画像情報特論 (1)
Advanced Image Information (1)

Introduction and Streaming Background

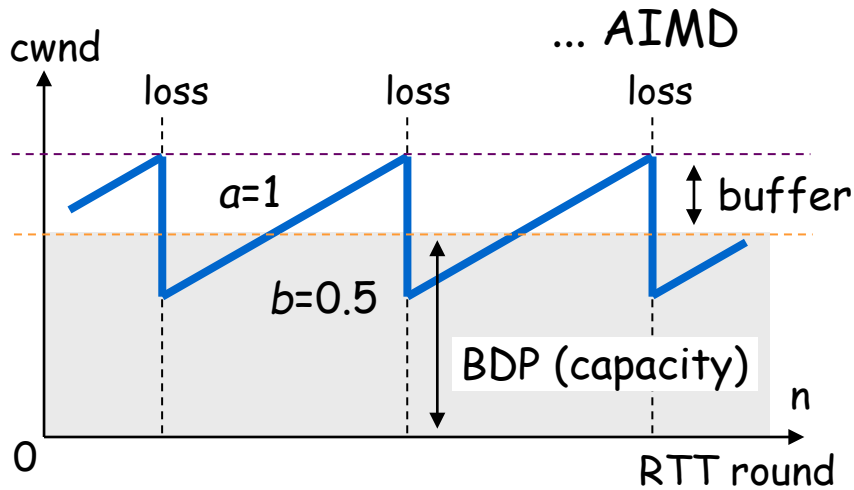情報理工・情報通信専攻　甲藤二郎

Dept. of Computer Science and Engineering, Jiro Katto
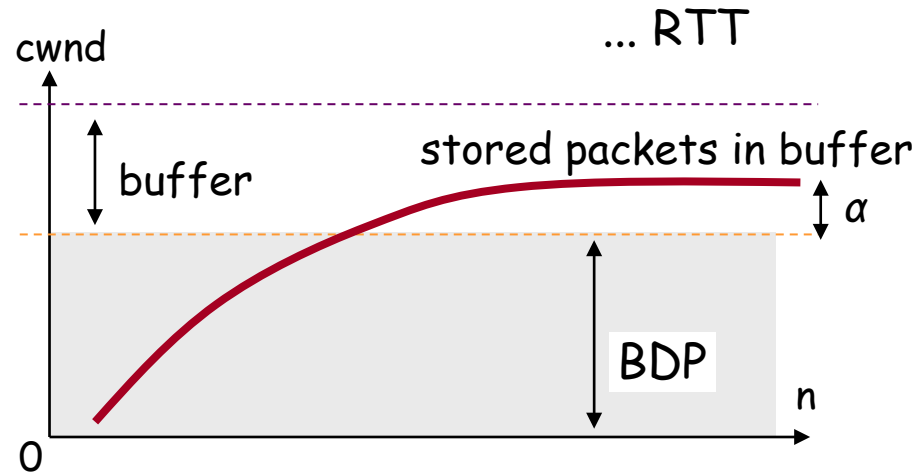
E-Mail: katto@waseda.jp

Introduction

TCP Variants

■ Loss-based

... AIMD

loss   loss   loss

cwnd

a=1

b=0.5

buffer

BDP (capacity)

0

n

RTT round

TCP-Reno, High-Speed TCP,
TCP-Westwood, CUBIC-TCP, ...

■ Delay-based

... RTT

cwnd

buffer

stored packets in buffer

α

BDP

0

n

TCP-Vegas, FAST-TCP

■ Hybrid   Compound TCP

■ TCP-BBR

BDP: Bandwidth-Delay Product

RTP and TFRC

- TFRC (over RTP/UDP)

slow start

transmission rate

congestion avoidance

packet losses

time

Modeling of steady-state
TCP behaviors

$$R = \frac{1}{RTT} \sqrt{\frac{3}{2p}}$$

p: packet loss rate

BDP/Buffer relationship
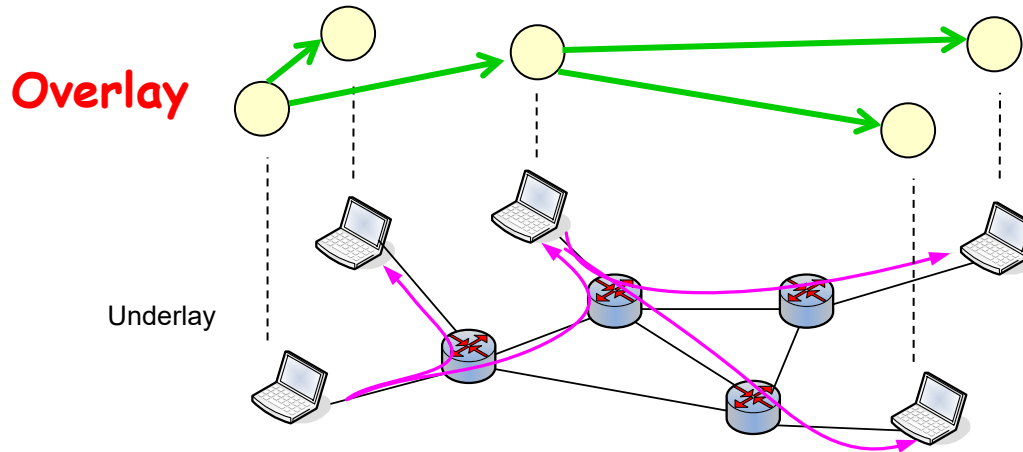
small buffer → ×efficiency
large buffer → ×delay

TFRC: TCP Friendly Rate Control

HTTP and MPEG-DASH

- ## MPEG-DASH: Dynamic Adaptive Streaming over HTTP

  - Multiple (bitrate, resolution) pairs ... representation
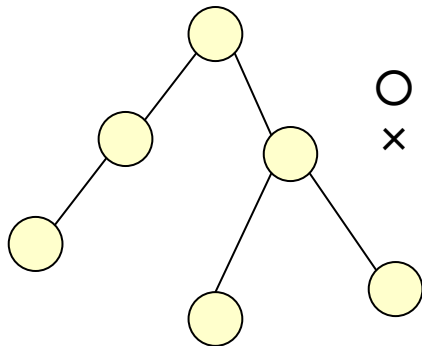 - Adaptive selection of representations

list of URLs

representations

I. Sodagar, et al. : The MPEG-DASH Standard for Multimedia Streaming Over the Internet, IEEE Multimedia, Apr.2011.

CDN, P2P & Cloud

- Overlay networks
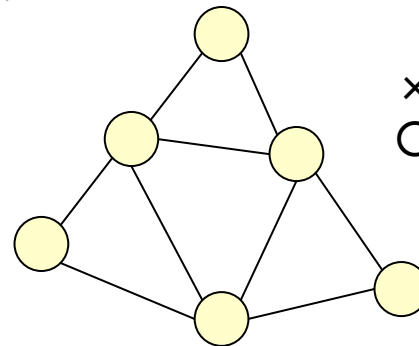
Overlay

Underlay

- tree

○ complexity
× robustness
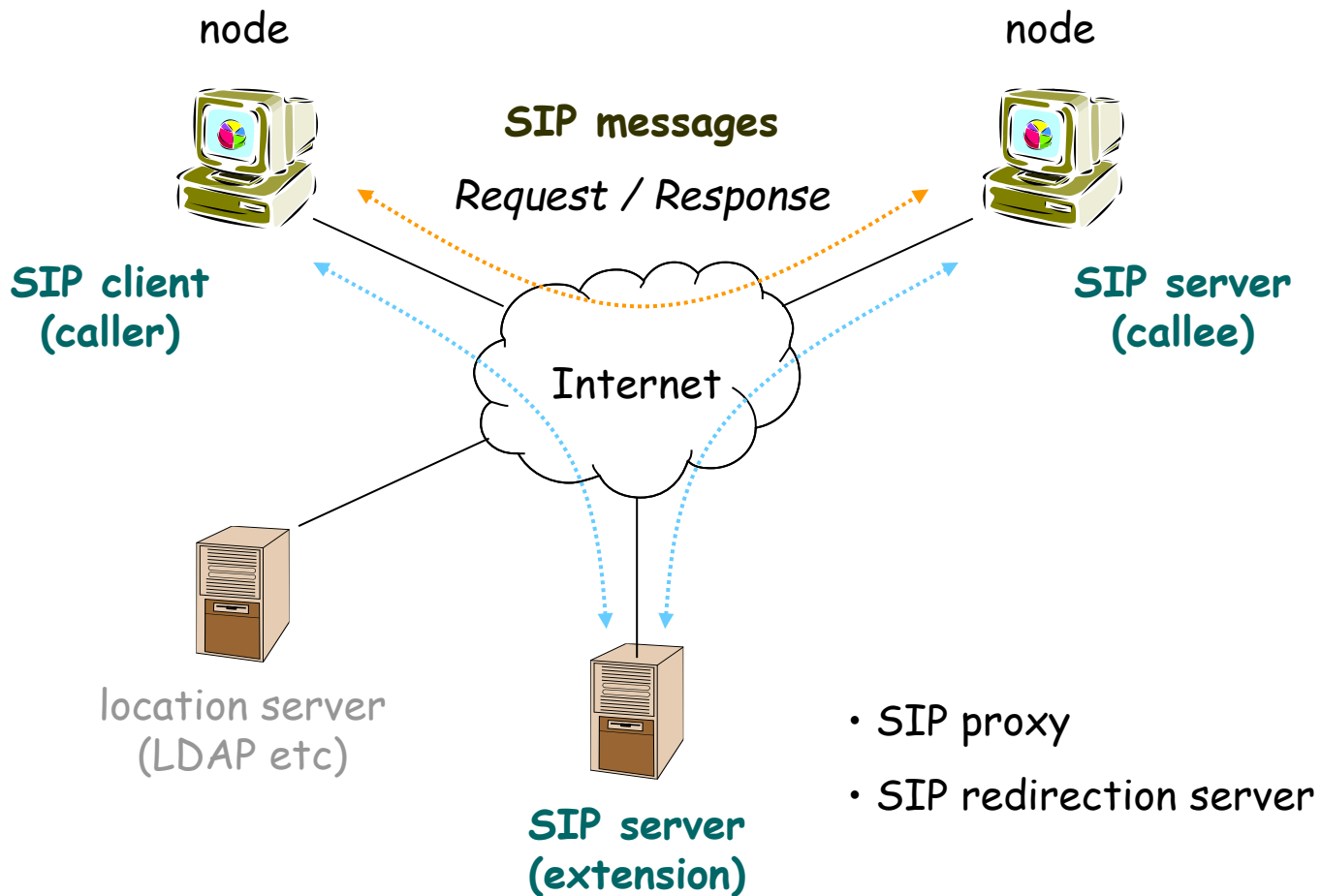
- mesh

× complexity
○ robustness

SIP and WebRTC

- SIP: Session Initiation Protocol

node   node

SIP messages

Request / Response

**SIP client
(caller)**

**SIP server
(callee)**

Internet

location server
(LDAP etc)

- SIP proxy

- SIP redirection server

**SIP server
(extension)**

RFC3261: Session Initiation Protocol

Video Compression Basics

T.Ebrahimi: "Trends and Challenges in Video Coding", June.2009.

Video Compression History

| 80 | 85 | 90 | 95 | 00 | 05 | 10 | 15 | 20 |

| MC | DCT | **MC+DCT** | **B-picture half-pel** | field | error resilience shape coding | **RD optimization deblocking filter** | **hierarchical B-picture** | **CU/PU/TU** |

ITU-T

| H.120 | Sub-rate | H.261 | | H.263/ H.263+ |

1.5M　　384k　　64k - 1.5M　　analog modem (9.6kb/s)

TV phone/conference　⇒ Internet, mobile
over ISDN

Joint (ITU-T & ISO/IEC)

| MPEG-2 / H.262 | | H.264/AVC | extentions | H.265/HEVC | H.266/VVC |

Digital Broadcasting　　Internet, mobile,　　3D, UHDTV, …　　360 degree,
DVD, HDTV (MPEG-3)　　Bluray　　　　　　　　　　　　　　　　AR/VR,
　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　…

ISO/IEC

| MPEG-1 | MPEG-4 |

CD-ROM, Karaoke　Internet, mobile

On2　　　　VC9　　　　AV1

AVC, HEVC and VVC

- ## HEVC (2013)

- ## VVC (2020)

Manege, 3840 x 2160, 60 Hz

HEVC (2013)
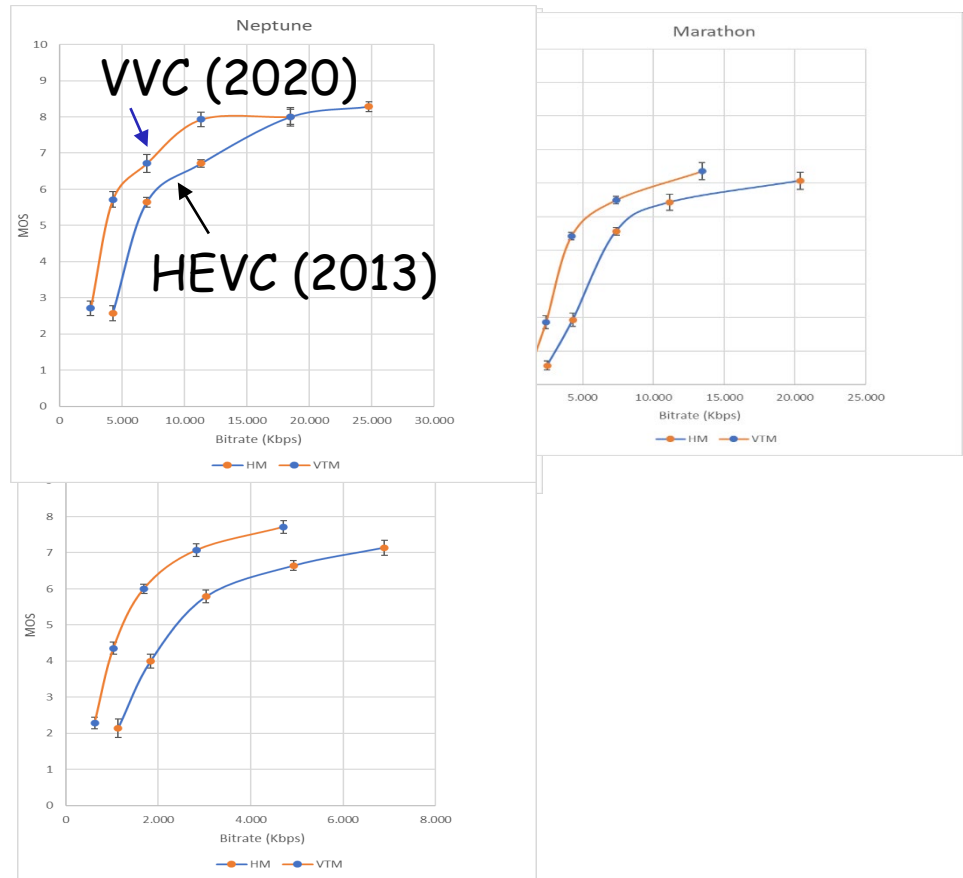
AVC(2003)

- HEVC MOS
- AVC MOS
- HEVC MOS BD-rate range
- AVC MOS BD-rate range
- HEVC PSNR
- AVC PSNR
- HEVC PSNR BD-rate range
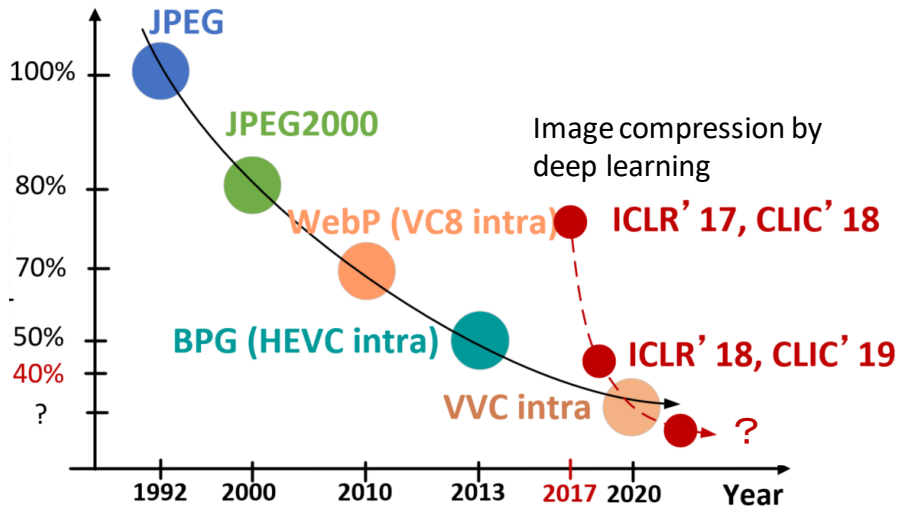- AVC PSNR BD-rate range

Neptune

VVC (2020)

HEVC (2013)

Marathon

Learned Image Compression

■ Active topics in these seven years

Compression performance

CLIC in CVPR and DCC

International standard for
image compression

http://www.compression.cc/

Image Quality Assessment

MOS Score vs PSNR

PSNR
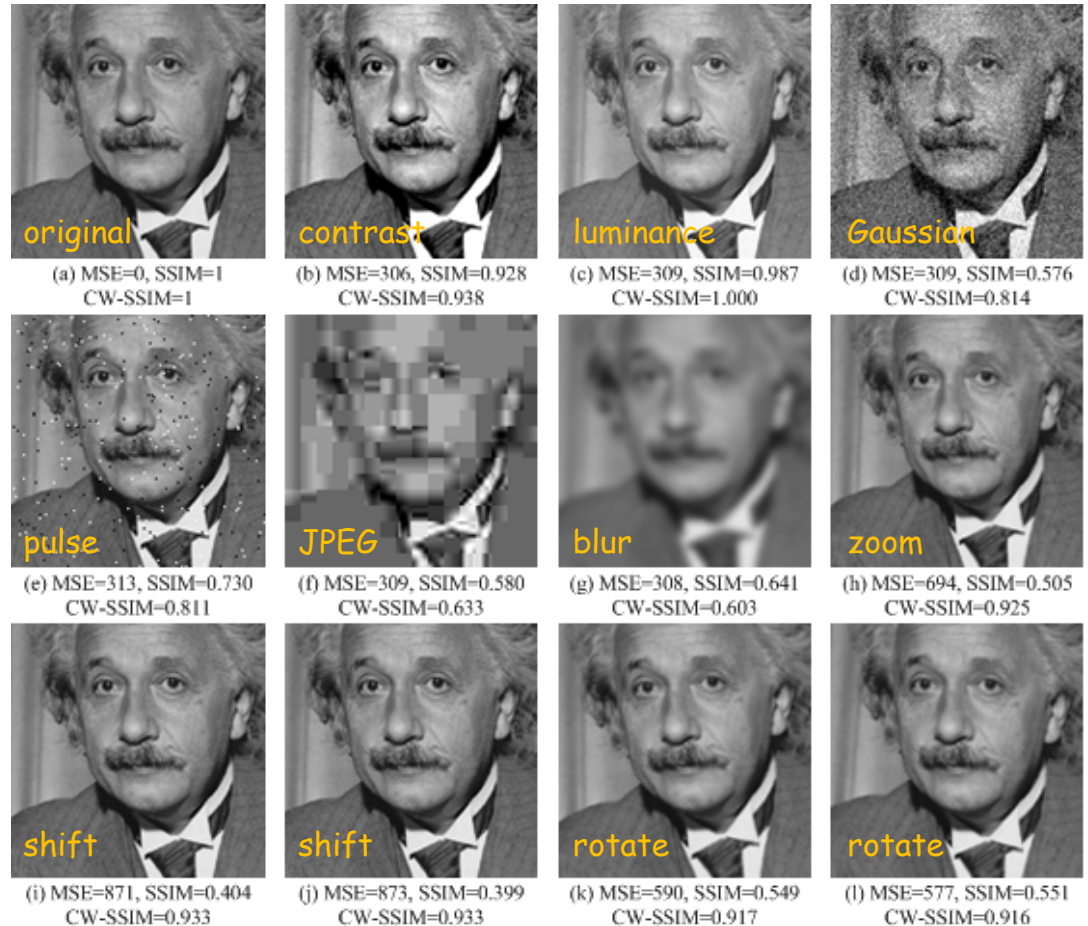
MOS

There exists strong correlation between MOS and MSE but not enough

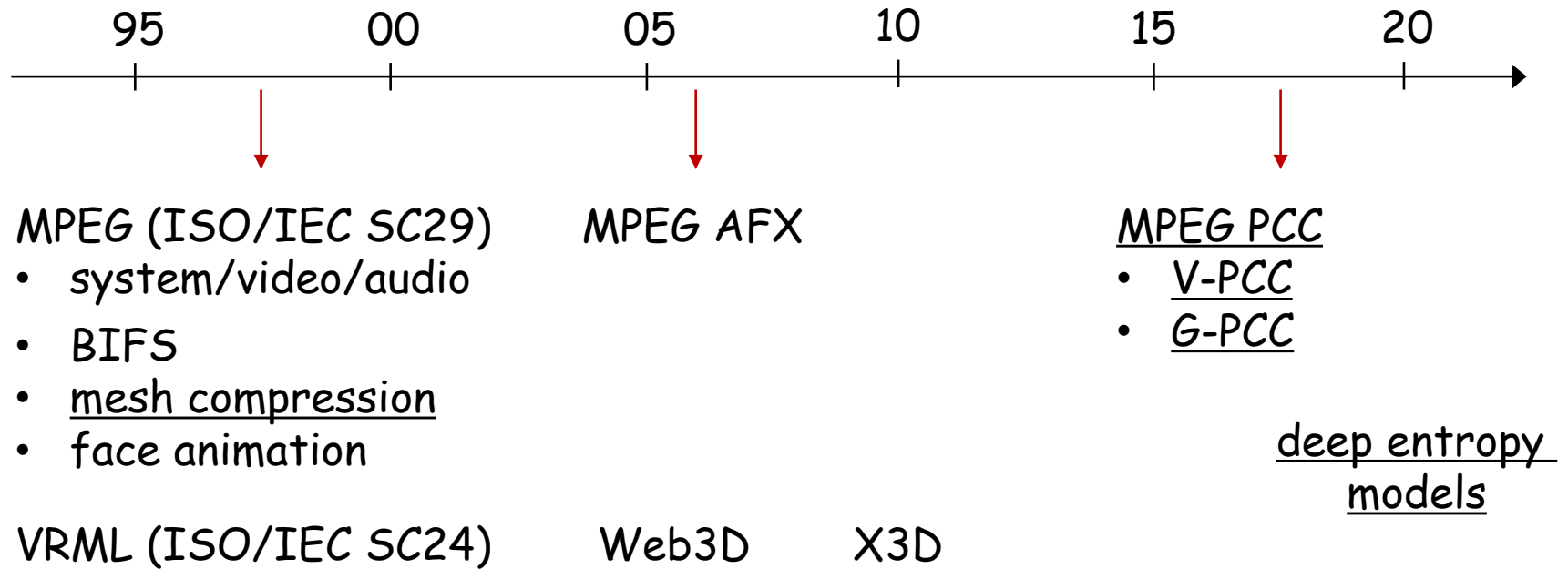(b)-(g) images have the same MSEs, but subjective impressions are different

(h)-(l) images are scaled, shifted or rotated, and have different MSEs

PSNR, SSIM, VMAF, LPIPS, …

original — (a) MSE=0, SSIM=1 CW-SSIM=1
contrast — (b) MSE=306, SSIM=0.928 CW-SSIM=0.938
luminance — (c) MSE=309, SSIM=0.987 CW-SSIM=1.000
Gaussian — (d) MSE=309, SSIM=0.576 CW-SSIM=0.814
pulse — (e) MSE=313, SSIM=0.730 CW-SSIM=0.811
JPEG — (f) MSE=309, SSIM=0.580 CW-SSIM=0.633
blur — (g) MSE=308, SSIM=0.641 CW-SSIM=0.603
zoom — (h) MSE=694, SSIM=0.505 CW-SSIM=0.925
shift — (i) MSE=871, SSIM=0.404 CW-SSIM=0.933
shift — (j) MSE=873, SSIM=0.399 CW-SSIM=0.933
rotate — (k) MSE=590, SSIM=0.549 CW-SSIM=0.917
rotate — (l) MSE=577, SSIM=0.551 CW-SSIM=0.916

[FIG2] Comparison of image fidelity measures for "Einstein" image altered with different types of distortions. (a) Reference image. (b) Mean contrast stretch. (c) Luminance shift. (d) Gaussian noise contamination. (e) Impulsive noise contamination. (f) JPEG compression. (g) Blurring. (h) Spatial scaling (zooming out). (i) Spatial shift (to the right). (j) Spatial shift (to the left). (k) Rotation (counter-clockwise). (l) Rotation (clockwise).

Z.Wang and A.Bovik: "Mean Squared Error: Love It or Leave It", IEEE Sig. Mag., Jan 2009.

Point Cloud Compression

| 95 | 00 | 05 | 10 | 15 | 20 |

MPEG (ISO/IEC SC29)
- system/video/audio
- BIFS
- mesh compression
- face animation

VRML (ISO/IEC SC24)

MPEG AFX

Web3D   X3D

MPEG PCC
- V-PCC
- G-PCC

deep entropy
models

BIFS: BInary Format for Scene description
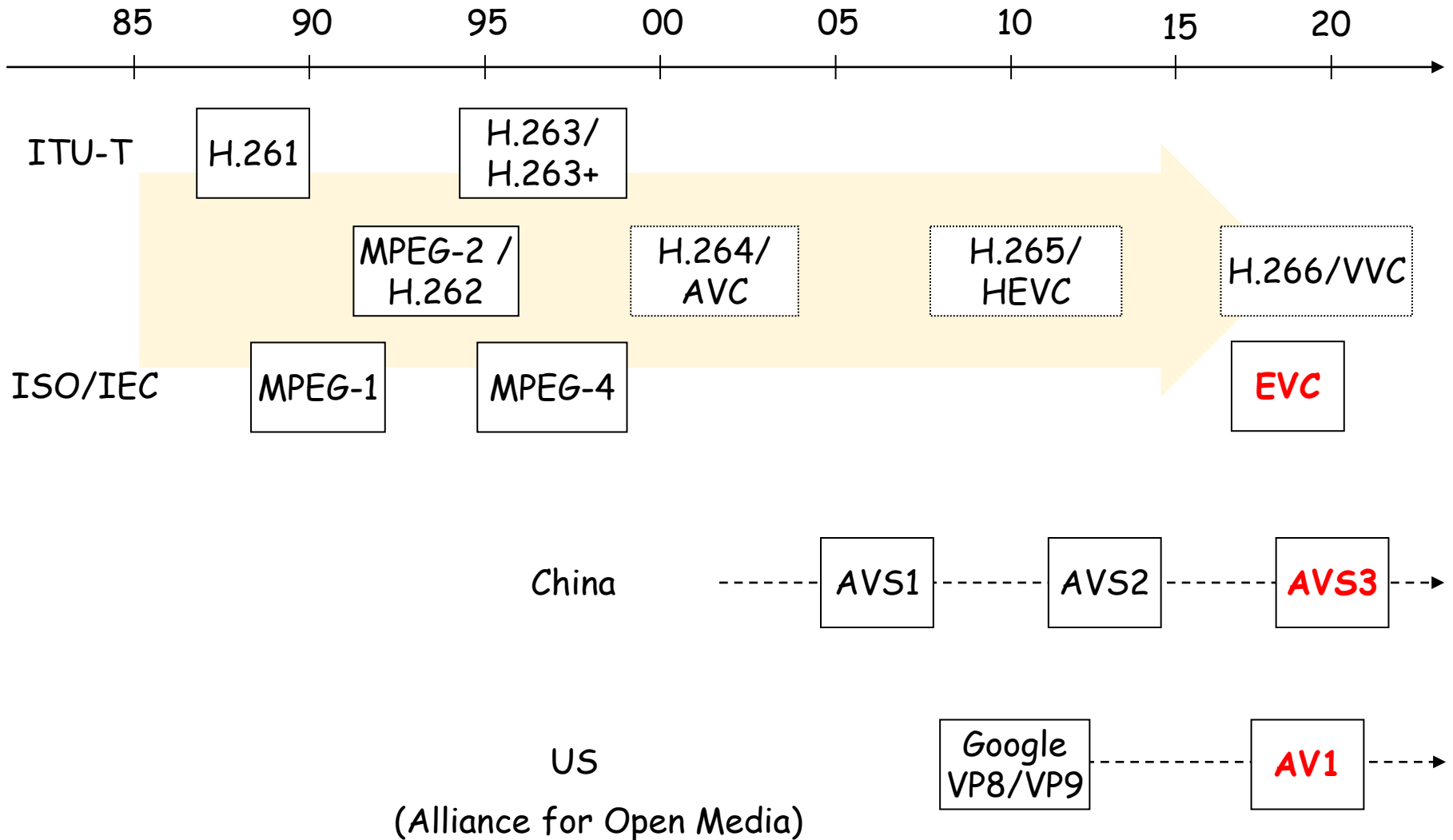VRML: Virtual Reality Modeling Language
AFX: Animation Framework eXtension

PCC: Point Cloud Compression
V-PCC: Video-based PCC
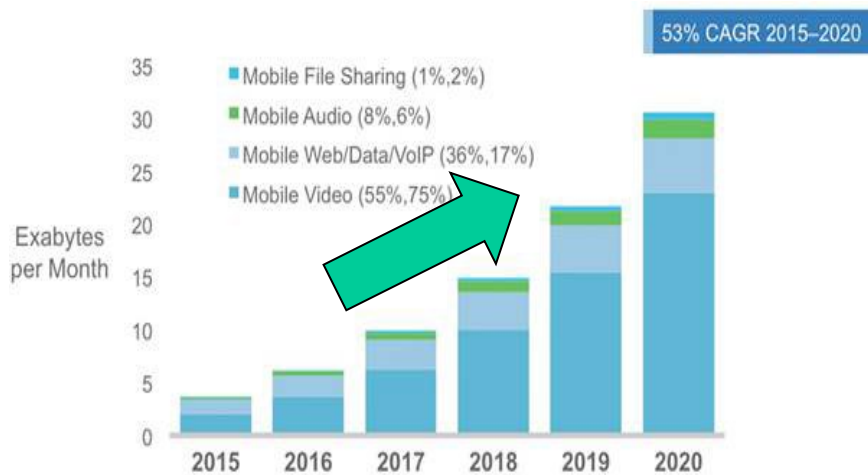G-PCC: Geometry-based PCC

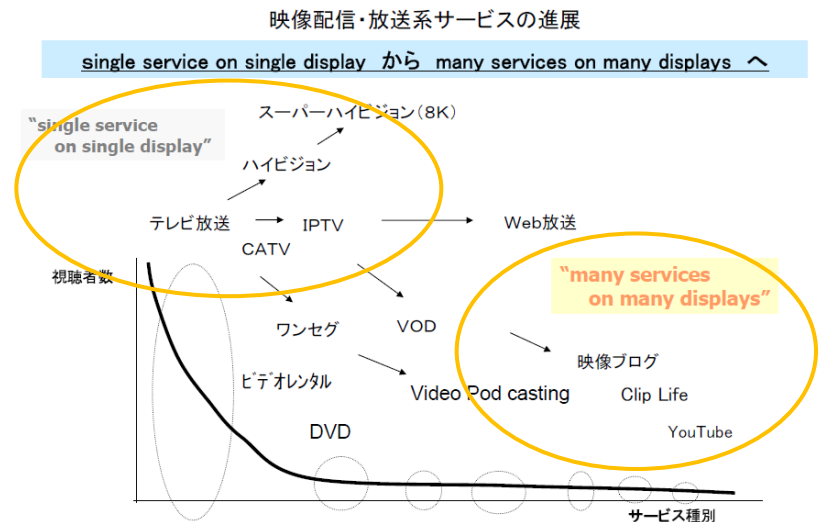# EVC, AVS, and AV1

Streaming Background

Recent Trends

- Drastic Increase of Video Traffic on Internet
  - more than 70%

- Evolution of Various Video Services
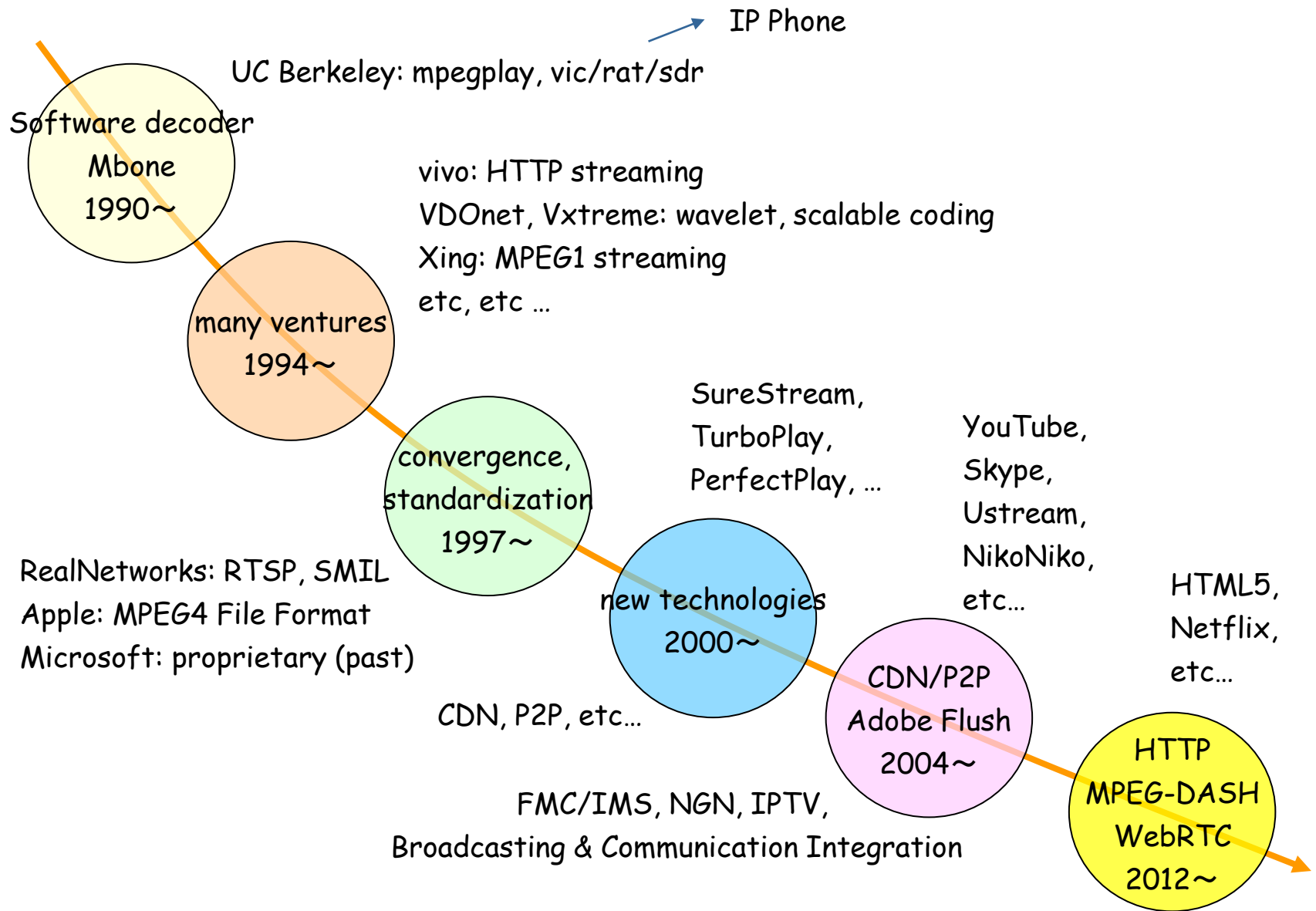 - higher resolution and personalization

(Cisco VNI, 2016)

(MIC Report, 2008)

History of Video Streaming

IP Phone

UC Berkeley: mpegplay, vic/rat/sdr

Software decoder
Mbone
1990~

vivo: HTTP streaming
VDOnet, Vxtreme: wavelet, scalable coding
Xing: MPEG1 streaming
etc, etc …

many ventures
1994~

SureStream,
TurboPlay,
PerfectPlay, …

YouTube,
Skype,
Ustream,
NikoNiko,
etc…

convergence,
standardization
1997~

RealNetworks: RTSP, SMIL
Apple: MPEG4 File Format
Microsoft: proprietary (past)

HTML5,
Netflix,
etc…

new technologies
2000~

CDN/P2P
Adobe Flush
2004~

CDN, P2P, etc…

HTTP
MPEG-DASH
WebRTC
2012~

FMC/IMS, NGN, IPTV,
Broadcasting & Communication Integration

Protocol Stack of RTP/UDP Video Streaming (and IP phone)

protocol stack for low-delay & interactive video streaming (e.g. conference)

application (L7)	video (H.264 etc…)	audio	SDP	layout (HTML. SMIL)
adaptation	RTP / RTCP		RTSP, SIP, SAP*	HTTP
transport (L4)	UDP / TCP / DCCP		TCP / UDP / SCTP	
network (L3)	IP (IPv4, IPv6, IP-multicast)			
datalink & physical (L2 & L1)	actual networks (802.3 (ethernet), 802.11 (WiFi), etc)			

* SAP: delivered by IP-multicast for program advertisement

Protocol Stack of HTTP Video Streaming

protocol stack for one-way video streaming

application (L7)	video (H.264 etc…)	audio	MPD (MPEG-DASH)	layout (HTML)
adaptation	HTTP			
transport (L4)	TCP			
network (L3)	IP (IPv4, IPv6)			
datalink & physical (L2 & L1)	actual networks (802.3 (ethernet), 802.11 (WiFi), etc)			

Protocol Stack of WebRTC

protocol stack for low-delay & interactive video streaming (e.g. conference)
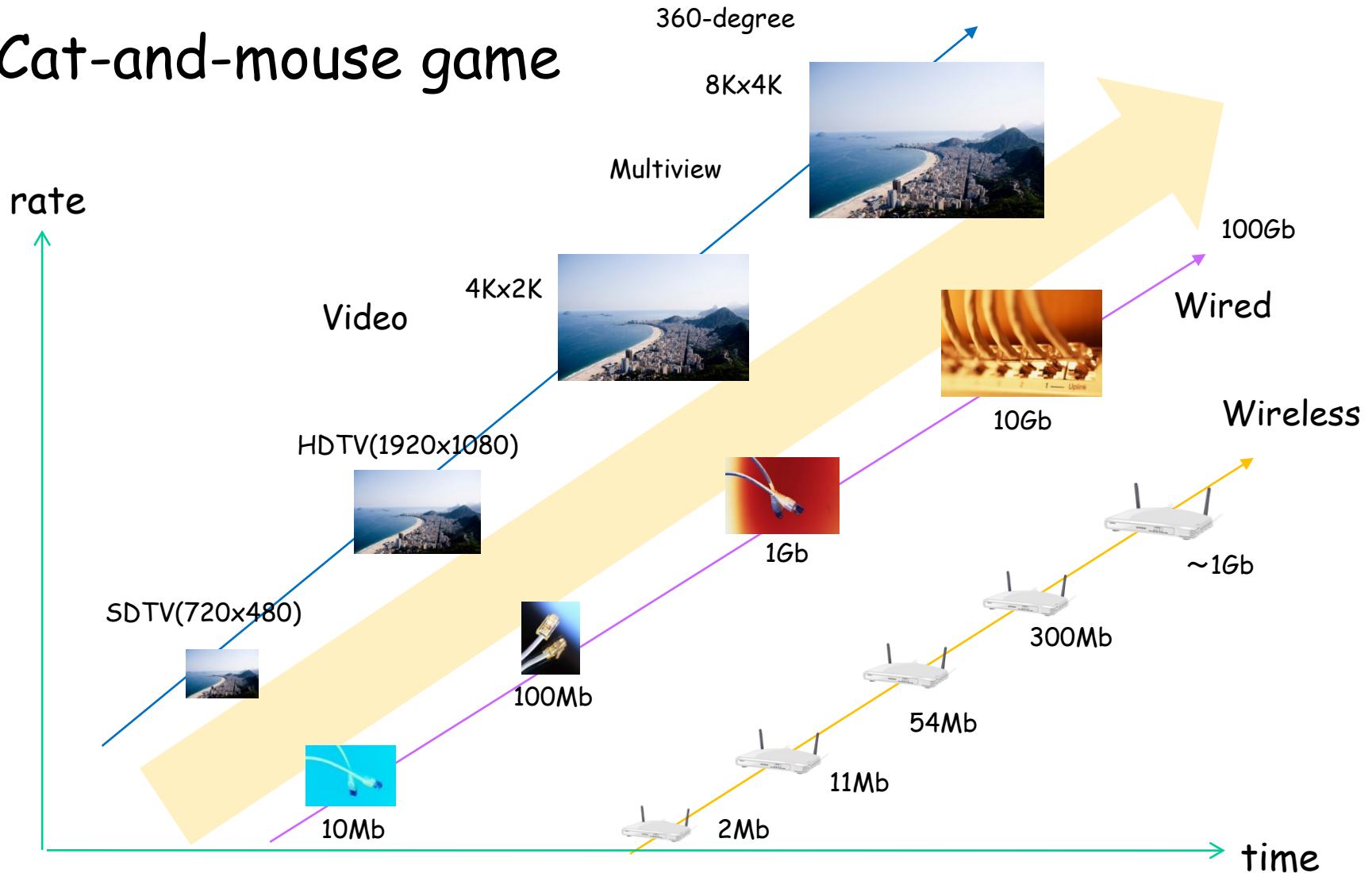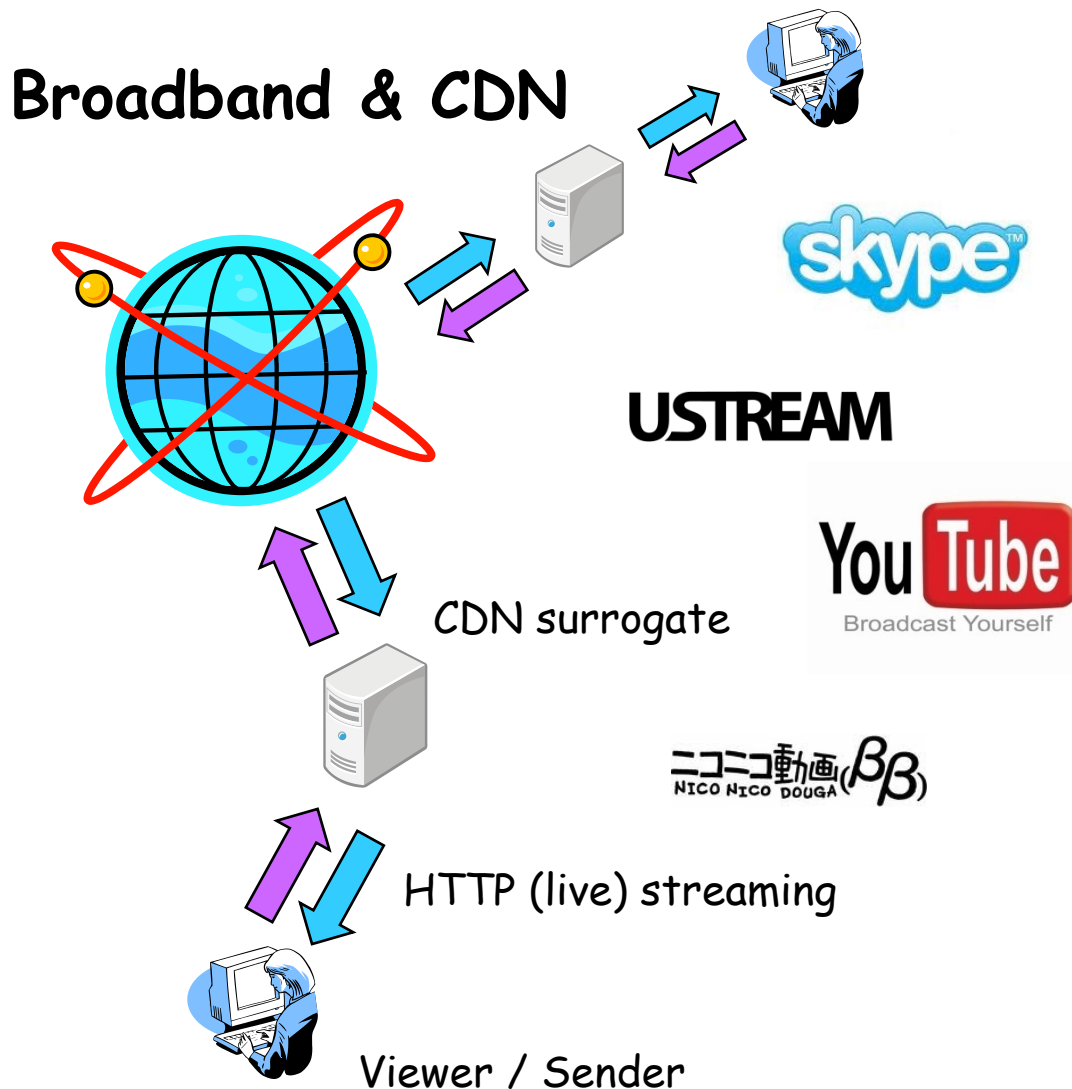
NAT traversal	media		data	signaling
	video	audio	data	SIP, SDP
STUN, TURN	SRTP		SCTP/DTLS	HTTP/TLS, WebSocket
UDP				TCP
IP				
MAC / PHY				

Networks and Multimedia

- Cat-and-mouse game

Broadband and CDN

Broadband & CDN

Viewer / Sender

CDN surrogate

HTTP (live) streaming

skype

USTREAM

You Tube Broadcast Yourself

ニコニコ動画(ββ) NICO NICO DOUGA

RTP/UDP & RTSP & TFRC

→ HTTP/TCP streaming

- Broadband
- CDN (Akamai, Lime Networks)
- Firewall (port 80)
- …

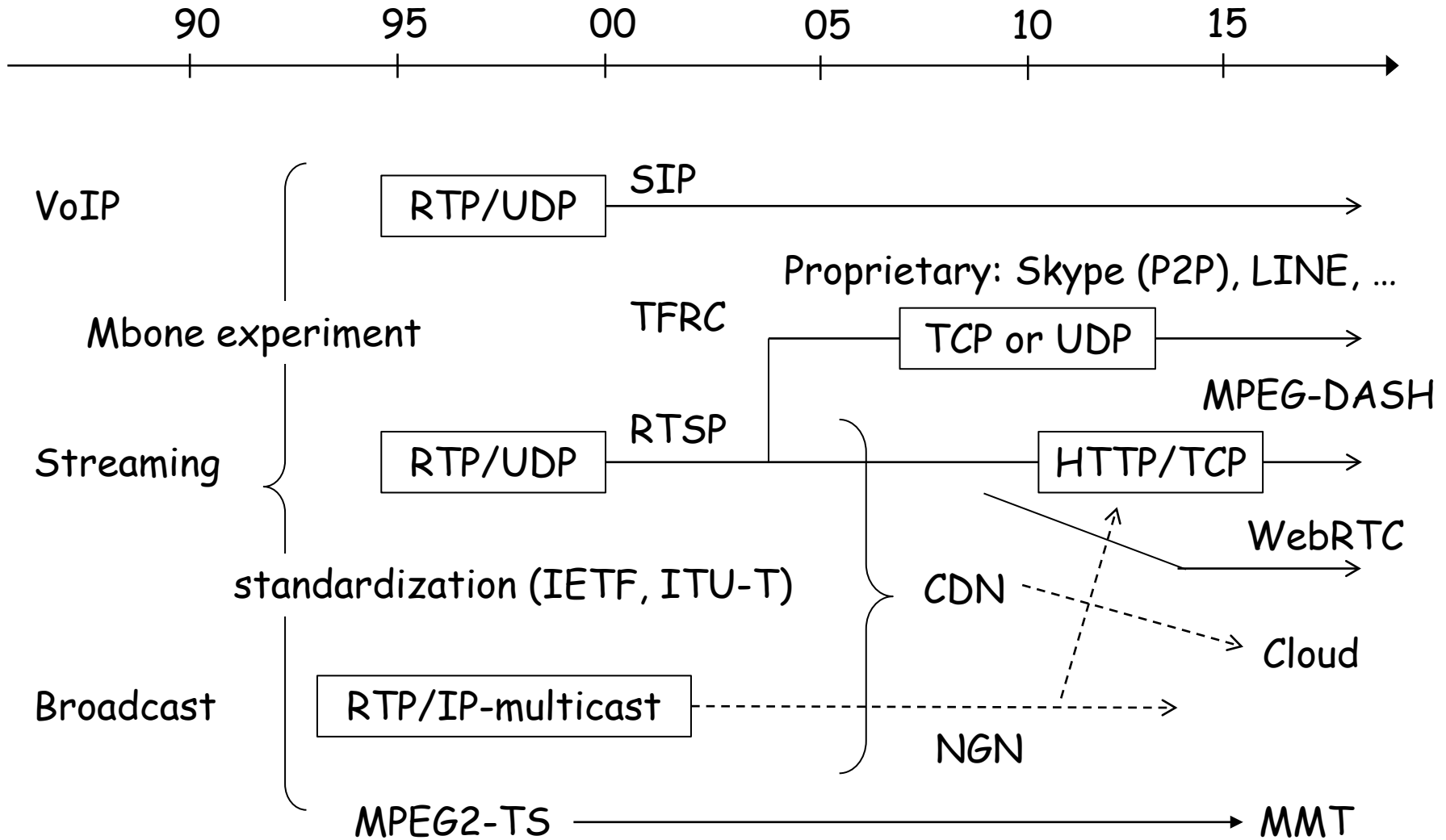One-way (on-demand / live)
- large buffer

Bi-directional (interactive)
- small buffer

CDN: Content Delivery Network ⇒ Cloud

IP Video Services

Services	Examples
IP phone & conference (interactive)	Telecommunication (SIP, H.323)
IPTV (one-way)	CATV, Telecommunication (MPEG-2 TS)
Web conferencing (interactive)	Zoom, Cisco WebEx, Skype, Google Hangout, etc …
Video streaming (one-way)	YouTube, Amazon Prime Video, Facebook, etc …

Protocol Transition

90 95 00 05 10 15

VoIP — RTP/UDP — SIP

Mbone experiment TFRC Proprietary: Skype (P2P), LINE, …

TCP or UDP

MPEG-DASH

Streaming — RTP/UDP — RTSP — HTTP/TCP

WebRTC

standardization (IETF, ITU-T) CDN

Cloud

Broadcast — RTP/IP-multicast

NGN

MPEG2-TS — MMT

TCP vs. UDP

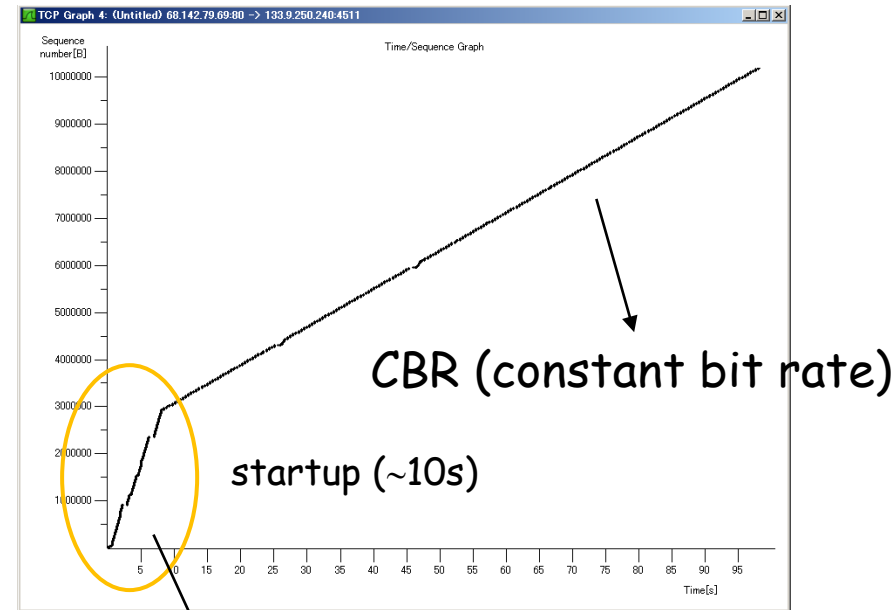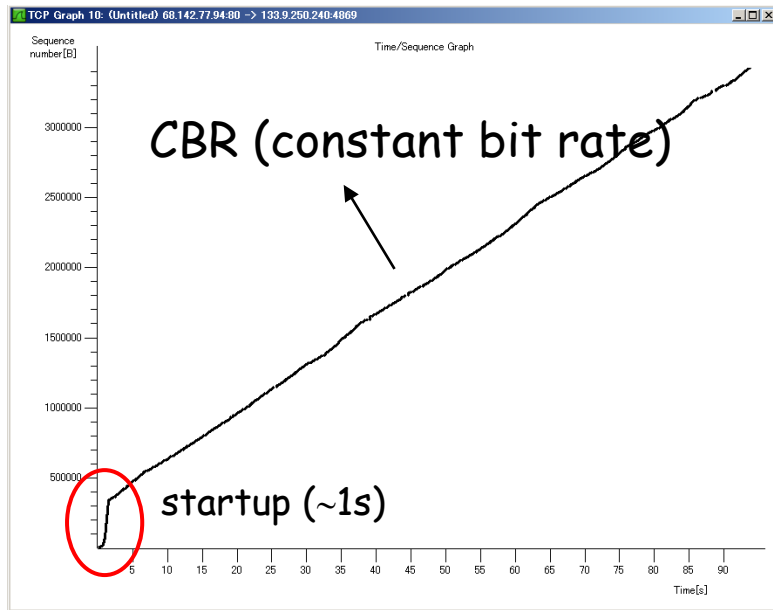|	Reliability	Low Delay	Congestion Control	Typical Application
TCP	◎ (ACK and lost packet retransmission)	✕ → ◯ (thanks to CDN & broadband network)	◯ → ◎ (TCP versions)	One way (on-demand) streaming
UDP	✕ (no ACK nor sequence number)	◎ (no ACK nor packet retransmission)	✕ → △ (RTP/RTCP and TFRC)	Interactive (bi-directional) phone & conference

prefetching & CBR

(prefetch, then CBR)

sequence
number

Live

On-Demand

CBR (constant bit rate)

CBR (constant bit rate)
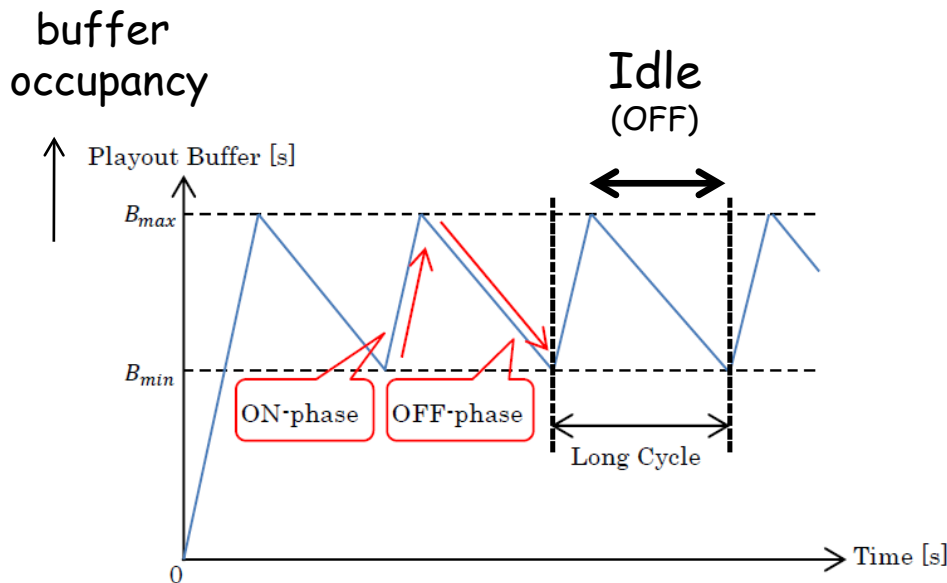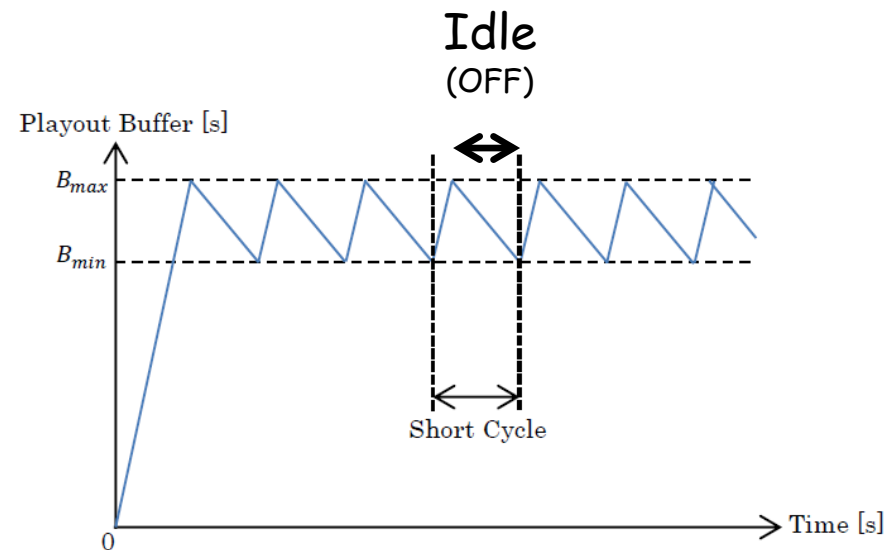
startup (~1s)

startup (~10s)

time

prefetching

ON/OFF cycles

(prefetch & idle cycles)

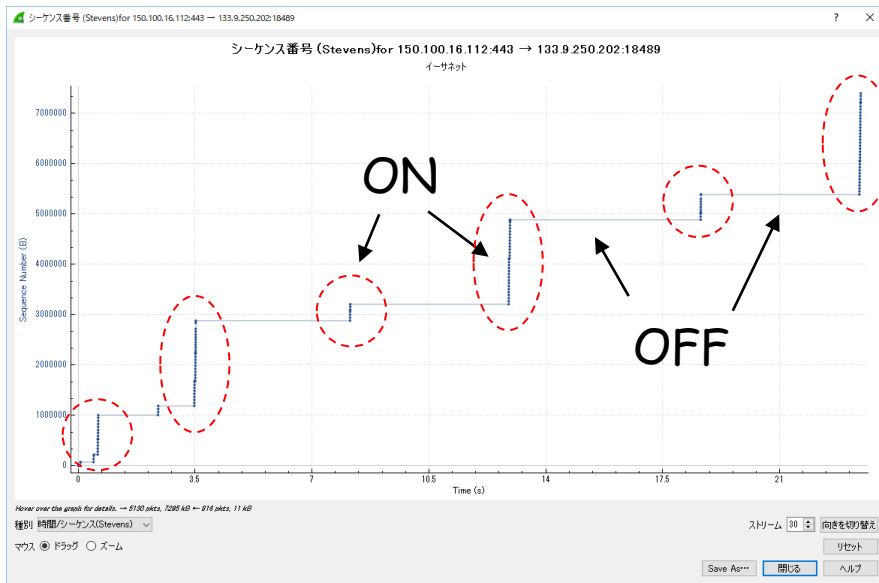- ## receiver buffer behaviors

(a) long ON-OFF Cycle (sawtooth)

(b) short ON-OFF Cycle (zippy pacing)
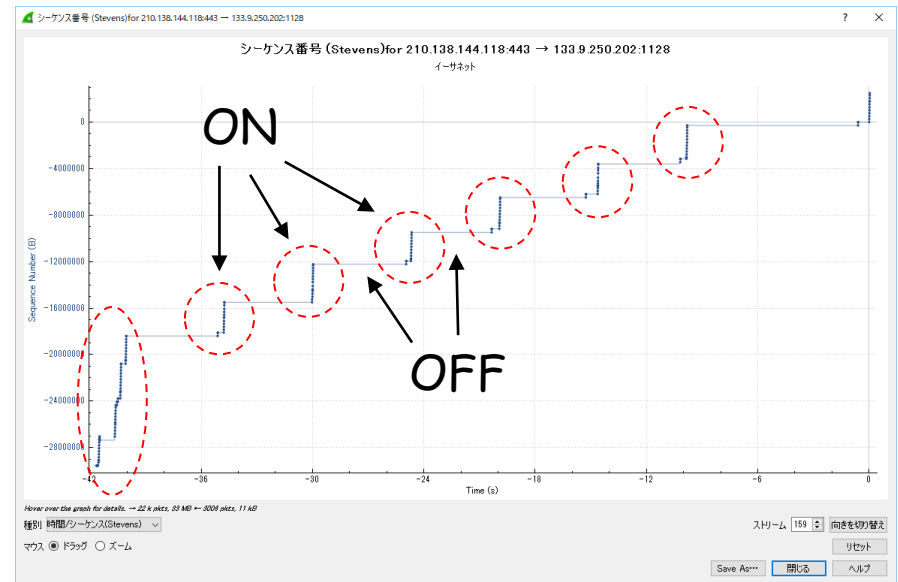
A. Rao, et al. ACM CoNEXT 2011

ON/OFF cycles

- sequence number behaviors

sequence
number

example 1 (YouTube)

example 2 (TVer)

This year's schedule (tentative)

This Year's Schedule

tentative

(Apr 12)	Class overview and backgrounds of video streaming
(Apr 19)	TCP variants
(Apr 26)	RTP and TFRC over UDP
(May 10)	HTTP and MPEG-DASH
(May 17)	CDN, P2P and Cloud
(May 24)	SIP and WebRTC
(May 31)	Other topics and <u>online test</u>
(June 07)	Video compression basics
(June 14)	H.264/AVC
(June 21)	HEVC/H.265 and VVC/H.266
(June 28)	Learned image compression
(July 05)	Image quality assessment
(July 12)	Point cloud compression
(July 19)	Class summary and <u>online test</u>
on Moodle	Final report