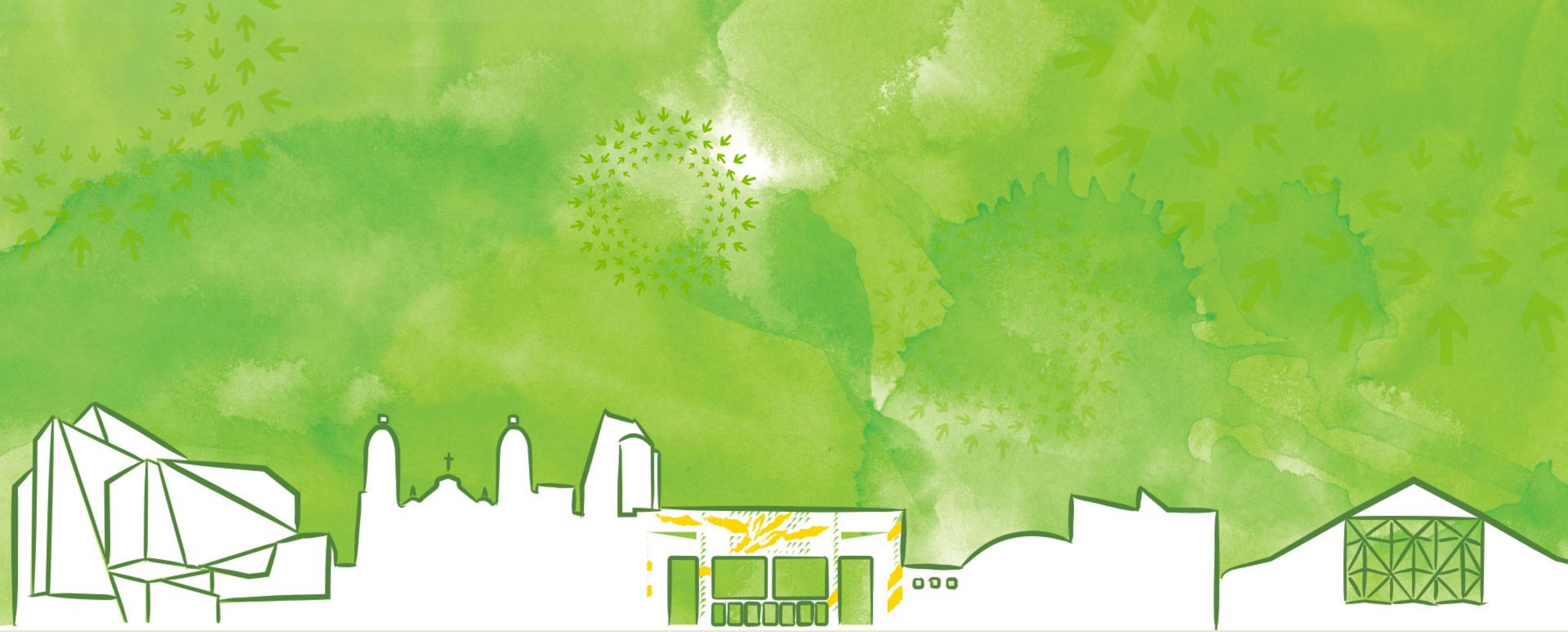


OCP Engineering Workshop **10 August 2016**| Durham, NH

CG-Open Rack-19

Specification Contribution

Andrew Alleman & Matt St. Peter

CTO

Contribution Overview

TD MP SC EP PS

- Radisys is contributing the Carrier-grade Open Rack concept to OCP in the form of a Rack + Sled interop specification
- Radisys also has a commercially available (and deployed) product family compliant with this specification, DCEngine, and there are multiple other vendors in various stages of development and production
- Collateral for the contribution includes the Specification, Installation Guide (for DCEngine), Product SKUs (for DCEngine), and supporting/marketing material

Recap: Concept & Philosophy

TD MP SC EP PS

- Leverage Open Compute and Open Rack as a base model and track continued innovations that take place over time
- Update and tune for telecom-oriented central offices, data centers & equipment practices
- Do not constrain with rigid backwards compatibility, but allow easy adaptation of existing COTS elements (OCP-contributed and others)
- Use agile-style approach to create deployable system instantiations and iterate quickly
- Collaborate tightly with customers all along the way
- Work closely with interested partners
- Create something useful, share, evolve, and expand

Recap: Broad Requirements

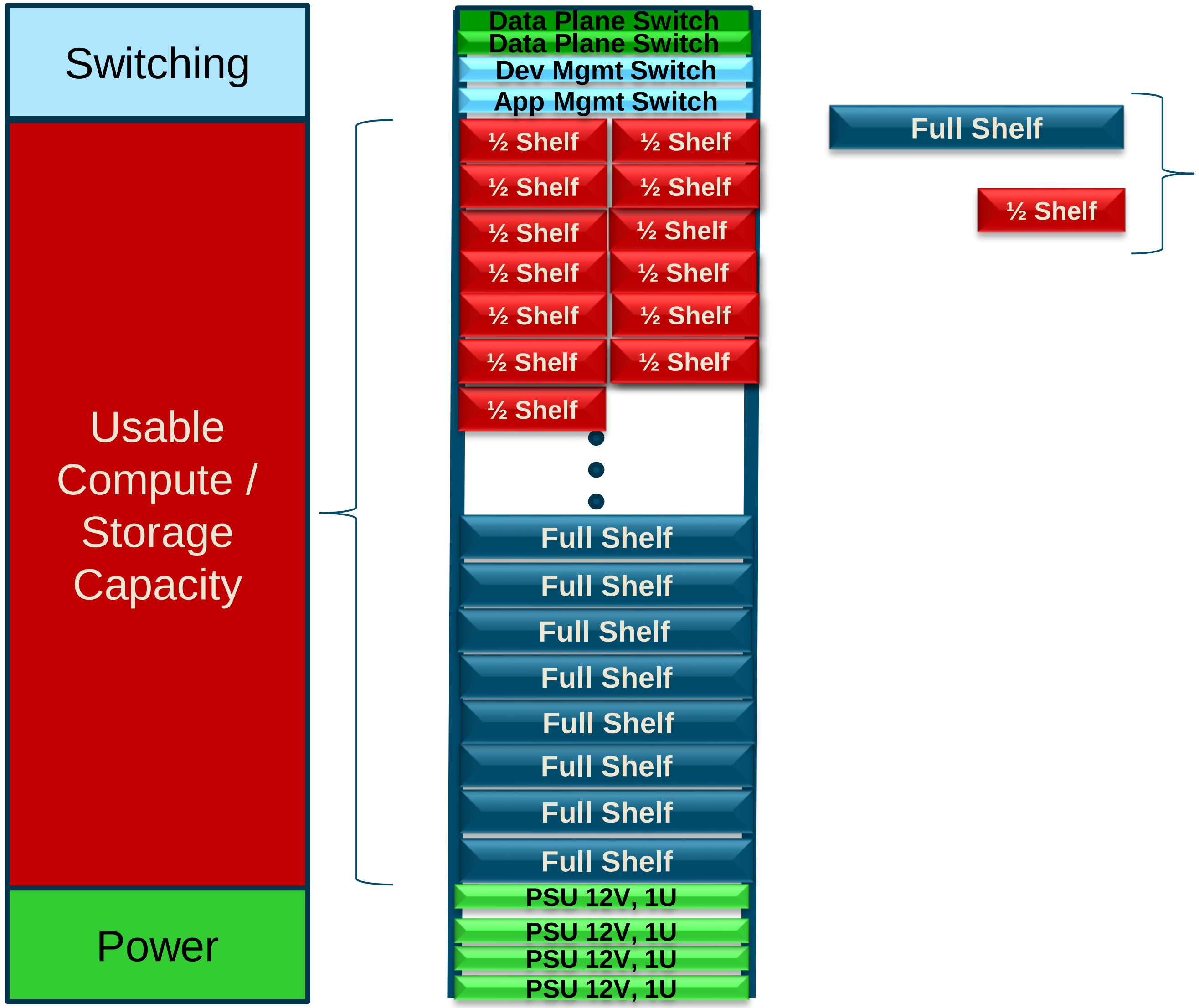
radisys

TD MP SC EP PS

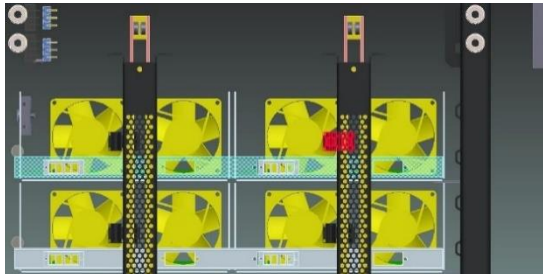
- **Physical**
  - Suitable for CO retrofit and new telco data center environments with standard and readily available dimensions
  - 19” rack width with 450mm equipment aperture
  - Standard “RU” spacing
  - 1000 to 1200mm cabinet depth, supporting GR-3160 floor spacing dimensions
- **Content/workload**
  - Heterogeneous compute and storage servers
  - Range of brawny and wimpy processing elements and storage technologies
  - Half and full rack width sleds, xRU high (2U typical)
- **Management**
  - Ethernet based OOB management network connecting all nodes via a TOR management switch
  - Optional rack level platform manager
- **Networking/Interconnect**
  - One or more Ethernet TOR networking switches for I/O aggregation to nodes
  - Fiber cables, cabled from the rear, blind-mate with flexible interconnect mapping. Front cabled solutions also acceptable.
  - Environment & Power Seismic & acoustic CO environmental requirements applicable
  - Safety and other certification standards also applicable
  - NEBS optional (L1/L3)
  - AC (120-250V) or DC (-48V) power to the rack.
  - Wide range of power consumption per rack – scalability required (from 5 to 35kW); typical deployments at 7-15kW.
  - Dual 12V (nominal) DC power bus bar for nodes; other options possible in future

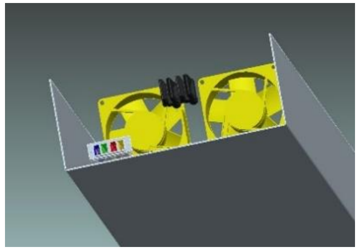
Recap: CG-OpenRack-19 Block Diagram

TD MP SC EP PS

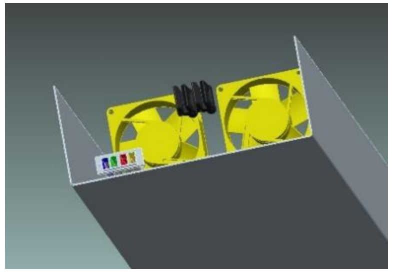


Standard 19" Rack





Vertical 12VDC bus bar in frame mates with power connector located on sled



4 x optical fiber ports via blind mate rear connector to sled

OCP Principals

1. Efficiency

TD MP SC EP PS

radisys

A key OCP tenant is efficient design. Aspects that can be considered include (but aren't limited to) power delivery and conversion, thermal efficiency, platform performance (per-W for example), reducing overall infrastructure costs, reducing code weight, reducing latencies and more.

- **Combined rack-level power conversion reduces overall conversion losses, combines redundancy components, and isolates source power changes to a single location.**
- **Localized cooling (per-sled thermal management) allows cooling only where and how much it is needed.**
- **Airflow impedance of one sled does not affect another, so there is no minimum impedance per sled, reducing the overall power consumed for cooling.**
- **Fan aggregation over multiple servers improves efficiency and airflow while reducing acoustic noise.**
- **Simplified thermal management reduces system complexity and coding overhead.**
- **Direct server-to-switch Ethernet connection minimizes latency. Built-in optical interfaces allow further latency reduction with higher speed dataplane link speeds.**

2. Scale

TD MP SC EP PS

radisys

OCP contributions must be scale-out ready. This means that the technology is designed with the right supporting features to allow for its maintenance in large scale deployments. This includes physical maintenance, remote management, upgradability, error reporting and appropriate documentation. Management tools should strive to adhere to the guideline provided by the OCP Hardware Management Project. Documentation should enable adopters towards a successful deployment, providing guidance on equipment installation, turn on and configuration, as well as physical and remote service.

- **Modular sled design is expandable from one to over 80 servers per rack with no system-level changes. Sled capacity, power capacity, and network size and capacity are entirely scalable to fill a rack or go beyond a single rack.**
- **Maintenance is simplified by local or network level component management, including component cataloging, monitoring, and event generation. All device management is based on open standard software.**
- **Single modules can be replaced in less than a minute.**
- **Alignment with the OCP Hardware Management Project implementation for BMCs over IPMI.**

3. Openness

TD MP SC EP PS

radisys

OCP encourages as much open source contribution as possible, but understands that in certain cases 100% open source contribution may not be possible. Whether fully open source or not, a contribution should strive to comply with a set of already existing open interfaces, at the very least provide one. Providing a solution compatible with already existing OCP contributions is one way to implement existing (open) interfaces.

- **Uses standard half-SSI form factor servers.**
- **Rack-level power distribution, modularity, and organization has many features in common with standard OCP architecture.**
- **The use of a 19” rack aperture allows the use of standard COTS switching and power conversion/distribution products. This includes open white-box switches.**
- **Additionally, servers use a standard half-width form factor that allows the use of existing products.**
- **Uses open source management software.**

4. Impact

TD MP SC EP PS

radisys

New OCP contributions must create meaningful positive impact within the OCP ecosystem. This can be attained by introducing efficiency gains, introducing new technologies and products that are valuable for scale out computing, creating a multiplier effect by building on top of already existing OCP solutions, and enabling a more robust supply chain by contributing alternative compatible solutions.

- **The rack and module construction provides a very high level of modularity and simplicity. This in turn greatly reduces system setup time and operator costs, while providing a very high level of performance per watt and performance per rack.**
- **The system architecture construction allows customers to meet specific requirements related to agency requirements, environments, and compatibility with existing site layout. These include RF emissions, acoustic noise, NEBS, seismic, and test suites such as NEBS.**

The Spec

Contents

TD MP SC EP PS

radisys

Overview

TD MP SC EP PS

Sled Physical Specifications

radisys

TD MP SC EP PS

Figure 2. **Sled Vertical Height (front view)**

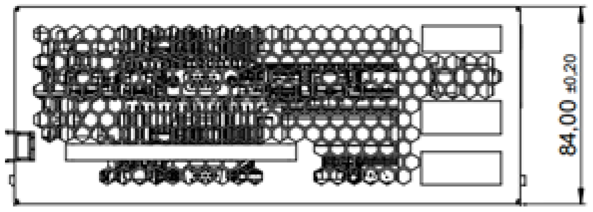


Figure 9. **Half-width sled horizontal width (front view)**

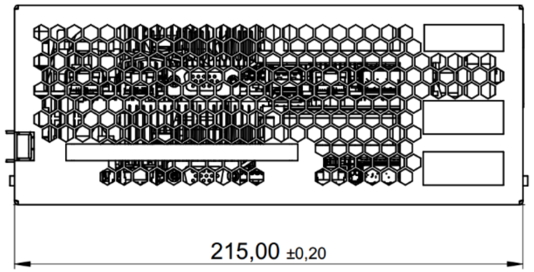


Figure 7. **Sled optical connector vertical position (rear view)**

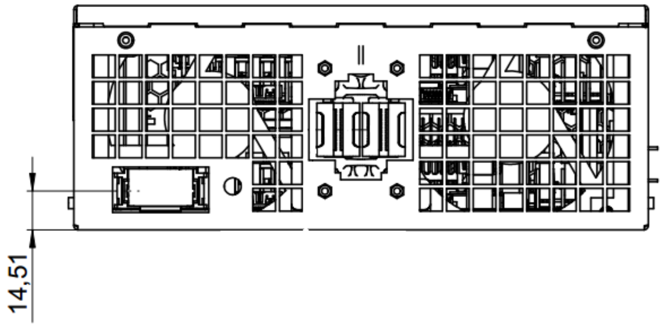


Figure 11. **Full-width sled horizontal width (front view)**

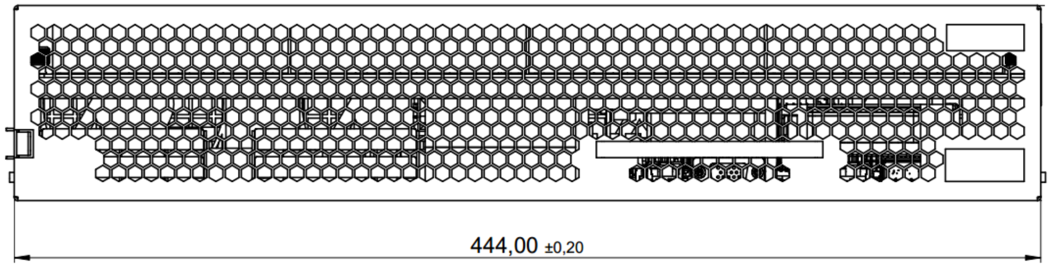
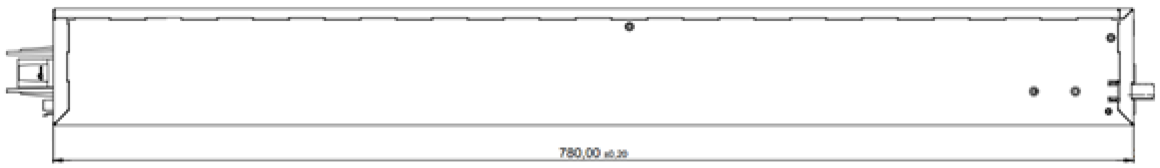


Figure 21. **Sled Operator Panel Art and Dimensions**

Figure 3. **Sled Horizontal Depth (side view)**



Rack/Bus Bar Physical Specifications

TD MP SC EP PS

Figure 22. Bus bar pair dimensions (top view)

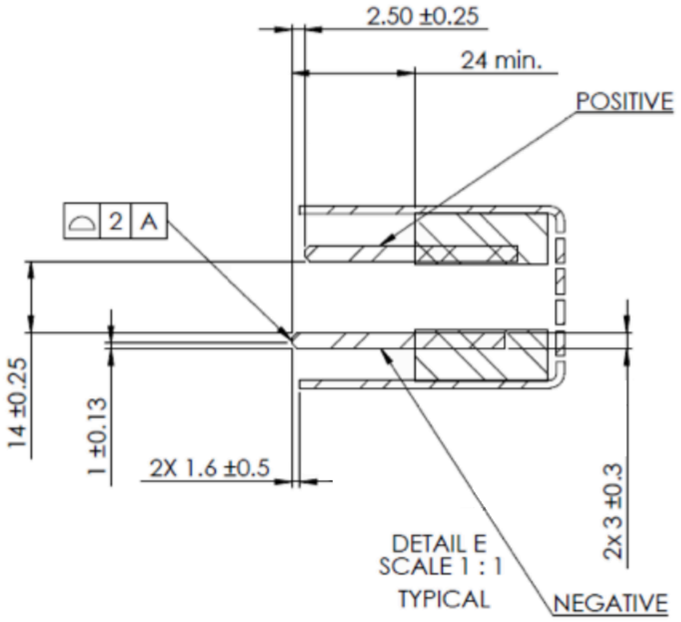
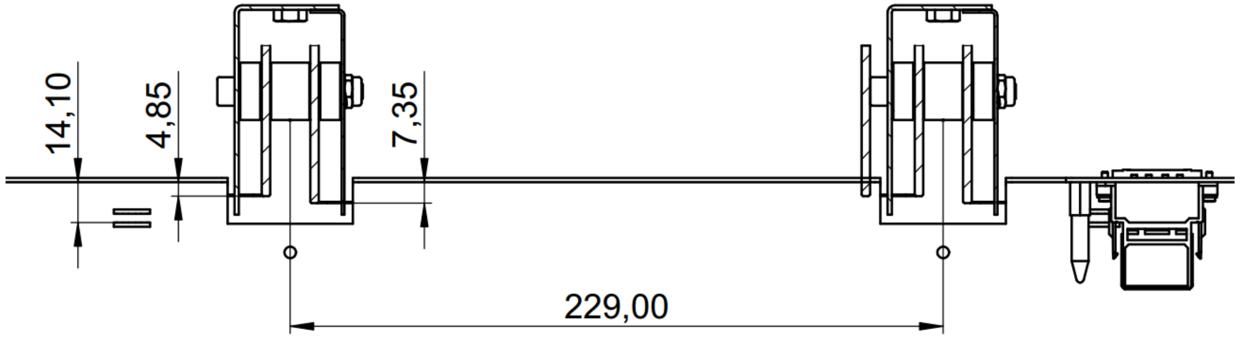


Figure 23. Bus bar pair horizontal location (top view)



Sled Optical Interconnect & Data Connections

TD MP SC EP PS

Figure 26. Sled blind-mate section assignments (rear view)

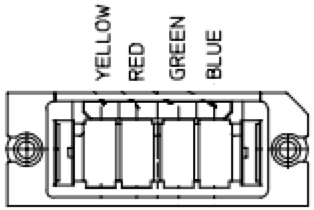


Figure 27. Example Data Plane and Management Plane Architecture

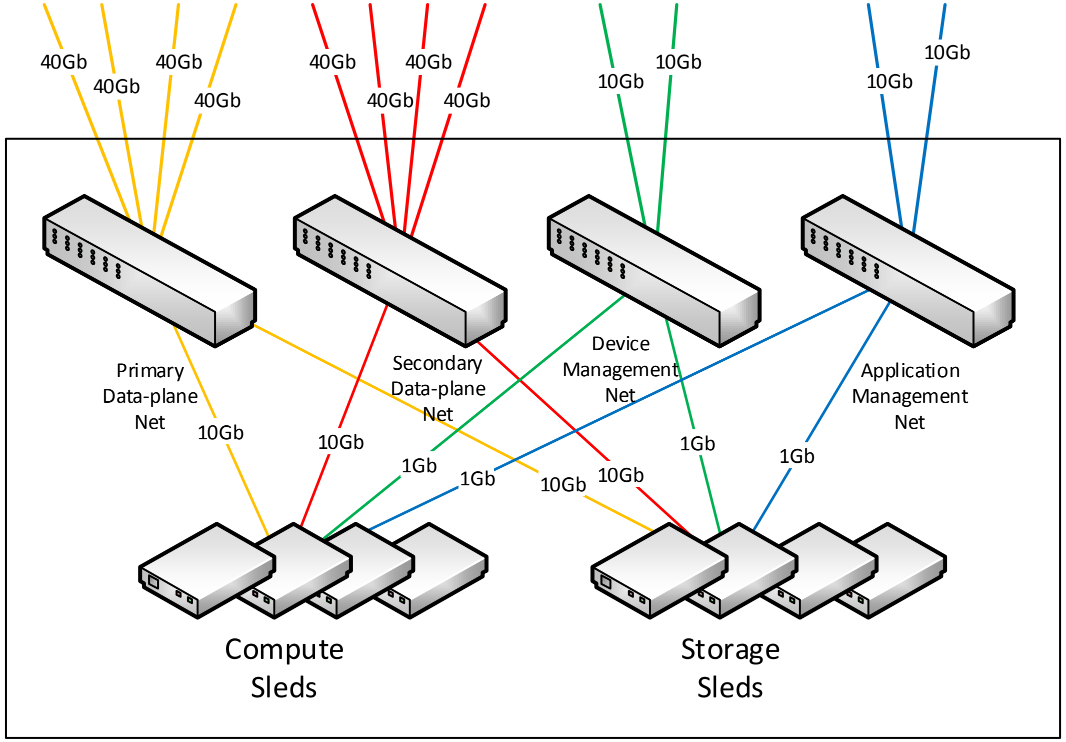


Figure 28. Network connection to sleds

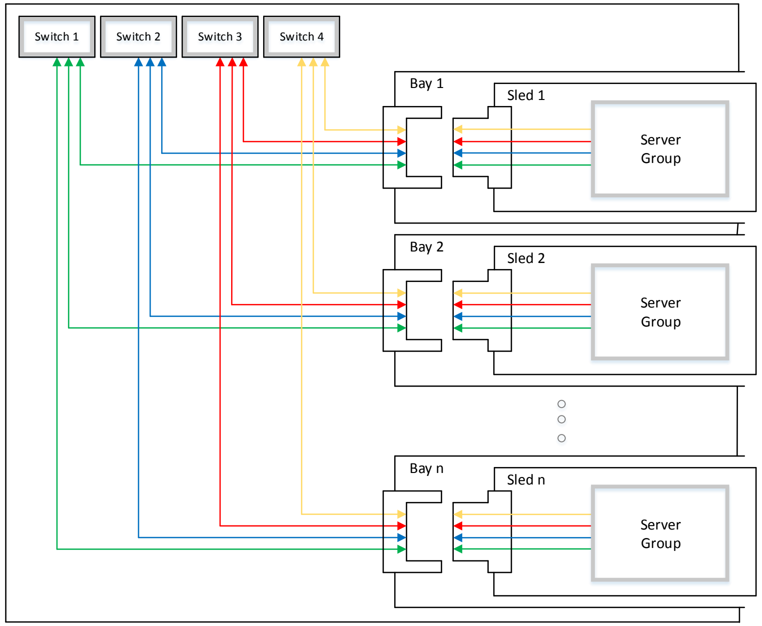
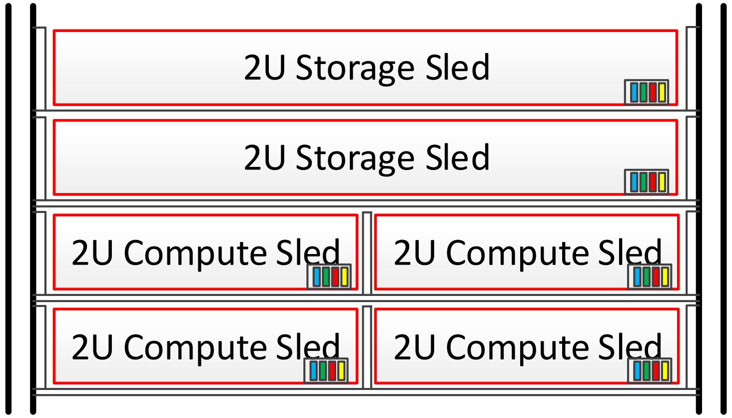


Figure 29. Sled Blind-Mate Optical Connections, Front View

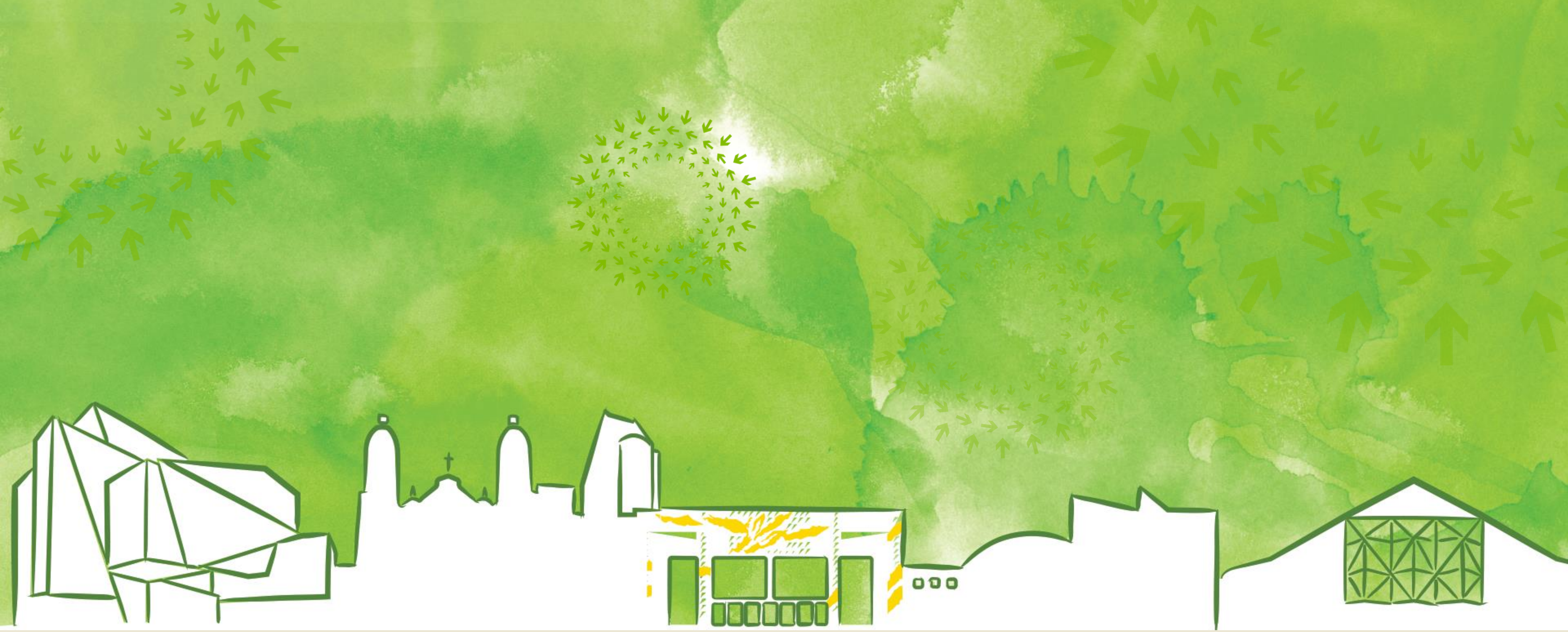


Management – Interop BMC focused only

radisys

TD MP SC EP PS

- Each CG-OpenRack-19 server shall have a dedicated BMC for various out-of-band platform management services.
- The health status of host processor should not affect the normal operation and network connectivity of the BMC.
- The BMC shall not share memory with host processor.
- The BMC management connectivity should work independently from host, and have no NIC driver dependency for Out-of-Band communication if using a shared-NIC.
- The BMC shall support a remote console, also known as Serial-Over-LAN (SOL) through the management network.
- POST codes should be displayed in SOL console during system POST.
- The BMC shall be fully IPMI 2.0 and DCMI 1.5 Compliant.
- BMC firmware shall support remote system power on/off/cycle and warm reboot through In-Band or Out-of-Band IPMI commands
- BMC firmware shall support a power-on policy selectable to be last-state, always-on, and always-off. The default setting is last-state.
- A change of power policy should be supported by IPMI command and take effect without BMC a firmware cold reset or a system reboot.
- Vendors should provide tool(s) to implement remote BMC firmware update, which will not require any physical access to the system. Remote update means either through Out-of-Band by management network or through In-Band by logging into local OS with data network.



OCP Engineering Workshop **10 August 2016**| Durham, NH