9

COMPUTING, STORAGE, AND NETWORKING RESOURCE MANAGEMENT IN DATA CENTERS

RONGHUI CAO[1], ZHUO TANG[1], KENLI LI[1] AND KEQIN LI[2]

[1] *College of Information Science and Engineering, Hunan University, Changsha, China*

[2] *Department of Computer Science, State University of New York, New Paltz, New York, United States of America*

9.1 INTRODUCTION

Current data centers can contain hundreds of thousands of servers [1]. It is no doubt that the performance and stability of data centers have been significantly impacted by resource management. Moreover, in the course of data center construction, the creation of dynamic resource pools is essential. Some technology companies have built their own data centers for various applications, such as the deep learning cloud service run by Google. Resource service providers usually rent computation and storage resources to users at a very low cost.

Cloud computing platform, which rent various virtual resources to tenants, is becoming more and more popular for resource service websites or data applications. However, with the increasing of virtualization technologies and various clouds continue to expand their server clusters, resource management is becoming more and more complex. Obviously, adding more hardware devices to extend the cluster scale of the data center easily causes unprecedented resource management pressures in data centers.

Resource management in cloud platforms refers to how to efficiently utilize and schedule the virtual resources, such as computing resources. With the development of various open-source approaches and expansion of open-source communities, multiple resource management technologies have been widely used in the date centers. OpenStack [2], KVM [3], and Ceph [4] are some typical examples developed over the past years. It is clear that these resource management methods are considered critical factors for data center creation.

However, some resource management challenges are still impacting the modern data centers [7]. The first challenge is how to integrate various resources (hardware resource and virtual resource) into a unified platform. The second challenge is how to easily manage various resources in the data centers. The third challenge is resource services, especially network services. Choosing an appropriate resource management method among different resource management platforms and virtualization techniques is hence difficult and complex. Therefore, the following criteria should be taken into account: ease of resource management, provisional storage pool, and flexibility in performing the network architectures (such as resource transmission across different instances).

In this chapter, we will first explain the resource virtualization and resource management in data centers. We will then elaborate on the cloud platform demands for data centers and the related open-source cloud offerings focusing mostly on cloud platforms. Next, we will elaborate on the single-cloud bottlenecks and the multi-cloud demands in data centers. Finally, we will highlight the different large-scale cluster resource management architectures based on the OpenStack cloud platform.

9.2 RESOURCE VIRTUALIZATION AND RESOURCE MANAGEMENT

9.2.1 Resource Virtualization

In computing, virtualization refers to the act of creating a virtual (rather than actual) version of something, including

---

Data Center Handbook: *Plan, Design, Build, and Operations of a Smart Data Center,* Second Edition. Edited by Hwaiyu Geng.
© 2020 John Wiley & Sons, Inc. Published 2020 by John Wiley & Sons, Inc.

virtual computer hardware platforms, storage devices, and computer network resources.

Hardware virtualization refers to the creation of virtual resources acts like the real computer with a full operating system. Software executed on these virtual resources is not directly running on the underlying hardware resources. For example, a computer that is running Microsoft Windows may host a virtual machine (VM) that looks like a computer with the Ubuntu Linux operating system; Ubuntu-based software can be run on the VM.

According to the different deployment patterns and operating mechanism, resource virtualization can be divided into two types: full virtualization and paravirtualization (Fig. 9.1). Full virtualization is also called primitive virtualization technology. It uses the VM to coordinate the guest operating systems and the original hardware devices. Some protected instructions must be captured and processed by the hypervisor. Paravirtualization is another technology that similar to the full virtualization. It uses hypervisor to share the underlying hardware devices, but its guest operating systems integrate the resource virtualization codes. In the past 5 years, the full virtualization technologies gained polarity with the rise of KVM, Xen, etc. KVM is open-source software, and the kernel component of KVM is included in mainline Linux, as of 2.6.20. The first version of KVM was developed at a small Israeli company, Qumranet, which has been acquired by Red Hat in 2008.

For the resource management, data centers must not only comprehensively consider various factors such as manufacturers, equipment, applications, users and technology, etc. but also consider the integration with operation maintenance process of data centers. Obviously, building an open, standardized, easy-to-expand, and interoperable unified intelligent resource management platform is not easy. The scale of data centers is getting larger and more complex, and the types of applications are becoming more and more complex, which makes the difficulty of resource management even more difficult:

- *Multitenant support:* Management of multiple tenants and their applied resources, applications, and operating systems in large-scale data centers with different contracts and agreements.
- *Multi-data center support:* Management of multiple data centers with different security levels, hardware devices, resource management approaches, and resource virtualization technologies.
- *Resource monitor:* Monitoring of various resources with different tenant requests, hardware devices, management platforms, and cluster nodes up to date.
- *Budget control:* Manage the cost of data centers and reduce budget as much as possible, where resources are procured based on "best cost"—regardless if it is deployed at the hardware devices or used for resource virtualization. Additionally, energy and cooling costs are also the principal aspects of budget reducing.
- *Application deploying:* Deploy new applications and services faster with limited understanding of resource availability as well as inconsistent policies and structure.

Data centers with heterogeneous architecture make the above problems particularly difficult since the resource management solutions with high scalability and performance are emergency needed. By tackling these problems, data services can be made more efficient and reliable, notably reducing the internal server costs and increasing the utilization of energy and resource in data centers.

As a result, various virtualization technologies and architectures have been used in data centers to simplify resource management. Without question, the wide use of virtualization brings many benefits for data centers, but it also incurs some costs caused by the virtual machine monitor (VMM) or called hypervisor. These costs usually come from various activities within the virtualization layer such as code rewriting, OS memory operations, and, most commonly, resource scheduling overhead. The hypervisor is the kernel of virtual resource
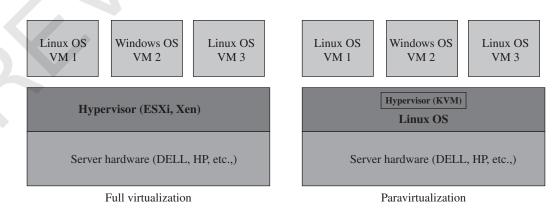

Full virtualization   Paravirtualization

FIGURE 9.1   Two resource virtualization methods. Source: Ronglui Cao.

management, especially for VM. It can be software, firmware, or hardware used to build and execute VM.

Actually, resource virtualization is not a new technology for the large-scale server cluster. It was largely used in the 1960s for mainframe and been widely used in early 2000 for resource pool creation and cloud platforms [5]. In a traditional concept of virtual servers, multiple virtual servers or VMs can be simultaneously operated on one traditional single physical server. As a result, the data centers can operate using VM to improve utilization of server resource capacity and consequently reduce the hardware device cost in data centers. With advances in virtualization technology, we are able to run over 100 VMs on one physical server node.

9.2.2 Resource Management

The actual overhead of resource management and scheduling in data centers vary depending on the virtualization technologies and cloud platforms being used. With greater resource multiplexing, hardware costs can be decreased by resource virtualization. While many data centers would like to move various applications to VMs to lower energy and hardware costs, this kind of transition should be ensured that will not be disrupted by correctly estimating the resource requirements. Fortunately, the disrupt problem can be solved by monitoring the workload of applications and attempt to configure the VMs.

Several earlier researches describe various implementations of hypervisor. The performance results showed that hypervisor can measure the overhead impact of resource virtualization on microbenchmark or macrobenchmark. Some commercial tools use trace-based methods to support server load balancing, resource management, and simulating placement of VMs to improve server resource utilization and cluster performance. Other commercial tools use the trace-based resource management solution that scales the resource usage traces by a given CPU multiplier. In addition, cluster system activities and application operations can incur additional overheads of CPUs.

9.2.2.1 VM Deployment

With the increasing task scale in data centers, breaking down a large serial task into several small tasks and assigning them to different VMs to complete the task in parallel is the main method to reduce the task completion time. Therefore, in modern data centers, how to deploy VMs has become one of the important factors that determine the task completion time and improve resource utilization.

When the VM deployment, the utilization of computation resource, and I/O resource are considered together, it may help to find a multi-objective optimization VM deployment model. Moreover, some VM-optimized deployment mechanisms based on the resource matching bottleneck can

also reduce data transmission response time in the data centers. Unfortunately, the excessive time complexity of these VM deployment algorithms will seriously affect the overall operation of data centers.

9.2.2.2 VM Migration

In order to meet the real-time changing requirements of the task, the VM migration technology is introduced in modern data centers. The primary application scenario is using VM migration to integrate resources and decrease energy consumption by monitoring the state of VMs. Green VM Migration Controller (GVMC) combines the resource utilization of the physical servers with the destination nodes of VM migration to minimize the cluster size of data centers. Classical genetic algorithm is often improved and optimized for VM migration to solve the energy consumption problem in data centers.

The VM migration duration is another interesting resource management issue for data centers. It is determined by many factors, including the image size of VM, the memory size, the choice of the migration node, etc. How to reduce the migration duration by optimizing these factors has always been one of the hot topics in data center resource management. Some researchers formalize the joint routing and VM placement problem and leverage the Markov approximation technique to solve the online resource joint optimization problem, with the goal of optimizing the long-term averaged performance under changing workloads.

Obviously, the traditional resource virtualization technologies or resource management mechanisms in data centers are both cannot meet the needs of the new generation of high-density servers and storage devices. On the other hand, the capacity growth of information technology (IT) infrastructure in data centers is severely constrained by floor space. The cloud platform deployed in data centers emerges as the resource management infrastructure to solve these problems.

9.3 CLOUD PLATFORM

The landscape of IT has been evolving ever since the first rudimentary computers were introduced at the turn of the twentieth century. With the introduction of the cloud computing model, the design and deployment of modern data centers have been transformed in the last two decades. Essentially, the difference between cloud service and traditional data service is that in the cloud platform, users can access their resources and data through the Internet. The cloud provider performs ongoing maintenance and updates for resources and services, often owning multiple data centers in several geographic locations to safeguard user data during outages and other failures. The resource management

in the cloud platform is a departure from traditional data center strategies since it provides a resource pool that can be consumed by users as services as opposed to dedicating infrastructure to each individual application.

9.3.1 Architecture of Cloud Computing

The introduction of the cloud platform enabled a redefinition of resource service that includes a new perspective—all virtual resources and services are available remotely. It offers three different model or technical use of resource service (Fig. 9.2): Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS).

Each layer of the model has a specific role:

- IaaS layer corresponds to the hardware infrastructure of data centers. It is a service model in which IT infrastructure is provided as a service externally through the network and users are charged according to the actual use of resources.
- PaaS is a model that is "laying" on the IaaS. It provides a computing platform and solution services and allows the service providers to outsource the middleware applications, databases, and data integration layer.
- SaaS is the final layer of cloud and deploys application software on the PaaS layer. It defines a new delivery method, which also makes the software return to the essence of service. SaaS changes the way traditional software services provided, reduces the large amount of upfront investment required for local deployment, and further highlights the service attributes of information software.

9.3.2 Common Open-Source Cloud Platform

Some open-source cloud platforms take a more comprehensive approach, all of which integrate all necessary functions (including virtualization, resource management, application interfaces, and service security) in one platform. If deployed on servers and storage networks, these cloud platforms can provide a flexible cloud computing and storage infrastructure (IaaS).

9.3.2.1 OpenNebula

OpenNebula is an interesting open-source application (under the Apache license) developed at Universidad Complutense de Madrid. In addition to supporting private cloud structures, OpenNebula also supports the hybrid cloud architecture. Hybrid clouds allow the integration of private cloud infrastructure with public cloud infrastructure, such as Amazon, to provide a higher level of scalability. OpenNebula supports Xen, KVM/Linux, and VMware and relies on libvirt for resource management and introspection [8].

9.3.2.2 OpenStack

OpenStack cloud platform was released in July 2010 and quickly became the most popular open-source IaaS solution. The cloud platform is originally combined of two cloud plans, namely, Rackspace Hosting (cloud files) and Nebula platform from NASA (National Aeronautics and Space Administration). It is a cloud operating system that controls large pools of compute, storage, and networking resources throughout a data center, all managed through a dashboard that gives administrators control while empowering their users to provision resources through a web interface [9].
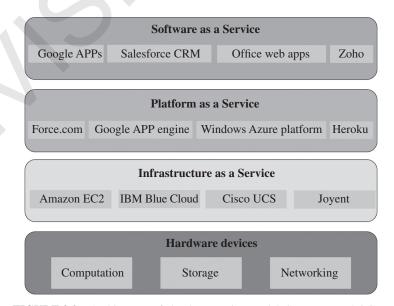

FIGURE 9.2 Architecture of cloud computing model. Source: Ronglui Cao.

9.3.2.3   *Eucalyptus*

Eucalyptus is one of the most popular open-source cloud solutions that used to build cloud computing infrastructure. Its full name is Elastic Utility Computing Architecture for Linking Your Programs to Useful Systems. The special of eucalyptus is that its interface is compatible with Amazon Elastic Compute Cloud (Amazon EC2—Amazon's cloud computing interface). In addition, Eucalyptus includes Walrus, a cloud storage application that is compatible with Amazon Simple Storage Service (Amazon S3—Amazon's cloud storage interface) [10].

9.3.2.4   *Nimbus*

Nimbus is another IaaS solution focused on scientific computing. It can borrow remote resources (such as remote storage provided by Amazon EC2) and manage them locally (resource configuration, VM deployment, status monitoring, etc.). Nimbus was evolved from the workspace service project. Since it is based on Amazon EC2, Nimbus supports Xen and KVM.

9.4   PROGRESS FROM SINGLE-CLOUD TO MULTI-CLOUD

With the ever-growing need of resource pool and the introduction of high-speed network devices, the data centers enable building scalable services through the scale-out model by utilizing the elastic pool of computing resources provided by such platforms. However, unlike native components, these extended devices typically do not provide specialized data services or multi-cloud resource management approaches. Therefore, enterprises have to consider the bottlenecks of computing performance and storage stability of single-cloud architecture. In addition, there is no doubt that traditional single-cloud platforms are more likely to suffer from single-point failures and vendor lock-in.

9.4.1   The Bottleneck of Single-Cloud Platform

Facing various resources as well as their diversity and heterogeneity, data center vendors may be confused about whether existing resource pools can completely meet the resource requirements of customer data. If not, no matter the level of competition or development, it is urgent for providers to extend hardware devices and platform infrastructures. To overcome the difficulties, data center vendors usually build a new resource pool under the acceptable bound of the risk and increase the number of resource nodes as the growing amount of data. However, when the cluster scales to 200 nodes, a request message will not respond until at least 10 seconds. David Willis, head of research and development at a UK telecom regulator, estimated that a lone OpenStack controller could manage around 500 computing nodes at most [6]. Figure 9.3 shows a general single-cloud architecture.

The bottlenecks of traditional single-cloud systems first lie in the scalability of architecture, which surely generates considerable expense of data migration. The extension of existing cloud platforms also makes customers suffer from service adjustments of cloud vendors that are not uncommon. For example, resource fluctuation in cloud platforms will affect the price of cloud services. Uncontrolled data availability further aggravates the decline in confidence of users. Some disruptions even lasted for several hours and directly destroy users' confidence. Therefore, vendors were confronted with a dilemma that they could do nothing but build a new cloud platform with a separate cloud management system.

9.4.2   Multi-cloud Architecture

Existing cloud resources exhibit great heterogeneities in terms of both performances and fault-tolerant requirements. Different cloud vendors build their respective infrastructures and keep upgrading them with newly emerging gears. Some multi-cloud architectures that rely on multiple cloud platforms for placing resource data have been used by current cloud providers (Fig. 9.4). Compared with the single-cloud storage, the multi-cloud platform can provide better service quality and more storage features. These features are extremely beneficial to the platform itself or cloud applications such as data backup, document archiving, and electronic health recording, which need to keep a large
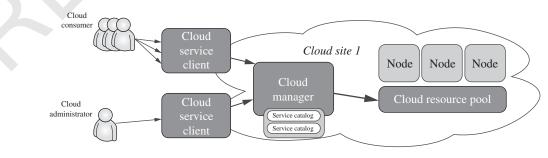

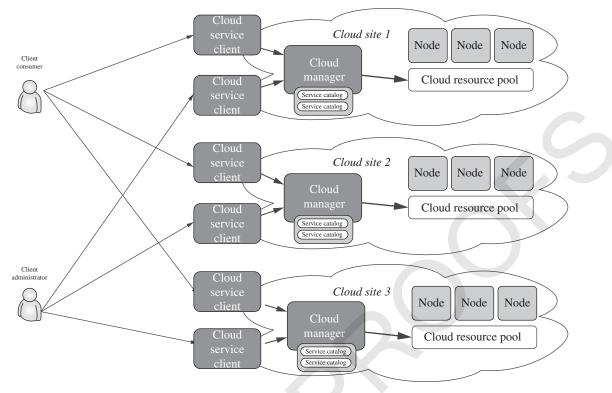**FIGURE 9.3**   A general single-cloud site. Source: Ronglui Cao.

FIGURE 9.4   Multi-cloud environment. Source: Ronglui Cao.

amount of data. Although the multi-cloud platform is a better selection, both administrators and maintainers are still inconvenienced since each bottom cloud site is managed by each provider separately and the corresponding resources are also independent.

Customers have to consider which cloud site is the most appropriate one to store their data with the highest cost effectiveness. Cloud administrators need to manage various resources in different manners and should be familiar with different management clients and configurations among bottom cloud sites. It is no doubt that these problems and restrictions can bring more challenges for resource storage in the multi-cloud environment.

9.5   RESOURCE MANAGEMENT ARCHITECTURE IN LARGE-SCALE CLUSTERS

When dealing with large-scale problems, naturally divide-and-conquer strategy is the best solution. It decomposes a problem of size N into K smaller subproblems. These subproblems are independent of each other and have the same nature as the original problem. In the most popular open-source cloud community, OpenStack community, there are three kinds of divide-and-conquer strategies for resource management in large-scale clusters: multi-region, multi-cell,

and resource cascading mechanism. The difference among them is the management concept.

9.5.1   Multi-region

In the OpenStack cloud platform, it supports to divide the large-scale cluster into different regions. The regions shared all the core components, and each of them is a complete OpenStack environment. When deploying in multi-region, the data center only needs to deploy a set of public authentication service of OpenStack, and other services and components can be deployed like a traditional OpenStack single-cloud platform. Users must specify a specific area/region when requesting any resources and services. Distributed resources in different regions can be managed uniformly, and different deployment architectures and even different OpenStack versions can be adopted between regions. The advantages of multi-region are simple deployment, fault domain isolation, flexibility, and freedom. It also has obvious shortcomings that every region is completely isolated from each other and the resources cannot be shared with each other. Cross-region resource migration can also not be supported. Therefore, it is particularly suitable for scenarios that the resources cross different data centers and distribute in different regions.

9.5.2   Nova Cells

The computation component of OpenStack provides nova multi-cell method for large-scale cluster environment. It is different from multi-region; it divides the large-scale clusters according to the service level, and the ultimate goal is to achieve that the single-cloud platform can support the capabilities of deployment and flexible expansion in data centers. The main strategy of nova cells (Fig. 9.5) is to divide different computing resources into cells and organize them in the form of a tree. The architecture of nova cells is shown as follows.

There are also some nova cell use cases in industry:

1. CERN (European Organization for Nuclear Research) OpenStack cluster may be the largest OpenStack deployment cluster currently disclosed. The scale of deployment as of February 2016 is as follows [11]:
   - Single region and 33 cells
   - 2 Ceph clusters
   - 5500 compute nodes, totaling 140k cores
   - More than 17,000 VMs
2. Tianhe-2 is one of the typical examples of the scale of China's thousand-level cluster, and it has been deployed and provided services in the National

Supercomputer Center in Guangzhou in early 2014. The scale of deployment is as follows [12].
- Single region and 8 cells.
- Each cell contains 2 control nodes and 126 computing nodes.
- The total scale includes 1152 physical nodes.

9.5.3   OpenStack Cascading

OpenStack cascading is a large-scale OpenStack cluster deployment supported by Huawei to support scenarios including 100,000 hosts, millions of VMs, and unified management across multiple data centers (Fig. 9.6). The strategy it adopts is also divide and conquer, that is, split a large OpenStack cluster into multiple small clusters and cascade the divided small clusters for unified management [13].

When users request resources, they first submit the request to the top-level OpenStack API. The top-level OpenStack will select a suitable bottom OpenStack based on a certain scheduling policy. The selected bottom OpenStack is responsible for the actual resource allocation.

This solution claims to support spanning up to 100 data centers, supports the deployment scale of 100,000 computing nodes, and can run 1 million VMs simultaneously. At present,
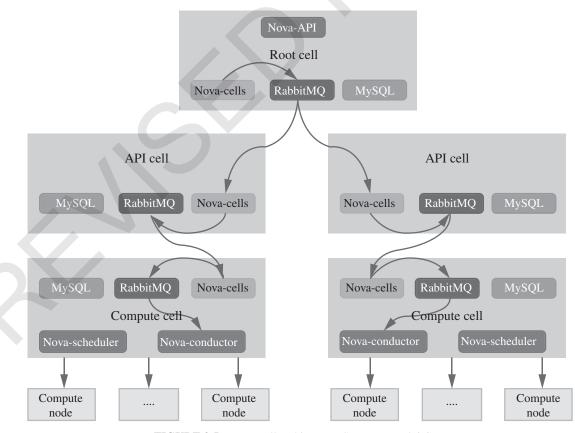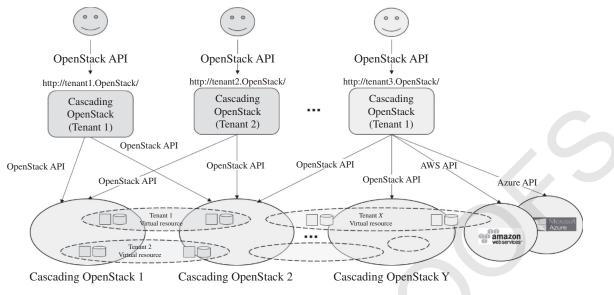

FIGURE 9.5   Nova cell architecture. Source: Ronglui Cao.

FIGURE 9.6   OpenStack cascading architecture. Source: Ronglui Cao.

the solution has separated two independent big-tent projects: one is Tricircle, which is responsible for network automation development in multi-cloud environment with networking component Neutron, and the other is Trio2o, which provides a unified API gateway for computation and storage resource management in multi-region OpenStack clusters.

9.6   CONCLUSIONS

The resource management of data centers is indispensable. The introduction of virtualization technologies and cloud platforms undoubtedly significantly increased in the resource utilization of data centers. Numerous scholars have produced a wealth of research on various types of resource management and scheduling in the data centers, but there is still further research value in many aspects. On the one hand, the resource integration limit still exists in a traditional data center and single-cloud platform. On the other hand, due to the defects of nonnative management of additional management plugins, existing multi-cloud architectures make resource management and scheduling often accompanied by high bandwidth and data transmission overhead. Therefore, the resource management of data centers based on the multi-cloud platform emerges at the historic moment under the needs of the constantly developing service applications.

REFERENCES

[1] Geng H. Chapter 1: Data Centers--Strategic Planning, Design, Construction, and Operations,*Data Center Handbook*. Wiley, 2014.

[2] Openstack. Available at http://www.openstack.org. Accessed on May 20, 2014.

[3] KVM. Available at http://www.linux-kvm.org/page/ Main_Page. Accessed on May 5, 2018.

[4] Ceph. Available at https://docs.ceph.com/docs/master/. Accessed on February 25, 2018.

[5] Kizza JM. Africa can greatly benefit from virtualization technology–Part II. Int J Comput ICT Res 2012;6(2).Available at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1. 372.8407&rep=rep1&type=pdf. Accessed on June 29, 2020.

[6] Cao R, et al. A scalable multi-cloud storage architecture for cloud-supported medical Internet of Things. IEEE Internet Things J, March 2020;7(3):1641–1654.

[7] Beloglazov A, Buyya R. Energy efficient resource management in virtualized cloud data centers. Proceedings of the 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing, Melbourne, Australia; May 17–20, 2010. IEEE. p. 826–831.

[8] Milojičić D, Llorente IM, Montero RS. Opennebula: a cloud management tool. IEEE Internet Comput 2011;15(2):11–14.

[9] Sefraoui O, Aissaoui M, Eleuldj M. OpenStack: toward an open-source solution for cloud computing. Int J Comput Appl 2012;55(3):38–42.

[10] Boland DJ, Brooker MIH, Turnbull JW. Eucalyptus Seed; 1980. Available at https://www.worldcat.org/title/eucalyptus-seed/oclc/924891653?referer=di&ht=edition. Accessed on June 29, 2020.

[11] Herran N. Spreading nucleonics: The Isotope School at the Atomic Energy Research Establishment, 1951–67. Br J Hist Sci 2006;39(4):569–586.

[12] Xue W, et al. Enabling and scaling a global shallow-water atmospheric model on Tianhe-2. Proceedings of the 2014 IEEE 28th International Parallel and Distributed Processing Symposium, Phoenix, AZ; May 19–23, 2014. IEEE. p. 745–754.

[13] Mayoral A, et al. Cascading of tenant SDN and cloud controllers for 5G network slicing using Transport API and Openstack API. Proceedings of the Optical Fiber Communication Conference. Optical Society of America, Los Angeles, CA; March 19–23, 2017. M2H. 3.