



(19) **United States**

(12) **Patent Application Publication**
Balle et al.

(10) **Pub. No.: US 2018/0027057 A1**

(43) **Pub. Date: Jan. 25, 2018**

(54) **TECHNOLOGIES FOR PERFORMING ORCHESTRATION WITH ONLINE ANALYTICS OF TELEMETRY DATA**

Publication Classification

(51) **Int. Cl.**
H04L 29/08 (2006.01)
G08C 17/02 (2006.01)
H04L 12/24 (2006.01)
H04L 12/911 (2006.01)
(52) **U.S. Cl.**
CPC *H04L 67/1008* (2013.01); *H04L 47/826* (2013.01); *H04L 67/12* (2013.01); *G08C 17/02* (2013.01); *H04L 41/147* (2013.01); *H04L 67/306* (2013.01); *G08C 2200/00* (2013.01)

(71) Applicants: **Susanne M. Balle**, Hudson, NH (US); **Rahul Khanna**, Portland, OR (US); **Nishi Ahuja**, University Place, WA (US); **Mrittika Ganguli**, Bangalore (IN)

(72) Inventors: **Susanne M. Balle**, Hudson, NH (US); **Rahul Khanna**, Portland, OR (US); **Nishi Ahuja**, University Place, WA (US); **Mrittika Ganguli**, Bangalore (IN)

(21) Appl. No.: **15/395,084**

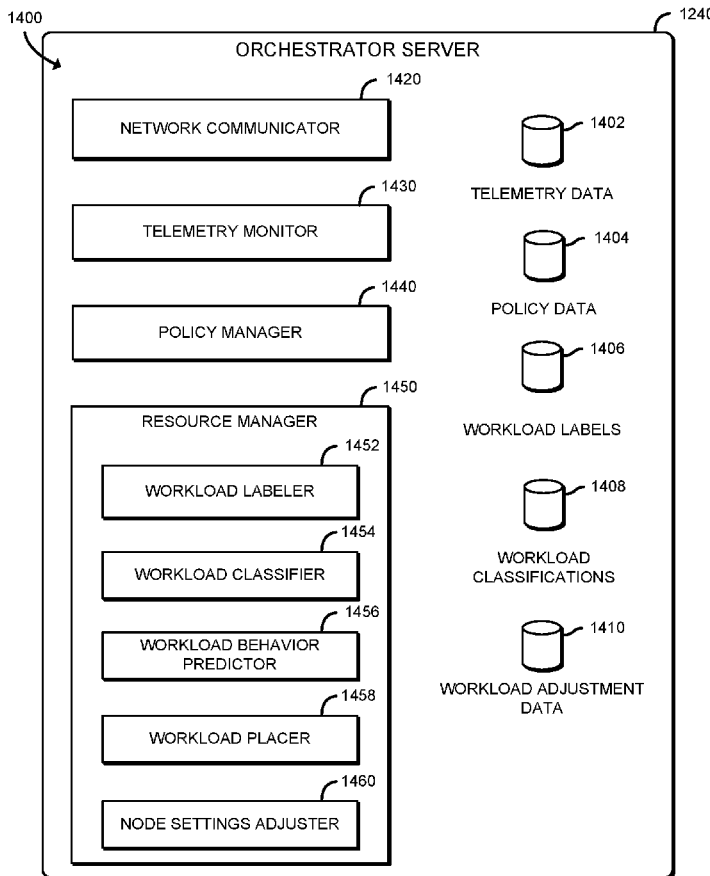
(22) Filed: **Dec. 30, 2016**

Related U.S. Application Data

(60) Provisional application No. 62/365,969, filed on Jul. 22, 2016, provisional application No. 62/376,859, filed on Aug. 18, 2016, provisional application No. 62/427,268, filed on Nov. 29, 2016.

(57) **ABSTRACT**

Technologies for performing orchestration with online analytics of telemetry data include an orchestrator server to assign workloads to each of a set of managed nodes, receive telemetry data indicative of resource utilization from the managed nodes as the workloads are performed, generate data analytics as a function of the telemetry data as the workloads are performed, determine, as a function of the data analytics, adjustments to the workload assignments to increase resource utilization among the managed nodes as the workloads are performed, and apply the determined adjustments to the managed nodes as the workloads are performed. Other embodiments are also described and claimed.



Data Center 100

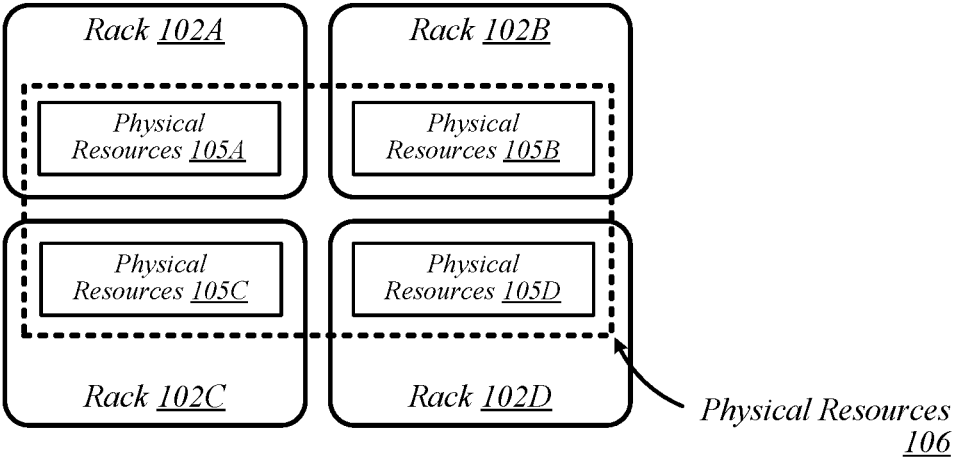


FIG. 1

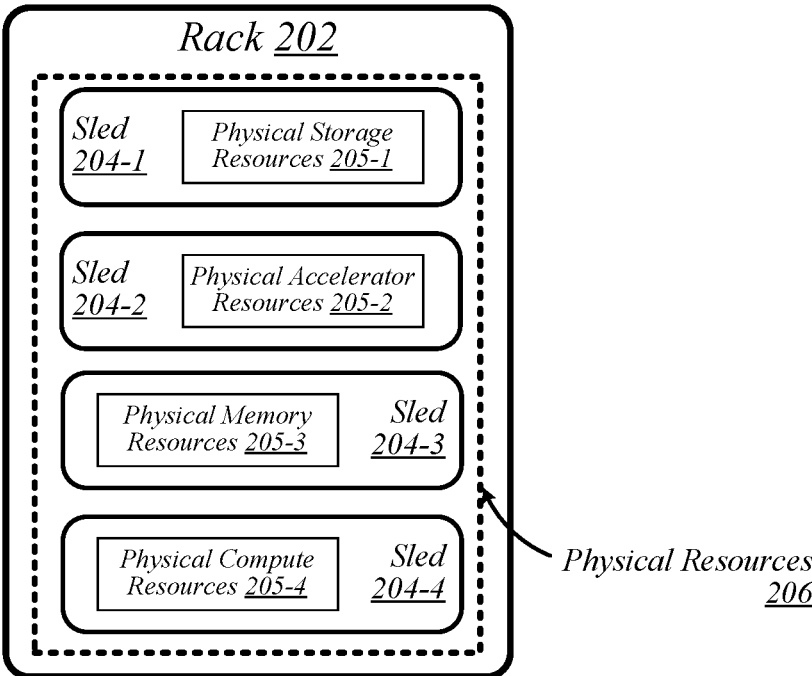


FIG. 2

Data Center 300

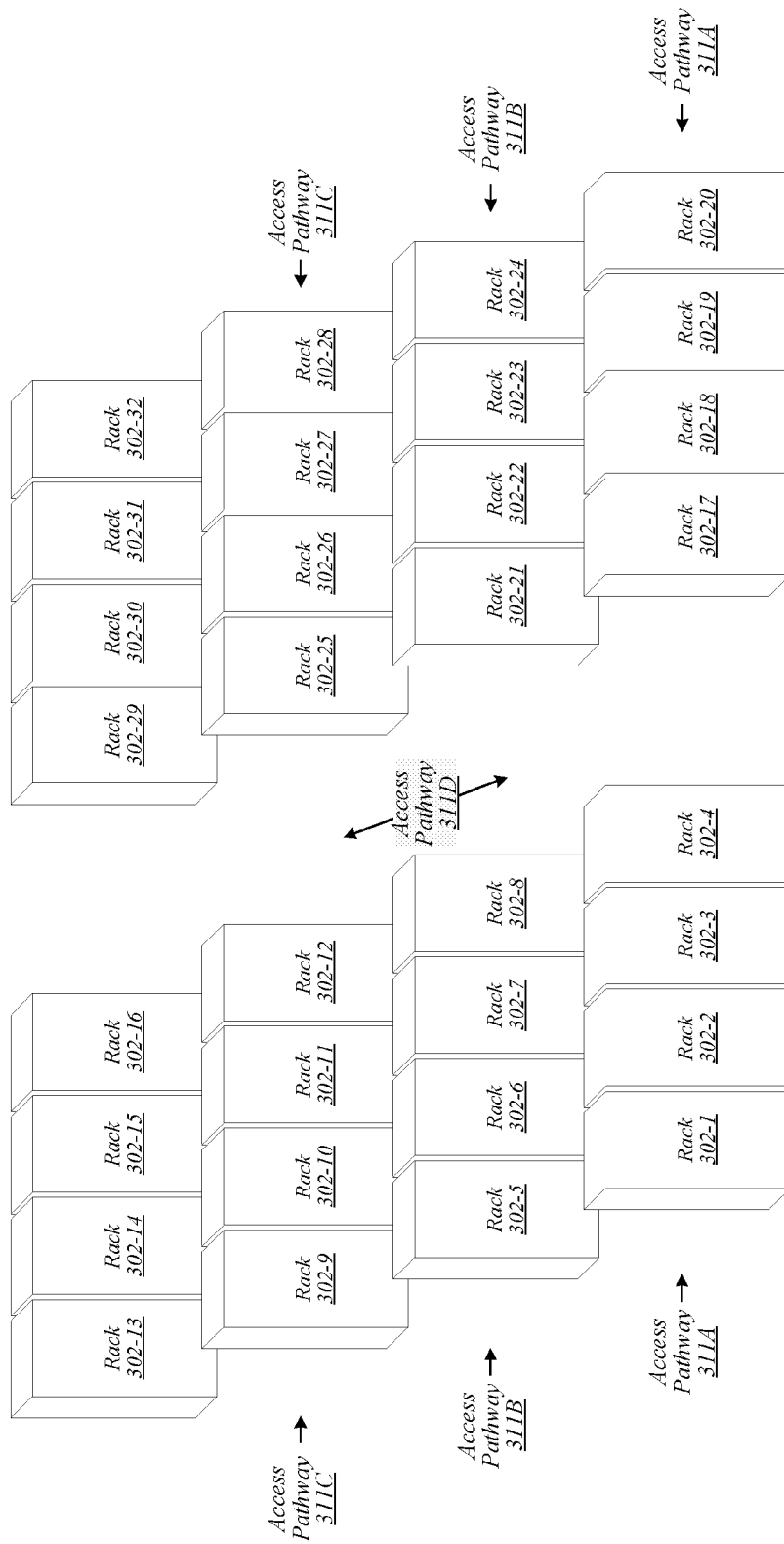


FIG. 3

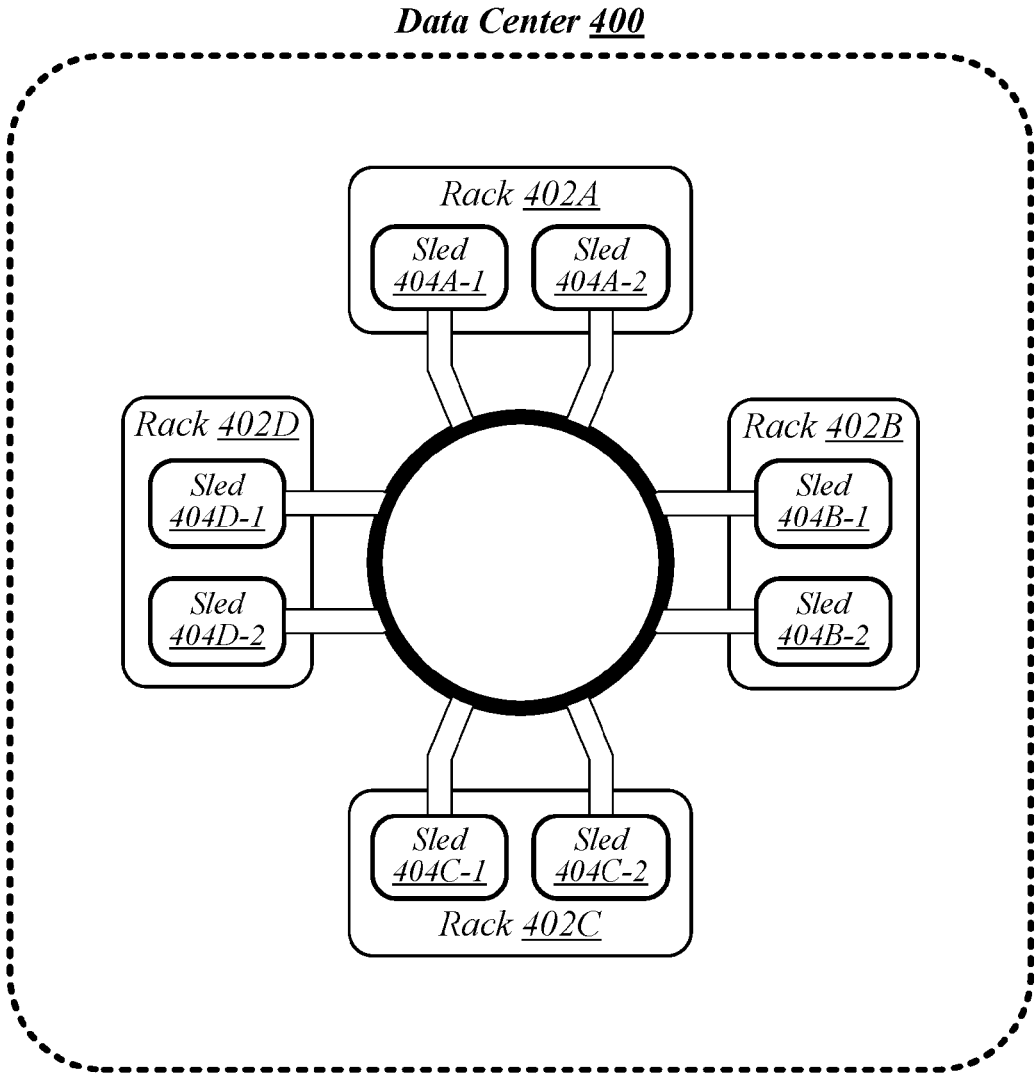


FIG. 4

500

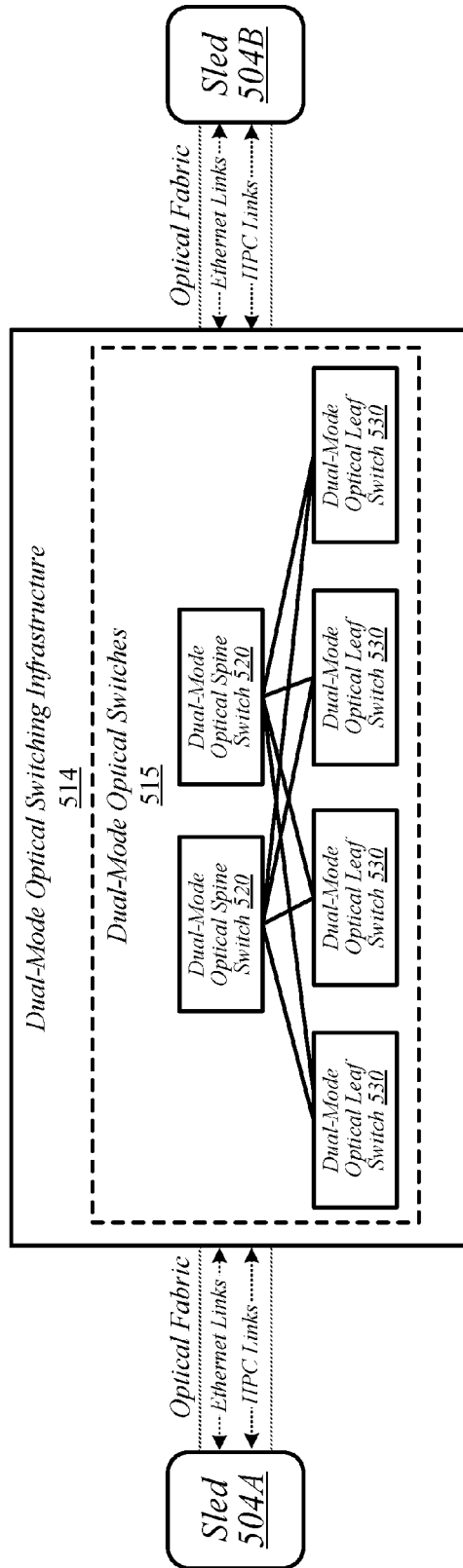


FIG. 5

600

Rack Access Region 601

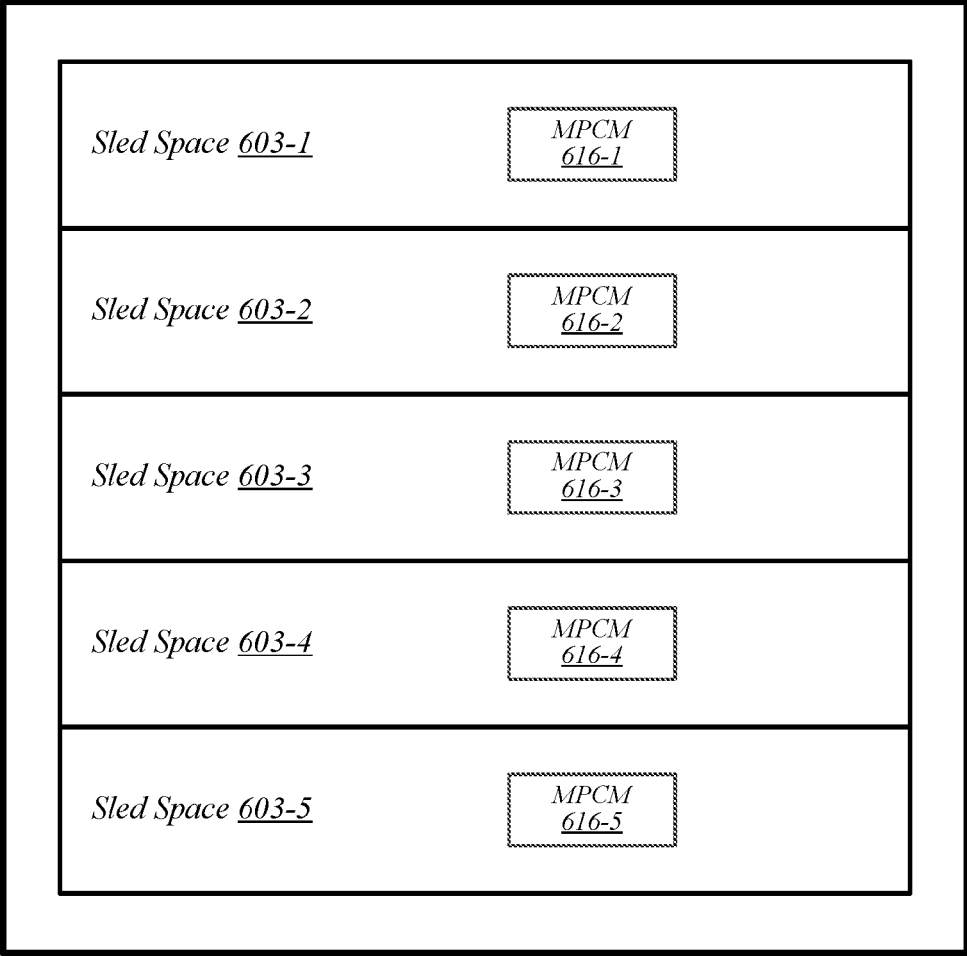


FIG. 6

Sled 704

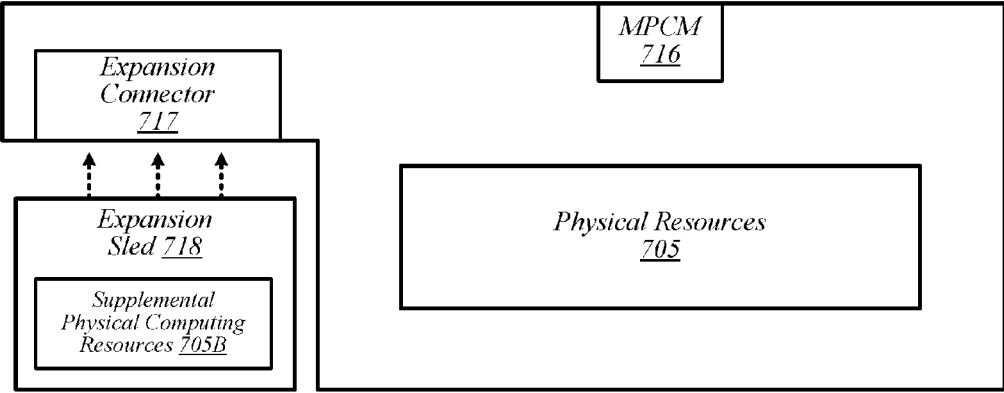


FIG. 7

800

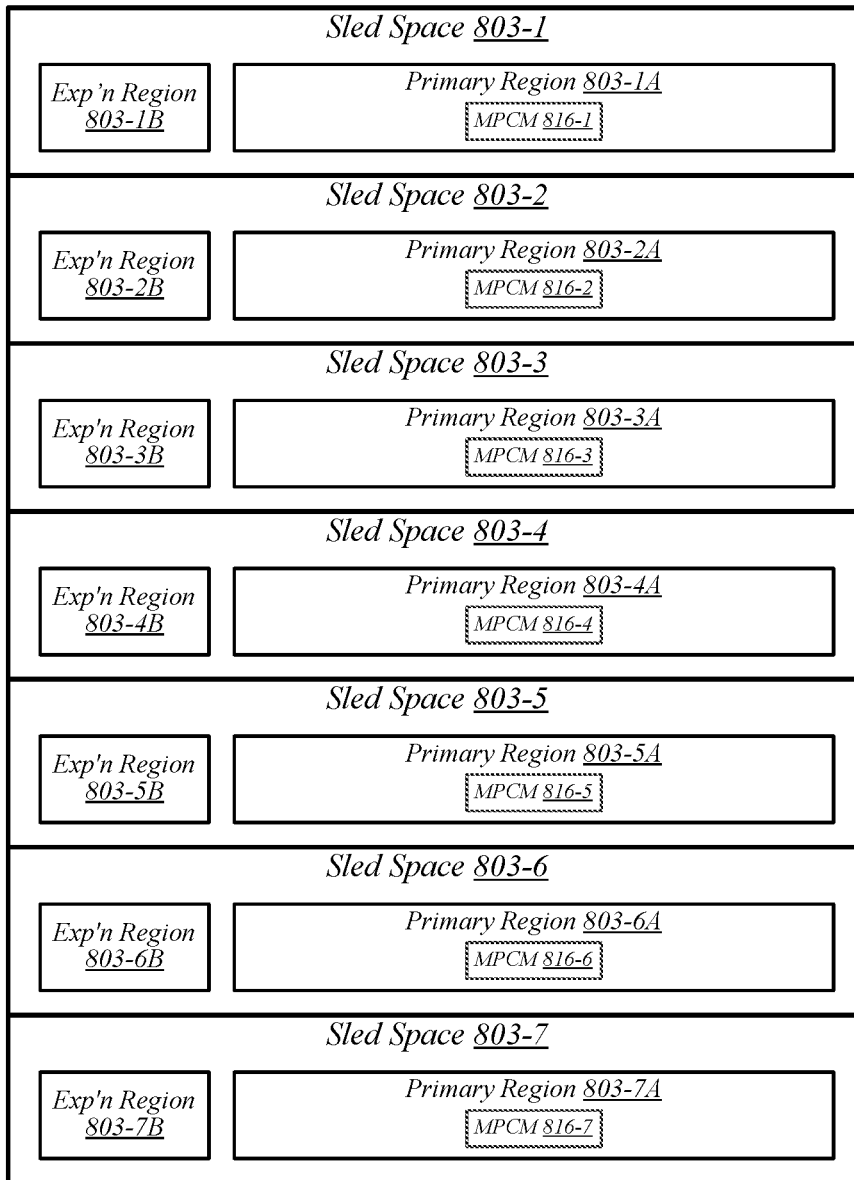


FIG. 8

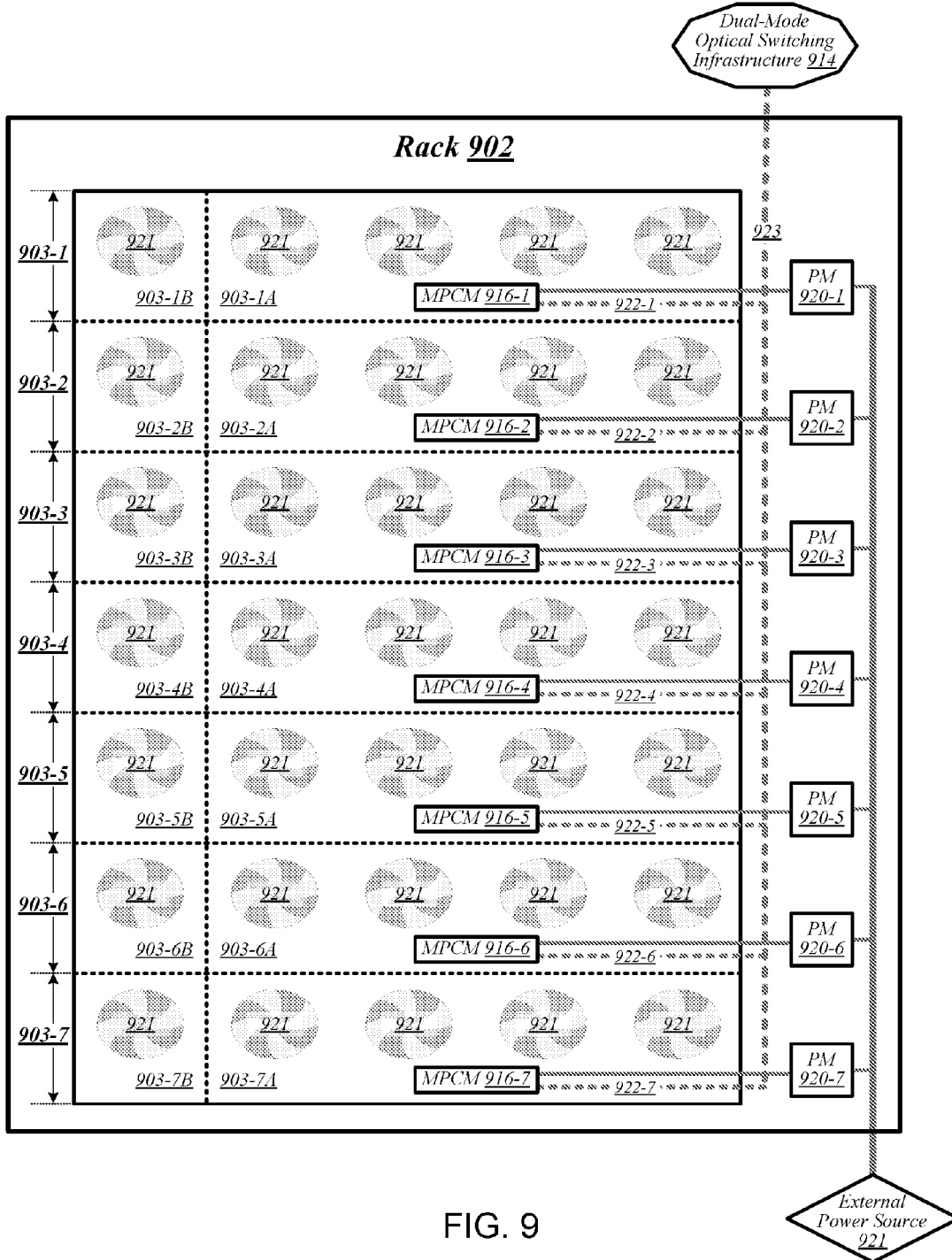


FIG. 9

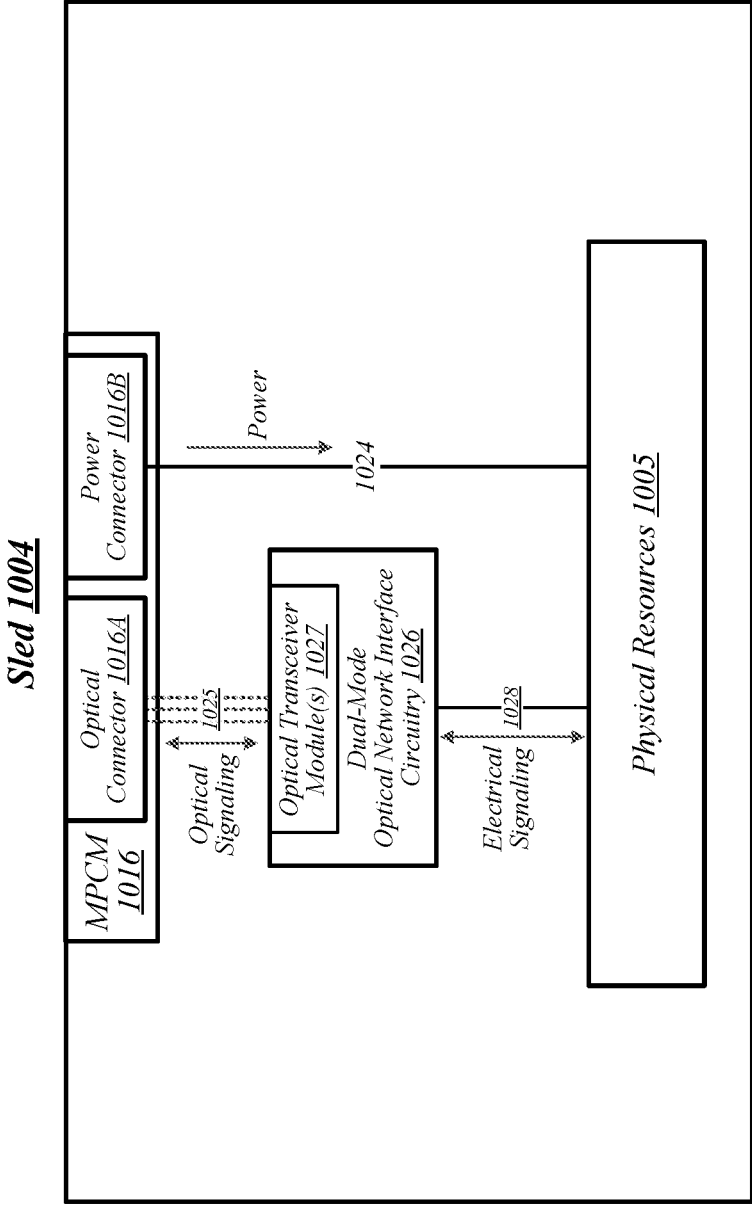


FIG. 10

Data Center 1100

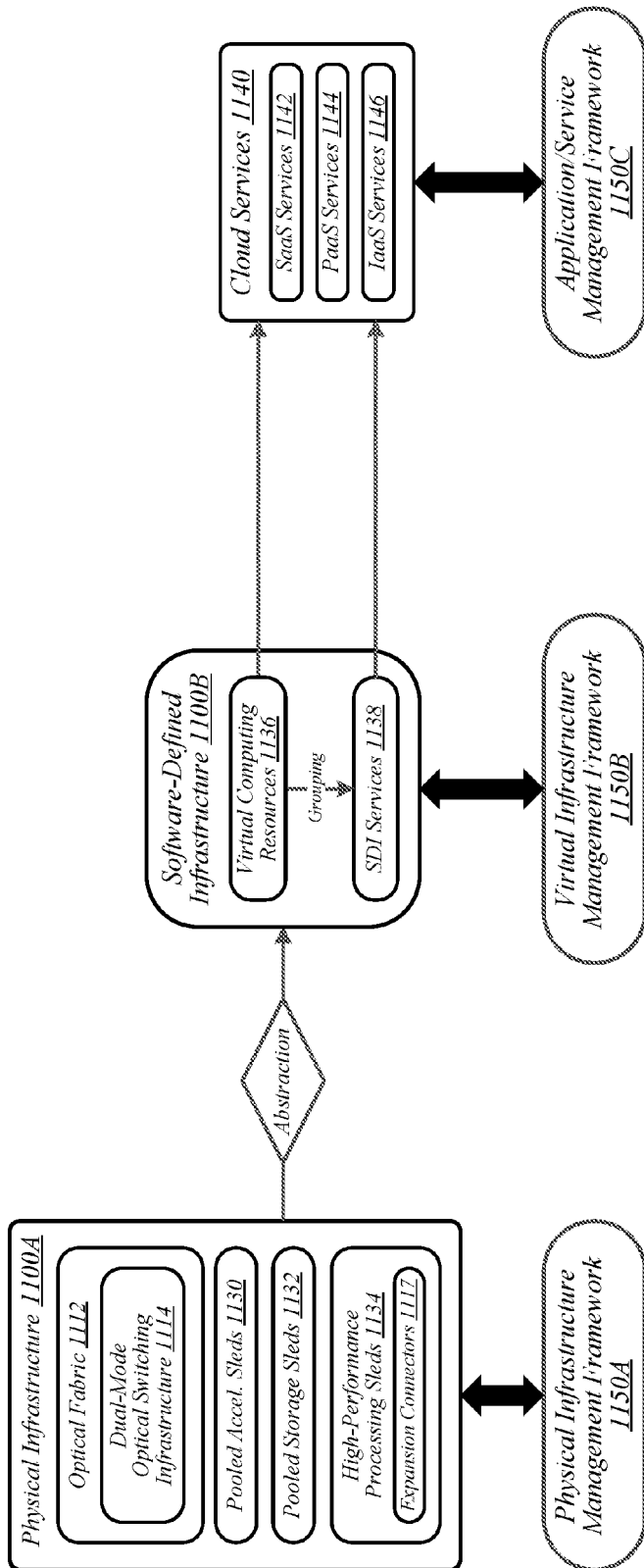


FIG. 11

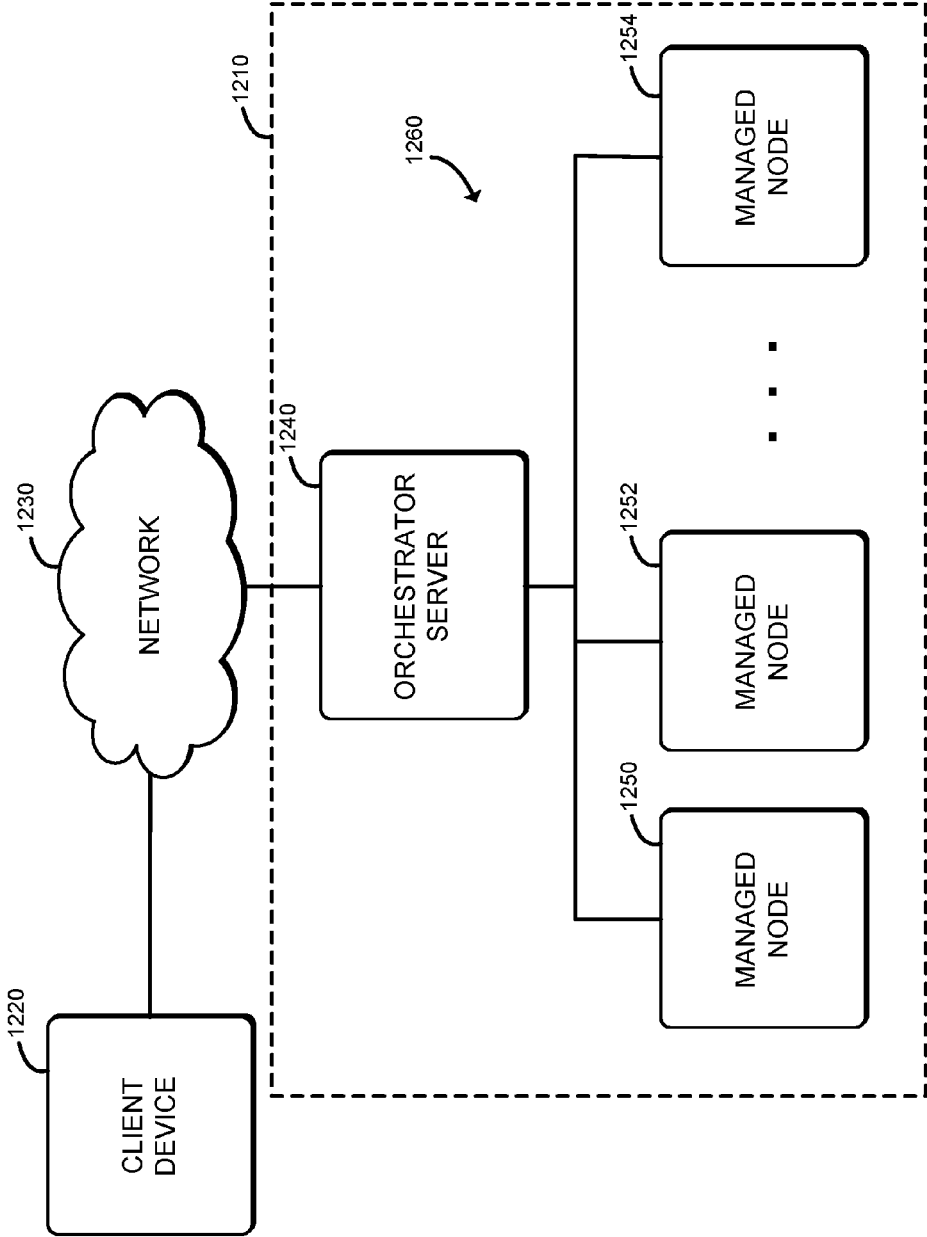


FIG. 12

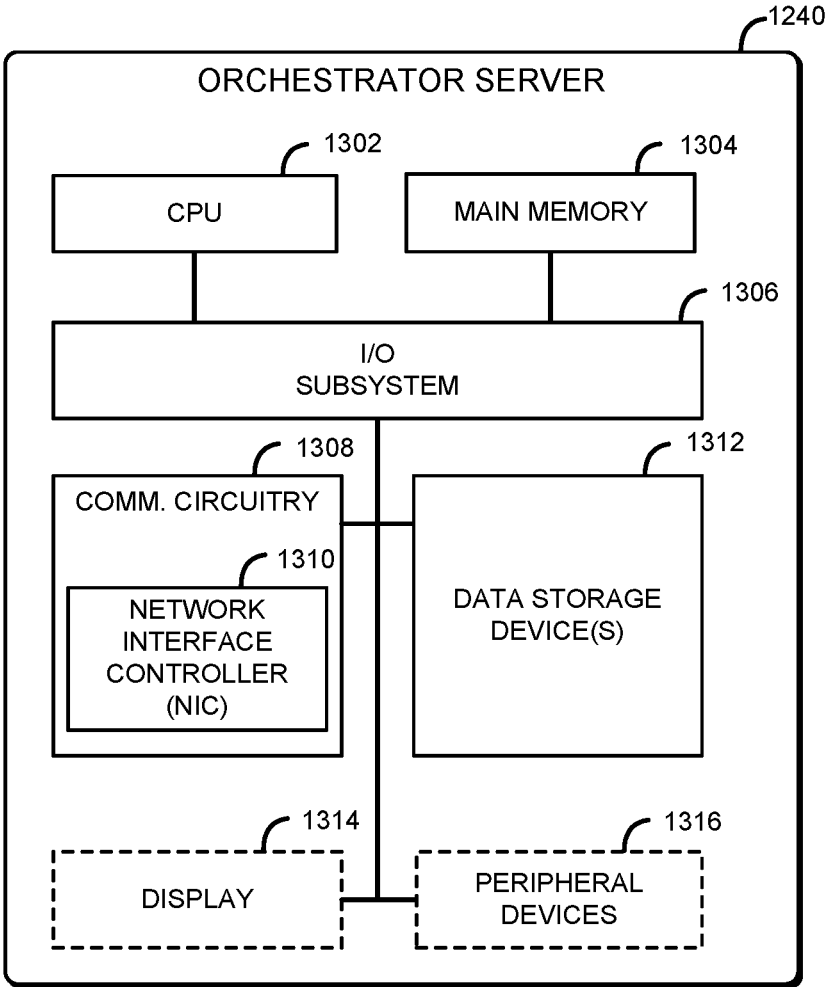


FIG. 13

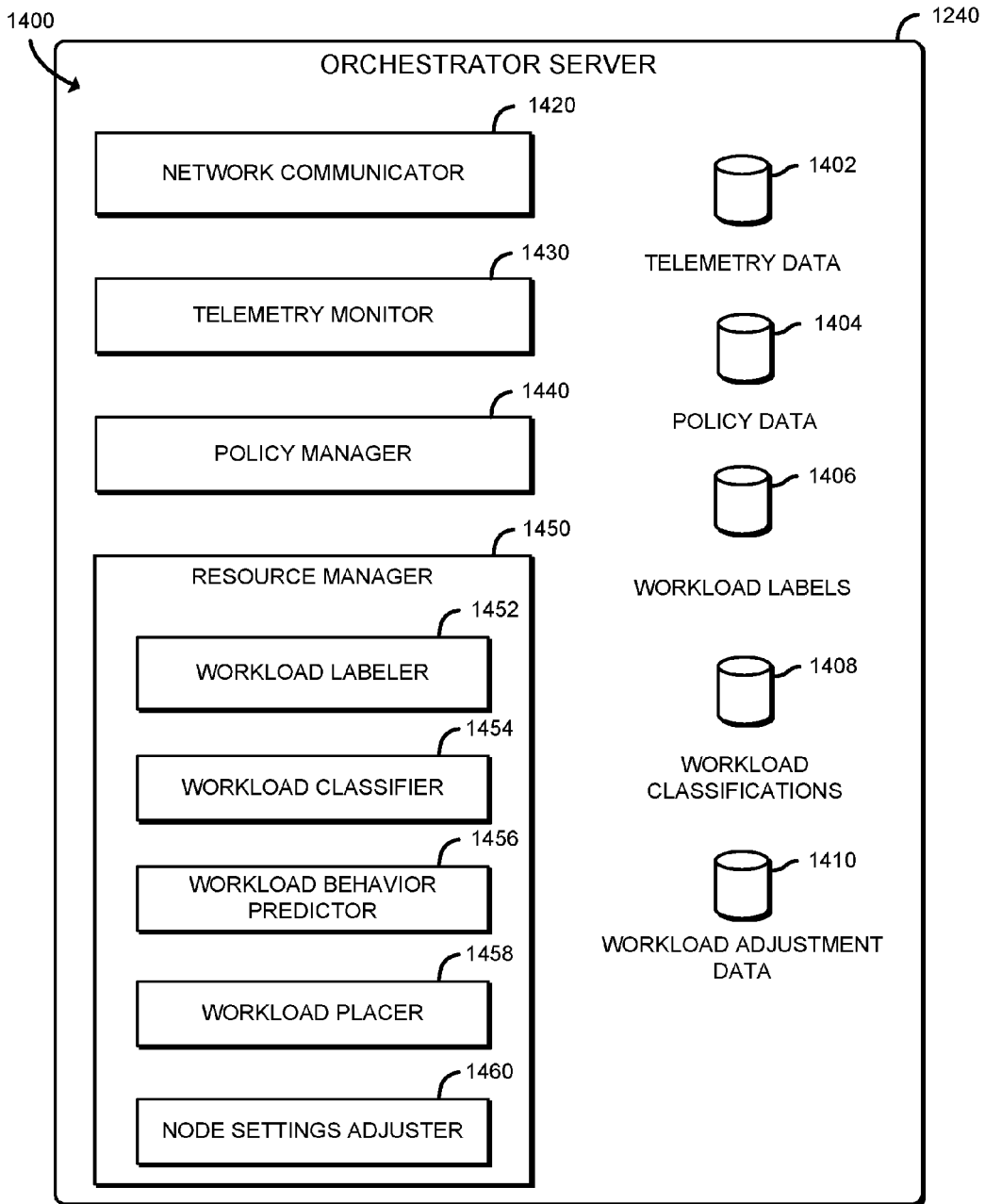


FIG. 14

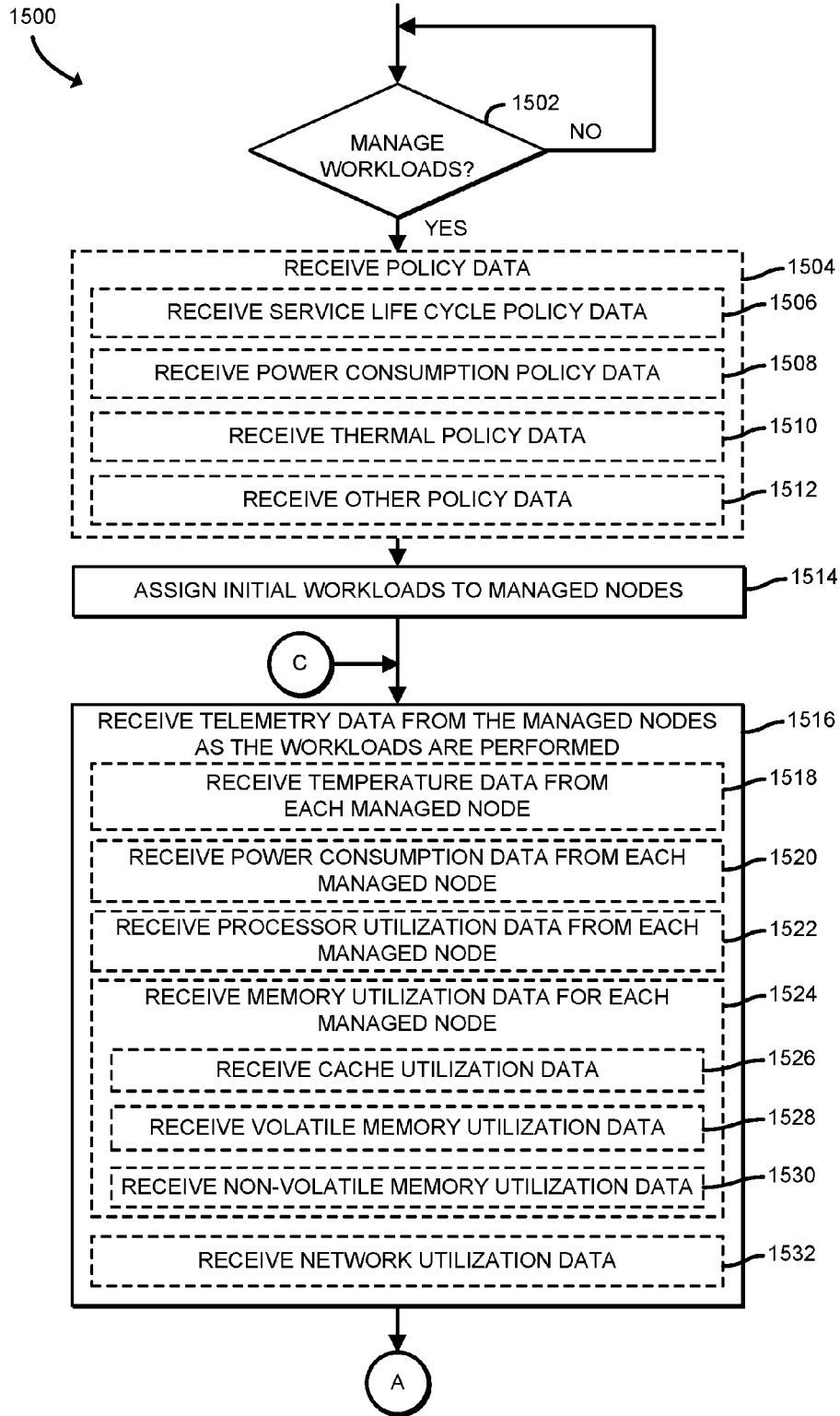


FIG. 15

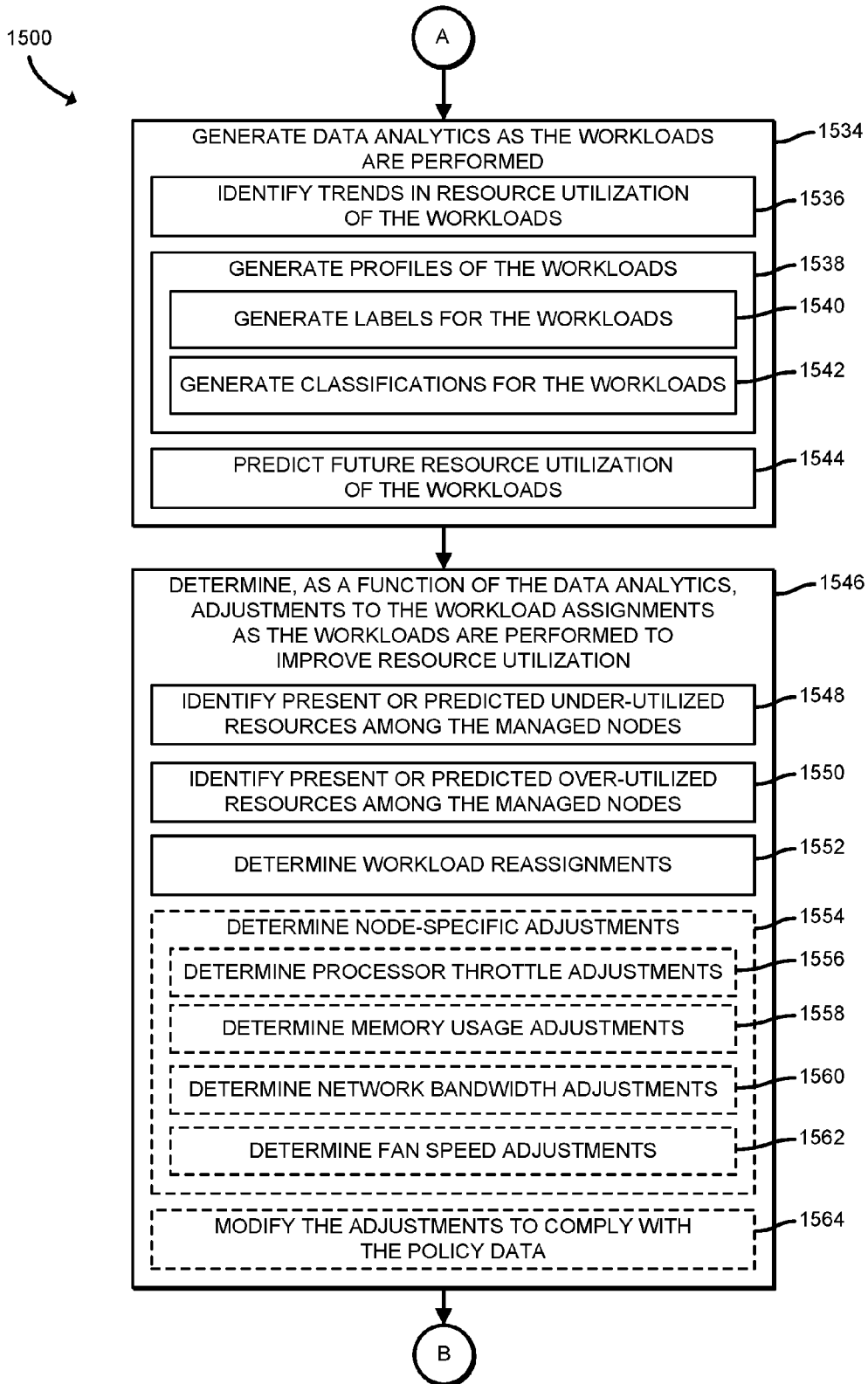


FIG. 16

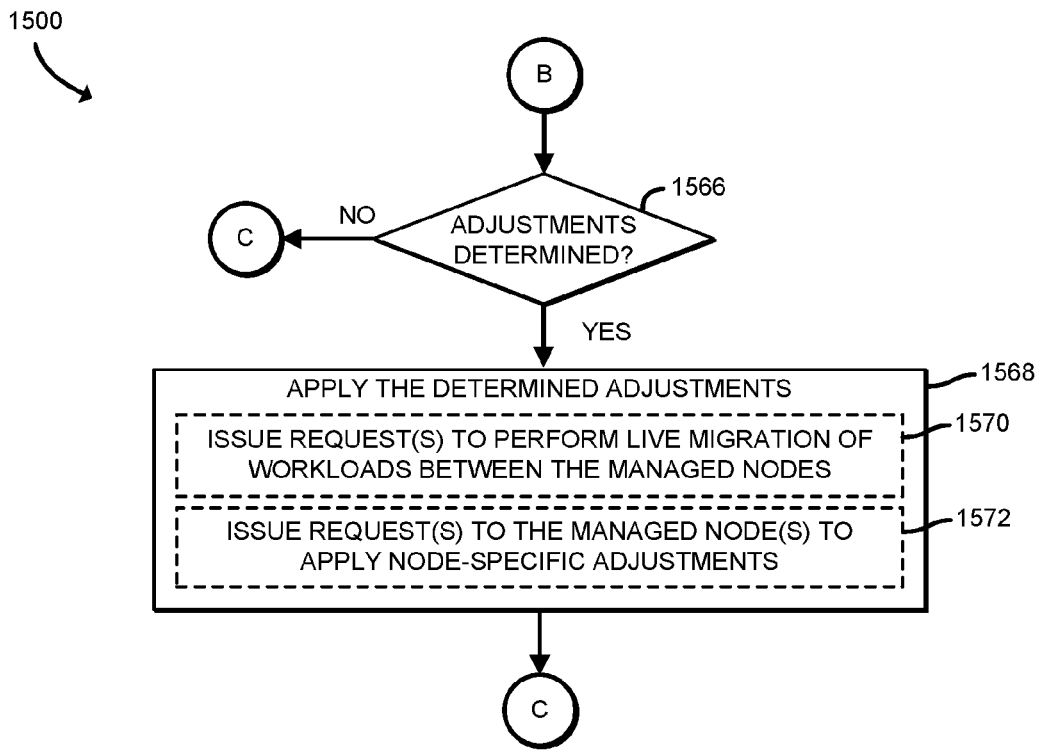


FIG. 17

TECHNOLOGIES FOR PERFORMING ORCHESTRATION WITH ONLINE ANALYTICS OF TELEMETRY DATA

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] The present application claims the benefit of U.S. Provisional Patent Application No. 62/365,969, filed Jul. 22, 2016, U.S. Provisional Patent Application No. 62/376,859, filed Aug. 18, 2016, and U.S. Provisional Patent Application No. 62/427,268, filed Nov. 29, 2016.

BACKGROUND

[0002] In cloud based computing environments, a server may be tasked with assigning workloads to various compute nodes in a network to provide services to one or more clients. In some environments, the server may receive logs from one or more of the compute nodes and analyze the logs after the workloads have been completed to determine information regarding the resource utilization (e.g., processor usage, volatile memory usage, non-volatile memory usage, network bandwidth usage) of the workloads. As an example, the log files may indicate that the processing capacity of one of the compute nodes was only slightly used while the processor of another compute node was overloaded during the same time period. After such an analysis, the server, or a human administrator, may determine new rules for assigning the workloads among the compute nodes in the future, to make better use of the resources available on each compute node. As such, the analysis and revised rules are effective only after a particular set of workloads have been executed to completion, or until execution has otherwise been halted.

BRIEF DESCRIPTION OF THE DRAWINGS

[0003] The concepts described herein are illustrated by way of example and not by way of limitation in the accompanying figures. For simplicity and clarity of illustration, elements illustrated in the figures are not necessarily drawn to scale. Where considered appropriate, reference labels have been repeated among the figures to indicate corresponding or analogous elements.

[0004] FIG. 1 is a diagram of a conceptual overview of a data center in which one or more techniques described herein may be implemented according to various embodiments;

[0005] FIG. 2 is a diagram of an example embodiment of a logical configuration of a rack of the data center of FIG. 1;

[0006] FIG. 3 is a diagram of an example embodiment of another data center in which one or more techniques described herein may be implemented according to various embodiments;

[0007] FIG. 4 is a diagram of another example embodiment of a data center in which one or more techniques described herein may be implemented according to various embodiments;

[0008] FIG. 5 is a diagram of a connectivity scheme representative of link-layer connectivity that may be established among various sleds of the data centers of FIGS. 1, 3, and 4;

[0009] FIG. 6 is a diagram of a rack architecture that may be representative of an architecture of any particular one of the racks depicted in FIGS. 1-4 according to some embodiments;

[0010] FIG. 7 is a diagram of an example embodiment of a sled that may be used with the rack architecture of FIG. 6;

[0011] FIG. 8 is a diagram of an example embodiment of a rack architecture to provide support for sleds featuring expansion capabilities;

[0012] FIG. 9 is a diagram of an example embodiment of a rack implemented according to the rack architecture of FIG. 8;

[0013] FIG. 10 is a diagram of an example embodiment of a sled designed for use in conjunction with the rack of FIG. 9;

[0014] FIG. 11 is a diagram of an example embodiment of a data center in which one or more techniques described herein may be implemented according to various embodiments;

[0015] FIG. 12 is a simplified block diagram of at least one embodiment of a system for performing orchestration with online analytics of telemetry data;

[0016] FIG. 13 is a simplified block diagram of at least one embodiment of an orchestrator server of the system of FIG. 12;

[0017] FIG. 14 is a simplified block diagram of at least one embodiment of an environment that may be established by an orchestrator server of FIGS. 12 and 13; and

[0018] FIGS. 15-17 are a simplified flow diagram of at least one embodiment of a method for managing workloads with online data analytics that may be performed by the orchestrator server of FIGS. 12-14.

DETAILED DESCRIPTION OF THE DRAWINGS

[0019] While the concepts of the present disclosure are susceptible to various modifications and alternative forms, specific embodiments thereof have been shown by way of example in the drawings and will be described herein in detail. It should be understood, however, that there is no intent to limit the concepts of the present disclosure to the particular forms disclosed, but on the contrary, the intention is to cover all modifications, equivalents, and alternatives consistent with the present disclosure and the appended claims.

[0020] References in the specification to “one embodiment,” “an embodiment,” “an illustrative embodiment,” etc., indicate that the embodiment described may include a particular feature, structure, or characteristic, but every embodiment may or may not necessarily include that particular feature, structure, or characteristic. Moreover, such phrases are not necessarily referring to the same embodiment. Further, when a particular feature, structure, or characteristic is described in connection with an embodiment, it is submitted that it is within the knowledge of one skilled in the art to effect such feature, structure, or characteristic in connection with other embodiments whether or not explicitly described. Additionally, it should be appreciated that items included in a list in the form of “at least one A, B, and C” can mean (A); (B); (C); (A and B); (A and C); (B and C); or (A, B, and C). Similarly, items listed in the form of “at least one of A, B, or C” can mean (A); (B); (C); (A and B); (A and C); (B and C); or (A, B, and C).

[0021] The disclosed embodiments may be implemented, in some cases, in hardware, firmware, software, or any

combination thereof. The disclosed embodiments may also be implemented as instructions carried by or stored on a transitory or non-transitory machine-readable (e.g., computer-readable) storage medium, which may be read and executed by one or more processors. A machine-readable storage medium may be embodied as any storage device, mechanism, or other physical structure for storing or transmitting information in a form readable by a machine (e.g., a volatile or non-volatile memory, a media disc, or other media device).

[0022] In the drawings, some structural or method features may be shown in specific arrangements and/or orderings. However, it should be appreciated that such specific arrangements and/or orderings may not be required. Rather, in some embodiments, such features may be arranged in a different manner and/or order than shown in the illustrative figures. Additionally, the inclusion of a structural or method feature in a particular figure is not meant to imply that such feature is required in all embodiments and, in some embodiments, may not be included or may be combined with other features.

[0023] FIG. 1 illustrates a conceptual overview of a data center 100 that may generally be representative of a data center or other type of computing network in/for which one or more techniques described herein may be implemented according to various embodiments. As shown in FIG. 1, data center 100 may generally contain a plurality of racks, each of which may house computing equipment comprising a respective set of physical resources. In the particular non-limiting example depicted in FIG. 1, data center 100 contains four racks 102A to 102D, which house computing equipment comprising respective sets of physical resources (PCRs) 105A to 105D. According to this example, a collective set of physical resources 106 of data center 100 includes the various sets of physical resources 105A to 105D that are distributed among racks 102A to 102D. Physical resources 106 may include resources of multiple types, such as—for example—processors, co-processors, accelerators, field-programmable gate arrays (FPGAs), memory, and storage. The embodiments are not limited to these examples.

[0024] The illustrative data center 100 differs from typical data centers in many ways. For example, in the illustrative embodiment, the circuit boards (“sleds”) on which components such as CPUs, memory, and other components are placed are designed for increased thermal performance. In particular, in the illustrative embodiment, the sleds are shallower than typical boards. In other words, the sleds are shorter from the front to the back, where cooling fans are located. This decreases the length of the path that air must travel across the components on the board. Further, the components on the sled are spaced further apart than in typical circuit boards, and the components are arranged to reduce or eliminate shadowing (i.e., one component in the air flow path of another component). In the illustrative embodiment, processing components such as the processors are located on a top side of a sled while near memory, such as DIMMs, are located on a bottom side of the sled. As a result of the enhanced airflow provided by this design, the components may operate at higher frequencies and power levels than in typical systems, thereby increasing performance. Furthermore, the sleds are configured to blindly mate with power and data communication cables in each rack 102A, 102B, 102C, 102D, enhancing their ability to be quickly removed, upgraded, reinstalled, and/or replaced.

Similarly, individual components located on the sleds, such as processors, accelerators, memory, and data storage drives, are configured to be easily upgraded due to their increased spacing from each other. In the illustrative embodiment, the components additionally include hardware attestation features to prove their authenticity.

[0025] Furthermore, in the illustrative embodiment, the data center 100 utilizes a single network architecture (“fabric”) that supports multiple other network architectures including Ethernet and Omni-Path. The sleds, in the illustrative embodiment, are coupled to switches via optical fibers, which provide higher bandwidth and lower latency than typical twister pair cabling (e.g., Category 5, Category 5e, Category 6, etc.). Due to the high bandwidth, low latency interconnections and network architecture, the data center 100 may, in use, pool resources, such as memory, accelerators (e.g., graphics accelerators, FPGAs, ASICs, etc.), and data storage drives that are physically disaggregated, and provide them to compute resources (e.g., processors) on an as needed basis, enabling the compute resources to access the pooled resources as if they were local. The illustrative data center 100 additionally receives usage information for the various resources, predicts resource usage for different types of workloads based on past resource usage, and dynamically reallocates the resources based on this information.

[0026] The racks 102A, 102B, 102C, 102D of the data center 100 may include physical design features that facilitate the automation of a variety of types of maintenance tasks. For example, data center 100 may be implemented using racks that are designed to be robotically-accessed, and to accept and house robotically-manipulatable resource sleds. Furthermore, in the illustrative embodiment, the racks 102A, 102B, 102C, 102D include integrated power sources that receive a greater voltage than is typical for power sources. The increased voltage enables the power sources to provide additional power to the components on each sled, enabling the components to operate at higher than typical frequencies.

[0027] FIG. 2 illustrates an exemplary logical configuration of a rack 202 of the data center 100. As shown in FIG. 2, rack 202 may generally house a plurality of sleds, each of which may comprise a respective set of physical resources. In the particular non-limiting example depicted in FIG. 2, rack 202 houses sleds 204-1 to 204-4 comprising respective sets of physical resources 205-1 to 205-4, each of which constitutes a portion of the collective set of physical resources 206 comprised in rack 202. With respect to FIG. 1, if rack 202 is representative of—for example—rack 102A, then physical resources 206 may correspond to the physical resources 105A comprised in rack 102A. In the context of this example, physical resources 105A may thus be made up of the respective sets of physical resources, including physical storage resources 205-1, physical accelerator resources 205-2, physical memory resources 204-3, and physical compute resources 205-5 comprised in the sleds 204-1 to 204-4 of rack 202. The embodiments are not limited to this example. Each sled may contain a pool of each of the various types of physical resources (e.g., compute, memory, accelerator, storage). By having robotically accessible and robotically manipulatable sleds comprising disaggregated resources, each type of resource can be upgraded independently of each other and at their own optimized refresh rate.

[0028] FIG. 3 illustrates an example of a data center 300 that may generally be representative of one in/for which one or more techniques described herein may be implemented according to various embodiments. In the particular non-limiting example depicted in FIG. 3, data center 300 comprises racks 302-1 to 302-32. In various embodiments, the racks of data center 300 may be arranged in such fashion as to define and/or accommodate various access pathways. For example, as shown in FIG. 3, the racks of data center 300 may be arranged in such fashion as to define and/or accommodate access pathways 311A, 311B, 311C, and 311D. In some embodiments, the presence of such access pathways may generally enable automated maintenance equipment, such as robotic maintenance equipment, to physically access the computing equipment housed in the various racks of data center 300 and perform automated maintenance tasks (e.g., replace a failed sled, upgrade a sled). In various embodiments, the dimensions of access pathways 311A, 311B, 311C, and 311D, the dimensions of racks 302-1 to 302-32, and/or one or more other aspects of the physical layout of data center 300 may be selected to facilitate such automated operations. The embodiments are not limited in this context.

[0029] FIG. 4 illustrates an example of a data center 400 that may generally be representative of one in/for which one or more techniques described herein may be implemented according to various embodiments. As shown in FIG. 4, data center 400 may feature an optical fabric 412. Optical fabric 412 may generally comprise a combination of optical signaling media (such as optical cabling) and optical switching infrastructure via which any particular sled in data center 400 can send signals to (and receive signals from) each of the other sleds in data center 400. The signaling connectivity that optical fabric 412 provides to any given sled may include connectivity both to other sleds in a same rack and sleds in other racks. In the particular non-limiting example depicted in FIG. 4, data center 400 includes four racks 402A to 402D. Racks 402A to 402D house respective pairs of sleds 404A-1 and 404A-2, 404B-1 and 404B-2, 404C-1 and 404C-2, and 404D-1 and 404D-2. Thus, in this example, data center 400 comprises a total of eight sleds. Via optical fabric 412, each such sled may possess signaling connectivity with each of the seven other sleds in data center 400. For example, via optical fabric 412, sled 404A-1 in rack 402A may possess signaling connectivity with sled 404A-2 in rack 402A, as well as the six other sleds 404B-1, 404B-2, 404C-1, 404C-2, 404D-1, and 404D-2 that are distributed among the other racks 402B, 402C, and 402D of data center 400. The embodiments are not limited to this example.

[0030] FIG. 5 illustrates an overview of a connectivity scheme 500 that may generally be representative of link-layer connectivity that may be established in some embodiments among the various sleds of a data center, such as any of example data centers 100, 300, and 400 of FIGS. 1, 3, and 4. Connectivity scheme 500 may be implemented using an optical fabric that features a dual-mode optical switching infrastructure 514. Dual-mode optical switching infrastructure 514 may generally comprise a switching infrastructure that is capable of receiving communications according to multiple link-layer protocols via a same unified set of optical signaling media, and properly switching such communications. In various embodiments, dual-mode optical switching infrastructure 514 may be implemented using one or more dual-mode optical switches 515. In various embodiments, dual-mode optical switches 515 may generally comprise

high-radix switches. In some embodiments, dual-mode optical switches 515 may comprise multi-ply switches, such as four-ply switches. In various embodiments, dual-mode optical switches 515 may feature integrated silicon photonics that enable them to switch communications with significantly reduced latency in comparison to conventional switching devices. In some embodiments, dual-mode optical switches 515 may constitute leaf switches 530 in a leaf-spine architecture additionally including one or more dual-mode optical spine switches 520.

[0031] In various embodiments, dual-mode optical switches may be capable of receiving both Ethernet protocol communications carrying Internet Protocol (IP packets) and communications according to a second, high-performance computing (HPC) link-layer protocol (e.g., Intel's Omni-Path Architecture's, Infiniband) via optical signaling media of an optical fabric. As reflected in FIG. 5, with respect to any particular pair of sleds 504A and 504B possessing optical signaling connectivity to the optical fabric, connectivity scheme 500 may thus provide support for link-layer connectivity via both Ethernet links and HPC links. Thus, both Ethernet and HPC communications can be supported by a single high-bandwidth, low-latency switch fabric. The embodiments are not limited to this example.

[0032] FIG. 6 illustrates a general overview of a rack architecture 600 that may be representative of an architecture of any particular one of the racks depicted in FIGS. 1 to 4 according to some embodiments. As reflected in FIG. 6, rack architecture 600 may generally feature a plurality of sled spaces into which sleds may be inserted, each of which may be robotically-accessible via a rack access region 601. In the particular non-limiting example depicted in FIG. 6, rack architecture 600 features five sled spaces 603-1 to 603-5. Sled spaces 603-1 to 603-5 feature respective multipurpose connector modules (MPCMs) 616-1 to 616-5.

[0033] FIG. 7 illustrates an example of a sled 704 that may be representative of a sled of such a type. As shown in FIG. 7, sled 704 may comprise a set of physical resources 705, as well as an MPCM 716 designed to couple with a counterpart MPCM when sled 704 is inserted into a sled space such as any of sled spaces 603-1 to 603-5 of FIG. 6. Sled 704 may also feature an expansion connector 717. Expansion connector 717 may generally comprise a socket, slot, or other type of connection element that is capable of accepting one or more types of expansion modules, such as an expansion sled 718. By coupling with a counterpart connector on expansion sled 718, expansion connector 717 may provide physical resources 705 with access to supplemental computing resources 705B residing on expansion sled 718. The embodiments are not limited in this context.

[0034] FIG. 8 illustrates an example of a rack architecture 800 that may be representative of a rack architecture that may be implemented in order to provide support for sleds featuring expansion capabilities, such as sled 704 of FIG. 7. In the particular non-limiting example depicted in FIG. 8, rack architecture 800 includes seven sled spaces 803-1 to 803-7, which feature respective MPCMs 816-1 to 816-7. Sled spaces 803-1 to 803-7 include respective primary regions 803-1A to 803-7A and respective expansion regions 803-1B to 803-7B. With respect to each such sled space, when the corresponding MPCM is coupled with a counterpart MPCM of an inserted sled, the primary region may generally constitute a region of the sled space that physically accommodates the inserted sled. The expansion region may

generally constitute a region of the sled space that can physically accommodate an expansion module, such as expansion sled **718** of FIG. 7, in the event that the inserted sled is configured with such a module.

[0035] FIG. 9 illustrates an example of a rack **902** that may be representative of a rack implemented according to rack architecture **800** of FIG. 8 according to some embodiments. In the particular non-limiting example depicted in FIG. 9, rack **902** features seven sled spaces **903-1** to **903-7**, which include respective primary regions **903-1A** to **903-7A** and respective expansion regions **903-1B** to **903-7B**. In various embodiments, temperature control in rack **902** may be implemented using an air cooling system. For example, as reflected in FIG. 9, rack **902** may feature a plurality of fans **919** that are generally arranged to provide air cooling within the various sled spaces **903-1** to **903-7**. In some embodiments, the height of the sled space is greater than the conventional “1U” server height. In such embodiments, fans **919** may generally comprise relatively slow, large diameter cooling fans as compared to fans used in conventional rack configurations. Running larger diameter cooling fans at lower speeds may increase fan lifetime relative to smaller diameter cooling fans running at higher speeds while still providing the same amount of cooling. The sleds are physically shallower than conventional rack dimensions. Further, components are arranged on each sled to reduce thermal shadowing (i.e., not arranged serially in the direction of air flow). As a result, the wider, shallower sleds allow for an increase in device performance because the devices can be operated at a higher thermal envelope (e.g., 250 W) due to improved cooling (i.e., no thermal shadowing, more space between devices, more room for larger heat sinks, etc.).

[0036] MPCMs **916-1** to **916-7** may be configured to provide inserted sleds with access to power sourced by respective power modules **920-1** to **920-7**, each of which may draw power from an external power source **921**. In various embodiments, external power source **921** may deliver alternating current (AC) power to rack **902**, and power modules **920-1** to **920-7** may be configured to convert such AC power to direct current (DC) power to be sourced to inserted sleds. In some embodiments, for example, power modules **920-1** to **920-7** may be configured to convert 277-volt AC power into 12-volt DC power for provision to inserted sleds via respective MPCMs **916-1** to **916-7**. The embodiments are not limited to this example.

[0037] MPCMs **916-1** to **916-7** may also be arranged to provide inserted sleds with optical signaling connectivity to a dual-mode optical switching infrastructure **914**, which may be the same as—or similar to—dual-mode optical switching infrastructure **514** of FIG. 5. In various embodiments, optical connectors contained in MPCMs **916-1** to **916-7** may be designed to couple with counterpart optical connectors contained in MPCMs of inserted sleds to provide such sleds with optical signaling connectivity to dual-mode optical switching infrastructure **914** via respective lengths of optical cabling **922-1** to **922-7**. In some embodiments, each such length of optical cabling may extend from its corresponding MPCM to an optical interconnect loom **923** that is external to the sled spaces of rack **902**. In various embodiments, optical interconnect loom **923** may be arranged to pass through a support post or other type of load-bearing element of rack **902**. The embodiments are not limited in this context. Because inserted sleds connect to an optical switching infrastructure via MPCMs, the resources typically spent

in manually configuring the rack cabling to accommodate a newly inserted sled can be saved.

[0038] FIG. 10 illustrates an example of a sled **1004** that may be representative of a sled designed for use in conjunction with rack **902** of FIG. 9 according to some embodiments. Sled **1004** may feature an MPCM **1016** that comprises an optical connector **1016A** and a power connector **1016B**, and that is designed to couple with a counterpart MPCM of a sled space in conjunction with insertion of MPCM **1016** into that sled space. Coupling MPCM **1016** with such a counterpart MPCM may cause power connector **1016** to couple with a power connector comprised in the counterpart MPCM. This may generally enable physical resources **1005** of sled **1004** to source power from an external source, via power connector **1016** and power transmission media **1024** that conductively couples power connector **1016** to physical resources **1005**.

[0039] Sled **1004** may also include dual-mode optical network interface circuitry **1026**. Dual-mode optical network interface circuitry **1026** may generally comprise circuitry that is capable of communicating over optical signaling media according to each of multiple link-layer protocols supported by dual-mode optical switching infrastructure **914** of FIG. 9. In some embodiments, dual-mode optical network interface circuitry **1026** may be capable both of Ethernet protocol communications and of communications according to a second, high-performance protocol. In various embodiments, dual-mode optical network interface circuitry **1026** may include one or more optical transceiver modules **1027**, each of which may be capable of transmitting and receiving optical signals over each of one or more optical channels. The embodiments are not limited in this context.

[0040] Coupling MPCM **1016** with a counterpart MPCM of a sled space in a given rack may cause optical connector **1016A** to couple with an optical connector comprised in the counterpart MPCM. This may generally establish optical connectivity between optical cabling of the sled and dual-mode optical network interface circuitry **1026**, via each of a set of optical channels **1025**. Dual-mode optical network interface circuitry **1026** may communicate with the physical resources **1005** of sled **1004** via electrical signaling media **1028**. In addition to the dimensions of the sleds and arrangement of components on the sleds to provide improved cooling and enable operation at a relatively higher thermal envelope (e.g., 250 W), as described above with reference to FIG. 9, in some embodiments, a sled may include one or more additional features to facilitate air cooling, such as a heatpipe and/or heat sinks arranged to dissipate heat generated by physical resources **1005**. It is worthy of note that although the example sled **1004** depicted in FIG. 10 does not feature an expansion connector, any given sled that features the design elements of sled **1004** may also feature an expansion connector according to some embodiments. The embodiments are not limited in this context.

[0041] FIG. 11 illustrates an example of a data center **1100** that may generally be representative of one in/for which one or more techniques described herein may be implemented according to various embodiments. As reflected in FIG. 11, a physical infrastructure management framework **1150A** may be implemented to facilitate management of a physical infrastructure **1100A** of data center **1100**. In various embodiments, one function of physical infrastructure management framework **1150A** may be to manage automated maintenance functions within data center **1100**, such as the use of

robotic maintenance equipment to service computing equipment within physical infrastructure 1100A. In some embodiments, physical infrastructure 1100A may feature an advanced telemetry system that performs telemetry reporting that is sufficiently robust to support remote automated management of physical infrastructure 1100A. In various embodiments, telemetry information provided by such an advanced telemetry system may support features such as failure prediction/prevention capabilities and capacity planning capabilities. In some embodiments, physical infrastructure management framework 1150A may also be configured to manage authentication of physical infrastructure components using hardware attestation techniques. For example, robots may verify the authenticity of components before installation by analyzing information collected from a radio frequency identification (RFID) tag associated with each component to be installed. The embodiments are not limited in this context.

[0042] As shown in FIG. 11, the physical infrastructure 1100A of data center 1100 may comprise an optical fabric 1112, which may include a dual-mode optical switching infrastructure 1114. Optical fabric 1112 and dual-mode optical switching infrastructure 1114 may be the same as—or similar to—optical fabric 412 of FIG. 4 and dual-mode optical switching infrastructure 514 of FIG. 5, respectively, and may provide high-bandwidth, low-latency, multi-protocol connectivity among sleds of data center 1100. As discussed above, with reference to FIG. 1, in various embodiments, the availability of such connectivity may make it feasible to disaggregate and dynamically pool resources such as accelerators, memory, and storage. In some embodiments, for example, one or more pooled accelerator sleds 1130 may be included among the physical infrastructure 1100A of data center 1100, each of which may comprise a pool of accelerator resources—such as co-processors and/or FPGAs, for example—that is globally accessible to other sleds via optical fabric 1112 and dual-mode optical switching infrastructure 1114.

[0043] In another example, in various embodiments, one or more pooled storage sleds 1132 may be included among the physical infrastructure 1100A of data center 1100, each of which may comprise a pool of storage resources that is available globally accessible to other sleds via optical fabric 1112 and dual-mode optical switching infrastructure 1114. In some embodiments, such pooled storage sleds 1132 may comprise pools of solid-state storage devices such as solid-state drives (SSDs). In various embodiments, one or more high-performance processing sleds 1134 may be included among the physical infrastructure 1100A of data center 1100. In some embodiments, high-performance processing sleds 1134 may comprise pools of high-performance processors, as well as cooling features that enhance air cooling to yield a higher thermal envelope of up to 250 W or more. In various embodiments, any given high-performance processing sled 1134 may feature an expansion connector 1117 that can accept a far memory expansion sled, such that the far memory that is locally available to that high-performance processing sled 1134 is disaggregated from the processors and near memory comprised on that sled. In some embodiments, such a high-performance processing sled 1134 may be configured with far memory using an expansion sled that comprises low-latency SSD storage. The optical infrastructure allows for compute resources on one sled to utilize remote accelerator/FPGA, memory, and/or SSD resources

that are disaggregated on a sled located on the same rack or any other rack in the data center. The remote resources can be located one switch jump away or two-switch jumps away in the spine-leaf network architecture described above with reference to FIG. 5. The embodiments are not limited in this context.

[0044] In various embodiments, one or more layers of abstraction may be applied to the physical resources of physical infrastructure 1100A in order to define a virtual infrastructure, such as a software-defined infrastructure 1100B. In some embodiments, virtual computing resources 1136 of software-defined infrastructure 1100B may be allocated to support the provision of cloud services 1140. In various embodiments, particular sets of virtual computing resources 1136 may be grouped for provision to cloud services 1140 in the form of SDI services 1138. Examples of cloud services 1140 may include—without limitation—software as a service (SaaS) services 1142, platform as a service (PaaS) services 1144, and infrastructure as a service (IaaS) services 1146.

[0045] In some embodiments, management of software-defined infrastructure 1100B may be conducted using a virtual infrastructure management framework 1150B. In various embodiments, virtual infrastructure management framework 1150B may be designed to implement workload fingerprinting techniques and/or machine-learning techniques in conjunction with managing allocation of virtual computing resources 1136 and/or SDI services 1138 to cloud services 1140. In some embodiments, virtual infrastructure management framework 1150B may use/consult telemetry data in conjunction with performing such resource allocation. In various embodiments, an application/service management framework 1150C may be implemented in order to provide QoS management capabilities for cloud services 1140. The embodiments are not limited in this context.

[0046] As shown in FIG. 12, an illustrative system 1210 for performing orchestration with online analytics of telemetry data includes an orchestrator server 1240 in communication with a set of managed nodes 1260. Each managed node 1260 may be embodied as an assembly of resources (e.g., physical resources 206), such as compute resources (e.g., physical compute resources 205-4), storage resources (e.g., physical storage resources 205-1), accelerator resources (e.g., physical accelerator resources 205-2), or other resources (e.g., physical memory resources 205-3) from the same or different sleds (e.g., the sleds 204-1, 204-2, 204-3, 204-4, etc.) or racks (e.g., one or more of racks 302-1 through 302-32). Each managed node 1260 may be established, defined, or “spun up” by the orchestrator server 1240 at the time a workload is to be assigned to the managed node 1260 or at any other time, and may exist regardless of whether any workloads are presently assigned to the managed node 1260. The system 1210 may be implemented in accordance with the data centers 100, 300, 400, 1100 described above with reference to FIGS. 1, 3, 4, and 11. In the illustrative embodiment, the set of managed nodes 1260 includes managed nodes 1250, 1252, and 1254. While three managed nodes 1260 are shown in the set, it should be understood that in other embodiments, the set may include a different number of managed nodes 1260. The system 1210 may be located in a data center and provide storage and compute services (e.g., cloud services) to a client device 1220 that is in communication with the system 1210 through a network 1230. The orchestrator server 1240 may support

a cloud operating environment, such as OpenStack, and the managed nodes **1250** may execute one or more applications or processes (i.e., workloads), such as in virtual machines or containers, on behalf of a user of the client device **1220**. As discussed in more detail herein, the orchestrator server **1240**, in operation, is configured to assign workloads to the managed nodes **1260** and receive telemetry data, which may be embodied as data indicative of the performance and conditions of each managed node **1260** as the managed nodes **1260** execute the workloads assigned to them. The orchestrator server **1240** is further configured to generate analytics based on the received telemetry data to identify trends in the resource usage of the workloads (e.g., a phase in which a workload utilizes a relatively high amount of processor capacity and relatively low memory, followed by another phase in which the workload utilizes a relatively high amount of memory and a relatively low amount of processor capacity, etc.), label or otherwise identify each workload, classify the workloads into categories indicative of their resource usage trends, and predict future resource utilizations of each workload.

[0047] Using the analytics data, the orchestrator server **1240**, in the illustrative embodiment, determines and applies adjustments to the distribution of the workloads among the managed nodes **1260** to make better use of (e.g., approximate a target utilization value, such as 90% or more) the available resources in each managed node **1260**. In the illustrative embodiment, and in contrast to typical systems, the orchestrator server **1240** generates the analytics, determines the adjustments, and applies the adjustments as the workloads are being executed, rather than waiting until the workloads have been completed and applying any adjustments for a subsequent set of workloads. By doing so, the orchestrator server **1240** obtains more accurate and timely data as to how workload adjustments affect the utilization of the available resources in the managed nodes **1260** and increases resource utilization, thereby enabling the same set of managed nodes **1260** to perform more workloads than they otherwise would in the same amount of time. As described in more detail herein, the orchestrator server **1240** may determine adjustments not only to optimize resource usage but also to comply with one or more policies indicative of a business's preferences as to the use of the managed nodes **1260**, such as heat production, power consumption, and/or wear on the components of the managed nodes **1260**.

[0048] Referring now to FIG. 13, the orchestrator server **1240** may be embodied as any type of compute device capable of performing the functions described herein, including issuing a request to have cloud services performed, receiving results of the cloud services, assigning workloads to compute devices, analyzing telemetry data indicative of performance and conditions (e.g., resource utilization, one or more temperatures, fan speeds, etc.) as the workloads are executed, and adjusting the assignments of the workloads to increase resource utilization as the workloads are performed. For example, the orchestrator server **1240** may be embodied as a computer, a distributed computing system, one or more sleds (e.g., the sleds **204-1**, **204-2**, **204-3**, **204-4**, etc.), a server (e.g., stand-alone, rack-mounted, blade, etc.), a multiprocessor system, a network appliance (e.g., physical or virtual), a desktop computer, a workstation, a laptop computer, a notebook computer, a processor-based system, or a network appliance. As shown in FIG. 13, the illustrative orchestrator server **1240** includes

a central processing unit (CPU) **1302**, a main memory **1304**, an input/output (I/O) subsystem **1306**, communication circuitry **1308**, and one or more data storage devices **1312**. Of course, in other embodiments, the orchestrator server **1240** may include other or additional components, such as those commonly found in a computer (e.g., display, peripheral devices, etc.). Additionally, in some embodiments, one or more of the illustrative components may be incorporated in, or otherwise form a portion of, another component. For example, in some embodiments, the main memory **1304**, or portions thereof, may be incorporated in the CPU **1302**.

[0049] The CPU **1302** may be embodied as any type of processor capable of performing the functions described herein. The CPU **1302** may be embodied as a single or multi-core processor(s), a microcontroller, or other processor or processing/controlling circuit. As discussed above, the managed node **1260** may include resources distributed across multiple sleds and in such embodiments, the CPU **1302** may include portions thereof located on the same sled or different sled. Similarly, the main memory **1304** may be embodied as any type of volatile (e.g., dynamic random access memory (DRAM), etc.) or non-volatile memory or data storage capable of performing the functions described herein. In some embodiments, the CPU **1302** may be embodied as, include, or be coupled to a field programmable gate array (FPGA), an application specific integrated circuit (ASIC), reconfigurable hardware or hardware circuitry, or other specialized hardware to facilitate performance of the functions described herein. Similarly, the main memory **1304** may be embodied as any type of volatile (e.g., dynamic random access memory (DRAM), etc.) or non-volatile memory or data storage capable of performing the functions described herein. In some embodiments, all or a portion of the main memory **1304** may be integrated into the CPU **1302**. In operation, the main memory **1304** may store various software and data used during operation such as telemetry data, policy data, workload labels, workload classifications, workload adjustment data, operating systems, applications, programs, libraries, and drivers. As discussed above, the managed node **1260** may include resources distributed across multiple sleds and in such embodiments, the main memory **1304** may include portions thereof located on the same sled or different sled.

[0050] The I/O subsystem **1306** may be embodied as circuitry and/or components to facilitate input/output operations with the CPU **1302**, the main memory **1304**, and other components of the orchestrator server **1240**. For example, the I/O subsystem **1306** may be embodied as, or otherwise include, memory controller hubs, input/output control hubs, integrated sensor hubs, firmware devices, communication links (e.g., point-to-point links, bus links, wires, cables, light guides, printed circuit board traces, etc.), and/or other components and subsystems to facilitate the input/output operations. In some embodiments, the I/O subsystem **1306** may form a portion of a system-on-a-chip (SoC) and be incorporated, along with one or more of the CPU **1302**, the main memory **1304**, and other components of the orchestrator server **1240**, on a single integrated circuit chip.

[0051] The communication circuitry **1308** may be embodied as any communication circuit, device, or collection thereof, capable of enabling communications over the network **1230** between the orchestrator server **1240** and another compute device (e.g., the client device **1220**, and/or the managed nodes **1260**). The communication circuitry **1308**

may be configured to use any one or more communication technology (e.g., wired or wireless communications) and associated protocols (e.g., Ethernet, Bluetooth®, Wi-Fi®, WiMAX, etc.) to effect such communication.

[0052] The illustrative communication circuitry **1308** includes a network interface controller (NIC) **1310**, which may also be referred to as a host fabric interface (HFI). The NIC **1310** may be embodied as one or more add-in-boards, daughtercards, network interface cards, controller chips, chipsets, or other devices that may be used by the orchestrator server **1240** to connect with another compute device (e.g., the client device **1220** and/or the managed nodes **1260**). In some embodiments, the NIC **1310** may be embodied as part of a system-on-a-chip (SoC) that includes one or more processors, or included on a multichip package that also contains one or more processors. In some embodiments, the NIC **1310** may include a local processor (not shown) and/or a local memory (not shown) that are both local to the NIC **1310**. In such embodiments, the local processor of the NIC **1310** may be capable of performing one or more of the functions of the CPU **1302** described herein. Additionally or alternatively, in such embodiments, the local memory of the NIC **1310** may be integrated into one or more components of the orchestrator server **1240** at the board level, socket level, chip level, and/or other levels. As discussed above, the managed node **1260** may include resources distributed across multiple sleds and in such embodiments, the communication circuitry **1308** may include portions thereof located on the same sled or different sled.

[0053] The one or more illustrative data storage devices **1312**, may be embodied as any type of devices configured for short-term or long-term storage of data such as, for example, memory devices and circuits, memory cards, hard disk drives, solid-state drives, or other data storage devices. Each data storage device **1312** may include a system partition that stores data and firmware code for the data storage device **1312**. Each data storage device **1312** may also include an operating system partition that stores data files and executables for an operating system.

[0054] Additionally, the orchestrator server **1240** may include a display **1314**. The display **1314** may be embodied as, or otherwise use, any suitable display technology including, for example, a liquid crystal display (LCD), a light emitting diode (LED) display, a cathode ray tube (CRT) display, a plasma display, and/or other display usable in a compute device. The display **1314** may include a touchscreen sensor that uses any suitable touchscreen input technology to detect the user's tactile selection of information displayed on the display including, but not limited to, resistive touchscreen sensors, capacitive touchscreen sensors, surface acoustic wave (SAW) touchscreen sensors, infrared touchscreen sensors, optical imaging touchscreen sensors, acoustic touchscreen sensors, and/or other type of touchscreen sensors.

[0055] Additionally or alternatively, the orchestrator server **1240** may include one or more peripheral devices **1316**. Such peripheral devices **1316** may include any type of peripheral device commonly found in a compute device such as speakers, a mouse, a keyboard, and/or other input/output devices, interface devices, and/or other peripheral devices.

[0056] The client device **1220** and the managed nodes **1260** may have components similar to those described in FIG. **13**. The description of those components of the orchestrator server **1240** is equally applicable to the description of

components of the client device **1220** and the managed nodes **1260** and is not repeated herein for clarity of the description. Further, it should be appreciated that any of the client device **1220** and the managed nodes **1260** may include other components, sub-components, and devices commonly found in a computing device, which are not discussed above in reference to the orchestrator server **1240** and not discussed herein for clarity of the description.

[0057] As described above, the client device **1220**, the orchestrator server **1240** and the managed nodes **1260** are illustratively in communication via the network **1230**, which may be embodied as any type of wired or wireless communication network, including global networks (e.g., the Internet), local area networks (LANs) or wide area networks (WANs), cellular networks (e.g., Global System for Mobile Communications (GSM), 3G, Long Term Evolution (LTE), Worldwide Interoperability for Microwave Access (WiMAX), etc.), digital subscriber line (DSL) networks, cable networks (e.g., coaxial networks, fiber networks, etc.), or any combination thereof.

[0058] Referring now to FIG. **14**, in the illustrative embodiment, the orchestrator server **1240** may establish an environment **1400** during operation. The illustrative environment **1400** includes a network communicator **1420**, a telemetry monitor **1430**, a policy manager **1440**, and a resource manager **1450**. Each of the components of the environment **1400** may be embodied as hardware, firmware, software, or a combination thereof. As such, in some embodiments, one or more of the components of the environment **1400** may be embodied as circuitry or a collection of electrical devices (e.g., network communicator circuitry **1420**, telemetry monitor circuitry **1430**, policy manager circuitry **1440**, resource manager circuitry **1450**, etc.). It should be appreciated that, in such embodiments, one or more of the network communicator circuitry **1420**, telemetry monitor circuitry **1430**, policy manager circuitry **1440**, or resource manager circuitry **1450** may form a portion of one or more of the CPU **1302**, the main memory **1304**, the I/O subsystem **1306**, and/or other components of the orchestrator server **1240**. In the illustrative embodiment, the environment **1400** includes telemetry data **1402** which may be embodied as data indicative of the performance and conditions (e.g., resource utilization, one or more temperatures, fan speeds, etc.) of each managed node **1260** as the managed nodes **1260** execute the workloads assigned to them. Additionally, the illustrative environment **1400** includes policy data **1404** indicative of user-defined preferences as to the heat production, power consumption, and life expectancy of the components of the managed nodes **1260**. Further, the illustrative environment **1400** includes workload labels **1406** which may be embodied as any identifiers (e.g., process numbers, executable file names, alphanumeric tags, etc.) that uniquely identify each workload executed by the managed nodes **1260**. In addition, the illustrative environment **1400** includes workload classifications **1408** which may be embodied as any data indicative of the resource utilization tendencies of each workload (e.g., processor intensive, network bandwidth intensive, etc.). Further, the illustrative environment **1400** includes workload adjustment data **1410** which may be embodied as any data indicative of reassignments (e.g., live migrations) of one or more workloads from one managed node **1260** to another managed node **1260** and/or adjustments to settings for components within each managed node **1260**, such as processor capacity (e.g., a

number of cores to be used, a clock speed, a percentage of available processor cycles, etc.) available to one or more workloads, memory resource capacity (e.g., amount of memory to be used and/or frequency of memory accesses to volatile memory and/or non-volatile memory) available to one or more workloads, and/or communication circuitry capacity available to one or more workloads.

[0059] In the illustrative environment 1400, the network communicator 1420, which may be embodied as hardware, firmware, software, virtualized hardware, emulated architecture, and/or a combination thereof as discussed above, is configured to facilitate inbound and outbound network communications (e.g., network traffic, network packets, network flows, etc.) to and from the orchestrator server 1240, respectively. To do so, the network communicator 1420 is configured to receive and process data packets from one system or computing device (e.g., the client device 1220) and to prepare and send data packets to another computing device or system (e.g., the managed nodes 1260). Accordingly, in some embodiments, at least a portion of the functionality of the network communicator 1420 may be performed by the communication circuitry 1308, and, in the illustrative embodiment, by the NIC 1310.

[0060] The telemetry monitor 1430, which may be embodied as hardware, firmware, software, virtualized hardware, emulated architecture, and/or a combination thereof as discussed above, is configured to collect the telemetry data 1402 from the managed nodes 1260 as the managed nodes 1260 execute the workloads assigned to them. The telemetry monitor 1430 may actively poll each of the managed nodes 1260 for updated telemetry data 1402 on an ongoing basis or may passively receive telemetry data 1402 from the managed nodes 1260, such as by listening on a particular network port for updated telemetry data 1402. The telemetry monitor 1430 may further parse and categorize the telemetry data 1402, such as by separating the telemetry data 1402 into an individual file or data set for each managed node 1260. The policy manager 1440, which may be embodied as hardware, firmware, software, virtualized hardware, emulated architecture, and/or a combination thereof as discussed above, is configured to receive and store the policy data 1404, which, as described above, is indicative of user-defined preferences as to the heat production, power consumption, and life expectancy (i.e., wear) of the components of the managed nodes 1260. The policy manager 1440 is further configured to provide the policy data 1404 to the resource manager 1450 to assist in determining adjustments to the assignment of workloads among the managed nodes 1260 and for adjusting settings within one or more of the managed nodes (e.g., processor capacity available to one or more workloads, memory resource capacity available to one or more workloads, and/or communication circuitry capacity available to one or more workloads) to optimize resource utilization, subject to the policies defined in the policy data 1404.

[0061] The resource manager 1450, which may be embodied as hardware, firmware, software, virtualized hardware, emulated architecture, and/or a combination thereof, is configured to generate data analytics from the telemetry data 1402, identify the workloads, classify the workloads, identify trends in the resource utilization of the workloads, predict future resource utilizations of the workloads, and adjust the assignments of the workloads to the managed nodes 1260 and the settings of the managed nodes 1260 to

increase the resource utilization (e.g., to reduce the amount of idle resources) while staying in compliance with the policy data 1404. To do so, the resource manager 1450 includes a workload labeler 1452, a workload classifier 1454, a workload behavior predictor 1456, a workload placer 1458, and a node settings adjuster 1460. The workload labeler 1452, in the illustrative embodiment, is configured to assign a workload label 1406 to each workload presently performed or scheduled to be performed by one or more of the managed nodes 1260. The workload labeler 1452 may generate the workload label as a function of an executable name of the workload, a hash of all or a portion of the code of the workload, or based on any other method to uniquely identify each workload. The workload classifier 1454, in the illustrative embodiment, is configured to categorize each labeled workload based on the resource utilization usage of each workload. For example, the workload classifier may categorize one set of labeled workloads as being consistently processor intensive, another set of labeled workloads as being consistently memory intensive, and another set of workloads as having phases of different resource utilization (high memory use and low processor use, followed by high processor use and low memory use, etc.).

[0062] The workload behavior predictor 1456, in the illustrative embodiment, is configured to analyze the telemetry data 1402 and the workload classification 1408 to predict future resource utilization needs of the various workloads based on their previous usage. In doing so, the workload behavior predictor 1456 may determine a present phase of a given workload and determine an amount of remaining time until the workload transitions to another phase having different resource utilization characteristics. The workload placer 1458, in the illustrative embodiment, is configured to initially assign workloads to the various managed nodes 1260 and determine, based on the telemetry data 1402, the workload classifications, and the policy data, whether the resources of the managed nodes 1260 could be more efficiently used (e.g., to reduce the amount of idle resources and to reduce the load on over-used resources) by reassigning the workloads among the managed nodes 1260, without violating the policies in the policy data (e.g., without generating more than a threshold amount of heat, without consuming more than a threshold amount of power, etc.). Similarly, the node settings adjuster 1460, in the illustrative embodiment, is configured to determine one or more adjustments to the settings within the managed nodes 1260 to provide or restrict the resources available to the workloads in accordance with the goal of optimizing resource usage and maintaining conformance with the policies in the policy data 1404. The settings may be associated with the operating system and/or the firmware or drivers of the components of the managed nodes 1260.

[0063] It should be appreciated that each of the workload labeler 1452, the workload classifier 1454, the workload behavior predictor 1456, the workload placer 1458, and the node settings adjuster 1460 may be separately embodied as hardware, firmware, software, virtualized hardware, emulated architecture, and/or a combination thereof. For example, the workload labeler 1452 may be embodied as a hardware component, while the workload classifier 1454, the workload behavior predictor 1456, the workload placer 1458, and the node settings adjuster 1460 are embodied as a virtualized hardware component or as some other combi-

nation of hardware, firmware, software, virtualized hardware, emulated architecture, and/or a combination thereof.

[0064] Referring now to FIG. 15, in use, the orchestrator server 1240 may execute a method 1500 for managing workloads with online data analytics. The method 1500 begins with block 1502, in which the orchestrator server 1240 determines whether to manage workloads using online data analytics. In the illustrative embodiment, the orchestrator server 1240 determines to manage workloads if the orchestrator server 1240 is powered on, in communication with the managed nodes 1260, and has received at least one request from the client device 1220 to provide cloud services (i.e., to perform one or more workloads). In other embodiments, the orchestrator server 1240 may determine whether to manage workloads based on other factors. Regardless, in response to a determination to manage workloads, in the illustrative embodiment, the method 1500 advances to block 1504 in which the orchestrator server 1240 receives policy data (e.g., the policy data 1404). In doing so, the orchestrator server 1240 may receive the policy data 1404 from a user (e.g., an administrator) through a graphical user interface (not shown), from a configuration file, or from another source. As indicated in block 1506, in receiving the policy data 1404, the orchestrator server 1240 may receive service life cycle policy data indicative of a target life cycle of one or more of the managed nodes 1260. Additionally or alternatively, as indicated in block 1508, the orchestrator server 1240 may receive power consumption policy data indicative of a target power usage or threshold amount of power usage of the managed nodes 1260 as they execute the workloads. As indicated in block 1510, the orchestrator server 1240 may additionally or alternatively receive thermal policy data indicative of a target temperature or a temperature threshold not to be exceeded by the managed nodes 1260 as they execute the workloads. Additionally or alternatively, as indicated in block 1512, the orchestrator server 1240 may receive other types of policy data indicative of thresholds or goals to be satisfied during the execution of the workloads.

[0065] After receiving the policy data 1404, in the illustrative embodiment, the method 1500 advances to block 1514 in which the orchestrator server 1240 assigns initial workloads to the managed nodes 1260. In the illustrative embodiment, the orchestrator server 1240 has not received telemetry data 1402 that would inform a decision as to where the workloads are to be assigned among the managed nodes 1260. As such, the orchestrator server 1240 may assign the workloads to the managed nodes 1260 based on any suitable method, such as assigning the each workload to the first available managed node that is idle (i.e., is not presently executing a workload), randomly assigning the workloads, or by any other method.

[0066] Having assigned the workloads, the method 1500 advances to block 1516 in which the orchestrator server 1240 receives telemetry data 1402 from the managed nodes 1260 as the workloads are performed (i.e., executed). In doing so, the orchestrator server 1240 may receive temperature data indicative of a temperature within each managed node 1260, as indicated in block 1518. Additionally or alternatively, the orchestrator server 1240 may receive power consumption data indicative of an amount of power consumed by each managed node 1260, as indicated in block 1520. In receiving the telemetry data 1402, the orchestrator server 1240 may additionally or alternatively receive processor utilization data indicative of an amount of processor

usage consumed by each workload performed by each managed node 1260, as indicated in block 1522. Moreover, as indicated in block 1524, the orchestrator server 1240 receive memory utilization data for each managed node 1260. The memory utilization data may be embodied as Intel Cache Allocation Technology (CAT) data, Intel Cache Monitoring Technology (CMT) data, Intel Memory Bandwidth Monitoring (MBM) data, and/or other data indicative of an amount or frequency of memory use by each workload performed by each managed node 1260.

[0067] In receiving the memory utilization data, the orchestrator server 1240 may receive cache utilization data indicative of a frequency of cache accesses associated with a workload and/or cache miss rate information, as indicated in block 1526. Additionally or alternatively, in receiving the memory utilization data, the orchestrator server 1240 may receive volatile memory utilization data indicative of an amount volatile memory (e.g., the main memory 1304) used, a frequency of accesses to the volatile memory, page fault data, and/or other information indicative of the utilization of the volatile memory within each managed node 1260, as indicated in block 1528. The orchestrator server 1240 may additionally or alternatively receive non-volatile memory utilization data indicative of the amount of data stored and/or retrieved from the data storage devices 1312 and/or a frequency at which each workload issues write requests and/or read requests to the data storage devices 1312 in each managed node 1260, as indicated in block 1530. In receiving the telemetry data 1402, the orchestrator server 1240 may additionally or alternatively receive network utilization data indicative of an amount of network bandwidth used by each workload performed by each managed node 1260, as indicated in block 1532. After receiving the telemetry data 1402, the method 1500 advances to block 1534 of FIG. 16, in which the orchestrator server 1240 generates data analytics as the workloads are being performed.

[0068] Referring now to FIG. 16, in generating the data analytics, the orchestrator server 1240 may identify trends in the resource utilization of the workloads, as indicated in block 1536. For example, the orchestrator server 1240 may identify patterns in which one or more of the workloads cycle through phases of high processor utilization with low memory usage, followed by low processor utilization and high memory usage, or other phases. As indicated in block 1538, in the illustrative embodiment, the orchestrator server 1240 generates profiles of the workloads. In doing so, in the illustrative embodiment, the orchestrator server 1240 generates the labels 1406 for the workloads to uniquely identify each workload, as indicated in block 1540. Additionally, in the illustrative embodiment, the orchestrator server 1240 generates the classifications 1408 of the workloads, as indicated in block 1542. In the illustrative embodiment, as indicated in block 1544, in generating the data analytics, the orchestrator server 1240 also predicts future resource utilization of the workloads, such as by comparing a present resource utilization of each workload to the trends identified in block 1536 to determine the present phase of each workload, and then identifying the upcoming phases of the workloads from the trends.

[0069] In block 1546, the orchestrator server 1240 determines, as a function of the data analytics, adjustments to the workload assignments as the workloads are performed, to improve resource utilization. In doing so, in the illustrative embodiment, the orchestrator server 1240 identifies pres-

ently and/or predicted under-utilized (e.g., usage is less than a predefined threshold value) resources among the managed nodes **1260**, as indicated in block **1548**. For example, the orchestrator server **1240** may determine that a particular managed node **1260** is executing a workload that is in a phase in which only 60% of the available processor capacity is being used. Alternatively, the orchestrator server **1240** may determine that while nearly 100% of the processor capacity is presently being used, the workload assigned to the managed node **1260** will soon transition to a phase in which only 60% of the processor capacity will be used, based on the future resource utilization predicted in block **1544**. As indicated in block **1550**, the orchestrator server **1240** may identify presently and/or predicted over-utilized (e.g., usage is greater than a predefined threshold value) resources among the managed nodes **1260**. In block **1552**, the orchestrator server **1240** may determine workload reassignments to reassign workloads having complementary resource utilizations (e.g., a workload with a high processor utilization and low memory utilization and another workload with low processor utilization and high memory utilization) to the same managed node **1260** to improve the resource utilization. In doing so, the orchestrator server **1240** may additionally determine to offset the timing of (e.g., temporarily pause the execution or delay the start time) of one workload to enable its resource usage trend to become complementary with the resource usage trend of another workload on the same managed node **1260**.

[0070] The orchestrator server **1240** may additionally determine node-specific adjustments, as indicated in block **1554**. The node-specific adjustments may be embodied as changes to settings within one or more of the managed nodes **1260**, such as in the operating system, the drivers, and/or the firmware of components (e.g., the CPU **1302**, the memory **1304**, the communication circuitry **1308**, the one or more data storage devices **1312**, etc.) to improve resource utilization. As such, in the illustrative embodiment, in determining the node-specific adjustments, the orchestrator server **1240** may determine processor throttle adjustments, such as clock speed and/or processor affinity for one or more workloads, as indicated in block **1556**. Additionally or alternatively, the orchestrator server **1240** may determine memory usage adjustments, such as allocations of volatile memory (e.g., the memory **1304**) and/or data storage capacity (e.g., capacity of the one or more data storage devices **1312**), memory bus speeds, and/or other memory-related settings. As indicated in block **1560**, the orchestrator server **1240** may determine network bandwidth adjustments, such as an available bandwidth of the communication circuitry **1308** to be allocated to each workload in the managed node **1260**. Additionally or alternatively, the orchestrator server **1240** may determine one or more fan speed adjustments to increase or decrease the cooling within the managed node **1260**, as indicated in block **1562**. In block **1564**, the orchestrator server **1240** may modify the adjustments to the assignments of the workloads and/or to the node-specific adjustments to comply with the policy data **1404**. As an example, the policy data **1404** may indicate that the power consumption is not to exceed a predefined threshold and, in view of the threshold, the orchestrator server **1240** determines to reduce the speed of the CPU **1302** to satisfy the threshold and reassign a processor-intensive workload away from the managed node **1260** because, at the reduced speed, the CPU **1302** would be unable to complete the processor-

intensive workload within a predefined time period (e.g., a time period specified in a Service Level Agreement (SLA) between the user of the client device **1220** and the operator of the system **1210**).

[0071] Referring now to FIG. **17**, in block **1566**, the orchestrator server **1240** determines whether adjustments were determined. If not, the method **1500** loops back to block **1516** of FIG. **15**, in which the orchestrator server **1240** again receives telemetry data **1402** from the managed nodes **1260** as the workloads are performed. Otherwise, if adjustments were determined, the method **1500** advances to block **1568** in which the orchestrator server **1240** applies the determined adjustments. In doing so, the orchestrator server **1240** may issue one or more requests to perform a live migration of a workload between two managed nodes **1260** (i.e., a workload reassignment), as indicated in block **1570**. In the illustrative embodiment, the migration is live because, rather than waiting until the workloads have been completed to analyze the telemetry data **1402**, the orchestrator server **1240** collects and analyzes the telemetry data **1402**, and makes adjustments online (i.e., as the workloads are being performed), as described above. Additionally or alternatively, as indicated in block **1572**, the orchestrator server **1240** may issue one or more requests to one or more of the managed nodes **1260** to apply the node-specific adjustments described above with reference to block **1554** of FIG. **16**. After applying the adjustments, the method **1500** loops back to block **1516** of FIG. **15**, in which the orchestrator server **1240** receives additional telemetry data **1402** from the managed nodes **1260**.

Examples

[0072] Illustrative examples of the technologies disclosed herein are provided below. An embodiment of the technologies may include any one or more, and any combination of, the examples described below.

[0073] Example 1 includes an orchestrator server to manage workloads among a plurality of managed nodes coupled to a network, the orchestrator server comprising one or more processors; communication circuitry coupled to the one or more processors; one or more memory devices having stored therein a plurality of instructions that, when executed by the one or more processors, cause the orchestrator server to assign the workloads to each of the managed nodes; receive telemetry data indicative of resource utilization from the plurality of managed nodes as the workloads are performed; generate data analytics as a function of the telemetry data as the workloads are performed; determine, as a function of the data analytics as the workloads are performed, adjustments to the workload assignments to increase resource utilization among the managed nodes; and apply the determined adjustments to the managed nodes as the workloads are performed.

[0074] Example 2 includes the subject matter of Example 1, and wherein the plurality of instructions, when executed by the one or more processors, further cause the orchestrator server to obtain policy data indicative of one or more goals to be achieved in the management of the workloads; and modify the adjustments as a function of the policy data.

[0075] Example 3 includes the subject matter of any of Examples 1 and 2, and wherein to obtain the policy data comprises to receive at least one of service life cycle policy data indicative of a target life cycle of one or more of the managed nodes, power consumption policy data indicative of a target power usage of one or more of the managed

nodes, or thermal policy data indicative of a target temperature of one or more of the managed nodes.

[0076] Example 4 includes the subject matter of any of Examples 1-3, and wherein to receive the telemetry data from the managed nodes comprises to receive at least one of temperature data indicative of a temperature within each managed node, power consumption data indicative of an amount of power consumed by each managed node, processor utilization data indicative of an amount of processor usage consumed by each workload performed by each managed node, memory utilization data indicative of an amount or frequency of memory use by each workload performed by each managed node, or network utilization data indicative of an amount of network bandwidth used by each workload performed by each managed node.

[0077] Example 5 includes the subject matter of any of Examples 1-4, and wherein to generate the data analytics comprises to identify trends in resource utilization of the workloads.

[0078] Example 6 includes the subject matter of any of Examples 1-5, and wherein to generate the data analytics comprises to generate profiles of the workloads, wherein the profiles are indicative of an identity of each workload and a resource usage classification of each workload.

[0079] Example 7 includes the subject matter of any of Examples 1-6, and wherein to generate the data analytics comprises to predict future resource utilization of the workloads.

[0080] Example 8 includes the subject matter of any of Examples 1-7, and wherein to determine adjustments to the workload assignments comprises to identify one or more of the managed nodes as under-utilized at a present time or at a future time.

[0081] Example 9 includes the subject matter of any of Examples 1-8, and wherein to determine adjustments to the workload assignments comprises to identify one or more of the managed nodes as over-utilized at a present time or at a future time.

[0082] Example 10 includes the subject matter of any of Examples 1-9, and wherein to determine adjustments to the workload assignments comprises to determine one or more workload reassignments among the managed nodes.

[0083] Example 11 includes the subject matter of any of Examples 1-10, and wherein to determine the adjustments comprises to determine one or more node-specific adjustments indicative of changes to an availability of one or more resources of a managed node to one or more of the workloads performed by the managed node.

[0084] Example 12 includes the subject matter of any of Examples 1-11, and wherein to determine the node-specific adjustments comprises to determine at least one of a processor throttle adjustment, a memory usage adjustment, a network bandwidth adjustment, or a fan speed adjustment.

[0085] Example 13 includes the subject matter of any of Examples 1-12, and wherein to apply the determined adjustments comprises to issue a request to perform a live migration of a workload between the managed nodes.

[0086] Example 14 includes the subject matter of any of Examples 1-13, and wherein to apply the determined adjustments comprises to issue a request to one of the managed nodes to apply one or more node-specific adjustments indicative of changes to an availability of one or more resources of the managed node to one or more of the workloads performed by the managed node.

[0087] Example 15 includes a method for managing workloads among a plurality of managed nodes coupled to a network, the method comprising assigning, by an orchestrator server, the workloads to each of the managed nodes; receiving, by the orchestrator server, telemetry data indicative of resource utilization from the plurality of managed nodes as the workloads are performed; generating, by the orchestrator server, data analytics as a function of the telemetry data as the workloads are performed; determining, by the orchestrator server, as a function of the data analytics and as the workloads are performed, adjustments to the workload assignments to increase resource utilization among the managed nodes; and applying, by the orchestrator server, the determined adjustments as the workloads are performed.

[0088] Example 16 includes the subject matter of Example 15, and further including obtaining, by the orchestrator server, policy data indicative of one or more goals to be achieved in the management of the workloads; and modifying, by the orchestrator server, the adjustments as a function of the policy data.

[0089] Example 17 includes the subject matter of any of Examples 15 and 16, and wherein receiving the policy data comprises receiving at least one of service life cycle policy data indicative of a target life cycle of one or more of the managed nodes, power consumption policy data indicative of a target power usage of one or more of the managed nodes, or thermal policy data indicative of a target temperature of one or more of the managed nodes.

[0090] Example 18 includes the subject matter of any of Examples 15-17, and wherein receiving the telemetry data from the managed nodes comprises receiving at least one of temperature data indicative of a temperature within each managed node, power consumption data indicative of an amount of power consumed by each managed node, processor utilization data indicative of an amount of processor usage consumed by each workload performed by each managed node, memory utilization data indicative of an amount or frequency of memory use by each workload performed by each managed node, or network utilization data indicative of an amount of network bandwidth used by each workload performed by each managed node.

[0091] Example 19 includes the subject matter of any of Examples 15-18, and wherein generating the data analytics comprises identifying trends in resource utilization of the workloads.

[0092] Example 20 includes the subject matter of any of Examples 15-19, and wherein generating the data analytics comprises generating profiles of the workloads, wherein the profiles are indicative of an identity of each workload and a resource usage classification of each workload.

[0093] Example 21 includes the subject matter of any of Examples 15-20, and wherein generating the data analytics comprises predicting future resource utilization of the workloads.

[0094] Example 22 includes the subject matter of any of Examples 15-21, and wherein determining adjustments to the workload assignments comprises identifying one or more of the managed nodes as under-utilized at a present time or at a future time.

[0095] Example 23 includes the subject matter of any of Examples 15-22, and wherein determining adjustments to

the workload assignments comprises identifying one or more of the managed nodes as over-utilized at a present time or at a future time.

[0096] Example 24 includes the subject matter of any of Examples 15-23, and wherein determining adjustments to the workload assignments comprises determining one or more workload reassignments among the managed nodes.

[0097] Example 25 includes the subject matter of any of Examples 15-24, and wherein determining the adjustments comprises determining one or more node-specific adjustments indicative of changes to an availability of one or more resources of a managed node to one or more of the workloads performed by the managed node.

[0098] Example 26 includes the subject matter of any of Examples 15-25, and wherein determining the node-specific adjustments comprises determining at least one of a processor throttle adjustment, a memory usage adjustment, a network bandwidth adjustment, or a fan speed adjustment.

[0099] Example 27 includes the subject matter of any of Examples 15-26, and wherein applying the determined adjustments comprises issuing a request to perform a live migration of a workload between the managed nodes.

[0100] Example 28 includes the subject matter of any of Examples 15-27, and wherein applying the determined adjustments comprises issuing a request to one of the managed nodes to apply one or more node-specific adjustments indicative of changes to an availability of one or more resources of the managed node to one or more of the workloads performed by the managed node.

[0101] Example 29 includes one or more machine-readable storage media comprising a plurality of instructions stored thereon that in response to being executed, cause an orchestrator server to perform the method of any of Examples 15-28.

[0102] Example 30 includes an orchestrator server to manage workloads among a plurality of managed nodes coupled to a network, the orchestrator server comprising one or more processors; communication circuitry coupled to the one or more processors; one or more memory devices having stored therein a plurality of instructions that, when executed by the one or more processors, cause the orchestrator server to perform the method of any of Examples 15-28.

[0103] Example 31 includes an orchestrator server to manage workloads among a plurality of managed nodes coupled to a network, the orchestrator server comprising resource manager circuitry to assign the workloads to each of the managed nodes; and telemetry monitor circuitry receive telemetry data indicative of resource utilization from the plurality of managed nodes as the workloads are performed; wherein the resource manager circuitry is further to generate data analytics as a function of the telemetry data as the workloads are performed, determine, as a function of the data analytics as the workloads are performed, adjustments to the workload assignments to increase resource utilization among the managed nodes, and apply the determined adjustments to the managed nodes as the workloads are performed.

[0104] Example 32 includes the subject matter of Example 31, and further including policy manager circuitry to obtain policy data indicative of one or more goals to be achieved in the management of the workloads, and wherein the resource manager circuitry is further to modify the adjustments as a function of the policy data.

[0105] Example 33 includes the subject matter of any of Examples 31 and 32, and wherein to obtain the policy data comprises to receive at least one of service life cycle policy data indicative of a target life cycle of one or more of the managed nodes, power consumption policy data indicative of a target power usage of one or more of the managed nodes, or thermal policy data indicative of a target temperature of one or more of the managed nodes.

[0106] Example 34 includes the subject matter of any of Examples 31-33, and wherein to receive the telemetry data from the managed nodes comprises to receive at least one of temperature data indicative of a temperature within each managed node, power consumption data indicative of an amount of power consumed by each managed node, processor utilization data indicative of an amount of processor usage consumed by each workload performed by each managed node, memory utilization data indicative of an amount or frequency of memory use by each workload performed by each managed node, or network utilization data indicative of an amount of network bandwidth used by each workload performed by each managed node.

[0107] Example 35 includes the subject matter of any of Examples 31-34, and wherein to generate the data analytics comprises to identify trends in resource utilization of the workloads.

[0108] Example 36 includes the subject matter of any of Examples 31-35, and wherein to generate the data analytics comprises to generate profiles of the workloads, wherein the profiles are indicative of an identity of each workload and a resource usage classification of each workload.

[0109] Example 37 includes the subject matter of any of Examples 31-36, and wherein to generate the data analytics comprises to predict future resource utilization of the workloads.

[0110] Example 38 includes the subject matter of any of Examples 31-37, and wherein to determine adjustments to the workload assignments comprises to identify one or more of the managed nodes as under-utilized at a present time or at a future time.

[0111] Example 39 includes the subject matter of any of Examples 31-38, and wherein to determine adjustments to the workload assignments comprises to identify one or more of the managed nodes as over-utilized at a present time or at a future time.

[0112] Example 40 includes the subject matter of any of Examples 31-39, and wherein to determine adjustments to the workload assignments comprises to determine one or more workload reassignments among the managed nodes.

[0113] Example 41 includes the subject matter of any of Examples 31-40, and wherein to determine the adjustments comprises to determine one or more node-specific adjustments indicative of changes to an availability of one or more resources of a managed node to one or more of the workloads performed by the managed node.

[0114] Example 42 includes the subject matter of any of Examples 31-41, and wherein to determine the node-specific adjustments comprises to determine at least one of a processor throttle adjustment, a memory usage adjustment, a network bandwidth adjustment, or a fan speed adjustment.

[0115] Example 43 includes the subject matter of any of Examples 31-42, and wherein to apply the determined adjustments comprises to issue a request to perform a live migration of a workload between the managed nodes.

[0116] Example 44 includes the subject matter of any of Examples 31-43, and wherein to apply the determined adjustments comprises to issue a request to one of the managed nodes to apply one or more node-specific adjustments indicative of changes to an availability of one or more resources of the managed node to one or more of the workloads performed by the managed node.

[0117] Example 45 includes an orchestrator server to manage workloads among a plurality of managed nodes coupled to a network, the orchestrator server comprising circuitry assigning the workloads to each of the managed nodes; circuitry for receiving telemetry data indicative of resource utilization from the plurality of managed nodes as the workloads are performed; means for generating data analytics as a function of the telemetry data as the workloads are performed; means for determining, as a function of the data analytics as the workloads are performed, adjustments to the workload assignments to increase resource utilization among the managed nodes; and means for applying the determined adjustments to the managed nodes as the workloads are performed.

[0118] Example 46 includes the subject matter of Example 45, and further including circuitry for obtaining policy data indicative of one or more goals to be achieved in the management of the workloads; and means for modify the adjustments as a function of the policy data.

[0119] Example 47 includes the subject matter of any of Examples 45 and 46, and wherein the circuitry for obtaining the policy data comprises circuitry for receiving at least one of service life cycle policy data indicative of a target life cycle of one or more of the managed nodes, power consumption policy data indicative of a target power usage of one or more of the managed nodes, or thermal policy data indicative of a target temperature of one or more of the managed nodes.

[0120] Example 48 includes the subject matter of any of Examples 45-47, and wherein the circuitry for receiving the telemetry data from the managed nodes comprises circuitry for receiving at least one of temperature data indicative of a temperature within each managed node, power consumption data indicative of an amount of power consumed by each managed node, processor utilization data indicative of an amount of processor usage consumed by each workload performed by each managed node, memory utilization data indicative of an amount or frequency of memory use by each workload performed by each managed node, or network utilization data indicative of an amount of network bandwidth used by each workload performed by each managed node.

[0121] Example 49 includes the subject matter of any of Examples 45-48, and wherein the means for generating the data analytics comprises means for identifying trends in resource utilization of the workloads.

[0122] Example 50 includes the subject matter of any of Examples 45-49, and wherein the means for generating the data analytics comprises means for generating profiles of the workloads, wherein the profiles are indicative of an identity of each workload and a resource usage classification of each workload.

[0123] Example 51 includes the subject matter of any of Examples 45-50, and wherein the means for generating the data analytics comprises means for predicting future resource utilization of the workloads.

[0124] Example 52 includes the subject matter of any of Examples 45-51, and wherein the means for determining adjustments to the workload assignments comprises means for identifying one or more of the managed nodes as under-utilized at a present time or at a future time.

[0125] Example 53 includes the subject matter of any of Examples 45-52, and wherein the means for determining adjustments to the workload assignments comprises means for identifying one or more of the managed nodes as over-utilized at a present time or at a future time.

[0126] Example 54 includes the subject matter of any of Examples 45-53, and wherein the means for determining adjustments to the workload assignments comprises means for determining one or more workload reassignments among the managed nodes.

[0127] Example 55 includes the subject matter of any of Examples 45-54, and wherein the means for determining the adjustments comprises means for determining one or more node-specific adjustments indicative of changes to an availability of one or more resources of a managed node to one or more of the workloads performed by the managed node.

[0128] Example 56 includes the subject matter of any of Examples 45-55, and wherein the means for determining the node-specific adjustments comprises means for determining at least one of a processor throttle adjustment, a memory usage adjustment, a network bandwidth adjustment, or a fan speed adjustment.

[0129] Example 57 includes the subject matter of any of Examples 45-56, and wherein the means for applying the determined adjustments comprises circuitry for issuing a request to perform a live migration of a workload between the managed nodes.

[0130] Example 58 includes the subject matter of any of Examples 45-57, and wherein the means for applying the determined adjustments comprises circuitry for issuing a request to one of the managed nodes to apply one or more node-specific adjustments indicative of changes to an availability of one or more resources of the managed node to one or more of the workloads performed by the managed node.

1. An orchestrator server to manage workloads among a plurality of managed nodes coupled to a network, the orchestrator server comprising:

- one or more processors;
- one or more memory devices having stored therein a plurality of instructions that, when executed by the one or more processors, cause the orchestrator server to:
 - assign the workloads to each of the managed nodes;
 - receive telemetry data indicative of resource utilization from the plurality of managed nodes as the workloads are performed;
 - generate data analytics as a function of the telemetry data as the workloads are performed;
 - determine, as a function of the data analytics as the workloads are performed, adjustments to the workload assignments to increase resource utilization among the managed nodes; and
 - apply the determined adjustments to the managed nodes as the workloads are performed.

2. The orchestrator server of claim 1, wherein the plurality of instructions, when executed by the one or more processors, further cause the orchestrator server to:

- obtain policy data indicative of one or more goals to be achieved in the management of the workloads; and
- modify the adjustments as a function of the policy data.

3. The orchestrator server of claim 2, wherein to obtain the policy data comprises to receive at least one of service life cycle policy data indicative of a target life cycle of one or more of the managed nodes, power consumption policy data indicative of a target power usage of one or more of the managed nodes, or thermal policy data indicative of a target temperature of one or more of the managed nodes.

4. The orchestrator server of claim 1, wherein to receive the telemetry data from the managed nodes comprises to receive at least one of temperature data indicative of a temperature within each managed node, power consumption data indicative of an amount of power consumed by each managed node, processor utilization data indicative of an amount of processor usage consumed by each workload performed by each managed node, memory utilization data indicative of an amount or frequency of memory use by each workload performed by each managed node, or network utilization data indicative of an amount of network bandwidth used by each workload performed by each managed node.

5. The orchestrator server of claim 1, wherein to generate the data analytics comprises to identify trends in resource utilization of the workloads.

6. The orchestrator server of claim 1, wherein to generate the data analytics comprises to generate profiles of the workloads, wherein the profiles are indicative of an identity of each workload and a resource usage classification of each workload.

7. The orchestrator server of claim 1, wherein to generate the data analytics comprises to predict future resource utilization of the workloads.

8. The orchestrator server of claim 1, wherein to determine adjustments to the workload assignments comprises to identify one or more of the managed nodes as under-utilized at a present time or at a future time.

9. The orchestrator server of claim 1, wherein to determine adjustments to the workload assignments comprises to identify one or more of the managed nodes as over-utilized at a present time or at a future time.

10. The orchestrator server of claim 1, wherein to determine adjustments to the workload assignments comprises to determine one or more workload reassignments among the managed nodes.

11. The orchestrator server of claim 1, wherein to determine the adjustments comprises to determine one or more node-specific adjustments indicative of changes to an availability of one or more resources of a managed node to one or more of the workloads performed by the managed node.

12. The orchestrator server of claim 11, wherein to determine the node-specific adjustments comprises to determine at least one of a processor throttle adjustment, a memory usage adjustment, a network bandwidth adjustment, or a fan speed adjustment.

13. One or more machine-readable storage media comprising a plurality of instructions stored thereon that, in response to being executed, cause an orchestrator server to:
assign the workloads to each of a plurality of managed nodes;

receive telemetry data indicative of resource utilization from the plurality of managed nodes as the workloads are performed;

generate data analytics as a function of the telemetry data as the workloads are performed;

determine, as a function of the data analytics as the workloads are performed, adjustments to the workload assignments to increase resource utilization among the managed nodes; and

apply the determined adjustments to the managed nodes as the workloads are performed.

14. The one or more machine-readable storage media of claim 13, wherein the plurality of instructions, when executed, further cause the orchestrator server to:

obtain policy data indicative of one or more goals to be achieved in the management of the workloads; and

modify the adjustments as a function of the policy data.

15. The one or more machine-readable storage media of claim 14, wherein to obtain the policy data comprises to receive at least one of service life cycle policy data indicative of a target life cycle of one or more of the managed nodes, power consumption policy data indicative of a target power usage of one or more of the managed nodes, or thermal policy data indicative of a target temperature of one or more of the managed nodes.

16. The one or more machine-readable storage media of claim 13, wherein to receive the telemetry data from the managed nodes comprises to receive at least one of temperature data indicative of a temperature within each managed node, power consumption data indicative of an amount of power consumed by each managed node, processor utilization data indicative of an amount of processor usage consumed by each workload performed by each managed node, memory utilization data indicative of an amount or frequency of memory use by each workload performed by each managed node, or network utilization data indicative of an amount of network bandwidth used by each workload performed by each managed node.

17. The one or more machine-readable storage media of claim 13, wherein to generate the data analytics comprises to identify trends in resource utilization of the workloads.

18. The one or more machine-readable storage media of claim 13, wherein to generate the data analytics comprises to generate profiles of the workloads, wherein the profiles are indicative of an identity of each workload and a resource usage classification of each workload.

19. The one or more machine-readable storage media of claim 13, wherein to generate the data analytics comprises to predict future resource utilization of the workloads.

20. The one or more machine-readable storage media of claim 13, wherein to determine adjustments to the workload assignments comprises to identify one or more of the managed nodes as under-utilized at a present time or at a future time.

21. The one or more machine-readable storage media of claim 13, wherein to determine adjustments to the workload assignments comprises to identify one or more of the managed nodes as over-utilized at a present time or at a future time.

22. The one or more machine-readable storage media of claim 13, wherein to determine adjustments to the workload assignments comprises to determine one or more workload reassignments among the managed nodes.

23. The one or more machine-readable storage media of claim 13, wherein to determine the adjustments comprises to determine one or more node-specific adjustments indicative of changes to an availability of one or more resources of a managed node to one or more of the workloads performed by the managed node.

24. The one or more machine-readable storage media of claim **23**, wherein to determine the node-specific adjustments comprises to determine at least one of a processor throttle adjustment, a memory usage adjustment, a network bandwidth adjustment, or a fan speed adjustment.

25. An orchestrator server to manage workloads among a plurality of managed nodes coupled to a network, the orchestrator server comprising:

 circuitry assigning the workloads to each of the managed nodes;

 circuitry for receiving telemetry data indicative of resource utilization from the plurality of managed nodes as the workloads are performed;

 means for generating data analytics as a function of the telemetry data as the workloads are performed;

 means for determining, as a function of the data analytics as the workloads are performed, adjustments to the workload assignments to increase resource utilization among the managed nodes; and

 means for applying the determined adjustments to the managed nodes as the workloads are performed.

26. A method for managing workloads among a plurality of managed nodes coupled to a network, the method comprising:

 assigning, by an orchestrator server, the workloads to each of the managed nodes;

 receiving, by the orchestrator server, telemetry data indicative of resource utilization from the plurality of managed nodes as the workloads are performed;

 generating, by the orchestrator server, data analytics as a function of the telemetry data as the workloads are performed;

 determining, by the orchestrator server, as a function of the data analytics and as the workloads are performed, adjustments to the workload assignments to increase resource utilization among the managed nodes; and

 applying, by the orchestrator server, the determined adjustments as the workloads are performed.

27. The method of claim **26**, further comprising:

 obtaining, by the orchestrator server, policy data indicative of one or more goals to be achieved in the management of the workloads; and

 modifying, by the orchestrator server, the adjustments as a function of the policy data.

28. The method of claim **27**, wherein receiving the policy data comprises receiving at least one of service life cycle policy data indicative of a target life cycle of one or more of the managed nodes, power consumption policy data indicative of a target power usage of one or more of the managed nodes, or thermal policy data indicative of a target temperature of one or more of the managed nodes.

* * * * *