International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

F.743.6
(08/2018)

SERIES F: NON-TELEPHONE TELECOMMUNICATION SERVICES

Multimedia services

Service requirements for next generation content delivery networks

Recommendation ITU-T F.743.6

ITU-T F-SERIES RECOMMENDATIONS

NON-TELEPHONE TELECOMMUNICATION SERVICES

TELEGRAPH SERVICE	
Operating methods for the international public telegram service	F.1–F.19
The gentex network	F.20–F.29
Message switching	F.30–F.39
The international telemessage service	F.40–F.58
The international telex service	F.59–F.89
Statistics and publications on international telegraph services	F.90–F.99
Scheduled and leased communication services	F.100–F.104
Phototelegraph service	F.105–F.109
MOBILE SERVICE	
Mobile services and multidestination satellite services	F.110–F.159
TELEMATIC SERVICES	
Public facsimile service	F.160–F.199
Teletex service	F.200–F.299
Videotex service	F.300–F.349
General provisions for telematic services	F.350–F.399
MESSAGE HANDLING SERVICES	F.400–F.499
DIRECTORY SERVICES	F.500–F.549
DOCUMENT COMMUNICATION	
Document communication	F.550–F.579
Programming communication interfaces	F.580–F.599
DATA TRANSMISSION SERVICES	F.600–F.699
MULTIMEDIA SERVICES	**F.700–F.799**
ISDN SERVICES	F.800–F.849
UNIVERSAL PERSONAL TELECOMMUNICATION	F.850–F.899
ACCESSIBILITY AND HUMAN FACTORS	F.900–F.999

For further details, please refer to the list of ITU-T Recommendations.

Recommendation ITU-T F.743.6

Service requirements for next generation content delivery networks

Summary

Recommendation ITU-T F.743.6 describes use cases, service requirements and challenges of the next generation content delivery network (CDN). A reference architecture named access CDN for meeting these requirements and challenges, which decouples the common functions and customized services, is also defined.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T F.743.6	2018-08-29	16	11.1002/1000/13657

Keywords

Content delivery network (CDN), next generation content delivery network (NG CDN).

* To access the Recommendation, type the URL http://handle.itu.int/ in the address field of your web browser, followed by the Recommendation's unique ID. For example, http://handle.itu.int/11.1002/1000/11830-en.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure, e.g., interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database at http://www.itu.int/ITU-T/ipr/.

Table of Contents

Introduction

With the popularity of video services such as 4K/8K television (TV), virtual reality/augmented reality (VR/AR) and live sports events, demands on network throughput are increasing by orders of magnitude along with the associated requirements for ultra-low latency. These bring new challenges for the current content delivery network (CDN), resulting in a trade-off between exponential increases in the network cost and the need for smooth growth in end-user access price. It will be difficult to cover the throughput magnitude gap by solely relying on a pure underlay network; thus the next generation CDN should solve backbone traffic convergence and improve end-user experience to resolve this gap, and accelerate the introduction of emerging services.

Recommendation ITU-T F.743.6

Service requirements for next generation content delivery networks

1 Scope

This Recommendation describes use cases as the driving force for service requirements, and challenges for the next generation content delivery network (NG CDN). It also provides a reference architecture to meet these service requirements.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-T Y.2019] Recommendation ITU-T Y.2019 (2010), *Content delivery functional architecture in NGN*.

3 Definitions

3.1 Terms defined elsewhere

None.

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

3.2.1 access CDN (A-CDN): A reference architecture for a highly distributed CDN. The architecture horizontally decouples the current vertically integrated CDN system into the content delivery network control part (CDN-CP) for end-to-end (E2E) service control and the content delivery network delivery part (CDN-DP) for distributed deployment.

3.2.2 CDN control part (CDN-CP): Provides customized services including dynamic content acceleration to meet customers' requirements, which is provided by the content delivery network service provider (CDN SP) and Internet content provider (ICP).

3.2.3 CDN delivery part (CDN-DP): Provides a generic content delivery function and open network capability, including real time multi-cast, to support content delivery network service provider (CDN SP) and Internet content provider (ICP) basic requirements provided by operators.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

A-CDN	Access CDN
AR	Augmented Reality
CDN	Content Delivery Network
CDN-CP	Content Delivery Network Control Part
CDN-DP	Content Delivery Network Delivery Part

CDN-SP	Content Delivery Network Service Provider
E2E	End-to-End
HD	High Definition
ICP	Internet Content Provider
IoT	Internet of Things
ISP	Internet Service Provider
IXP	Internet Exchange Point
NG CDN	Next Generation of CDN
OTT	Over the Top
QoE	Quality of Experience
QoS	Quality of Service
ROI	Range of Interest
SD	Standard Definition
TV	Television
VR	Virtual Reality

5 Conventions

None.

6 Next generation content delivery network use cases and requirements

The existing hot services and merging of new technologies and services such as 4K/8K television (TV), virtual reality/augmented reality (VR/AR), live streaming, and Internet of things (IoT) drive the next generation of CDN (NG CDN) with requirements entering a new decade of development. As a driver, some of the typical use cases of NG CDN are briefly discussed below.

6.1 4K/8K/VR/AR

For VR/AR applications, latency is the most important performance parameter. VR/AR is latency sensitive because human perception requires accurate and smooth movements in vision. Increased latency can lead to a detached VR experience and contribute to a motion-sickness sensation. VR/AR industries agree that application latency should be less than 20 ms for the motion-to-photon latency to be imperceptible. Range of interest (ROI) needs to be regenerated once a user's head turns, with latency less than 20ms. Content cannot be downloaded wholly with limited access bandwidth, while the terminal requires content delivery at a dynamic ROI zone.

The network bandwidth for VR/AR is the throughput required to stream 360 degree video to a user, which is another critical parameter for VR/AR applications. 8K quality and above is necessary for VR, as 4K VR 360 degree video is only equivalent to 240p on TV. VR 360 degree video experience may evolve through the stages listed in Table 1, requiring a throughput of 400 Mbit/s and above, more than 100 times higher than the current throughput of high definition (HD) video services

Table 1 – VR network requirements (bandwidth and latency)

VR resolution		Equivalent TV resolution	Bandwidth	Latency
Early stage VR (current)	1K*1K@visual field 2D_30fps_8bit_4K	240p	25 Mbit/s	40 ms
Entry-Level VR	2K*2K@visual field 2D_30fps_8bit_8K	SD	100 Mbit/s	30 ms
Advanced VR	4K*4K@visual field 2D_60fps_10bit_12K	HD	400 Mbit/s	20 ms
Extreme VR	8K*8K@visual field 3D_120fps_12bit_24K	4K	1 Gbit/s (smooth play) 2.35 Gbit/s (interactive)	10 ms

Due to the acute traffic increase, change in convergence model and high requirement of VR videos on bandwidth, latency and packet loss, traditional high-convergence networks face the following challenges:

– low network efficiency: Expansion of end-to-end (E2E) network equipment is needed as video traffic increases. CDN is deployed in a higher position and away from the end user; service traffic travels through more network equipment, which therefore occupies a greater number of network resources. In addition, more equipment increases the probability of bottlenecks, blocking and E2E latency;

– poor user experience: With multi-service concurrency, the network utilization ratio increases and packet loss and delay also increase.

High resolution and bitrates of 4K/8K/VR cause a rapid rise of delivery cost, with standard definition (SD) delivery already occupying 40% - 60% of the cost, which becomes the biggest obstacle for users to enjoy 4K/VR. NG CDN requires low cost with high performance high bit-rate content delivery.

6.2 Live streaming

Live video brings a much higher demand on concurrency, continuity and low latency. Taking user generated content (UGC) mode for example, hosts are distributed; a traditional centralized CDN could cause extra latency and packet loss while content is uploading. Future live streaming emphasizes liveness and interaction, with video lag ratio under 2%, and latency less than 3s. Spiky unicast traffic of live events creates congestion in transit, core and aggregation. Traditional centralized CDN nodes cannot meet the requirement of live streaming service, since:

– CDN is a static multi-level cache system with circuitous traffic. It does not take network and topology status into considerations, making it difficult to meet the real time and interactive demand of live streaming;

– live streaming only requires forward instead of traditional CDN storage-forward mode. Further, the traditional server architecture has inferior performance for live streaming.

Thus, a new architecture is required for CDN services with high performance and edge live streaming as a service.

6.3 IoT/edge computing

The rapid emergence of IoT will connect numerous devices to the network, generating high concurrency data analysis, storage and processing in the network edge. The core router will be heavily loaded and will become a bottleneck if all traffic flows completely back to the data centre. Distribution of CDN nodes enables and strengthens IoT in data pre-processing, storage and forwarding that is specific to large scale thing-to-thing communication. Data could be filtered, merged and lightweight computed in the CDN edge node and then sent back to the source centre.

From an end-user perspective, users are more and more willing to pay for a better experience. With the development of IoT, NG CDN needs to guarantee user experience in not only traffic and speed, but also in numerous devices, and large numbers of computing and learning tasks.

CDN faces the following challenges when adapting to IoT:

– capacity at the edge: Edge computing enables IoT due to the lack of bandwidth infrastructure. Points of presence placed closer to geographic location allows users to interact more quickly with content;

– technology to move information faster: Data has to be moved reliably and on time consistently. CDN should be able to securely accelerate enterprise applications;

– data security: With growing IoT units, the number of distributed denial of service (DDoS) attacks is expected to increase exponentially in the coming years. Research shows that 85% of enterprises around the world are deploying or intend to deploy IoT, while only 10% feel confident in securing devices.

7 NG CDN requirements summary

As HD/ultra-high definition (UHD) and VR/AR video, live streaming become basic services of operators, network infrastructure needs to support these high-quality services. Thus, the CDN caching, storage and computing abilities should be taken into consideration. For a video service-centric featured future network, CDN service capabilities will become a basic infrastructure for operators, thus becoming a key feature for network reconfiguration and redesign. It is possible that "pay by traffic mode" will enable CDN capabilities to become an operators' main income source (rather than bandwidth offering paid per month). These capabilities are illustrated in Figure 7-1.

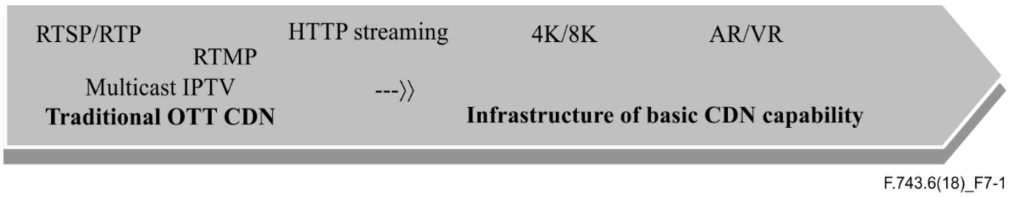

Figure 7-1 – CDN capabilities becomes basic infrastructure of video service-centric featured future network

The emergence of new technologies and services such as high-definition television (HDTV), VR/AR and IoT, require NG CDN to be able to meet the requirements of ultra-low latency, high throughput, multi-direction traffic, and experience lossless content delivery. These considerations, as illustrated in Figure 7-2, include:

– ultra-low latency, improving from seconds to milliseconds, meeting challenges brought by live streaming, live events, network gaming, IoT, etc.;

– ultra-high throughput, improving from kilobits/s or megabits/s to gigabits/s meeting challenges brought by 4K/8K TV, VR/AR, etc.;

– multi-direction delivery from a few sources in the source-tree to everyone as source in a source mesh;

– user experience of lossless delivery to guarantee the content and latency in high reliability.

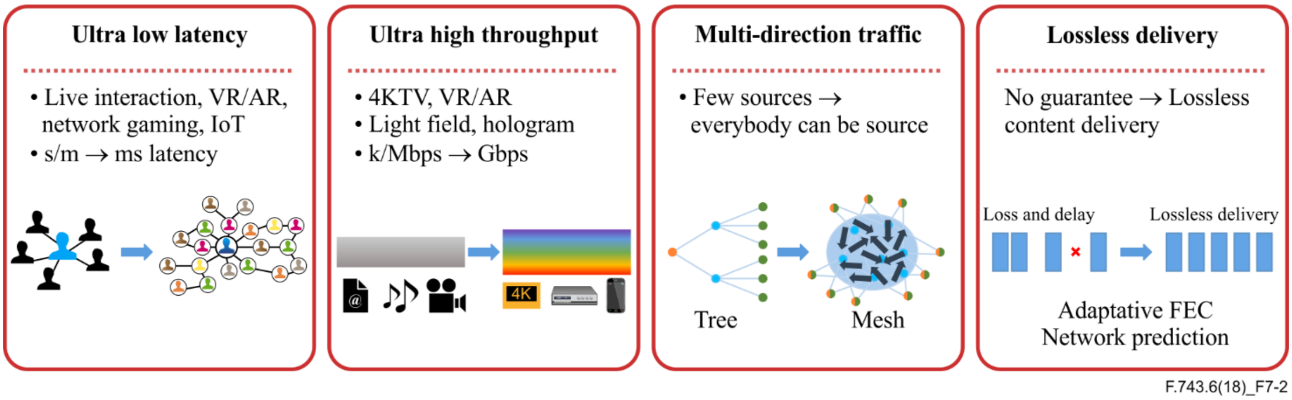

Figure 7-2 – NG CDN requirements

8 Challenges of NG CDN

With the emergence of new technologies and services such as VR/AR, IoT and holographic interaction, CDN faces issues for accelerating and delivering dynamic and interactive content instead of traditional web services. Challenges of current CDN include:

– the backbone/metro network construction and expansion cannot meet the requirements from the increase of video and new emerging services. Current CDN nodes are often highly placed at the backbone/metro network. This type of CDN network architecture cannot fit these new services well since it poses a great challenge on the backbone/metro network capacity and the E2E delivery performance;

– the traditional storage and forwarding models have almost no considerations for dynamic routing to guarantee high bandwidth and low latency experiences, and no mechanisms to optimize traffic direction and mesh network capabilities;

– the single node of current CDN is usually located in the backbone/metro network, and in low-cost regions geographically far from users. This makes it hard for current CDN to cover high-rate/high burst traffic and strongly interactive services;

– the lack of data sharing and CDN interworking between CDN/over-the-top (OTT) and network operators does not fully take advantage of the potential of the access network; that is, the important "last mile" role in the CDN value chain;

– disorderly traffic flow and out-of-order traffic causes high network transmission costs; this is undesirable from the perspective of network operators;

– storage capacity and throughput are the main contradictions which CDN should handle. There is a gap between the requirement of CDN edge nodes and source nodes. How the most suitable storage media for different storage capacity and throughput requirements of different content types should be considered. Table 2 provides examples of storage capacity and throughput for different content types.

Table 2 – Examples of storage capacity and throughput for different content types

Content type	Storage capacity	Throughput
Image	3.4 PB	7 Tbit/s
Standard definition video	403 TB	30 Tbit/s
4K	24 PB	1.8 Pbit/s
VR	2.4 EB	180 Pbit/s

9 Reference architecture to meet service requirements

As the background analysis in Appendix I shows, even though major players in the CDN industry, including cloud providers, CDN service provides (SPs) and operators, are tending to bring CDN nodes to the edge in different ways, there are still issues facing the requirements of NG CDN.

The root cause is that CDN, as a typical overlay network, cannot perceive the underlay network status. An overlay network that ignores the properties of the underlay infrastructure can lead to congestion and reliability issues.

– some CDN SPs are calling for operators for open network capabilities such as topology information and dynamic running data for targeting performance optimization, quality of service (QoS) guarantees for insuring differentiated services and E2E service ability via tight coordination with underlay network;
– from an operators' perspective, they believe it is not appropriate to open the network to provide this key information. However, operators are suffering from pressures brought by rapid growing bandwidth, and also disordered traffic.

The existing ITU-T CDN Recommendation [ITU-T Y.2019], published in 2010, describes the content delivery functional architecture in NGN. It has no consideration for the trend of bringing CDN nodes to the edge, which is driven by modern emerging services such as 4K, 8K, AR, VR etc., and with the feature of ultra-high throughput and ultra-low latency.

Thus, there is a demand for an updated solution to meet both operators' and CDN SP's requirements. This new technology and architecture should be considered through tight coordination between overlay and underlay networks to improve the delivery quality while not bring backbone network heavy investment pressure to the operators.

A reference architecture for highly distributed CDN, termed access CDN (A-CDN) is illustrated in Figure 9-1. The architecture horizontally decouples the current vertically integrated CDN system into the content delivery network control part (CDN-CP) for E2E service control and the content delivery network delivery part (CDN-DP) for distributed deployment. The key features of this architecture are:

– CDN-CP: Provides customized services including dynamic content acceleration to meet customers' requirements, which is provided by a CDN SP and Internet content provider (ICP);
– CDN-DP: Provides a generic content delivery function and open network capability, including real time multicast to support the CDN SP and ICP's basic requirement, which is provided by operators;
– A-CDN underlay network: Enhances the underlay network capability for a CDN service, which is provided by operators;
– both CDN-DP and A-CDN underlay networks combine to be multi-telecom A-CDN, an open platform for distributed CDN. Open application programming interfaces (APIs) are provided by the platform via which network capabilities such as multicast can be opened to the CDN SP and ICP.

This architecture can utilize each role's strength. For SPs, they are strong at CDN service customization. For operators, they have rich central office resources and can provide QoS guarantees through coordination with the underlay.

The benefits for SPs are footprint extension and user equipment (UE) enhancement. The benefits for operators are high throughput video traffic offload and localization.

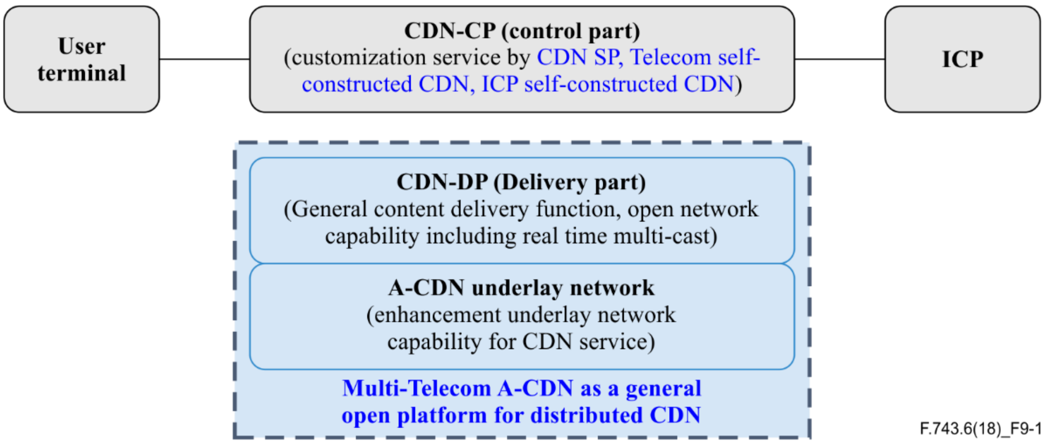

Figure 9-1 – Access-CDN - a recommended architecture for highly distributed CDN

The values generated through coordination between an operator A-CDN platform and the underlying network are illustrated in Figure 9-2. Some of these values include:

– overlay and underlay coordination for more appropriate content placement (to ensure the hit rate and reuse rate);

– routing, distribution and scheduling considering the underlay network topology (local requests and local scheduling based on network status);

– within the same administration domain inside one telecom provider, the operator can deliver all types of services by opening the underlay capability, such as multicast services, QoS guaranteed services, security guaranteed services or other personalized services.

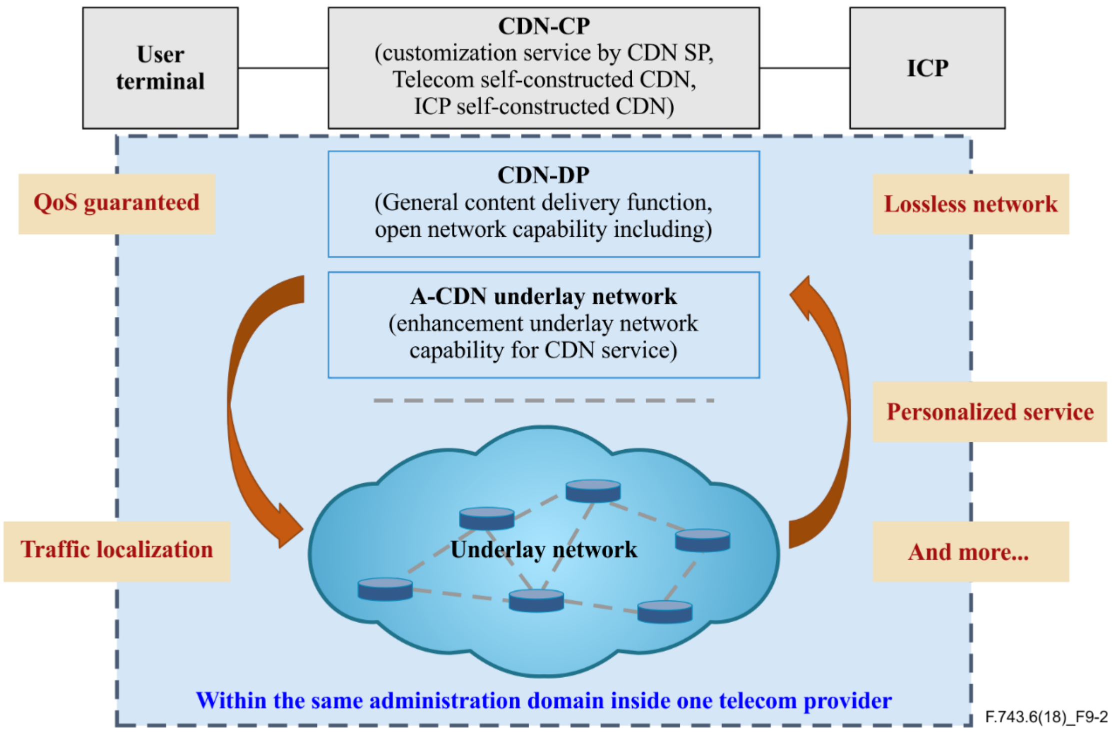

Figure 9-2 – Coordination between access-CDN platform and the underlay network

Appendix I

Existing modes for bringing CDN to the edge

(This appendix does not form an integral part of this Recommendation.)

CDN nodes are shifting from the backbone network (the traditional CDN deployment position) to the edge, which implies a metro network architecture as illustrated in Figure I.1, with lower convergence ratio; this has become common among CDN SPs and operators.

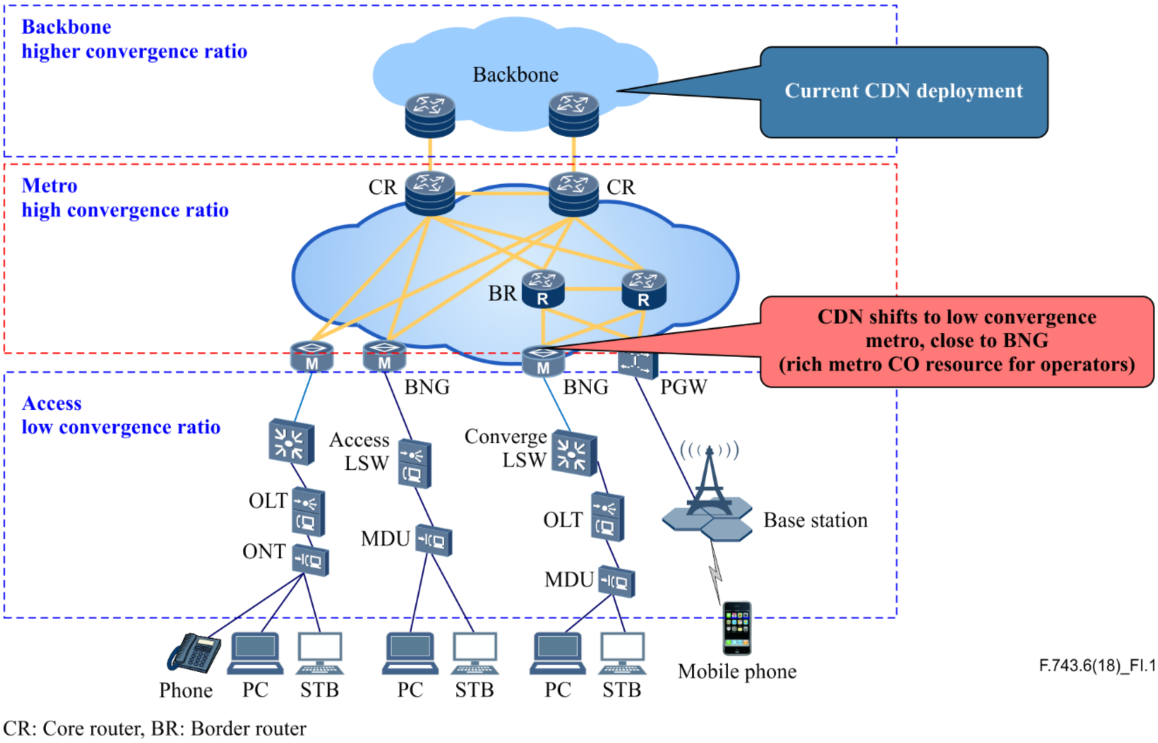

Figure I.1 – CDN nodes shift from backbone network to metro network, which is close to the border network gateway (BNG)

There are several benefits. First, significant saving on backbone bandwidth for the Internet service provider (ISP). Second, quality of experience (QoE) of ISP's customers is improved with a distributed architecture closer to the user.

An analysis of current mainstream modes to bring CDN to the edge are described below.

Mode 1 – peering (some CDN service providers and also operators who build global CDN):

– for footprint extension, some CDN SPs are peering with thousands of Internet exchange points (IXP) points of presence to interconnect with ISPs that provide direct paths to access networks;

– for better user experience, some CDN SPs also deploy servers in ISP network utilizing operators' site and bandwidth for lower latency and packet loss rate;

– they have deployed servers in networks across countries, able to manage routing and delivery quality to geographically diverse users. Servers are typically placed in tier 2, and tier 1 networks;

– some operators' global CDN construction also achieve broad coverage by peering mode, by interconnecting with thousands of global ISPs via IXP mode;

– pros: Low cost and easy deploying for inter-region, inter-network coverage;

– cons: Requires heavy investment on backbone for dozens of times bandwidth growth. But operators are unwilling since there is no profit model. Thus, peering mode will cause packet loss due to congestion and latency due to long distances from users. Thus, the future requirements cannot be met. 4K/8K/VR video requires high throughput as bandwidths move from Mbit/s to Gbit/s for a single user.

Mode 2 – self-build CDN + ISP cooperation mode (global OTT service providers):

– some large OTT service providers are running their own CDN nodes for the following reasons:
  - grow faster: traditional CDN SP cannot meet the requirement of fast growing CDN bandwidth;
  - reduced cost: especially when it comes to a certain scale;
  - enhanced video QoE, including customized services;

– deployed CDN nodes down and enhance service quality by cooperation with ISPs for better user experience:
  - OTT service providers provide CDN server and software, ISP provide site and bandwidth. No need to go to the traffic bottlenecks such as backbone network and IXP nodes when ISP user is accessing their service;
  - although OTT service providers spend almost no money on bandwidth, CDN server costs are very high in order to offload up to 80% of traffic according to ISP's requirement;

– pros: Solving the problem of ISP backbone bandwidth expansion;

– cons: Since there are numerous ISPs worldwide, CDN node number will be huge to deploy down to the edge. Furthermore, massive storage should be deployed to offload 80% traffic to meet the requirement of ISP, which causes the storage cost to be very high.

Mode 3 – self-build CDN (some CDN SPs and OTT service providers in China):

– China CDN SPs and OTT service providers have already brought CDN to the edge (e.g., city Internet data center (IDC)). Yet disordered traffic still exists since the focus is on usage rate of CDN servers to reduce storage cost rather than localized scheduling. Statistics by China Telecom in 2015 indicate that less than 20% of traffic flows are localized.

– cons: Disordered traffic causes high network transmission costs to the operator; also no experience guarantee is provided:
  – insufficient number of nodes, limited coverage, with no reliable user experience;
  – dispatching service off-site means inter-network traffic settlement cost is high;
  – CDN service scheduling is not aware of the underlying network status. Traffic flow with no reasonable guidance and experience is not guaranteed.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	Tariff and accounting principles and international telecommunication/ICT economic and policy issues
Series E	Overall network operation, telephone service, service operation and human factors
Series F	**Non-telephone telecommunication services**
Series G	Transmission systems and media, digital systems and networks
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Environment and ICTs, climate change, e-waste, energy efficiency; construction, installation and protection of cables and other elements of outside plant
Series M	Telecommunication management, including TMN and network maintenance
Series N	Maintenance: international sound-programme and television-transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling, and associated measurements and tests
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks, open system communications and security
Series Y	Global information infrastructure, Internet protocol aspects, next-generation networks, Internet of Things and smart cities
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2018