Cloud Data Center Network Architectures and Technologies

Data Communication Series

Cloud Data Center Network Architectures and Technologies
Lei Zhang and Le Chen

Campus Network Architectures and Technologies
Ningguo Shen, Bin Yu, Mingxiang Huang and Hailin Xu

Enterprise Wireless Local Area Network Architectures and Technologies
Rihai Wu, Xun Yang, Xia Zhou and Yibo Wang

Software-Defined Wide Area Network Architectures and Technologies
Cheng Sheng, Jie Bai and Qi Sun

SRv6 Network Programming: Ushering in a New Era of IP Networks
Zhenbin Li, Zhibo Hu and Cheng Li

For more information on this series please visit: https://www.routledge.com/Data-Communication-Series/book-series/DCSHW

Cloud Data Center Network Architectures and Technologies

Lei Zhang and Le Chen

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

人民邮电出版社
POSTS & TELECOM PRESS

First edition published 2021
by CRC Press
6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742

and by CRC Press
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

Translated by Yongdan Li

CRC Press is an imprint of Taylor & Francis Group, LLC

English Version by permission of Posts and Telecom Press Co., Ltd.

Library of Congress Cataloging-in-Publication Data
Names: Zhang, Lei (Engineering teacher), author. | Chen, Le
(Writer on computer networking), author.
Title: Cloud data center network architectures and technologies / Lei Zhang, Le Chen.
Description: First edition. | Boca Raton : CRC Press, 2021. | Summary: "This book has been written with the support of Huawei's large accumulation of technical knowledge and experience in the data center network (DCN) field, as well as its understanding of customer service requirements. This book describes in detail the architecture design, technical implementation, planning and design, and deployment suggestions for cloud DCNs based on the service challenges faced by cloud DCNs. This book starts by describing the overall architecture and technical evolution of DCNs, with the aim of helping readers understand the development of DCNs. It then proceeds to explain the design and implementation of cloud DCNs, including the service model of a single data center (DC), construction of physical and logical networks of DCs, construction of multiple DCNs, and security solutions of DCs. Next, this book dives deep into practices of cloud DCN deployment based on real-world cases to help readers better understand how to build cloud DCNs. Finally, this book introduces DCN openness and some of the hottest forward-looking technologies. In summary, you can use this book as areference to help you to build secure, reliable, efficient, and open cloud DCNs. It is intended for technical professionals of enterprises, research institutes, information departments, and DCs, as well as teachers and students of computer network-related majors in colleges and universities"— Provided by publisher.
Identifiers: LCCN 2020048556 (print) | LCCN 2020048557 (ebook) |
ISBN 9780367695705 (hardcover) | ISBN 9781003143185 (ebook)
Subjects: LCSH: Cloud computing. | Computer network architectures.
Classification: LCC QA76.585 .Z429 2021 (print) | LCC QA76.585 (ebook) |
DDC 004.67/82—dc23
LC record available at https://lccn.loc.gov/2020048556
LC ebook record available at https://lccn.loc.gov/2020048557

ISBN: 978-0-367-69570-5 (hbk)
ISBN: 978-0-367-69775-4 (pbk)
ISBN: 978-1-003-14318-5 (ebk)

Typeset in Minion
by codeMantra

Contents

Chapter 11 ▪ Cutting-Edge Technologies 373

Chapter 12 ▪ Components of the Cloud DCN Solution 425

Summary

This book has been written with the support of Huawei's large accumulation of technical knowledge and experience in the data center network (DCN) field as well as its understanding of customer service requirements. It describes in detail the architecture design, technical implementation, planning and design, and deployment suggestions for cloud DCNs based on the service challenges faced by cloud DCNs. It starts by describing the overall architecture and technical evolution of DCNs, with the aim of helping readers understand the development of DCNs. It then proceeds to explain the design and implementation of cloud DCNs, including the service model of a single data center (DC), construction of physical and logical networks of DCs, construction of multiple DCNs, and security solutions of DCs. Next, it dives deep into practices of cloud DCN deployment based on real-world cases to help readers better understand how to build cloud DCNs. Finally, it introduces DCN openness and some of the hottest forward-looking technologies.

In summary, you can use this book as a reference to help you to build secure, reliable, efficient, and open cloud DCNs. It is intended for technical professionals of enterprises, research institutes, information departments, and DCs, as well as teachers and students of computer network-related majors in colleges and universities.

DATA COMMUNICATION SERIES

Members

Shaowei Liu, President of Huawei Data Communication Product Line R&D Department

Zhipeng Zhao, Director of Huawei Data Communication Marketing Department

Xing Li, President of Campus Network Domain, Huawei Data Communication Product Line

Xiongfei Gu, President of WAN Domain, Huawei Data Communication Product Line

Leon Wang, President of Data Center Network Domain, Huawei Data Communication Product Line

Mingsong Shao, Director of Switch & Enterprise Gateway Product Department, Huawei Data Communication Product Line

Mingzhen Xie, Director of Information Digitalization and Experience Assurance Department, Huawei Data Communication Product Line

Jianbing Wang, Director of Architecture & Design Department, Huawei Data Communication Product Line

INTRODUCTION

This book first looks at the service characteristics of cloud computing and describes the impact of cloud computing on DCNs, evolution of the overall architecture and technical solution of DCs, and physical network, logical network, multi-DC, and security design solutions of DCs. Then, based on practical experiences of cloud DC deployment, it provides the recommended planning before deployment and key steps in implementation. Finally, it explains the hottest technologies of DCNs and the construction solution of Huawei cloud DCNs.

This book is a useful guide during SDN DCN planning and design, as well as engineering deployment, for ICT practitioners such as network engineers. For network technology enthusiasts and students, it can also be used as a reference for learning and understanding the cloud DCN architecture, common technologies, and cutting-edge technologies.

How Is the Book Organized

This book consists of 12 chapters. Chapter synopses follow below.

Chapter 1: Introduction to Cloud DCNs

This chapter covers basic features of cloud computing, development and evolution of virtualization technologies, and basics of cloud DCNs. It also describes characteristics of SDN network development and the relationship between orchestration and control, before explaining the concerns of enterprises in selecting different solutions from the service perspective.

Chapter 2: DCN Challenges

This chapter describes the five challenges faced by DCNs in the cloud computing era: large pipes for big data; pooling and automation for networks; security as a service (SECaaS) deployment; reliable foundation for networks; and intelligent O&M of DCs.

Chapter 3: Architecture and Technology Evolution of DCNs

This chapter describes the general architecture of physical networks in DCs and the evolution of major network technologies. The architecture and evolution process of typical DCNs are then described, with financial services companies and carriers used as examples.

Chapter 4: Functional Components and Service Models of Cloud DCNs

This chapter describes the orchestration model and service process of network services provided by the cloud platform and by the SDN controller (for when no cloud platform is available).

Chapter 5: Constructing a Physical Network (Underlay Network) in a DC

This chapter describes the typical architecture and design principles of the physical network, as well as comparison and selection of common network technologies.

Chapter 6: Constructing a Logical Network (Overlay Network) in a DC

This chapter describes basic concepts of logical networks, basic principles of mainstream VXLAN technologies, and how to use VXLAN to build logical networks.

Chapter 7: Constructing a Multi-DC Network

Multiple DCs need to be deployed to meet service requirements due to service scale expansion and service reliability and continuity requirements. This chapter describes the service requirement analysis and recommended network architecture design for multiple DCNs.

Chapter 8: Building E2E Security for Cloud DCNs

This chapter describes security challenges faced by cloud DCs and the overall technical solution at the security layer. It also describes specific security technologies and implementation solutions in terms of virtualization security, network security, advanced threat detection and defense, border security, and security management.

Chapter 9: Best Practices of Cloud DCN Deployment

This chapter describes the recommended planning methods for service and management networks of typical cloud DCs based on cloud DC deployment practices, and also explains key configuration processes during deployment, operation examples, and common service provisioning process based on the SDN controller.

Chapter 10: Openness of DCN

This chapter describes the necessity of DCN openness and the capabilities and benefits brought by openness of SDN controllers and forwarders.

Chapter 11: Cutting-Edge Technologies

This chapter describes some new DCN technologies and development trends that attract industry attention, including basic concepts and mainstream solutions of container networks, hybrid clouds, and AI Fabric.

Chapter 12: Components of the Cloud DCN Solution

This chapter describes the positioning, features, and functional architecture of some components that can be used during cloud DCN construction, including CloudEngine DC switches, CloudEngine virtual switches, HiSecEngine series firewalls, iMaster NCE-Fabric, and the SecoManager.

Icons Used in This Book

Acknowledgments

This book has been jointly written by the Data Communication Digital Information and Content Experience Department and Data Communication Architecture & Design Department of Huawei Technologies Co., Ltd. During the process of writing the book, high-level management from Huawei's Data Communication Product Line provided extensive guidance, support, and encouragement. We are sincerely grateful for all their support.

The following is a list of participants involved in the preparation and technical review of this book.

Editorial board: Lei Zhang, Le Chen, Fan Zhang, Shan Chen, Xiaolei Zhu, Zhongping Jiang, and Xuefeng Wu

Technical reviewers: Jun Guo, Jianbing Wang, Lei Zhang, Le Chen, and Fan Zhang

Translators: Yongdan Li, Zhenghong Zhu, Kaiyan Zhang, Mengyan Wang, Michael Chapman, and Fionnuala Magee

While the writers and reviewers of this book have many years of experience in ICT and have made every effort to ensure accuracy, it may be possible that minor errors have been included due to time limitations. We would like to express our heartfelt gratitude to the readers for their unremitting efforts in reviewing this book.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Authors

Mr. Lei Zhang is the Chief Architect of Huawei's DCN solution. He has more than 20 years' experience in network product and solution design, as well as a wealth of expertise in product design and development, network planning and design, and network engineering project implementation. He has led the design and deployment of more than ten large-scale DCNs for Fortune Global 500 companies worldwide.

Mr. Le Chen is a Huawei DCN solution documentation engineer with eight years' experience in developing documents related to DCN products and solutions. He has participated in the design and delivery of multiple large-scale enterprise DCNs. He has written many popular technical document series such as *DCN Handbook* and *BGP Topic*.

Taylor & Francis

Taylor & Francis Group

http://taylorandfrancis.com

CHAPTER 1

Introduction to Cloud DCNs

A CLOUD DATA CENTER (DC) is a new type of DC based on cloud computing architecture, where the computing, storage, and network resources are loosely coupled. Within a cloud DC, various IT devices are fully virtualized, while also being highly modularized, automated, and energy efficient. In addition, a cloud DC features virtualized servers, storage devices, and applications, enabling users to leverage various resources on demand. Automatic management of physical and virtual servers, service processes, and customer service charging is also provided. Starting with cloud computing and virtualization, this chapter describes the software-defined networking (SDN) technology used by cloud data center networks (DCNs) to tackle the challenges introduced by this new architecture.

1.1 CLOUD COMPUTING

Before examining cloud DCNs in more detail, we should first take a closer look at cloud computing. The pursuit of advanced productivity is never ending. Each industrial revolution has represented a leap in human productivity, as our society evolved from the mechanical and electric eras through to the current automatic and intelligent era.

Since the 1980s, and owing to the advances of global science and technology, culture, and the economy, we have gradually transitioned from an industrial society to an information society. By the mid-1990s, economic globalization had driven the rapid development of information technologies, with the Internet becoming widely applied by all kinds of businesses. As the global economy continues to grow, cracks have begun to appear in the current processes of enterprise informatization. Constrained by complex management modes, spiraling operational expenses, and weak scale-out support, enterprises require effective new information technology solutions. Such requirements have driven the emergence of cloud computing.

The US-based National Institute of Standards and Technology (NIST) defines the following five characteristics of cloud computing:

- On-demand self-service: Users can leverage self-services without any intervention from service providers.
- Broad network access: Users can access a network through various terminals.
- Resource pooling: Physical resources are shared by users, and resources in a pool are region-independent.
- Rapid elasticity: Resources can be quickly claimed or released.
- Measured service: Resource measurement, monitoring, and optimization are automatic.

"On-demand self-service" and "broad network access" express enterprises' desire for higher productivity and particularly the need for service automation. "Resource pooling" and "rapid elasticity" can be summarized as flexible resource pools, while "measured services" emphasize that operational support tools are required to tackle the considerable challenges of automation and virtualization. More intelligent and refined tools are also required to reduce the operating expense (OPEX) of enterprises.

Cloud computing is no longer just a term specific to the IT field. Instead, it now represents an entirely new form of productivity, as it creates a business model for various industries, drives industry transformation, and reshapes the industry chain. Cloud computing introduces revolutionary changes to traditional operations and customer experience, and seizing the opportunities of cloud computing will boost growth throughout the industry.

1.2 VIRTUALIZATION TECHNOLOGIES INTRODUCED BY CLOUD COMPUTING

Virtualization is a broad term. According to the Oxford Dictionary, “virtual” refers to something that is “physically non-existent, but implemented and presented through software.” Put another way, a virtual element is a specific abstraction of an element. Virtualization simplifies the expression, access, and management of computer resources, including infrastructures, systems, and software, and provides standard interfaces for these resources. Virtualization also reduces the dependency of service software on the physical environment, enabling enterprises to achieve higher stability and availability based on simplified operation processes, improve resource utilization, and reduce costs.

Throughout the years, virtualization technologies have flourished in the computing, network, and storage domains, and have become interdependent on one another. The development of computing virtualization technologies is undoubtedly critical, while the development of network and storage virtualization technologies is intended to adapt to the changes and challenges introduced by the former. In computing virtualization, one physical machine (PM) is virtualized into one or more virtual machines (VMs) using a Virtual Machine Manager (VMM), which increases utilization of computer hardware resources and improves IT support efficiency.

A VMM is a software layer between physical servers and user operating systems (OSs). By means of abstraction and conversion, the VMM enables multiple user OSs and applications to share a set of basic physical hardware. Consequently, the VMM can be regarded as a meta OS in a virtual environment. It can allocate the correct amount of logical resources (such as memory, CPU, network, and disk) based on VM configurations,

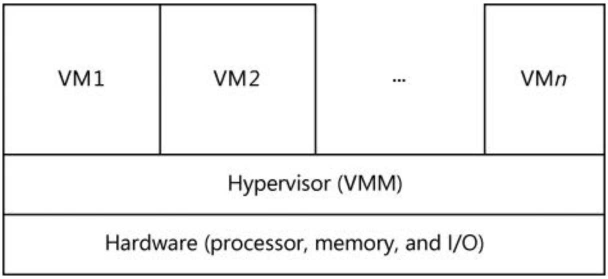

FIGURE 1.1 Virtualization.

load the VM's guest OS, and coordinate access to all physical devices on the VM and server, as shown in Figure 1.1.

The following types of VMMs are available:

- Hypervisor VM: runs on physical hardware and focuses on virtual I/O performance optimization. It is typically used for server applications.
- Hosted VM: runs on the OS of a PM and provides more upper-layer functions such as 3D acceleration. It is easy to both install and use, and is typically utilized for desktop applications.

While multiple computing virtualization technologies exist, they often use different methods and levels of abstraction to achieve the same effect. Common virtualization technologies include the following:

1. **Full virtualization**

   Also known as original virtualization. As shown in Figure 1.2, this model uses a VM as the hypervisor to coordinate the guest OS and original hardware. The hypervisor obtains and processes virtualization-sensitive privileged instructions so that the guest OS can run without modification. As all privileged instructions are processed by the hypervisor, VMs offer lower performance than PMs. While such performance varies depending on implementation, it is usually sufficient to meet user requirements. With the help of hardware-assisted virtualization, full virtualization gradually

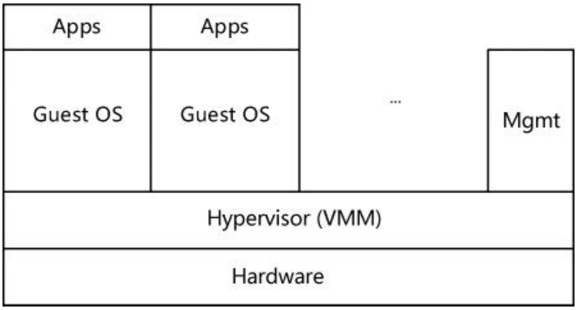

FIGURE 1.2 Full virtualization.

overcomes its bottleneck. Typical hardware products include IBM CP/C MS, Oracle VirtualBox, KVM, VMware Workstation, and ESX.

2. **Paravirtualization**

   Also known as hyper-virtualization. As shown in Figure 1.3, paravirtualization, similar to full virtualization, uses a hypervisor to implement shared access to underlying hardware. Unlike full virtualization, however, paravirtualization integrates virtualization-related code into the guest OS so that it can work with the hypervisor to implement virtualization. In this way, the hypervisor does not need to recompile or obtain privileged instructions, and can achieve performance close to that of a PM. The most well-known product of this type is Xen. As Microsoft Hyper-V uses technologies similar to Xen, it can also be classified as paravirtualization. A weakness of paravirtualization is its requirement that a guest OS be modified, and only a limited number of guest OSs are supported, resulting in a poor user experience.

3. **Hardware emulation**

   The most complex virtualization technology is undoubtedly hardware emulation. As shown in Figure 1.4, hardware emulation creates a hardware VM program on the OS of a PM in order to emulate the required hardware (VM) and runs this on the VM program. If hardware-assisted virtualization is not available, each instruction must be emulated on the underlying hardware, reducing operational performance to less than one percent of that of a PM in some cases. However, hardware emulation can enable an OS designed for PowerPC to run on an ARM processor host without any

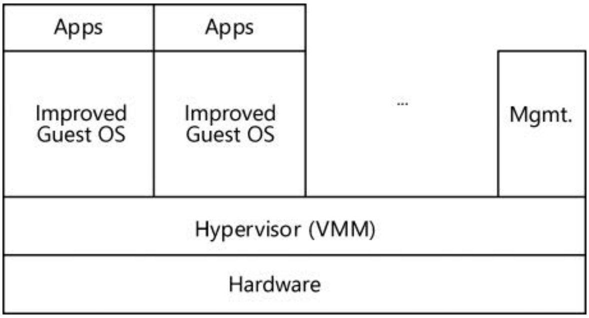

FIGURE 1.3 Paravirtualization.

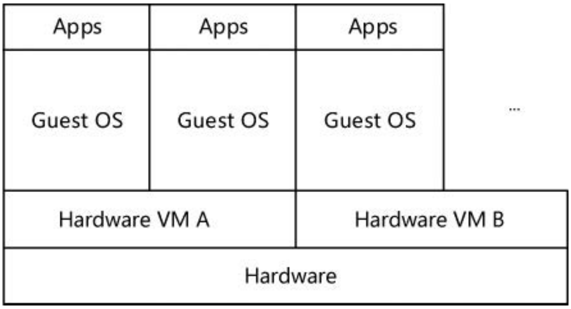

FIGURE 1.4 Hardware emulation.

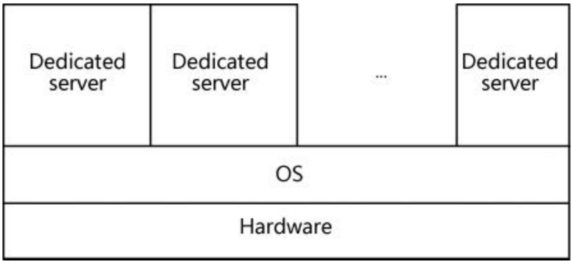

FIGURE 1.5 OS-level virtualization.

modifications. Typical hardware emulation products include Bochs and quick emulator (QEMU).

4. **OS-level virtualization**

   As shown in Figure 1.5, this technique implements virtualization by simply isolating server OSs. As a result, OS-level virtualization can achieve smaller system overheads, preemptive compute resource scheduling, and faster elastic scaling. However, its weaknesses include resource isolation and security. Container technology, as a typical OS-level virtualization technology, is becoming increasingly popular.

5. **Hardware-assisted virtualization**

   Hardware vendors such as Intel and AMD improve virtualization performance by implementing software technologies used in full virtualization and paravirtualization based on hardware. Hardware-assisted virtualization is often used to optimize full virtualization and paravirtualization, rather than operating as a parallel. The best-known example of this is VMware Workstation which, as a full-virtualization platform, integrates hardware-assisted virtualization in VMware 6.0 (including Intel VT-x and AMD-V). Mainstream full virtualization and paravirtualization products support hardware-assisted virtualization and include VirtualBox, KVM, VMware ESX, and Xen.

While the above computing virtualization technologies are not perfect, driven by the changing upper-layer application requirements and hardware-assisted virtualization, they have seen widespread application for a number of years. In 2001, VMware launched ESX, which reshaped the virtualization market. Two years later, Xen1.0 was released and open-sourced. In 2007, KVM was integrated into Linux 2.6.20. And in 2008, Microsoft and Citrix joined forces to launch Hyper-V, while Kubernetes was developing into a mature container technology. Today, these virtualization technologies are still developing rapidly, having not yet reached maturity. In terms of business models, the competition between open source and closed source continues unabated.

Table 1.1 describes the strengths and weaknesses of each virtualization technology.

In addition to computing virtualization, DCN virtualization technologies are also evolving rapidly due to the changes in computing and storage. Virtualization technologies evolved from N-to-1 (horizontal/vertical virtualization) and 1-to-N (virtual switch) technologies to overlay technologies, resulting in large-scale virtual Layer 2 networks (multiple Layer 2 networks combined) capable of delivering extensive compute resource pools for VM migration. SDN was then developed, which associates computing and network resources to abstract, automate, and measure network functions. This technological advancement is driving DCNs toward autonomous driving networks or intent-driven networks. The following chapters will elaborate on these technologies.

1.3 SDN FOR CLOUD COMPUTING

In a cloud DC, virtualized resources are further abstracted as services for flexible use. Cloud computing services can be classified into Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS), which correspond to hardware, software platform, and application resources, respectively.

TABLE 1.1 Strength and Weakness of Virtualization Technologies

	Virtualization Technology				
Item	**Full Virtualization**	**Paravirtualization**	**Hardware Emulation**	**OS-level Virtualization**	**Hardware-Assisted Virtualization**
(e) Speed (compared with physical servers)	30%–80%	More than 80%	Less than 30%	80%	More than 80%
Strengths	The Guest OS does not need to be modified. It is fast and easy to use, and provides useful functions	Compared with full virtualization, it offers a more simplified architecture, which enables a faster speed	The Guest OS does not need to be modified. Typically applicable to hardware, firmware, and OS development	Highly cost-effective	Centralized virtualization provides the fastest speed
Weaknesses	Performance, especially I/O, is poor in hosted mode	The Guest OS must be modified, which affects user experience	Very slow speed. (In some cases, speeds are lower than one percent of that of the physical server)	Limited OS support	The hardware implementation requires more optimization
Trend	Becoming mainstream	Significant use	Phasing out, but still in use	Used for specific applications, such as VPS	Widely used

The requirements for quality attributes vary according to the cloud service layer. IaaS is dedicated to providing high-quality hardware services, while SaaS and PaaS emphasize software flexibility and overall availability. Based on the hierarchical decoupling and mutual distrust principles, they decrease the reliability requirement for a single service to 99.9%, meaning a service can only be interrupted for less than 8.8 hours over the course of a year. For example, users may encounter one or two malfunctions in the email system, instant messaging (IM) software, or even the OS, but no real faults in hardware systems or driver software.

In terms of software technologies, the software architecture and technology selection of cloud services offer varying quality attributes. Legacy software can be classified into IT and embedded software. IT software is applicable to the SaaS and PaaS layers, and focuses on elastic expansion and fast rollout. In fault recovery scenarios, or those that require high reliability, IT software uses methods such as overall rollbacks and restarts. Embedded software focuses more on the control of software and hardware statuses to achieve higher reliability, and is more widely applicable to the IaaS layer.

As IaaS systems become more automated and elastic, and as new software technologies such as distribution, service-orientation, Cloud Native, and Service Less continue to emerge, SaaS/PaaS systems are being subverted, and IT software is undergoing an accelerated transformation to Internet software. Based on the DevOps agile development mode, as well as technical methods such as stateless services and distributed computing, SaaS/PaaS systems provide self-service, real-time online, and quick rollout capabilities for services. This transformation drives the development of new Internet business models.

At the same time, users are beginning to re-examine whether IaaS systems have high requirements on real-time performance across all scenarios and whether refined control over systems is required. This kind of thinking also influences the development of SDN, splitting system development along two different paths: control-oriented and orchestration-oriented.

1. In the current phase, upper-layer services cannot be completely stateless, and the network still needs to detect the computing migration status. In addition, a unified management platform is required to implement fine-grained status management for routing protocols and other information, in order to meet the requirements for fast network switchover in fault scenarios and ensure no impact on

upper-layer services. During the enterprise cloudification process, the following three solutions are provided for the control-oriented path, each of which can be chosen depending on software capabilities and the organizational structure of individual enterprises.

- Cloud-network integration solution: This solution utilizes the open-source OpenStack cloud platform and commercial network controllers to centrally manage network, computing, and storage resources, and implement resource pooling. The cloud platform delivers network control instructions to the network controller through the RESTful interface, and the network controller deploys the network as instructed.
- End-to-end cloud-network integration solution: This solution is also based on a combination of the cloud platform and network controller, and is typically used by commercial cloud platforms such as VMware and Azure. Compared with the open-source cloud platform, the commercial cloud platform offers improved availability and maintainability. As a result, most enterprises use this solution, with only those possessing strong technical capabilities preferring the open-source cloud platform.
- Virtualization solution based on a combination of computing virtualization and network controller: This solution is based on a combination of the virtualization platform (such as vCenter and System Center) and network controller, with no cloud platform used to centrally manage computing and network resources. After delivering a computing service, the virtualization platform notifies the network controller, which then delivers the corresponding network service.

2. The orchestration-oriented path depends on future upper-layer service software architecture, and the expectation that IaaS software will be further simplified after SaaS/PaaS software becomes stateless. The status of the IaaS software does not need to be controlled. Instead, IaaS software only needs to be orchestrated using software tools, similar to the upper-layer system.

   Selecting the appropriate path or solution depends on many factors such as the enterprise organization structure, existing software architecture, technology and resource investment, and cloudification progress. Although all four solutions have their strengths and weaknesses, they can support enterprise service cloudification over a long term. The following describes what features each solution offers and why an enterprise might select them.

1. Cloud-network integration solution: (OpenStack + network controller): This solution has the following strengths:
   - As OpenStack is the mainstream cloud platform, the open-source community is active and provides frequent updates, enabling rapid construction of enterprise cloud capabilities.
   - Customized development can be applied to meet enterprise business requirements. These are subject to independent intellectual property rights.
   - OpenStack boasts a healthy ecosystem. Models and interfaces are highly standardized, and layered decoupling facilitates multi-vendor interoperability.

   This solution has the following weaknesses:

   - Commercialization of open-source software requires software hardening and customization. As such, enterprises must already possess the required technical reserves and continuously invest in software development.
   - During the current evolution from single to multiple DCs, and to hybrid clouds, a mature multi-cloud orchestrator does not yet exist within the open-source community. This will need to be built by enterprises themselves.
   - Enterprises that choose this solution generally possess advanced software development and integration capabilities, are concerned about differentiated capabilities and independent intellectual property rights, and require standardization and multi-vendor interconnection. Currently, this solution is mainly used for carriers' telco clouds, and for the DCs of large financial institutions and Internet enterprises.

2. End-to-end cloud-network integration solution (Huawei FusionCloud, VMware vCloud, and Microsoft Azure Stack):

   The advantages of this solution are its full support for single/multiple DCs, hybrid clouds, and SaaS/PaaS/IaaS end-to-end delivery, which enable the rapid construction of cloudification capabilities for enterprises.

   However, areas where this solution requires improvement include openness and vendor lock-in.

   Enterprises that choose this solution are often in urgent need of cloud services. They need to quickly build these services in a short period of time to support new business models and quickly occupy the market, but lack the required technical reserves. As such, this solution is typically employed for DCs of small- and medium-sized enterprises.

3. Virtualization solution (computing virtualization+network controller):

   This solution has the following strengths:

   - In the virtualization solution, computing and network resources are independent of one another, and the enterprise organization (IT and network teams) does not require immediate restructuring.
   - A large number of cloud and PaaS/SaaS platforms are currently available, and the industry is considered to be mature. Consequently, the association between computing virtualization and network controllers enables the rapid construction of IaaS platforms capable of implementing automation and satisfying service requirements. This approach is low risk and is easy to initiate.
   - The open architecture of the IaaS layer and layered decoupling allow flexible selection of the cloud and PaaS/SaaS platforms.
   - This solution is based on mature commercial software, which requires no customized development and offers high reliability.

     A weakness of this solution is that IaaS/PaaS/SaaS software models need to be selected in rounds, which slows down the cloudification of enterprises.

Enterprises that choose this solution have complex organizational structures and fixed service applications. They do not want to lock vendors, but have major concerns relating to solution stability and reliability. As such, this solution is typically implemented in the DCs of medium- and large-sized enterprises in the transportation and energy industries.

Following the development of container technologies, compute and storage resources are becoming less dependent on networks. In this solution, the network controller can evolve into an orchestrator, and the overall solution can evolve to tool-based orchestration.

4. Tool-based orchestration solution:

   This solution has the following strengths:

   - Provides customized service orchestration capabilities to adapt to enterprise service applications.
   - Easy to develop and quick to launch, as it is based on script or graphical orchestration tools.

   This solution has the following dependencies or weaknesses:

   - The IaaS, PaaS, and SaaS software must comply with the stateless principle, and service reliability is independent of regions and does not rely on the IaaS layer (VM migration).
   - Service application scenarios are relatively simple and services are independent of each other, preventing conflicts and mutual coverage impacts.

   Enterprises that choose this solution do not rely heavily on legacy service software, or can begin restructuring at low costs. They operate clear service scenarios, and require rapid responses to service changes. In addition, they can apply strict DC construction specifications to ensure service independence. This solution is typically implemented in the DCs of Internet enterprises.

Enterprises have varying concerns about their network capabilities when choosing from the available solutions, as shown in Table 1.2.

TABLE 1.2 Enterprise Concerns When Choosing Cloudification Solutions

	Single DC		Multiple DCs and Hybrid Cloud	
Solution Chosen	**Enterprise DC Scenario**	**Telco Cloud Scenario**	**Enterprise DC Scenario**	**Telco Cloud Scenario**
Cloud-network integration solution (OpenStack+ network controller)	• OpenStack interconnection automation capability • Networking and forwarding performance (network overlay) • VAS multi-vendor automation capability • PaaS software integration capability • O&M capabilities: Zero Touch Provisioning (ZTP), dialing test methods, and underlay network automation • IPv6 capability	• Networking diversity and layered decoupling (hybrid overlay) • OpenStack interconnection automation • Capability (Layer 2, Layer 3, and IPv6) • SFC capability • Routing service (BGP and BFD) • Standardization (BGP-EVPN, and MPLS) • Quality of Service (QoS)	• Multiple OpenStack platforms • Unified resource management • Security automation • Multi-DC O&M capabilities	• Multi-DC inter-operation standardization • Automation of interconnection between MANs and DCs
End-to-end cloud-network integration solution	• Physical server access capability (physical switches connected to the VMware NSX Controller to implement physical server automation) • Rapid VMware vRealize integration capability • Automated the underlay network O&M capability (hardware switches connected to VMware vRNI and Azure Stack) • Hybrid cloud			
Virtualization solution (computing virtualization platform+network controller)	• Fine-grained security isolation (microsegmentation) • Multi-vendor VAS device automation capability (SFC) • Underlay network O&M (ZTP and configuration automation) • Multicast function • IPv6 capability		• Association with VMware and network DR and switchover • Unified management of multiple DCs • Forward compatibility and evolution	
Tool-based orchestration solution	• Interconnection between network devices and orchestration tools such as Ansible and Puppet • Openness of O&M interfaces on network devices • Response speed of customized interfaces on network devices			

To summarize, enterprises should choose a cloud-based transformation solution capable of matching their specific service requirements and technical conditions.

1.4 DCN PROSPECTS

1. Intent-driven network

   According to Gartner's technology maturity model, shown in Figure 1.6, SDN/NFV technologies are now ready for large-scale commercial use following many years of development.

   In the future, as the development of automation, big data, and cloud technologies continues, autonomous networks (ANs) will gradually be put into practice, once again driving the rapid development and evolution of the entire industry. However, SDN still has a long way to go before reaching the AN goal. While SDN technology activates physical networks through automation, there are still broad gaps separating business intent and user experience. For example, an enterprise's business intent is to quickly expand 100 servers due to the expected service surge of a big event. To address this intent, the enterprise needs to perform a series of operations, such as undoing interface shutdown, enabling LLDP, checking the topology, and enabling the server to manage the network. There are a lot of

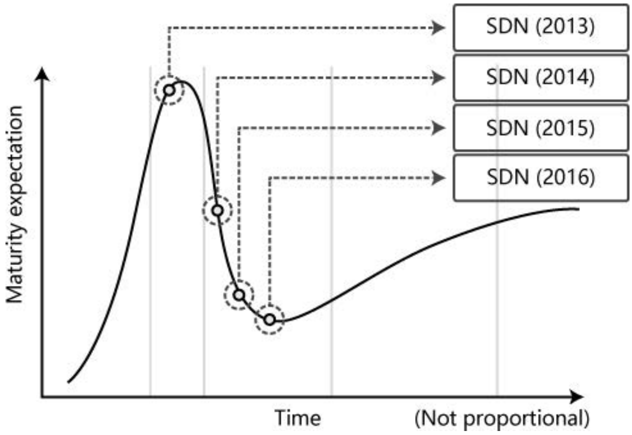

FIGURE 1.6 Gartner's technology maturity model.

expenses involved in implementing this business intent based on the network.

In this case, a digital world must be constructed over the physical network. This not only digitizes a physical network element (NE), but also quickly maps business intents to network requirements and digitizes user experience and applications on the network. Based on automation, big data, and cloud technologies, the digital world transforms from device-centric to user-centric, and bridges customers' business intents and physical networks. Intent-driven networks can quickly provide services based on the digital world, improve user experience, and enable preventive maintenance.

2. AI Fabric

   As Internet technologies are developing, upper-layer DC services are shifting their focus from service provisioning efficiency to data intelligence and business values. The three core elements of AI applications are algorithms, computing power, and data. Among the core elements, data is the most critical. All AI applications use advanced AI algorithms to mine intelligence from data and extract useful business value. This poses higher requirements on a DC's IaaS platform.

   To start with, AI applications require the IaaS layer to provide a high-performance distributed storage service capable of carrying massive amounts of data, and the AI algorithms require a high-performance distributed computing service capable of massive data computing. It is estimated that by 2025, the amount of data generated and stored worldwide will reach 180 ZB. Such an incredible volume of data will be beyond the processing capacity of humans, leaving 95% to be processed by AI. Service requirements drive the rapid development of Solid-State Disks (SSDs) and AI chips, and the sharp increase in communication services between distributed nodes leads to more prominent network bottlenecks.

   - Current storage mediums (SSD) deliver access speeds 100 times faster than conventional distributed storage devices (such as hard disk drives). In addition, network delay rates have increased from less than 5% to about 65%. There are two types of network delay: delay caused by packet loss (about 500 μs) and queuing delay caused by network congestion (about 50 μs). Avoiding packet loss and congestion is a core objective for improving input/output operations per second (IOPS).

- AI chips can be anywhere between 100 and 1000 times faster than legacy CPUs. In addition, the computing volume of AI applications increases exponentially. For example, distributed training for a large-scale speech recognition application results in the training quantity for a computing task reaching approximately 20 exaFLOPS, requiring 40 CPU-installed servers to calculate more than 300 million parameters (4 bytes for a single parameter). In each iterative calculation, the CPU queuing delay (approximately 400 ms) exceeds the CPU calculation delay (approximately 370 ms). If millions of iterative calculations are used during one training session, it will last for an entire month. Reducing the communication waiting time and shortening AI training have become core requirements of AI distributed training.

To meet the requirements of AI applications, network protocols and hardware have been greatly improved. In terms of protocols, Remote Direct Memory Access (RDMA) and RDMA over Converged Ethernet (RoCE) alleviate TCP problems such as slow start, low throughput, multiple copies, high latency, and excessive CPU consumption. In terms of hardware, Ethernet devices have made great breakthroughs in lossless Ethernet.

- Virtual multi-queue technology is used to precisely locate back pressure in congestion flows, which prevents impacts on normal traffic.
- The congestion and back pressure thresholds are dynamically calculated and adjusted in real time, ensuring maximum network throughput without packet loss.
- The devices proactively collaborate with the NIC to schedule traffic to the maximum quota and to prevent congestion. As such, next-generation lossless Ethernet adaptive to AI equals, or even exceeds, the InfiniBand (IB) network in terms of forwarding performance, throughput, and latency. From the perspective

of overall DC operation and maintenance, a unified converged network (convergence of the storage network, AI computing network, and service network), which is considerably more cost-effective, can be built based on Ethernet.

AI Fabric is a high-speed Ethernet solution based on lossless network technologies. It provides network support for AI computing, high-performance computing (HPC), and large-scale distributed computing. AI Fabric uses two-level AI chips and a unique intelligent algorithm for congestion scheduling to achieve zero packet loss, high throughput, and ultra-low latency for RDMA service flows, improving both computing and storage efficiency in the AI era. Private network performance is now available at the cost of Ethernet, delivering a 45-fold increase in overall Return on Investment (ROI). For more on AI Fabric, see Chapter 11.