Contents

1 Introduction .............................................................................................................................................. 6
   Disclaimer .................................................................................................................................................. 6
   Scope .......................................................................................................................................................... 6
   Intended audience ..................................................................................................................................... 6
   Terminology .............................................................................................................................................. 6
   Conventions ............................................................................................................................................ 7
   1.1 Platform overview ............................................................................................................................... 7
   1.2 Software interfaces .................................................................................................................................. 8
   1.3 Recommended essential hardware elements .......................................................................................... 9
   1.5 Platform hardware design guideline summary ...................................................................................... 9
   1.6 Hardware Design General Specifications ............................................................................................ 11
   1.6.2 Drawer must have a PSME .................................................................................................................. 11
   1.6.3 Multiblade chassis capable of populating more than one blade inside the chassis ..................... 11
   2.1.1 Shared power .................................................................................................................................. 11
   2.1.2 Shared power bus bar across entire rack ........................................................................................ 11
   2.1.3 Shared cooling ................................................................................................................................. 11
   2.1.4 Centralized cooling across entire rack ............................................................................................ 12
   2.1.5 JBOD support .................................................................................................................................. 12
   2.1.7 Compute blade with one or more HDDs .......................................................................................... 12
   2.1.8 Composed node with M.2 drive ..................................................................................................... 12
   2.1.9 At least one Intel® Rack Scale Architecture compute blade node (drive optional) ..................... 12
   2.1.11 Compute blade serviceability independence .................................................................................. 12
   2.1.12 Ethernet-based fabric ................................................................................................................... 12
   2.1.13 At least one Ethernet switch per rack ........................................................................................... 13
   2.1.14 At least one Ethernet switch in the pod ......................................................................................... 13
   2.2 Network switch support for network software agent ........................................................................ 13
   2.2.1 Component location identification support .................................................................................... 13
   2.2.2 Unique drawer ID number within the rack .................................................................................... 14
   2.2.4 Drawer ID numbering from bottom to top ..................................................................................... 14
   2.2.5 Unique module ID number within the drawer ............................................................................... 14
   2.2.6 Module ID numbering from left to right ......................................................................................... 14
   2.2.8 Unique blade ID number within the module .................................................................................. 14
   2.2.9 Blade ID numbering from left to right or front to back ................................................................ 15
   2.3 Power supply position numbering left to right or top to bottom .................................................... 15
   2.3.1 Consistent and sequential fan position numbering left to right or top to bottom ....................... 15
   3.1 Unique ID for all FRUs ....................................................................................................................... 15
   3.1.1 Intel® Rack Scale Architecture fabric and secure management network configuration ................ 15
   3.1.2 Management network and production network separation ......................................................... 15
   3.1.3 Composed Node Design Guidelines ............................................................................................. 16
   4.2.1 Node reset, power, and performance .............................................................................................. 16
   4.3 Node reset support .................................................................................................................................. 16
   4.3.1 Power monitoring and power budget support (Intel® NM 3.0) ........................................................ 16
   4 PSME Design Guidelines .......................................................................................................................... 17
   PSME overview .......................................................................................................................................... 17
   PSME configuration management .............................................................................................................. 17
   Private rackwide network ........................................................................................................................... 17
   PSME managed assets ............................................................................................................................... 17
   Blade presence identification ..................................................................................................................... 17
5 RMM Design Guidelines .................................................................................................................................................... 18
   RMM overview ................................................................................................................................................................. 18
   Discrete RMM ................................................................................................................................................................. 18
      Discrete RMM – private rackwide management network ......................................................................................... 18
   RMM general support .................................................................................................................................................... 18
      High availability RMM support .................................................................................................................................. 18
6 POD Manager Design Guidelines ........................................................................................................................................ 19
  5.1 PODM overview .......................................................................................................................................................... 19
  5.2 PODM configuration management ........................................................................................................................... 19
  5.3 Secure communication channel for management network ....................................................................................... 19
A Intel® Rack Scale Architecture Platform Hardware Alignment Checklist ....................................................................... 20

Figures

Figure 1 Pod logical hierarchy ............................................................................................................................................... 7
Figure 2 API blocks ............................................................................................................................................................. 8
Figure 3 Intel® Rack Scale Architecture component location identification ................................................................. 14
Figure 4 Pod manager logical view ..................................................................................................................................... 19

Tables

Table 1 Terminology ............................................................................................................................................................ 6
Table 2 Intel® Rack Scale Architecture platform design guideline summary ................................................................. 10
Table 3 Essential criteria ...................................................................................................................................................... 20
## Revision History

<table>
<thead>
<tr>
<th>Revision</th>
<th>Description</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>001</td>
<td>Initial public release.</td>
<td>August 17, 2015</td>
</tr>
<tr>
<td>002</td>
<td>Addressed errors and user content feedback.</td>
<td>August 21, 2015</td>
</tr>
</tbody>
</table>
This document provides general hardware guidelines for Intel’s development partners to align with Intel® Rack Scale Architecture specifications for hardware. The complete Intel® Rack Scale Architecture platform design specifications are available to Intel partners and customers under NDA from your Intel field representative.

1 Introduction

Disclaimer

Technologies and concepts expressed in this document are in pathfinding and may go through substantial changes before being productized. No product commitment or even readiness should be inferred from this document.

1.1 Scope

The scope of this document is platform hardware design for Intel® Rack Scale Architecture hardware components.

1.2 Time Horizon: Guidance listed here covers 2015 designs that target 2016 or later commercial platforms.

1.3 Intended audience

The intended audiences for this document are hardware vendors (for example, OEMs/ODMs) who will build Intel® Rack Scale Architecture platforms or components that would be integrated into the platform.

1.4 Terminology

Table 1 Terminology

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>API</td>
<td>Application program interface. A set of routines, protocols, and tools for building software applications. API defines operations, inputs, and outputs.</td>
</tr>
<tr>
<td>BIOS</td>
<td>Basic input/output system. To initialize and test compute/storage node hardware components, and to load a boot loader or an operating system from a mass memory device. The BIOS supports UEFI interface.</td>
</tr>
<tr>
<td>BMC</td>
<td>Baseboard management controller. A specialized service processor that monitors the physical state of a computer and provides services to monitor and control certain node operations. The BMC supports Intelligent Platform Management Interface (IPMI).</td>
</tr>
<tr>
<td>composed node</td>
<td>Node composed by PODM. PODM composes nodes within the rack by communicating with PSME for compute node and storage allocation based on the user input.</td>
</tr>
<tr>
<td>CPP</td>
<td>Control Plane Process or is the Switch Management CPU. This is the host that Intel uses to run the PSME on the Intel Software Delivery Vehicle reference platform.</td>
</tr>
<tr>
<td>DMC</td>
<td>Drawer management controller. Controller that manages the drawer, where the PSME functionality is normally implemented.</td>
</tr>
<tr>
<td>EORS</td>
<td>End-of-row switch. Switch that is connected to the end of the row, either through copper or fiber.</td>
</tr>
<tr>
<td>HA</td>
<td>High availability. Generally a redundancy component available. For example, if a rack supports HA RMM, then more than one RMM present in the rack and if primary RMM fails, secondary RMM provides the RMM functionality.</td>
</tr>
<tr>
<td>iPXE</td>
<td>An open-source implementation of the PXE client firmware and bootloader.</td>
</tr>
<tr>
<td>MMC</td>
<td>Module management controller. The controller that manages the blades in the module.</td>
</tr>
<tr>
<td>node</td>
<td>Any compute node, such as an Intel® Xeon® or Intel® Atom™ processor, in the module under a drawer.</td>
</tr>
<tr>
<td>pod</td>
<td>A collection of racks within a shared infrastructure management domain.</td>
</tr>
<tr>
<td>PODM</td>
<td>Pod manager. Logical management functionality across all infrastructure in a pod.</td>
</tr>
<tr>
<td>PSME</td>
<td>Pooled system management engine. A microcontroller responsible for configuration of shared and pooled memory controlled by the SMC, pooled storage by the PNC, the nodes and network (SDN) of the compute nodes and switch. Also known as DMC (drawer management controller).</td>
</tr>
<tr>
<td>PTAS</td>
<td>Power thermal aware scheduling.</td>
</tr>
<tr>
<td>PXE</td>
<td>Preboot eXecution Environment. A specification that allows devices to boot over a network.</td>
</tr>
</tbody>
</table>
Conventions

The key words/phrases "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119.

1.5 Platform overview

Figure 1 illustrates the various elements for the platform in a pod logical view.

1.6 Figure 1 Pod logical hierarchy
Software interfaces

A complete hardware solution will require a management software interface with a set of APIs that facilitate use cases. An example of a REST based API framework for the Intel® Rack Scale Architecture PODM, RMM, and PSME is shown in Figure 2.

Figure 2   API blocks

Intel will make available examples of reference software, APIs, and code under NDA through your Intel representative or available via Open Source at 01.org in late 2015.
Recommended essential hardware elements

Intel recommends that all of the following items be incorporated into 2015 Intel® Rack Scale Architecture hardware designs in order to be aligned with the Intel vision and intent. The latter part of this document describes these criteria in more detail.

Intel® Rack Scale Architecture systems should incorporate the following essential elements:

1.6.2
- A PSME for each drawer.
- Efficient power.
- Efficient cooling.
- At least one compute blade node (drive optional).
- Ethernet-based fabric.
- At least one Ethernet switch in the pod.
- Unique drawer ID number within the rack.
- Unique module ID number within the drawer.
- Unique blade ID number within the module.
- Management network and production network separation.
- Node reset support.
- Power monitoring and power budget support (NM3.0).
- Secure private rackwide network.
- Blade presence identification.
- Discrete RMM – private rackwide management network.
- Secure communication channel for management network.

1.6.3 Platform hardware design guideline summary

This section lists the complete requirements for defining how the platform shall be designed to conform to the Intel® Rack Scale Architecture design. Not all requirements are equal, however; some elements may not be immediately available in 2015 designs or are not as critical in the early phases of product rollout.

Intel considers a hardware system to be “Intel® Rack Scale Architecture Aligned” if 75% of the items in this section are intended to be implemented by the partner by the end of 2016 as well as all of the items listed in section 1.6. Items listed as “Required” are more critical to use cases in the near term as compared to “Optional” or “Recommended” items.

The items listed in the previous section are essential for Intel® Rack Scale Architecture Alignment as fundamental currently available building blocks. The remaining items in this section are highly recommended to be incorporated into the partner’s roadmap in the 2016 time horizon.

Note that each row of the Table 2 contains a reference to a section in the appendix which contains a more complete description of the design criteria.
<table>
<thead>
<tr>
<th>Section</th>
<th>Design criteria</th>
<th>Design criticality</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1.1</td>
<td>Drawer must have a PSME</td>
<td>Required</td>
</tr>
<tr>
<td>2.1.2</td>
<td>Multiblade chassis capable of populating more than one blade inside the chassis</td>
<td>Recommended</td>
</tr>
<tr>
<td>2.1.3</td>
<td>Shared power</td>
<td>Required</td>
</tr>
<tr>
<td>2.1.4</td>
<td>Shared power bus bar across entire rack</td>
<td>Recommended</td>
</tr>
<tr>
<td>2.1.5</td>
<td>Shared cooling</td>
<td>Required</td>
</tr>
<tr>
<td>2.1.6</td>
<td>Centralized cooling across entire rack</td>
<td>Recommended</td>
</tr>
<tr>
<td>2.1.7</td>
<td>JBOD support</td>
<td>Optional</td>
</tr>
<tr>
<td>2.1.8</td>
<td>Compute blade with one or more HDDs</td>
<td>Optional</td>
</tr>
<tr>
<td>2.1.9</td>
<td>Composed node with M.2 drive</td>
<td>Optional</td>
</tr>
<tr>
<td>2.1.10</td>
<td>At least one Intel® Rack Scale Architecture compute blade node (drive optional)</td>
<td>Required</td>
</tr>
<tr>
<td>2.1.11</td>
<td>Compute blade serviceability independence</td>
<td>Required</td>
</tr>
<tr>
<td>2.1.12</td>
<td>Ethernet-based fabric</td>
<td>Required</td>
</tr>
<tr>
<td>2.1.13</td>
<td>At least one Ethernet switch per rack</td>
<td>Optional</td>
</tr>
<tr>
<td>2.1.14</td>
<td>At least one Ethernet switch in the pod</td>
<td>Required</td>
</tr>
<tr>
<td>2.1.15</td>
<td>Network switch support for network software agent</td>
<td>Required</td>
</tr>
<tr>
<td>2.2.1</td>
<td>Unique drawer ID number within the rack</td>
<td>Required</td>
</tr>
<tr>
<td>2.2.2</td>
<td>Drawer ID numbering from bottom to top</td>
<td>Recommended</td>
</tr>
<tr>
<td>2.2.3</td>
<td>Unique module ID number within the drawer</td>
<td>Required</td>
</tr>
<tr>
<td>2.2.4</td>
<td>Module ID numbering from left to right</td>
<td>Required</td>
</tr>
<tr>
<td>2.2.5</td>
<td>Unique blade ID number within the module</td>
<td>Required</td>
</tr>
<tr>
<td>2.2.6</td>
<td>Blade ID numbering from left to right or front to back</td>
<td>Recommended</td>
</tr>
<tr>
<td>2.2.7</td>
<td>Power supply position numbering left to right or top to bottom</td>
<td>Required</td>
</tr>
<tr>
<td>2.2.8</td>
<td>Consistent and sequential fan position numbering left to right or top to bottom</td>
<td>Required</td>
</tr>
<tr>
<td>2.2.9</td>
<td>Unique ID for all FRUs</td>
<td>Recommended</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Management network and production network separation</td>
<td>Required</td>
</tr>
<tr>
<td>3.1.1</td>
<td>Node reset support</td>
<td>Required</td>
</tr>
<tr>
<td>3.1.2</td>
<td>Power monitoring and power budget support (Intel® NM 3.0)</td>
<td>Required</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Private rackwide network</td>
<td>Required</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Blade presence identification</td>
<td>Required</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Discrete RMM – private rackwide management network</td>
<td>Required</td>
</tr>
<tr>
<td>5.3.1</td>
<td>High availability RMM support</td>
<td>Optional</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Secure communication channel for management network</td>
<td>Required</td>
</tr>
</tbody>
</table>
This section describes the Intel® Rack Scale Architecture platform-level feature design guidelines. The subsequent sections describe the platform subcomponent level design guidelines that apply to the platform hardware. The Intel® Rack Scale Architecture platform meets the following generic requirements as stipulated (optional, recommended, or required) in each section.

2 Hardware Design General Specifications

2.1 Drawer must have a PSME

Required

A rack is made up of drawers. An Intel® Rack Scale Architecture system must provide a mechanism to manage each rack level end point components down to the drawer level. Therefore, the drawers must support management by having a PSME.

In some cases, a PSME may service multiple drawers, as long as the drawer is uniquely addressable and provides the necessary instrumentation. For example, if each drawer has a microcontroller to provide the necessary instrumentation for all drawer requirements (such as module presence detection) and is interfaced to the RMM, then the drawer PSME could physically run in the RMM and represent each drawer instance to meet this requirement.

2.1.2 Multiblade chassis capable of populating more than one blade inside the chassis

Recommended

We recommend having a drawer to support multiple blades and share a power supply across multiple blades in the platform. This feature allows modularity for upgrade, serviceability, and to effectively utilize the real estate footprint.

2.1.3 Shared power

Required

Compute blades must support cost-effective, efficient, and manageable shared power. It can be achieved either by sharing power across two or more blades or having a nonredundant >90% efficient (delivered power to blade/input) AC power. Efficient data center power delivery is necessary to reduce operating cost, which is a core value proposition for Intel® Rack Scale Architecture.

2.1.4 Shared power bus bar across entire rack

Recommended

Use of shared power bus bar across entire rack would allow consolidation of the power supplies, reduce real estate usage, and reduce cost.

2.1.5 Shared cooling

Required

Compute blades must support cost-effective and manageable shared cooling to positively affect TCO. The platform must support shared cooling across two or more blades. One option to achieve efficient cooling is to use a shared cooling fan larger than 2U (3.5 inches) in diameter that can cool multiple modules simultaneously. If the system does not utilize fans, the system could implement a more efficient system than shared cooling, such as liquid cooled system.
Centralized cooling across entire rack

Recommended.

Centralized cooling generally increases efficiency and reduces cost when the racks are fully populated. We recommend that the platforms be designed with centralized cooling across the entire rack. In an Intel® Rack Scale Architecture design, centralized cooling increases the platform cooling efficiency and operating cost.

2.1.6

JBOD support

Optional.

JBOD allows for efficient pooling of storage resources. Composed nodes could have local storage, JBOD, network-based storage, or a combination of any of these. This feature allows efficient pooling of storage resources as well as more efficient management of storage resources.

2.1.7

Compute blade with one or more HDDs

Optional.

An Intel® Rack Scale Architecture compute blade has iSCSI or iPXE support; hence, the boot or storage could be accessed through the network. Compute blades sometimes use local storage such as HDD for boot or delta-file storage.

The benefit of local compute HDD is that, in the event of network errors, the system may be able to store the errors as well as the current data in the local storage, if local storage is present.

2.1.8

Composed node with M.2 drive

Optional.

Use of M.2 as local storage would provide better performance, compared to an HSD.

In an Intel® Rack Scale Architecture design, the M.2 drive will require less rack real estate as compared to drive bay based local storage.

2.1.9

At least one Intel® Rack Scale Architecture compute blade node (drive optional)

Required.

The pod must have at least one compute drawer with at least one compute blade. It is possible that some racks support storage and some support compute only, but a pod must have at least one composed blade. A compute blade is essential as it is the most basic building block for the rack.

2.1.10

Compute blade serviceability independence

Required.

Compute blades must support modular CPU and memory that can be serviced or upgraded independent of shared resources (shared power, shared cooling, shared network, shared storage).

2.1.11

Ethernet-based fabric

Required.

The platform must support Ethernet-based fabric for network connectivity.
The use of a common ubiquitous transport is important for standard Intel® Rack Scale Architecture designs and interoperability. Ethernet is a network used widely today and Intel® Rack Scale Architecture devices need to support widely used network technologies.

**At least one Ethernet switch per rack**

Optional.

The network in the rack is generally connected via a disaggregated switch or TORS, but it is possible to build an Intel® Rack Scale Architecture system with the network connected to an EORS with a patch panel in the compute rack. The Ethernet switch is necessary to provide connectivity from the rack to the rest of the datacenter and the user workloads.

**At least one Ethernet switch in the pod**

Required.

The pod must have at least one Ethernet switch component to connect from the pod to the external network. The Ethernet switch is necessary to provide PODM connectivity from the rack to the user workloads.

**Network switch support for network software agent**

Required.

Network switch components (such as TORS and disaggregated switches) must support running a network software agent to monitor and configure the switches. The combined network switch and network software agent provide a mechanism to deliver security fixes, performance fixes, and bug fixes, which will keep the system up to date.

**Component location identification support**

A key attribute of Intel® Rack Scale Architecture management is location-aware discovery. A datacenter manager should be able to identify the physical location of hardware components to service the hardware components. The requirements below suggest a specific identification method (for example, left to right). An alternative identification method can use as long as it is unique, communicates the location, is consistently used across the entire implementation, and is supported in the management interfaces (for example, right to left).

Identification support is essential to Intel® Rack Scale Architecture use cases because they enable the consistent disaggregation, assembly, and disassembly of assets in pooled system management. The Intel® Rack Scale Architecture system requires a mechanism to identify the components. The specifications in this section propose a mechanism for consistent tagging that will enable efficient management.

To support the location identification, each Intel® Rack Scale Architecture component must meet the following requirements, as shown in Figure 3.
2.2.1 Unique drawer ID number within the rack

Required.

The drawers within the rack must have a unique number based on where the drawer is populated in the rack. The scope of the drawer ID is the rack.

2.2.2 Drawer ID numbering from bottom to top

Recommended.

We recommend that drawer numbers use the base as 1, and be numbered from bottom to top within the rack based on the physical location. If any location is not populated, then that location is skipped.

2.2.4 Unique module ID number within the drawer

Required.

The drawers within the rack must have a unique number based on where the drawer is populated in the rack. The scope of the module ID is the drawer.

2.2.5 Module ID numbering from left to right

Recommended.

We recommend that module numbers use the base as 1 and be numbered from left to right within the drawer, based on the physical location. If any module is not populated, then that location indicates module not present. If the system modules are aligned vertically, then they should be numbered as described, bottom to top.

2.2.6 Unique blade ID number within the module

Required.
The blades within the module must have a unique number based on where the blade is populated in the drawer. The scope of the blade ID is the module.

### Blade ID numbering from left to right or front to back

Recommended.

We recommend that blade numbers use the base as 1 and be numbered from left to right or front to back within the module, based on physical location. If a blade is not populated, then that location shall indicate blade not present.

### Power supply position numbering left to right or top to bottom

Required.

Service personnel should be able to easily identify the location of a power supply failure for replacement. We recommend that the power supply position location label use the base as 1 and be numbered from left to right or top to bottom, within each subsystem (rack, drawer, or module). The location information could be implemented using straps or switches.

### Consistent and sequential fan position numbering left to right or top to bottom

Required.

Service personnel should be able to easily identify the location of a failed fan for replacement. We recommend that the fan position location label use the base as 1 and be numbered from left to right or top to bottom, within each subsystem (rack, drawer, or module). The location information could be implemented using straps or switches.

### Unique ID for all FRUs

Recommended.

We recommend that each major field replaceable unit (FRU) in the rack (such as TORS, EORS, modules, and blades) have a GUID (such as manufacturer ID, model number, and serial numbers) to help identify components for service.

### Intel® Rack Scale Architecture fabric and secure management network configuration

The RMM, PSME, CPP, TORS, and EORS should be able to communicate to their upstream ports, come up with IP address/port mappings, and respond to Intel® Rack Scale Architecture API requests.

### Management network and production network separation

Required.

The platform must support a separate IP address for Intel® Rack Scale Architecture management-related communication (control plane) such as PODM, RMM, and PSME configuration, and for production network (data plane).

A separate management network is essential to preventing unauthorized access from general access. Normally during the component rests, the components go through initialization and authentication process in which the network component may be vulnerable for attach. This requirement attempts to prevent those attacks.
This section describes the Intel® Rack Scale Architecture platform compute/storage node design guidelines. A composed node is generally a combination of a compute blade and a module with network and/or storage connectivity. In some cases, the module and blade may be a single hardware element.

3 Composed Node Design Guidelines

Node reset, power, and performance

Node reset support

3.1 Required.

3.1.1 The PSME should provide individual node reset support.

Power monitoring and power budget support (Intel® NM 3.0)

3.1.2 Required.

Intel® Rack Scale Architecture composed nodes must implement power measuring and power control logic for the BMC/PSME to determine the power consumed by the node and limit the power to the nodes. The node must support the specifications in Intel® Intelligent Power Node Manager 3.0 External Interface Specification Using IPMI (332200) or a later version or the functional equivalent. The following Intel® NM features must be supported:

- Platform power capping.
- Power threshold alerting.
- Changing P/T-states available.

We recommend PTAS support be provided at the drawer, module, and blade level.
This section describes Intel® Rack Scale Architecture platform PSME design guidelines.

## PSME overview

The PSME is responsible for drawer identification management, as well as supporting the PSME Intel® Rack Scale Architecture API, and communicating with the BMC to perform node-level management. If the RMM is not present in the rack, one of the PSMEs in the rack would provide the RMM functionality. On-storage bricks may not have a PSME, in which case the BMC would perform PSME functionality and provide PSME Intel® Rack Scale Architecture API support.

### PSME configuration management

#### Private rackwide network

**Required.**

**4.2.1**

The PSME must be connected to a private rackwide management network that connects PSME-RMM for the reasons outlined in section 2.3.1. The PSME must get its IP address from the datacenter admin before it communicates any management information.

### PSME managed assets

#### Blade presence identification

**Required.**

The PSME should provide blade presence and location.
This section describes the Intel® Rack Scale Architecture platform RMM design guidelines.

## RMM overview

The rack manager module (RMM) is responsible for handling infrastructure functions such as power, cooling, and assigning PSME IDs.

### 5 Discrete RMM

Optional.

If the rack contains a discrete RMM, the following requirements apply to the RMM.

#### 5.2 Discrete RMM – private rackwide management network

Required.

The RMM must be connected to a rackwide private network for the reasons outlined in section 2.3.1. The RMM must be able to get its IP address from the datacenter administrator before it communicates to the PODM.

### 5.3 RMM general support

#### 5.3.1 High availability RMM support

Optional.

The racks that require high availability (HA), multiple RMMs could be incorporated in the rack. We recommend the following for HA RMM implementation:

- When multiple RMMs are present, there must be only one active RMM; others are passive RMMs.
- The active RMM and passive RMMs must coordinate such that a copy of the active RMM's data will be in sync with the passive RMMs.
- All RMMs must be on the rackwide private network.
- A passive RMM could implement a heartbeat mechanism to detect a failure of the active RMM. When multiple RMMs are present, generally a priority order will be assigned to determine who will take over upon failure of the active RMM.
- If an RMM is hot-added, the active RMM must provide the current configuration information, such as PSME ID, to the newly added RMM.

The use of multiple RMMs enhances system availability through redundancy.
This section describes the Intel® Rack Scale Architecture platform PODM design guidelines.

**PODM overview**

The pod manager, as shown in Figure 4, is responsible for discovery of resources in the pod, configuring the resources, power and reset control, power management, fault management, monitoring the resources usage, and composing a server based on Intel® Rack Scale Architecture specifications.

The pod manager interacts with RMMs in the rack, PSMEs, and CPPs to create representation of the Intel® Rack Scale Architecture pod.

The pod manager allows composing a physical node to match the logical node requirements specified by the solution stack. Such composition is able to specify a system at a sub-composed node granularity, as shown in Figure 4.

**PODM configuration management**

Secure communication channel for management network

Required.

The PODM must be connected to the RMM and PSME through a private network for the reasons outline in section 2.3.1. Any management related activity such as reconfiguration must be performed only after establishing secure communication channel between the PODM to the PSME and the PODM to the RMM.
Table 3 lists the essential criteria for incorporating into a 2015 Intel® Rack Scale Architecture platform design. Hardware manufacturers and designers can use this checklist as a guide to test how closely the design conforms to the most critical requirements.

<table>
<thead>
<tr>
<th>Section</th>
<th>Validation criteria</th>
<th>Design checklist</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1.1</td>
<td>Drawer must have a PSME</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1.3</td>
<td>Shared power</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1.5</td>
<td>Shared cooling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1.10</td>
<td>At least one Intel® Rack Scale Architecture compute blade node (drive optional)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1.12</td>
<td>Ethernet-based fabric</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1.14</td>
<td>At least one Ethernet switch in the pod</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.2.1</td>
<td>Unique drawer ID number within the rack</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.2.3</td>
<td>Unique module ID number within the drawer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.2.5</td>
<td>Unique blade ID number within the module</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.3.1</td>
<td>Management network and production network separation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.1.1</td>
<td>Node reset support</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.1.2</td>
<td>Power monitoring and power budget support (Intel® NM 3.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.2.1</td>
<td>Private rackwide network</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.3.1</td>
<td>Blade presence identification</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.2.1</td>
<td>Discrete RMM – private rackwide management network</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.2.1</td>
<td>Secure communication channel for management network</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>