Low Encoding Overhead Ultra Low Latency Streaming via HESP Through Sparse Initialization Streams

Pieter-Jan Speelmans
CTO

MHV'22, Mar. 1-3, 2022, Denver, CO, USA

Background on the High Efficiency Streaming Protocol

Sub-second latency

- ➔ Near real-time interactivity
- ➔ In-sync delivery across devices
- ➔ Interactive user experiences

Up to 20% reduction in bandwidth & costs

- ➔ Cost savings for content distributors
- ➔ Re-invest bandwidth savings in higher quality media delivery

Scales over existing infrastructure

- ➔ Efficient delivery over existing HTTP infrastructure & CDNs
- ➔ CMAF compatible with standard encoders
- ➔ Cross platform playback

Enhanced viewer experience

- ➔ Instant zapping and seeking times
- ➔ Full adaptive bitrate to respond to varying network conditions

MINIMALISTIC MANIFEST

- Manifest is not needed to start playing the stream.
- Low frequency updates

HTTP DELIVERY

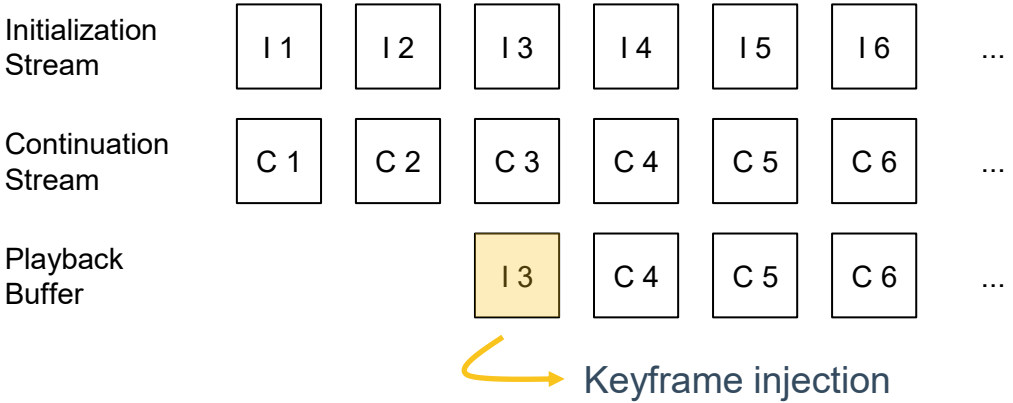
- HTTP/1.1 based (HTTP/1.1 CTE & Open-Ended Range Requests)
- HTTP/2 & HTTP/3 (QUIC) frame based streaming

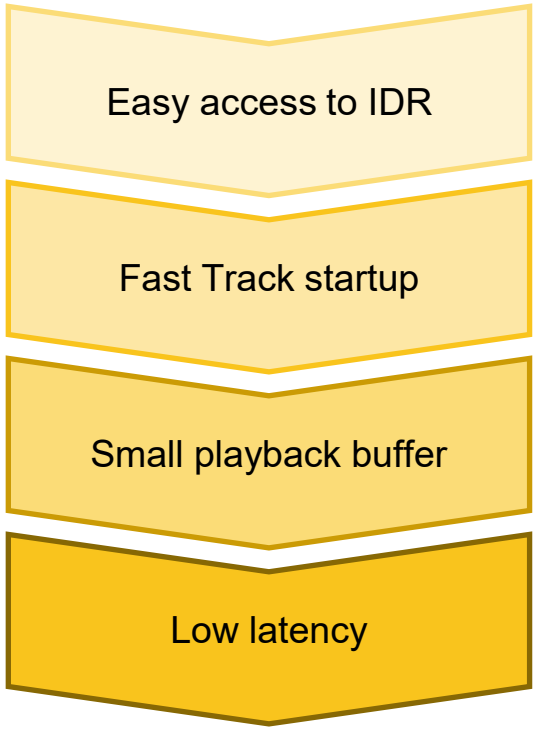
TWO COMPLEMENTARY STREAMS

- Initialization Stream: can request images at any moment to start playback
- Continuation Stream: can continue playback after an Initialization Stream image

Specification https://www.ietf.org/id/draft-theo-hesp-01.html

HESP Initialization Streams





Sparse Initialization Streams

> *A Sparse Initialization Stream is an Initialization Stream which is generated at a frame rate lower than the frame rate of the Continuation Stream.*

Sparse Initialization Streams

Track id	Bitrate	Resolution	Frame Rate
1080p	4000 kbps	1920 x 1080	30 fps
720p	2500 kbps	1280 x 720	30 fps
540p	1000 kbps	960 x 540	30 fps
360p	400 kbps	640 x 360	30 fps

4 Tracks (bitrates)

Figure 2: **Overview of Continuation Stream encoding profiles**

Scenario	Initialization Stream profile			
	1080p	720p	540p	360p
1	-	-	-	-
2	30 fps	30 fps	30 fps	30 fps
3	-	30 fps	30 fps	30 fps
4	-	-	30 fps	30 fps
5	-	-	-	30 fps
6	10 fps	10 fps	10 fps	30 fps
7	5 fps	5 fps	5 fps	30 fps
8	2 fps	2 fps	2 fps	30 fps
9	1 fps	1 fps	1 fps	30 fps
10	1 fps	30 fps	1 fps	30 fps
11	1 fps	1 fps	30 fps	30 fps

DASH reference

Full HESP stream

Reusing key frames

Reduced frame rate

Mixed configurations

Figure 3: **Overview of test scenarios and their Initialization Stream profiles**

Impact on encoding

Scenario	Initialization Stream profile			
	1080p	720p	540p	360p
1	-	-	-	-
2	30 fps	30 fps	30 fps	30 fps
3	-	30 fps	30 fps	30 fps
4	-	-	30 fps	30 fps
5	-	-	-	30 fps
6	10 fps	10 fps	10 fps	30 fps
7	5 fps	5 fps	5 fps	30 fps
8	2 fps	2 fps	2 fps	30 fps
9	1 fps	1 fps	1 fps	30 fps
10	1 fps	30 fps	1 fps	30 fps
11	1 fps	1 fps	30 fps	30 fps

Figure 3: **Overview of test scenarios and their Initialization Stream profiles**

Scenario	Instructions / s	Increase vs Scenario 1	Decrease vs Scenario 2	
1	7,922,246,220	-	40,8 %	Reference
2	13,380,214,610	+68,9 %	-	
3	10,357,872,944	+30,7 %	22,6 %	Top quality
4	9,126,113,383	+15,2 %	31,8 %	
5	8,164,052,126	+3,1 %	39,0 %	
6	10,124,430,951	+27,8 %	24,3 %	10fps
7	9,126,385,922	+15,2 %	31,8 %	
8	8,403,728,247	+6,1 %	37,2 %	2fps
9	8,197,873,181	+3,5 %	38,7 %	
10	9,775,632,742	+23,4 %	26,9 %	
11	9,093,934,438	+14,8 %	32,0 %	

Figure 4: **Average number of instructions per second.**

Impact on startup time

Worst case impact:

- Either additional latency at startup
- Or additional startup time

How much?

Average: $\frac{(\frac{1}{f_i} - \frac{1}{f_{nominal}})}{2}$

Worst case: $(\frac{1}{f_i} - \frac{1}{f_{nominal}})$

Scenario	Average start-up time (s) Maximum start-up time (s)					
	1080p	720p	540p	360p		
1	1.452 2.854	1.407 2.633	1.459 2.810	1.384 3.003		Reference
2	0.376 0.539	0.389 0.532	0.386 0.523	0.360 0.515		
3	1.387 2.704	0.383 0.534	0.386 0.489	0.375 0.558		Top quality (+30.7%)
4	1.389 2.842	1.399 2.844	1.346 2.669	0.381 0.519		
5	1.401 2.473	1.419 2.377	1.412 2.256	0.373 0.453		
6	0.435 0.655	0.435 0.544	0.412 0.553	0.386 0.461	+33ms +67ms	10fps (+27.8%)
7	0.500 0.696	0.454 0.658	0.469 0.735	0.392 0.559		
8	0.641 1.142	0.607 0.989	0.633 1.119	0.368 0.495	+233ms +467ms	2fps (+6.1%)
9	0.883 1.519	0.855 1.702	0.845 1.673	0.368 0.491		
10	0.842 1.587	0.359 0.546	0.868 1.500	0.384 0.566		
11	0.894 1.689	0.900 1.464	0.371 0.447	0.378 0.479		

Figure 6: **Average and maximum start-up time results**

Impact on ABR switching speed?

Same as startup time

- Not an issue for switching up
  - Delay in higher quality should not have a significant impact
- BIG issue for switching down:
  - Must have enough buffer to: worst case scenario

Must switch to lower quality with Full Initialization Stream to guarantee latency

Scenario	Average start-up time (s) Maximum start-up time (s)					
	1080p	720p	540p	360p		
1	1.452 2.854	1.407 2.633	1.459 2.810	1.384 3.003		Reference
2	0.376 0.539	0.389 0.532	0.386 0.523	0.360 0.515		
3	1.387 2.704	0.383 0.534	0.386 0.489	0.375 0.558	+ 983ms +1967ms	Top quality (+30.7%)
4	1.389 2.842	1.399 2.844	1.346 2.669	0.381 0.519		
5	1.401 2.473	1.419 2.377	1.412 2.256	0.373 0.453		
6	0.435 0.655	0.435 0.544	0.412 0.553	0.386 0.461	+33ms +67ms	10fps (+27.8%)
7	0.500 0.696	0.454 0.658	0.469 0.735	0.392 0.559		
8	0.641 1.142	0.607 0.989	0.633 1.119	0.368 0.495	+233ms +467ms	2fps (+6.1%)
9	0.883 1.519	0.855 1.702	0.845 1.673	0.368 0.491		
10	0.842 1.587	0.359 0.546	0.868 1.500	0.384 0.566	+483ms +967ms	(+23.4%)
11	0.894 1.689	0.900 1.464	0.371 0.447	0.378 0.479	+483ms +967ms	(+14.8%)

Figure 6: **Average and maximum start-up time results**

Conclusions

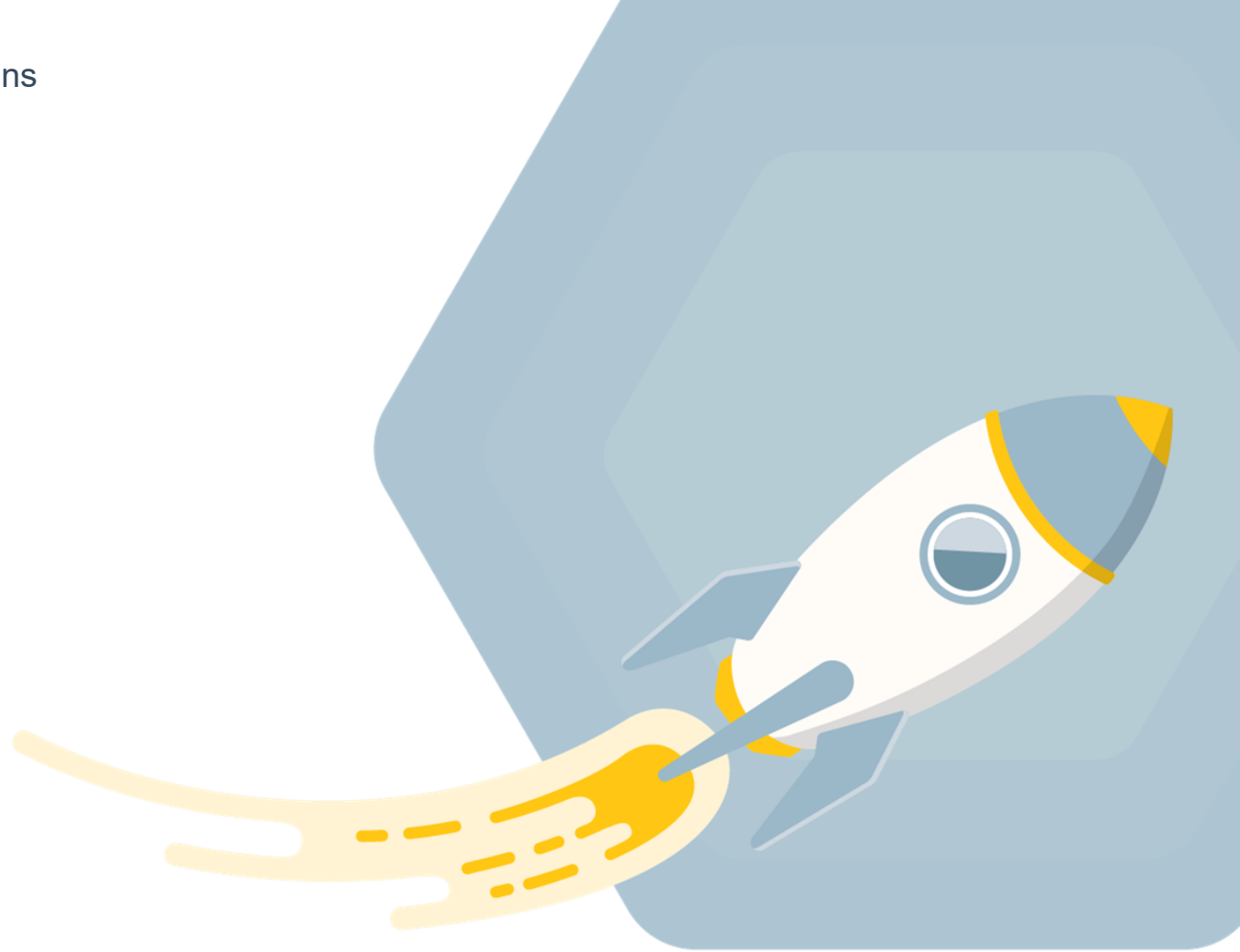
1. Allows for encoding cost close to MPEG-DASH and HLS

2. Minimum buffer size can remain the same: latency is not impacted

3. Small to no impact on QoE

Pieter-Jan Speelmans
ps@theoplayer.com

Questions?



HESP ALLIANCE

THEO