White Paper on the Experience-driven Bearer Network

Issue 01

Date 2016-04-11

HUAWEI TECHNOLOGIES CO., LTD.

Trademarks and Permissions

Notice

Huawei Technologies Co., Ltd.

Address: Huawei Industrial Base
Bantian, Longgang
Shenzhen 518129
People's Republic of China

Website: http://www.huawei.com

Email: support@huawei.com

Contents

Tables

Figures

White Paper on the Experience-driven 4K Bearer Network

Keywords: 4K, experience, U-vMOS, bearer network, high throughput, 100 Mbps, 20 ms, 10^{-5}, Wi-Fi, FTTH, flatten, HTR, OTN, CDN, KQI

Abstract:

Video continuously drives the ultra-broadband development and changes user behavior and demands. Carriers have begun to focus more on experience instead of connections, with networks being reconstructed to be service-driven and experience-driven. A user experience-centric ultra-broadband network is all in need to provide best-experience 4K video experience for users. This white paper defines the experience-driven 4K bearer network and describes the target network architecture and deployment recommendations. Carriers can select a suitable deployment solution based on their own network infrastructure and service development strategies.

About This Document

Carriers have been on the search for a key service offering able to satisfy user requirements, fully demonstrate the capabilities of ultra-broadband networks, and create business value. 4K, the new benchmark in video, offers carriers a compelling point for growth.

Single-channel 4K requires 30–50 Mbps throughput. As the prices for 4K terminals continue to come down, families are beginning to purchase multiple 4K terminals for their home viewing pleasure. This means that 100 Mbps up to 1000 Mbps bandwidth will become the basic expectation for home broadband users. 4K video releases the value of ultra-broadband and provides a better service experience to users. Users are rather willing to pay for a better service experience, which in turn means that consumers, carriers, and content providers all win and a positive business cycle is formed. Seeing the new opportunity in 4K video technologies, more than 60 carriers worldwide have announced a video strategy. Video is fast becoming a basic service offering for carriers.

There are, however, concerns in experience with the rising popularity of video services. According to a report from Conviva, 35% of responding users stated that the viewing experience is the number one determining factor in their selection of a service provider (higher than that of the video content being offered). One-third of the users stated they become frustrated when video stalling occurs and stop watching immediately while 84% stop watching if the picture quality turns poor for longer than one minute. Considering these findings, ensuring a better viewing experience is crucial to commercial success.

1 Experience-centric 4K Bearer Network

1.1 Experience-centric Network Construction Concept

The New Drivers: Service and Experience

Bandwidth has become the "new utility" essential to work and life. Consumer expectations on bandwidth are changing from merely increases in bandwidth to include superior service experience and quality. Carriers have begun to focus more on experience instead of connections, with networks being reconstructed to be service-driven and experience-driven.

The 4K Video Era in Here!

Many carriers worldwide have included video services in their strategic ICT transformations. With the popularity of 4K terminals, provisioning of 4K services is crucial to carriers wanting to build differentiated competitive advantages into their offering portfolios. For example, Deutsche Telecom is taking on paid TV programming providers with 4K video services, and the carrier is making inroads. British Telecom has built an NGA ultra-broadband network and launched its BT Sport video service, achieving 23% business growth. In yet another example, China Telecom Sichuan now has over five million users to its video service, and 1.2 million of those are 4K video users and that number continues to grow rapidly.

Higher Requirements on Networks from 4K Video

4K does not only mean an enhancement in video resolution, it also entails other major improvements in video quality. These improvements include clearer image quality with the 3840 x 2160 resolution in 4K (four times that of HD); smoother playback with rates at 50, 60, and even up to 120 frames per second as compared to the 24 in HD; more realistic colors with gradation improving from 8 bits to 10 and 12 bits; color gamut improving by more than 50% versus that of HD, offering more natural colors; and a streaming rate that is 5–10 times higher than HD video.

There are three definitions in 4K: quasi 4K, carrier-grade 4K, and ultra 4K. The different grades have specific requirements on the image frame rate, sampling rate, compression ratio, and network transmission bandwidth. The higher the requirements, the more rich the color experience and more vivid the display. **Carrier-grade 4K is considered the most appropriate choice for commercial deployments.**

Table 1-1 Comparison between the three 4K grades

Item		Quasi 4K	Carrier-grade 4k	Ultra 4K	8K
Resolution		3840 x 2160	3840 x 2160	3840 x 2160	7680 x 4320
Frame rate		25/30P	50/60P	100/120P	120P
Sampling bits		8	10	12	12
Compression		HEVC main profile	HEVC main 10	HEVC Range Extension	HEVRange Extension
Average bit rate	VOD	12–16 Mbps	20–30 Mbps	30–45 Mbps	70–90 Mbps
	BTV	25–30 Mbps	25–35 Mbps	40–55 Mbps	80–100 Mbps
Operation rate	VOD	18–24 Mbps	30–45 Mbps	45–67.5 Mbps	105–135 Mbps
	BTV	25–30 Mbps	32.5–45.5 Mbps	52–71.5 Mbps	104–130 Mbps

NOTE

- VOD services use the variable bit rate (VBR). Considering bit rate variations, the operation bandwidth for smooth playback must be at least 1.5 times the average bit rate.
- Multicast services use the constant bit rate (CBR). The operation rate is considered with FCC deployed. Minimum bandwidth = CBR x 1.3.

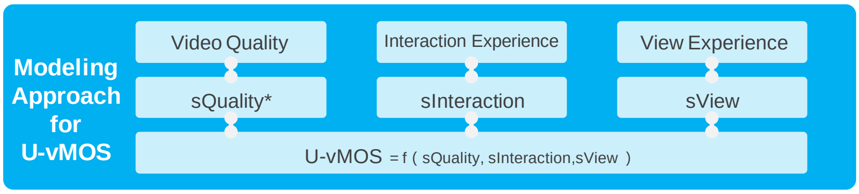
1.2 U-vMOS

U-vMOS (User, Unified, Ubiquitous-Mean Opinion Score for Video) is a video experience measurement system developed by Huawei to evaluate video experience for network optimization. Huawei iLab carried out video-centric human factors engineering tests to track individual reactions to video watching. The information collected using test instruments (such as the eye tracker and physiograph) and the test reports are used to establish a mathematical model and define U-vMOS scoring standards, with an aim to show subjective viewing experience in an objective way.

U-vMOS scores are determined by three factors: video quality (sQuality), interactive experience (sLoading), and viewing experience (sStalling). These factors cover the video resolution, the number of program sources, screen size, operating experience, and video playback fluency. U-vMOS scores range from 1 to 5, where 5 is excellent, 4 good, 3 fair, 2 poor, and 1 bad. A higher U-vMOS score comes from a larger screen size, a higher resolution of program sources, and more fluent video playback.

Figure 1-1 U-vMOS modeling

$$U - vMOS = f\left(sQuality, sInteraction, sView\right)$$



sQuality

sQuality is determined by the following factors: DisplaySize, VideoComplexity, Resolution, BitRate, CodecType, and VideoFrameRate.

Figure 1-2 Factors of sQuality

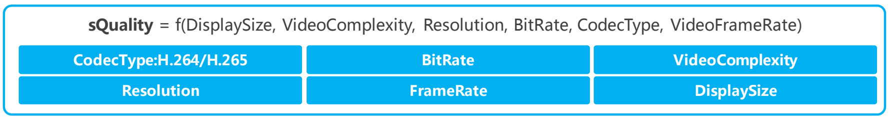


Table 1-2 lists the maximum sQuality values for videos of varying resolutions on typical screens.

Table 1-2 Maximum sQuality values for videos of varying resolutions on typical screens

Resolution	Screen Size						
	4.5 inch	5.5 inch	7 inch	9.7 Inch	42 inch	84 inch	100 inch
8K	5.0	5.0	5.0	5.0	5.0	4.9	4.9
5K	4.96	4.95	4.93	4.91	4.90	4.81	4.78
4K	4.90	4.88	4.86	4.82	4.78	4.66	4.62
2K	4.77	4.73	4.69	4.63	4.53	4.31	4.25
1080P	4.62	4.58	4.52	4.44	4.25	3.96	3.87
720P	4.32	4.26	4.17	4.05	3.69	3.29	3.18
480P	3.89	3.79	3.68	3.52	2.95	2.48	2.37
360P	3.49	3.38	3.25	3.06	2.36	1.91	1.80

sInteraction

sInteraction indicates the zapping time for BTV and the loading time for VOD.

Figure 1-3 Factors of sInteraction

BTV

$sInteraction = f\,(sZapping)$

$sZapping = f\,(zapping\ time)$

VoD

$sInteraction = f\,(sLoading)$

$sLoading = f\,(loding\ Time, DisplaySize)$

Table 1-3 and Table 1-4list the maximum sZapping and sLoading values, respectively.

Table 1-3 Typical sZapping values

sZapping	
Score	**Zapping Time (ms)**
Excellent (5)	<=100
Good (4)	500
Fair (3)	1000
Poor (2)	2000
Bad (1)	>4000

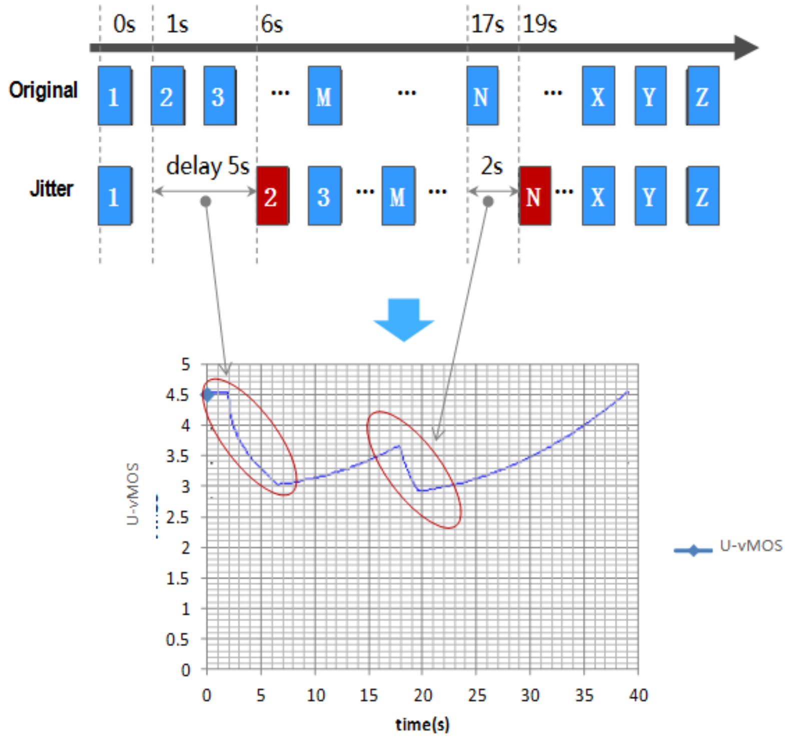
Table 1-4 Typical sLoading values

sLoading		
Score	**sLoading Time@TV**	**sLoading Time@phone**
Excellent (5)	<=100	<=100
Good (4)	1000	1000
Fair (3)	2000	3000
Poor (2)	5000	5000
Bad (1)	8000	10000

sView

Buffers that occur because data packets do not reach the destination in time may lead to video stalls, which have great impacts on video experience. In general, video quality deterioration is closely related to the stall duration and stall interval. For longer video playback, video quality deterioration is closely related to the stall frequency and stall duration. When video playback is restored from a stall, video experience experiences a slow recovery. If the video can normally play afterward, the real-time video quality experience is restored to normal. However, if another stall occurs, the video quality experience will be affected by both the stall duration and stall interval. Figure 1-4 shows real-time changes of U-vMOS scores during a video stall.

Figure 1-4 Real-time changes of U-vMOS scores during a video stall



sView is determined by the following factors: Duration (average time required for a video stall), Interval (average time between two video stalls), and Frequency (number of video stalls).

Figure 1-5 Factors of sView

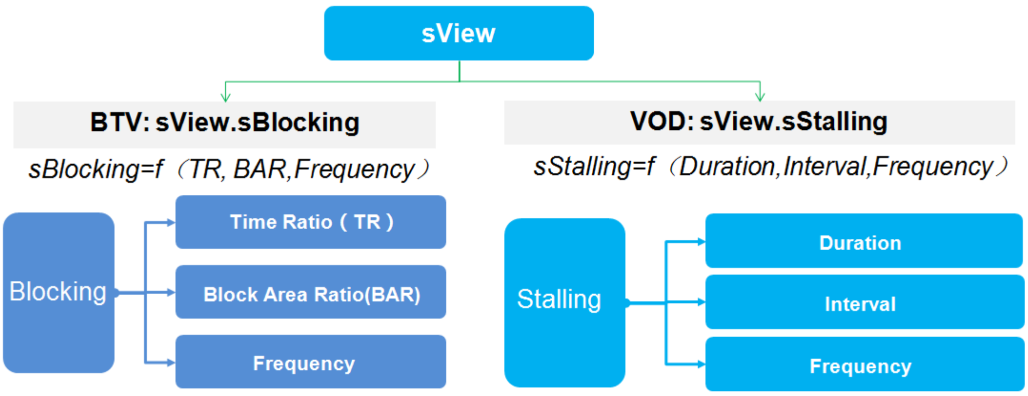


Table 1-5, Table 1-6, and Table 1-7 list the typical sView values of videos on smartphones/Pads and TVs.

Table 1-5 Typical sStalling values of videos on smartphones/Pads in a one-minute period

Typical sStalling Values for Smartphones/Pads				
Score	Stall Frequency	Stall Interval (s)	Stall Duration (s)	Proportion of Stall Duration
5	0	0	0	0%
4	1	0	2.7	5%
3	2	>10	3	10%
2	>2	<5	>5	15%
1	>3	<2	>10	30%

Table 1-6 Typical sStalling values of videos on TVs in a 45-minute period

Typical sStalling Values for TV				
Score	Stall Frequency	Stall Interval (s))	Stall Duration (s)	Proportion of Stall Duration
5	0	0	0	0%
4	1	0	2.7	0. 1%
3	3	>30s	9	1%
2	6	>30s	22.5	5%
1	>10	>30s	>27	10%

Table 1-7 Typical sBlocking values

Typical sBlocking Values			
Score	Proportion of Artifact Duration	Proportion of Artifact Area	Number of Artifacts
5	0%	0%	0
4	4%	35%	1
3	10%	45%	2
2	15%	35%	6
1	50%	95%	12

1.3 Experience-driven 4K Bearer Network

Definition: A 4K bearer network of the best experience is an experience-driven, end-to-end network that provides 4K service experience with a U-vMOS score ≥ 4.0.

- To have a U-vMOS score ≥ 4.0, sInteraction should be ≥ 4 and sView should be 5 (vital to video experience).

1.3.1 Network Requirements for VOD Services with a U-vMOS Score ≥ 4.0

Network Requirements for sQuality

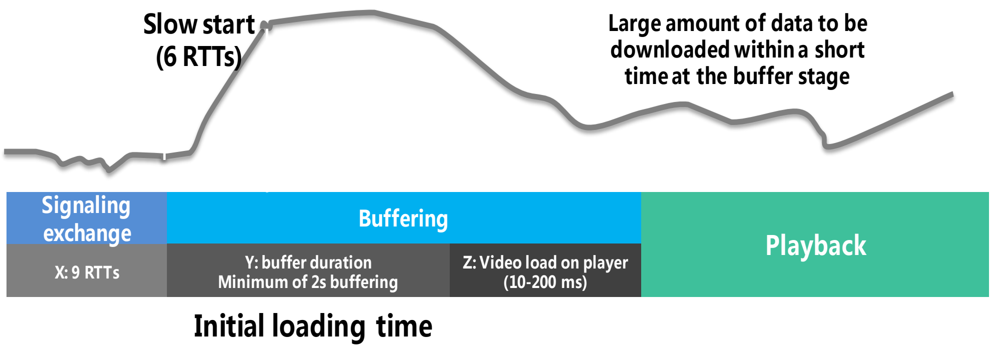
For 4K VOD services, sQuality should be ≥ 4, and the average bit rate should be ≥ 25 Mbps (H.265, VBR).

Network Requirements for sInteraction

sInteraction indicates sLoading for VOD services. To ensure that sInteraction is ≥ 4, the initial loading time should be ≤ 1s.

The most widely used video streaming technology is HTTP Live Streaming (HLS) where initial buffering comes in three stages: signaling exchange (X: ≤ 9 RTTs), download 2s video data before play (Y:slow start; Z: main download), and video load on a player (Z: usually 10 ms to 200 ms). To ensure that the initial loading time is ≤ 1s, X+Y+Z should be ≤ 1s.

Figure 1-6 HLS initial buffering time



Considering that TCP slow start requires six RTTs and the download rate is typically low within this period, the entire download time requires at least seven RTTs. Specifically, the initial buffering time should be 9 RTTs + (6 + 1) RTTs + 200 ms ≤ 1000 ms, with each RTT ≤ 50 ms.

Given that the video load on a player takes 200 ms and the video download for slow start is not counted (T = 1000 ms – 9 RTTs – 6 RTTs – 200 ms = 800 ms – 15 RTTs), a single TCP thread needs to download media data at 50 Mbps (25 Mbps x 2s). Table 1-8 lists the TCP throughput requirements for varying RTTs (≤ 50 ms).

Table 1-8 TCP throughput requirements for varying RTTs

E2E RTT	TCP Throughput Requirement
10 ms	77 Mbps
20 ms	100 Mbps
30 ms	143 Mbps
40 ms	250 Mbps
50 ms	1 Gbps

The E2E delay of 10 ms and the throughput of over 100 Mbps per user are a little bit restrictive requirements. Therefore, the U-vMOS system takes the E2E delay of 20 ms and the throughput of 100 Mbps for a single TCP thread per user as the requirements for sLoading ≥ 4. The following is the calculation formula for TCP throughput:

$$\mathbf{Throughput} \leq \mathbf{min}\left(\mathbf{Max}(\mathbf{BW}), \frac{\mathbf{WSS}}{\mathbf{RTT}}, \frac{\mathbf{MSS}}{\mathbf{RTT}} \times \frac{\mathbf{1}}{\sqrt{\mathbf{p}}}\right)$$

With a throughput of 100 Mbps and an RTT of 20 ms, the required bandwidth is ≥ 100 Mbps and the required PLR is ≤ 3.4×10^{-5}.

NOTE

The HTTP streaming solution is being optimized continuously to improve service experience while at the same time reducing network requirements. For example, a mainstream OTT uses UDP-based QUIC for signaling so that less than five RTTs are used during interaction. At the same time, the bit rate is set to slow for the loading of titles, which reduces network RTT requirements and the amount of data to be downloaded for minimum buffering. In typical scenarios, the OTT videos can have the initial loading time of ≤ 1s in normal network conditions. In the event of fast forward or rewind, the video loading time will exceed 1s.

If terminal + cloud optimization is not taken into account, the network requirements for HTTP VOD services with sInteraction ≥ 4 are as follows:

- E2E bandwidth ≥ 100 Mbps
- RTT ≤ 20 ms
- PLR ≤ 3.4×10^{-5}

Network Requirements for sView

sView indicates sStalling for VOD services.

Video playback fluency is of vital importance to user experience. Therefore, a network with a U-vMOS score ≥ 4 requires sView = 5. That is, no stall occurs during video playback.

Based on Huawei iLab test results, the throughput of a single TCP connection ≥ 1.5 times of the average bit rate can ensure playback fluency for 95% of 4K videos. With sQuality of 4, the average bit rate of 4K@H.265 videos is ≥ 25 Mbps. With sView of 5, the required TCP throughput is > 37.5 Mbps. The following is the calculation formula for TCP throughput:

$$\mathbf{Throughput} \leq \mathbf{min}\left(\mathbf{Max}(\mathbf{BW}), \frac{\mathbf{WSS}}{\mathbf{RTT}}, \frac{\mathbf{MSS}}{\mathbf{RTT}} \times \frac{\mathbf{1}}{\sqrt{\mathbf{p}}}\right)$$

With an RTT of 20 ms, the required bandwidth is ≥ 37.5 Mbps and the required PLR is ≤ 2.4 x 10^{-4}.

In summary, if cloud + terminal optimization is not taken into consideration, the network requirements for VOD services with a U-vMOS score ≥ 4 are as follows:

- E2E bandwidth ≥ 100 Mbps
- RTT ≤ 20 ms
- PLR ≤ 3.4 x 10^{-5}

If cloud + terminal optimization is performed, there is no requirement for the initial loading time, and the network requirements for VOD services with a U-vMOS score ≥ 4 are as follows:

- Bandwidth ≥ 37.5 Mbps
- RTT ≤ 20 ms
- PLR ≤ 2.4 x 10^{-4}

1.3.2 Network Requirements for BTV Services with a U-vMOS Score ≥ 4.0

Network Requirements for sQuality

The mainstream BTV services transmit real-time streams by means of UDP, and Constant Bit Rate (CBR) is usually used to prevent network loss from network traffic bursts. For 4K BTV services, sQuality should be ≥ 4, and the average bit rate should be ≥ 30 Mbps (H.265, CBR).

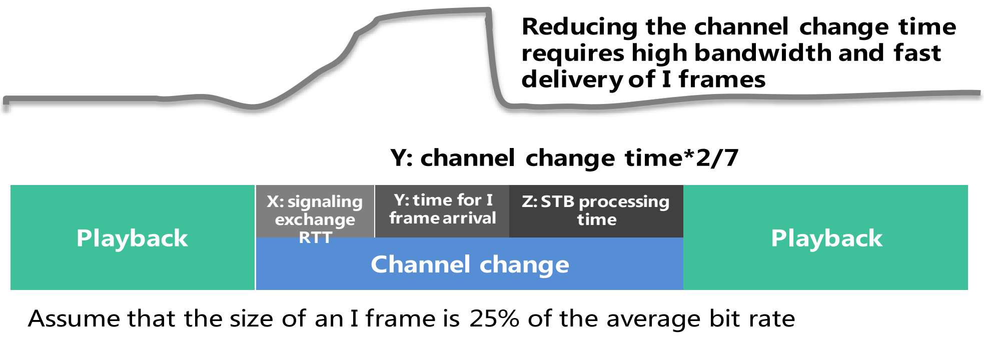
Network Requirements for sInteraction

sInteraction indicates sZapping for BTV services.

To ensure that sZapping is ≥ 4, the channel change time should be ≤ 500 ms.

For mainstream UDP BTV systems, channel change comes in three stages: signaling exchange (X: 1 RTT), download of a complete I frame (Y), and video load on a player (Z: usually 10 ms to 200 ms). To ensure that the channel change time is ≤ 500 ms, X+Y+Z should be ≤ 500 ms.

Figure 1-7 Initial buffering time for UDP BTV



Within a period of T (500 ms – RTT – 200 ms = 300 ms – RTT), an STB needs to load media data at 15 Mbps (30 Mbps*2s*25%). The typical GOP of BTV is 2s, and the size of an I frame is typically ≤ 25% of a GOP. Typically, the UDP BTV system is insensitive to delay, and the E2E RTT for BTV services is not as restrictive as that for VOD services. Taking an example of the typical E2E RTT as 30 ms, the required bandwidth per user is ≥ 56 Mbps.

In addition, the fast channel change (FCC) solution is deployed to speed up channel change. To ensure normal working of the FCC solution, the E2E bandwidth per user should be ≥ 1.3 times of the average bit rate (30 Mbps * 1.3 = 39 Mbps).

In summary, for UDP BTV services with sInteraction ≥ 4, the required E2E bandwidth is ≥ 56 Mbps.

Network Requirements for sView

For BTV services, sView indicates sBlocking.

Viewing experience with no image damage is of vital importance to end user experience. A U-vMOS score ≥ 4 requires sView = 5, that is, no artifacts during video playback. Based on TR-126 requirements, 4K BTV services with no image damage require a PLR of < 10^{-6}. Currently, technologies such as RET and FCC are used at the application layer to reduce video requirements for PLR to around 10^{-4}.

In summary, the network requirements for 4K BTV services with a U-vMOS score ≥ 4.0 are as follows:

- E2E bandwidth ≥ 56 Mbps
- PLR ≤ 10^{-6} with RET not taken into account; PLR ≤ 10^{-4} with RET taken into account

1.3.3 Key Characteristics of Experience-driven 4K Bearer Networks

Network requirements for a U-vMOS score ≥ 4.0

		VOD Service		BTV	
		Without Cloud + Terminal Optimization	With Cloud + Terminal Optimization	Without RET	With RET
Average Bit Rate		≥ 25 Mbps	≥ 25 Mbps	≥ 30 Mbps	≥ 30 Mbps
Network indicators	E2E bandwidth	≥ 100 Mbps	≥ 37.5 Mbps	≥ 56 Mbps	≥ 56 Mbps
	RTT	≤ 20 ms	≤ 20 ms		
	PLR	≤ $3.4 * 10^{-5}$	≤ $2.4 * 10^{-4}$	≤ 10^{-6}	≤ 10^{-4}

In summary, to ensure a U-vMOS of ≥ 4, the network requirements are an E2E bandwidth of ≥ 100Mbps, an RTT of 20 ms, and a PLR of 10^{-5}.

O&M is also an indispensable parts of video experience guarantee. Highly perceivable 4K services impose even higher requirements for service experience monitoring and troubleshooting. From the O&M perspective, the 4K bearer network should be perceivable and manageable.

Therefore the key characteristics of 4K bearer networks of the best experience are as follows:

- For a U-vMOS score ≥ 4.0: E2E bandwidth ≥ 10 Mbps, RTT of 20 ms, and PLR of 10^{-5}
- Maintenance: perceivable experience and fast fault demarcation and location

In some scenarios that have access bandwidth limitations, the 4K bearer network can be designed based on a U-vMOS score 3.5 (sView = 5 and sInteraction ≥ 3), and the key characteristics are as follows:

- For a U-vMOS score ≥ 3.5: E2E bandwidth ≥ 50 Mbps, RTT of 40 ms, and PLR of 10^{-5}
- There are two options to deploy 4K services in the preceding network conditions. One is to reduce the video bit rate. The other is a sacrifice of sInteraction (for example, ≥ 3) to ensure video playback fluency (sView = 5) by not choosing forward/rewind or, alternatively, the use of cloud + terminal optimization to ensure fast start.

2 Experience-driven 4K Bearer Network

2.1 Target Architecture for the Experience-driven 4K Bearer Network

2.1.1 Challenges Faced by Legacy Network Architecture

The network architecture is determined by service and its features.

- Legacy HSI services (e.g. web browsing)
  - Service characteristics: **low concurrency rate, and low network perception,** undemanding delay
  - Network KPI: SLA = "bandwidth"
  - Network architecture characteristics: high convergence, multi-level aggregation
- 4K video
  - Service characteristics: **high bandwidth, high concurrency, large burst, and high network perception,** strict with delay
  - Network KPI: SLA = "bandwidth, delay, and packet loss rate"

As 4K services are rolling out, the following conditions must be met when legacy networks are transforming into experience-driven 4K bearer networks:

- The multi-layer legacy network model with high convergence is simplified to keep pace with the 4K service model.
- End-to-end experience is guaranteed because the 4K network experience is not limited.

The details are as follows:

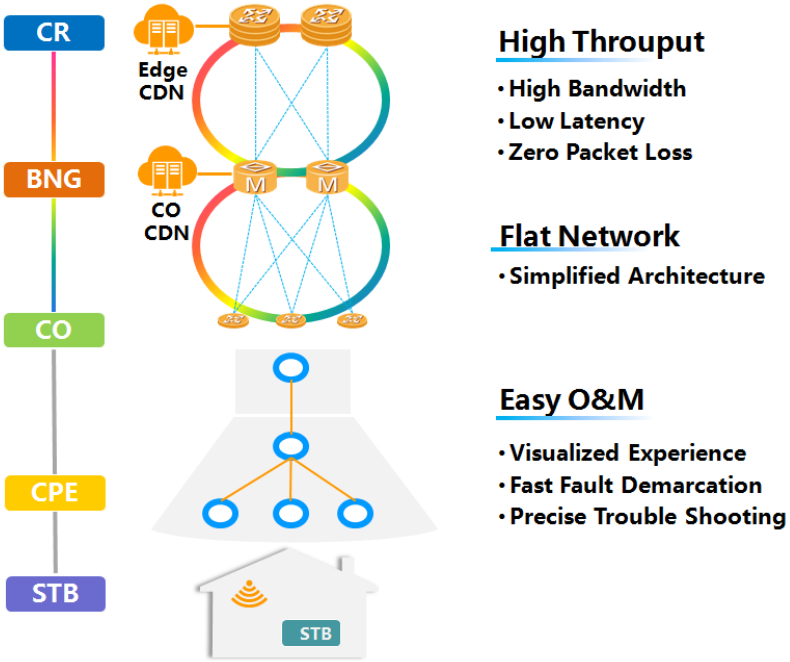
- The home network hits a bottleneck in high bandwidth and seamless coverage, failing to meet multi-screen user experience.
- Low-speed copper access networks, such as DSL networks, also have a bandwidth bottleneck.
- The metro aggregation layers have multi-level aggregation, rendering multiple congestion points, poor scalability, and complex O&M.

- Though video is a basic service, video experience cannot be detected, making it difficult to identify experience problems.

2.1.2 Target Architecture for the Best-Experience 4K Bearer Network

To support scale development of 4K services, architecture transformation is inevitable. The target architecture for the experience-driven 4K bearer network must be a flat network that is able to provide high throughput and is easy for O&M.

Target architecture for the best-experience 4K bearer network



High throughput

- Bandwidth (E2E bandwidth ≥100M/screen)
  - Home network: bandwidth ≥ 100 Mbps; seamless coverage
  - Access network: bandwidth per screen ≥ 100 Mbps
  - OLT to the edge CDN: Plan bandwidth based on different convergence rates given specific user scale, 4K penetration rate, and 4K concurrency rate. Network devices must have the evolution capability for non-convergence.
- Delay (RTT ≤ 20 ms)
  - On CDNs that are mainly deployed on the metro network (with a coverage of less than 200 km), the delay comprises respective delay on the home, access and metro networks. The home and metro network delay has little impact when no traffic congestion occurs. As such, the delay is mostly generated on the access network.
  - On access networks where FTTH or G.fast is deployed, the delay is less than 6 ms, which meets the 4K service requirements.
  - On access networks where VDSL2, vectoring, or super vectoring is deployed, the delay is between 10 ms and 20 ms. The E2E delay may exceed 20 ms after the delay

on the home and metro networks, though very low, is accumulated. On such networks, delay optimization solutions, such as TCP acceleration, should be deployed.

- Packet loss rate ($\leq 10^{-5}$)
  - Packet loss mainly occurs due to poor line quality and network node congestion. Network node congestion must be mainly considered during architecture design.
  - The downlink from the OLT to the user side is non-blocking. Traffic congestion mainly occurs between the OLT to the upstream CDN. As such, QoS, light load, and a low convergence rate of 1:5 to 1:2 should be planned.
  - Because congestion caused by traffic bursts is inevitable, devices must be able to buffer burst traffic to prevent or minimize impact on user experience. Devices, especially Layer 2 network devices, must have sufficient buffer resources.
  - Network and device upgrades: Modernization is needed on access networks, and large-capacity OLTs, routers, and transport devices must be upgraded.

Flat network

Networks must be low convergence-designed for network KPI fulfillment. If the aggregation network still has multiple layers, network construction costs will rise rapidly as the user scale grows, and many nodes may experience congestion. Therefore, the metro network architecture must be simplified to improve network performance and cost-effectiveness.

- CDN&BNG moved downstream to the edge
- Networks simplified from five layers to three layers
- OTN to CO

Easy O&M

- Visualized experience
- Fast fault demarcation
- Precise fault locating

2.2 Stepped Evolution to 4K Bearer Networks

Rome wasn't built in a day. The best-experience 4K bearer network does not need to come into place all at once. Carriers can gradually evolve their networks to best-experience 4K bearer networks in line with the particulars of their business developments. The following outlines the measured steps to delivering 4K video services featuring ultimate user experience.

Initial 4K video service offering: Focus on quickly attracting users and establishing a positive business cycle while ensuring a good basic experience. Launching the 4K service is the crucial first step. The target metrics for this phase include smooth playback with sView reaching the 5.0 standard, sInteraction ≥ 3.0, and loading taking two seconds or less (same as that of cable). These requirements mean that the network must ensure E2E bandwidth ≥ 50 Mbps, RTT ≤ 40 ms, and PLR ≤ 10^{-5}.

Stepped improvement in 4K video service: As users increase and the positive business cycle becomes well established, the next step is improving the service experience. This means sView reaches the 5.0 standard, 4K interactive experience (sInteraction) reaches the 4.0 standard (initial loading time and loading time at each dragging of the progress bar of less than one second), and the network must ensure E2E bandwidth ≥ 100 Mbps and RTT ≤ 20 ms.

At-scale development of 4K video services: Once there are a large number of users, the high bandwidth and concurrency attributes of the 4K video service places much added strain on the network. Expansion of legacy networks would entail high costs and may not solve experience deterioration issues. Once video users reach scale, optimizations in the network architecture must be considered, requiring a network transformation to deliver best-experience 4K bearer networks.

2.3 Home Network with High-Speed and High Performance Wi-Fi

Key Challenges Faced by Home Wi-Fi Networks for 4K Service Bearer

- Service impact: Video, download, and Internet services affect 4K services.
- Wi-Fi coverage in a home: Because the housing structure varies and Wi-Fi signals attenuate after passing through walls, 100% Wi-Fi coverage is impossible.
- Interference from radio signals: The 2.4 GHz frequency band has large radio interference. If other home appliances also use Wi-Fi frequency band, 4K video transmitted using Wi-Fi will also be affected.

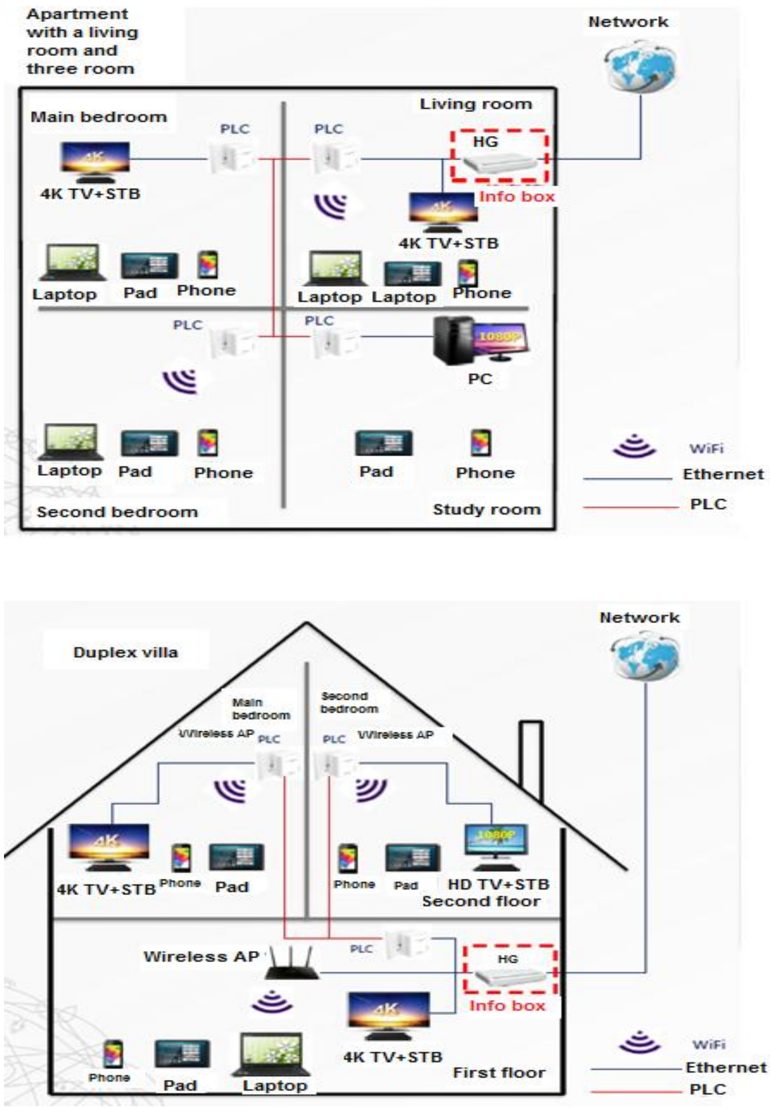
High-Speed and High-Performance Home Wi-Fi Network Solution

- Separate service bearer: Dual-frequency Wi-Fi gateways are used to provide dual SSIDs, allowing 2.4 GHz Wi-Fi to carry Internet/download services and 5 GHz Wi-Fi to carry video services. This decreases interference from 2.4 GHz frequency band on video services carried over 5 GHz frequency band, and accordingly reduces packet loss.
- Distributed Wi-Fi solution: To allow an access rate of over 100 Mbps at any location in a home, the home Wi-Fi network must be deployed in distributed but not centralized mode. Specifically, deploy multiple Wi-Fi access points (APs) in a home for Wi-Fi access at any location in a home. Distributed Wi-Fi can be deployed using the Wi-Fi repeater, PLC Wi-Fi AP, or Ethernet cascading router, allowing flexible scalability, plug and play, and seamless handover.
- The following recommendations are provided to decrease radio signal interference on home networks:
  - In the case of a large floor area where the home gateway cannot provide 4K video directly through network cables or Wi-Fi, APs can be connected to network cables, PLC, or coaxial cables or Wi-Fi modems to expand the home network coverage, allowing 4K video at any location in the home.

- Use G.hn PLC for interconnection between rooms to reduce inter-AP interference. PLC APs cannot be inserted into lightning protection sockets. The power line between master and slave PLC APs must not be greater than 10m.
- Each room has only one AP deployed, but more than two antennas. 5 GHz Wi-Fi is used, with a channel bandwidth of 40 MHz. APs deployed in different rooms have different frequencies. Retain a distance of less than 10m between each AP and the STB and TV.
- Avoid concurrent 4K video playback on more than one TV and one Pad using 5 GHz Wi-Fi. Avoid concurrent playback of videos on three or more terminals
- Do not use home appliances without EMI certificates. Such home appliances will interfere with Wi-Fi.

The following figures show typical 4K deployments on home networks:

Figure 2-1 4K deployment for home networks in two typical apartment layouts



Conclusion: 4K services have the following requirements on home Wi-Fi networks:

- 4K services are separated from other services and do not affect one another, with the home gateway providing dual SSIDs.
- Full Wi-Fi coverage, ensuring 50 Mbps to 100 Mbps bandwidth
- Home appliances and Wi-Fi signals do not interfere. To do so, do not use home appliances without EMI certificates, and use 5 GHz Wi-Fi to carry 4K services.

2.4 Access Network

2.4.1 Best 4K Experience Based on FTTH

Key Challenges of 4K Service Provisioning over an FTTH Network

Challenges of 4K service provisioning over an FTTH network are: whether bandwidths meet requirements, whether the OLT architecture meets the requirement, whether fast program zapping is supported, and how to precisely locate a fault point.

- How to provide 100 Mbps assured bandwidth under different split ratios according to different optical access technologies, including EPON, GPON, and 10G PON? Huawei suggests the following:

1. **EPON access:** Assume that the downstream bandwidth of a single port is 1 Gbps. If the concurrent user online rate is 50% and every program occupies 50 Mbps bandwidth, each user can watch 2 programs at the same time. In this case, every EPON port supports a maximum of 21 users. Therefore, the recommended split ratio is 1:32. In the upstream direction, 50% traffic is calculated for port capability. If each EPON port houses 16 ports, 8 Gbps bandwidth is required in the upstream direction for a board. In actual service provisioning, the upstream traffic is monitored for you to determine whether network expansion is needed. If the upstream traffic exceeds 60% of the port capability, more upstream ports are deployed. Assume that the OLT supports 16 service slots. 14 service boards are recommended, which support 7168 (14 * 16* 32) EPON ports. Another 2 slots are reserved for upstream interface boards to provide up to 112 Gbps upstream transmission capabilities.
2. **GPON access:** Assume that the downstream bandwidth of a single port is 2.5 Gbps. According to the preceding assumption, each GPON port supports a maximum of 51 users and split ratio 1:64 can be configured. In the upstream direction, 20 Gbps bandwidth is required for 50% convergence rate. Assume that the GPON OLT supports 16 slots. 14 service boards are recommended, which support 14336 (14*16*64) GPON users. Another 2 slots are reserved for upstream interface boards to provide up to 280 Gbps upstream transmission capability. Some countries or areas have high standards on network device security. This requires disaster recovery if users exceed the preset quantity. For example, it is better to deploy 2 devices in different places for more than 10,000 users. Previously, the voice service requires such a disaster recovery design. Considering ODN sharing and long-term development, split ratio 1:32 is recommended for a GPON network. In this case, a single-subrack OLT supports up to 7168 (14*16*32) users and the reserved upstream bandwidth is 280 Gbps.

3. **10G PON+LAN for best 4K experience:** 10G PON is adopted in the upstream direction. A LAN port supports 100 Mbps, meeting 4K video requirements. If every program occupies 50 Mbps bandwidth, every 10G PON port supports a maximum of 8 LAN devices. Therefore, split ratio 1:8 is recommended. Every 10G PON port supports a maximum of 256 concurrent online users.

- **OLT architecture suggestion:** Traditional OLTs use the centralized architecture. In big video era, the chip of a control board becomes the bottleneck that affects the performance of the entire system. Therefore, OLTs adopting the distributed architecture are recommended.
- **Distributed buffering for program zapping:** Traditional centralized OLTs have their buffers on the control boards. Such a design limits the device capacity and can only be used for scenarios having low traffic burst rate. Distributed OLTs have their buffers on service boards. In this case, the general buffering capabilities of both the upstream and downstream ports are improved together with the service board expansion. The distributed OLT features over 10 times capabilities compared with the centralized OLT and addresses 4K demands on high traffic burst rate, low delay, and low packet loss rate.

Requirements on FTTH by the 4K Service

- **EPON:** Each port supports up to 21 4K users. The maximum split ratio can be 1:32. In the upstream direction, if the traffic convergence rate is 50%, the board housing 16 ports needs 8 Gbps bandwidth.
- **GPON**: Each port supports up to 51 users. The split ratio can be 1:64. In the upstream direction, if the traffic convergence rate is 50%, the board housing 16 ports needs 20 Gbps bandwidth.
- **10GPON**: Every 10G PON port supports up to 8 LAN devices. Split ratio 1:8 is recommended. Every 10G PON port supports a maximum of 256 concurrent online users.
- **OLT**: The distributed OLT is recommended, which supports buffering of its service boards.

2.4.2 Best 4K Experience Based on the Superfast Copper Solution

Using copper technologies, bandwidth is the key factor that affects the 4K service. Factors that affect the bandwidth are: adopted copper technology, copper line distance, and copper line quality (involves in packet loss rate and delay).

Table 2-1 Typical network KPIs for different access technologies

Access Technology	Bandwidth (DS)	RTT (ms)	PLR
VDSL2	50M@ < 1000 m	10–20	$10^{-4\sim-5}$
Vectoring	50–120M@ < 800 m	10–20	$10^{-4\sim-5}$
SuperVector	100–300M@300–500 m	10–20	$10^{-4\sim-5}$
G.Fast	200M–1.2G@100–500 m	2–6	$10^{-4\sim-5}$

- ADSL2+ supports access bandwidth of 24 Mbps at most. The 4K service is not recommended to be provisioned under ADSL2+.
- VDSL2 supports 100 Mbps downstream bandwidth. Vectoring that improves the line rate through far-end crosstalk (FEXT) cancellation supports high access bandwidth of 120 Mbps within 300 m using 0.5 mm-diameter copper line. It supports 100 Mbps access bandwidth at typical access distance of 500 m.
- SuperVector expands the working frequency from 17 MHz to 35 MHz. Using vectoring for crosstalk cancellation, SuperVector speeds up the bandwidth to 300 Mbps at the distance of 300 and 100 Mbps 700 m. Therefore, SuperVector is recommended for 300–700 m.
- G.fast brings copper lines into Gigaband era, with which, 400 Mbps bandwidth can be reached at the distance of 300 m. Therefore, G.fast is recommended for 0–300 m.

Best 4K Experience under Fulfilled the requirements for Packet Loss Rate and Bidirectional Delay

In the lab conditions, the packet loss rate can be 10^{-5} to 10^{-7}. In a live network, the E2E packet loss rate can be 10^{-3} to 10^{-5}. DSL signals are transmitted over lines at the rate of 64% of the speed of light. Therefore, the line delay value is very small. For example, VDSL2/vectoring bidirectional delay is 10–20 ms, SuperVector 10–30 ms, and G.fast 2–6 ms. In actual 4K video deployment, after retransmission and block interleaving are used as recommended, the bidirectional delay will be dramatically decreased to 8–12 ms.

Best 4K Experience Based on Deeper FTTC/FTTB Sites

Assume that an FTTC site connects to 384 users, VDSL2 is used, only one program is provisioned, and traffic convergence rate of the upstream port is 50%. About 7.3 Gbps (384* 38M/2) is recommended to be reserved. G.fast is usually used for FTTB because it takes effect at a short distance for fewer users. Therefore, a G.fast device supports up to 96 users. In addition, G.fast features higher bandwidth. Therefore, its upstream port needs to ensure that every user can be provisioned with the lowest bandwidth of 100 Mbps. This says it can fully relieve bandwidth pressures brought by 4K video.

Copper lines are greatly affected by distance and crosstalk. Therefore, the copper line diagnosis system is recommended. Using this system, you can pre-evaluate the line performance before service provisioning to determine the service provisioning rate and then the packages suitable for users.

Conclusion: Requirements on a Copper Network

- The copper line diagnosis system is installed. Using this system, you can pre-evaluate the line performance before service provisioning to determine the service provisioning rate and then the packages suitable for users.
- VDSL2, SuperVector, and G.fast are used to support at least 50 Mbps bandwidth.
- Retransmission and block interleaving are used to reduce delay to 8–12 ms.
- FTTC/FTTB sites can be moved closer to users when copper line bandwidth is insufficient.

- TCP can be used if the physical bandwidth provided through copper lines can be reached but throughput is insufficient. For details, see section 2.5.4 "TCP Acceleration for Network Throughput Improvement."

2.4.3 Best 4K Experience Based on D-CCAP

Traditional telecommunications carriers launch video and 4K services, bringing great pressures to MSOs. Furthermore, traditional broadcasting TV is now challenged, that is, it occupies a great amount of spectrums and new IP-based services (such as place shifting) cannot be supported. Therefore, many cable carriers turn to IPTV.

Distributed D-CCAP Based on DOCSIS 3.0

DOCSIS 3.0 supports 32 channels for 16 Gbps downstream bandwidth at most. Assume that every program occupies 50 Mbps bandwidth and 50% service provisioning rate is used. Every D-CCAP supports a maximum of 66 4K users (recommended). In nowadays, a majority of cable carriers deploying DOCSIS 3.0 networks use such networks for Internet access only, and TV programs are transmitted in the traditional broadcast way. A large amount of DOCSIS 3.0 spectrums are reserved for broadcasting videos. Therefore, you are not advised to provision lots of 4K videos over DOCSIS 3.0 networks. However, you can use DOCSIS 3.0 networks for few programs.

Distributed D-CCAP Based on DOCSIS 3.1

DOCSIS 3.1 now supports 4 Gbps downstream bandwidth. After spectrum expansion in the future, it supports 10 Gbps downstream bandwidth, which is applicable to IPTV and 4K video. Assume that every program occupies 50 Mbps bandwidth and 50% service provisioning rate is used. Every D-CCAP supports 164–410 4K users, which is consistent with the number of users supported by the remote fiber node.

Now, it is discussing to use G.fast over coaxial cables. Therefore, in the future, copper lines and coaxial cables are only media resources for broadband access. All spectrums will be used for broadband improvement and traditional broadcasting video will be replaced with the IP-based multicast and VOD.

Conclusion: Requirements on a Coaxial Network

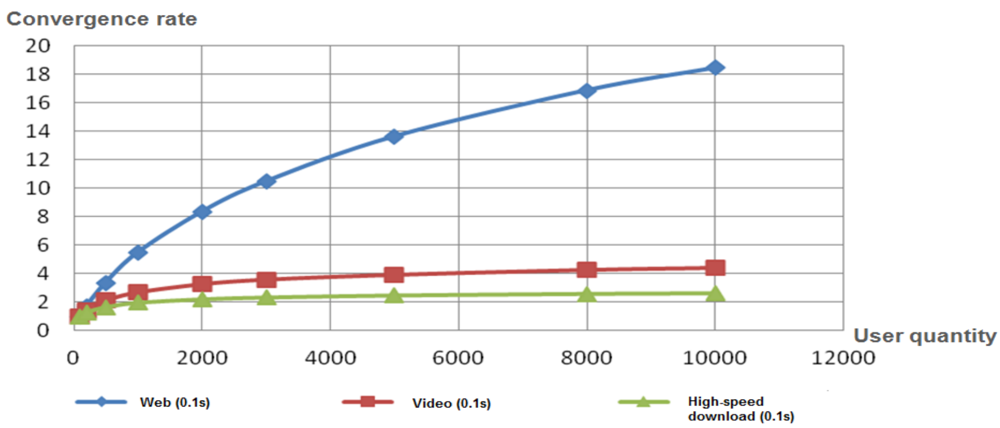
- You are not advised to use DOCSIS 3.0 for a large quantity of 4K videos.
- Every distributed D-CCAP based on DOCSIS 3.1 supports 164–410 4K users.

2.5 Metro/Backbone Network

As the 4K video service is developed, traffic on metro networks will mushroom. Industry experience shows within two to three years of 4K service launch, traffic increases three to seven times and that number increases to 10 times or more after five years.

Launch of 4K video services creates a surge in traffic and reduces the convergence rate. Compared to web browsing and other home broadband services, the waveform fluctuations in video traffic are small as is the convergence rate under the same concurrency rate.

Figure 2-2 Convergence rates of different services



Faced by the traffic surge and change in the convergence model, the legacy network that has high aggregation and convergence rates present the following characteristics when bearing 4K video:

- **Low network efficiency**

  When video traffic is increasing, capacity expansion is needed for E2E network devices. The more the aggregation layers, the lower the convergence rate, the more E2E devices need capacity expansion. If CDNs are deployed upstream of the network where contents are far away from end users, service traffic has to traverse multiple network devices before reaching the end users, consuming a large number of network resources.

- **Poor user experience**

  In the case of multiple concurrent services, a higher network usage will cause a higher packet loss rate and delay, which accordingly deteriorate video service experience. On light-loaded networks, 98.7% of burst packet loss occurs on aggregation nodes where high bandwidth consumption transitions to low bandwidth consumption. The increase of packet loss also deteriorates video service experience.

Figure 2-3 Relationship between the delay and link utilization

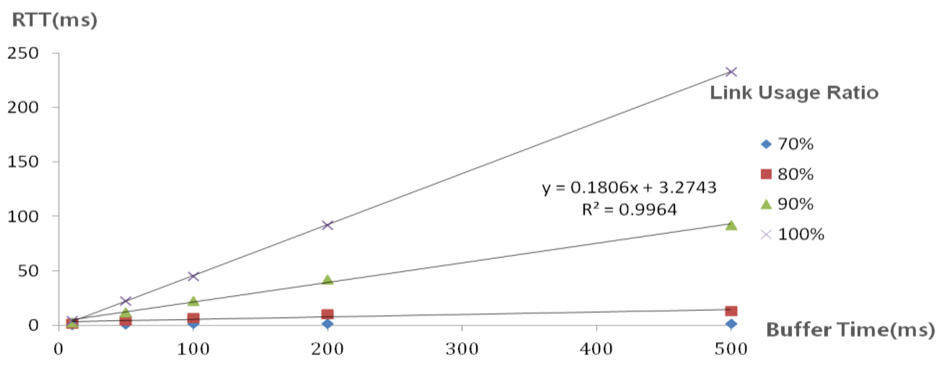
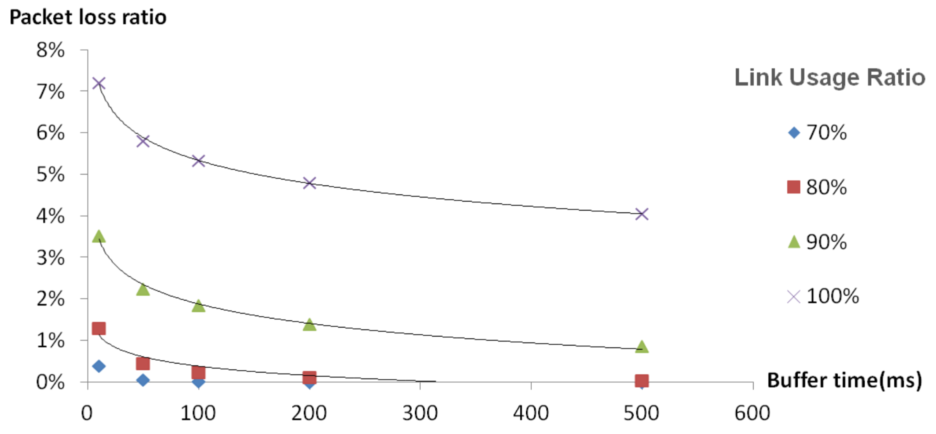


Figure 2-4 Relationship between the packet loss rate and link utilization



Network delayering and simplification and proper CDN deployment will remarkably improve the network efficiency.

- Simplified IP network layers
- OTN to CO
- Distributed CDN deployment aligning with the service scale

In addition, service experience can be remarkably improved if video services are more tolerable of network delay and packet loss.

2.5.1 Simplified IP Network Layers

As video and 4K services are increasing intensively deployed, the practice of deploying BRASs at a higher layer of the network has become the key obstacle to user experience improvement and deployment cost reduction.

- **Video experience**

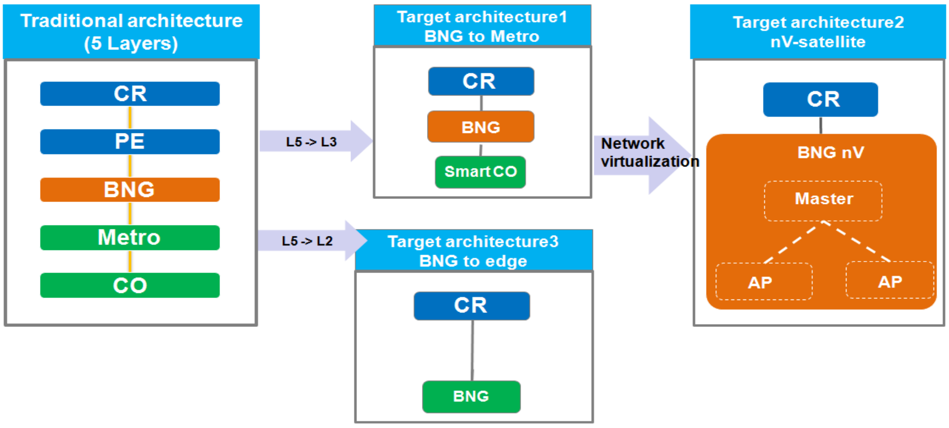
  Video services are delay-sensitive. Transmission delay is natural for packets forwarded hop by hop on a network. When the network is free of congestion, the transmission delay is less than 50 us. Once the network is congested, even high-priority services will experience a millisecond-level or even second-level delay on a single device. Therefore, deploying the CDN at a lower layer of the network helps reduce the number of hops between video sources and terminals, lower the congestion probability, and minimize the E2E delay. When the CDN moves down, BNGs functioning as user gateways also need to do so. Otherwise, deeper CDN will be meaningless.

- **Deployment costs**

  In scenarios where BNGs are deployed at a higher layer of the network, multi-level aggregation exists between BNGs and the access network. This deployment scheme is quite cost-effective in an era dominated by Internet services, because a web page visit or file download is usually completed within a short time due to the statistical multiplexing characteristic of Internet services. Network resources are often used by different users at different times. Traffic aggregates hop by hop from users to the access network, and then to IP gateways and CRs. Therefore, traditional network design places one or multiple layers of aggregation devices between the access network and BNGs to reduce BNG bandwidth requirements (BNG bandwidth costs are higher than switch bandwidth costs). In the 4K era, video traffic is dominant. Video service users are often stably online for a long time, rendering BNGs' bandwidth statistical multiplexing characteristic less useful. Meanwhile, the value of large Layer 2 aggregation decreases. In extreme cases, large Layer 2 aggregation may even drive up network costs. In addition, many sites on the live network use PPPoE to bear BTV services, and BNGs serve as the last-hop multicast replication points for PPPoE users. Higher BNG location means higher E2E bandwidth requirements and network construction costs.

The following describes some flat network architectures.

Figure 2-5 Flat network architectures



Target architecture 1: BNG to metro (network delayed from five to three layers)

After moving downstream to the Metro network, BNGs will be integrated with NPEs on the Layer 2 metro network. After BNGs are further integrated with upstream PEs, the network is delayered from five to three layers.

This architecture requires BNGs to terminate L2VPN and user services on the metro network and forward user service traffic to the upstream network-side L3VPN or Internet based on service features.

An estimated metro network model for target architecture 1 is as follows:

- OLT upstream-downstream convergence rate: 1:2
- CO upstream-downstream convergence rate: 1:2
- BNG upstream-downstream convergence rate: 1:2
- CR upstream-downstream convergence rate: 1:2
- 4000 users per OLT
- 4 OLTs per CO
- 10 COs per BNG
- 10 BNGs per CR

Table 2-2 Traffic model for target architecture 1

Item		Quasi 4K	Carrier-grade 4k	Ultra 4K	8K
User bandwidth requirement (Mbps)		30	50	75	100
Penetration rate		30%	45%	60%	75%
Concurrency rate		20%	30%	40%	55%
Bandwidth requirement	CR upstream (Gbps)	180	675	1800	4125
	CR downstream (Gbps)	360	1350	3600	8250
	BNG upstream (Gbps)	36	135	360	825
	BNG downstream (Gbps)	72	270	720	1650
	CO upstream (Gbps)	7.2	27	72	165
	CO downstream (Gbps)	14.4	54	144	330
	CR (Gbps)	540	2025	5400	12375
	BNG (Gbps)	108	405	1080	2475
	CO (Gbps)	21.6	81	216	495

The bandwidth requirements for each device in the traffic model are as follows:

- CO (device level): 100 Gbps at least, 200 Gbps recommended, 400 Gbps preferred, and capable of smoothly evolving to 2 Tbps
- BNG (board level): 100 Gbps at least, 200 Gbps recommended, 400 Gbps preferred, and capable of smoothly evolving to 2 Tbps
- CR (board level): 400 Gbps at least, 1 Tbps recommended, 2 Tbps preferred, and capable of smoothly evolving to 8 Tbps

BNGs must be capable of:

- Providing FMC integrated access capabilities
- Bearing leased line services, home broadband services, and video services
- Terminating L2VPN and providing access to L3VPN/Internet/new L2VPN

Target architecture 2: Based on target architecture 1, virtual access is used to map CO devices into a BNG's remote boards, which are not independently presented.

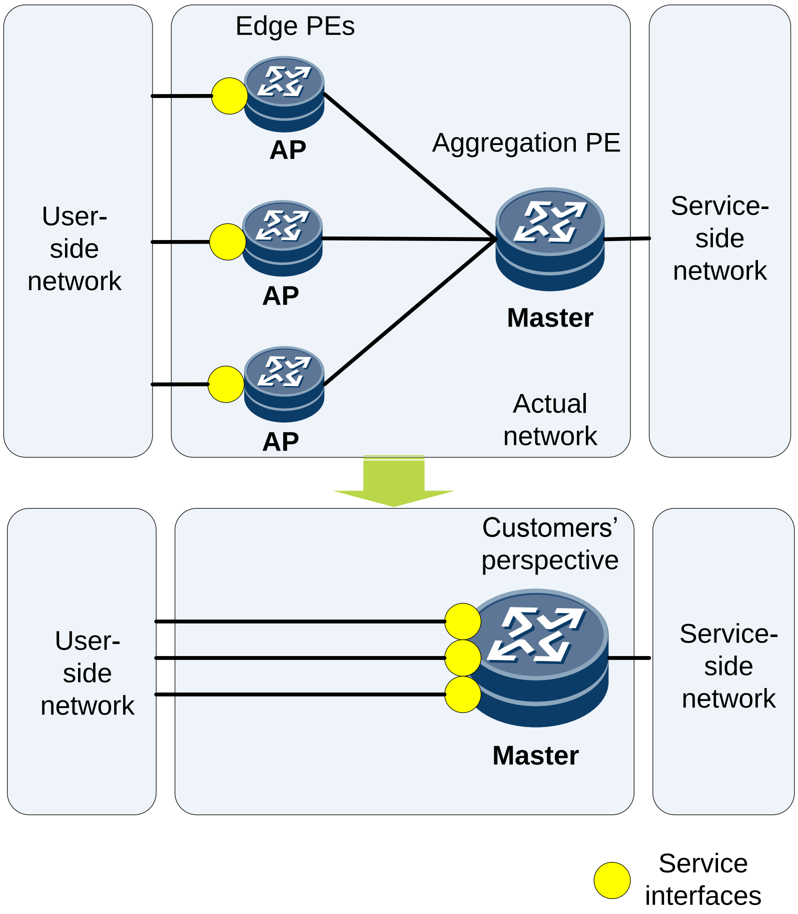
After a BNG is moved downstream to the metro edge, a large Layer 2 network still exists and there are still many aggregation NEs. Carriers must manage these NEs, which results in heavy maintenance workloads.

Virtual access can further simplify the network architecture so that CO devices are managed as a BNG's remote boards. The CO devices are not independently visible to a network management system (NMS), which reduces the number of devices managed by carriers and delayers the network to two layers logically.

Figure 2-6 shows the virtual access solution. After this solution is deployed, all CO devices can be considered as a BNG's remote boards, whereas the CO devices' service interfaces to the user-side network can be considered as the BNG's interfaces. Virtual access enables customers to configure and manage services only on the BNG, which significantly simplifies service deployment and network O&M. In the virtual access solution, the CO devices are called access points (APs), and the BNG is called a master.

- AP: an access node in the virtual access system. An AP can be considered as a master's board and is automatically discovered and managed by a master. An AP receives configuration and forwarding entries delivered by a master and provides external communication interfaces.
- Master: a control node in the virtual access system. A master establishes a tunnel, delivers a service, controls traffic, and manages connected APs.

Figure 2-6 Virtual access solution



Virtual access offers the following benefits:

- Simplified network planning

After virtual access is used, a master automatically calculates forwarding paths and delivers entries in the virtual access system. IS-IS automatically collects topology information in the virtual access system, and no IP address needs to be configured on the interfaces between a master and AP. Therefore, customers do not need to plan IP addresses.

- Simplified service deployment

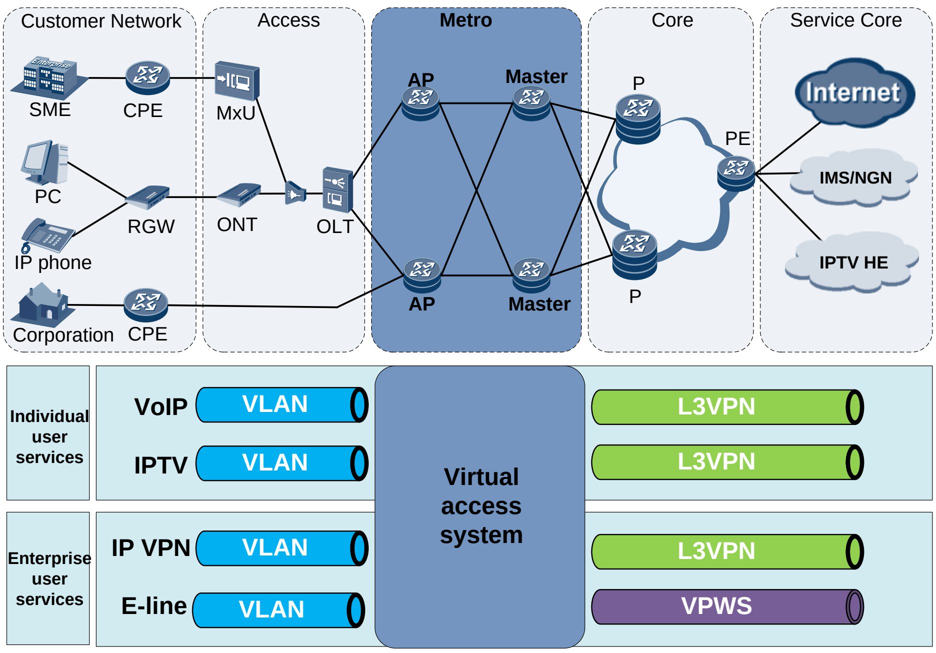
The control and management planes of the virtual access system are centralized on a master, and services are configured and queried on a master or NMS. Therefore, customers do not need to configure and manage services on APs. In addition, dynamic protocols (such as BGP, LDP, and RSVP-TE) do not need to run in the virtual access system, which significantly reduces configuration workloads.

- Simplified O&M and management

Virtual access supports plug-and-play (PnP) for APs. After APs go online, a master automatically discovers and manages them, which simplifies network O&M and management. In addition, a master centrally reports service alarms in the virtual access system to an NMS and APs report only their own fault alarms, which reduces the number of alarms and facilitates fault locating.

The virtual access solution can be used for fixed broadband (FBB) services, including 4K video services and metro services (such as VoIP, IP VPN, E-Line, and Internet services). Figure 2-7 shows a comprehensive bearer solution.

Figure 2-7 Typical application of virtual access



Requirements for key devices' capabilities in virtual access scenarios:

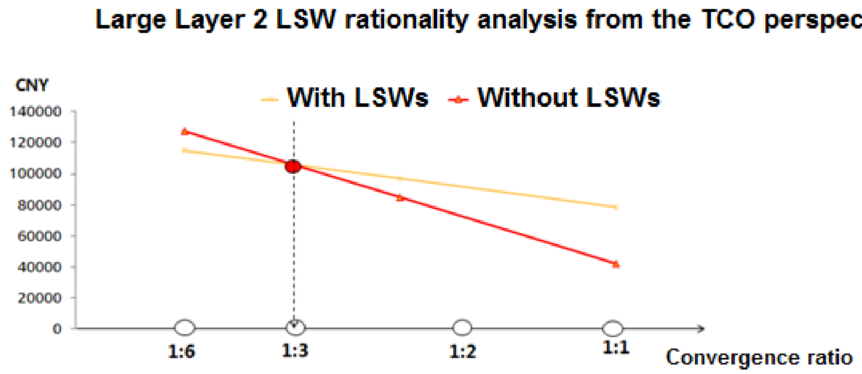
In addition to requirements for BNG/CR/CO devices' capabilities in target architecture 1, BNG and CO devices provide virtual access capabilities and are externally presented as a unique logical NE.

Target architecture 3: BNG to edge (network delayered from five to two layers)

Compared with target architecture 1, this architecture further flattens the network by directly connecting OLTs to BNGs without intermediate convergence or aggregation.

As aforementioned, in the Internet era, adding more switches helps enhance network convergence, reducing the number of required router ports. Because router costs are far higher than switch costs, this implementation reduces TCO. However, as video services are increasingly intensively deployed, the bandwidth convergence rate is becoming lower and lower. Switch aggregation can do little to conserve router ports now. According to a survey on carrier C's TCO in province Y, when the convergence rate between OLT upstream bandwidth and BRAS upstream bandwidth equals 1:3, adding a layer of switches between OLTs and BRASs equals directly connecting OLTs to BRASs in TCO. As the convergence rate decreases further, directly connecting OLTs to BRASs turns out to be more cost-effective. In addition, the upstream physical interface bandwidth of a mainstream OLT today can reach 10 Gbps, well matching the downstream interface bandwidth of a BNG. Therefore, there is no need to add aggregation devices for bandwidth adaptation.

Figure 2-8 TCO analysis



BNGs can move downstream in two modes. One is to leave BNGs at POPs but remove the aggregation layer between OLTs and BNGs.

An estimated metro network model for this architecture is as follows:

- OLT upstream-downstream convergence rate: 1:2
- BNG upstream-downstream convergence rate: 1:2
- CR upstream-downstream convergence rate: 1:2
- 4000 users per OLT
- 20 OLTs per BNG
- 10 BNGs per CR

Table 2-3 Traffic model 1 for target architecture 3

Item		Quasi 4K	Carrier-grade 4k	Ultra 4K	8K
User bandwidth requirement (Mbps)		30	50	75	100
Penetration rate		30%	45%	60%	75%
Concurrency rate		20%	30%	40%	55%
Bandwidth requirement	CR upstream (Gbps)	180	675	1800	4125
	CR downstream (Gbps)	360	1350	3600	8250
	BNG upstream (Gbps)	36	135	360	825
	BNG downstream (Gbps)	72	270	720	1650
	OLT upstream (Gbps)	3.6	13.5	36	82.5
	OLT downstream (Gbps)	7.2	27	72	165
	CR (Gbps)	540	2025	5400	12375
	BNG (Gbps)	108	405	1080	2475
	OLT (Gbps)	10.8	40.5	108	247.5

The bandwidth requirements for each device in the traffic model are as follows:

- BNG (board level): 100 Gbps at least, 200 Gbps recommended, 400 Gbps preferred, and capable of smoothly evolving to 2 Tbps
- CR (board level): 400 Gbps at least, 1 Tbps recommended, 2 Tbps preferred, and capable of smoothly evolving to 8 Tbps

BNGs must be capable of:

- Providing FMC integrated access capabilities
- Bearing leased line services, home broadband services, and video services
- Terminating L2VPN and providing access to L3VPN/Internet/new L2VPN

The other mode is to move BNGs from POPs downstream to CO sites by deploying BNGs and access devices in the same equipment room. After being moved downstream, BNGs will face numerous diversified access devices and access interfaces.

An estimated metro network model for this architecture is as follows:

- OLT upstream-downstream convergence rate: 1:2
- BNG upstream-downstream convergence rate: 1:2
- CR upstream-downstream convergence rate: 1:2
- 4000 users per OLT
- 5 OLTs per BNG
- 40 BNGs per CR

Table 2-4 Traffic model 2 for target architecture 3

Item		Quasi 4K	Carrier-grade 4k	Ultra 4K	8K
User bandwidth requirement (Mbps)		30	50	75	100
Penetration rate		30%	45%	60%	75%
Concurrency rate		20%	30%	40%	55%
Bandwidth requirement	CR upstream (Gbps)	180	675	1800	4125
	CR downstream (Gbps)	360	1350	3600	8250
	BNG upstream (Gbps)	9	33.75	90	206.25
	BNG downstream (Gbps)	18	67.5	180	412.5
	OLT upstream (Gbps)	3.6	13.5	36	82.5
	OLT downstream (Gbps)	7.2	27	72	165
	CR (Gbps)	540	2025	5400	12375
	BNG (Gbps)	27	101.25	270	618.75
	OLT (Gbps)	10.8	40.5	108	247.5

The bandwidth requirements for each device in the traffic model are as follows:

- BNG (device level): 100 Gbps at least, 200 Gbps recommended, 400 Gbps preferred, and capable of smoothly evolving to 2 Tbps
- CR (board level): 400 Gbps at least, 1 Tbps recommended, 2 Tbps preferred, and capable of smoothly evolving to 8 Tbps

NOTE

The BNG must be 300 mm deep and can reside in the same cabinet with the OLT.

- After moving downstream to CO sites, BNGs must provide the any access capability (E1/ATM/POS/GE/10GE/40GE).

2.5.2 OTN to CO

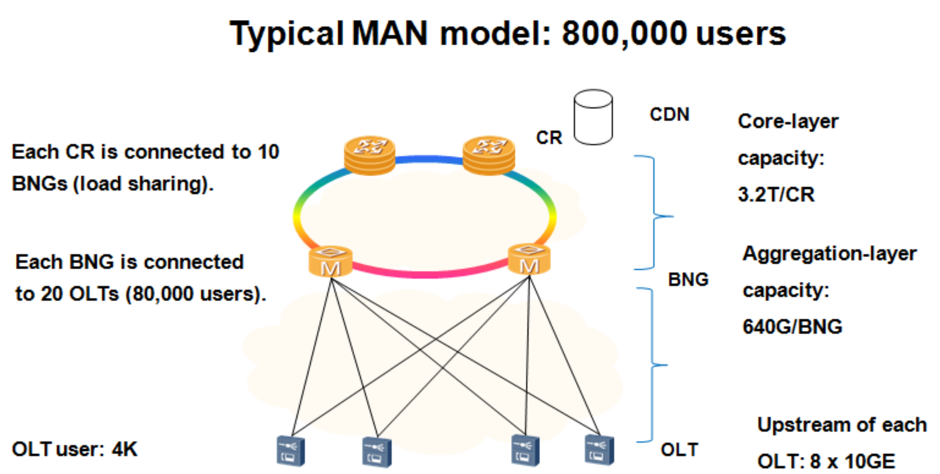
Challenges to Traditional MANs

On a traditional MAN, Ethernet or MPLS switches are used for convergence and optical fibers are used for transmission between BNGs and CRs and between OLTs and BNGs. Such a MAN faces the following challenges during the fast development of 4K video services:

- Challenge 1: Limited trunk fiber resources cannot satisfy increasing MAN bandwidth requirements of 4K video services.

Take the Carrier-grade 4K video service in Table 2-4 as an example. Figure 2-9 shows a typical model of a MAN with 800,000 users. The capacities at the core layer and the aggregation layers can reach 640G and 3.2T respectively.

Figure 2-9 Typical MAN model



On a traditional network with direct fiber connections, each OLT uses four 10GE ports to connect to each BNG. Therefore, 80 pairs of fibers (20* 4) are required between each BNG and OLT, and a total of 800 (10 * 80) pairs of fibers are required.

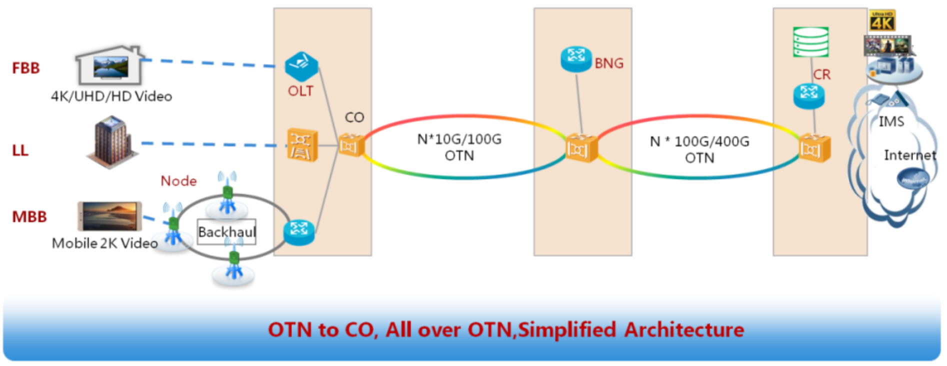
- Challenge 2: Time-consuming fiber routing, connection, and adjustment cannot satisfy demands for fast bandwidth allocation to 4K video services.

  Although the traffic topology of a MAN is in a tree shape, a ring topology with optical cables is usually used between lower-layer nodes and upper-layer nodes because MANs between OLTs and BNGs and between BNGs and CRs are constructed hierarchically along roads. Multiple regeneration sites for fiber jumping are required between COs and BNGs. If the direct fiber connection mode is used, fiber resources need to be coordinated during each service adjustment or capacity expansion for rapidly growing video services, which is time-consuming. Even though existing optical fibers are available and only manual fiber adjustment is required, it still takes several weeks to manually adjust the fiber connections. If new fibers need to be routed, it may even take several months. Therefore, the traditional fiber capacity expansion mode cannot adapt to fast growing video traffic.

OTN to CO Satisfying Requirements of Heavy Traffic and Numerous Connections on 4K Bearer Networks

Video service development drives network flattening (the number of network layers is reduced from 5 to 3). As a result, the distances between OLTs and BNGs and between BNGs and CRs are extended, and interconnection bandwidth and connection quantity grow explosively. The OTN to CO solution can provide basic interconnection pipes that provide ultra-large single-fiber bandwidth, adapt to long distances, have no traffic convergence, and support fast bandwidth on demand (BoD). Figure 2-10 shows the typical networking.

Figure 2-10 Typical OTN to CO networking



Some carriers in the industry, such as China Telecom and China Mobile, have deployed OTN devices to CO sites to carry 4K video services.

Values of OTN to CO

- **Transmission of n*Tbps data over one pair of fibers:** Each aggregation ring requires only one pair of fibers, greatly saving optical fibers.
- **Plug and Play (PnP) bandwidth pool**: OTN devices provide high-quality hard pipes for OLTs, BNGs, CRs, and CDNs. The directions and bandwidth can be flexibly adjusted. ROADM based on wavelengths and the large-capacity (at the Tbps level) hybrid grooming technology can achieve on-demand bandwidth allocation to video services in each direction. Users only need to add service port connections before provisioning services. The PnP and remote configuration features enable the users to groom transmission network services and expand bandwidth within minutes.
- **L0/L1 hard pipe (wavelength/ODUk sub-wavelength):** OTN networking ensures zero packet loss in the case of full load and a single-node latency of less than 50 μs because service pass-through is based on L0/L1 hard pipes, and supports point-to-multipoint (P2MP) connections, satisfying the requirements of broadcast video services.

Function Requirements of OTN to CO

- The OTN devices that carry OLT/IP services at different network layers need to meet requirements. (Referring to Table 2-4, at least 50% margin of the OTN capacity need to be reserved.)
  - OTN device at CO sites: system capacity at large CO sites in urban areas > 3 Tbps, system capacity at medium CO sites in counties > 600 Gbps, and system capacity at small CO sites in towns > 100 Gbps
  - OTN device at BNG sites: a single-slot capacity of 400 Gbps and system capacity of more than 6 Tbps
  - OTN device at CR sites: a single-slot capacity of 400 Gbps and system capacity of more than 12 Tbps

- Metro 100G WDM devices support pluggable CFP optical modules whose power consumption is less than 50% of the power consumption of 100G MSA modules.
- Photonics integrated device (PID) needs to be supported. The silicon photonics technology is used to integrate separate traditional WDM optical-layer boards, such as OA, MUX, and DCM boards, into one board and implement automatic optical-layer configuration, simplifying O&M.
- Programmable 100G/200G/400G OTN interfaces (FlexRate) are based on the same hardware, and software configurations can achieve fast bandwidth acceleration, meeting different bandwidth requirements.
- The Optical Doctor/Fiber Doctor (OD/FD) for monitoring online optical performance and fiber quality can work with U-vMOS to locate physical-layer faults.

2.5.3 Distributed CDN Aligning with the Service Scale

Well Planned Distributed CDN Helps Improve Service Experience and Reduce Network Construction Costs

The rapid development of video services also raises requirements for storage and computation resources on the network edge. To ensure video experience, multiple levels of CDNs (including core, edge, and regional CDNs) have to be constructed. The closer the CDN is to a user, the better fuser experience.

The CDN location determines the direction, delay, and packet loss rate of video traffic and factors into 4K service provisioning costs. Trading storage resources for bandwidth is a common practice in the industry, but this practice cannot be used under certain circumstances. CDN deployment must align tightly with network deployment to achieve optimal TCO.

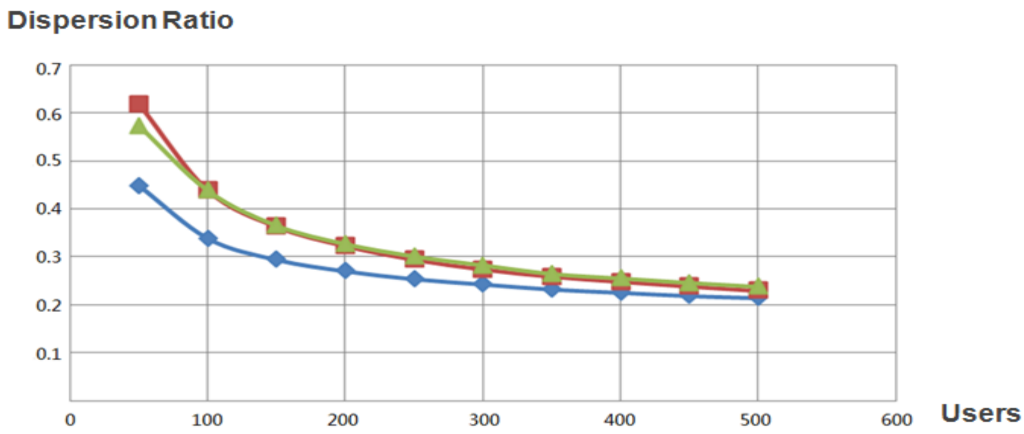
User Quantity of a CDN Affects the CDN Hit Rate and Utilization

CDN costs can be divided into three parts: streaming server costs, storage server costs, and construction costs. CDN utilization is reversely proportional to CDN costs. The lower the CDN utilization, the higher the overall CDN costs.

CDN utilization is determined by the CDN hit rate. CDN hit rate = (1 – Number of source retrieving times/Number of requests) x 100%. If a CDN node has cached the content requested by a user, the CDN node directly provides the service flow upon request. This is called a direct CDN hit. In this case, the number of source retrieving times remains unchanged, but the number of requests increases by 1. If a CDN node has not cached the content requested by a user, the CDN node will obtain the content from its upstream CDN node before providing the content to the user. Meanwhile, the CDN node caches the content locally. In this case, the number of source retrieving times and number of requests both increase by 1.

Based on empirical data analysis results, the CDN hit rate is relevant to the number of concurrent users. When there are only a few users, the requested content is highly dispersed. As the number of users increases, the content dispersion degree gradually decreases. After the number of users reaches a certain level, the content dispersion degree stays relatively stable.

Figure 2-11 Relationship between the number of users and the content dispersion degree

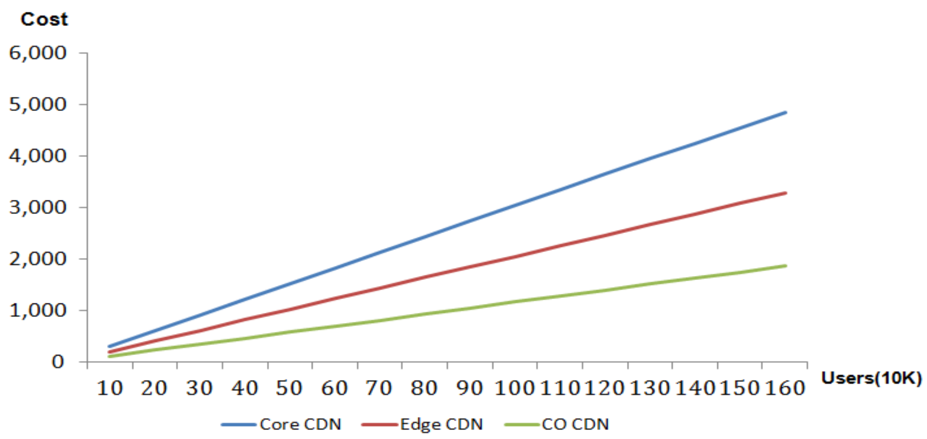


When the user quantity is small, deploying the CDN at a lower layer of the network will significantly reduce the CDN hit rate.

Moving the CDN Downstream Drives Down Network Construction Costs, But Drives Up CDN Construction Costs

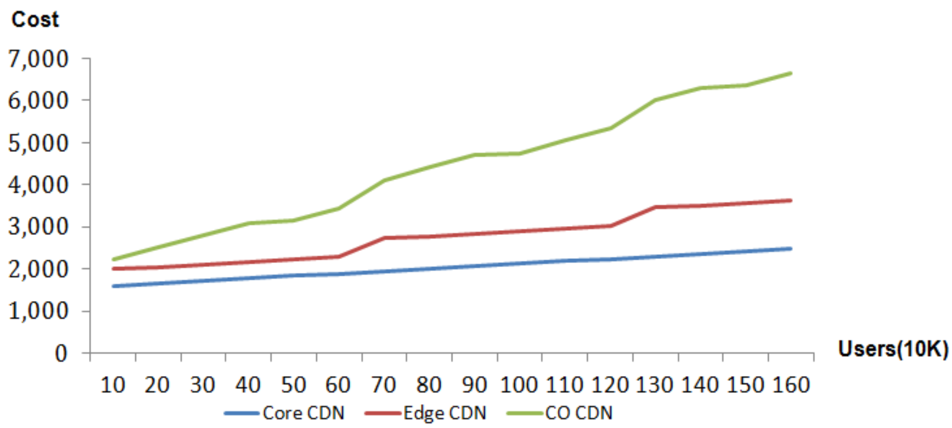
After the CDN is deployed at a lower layer of the network, the utilization of some devices, such as streaming servers, will decrease and new storage servers have to be deployed. As a result, CDN construction costs will rise.

Figure 2-12 Comparison of CDN construction costs



Moving the CDN downstream reduces traffic forwarding hops and intermediate network devices, thereby driving down network expansion costs.

Figure 2-13 Comparison of CDN construction costs after moving the CDN downstream



After analyzing the tradeoff between saved network construction costs and increased CDN construction costs, we find that moving the CDN downstream will be cost-effective only after the number of users reaches a certain level. According to empirical data analysis results, hierarchical CDNs should be deployed as follows:

- Start deploying the level 2 CDN when the number of video service users exceeds 150,000. In this case, a level 2 CDN node hosts not less than 30,000 users.
- Start deploying the level 3 CDN when the number of video service users exceeds 400,000. In this case, a level 3 CDN node hosts not less than 80,000 users.

2.5.4 TCP Acceleration for Network Throughput Improvement

In scenarios where the packet loss rate, delay, and jitter are high, TCP acceleration can effectively improve network throughput and accordingly 4K VOD user experience while shortening the network expansion period and reducing the network expansion scale and costs.

Table 2-5 TCP acceleration effects in the case of 100 Mbps bandwidth, 4 ms delay, and increasing packet loss rates

No.	Delay (ms)	Packet Loss Rate	FTP Download Rate (Mbps) (Not Accelerated)	FTP Download Rate (Mbps) (Accelerated)
1	4	0.010%	100	100
2	4	0.025%	69	100
3	4	0.050%	48	100
4	4	0.100%	34	100
5	4	0.150%	29	100
6	4	0.200%	25	100
7	4	0.500%	17	100
8	4	1.000%	11	100

Table 2-6 TCP acceleration effects in the case of 100 Mbps bandwidth, 0.01% packet loss rate, and increasing delay

No.	Delay (ms)	Packet Loss Rate	FTP Download Rate (Mbps) (Not Accelerated)	FTP Download Rate (Mbps) (Accelerated)
1	4	0.010%	100	100
2	10	0.010%	53	**100**
3	15	0.010%	37	**100**
4	20	0.010%	30	**100**
5	30	0.010%	21	**100**
6	50	0.010%	16	**99.8**
7	100	0.010%	15	**98**
8	150	0.010%	14	**92**

Two mainstream TCP acceleration methods are currently available: network acceleration and CDN acceleration.

Network acceleration

TCP acceleration devices are deployed flexibly on the network to accelerate video service traffic. The acceleration process is transparent, and terminals and service servers are unaware of either acceleration devices or the acceleration process. TCP acceleration effectively reduces impact of network delay and packet loss on the throughput where it takes effect, retaining a high throughput. TCP acceleration can be implemented for traffic of specified users or services, monetizing differentiated services.

CDN acceleration

- In-house developed TCP acceleration algorithms are employed or commercial TCP acceleration software plug-ins are installed on service servers to accelerate video service traffic.
  - If in-house developed TCP acceleration algorithms are to be used, each service provider must develop their own TCP acceleration algorithms and perform software upgrade on each service server. In-house developed TCP acceleration algorithms must be effective.
  - If commercial TCP acceleration software plug-ins are to be used, each service provider must purchase and install such a plug-in on each service server.

Table 2-7 Comparison between TCP acceleration methods

Item	Network Acceleration	CDN Acceleration
Value-added operation	Traffic acceleration can be performed for specified users or services, implementing refined user throughput management and differentiated user experience.	Service-level TCP acceleration is allowed. Refined user-level management and multi-service coordination control are difficult to achieve.
Solution deployment cost	Network acceleration devices can be shared by multiple services and allow on-demand expansion.	Traffic acceleration is specific to services on each service server.
Solution deployment complexity	Only network reconstruction is needed, without modification on terminals or servers. Network acceleration devices are deployed in standalone mode, decoupled from services. The deployment is simple and quick.	All service servers in use must be upgraded, which takes a long time for deployment. Terminal upgrade is also needed in some solutions, which is complex.

Network acceleration and CDN acceleration solutions have their own advantages in specific scenarios. You can consider the operation, costs, and deployment complexity and select one based on the live-network conditions and service targets.

3 O&M: Visualized Video Experience Management

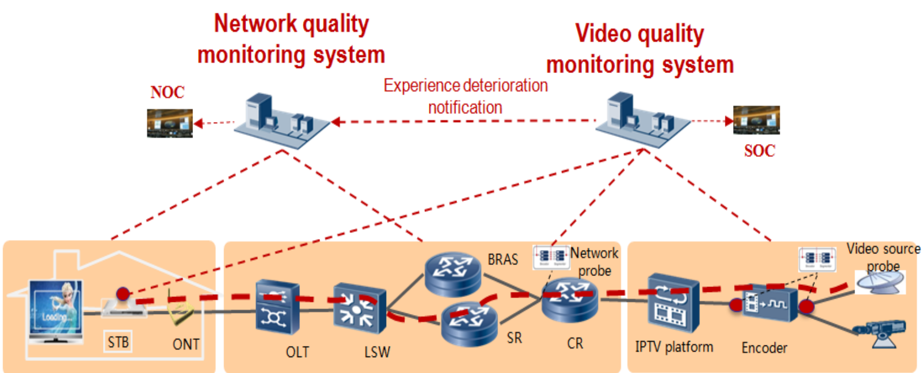
From the O&M perspective, experience-driven 4K bearer networks should be perceivable and manageable.

Perceivable: 4K video user experience can be perceived.

Manageable: Faults can be quickly demarcated and located when video experience deteriorates.

Such a perceivable and manageable high-quality 4K bearer network requires the mapping O&M architecture.

Figure 3-1 Experience-driven 4K bearer network O&M architecture



The **video quality monitoring system** supports media service operations center (SOC) in the following aspects:

- Monitoring user video U-vMOS/KQIs and providing warnings as required
- Monitoring video source quality, providing warnings as required, and providing multi-dimensional analysis and display
- Monitoring platform and STB running conditions and providing multi-dimensional analysis and display

- Locating STB/platform/video source faults and providing analysis of massive quality-related faults

The **network quality monitoring system** supports the network operations center (NOC) in the following aspects:

- Monitoring second-level period port performance for key areas on MAN/core networks, providing warnings as required, and providing multi-dimensional analysis and display
- Automatically discovering user video paths, supporting hop-by-hop detection of KPIs (delay/packet loss/traffic), and providing multi-dimensional analysis and display
- Monitoring access devices' and ONUs' traffic, packet loss, and optical power and providing multi-dimensional analysis and display
- Demarcating and locating faults for video experience quality deterioration

3.1 Video Experience Perception

Perceptible video experience U-vMOS/KQIs is the basic requirement for experience-driven 4K bearer network O&M. Specifically, key indicators such as the U-vMOS, number of video stalls, video stall duration, initial loading time, fast forward and rewind response time, channel change time, and E2E delay, packet, jitter, and throughput should be perceptible for each user and each live channel. In addition, summary and trend analysis can be performed in area, device, and time dimensions.

3.1.1 Carrier Video Experience Perception

The video quality monitoring system can be deployed to obtain user experience for carrier's self-operating videos. That is, software and hardware probes are deployed on the carrier video service platform and STBs to obtain U-vMOS/KQIs.

Table 3-1 Deployment modes for the video quality monitoring system

Service Type	Deployment Position	Probe Type	Deployment Mode	
BTV	Headend encoder egress	Hardware	Traffic direction	For example, traffic directed to an x86 server
	STB	Software	Embedded	For example, embedded in an STB
VOD&OTT	Headend HMS egress	Software	Embedded/traffic direction	For example, embedded in an HMS
	STB	Software	Embedded	For example, embedded in an STB

- Advantages:

  - Really reflects user experience, with a high test precision.
  - Applies to BTV, VOD, and OTT services in both video encryption and non-encryption scenarios.
- Disadvantage:
  - Lacks a unified standard. The implementation may vary with video platform manufacturers.

3.1.2 OTT Video Experience Perception

For OTT videos that are not owned by carriers, third-party OTT services may lack a video quality monitoring system. In this case, a carrier can deploy probes on CRs/BRASs to collect U-vMOS/KQIs.

- Advantages:
  - Easy to deploy: Probes can be deployed at key network nodes.
  - Unified measurement: Different content providers can use a unified solution for experience measurement.
- Disadvantages:
  - The probes need to work with the headend and terminals to monitor user experience when monitoring the Interaction experience or encrypted video quality.
  - Only TCP transmission is supported. For UDP transmission, only the service experience at probe deployment positions, instead of end user experience, can be measured.

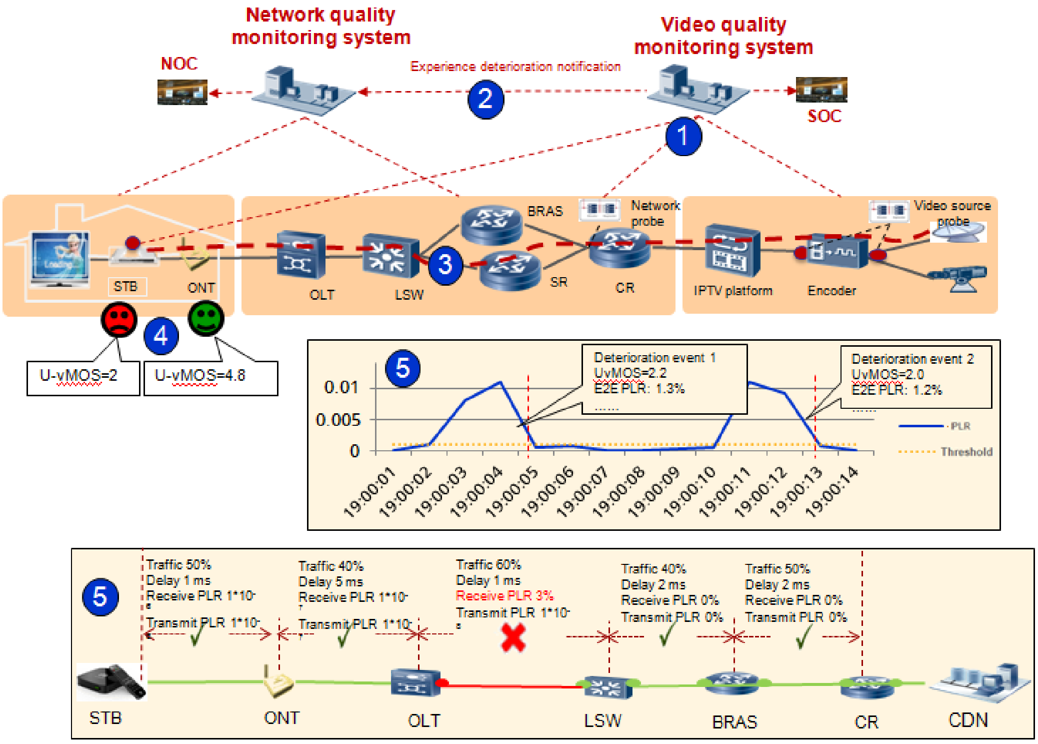
3.2 Video Trouble Shooting

Video services are cross-network basic services for carriers. Fast demarcating and locating of video quality deterioration face the following challenges:

- Randomness: Video mosaic and stalls occur randomly, and fail to recur.
- Instantaneous change: Experience deterioration is closely related to actual paths of video streams and traffic congestion on the paths at the watching time. Without historical information, faults are hard to be located.
- Cross-domain: A fault may involve multiple areas such as the service platform, bearer network, and home network, which requires joint efforts.

To address the preceding challenges, an end-to-end collaboration roadmap is required as follows: experience deterioration detection > system association > real-time path discovery > comparison of key nodes' U-vMOS scores (for demarcation) > hop-by-hop network KPI analysis (for location), as shown in the following figure:

Figure 3-2 Troubleshooting process



Step 2 Perceive experience deterioration.

Use the video quality monitoring system to regularly monitor key indicators such as the U-vMOS, MDI, throughput, rate, and E2E delay and packet loss. A low U-vMOS score is an indication of user experience deterioration.

Step 3 Perform system collaboration.

When a user starts or stops video watching and experience deteriorates, a mapping notification will be sent to the network quality monitoring system. This notification contains the video streams' quintuple information (source and destination IP addresses, source and destination port numbers, and transmission protocol) and U-vMOS/KQIs upon experience deterioration.

Step 4 Discover paths in real time.

After receiving quintuple information of video services, the network quality monitoring system discovers the related E2E paths (ONT-OLT-LSW-BAS-BR-CR) in real time by querying information such as the Layer 3 route forwarding table, ARP tables on BAS devices, Layer 2 device MAC address tables, and NAT table. Scenarios such as Native ETH, L3VPN, L2VPN, and multi-layer NAT, and link load balancing can be coordinated based on actual networking during real-time path discovery.

Step 5 Compare U-vMOS scores of key nodes for fault demarcation.

The network quality monitoring system starts U-vMOS monitoring for key network nodes such as STBs, ONTs, and CRs and helps demarcate faults quickly by comparing their

U-vMOS scores. For example, if the U-vMOS score is 5 for a CR, 4.8 for an ONT, and 2.0 for an STB, quality deterioration occurs on the home network.

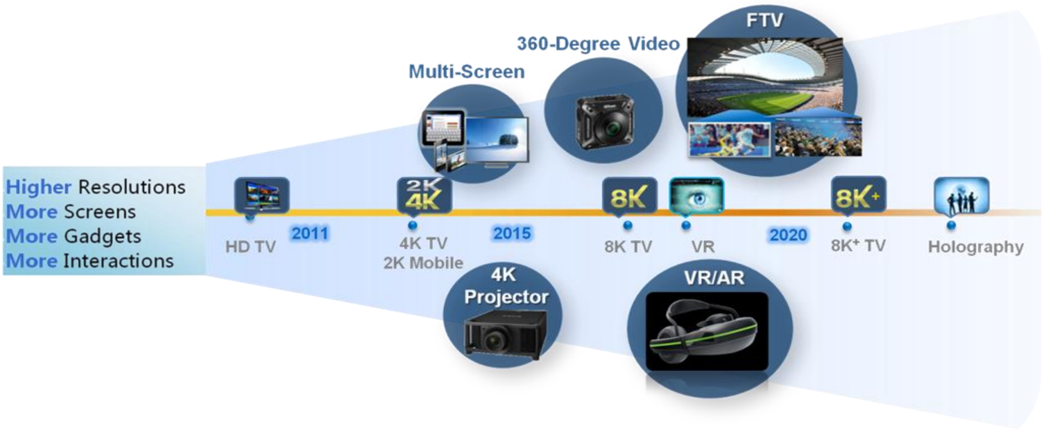
Step 6 Perform hop-by-hop network KPI analysis for fault location.

Second-level period port performance monitoring and video stream performance measurement are started on all upstream and downstream interfaces on the path (ONT-OLT-LSW-BAS-BR-CR) to obtain KPIs hop by hop. In the KPI trend chart, video quality deterioration events and U-vMOS/KQIs are also displayed to support fault locating.

4 Prospect

First, a better service experience, with higher resolution, more screens, and more gadgets and interactions, will be in demand at all times for end users, driving the video traffic growth.

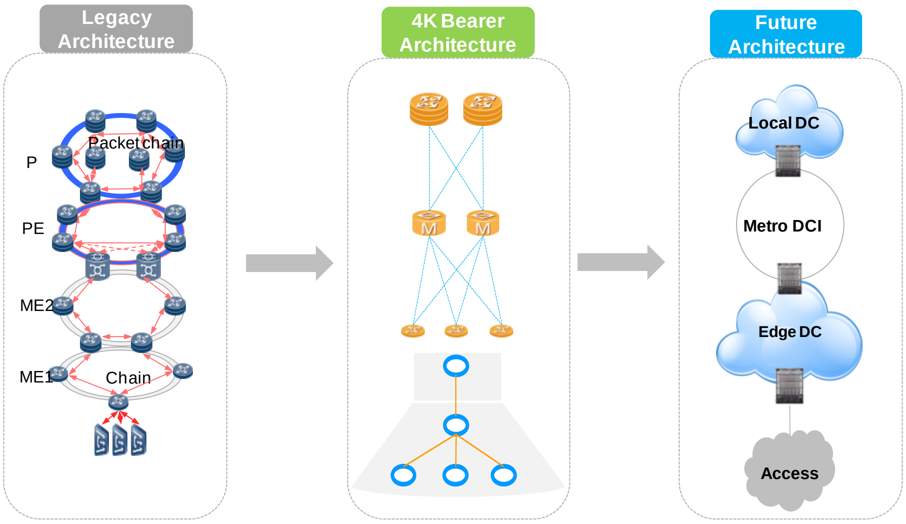
Figure 4-1 The Pursuit of Better Experience Is a Never-Ending Process



Secondly, as BTV services are more commonly carried over OTT, unicast transmission of BTV services renders deteriorated service experience and high bandwidth lease costs. Some carriers have attempted to open multicast capabilities to OTT providers. This attempt builds new revenue streams and relieves pressure on capacity expansion, as well as reduces costs for OTT providers.

As cloud services grow, the cloud storage and computing demand is surging. To cope with this trend, carriers are increasing DC infrastructure construction, and the DC-centric network architecture has been widely recognized by carriers. AT&T's Domain 2.0, China Unicom's Cube-Net, and Telefonica's UNICA are typical programs proposing a DC-centric network architecture. In the long run, service cloudification and edge DC construction both drive the reconstruction of network architecture, and all service models are keeping pace with the DC-centric architecture. Video will be no exception.

Figure 4-2 Smooth evolution to the future architecture



Experience-driven 4K bearer architecture has smooth evolution capability to evolve to the future architecture.

- E2E high-throughput pipes provide resource pools with sufficient bandwidth for the development of 4K and cloud services.
- BNG to edge is applicable to anywhere CDN and edge DC deployment.

A References

- White Paper on Huawei U-vMOS Video Experience Standards v1.0
- TR 126: Triple-play Services Quality of Experience (QoE)
- VRP V800R011C00 Technical White Paper on Virtual Access

B Acronyms and Abbreviations

Acronym and Abbreviation	Full Name
AP	access point
CBR	constant bit rate
CDN	content delivery network
CO	central office
FCC	fast channel change
FD	fiber doctor
HLS	HTTP Live Streaming
KPI	key performance indicator
KQI	key quality indicator
NOC	network operation center
OD	optical doctor
PID	photonics integrated device
PLC	power line communication
RET	retransmission
RTT	round-trip time
SLA	service level agreement
SOC	service operation center
SSID	service set identifier
STB	set top box
UDP	User Datagram Protocol
VBR	variable bit rate