

OTT STREAMING FUNDAMENTALS

2-DAY ONSITE TECHNICAL TRAINING

- **Instructor:** Clement Duval (clement@videostreamstack.com)
- **Duration:** 14 hours over 2 days, 9:00am–5:30pm, with breaks/lunch
- **Language:** English or French
- **Requirements:** basic knowledge of networking and IPTV (desirable)
- **Audience:** Engineers and developers wanting to dive deep into the OTT protocols
- **Location:** Delivered on-site to the customer's facilities, 15 students maximum
- **Format:** a total of **100 interactive quizzes** and hands-on **labs on real-world**, live manifests.

DAY 1 – FOUNDATIONS

Module 1 – HTTP Adaptive Streaming

IPTV vs OTT

- What OTT is called in such a way
- What separates OTT from IPTV in network type, transmission protocol, interoperability, and display mode

Challenges of Internet Streaming

- Why does bandwidth fluctuate on the public internet
- How do multiple renditions let a player adapt to it

HTTP Streaming

- How is a stream split into independent, decodable segments
- Why are they delivered as HTTP file transfers rather than a continuous flow
- Why is HTTP the chosen protocol
- How does HTTP/TCP delivery differ from traditional UDP streaming

The OTT Manifest

- Why is an index file needed
- What must it tell the player about variants and segment locations

OTT Delivery Chain: Components & Roles

- What does each element (production, transcoder, packager/origin server, CDN, player) contribute
- How do fragment/segment durations shape the workflow

Birth of Adaptive Streaming

- What limitations of RTMP/RTSP pushed the industry toward HTTP
- How did CDNs let video scale like web files
- Why didn't adaptive streaming scale before the iPhone 3G era

Dominant OTT Protocols: HLS vs MPEG-DASH

- Brief overview of HLS and DASH
- How do they differ in manifest format, media container, and vendor support

Module 2 – Video Encoding for OTT

Video Encoding Basics

- What does video encoding trade off to cut bitrate so drastically
- Is it correct to call it “compression”

Frame Types: I, P, B and Stream Structure

- How many references (past/future) do each of Intra, Predicted, and Bi-predicted frames use
- What is encoding efficiency, and which frame type is most and least efficient
- What setting determines a real-world stream's complexity

Group of Pictures (GoP) and Configurations

- Which three streaming issues does GoP address
- Why it enables increased video quality in certain cases
- Example of variable GoP and open GoP, and when they are used
- Device interoperability vs. video quality
- Typical GoP configurations used in OTT and why

GoP Duration & Video Quality

- How does GoP duration affect video quality
- “Long GoP” vs. “I-frame only”: how they compare in bitrate needed for equal quality

Encoding vs. Display Order

- The encoding and decoding process
- Why would encoding/transmission order differ from display order
- How it may introduce latency

Encoding OTT Fragments & IDR

- What marks an OTT fragment's start and end
- Definition of an IDR frame and why it was created
- Analysis of four real-world scenarios of OTT encoding parameters
- What makes an OTT encoder different from a typical IPTV encoder

Module 3 — CMAF & Fragmented MP4

ISOBMFF History and High-Level Structure

- Where did MP4 originate, and how did it become an ISO standard
- How does ISOBMFF differ from MPEG-TS, and why separate metadata from media data

Non-Fragmented MP4

- Understanding the MOOV and MDAT boxes
- Why is classic MP4 unfit for adaptive streaming

MOOV Box Deep Dive

- What do MVHD, STSD, STSZ, and CTTS describe
- How does timescale affect duration

Fragmented MP4 (MOOF & MDAT)

- How does fragmenting solve live-stream indexing
- What role does MOOF/MDAT fragmenting play in adaptive streaming

MOOF Box Deep Dive

- How does TRUN describe frame timing and byte offsets
- What is the composition time offset for, illustrated with an example of two consecutive B-frames

OTT Media Init Segment & Streaming with Fragmented MP4

- What does FTYP declare about a file: brands vs. compatible brands, and examples
- Why does OTT streaming need a separate init segment
- How does fragmented MP4 enable bitrate switching

Fragment Timing & Ordering

- What do sequence number and base media decode time ensure
- Understanding the computation of these timestamps
- How is audio/video sync kept across variants

CMAF Features

- What problem did CMAF solve for OTT delivery
- How does CMAF relate to ISOBMFF
- CMAF-specific features

DAY 2 — PROTOCOLS IN PRACTICE

Module 4 — HTTP LIVE STREAMING with CMAF

Understanding What HLS Is

- Playlist versus media container: what is the difference
- Why is HLS the most widely used protocol

Master and Variant Playlist Hierarchy

- How are master and variant playlists related, and what does each provide to the player

Master Playlist Anatomy

- Core tags in a master playlist
- Video variant definition
- The average bandwidth attribute
- How is the bandwidth ladder structured
- Bandwidth vs. throughput, and how does a player compute throughput
- Audio rendition groups: what they are and why they were invented
- How do default and auto-select affect track selection

Variant Playlist Structure

- Absolute URL vs. relative path URL, and how relative paths save playlist size
- How to reconstruct an absolute URL from relative paths
- Version, target duration, and segment duration
- What does live window size control

Timing & Sequencing Attributes

- How is a segment's wall-clock time computed
- What role does the sequence number play

Adaptive Bitrate Switching

- How should a player pick its starting variant
- How does throughput drive an upgrade or downgrade

Module 5 — MPEG-DASH with CMAF

HLS vs MPEG DASH

- Why is DASH considered more efficient than HLS
- What separates DASH implicit addressing from HLS explicit addressing

MPD Structure and Representation

- How is an MPD organized into Period, Adaptation Set, Representation
- What attributes define a Representation, and why players need them before downloading media

Segment Templates and Path Navigation

- How does a Segment Template generate init and media segment URLs
- What do timescale and duration values represent
- How does relative path navigation reduce the size of an MPD

The Adaptation Set

- What must representations share to belong to the same Adaptation Set

Synchronization and Time Addressing

- How do players and packagers agree on a common origin of time
- What role does availabilityStartTime play in computing segment start times
- How is a segment's absolute start time derived from Time addressing

Time Shift Buffer Depth

- What does timeShiftBufferDepth define and limit for a player

Variable Segment Durations

- Why do segment durations vary, and how does a Segment Timeline express this
- Why do audio and video segments not always share the same duration

Number Segment Addressing

- How does Number addressing differ from Time addressing

Video Stream stack