SVT-AV1 Bitrate Control Modes

Jan Ozer
Streaming Learning Center
jan.ozer@streaminglearningcenter.com

Agenda

- Overview
- The command strings
- Performance
- Test Files
  - *Easy Hard*
  - *Test*
  - *Football*
  - *Meridian*
- Conclusions

Overview

There are four bitrate control techniques worth considering for live encoding/ transcoding using version 1.7 of the SVT-AV1 codec. These are:

- Capped CRF
- VBR
- Capped VBR
- Constrained VBR

With capped CRF, you supply a CRF value (42) to set quality, and a maximum bitrate (mbr=4500) to set the cap. For more background on capped CRF, see here.

With VBR, you choose VBR rate control (rc=1) and set a target bitrate (tbr=4500). For more on VBR, see here.

With capped VBR, you choose VBR rate control (rc=1) and set a target (tbr=4500) and a maximum bitrate (mbr=4501). The maximum bitrate is set at 4501 here to produce a relatively consistent stream (like CBR) for live transcoding.

Quoting one of the SVT-AV1 developers, constrained VBR is a new technique designed to “keep the same rate for each gop. You can think of it as a CBR per gop, not on a sliding window.” You set the rate control as VBR (rc=1), the target (tbr=4500), and then include the gop-constraint-rc=1 switch.

According to the same developer, “CBR is only implemented for the low delay use case (video conferencing, very low latency live use cases), where the other modes are deployed for other use cases such as broadcasting and live streaming where latency isn't an issue.” My tests revealed CBR to be a poor choice for normal broadcasting, so I didn’t include CBR in these tests.

Capped CRF

```
ffmpeg -y -i test.mp4 -c:v libsvtav1  -g 120 -preset 8 -crf 42 -svtav1-params mbr=4500  test_capped_CRF.mp4
```

VBR

```
ffmpeg -y -i test.mp4 -c:v libsvtav1  -g 120 -preset 8  -svtav1-params rc=1:tbr=4500:enable-force-key-frames=0 test_VBR.mp4
```

Capped VBR

```
ffmpeg -y -i test.mp4 -c:v libsvtav1  -g 120 -preset 8  -svtav1-params rc=1:tbr=4500:mbr=4501:enable-force-key-frames=0 test_capped_VBR.mp4
```

Constrained VBR

```
ffmpeg -y -i test.mp4 -c:v libsvtav1  -g 120 -preset 8  -svtav1-params rc=1:tbr=4500:enable-force-key-frames=0:gop-constraint-rc=1 test_contrained_VBR.mp4
```

Key Findings - Overall

Test procedure: I tested four 1080p30 files with summary results shown on the right. This was for the live use case using preset 8 (VOD results using a higher quality preset may be different).

Overall:

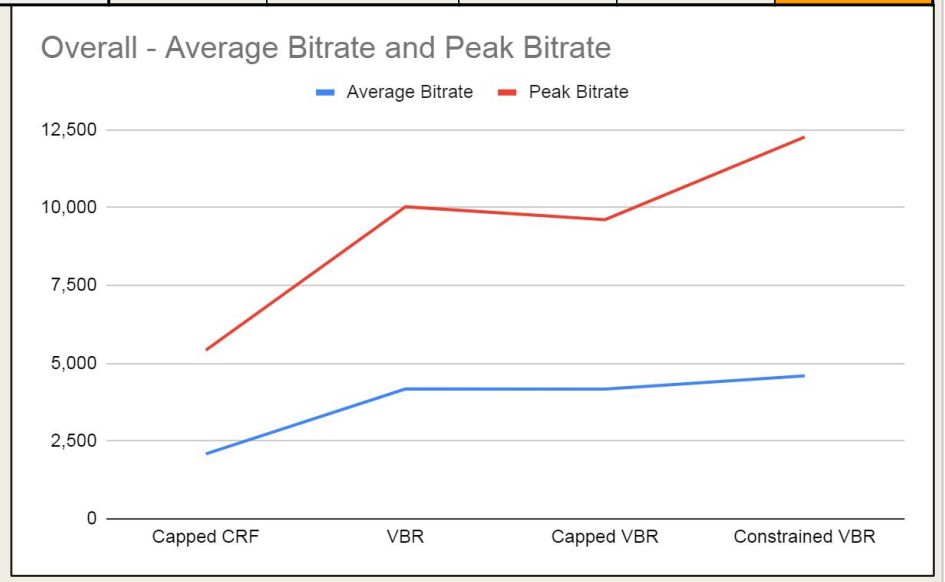
- Capped CRF shows promise with significant bitrate savings, good quality retention, and the best overall performance by ~10 - 25% (meaning more streams from the same hardware).
- Capped VBR was very similar to VBR with no significant advantages (again, in the live use case).
- Constrained VBR produced a more consistent bitrate, but had:
  - The highest peak bitrates (potential deliverability issues).
  - The lowest average VMAF score of the three VBR variants.
  - The lowest low-frame quality scores, indicating the potential for transient quality issues.

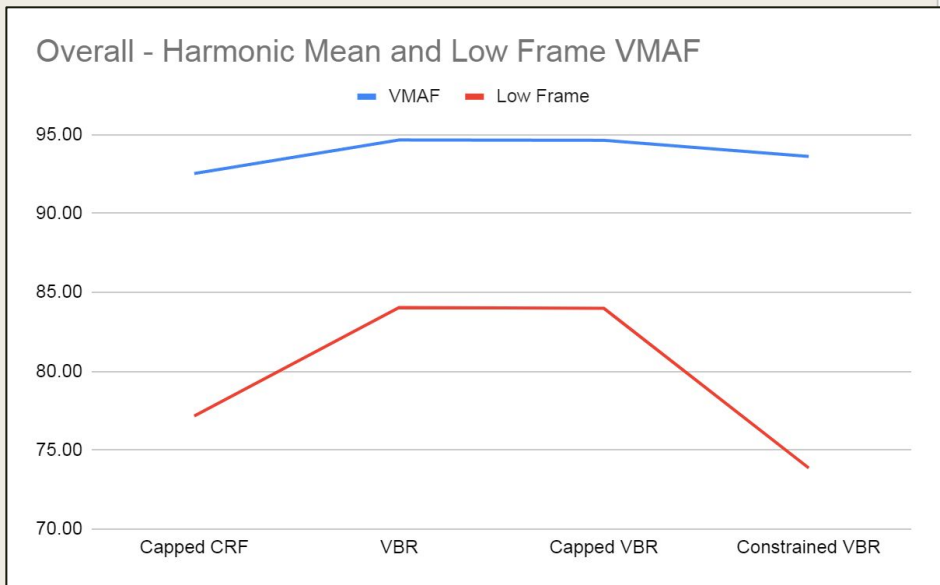
Bottom line: Either use VBR or capped CRF.

Final thoughts: The bitrate savings delivered by capped CRF relates directly to the VBR bitrate. If I tested VBR at 3 Mbps, the savings would have been much lower. That said, 4.5 Mbps for 1080p30 video files, particularly sports, isn't that conservative.

In addition, for the record, capped CRF doesn't deliver "better compression" than any of the VBR variants. It's a form of per-title encoding that adopts the bitrate to the content. If you encoded using VBR at the same bitrate as capped CRF, the quality would be similar. Of course, in a live setting, you don't have that luxury.

Overall	Average Bitrate	Capped CRF Savings	Peak Bitrate	VMAF	Low Frame
Capped CRF	2,089		5,421	92.54	77.17
VBR	4,174	49.96%	10,029	94.66	84.05
Capped VBR	4,169	49.90%	9,616	94.64	83.99
Constrained VBR	4,594	54.54%	12,272	93.62	73.87





Online Course for New Streaming Professionals

Streaming Media 101:
Technical Onboarding for Streaming Media Professionals

Learn the technologies, techniques, and skills to succeed in a streaming media-related role, whether it's producing and distributing streaming video or creating the tools and services necessary to do so.

https://bit.ly/StreamingMedia101

What You Will Learn

In about 11 hours, this online course will teach you the terms, technologies, best practices, and skills needed to excel in a technical role in the streaming media industry. You will learn:

- How to encode and deploy streaming video using the H.264, HEVC, VP9, and AV1 codecs
- How to encode for single file and adaptive bitrate encoding and packaging for HLS, DASH, and CMAF
- About digital rights management (DRM) and distribution issues like choosing a CDN and how to measure and ensure Quality of Service and Quality of Experience
- Critical production-level decisions, like whether to encode on-premise or in the cloud, how to choose a per-title encoding technology and cloud encoder, and how to compute the breakeven on deploying an advanced codec like HEVC or AV1

You will learn to:

- Analyze files with MediaInfo, Bitrate Viewer, Apple's AVQT, and the Moscow State University Video Quality Measurement Tool
- Encode in FFmpeg and Handbrake
- Produce mezzanine files for upload to a streaming service
- Connect to YouTube Live and Facebook Live
- Embed a live or on-demand video into a web page

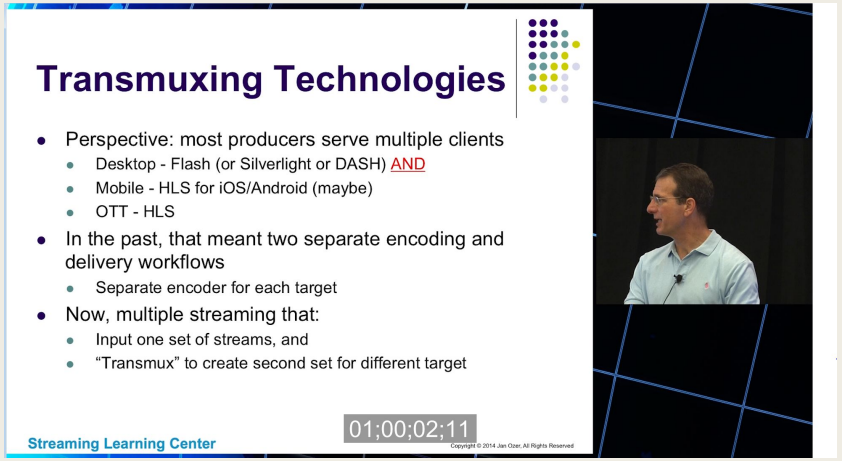
Test Files

- EasyHard
- Test
- Football
- Meridian

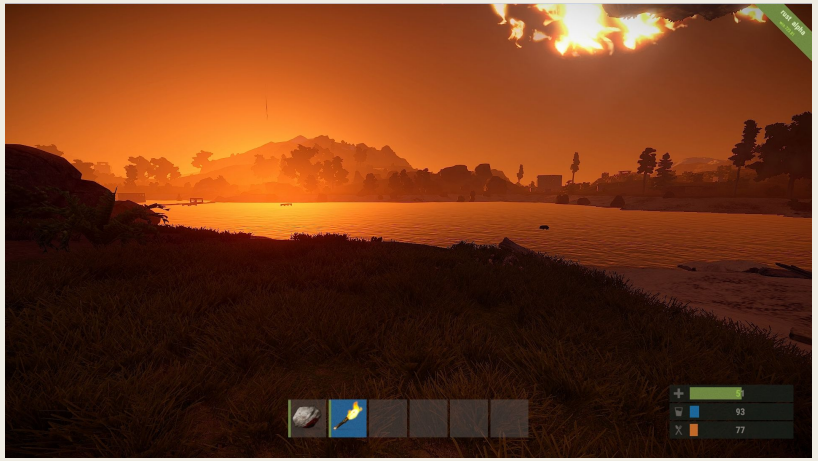
Easy Hard Clip

- Content description
- Bitrate/quality results
- Bitrate visualizations
- Quality visualization

Test 1 - Easy Hard



6 seconds very easy



9 seconds ridiculously hard

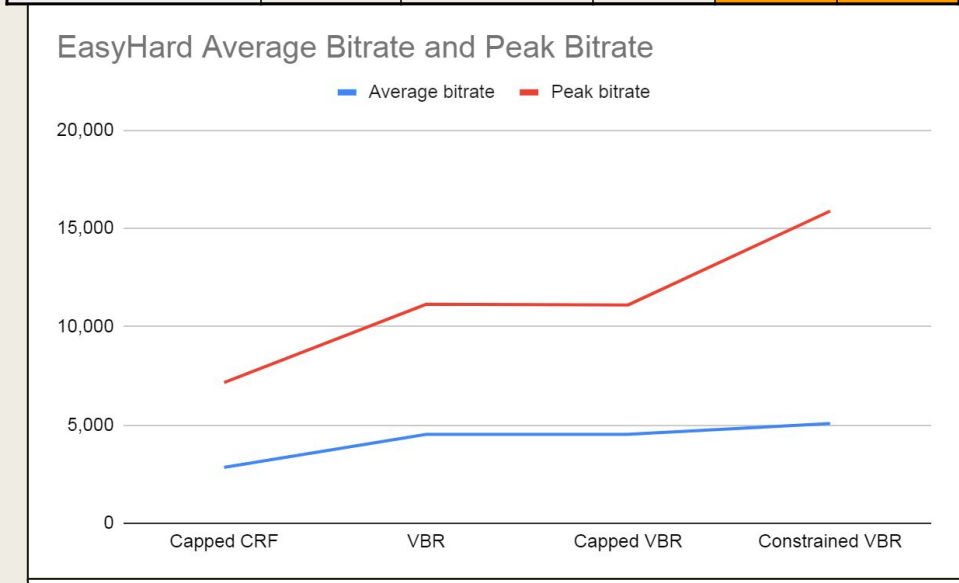
This file is useful for testing bitrate control mechanisms because the content is so variable between the two components.

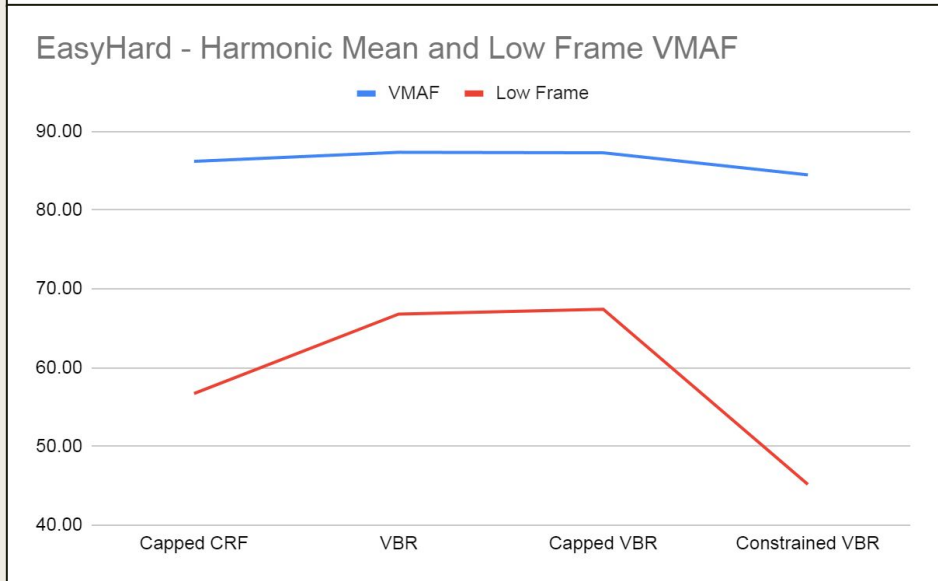
Data - Easy Hard

Observations:

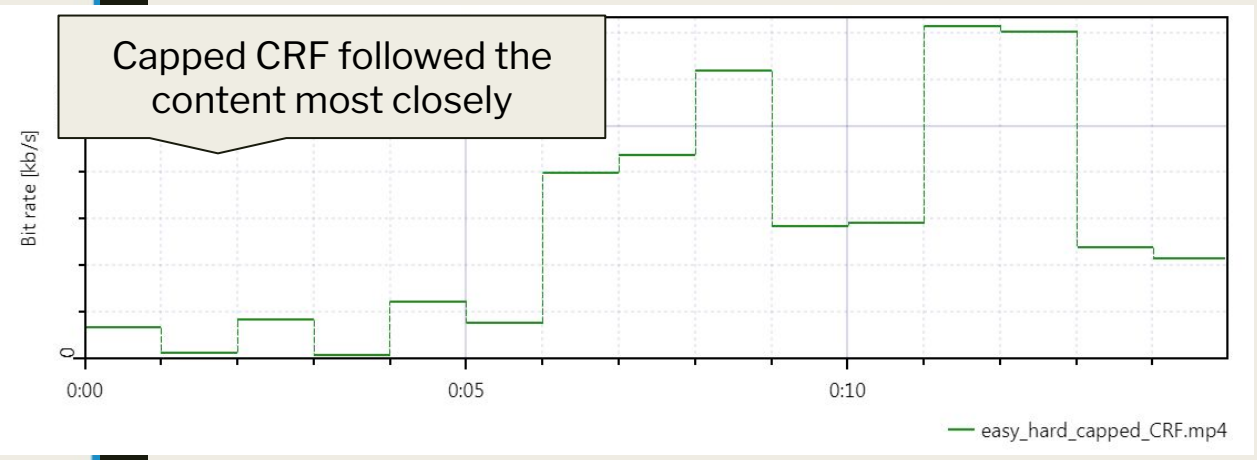
- **Capped CRF** - Followed content most closely producing significant savings with minimum quality delta
- **VBR** - good overall performer
- **Capped CBR** - ditto - nothing to distinguish from VBR
- **Constrained VBR**
  - Highest average bitrate
  - Highest peak bitrate
  - Lowest overall VMAF
  - Lowest low-frame VMAF
- Let's look at the graphs and see why

EasyHard	Average bitrate	Capped CRF Savings	Peak bitrate	VMAF	Low Frame
Capped CRF	2,844		7,156	86.20	56.71
VBR	4,523	37.12%	11,142	87.35	66.81
Capped VBR	4,526	37.16%	11,109	87.29	67.41
Constrained VBR	5,074	43.95%	15,886	84.48	45.17

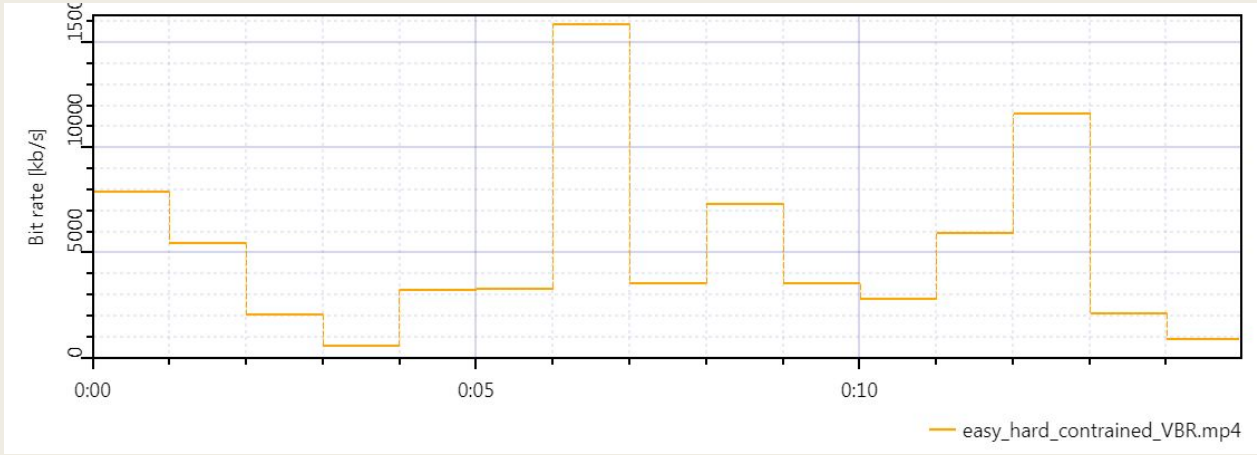




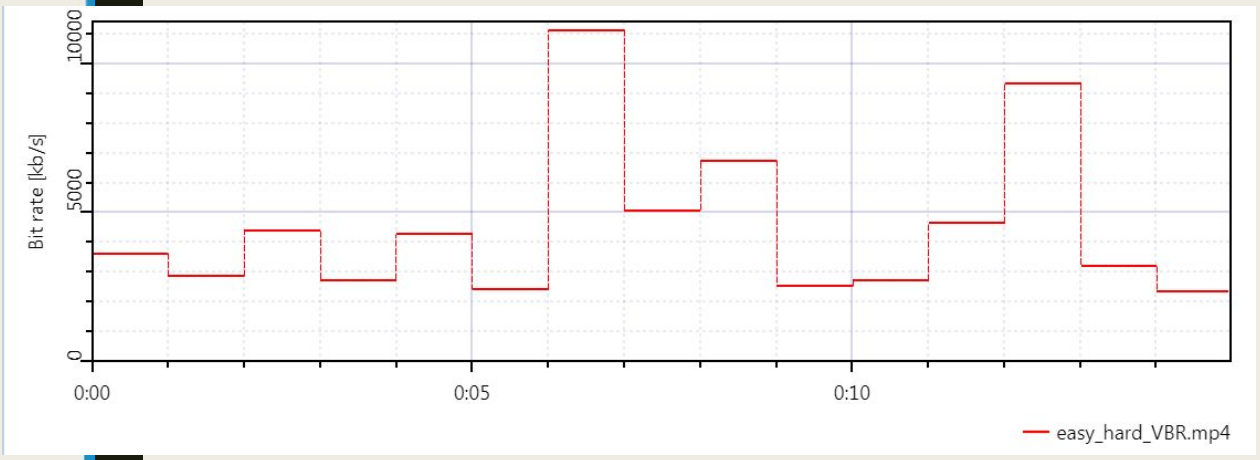
EasyHard - **All Bitrate Graphs**



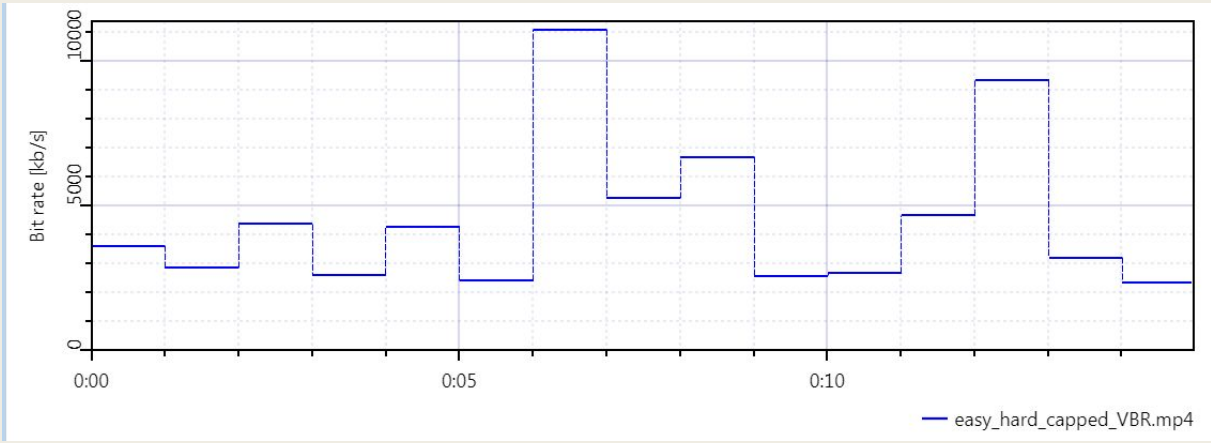
Capped CRF



Constrained VBR

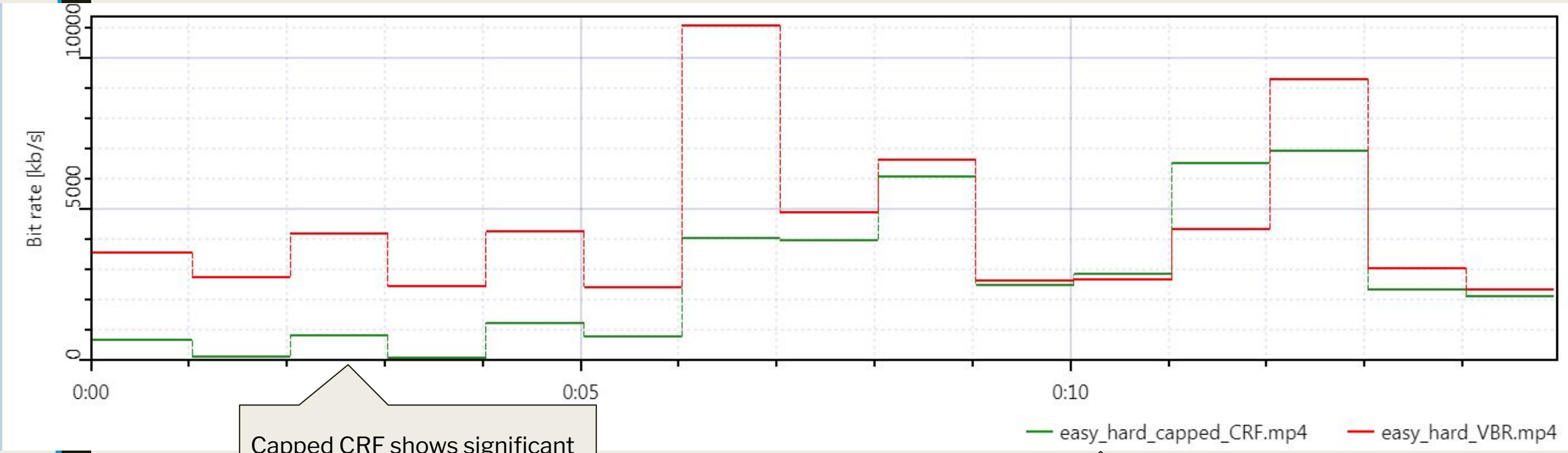


VBR



Capped VBR

EasyHard - **Capped CRF/VBR**

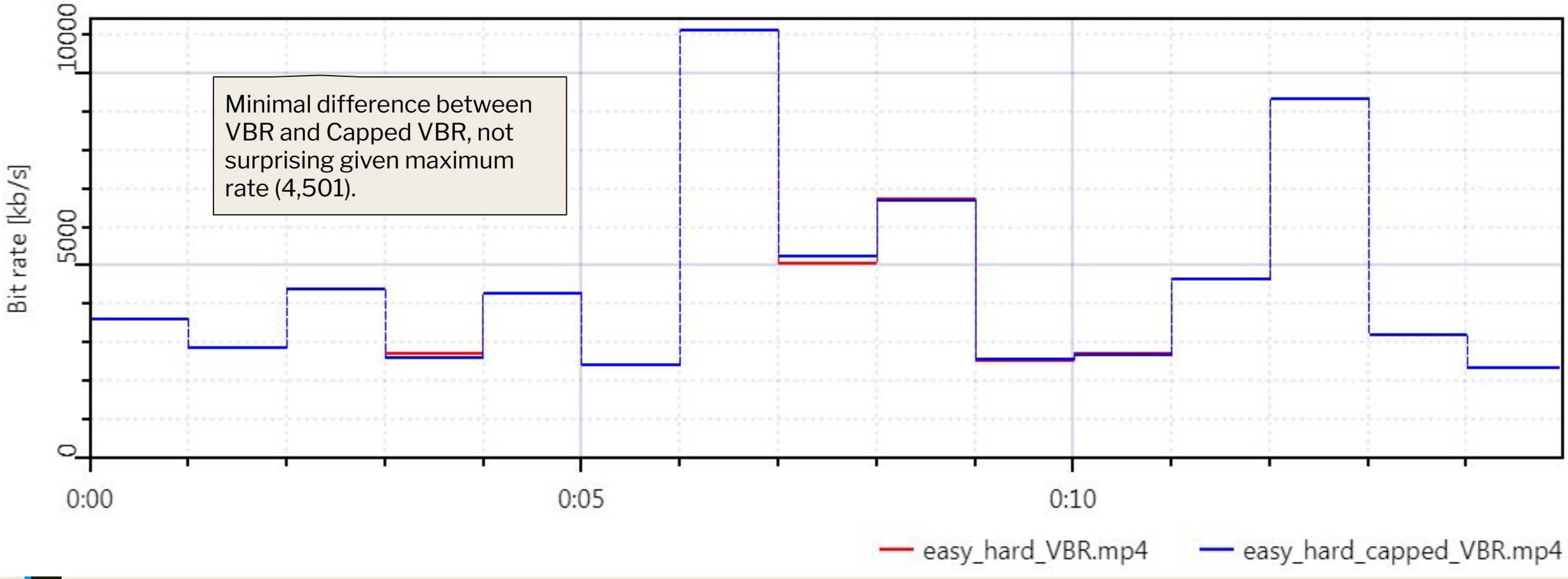


Capped CRF shows significant bitrate savings in the easy to encode region

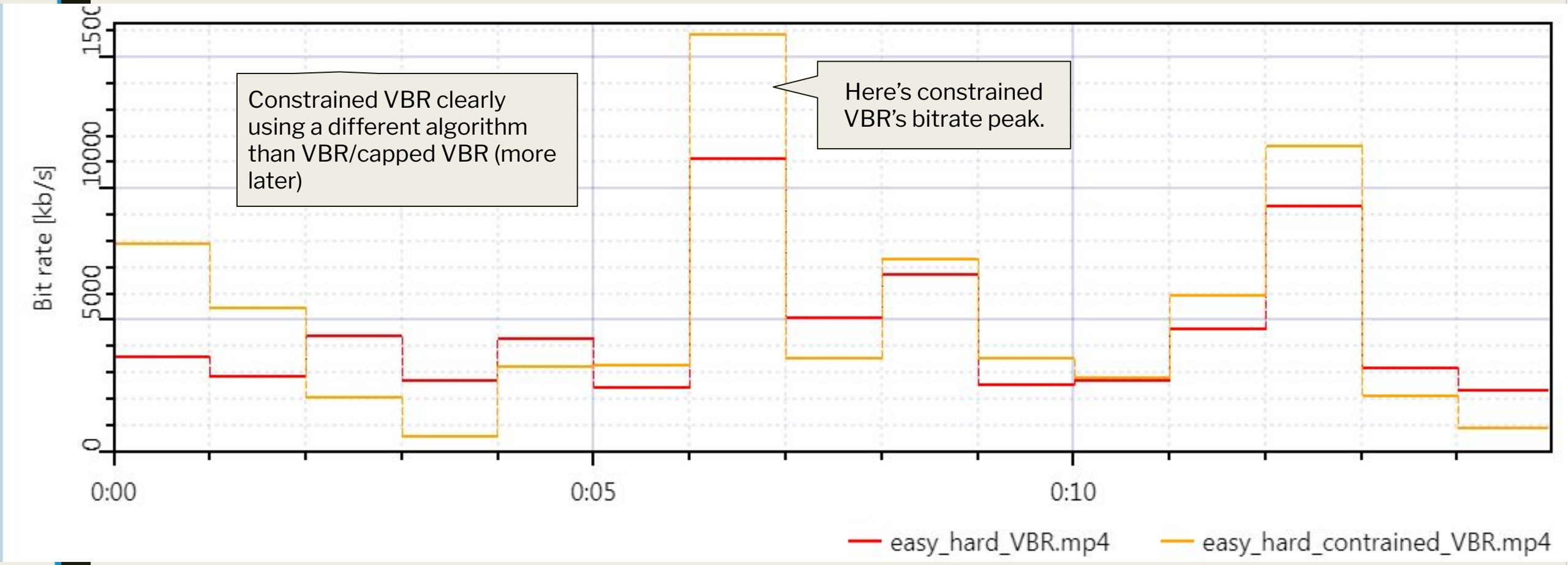
Capped CRF showed about the same bitrate in the hard to encode region as VBR. Better matching of bitrate and complexity than VBR.

Legend

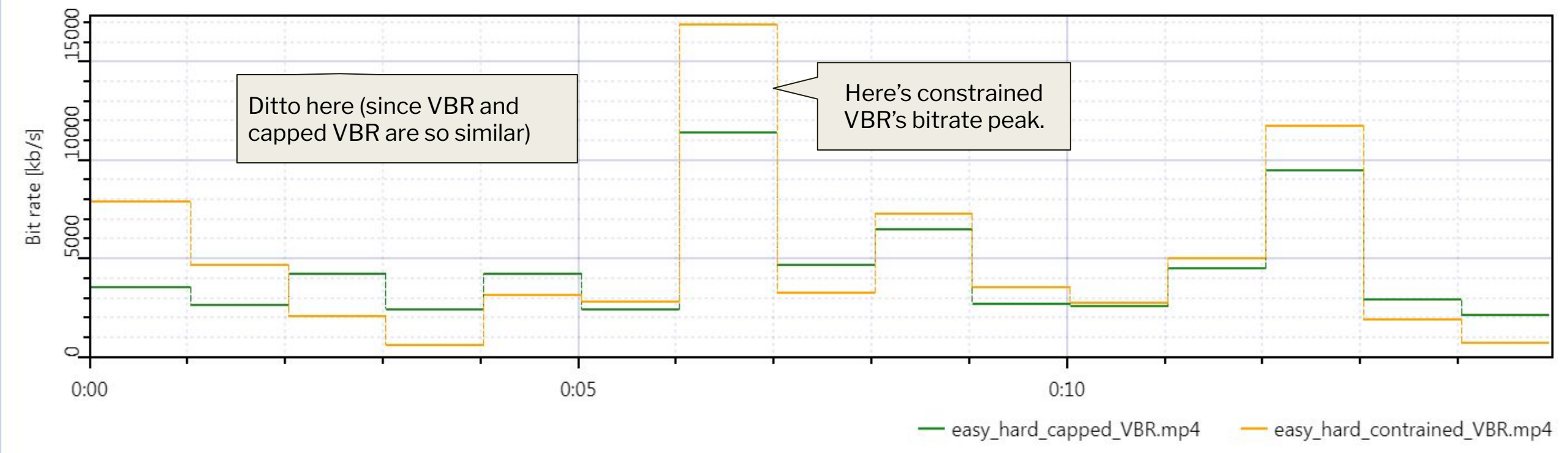
EasyHard - **VBR/Capped VBR**

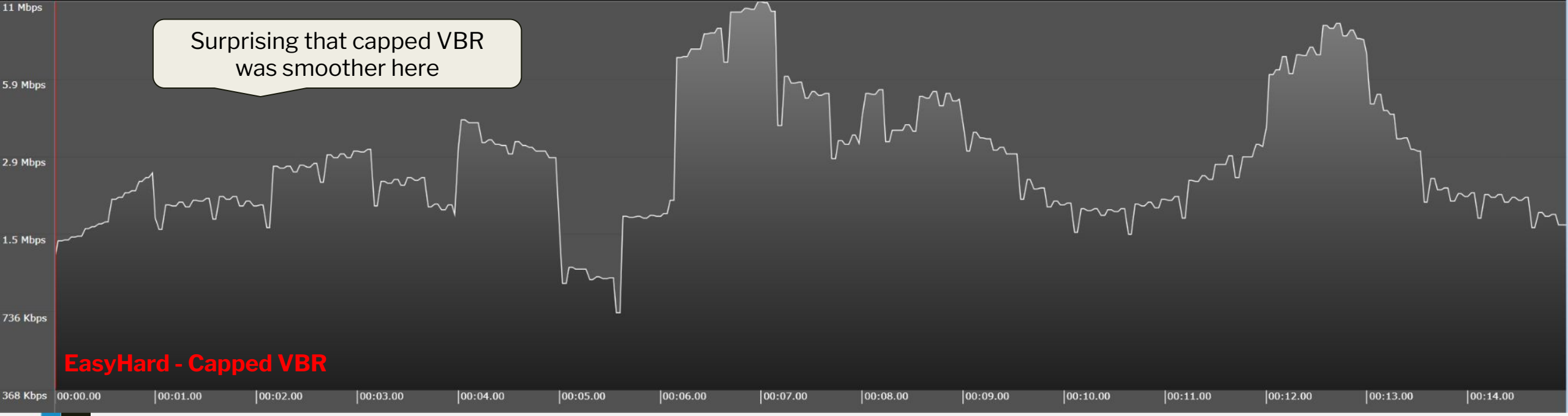


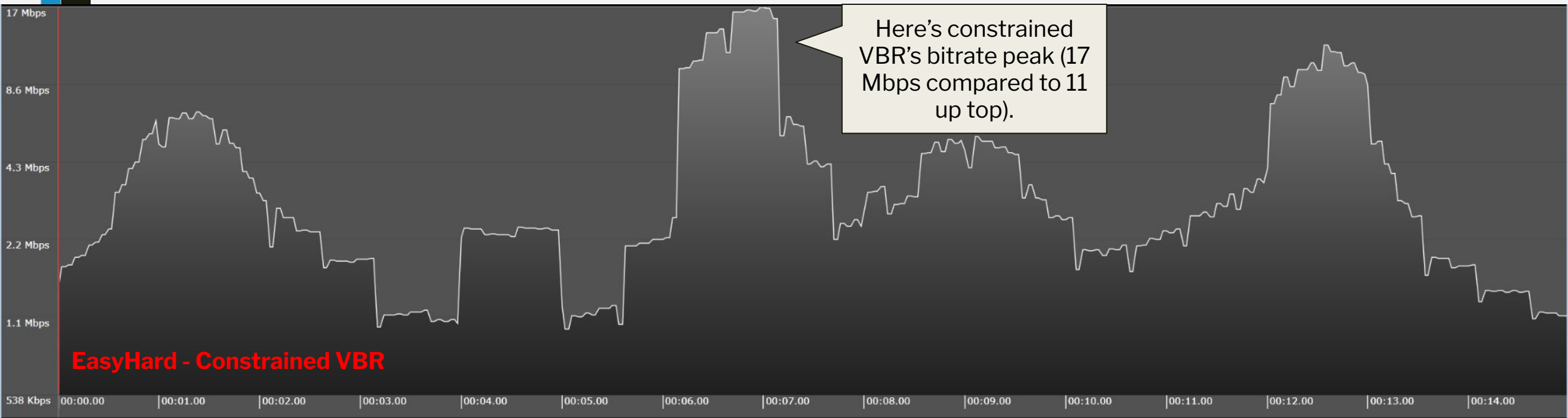
EasyHard - **VBR/Constrained VBR**



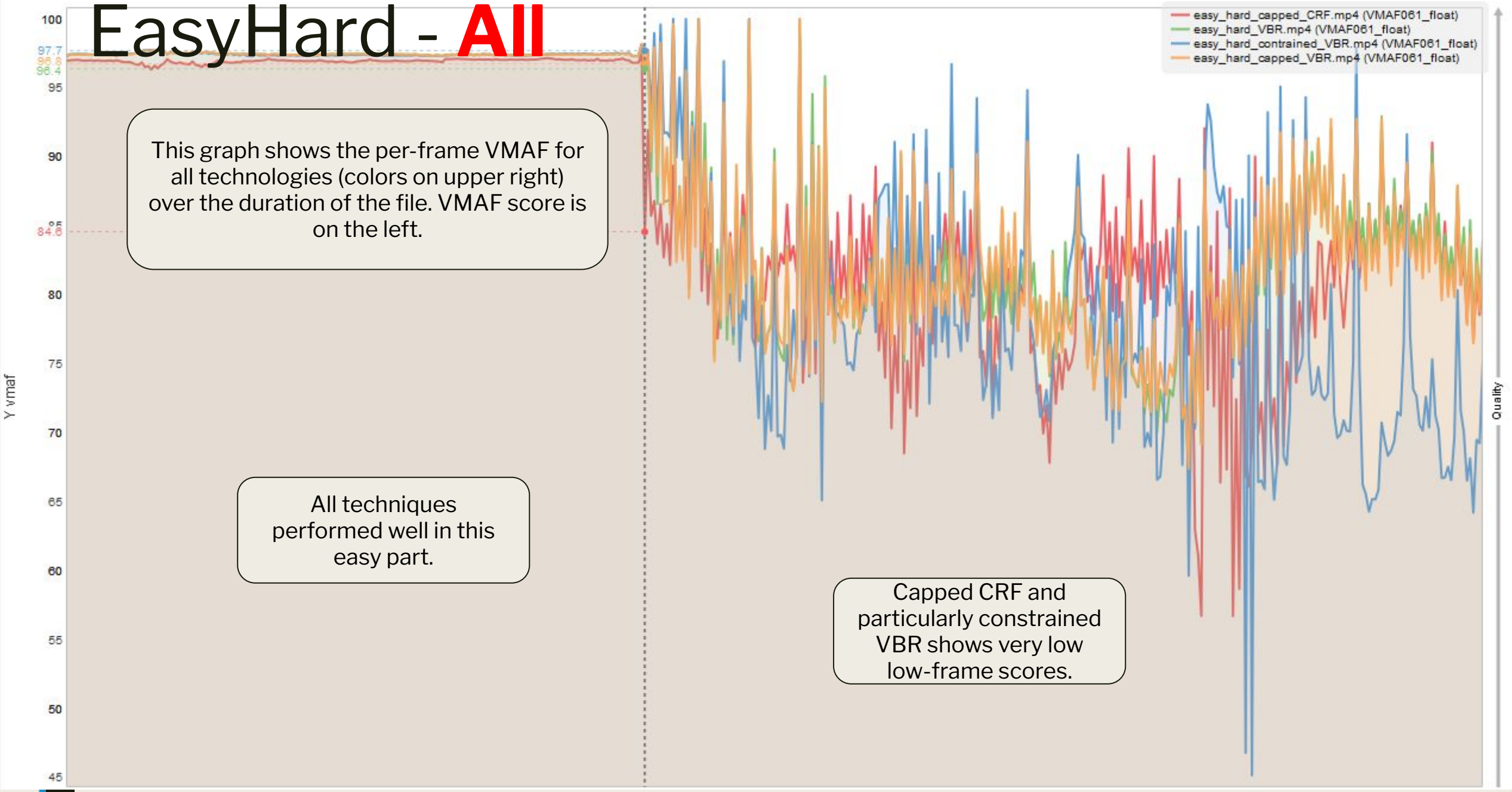
EasyHard - Capped VBR/Constrained VBR




11 Mbps
5.9 Mbps
2.9 Mbps
1.5 Mbps
736 Kbps
368 Kbps
Surprising that capped VBR was smoother here
EasyHard - Capped VBR
00:00.00
00:01.00
00:02.00
00:03.00
00:04.00
00:05.00
00:06.00
00:07.00
00:08.00
00:09.00
00:10.00
00:11.00
00:12.00
00:13.00
00:14.00


17 Mbps
8.6 Mbps
4.3 Mbps
2.2 Mbps
1.1 Mbps
538 Kbps
Here's constrained VBR's bitrate peak (17 Mbps compared to 11 up top).
EasyHard - Constrained VBR
00:00.00
00:01.00
00:02.00
00:03.00
00:04.00
00:05.00
00:06.00
00:07.00
00:08.00
00:09.00
00:10.00
00:11.00
00:12.00
00:13.00
00:14.00

EasyHard - All

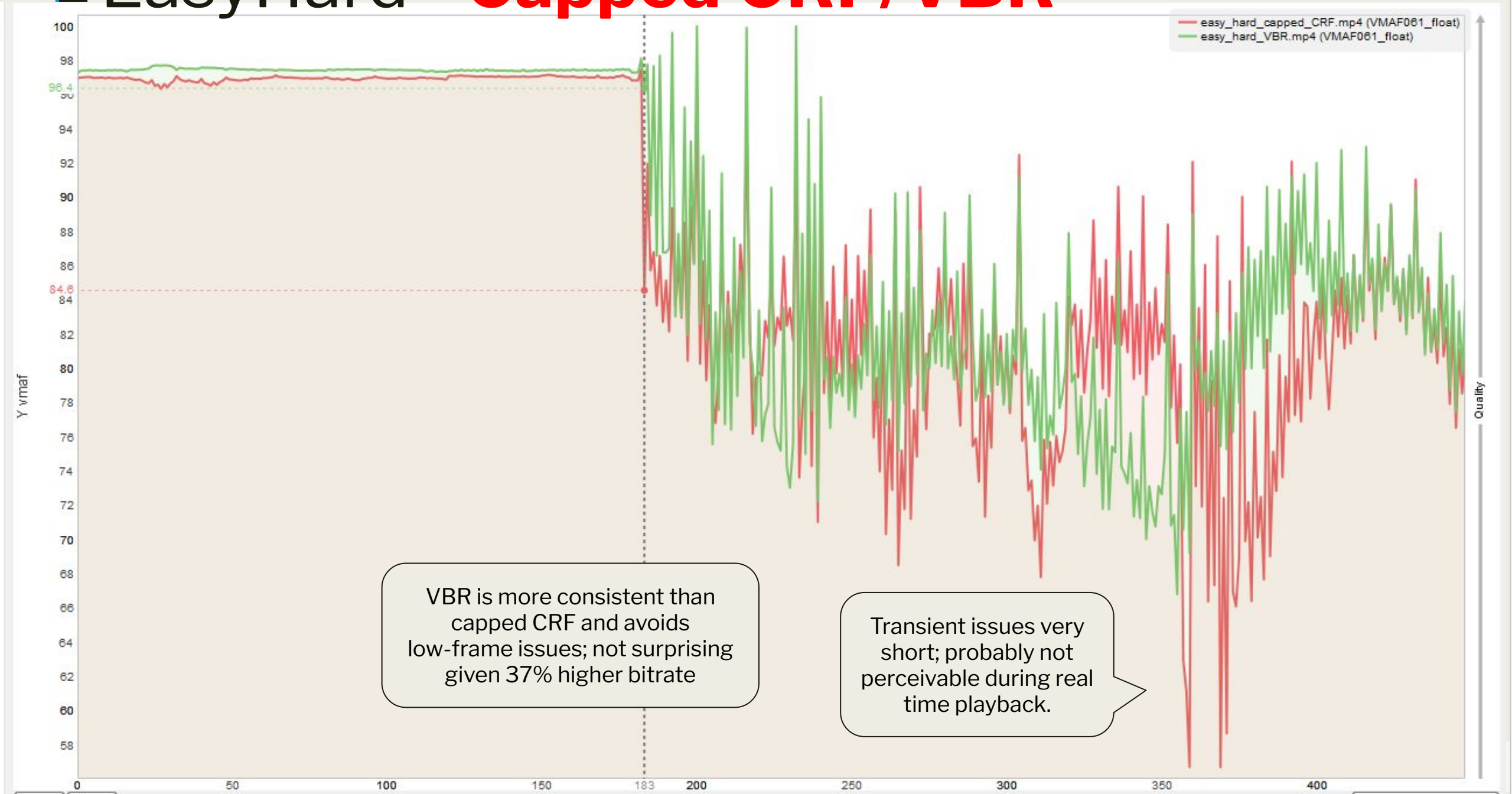


This graph shows the per-frame VMAF for all technologies (colors on upper right) over the duration of the file. VMAF score is on the left.

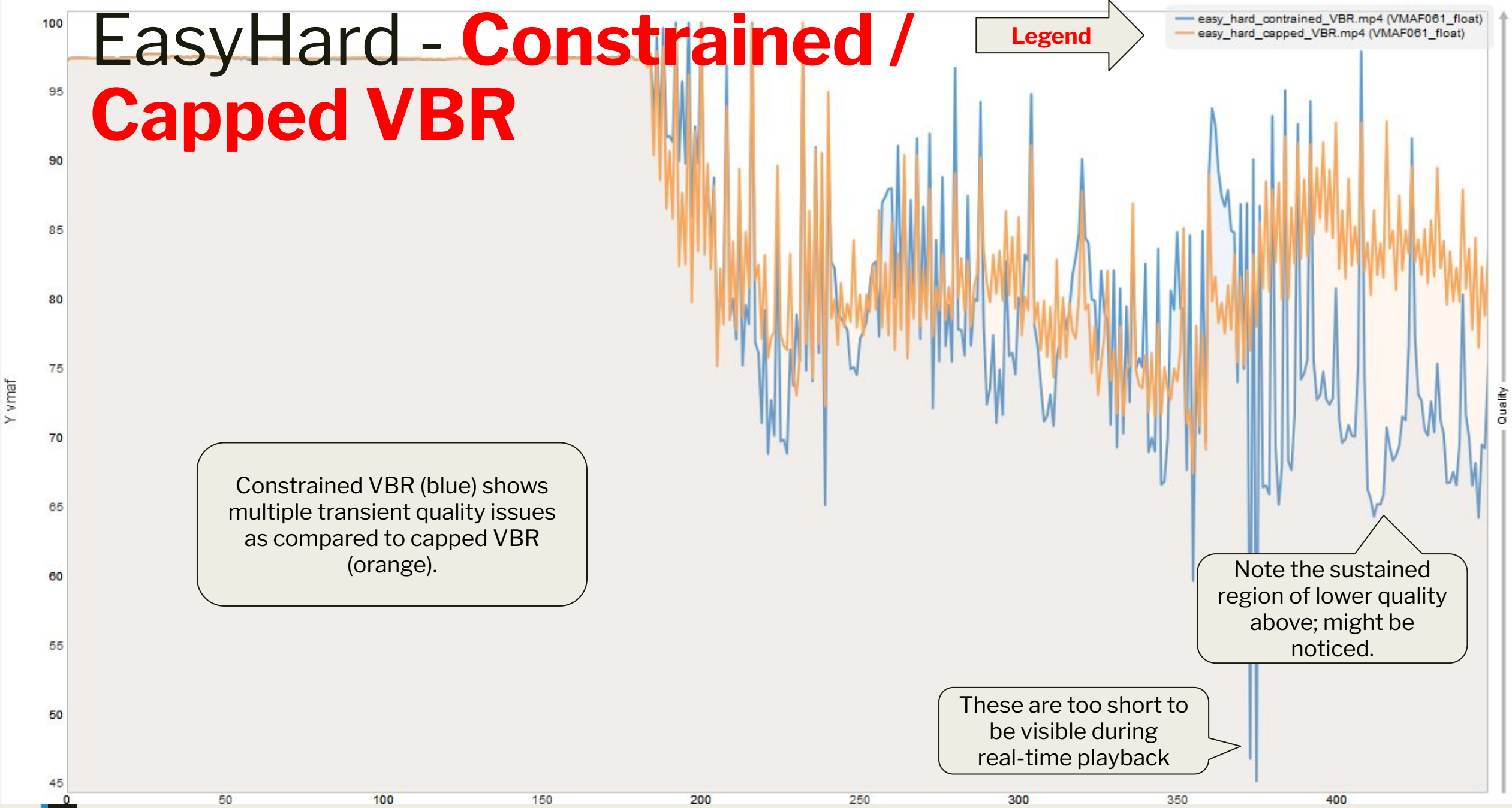
All techniques performed well in this easy part.

Capped CRF and particularly constrained VBR shows very low low-frame scores.

EasyHard - Capped CRF/VBR



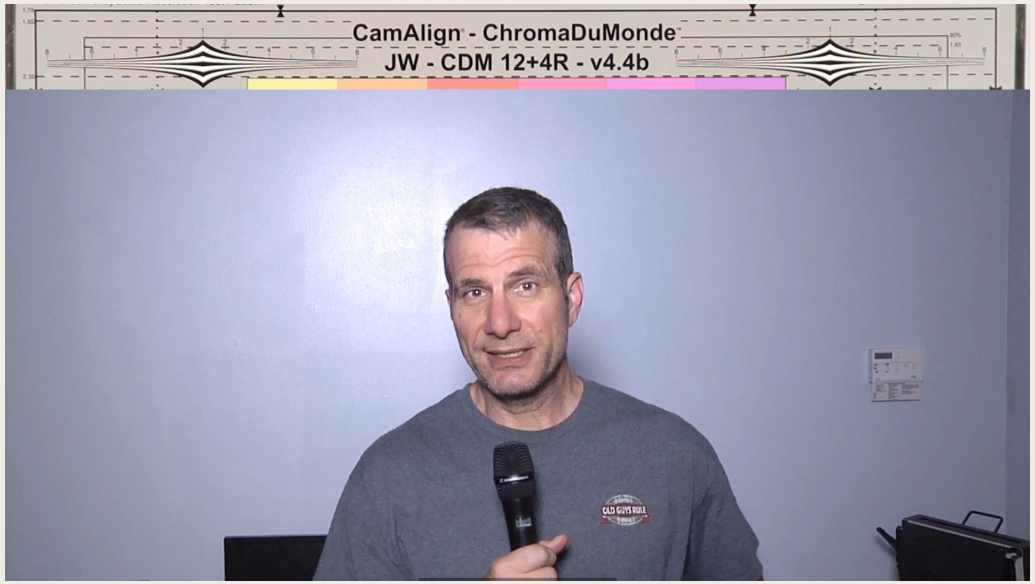
EasyHard - Constrained / Capped VBR



Test Clip

- Content description
- Bitrate/quality results
- Bitrate visualizations
- Quality visualization

Test



30 seconds talking head

30 seconds ballet

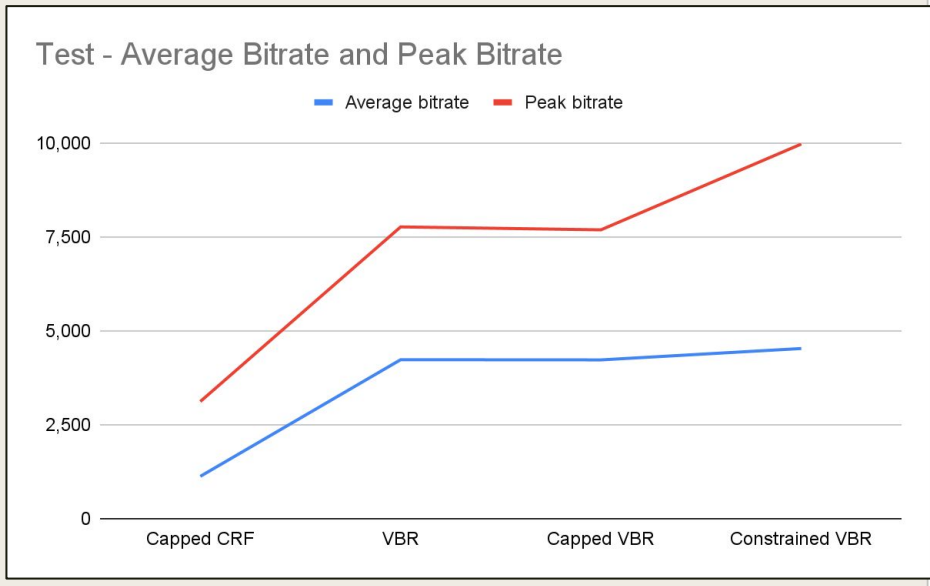
This clip is 8 minutes long, with 8 sequences of 30 second talking head followed by 30 seconds of ballet

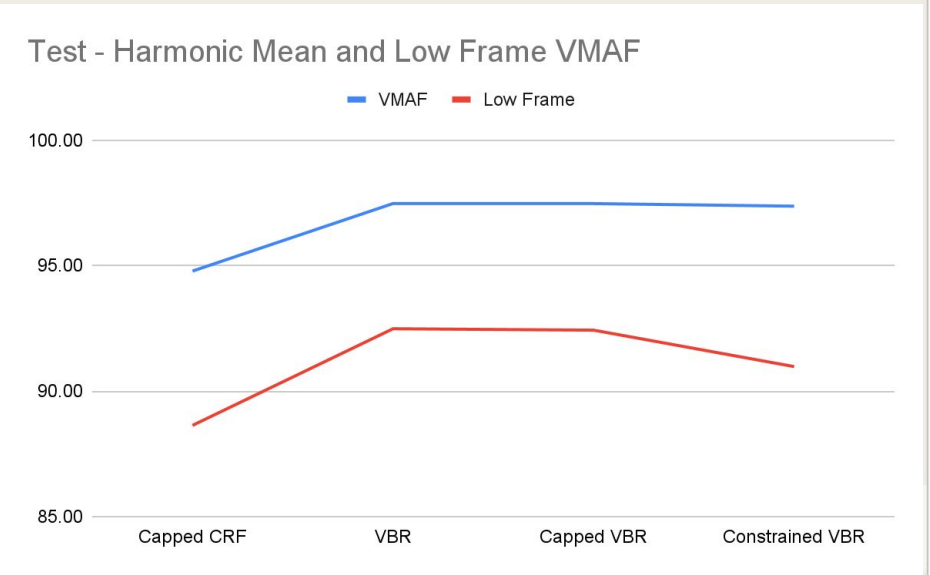
Data - Test

Observations:

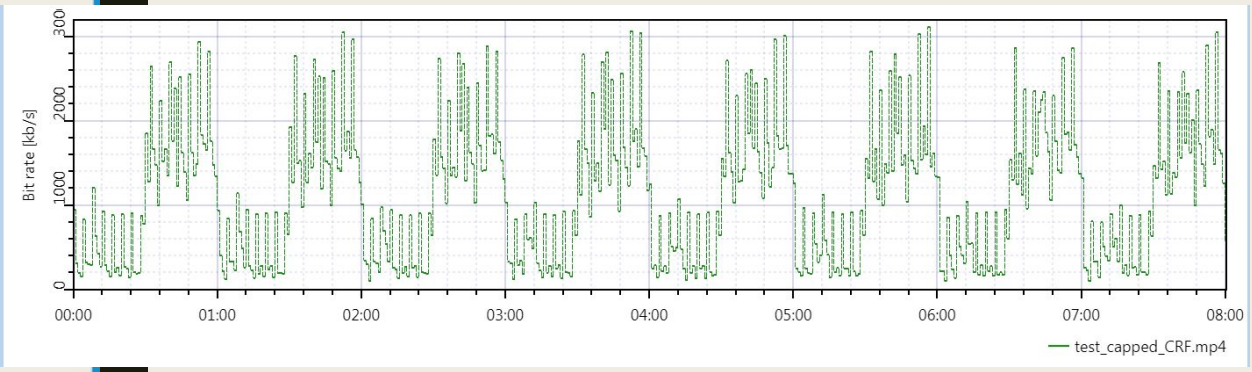
- **Capped CRF** - Followed content most closely producing very significant savings with minimum quality delta. Lowest overall VMAF but still at 94.78, within the typical 93-95 VMAF target.
- **VBR** - Best overall VMAF and low-frame, though bitrate is significantly higher than capped CRF.
- **Capped CBR** - Nothing to distinguish from VBR
- **Constrained VBR**
  - Highest average bitrate
  - Highest peak bitrate
  - Lowest VMAF of VBR variants
  - Lowest low-frame VMAF of VBR variants
- Let's look at the graphs

Test	Average bitrate	Capped CRF Savings	Peak bitrate	VMAF	Low Frame
Capped CRF	1,126		3,117	94.78	88.63
VBR	4,230	73.38%	7,764	97.47	92.48
Capped VBR	4,225	73.35%	7,686	97.47	92.43
Constrained VBR	4,528	80.43%	9,969	97.37	90.98

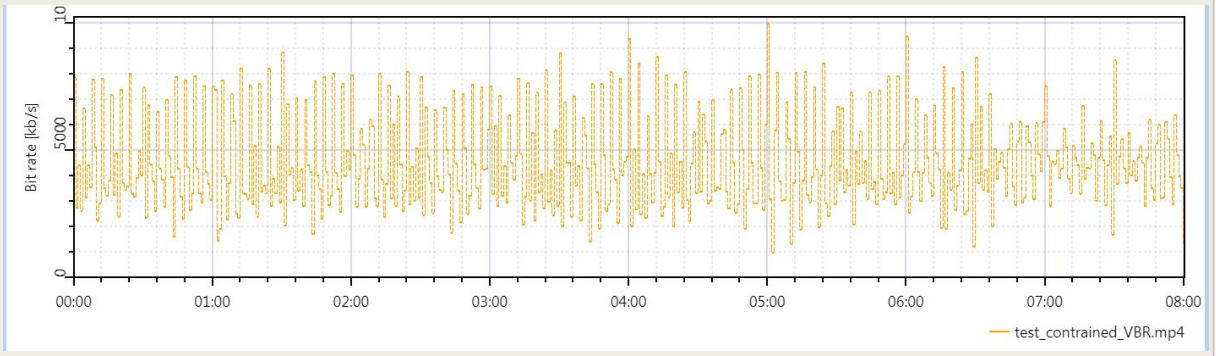




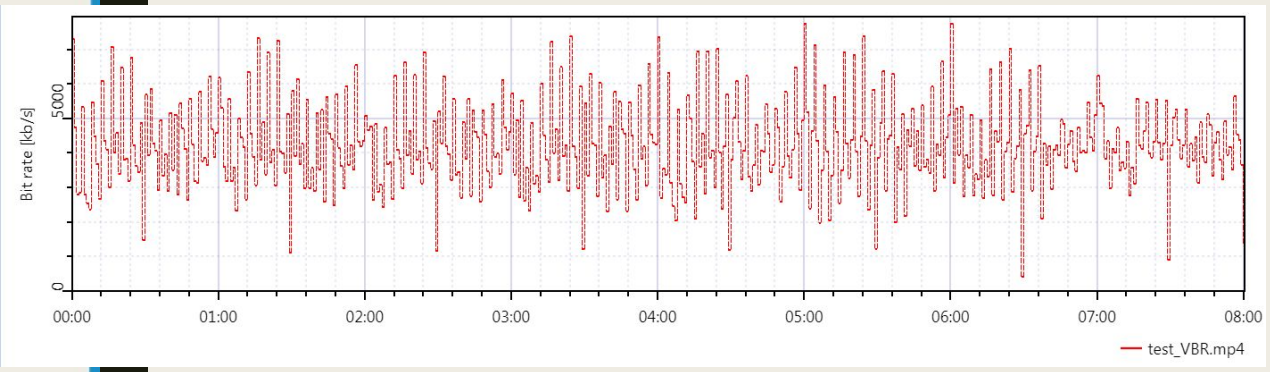
Test - Bitrate Graphs



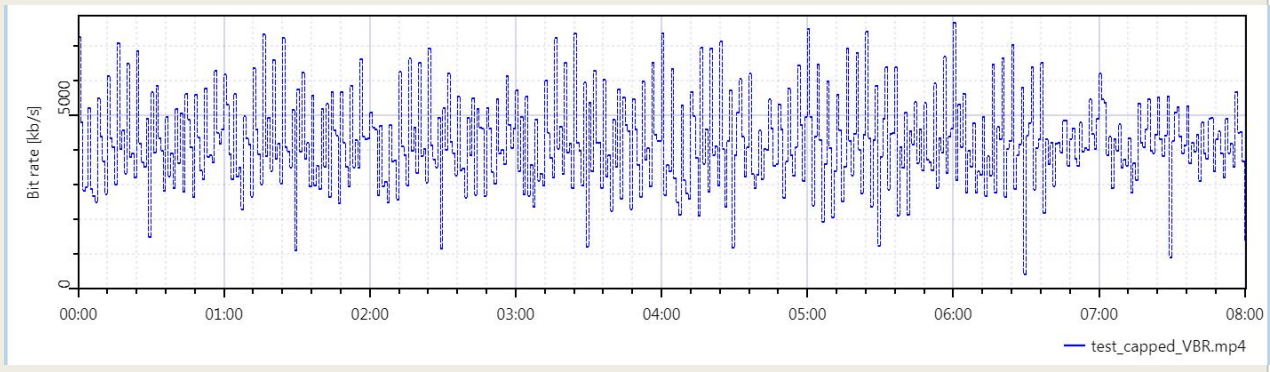
Capped CRF



Constrained VBR

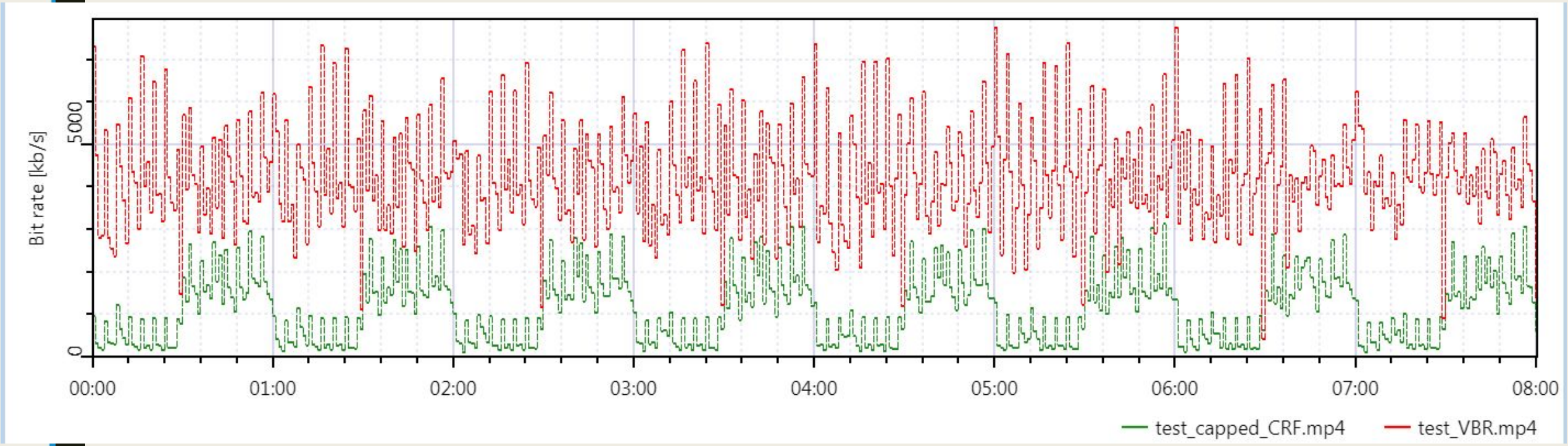


VBR



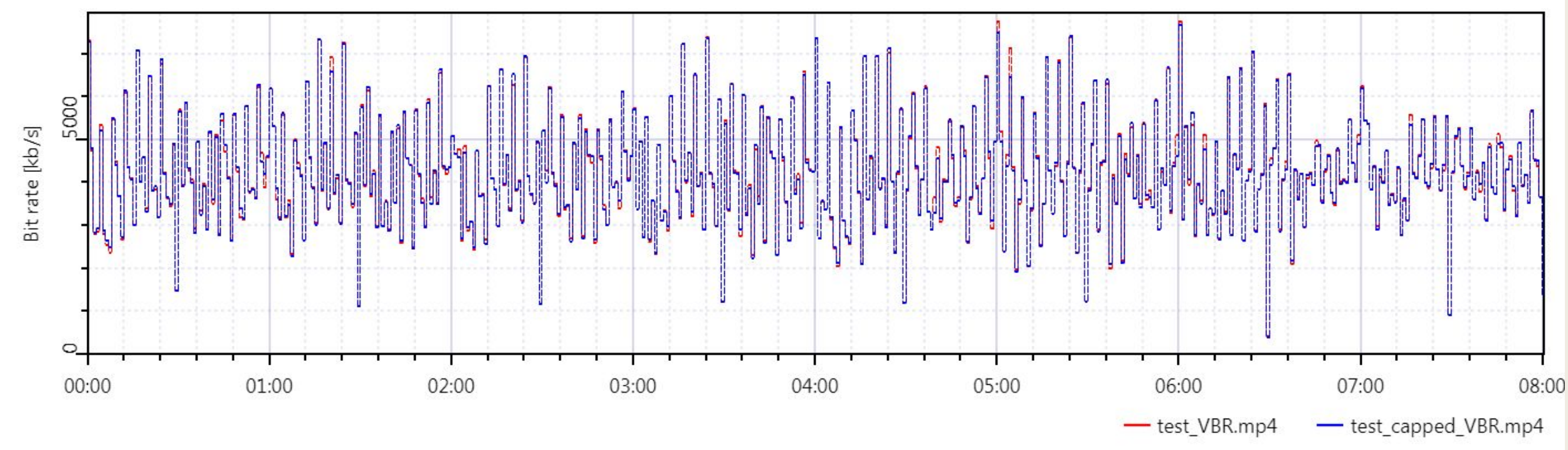
Capped VBR

Test - Capped CRF/VBR



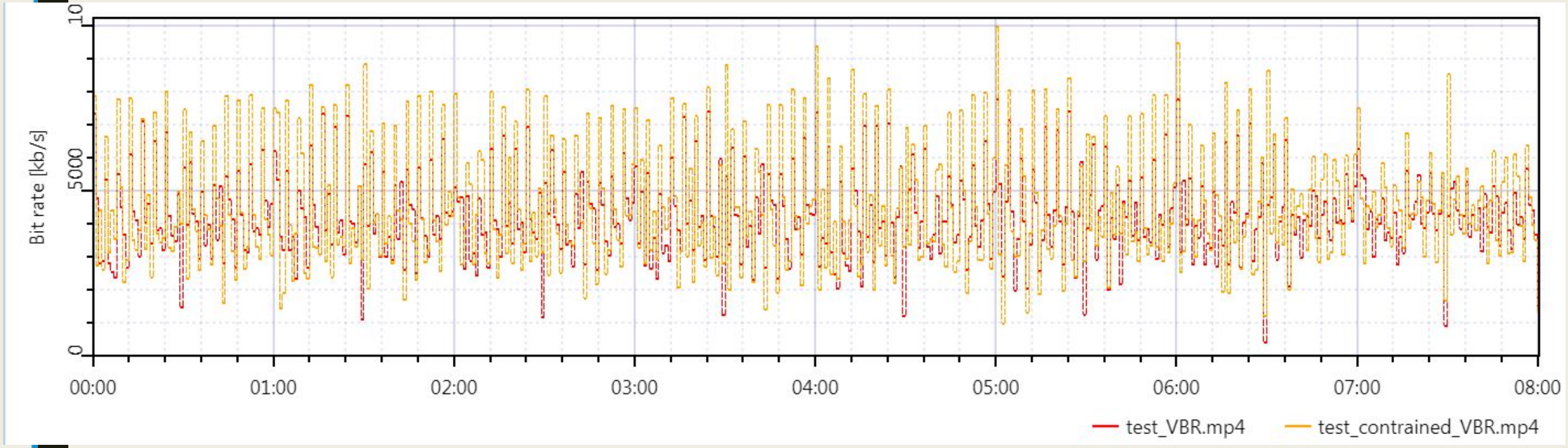
VBR is not looking VBR-ish at all, and doesn't significantly react to major content changes.

Test - VBR/Capped VBR



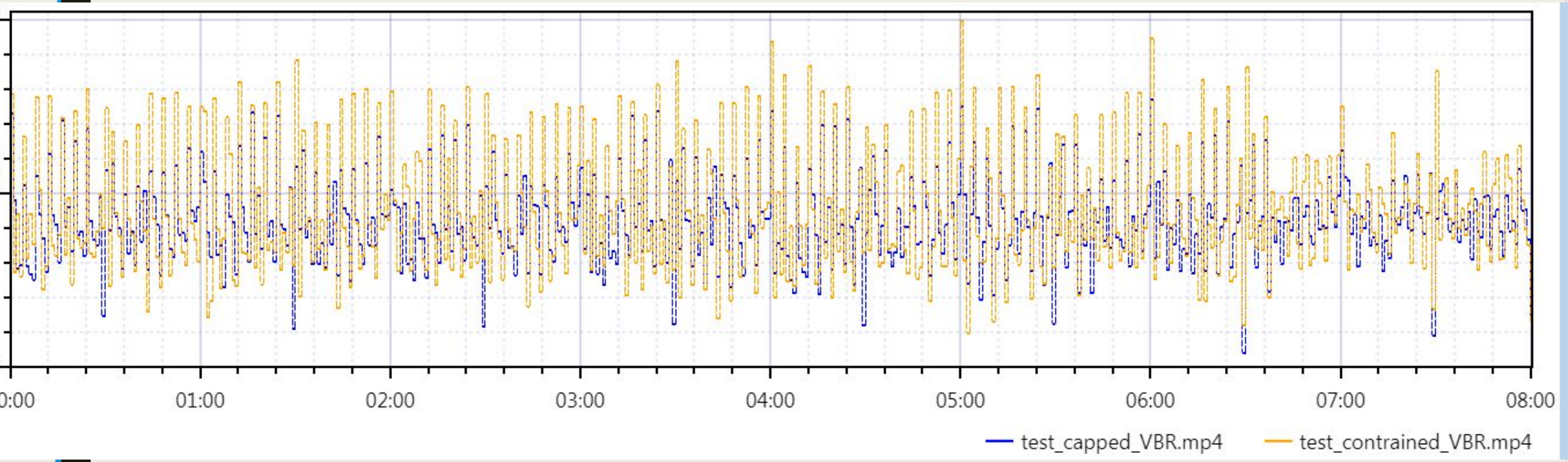
Again, capped VBR very similar to VBR (not surprising given 4501 mbr)

Test - VBR/Constrained VBR

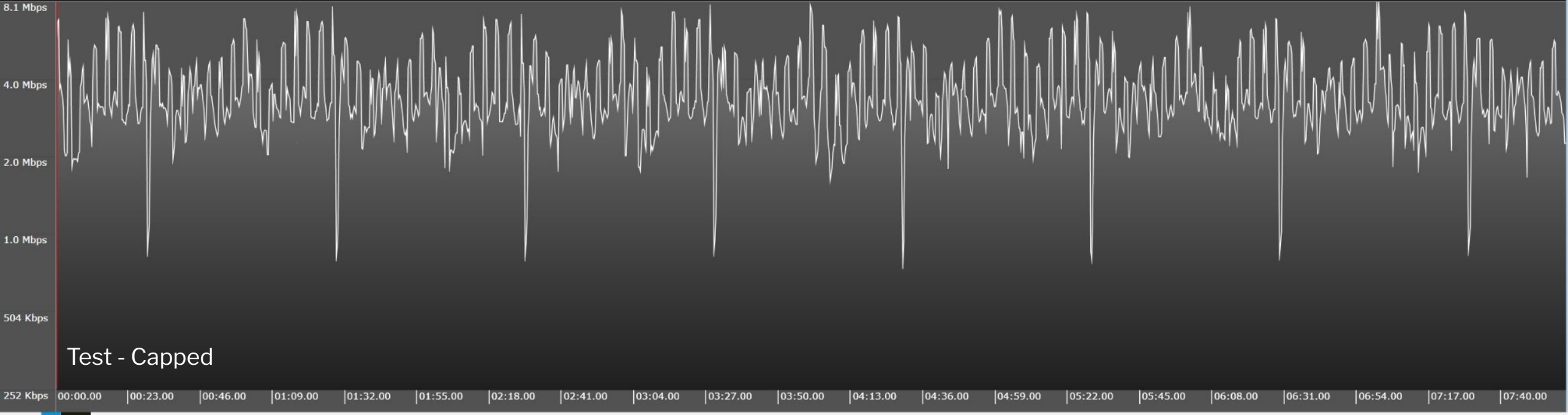


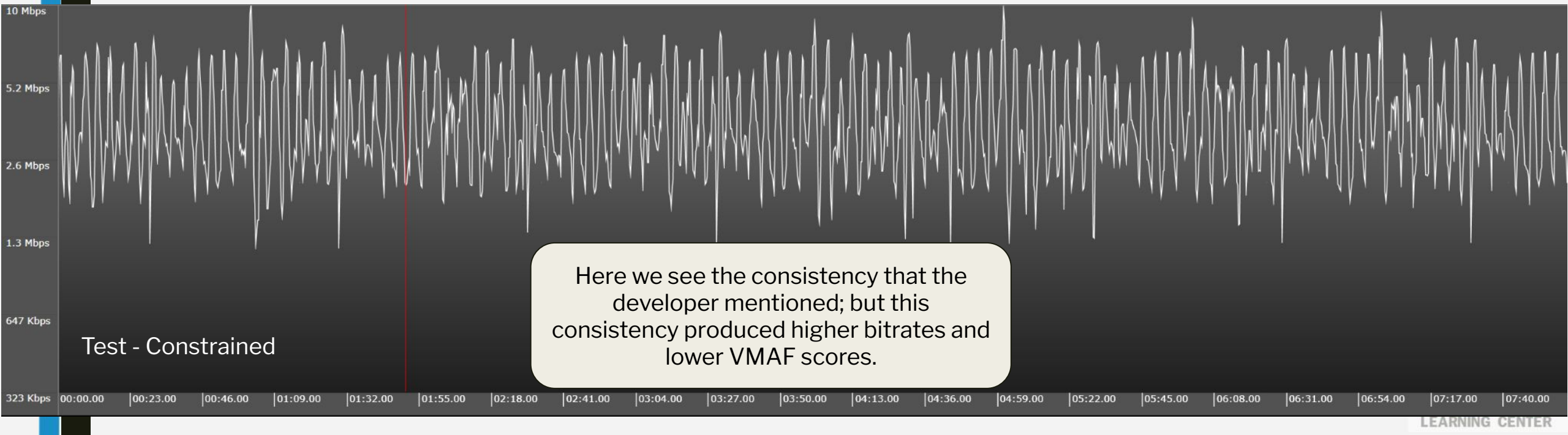
Again, clearly different algorithms; more later.

Test - Capped VBR/Constrained VBR



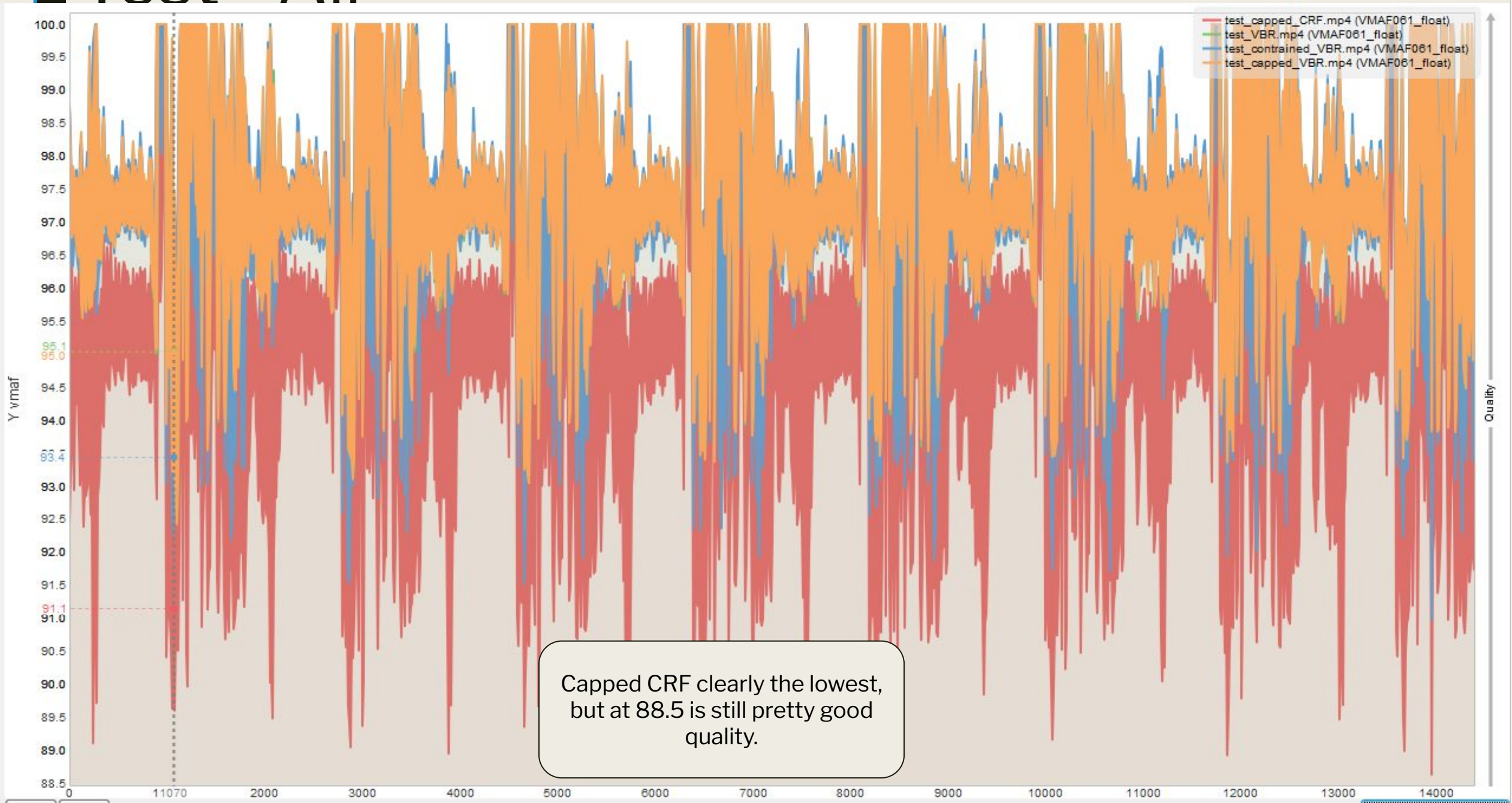
Again, clearly different algorithms; more later.





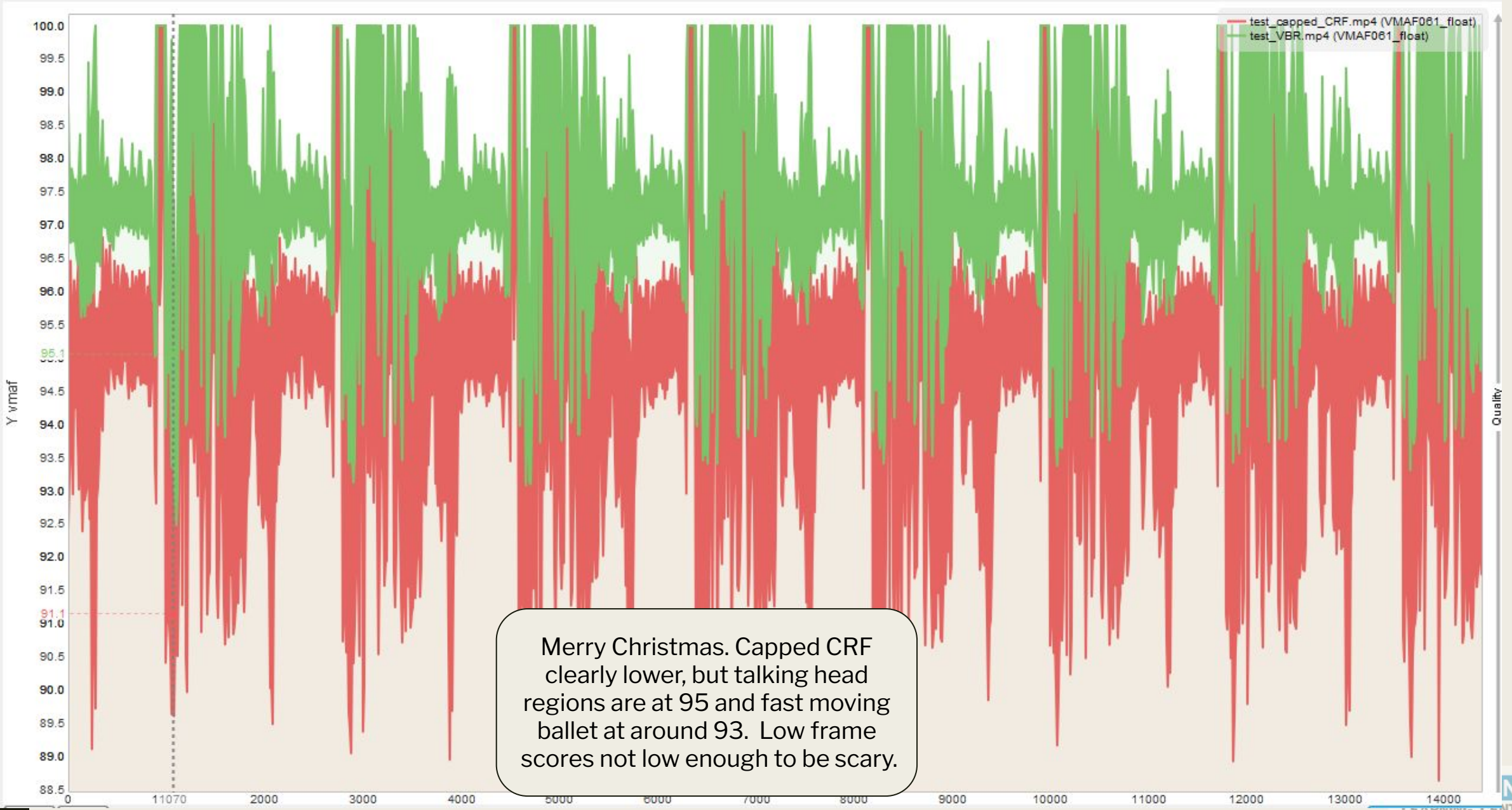
Here we see the consistency that the developer mentioned; but this consistency produced higher bitrates and lower VMAF scores.

Test - All



Capped CRF clearly the lowest, but at 88.5 is still pretty good quality.

Test - Capped CRF/VBR



Merry Christmas. Capped CRF clearly lower, but talking head regions are at 95 and fast moving ballet at around 93. Low frame scores not low enough to be scary.

Test - Constrained VBR / Capped VBR

Procedure - Test Clip

- Content description
- Performance tests (new)
- Bitrate/quality results
- Bitrate visualizations
- Quality visualization

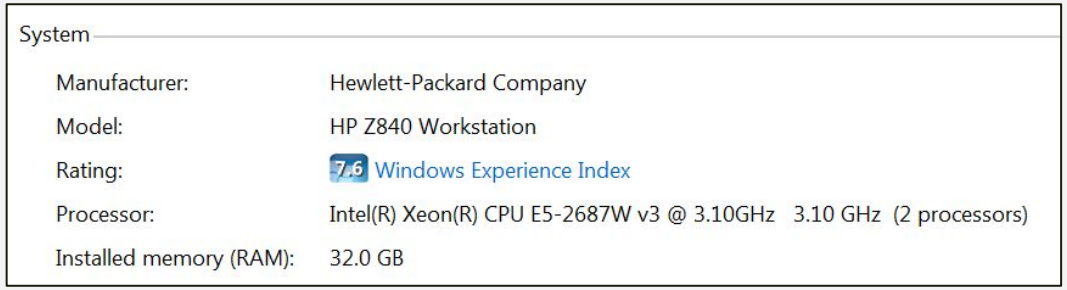
Football Test Clip

High Motion

A two-minute segment of Harmonic's iconic football test clip.

Throughput

	Football - No LP		Football - 8 LP	
	Encoding Speed	**CCRF increase**	**Encoding Speed**	**CCRF increase**
Capped CRF	1.7		1.26	
VBR	1.38	23.2%	1.15	9.6%
Capped VBR	1.34	26.9%	1.09	15.6%
Constrained VBR	1.37	24.1%	1.15	9.6%



I tested on my old 40-core HP Z840 workstation without limiting the logical processors. You see that with no logical processor limit, capped CRF encoded at 1.7x real time, about 25% faster than the VBR alternatives.

I then used the lp switch to limit the logical processors used to 8 and the advantage dropped to about 10-15%.

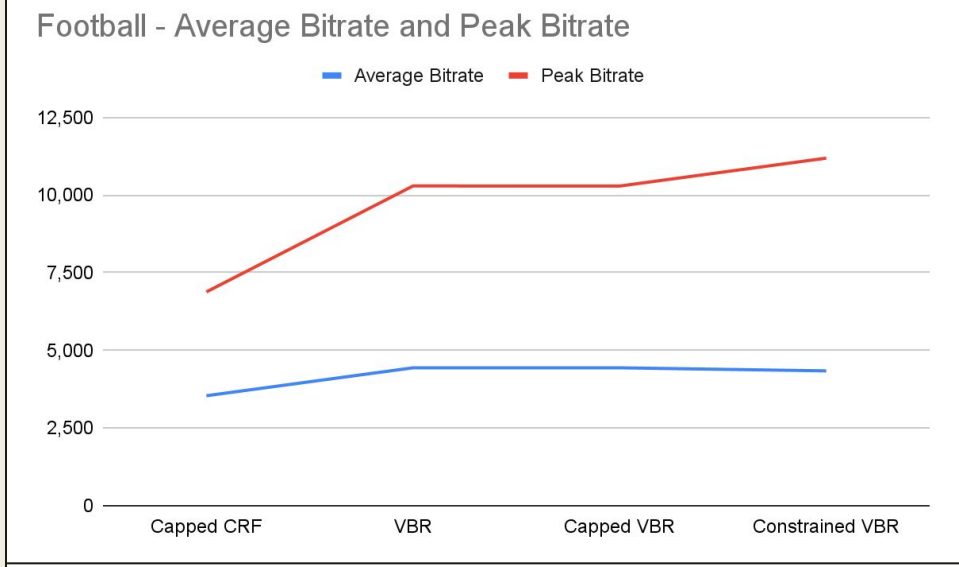
My contact on the SVT-AV1 development team told me that at least part of capped CRF's speed advantage related to the lower bitrate, as opposed to the technique itself. The bottom line is that with challenging footage, the bitrate differential should be modest as compared to less challenging footage, but that capped CRF should deliver slightly more throughput than any of the VBR alternatives.

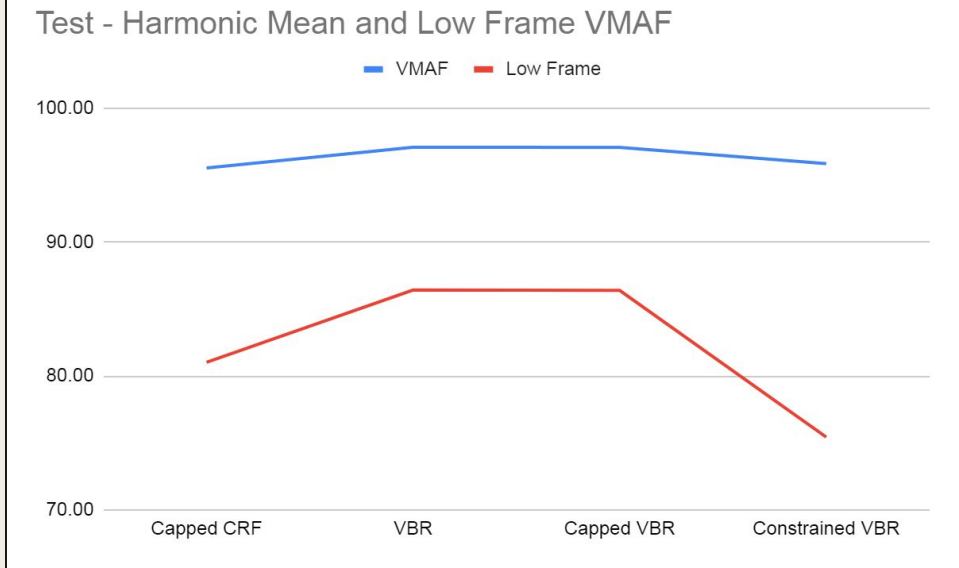
Data - Football

Observations:

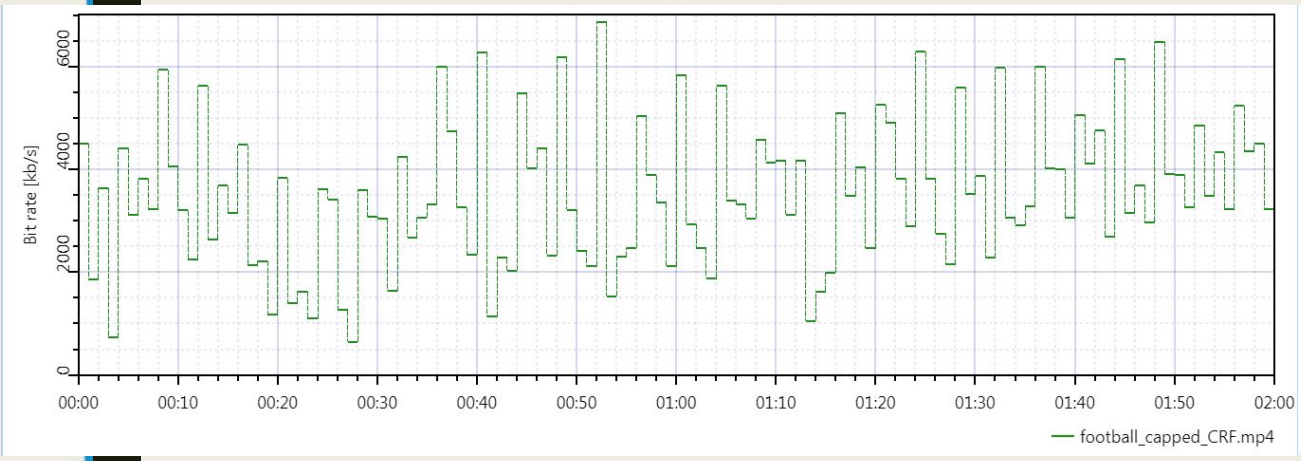
- **Capped CRF** - As you would expect, the capped CRF bitrate advantage dropped with challenging footage. Still, capped CRF delivered a ~20% savings while keeping overall VMAF above 95.
- **VBR** - best overall performer of the three VBR variants.
- **Capped CBR** - Nothing to distinguish from VBR
- **Constrained VBR**
  - Highest peak bitrate
  - Lowest VMAF of VBR variants
  - Lowest low-frame VMAF of VBR variants
- Let's look at the graphs

Football	Average Bitrate	Capped CRF Savings	Peak Bitrate	VMAF	Low Frame
Capped CRF	3,540		6,882	95.53	81.02
VBR	4,439	20.25%	10,293	97.07	86.42
Capped VBR	4,439	20.25%	10,290	97.06	86.39
Constrained VBR	4,340	18.43%	11,188	95.85	75.44

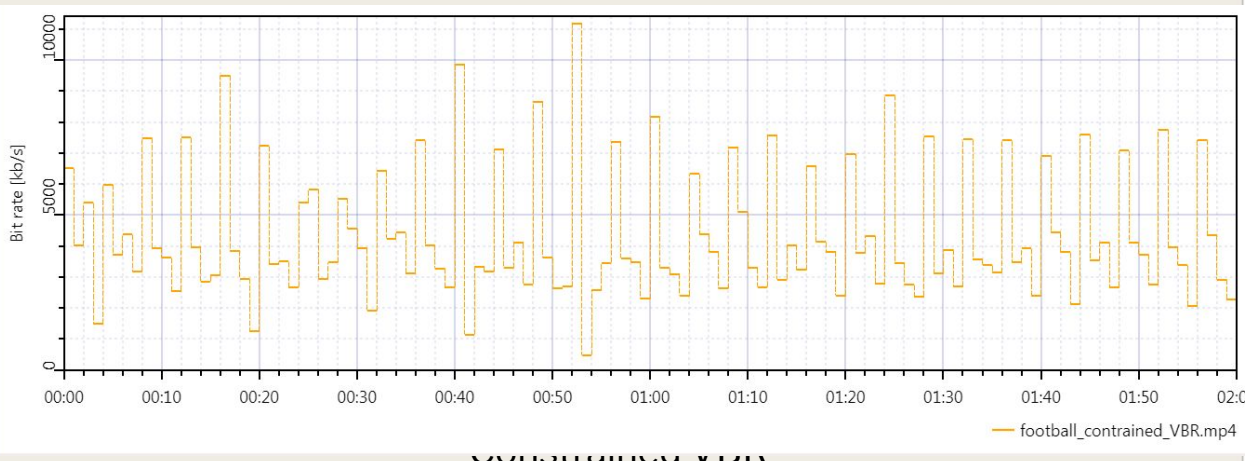




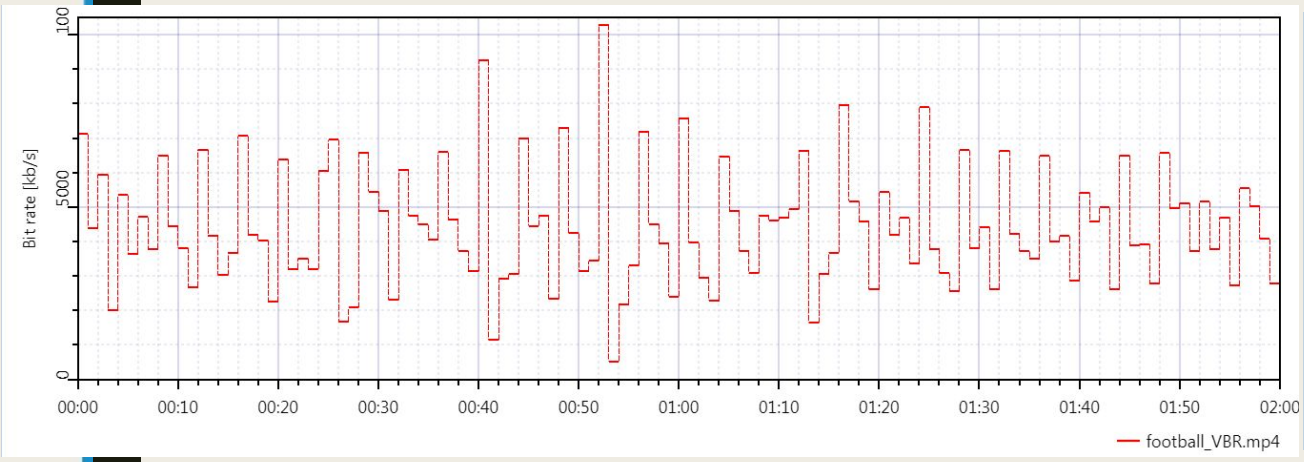
Football - Bitrate Graphs



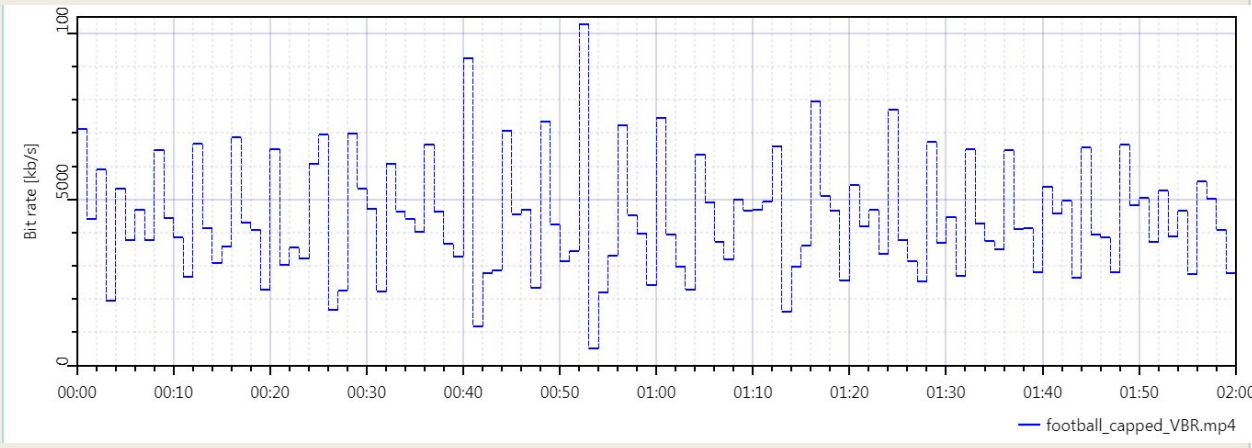
Capped CRF



Constrained VBR

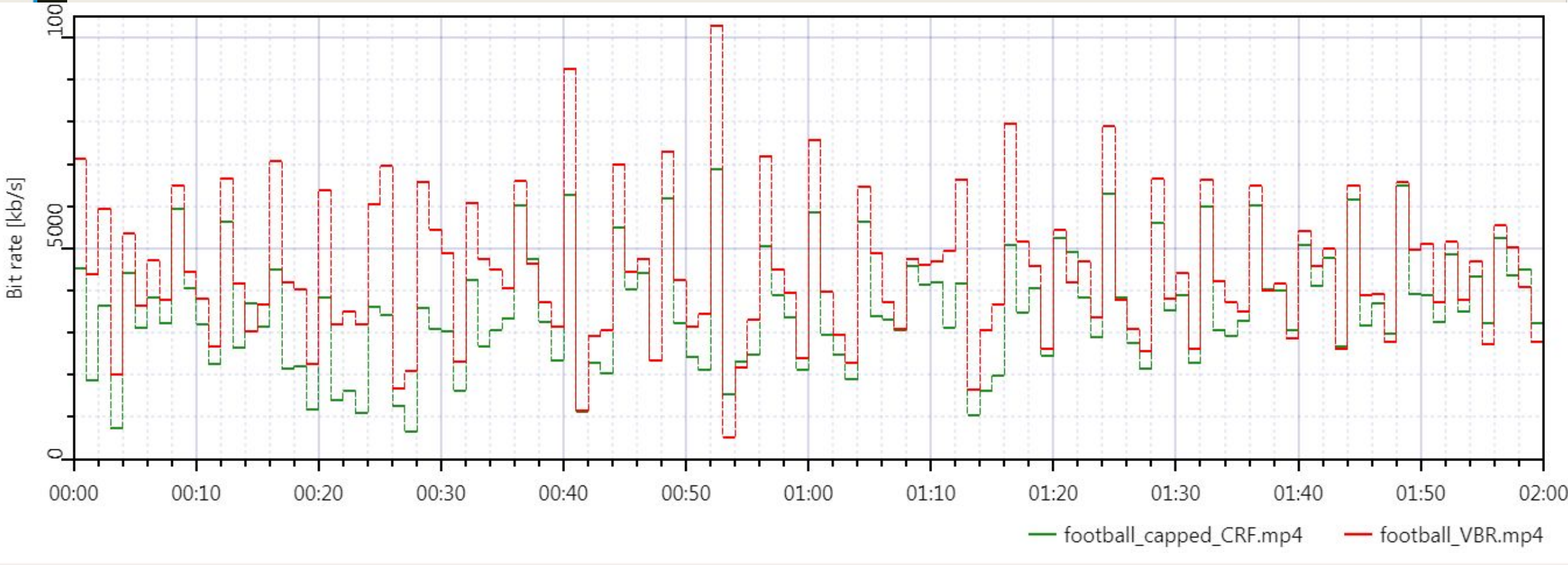


VBR



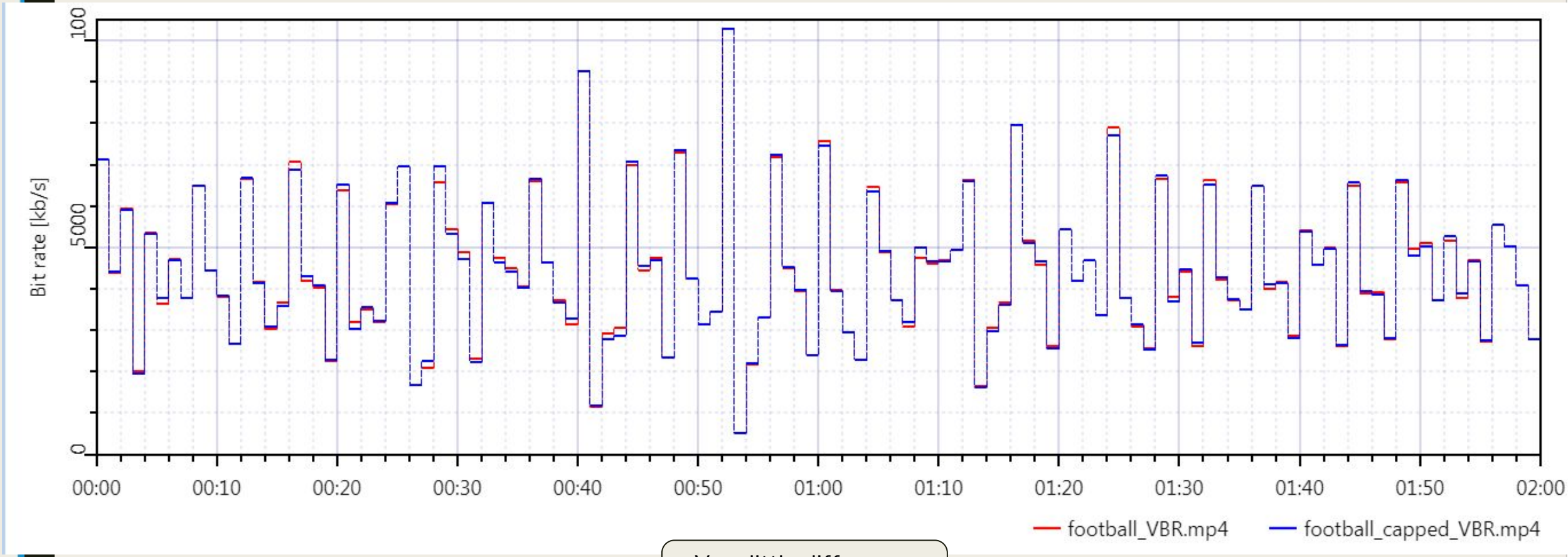
Capped VBR

Football - Capped CRF/VBR



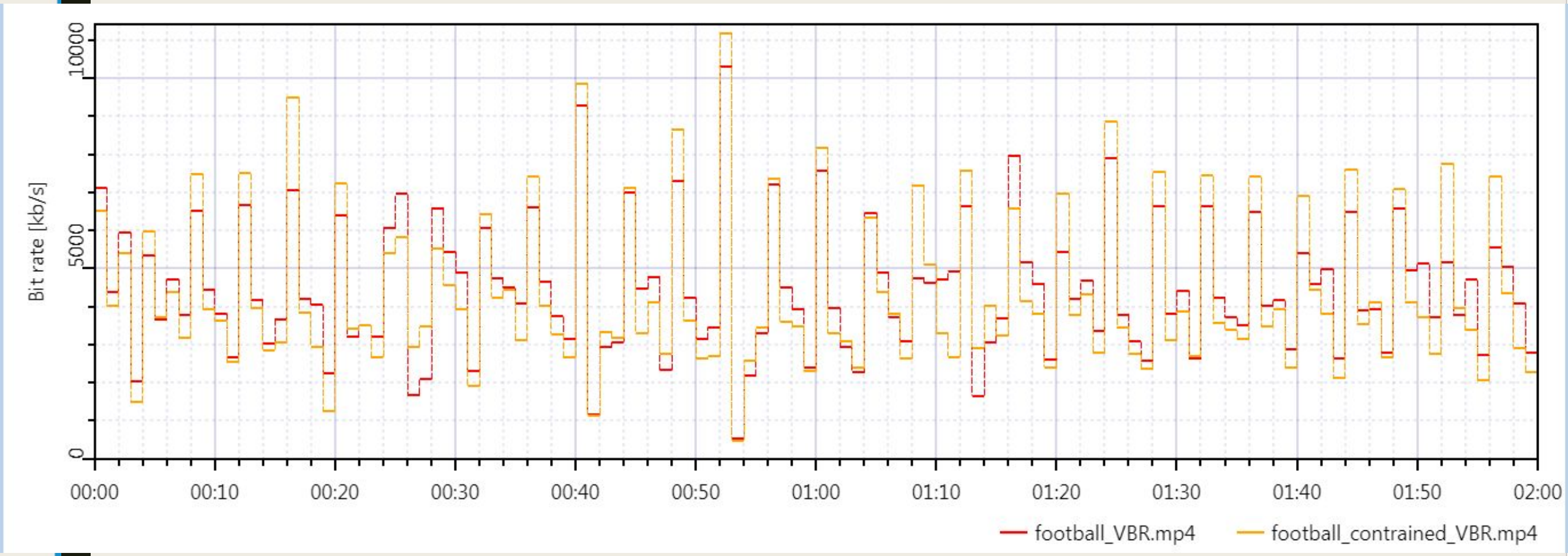
Capped CRF generally lower throughout

Football - VBR/Capped VBR



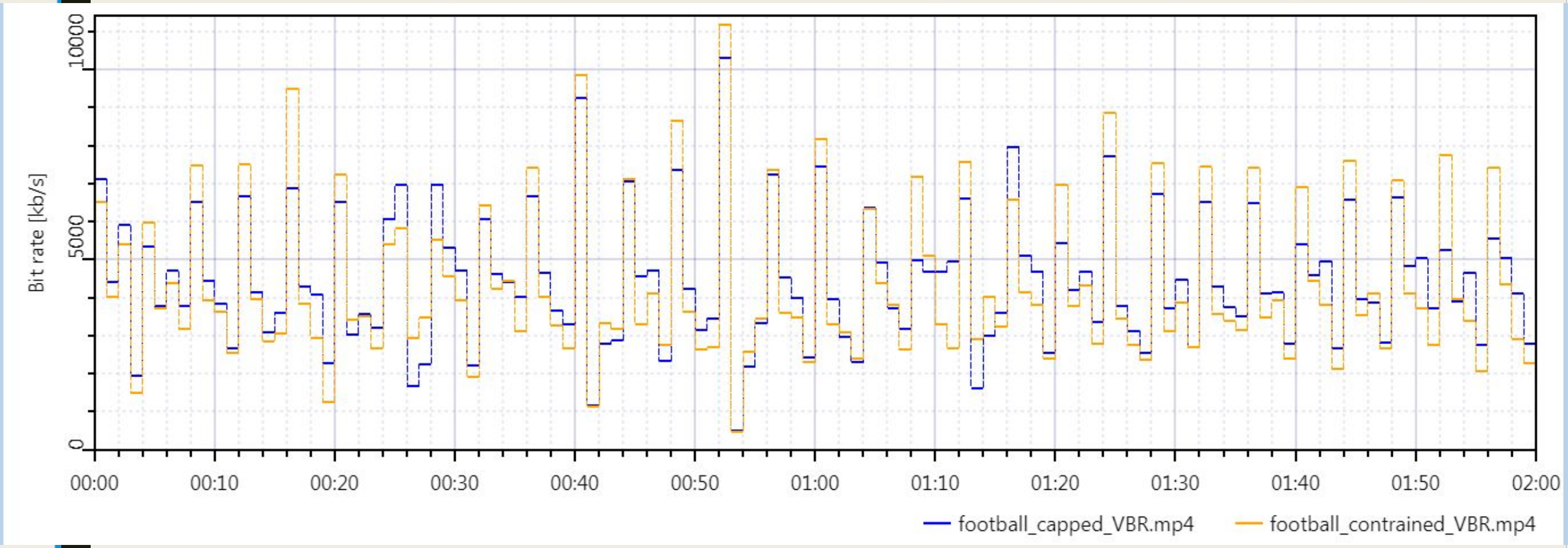
Very little difference here.

Football - VBR/Constrained VBR

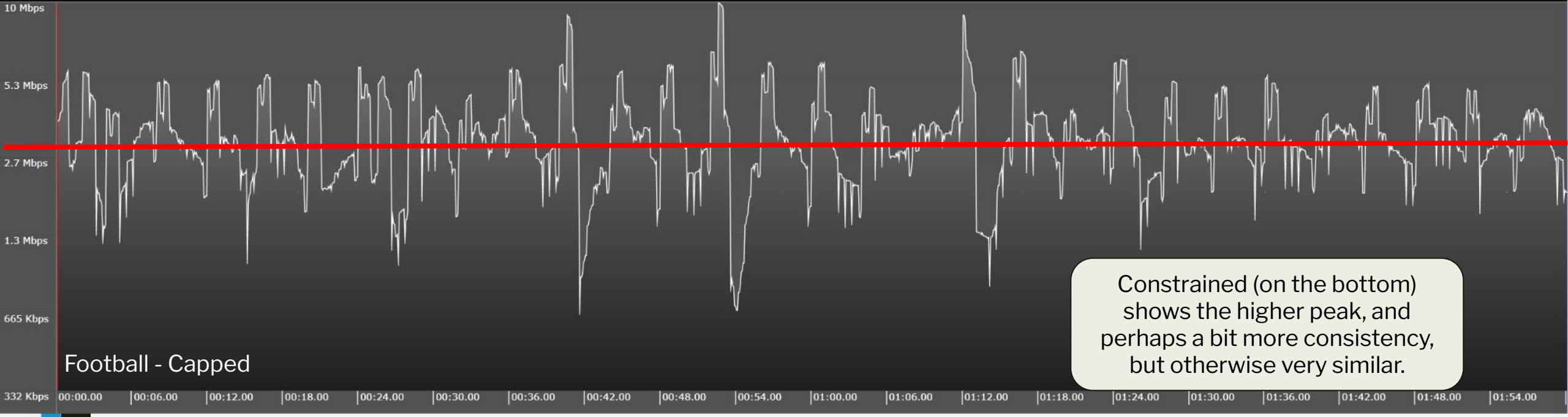


Again, different algorithms working.

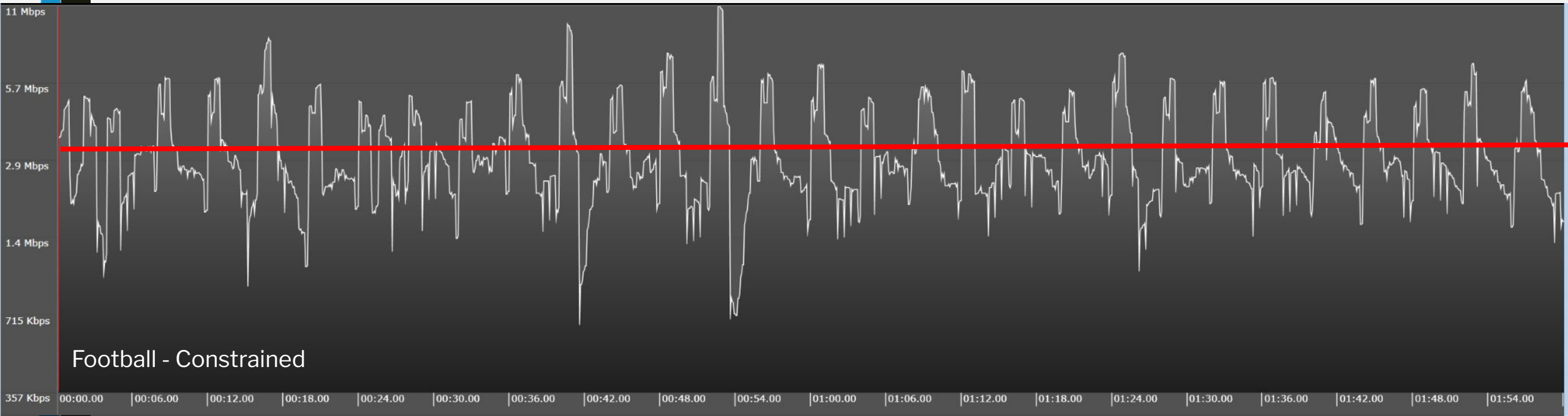
Football - Capped VBR/Constrained VBR



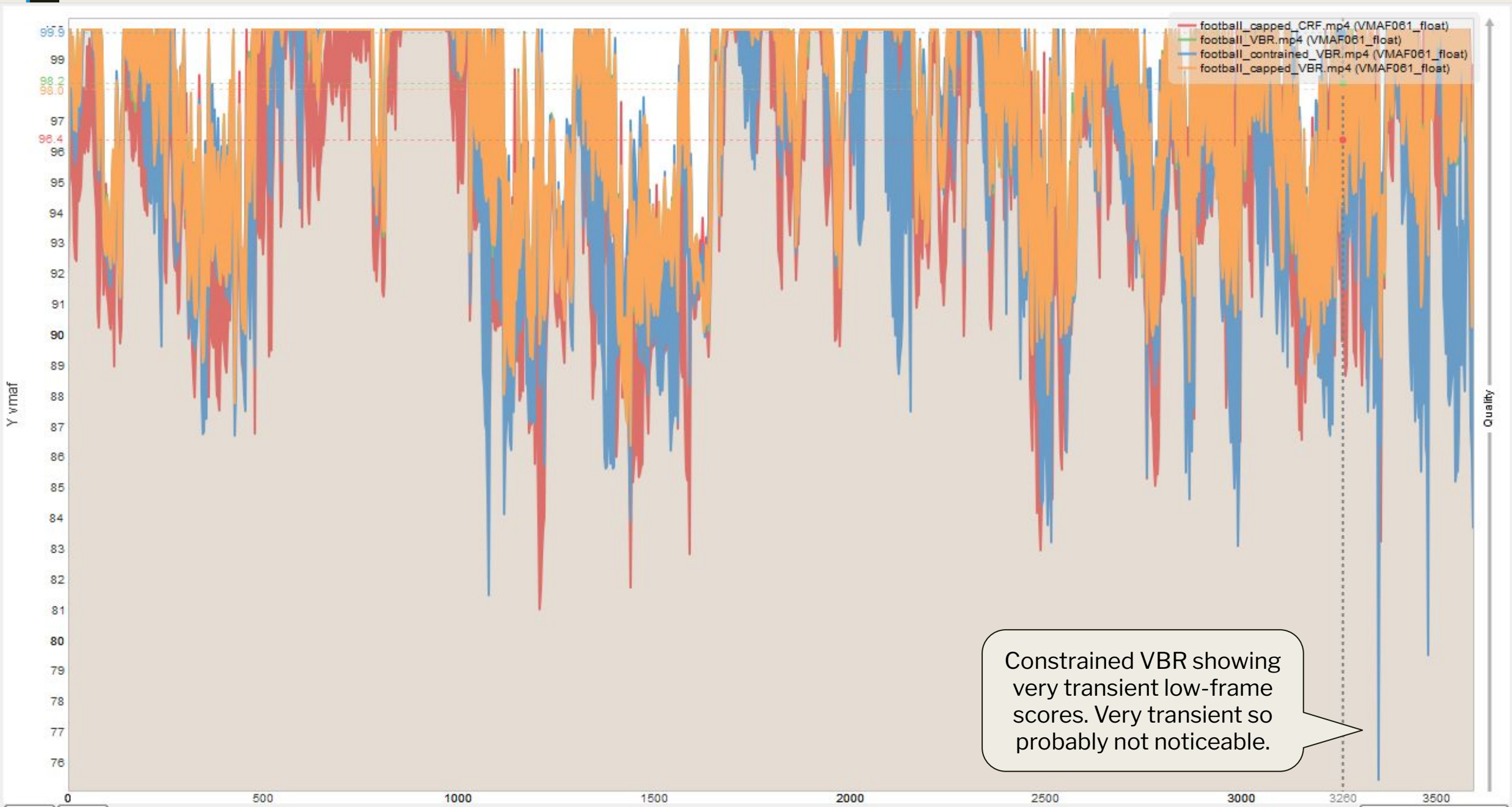
Here too.



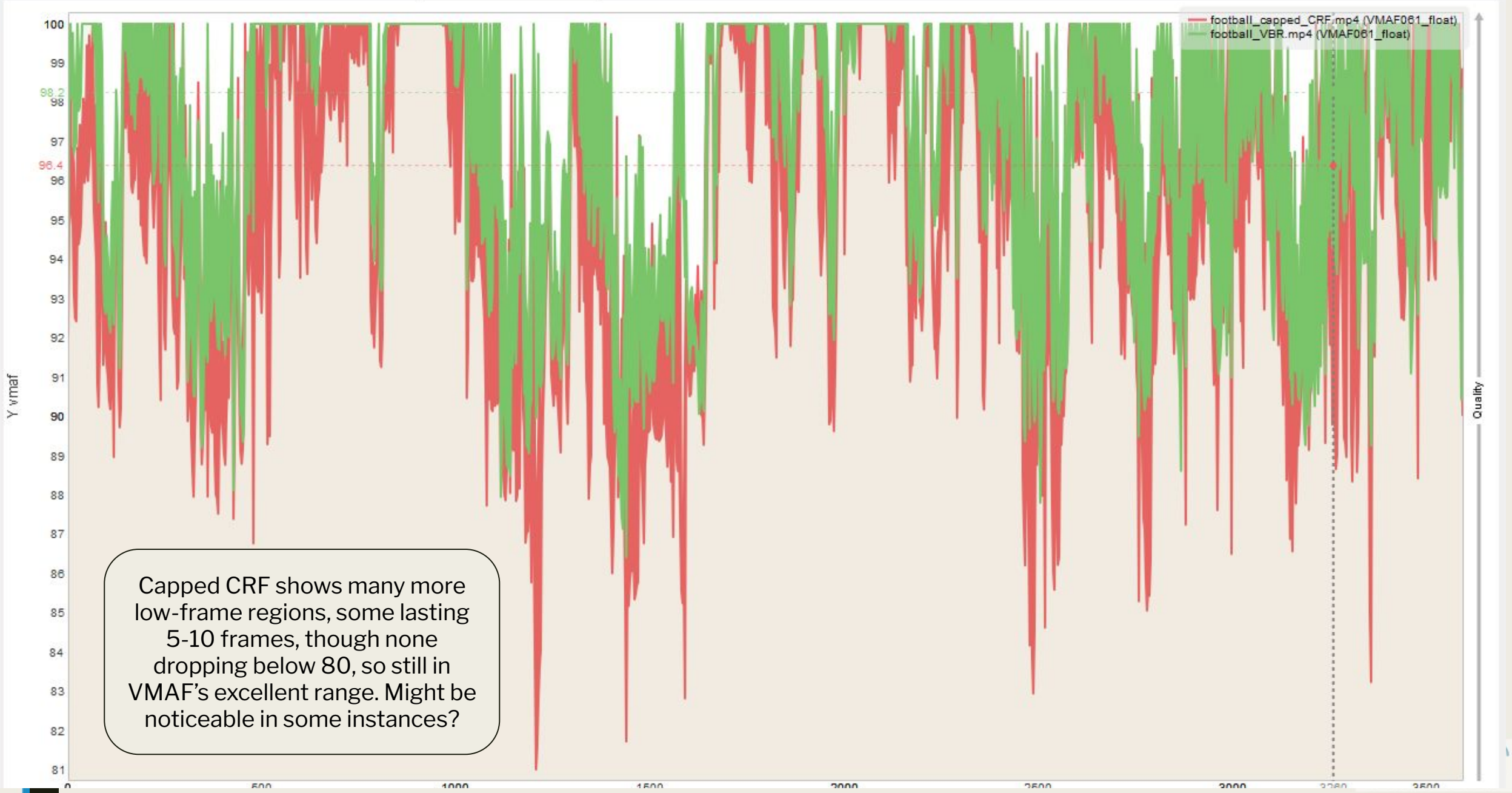
Constrained (on the bottom) shows the higher peak, and perhaps a bit more consistency, but otherwise very similar.



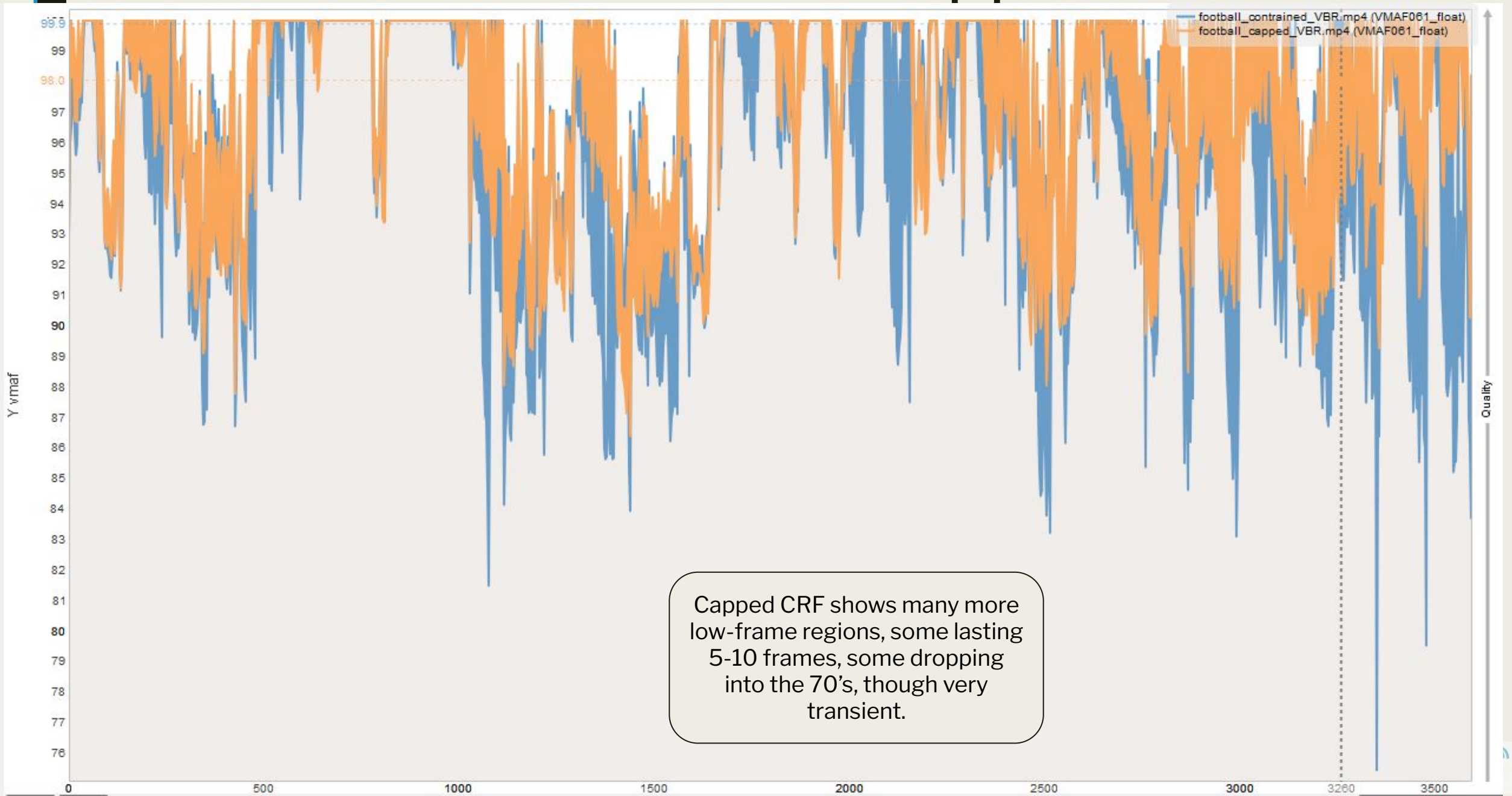
Football - All



Football - Capped CRF/VBR



Football - Constrained / Capped VBR



Meridian

- Content description
- Performance tests
- Bitrate/quality results
- Bitrate visualizations
- Quality visualization

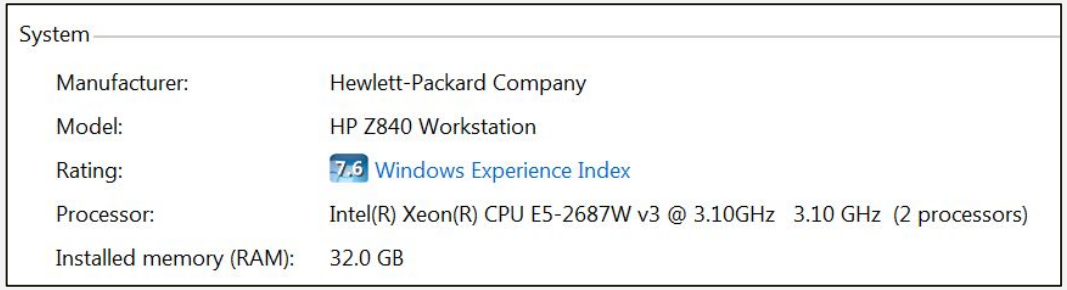
Test 4 - Meridian Test Clip

High Motion

When you're Netflix, you don't find test clips, you create your own. This is a 3-minute extract from Netflix's Meridian.

Throughput

	Meridian - No LP		Meridian - 8 LP	
	Encoding Speed	**CCRF increase**	**Encoding Speed**	**CCRF increase**
Capped CRF	2.63		1.95	
VBR	2.03	29.6%	1.65	18.2%
Capped VBR	2.01	30.8%	1.62	20.4%
Constrained VBR	2.01	30.8%	1.66	17.5%



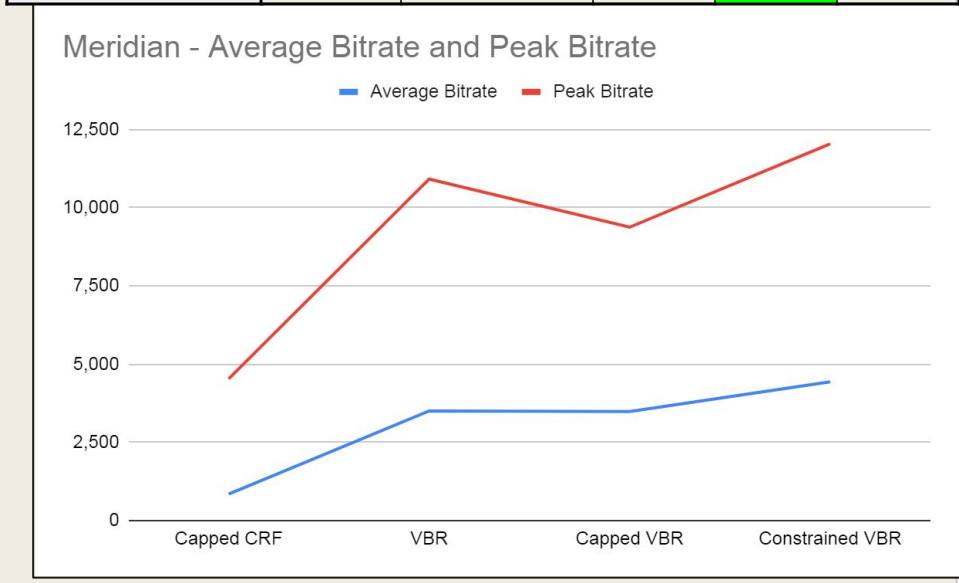
Same CPU, and without logical processor limits, capped CRF delivered ~30% more throughput than any of the VBR alternatives. When limited to 8 logical processors, the differential dropped down to around 20% depending upon the VBR variant.

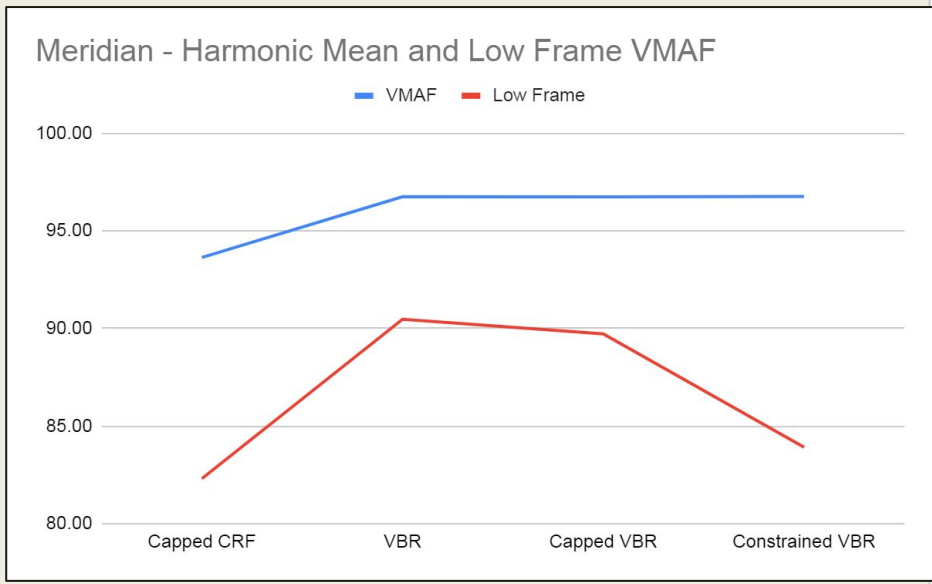
Live event producers should find the performance disparity between the Football clip and the Meridian clip concerning because CPU utilization will vary with the footage.

This means that you need to leave plenty of performance headroom for any clip with mixed complexity or (gulp) consider ASIC-based hardware which isn't impacted by content complexity.

Data - Meridian

Meridian	Average Bitrate	Capped CRF Savings	Peak Bitrate	VMAF	Low Frame
Capped CRF	844		4,530	93.64	82.31
VBR	3,504	75.91%	10,917	96.75	90.47
Capped VBR	3,484	75.77%	9,379	96.75	89.73
Constrained VBR	4,435	80.97%	12,043	96.77	83.91

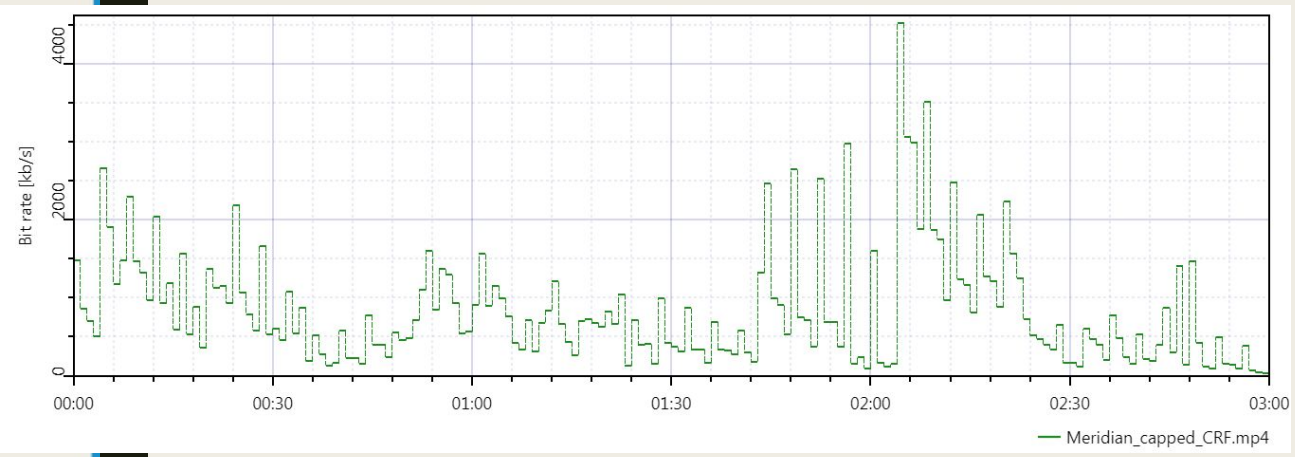




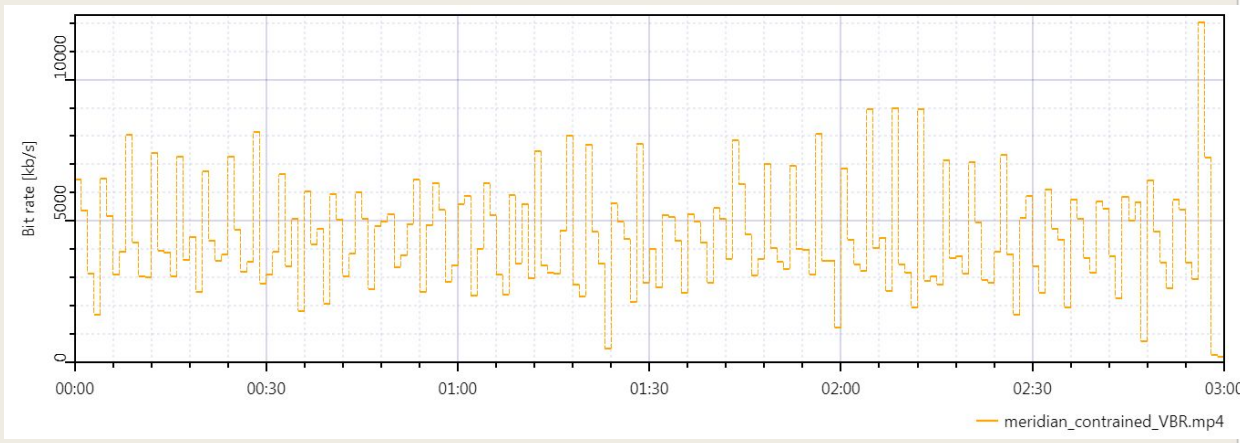
Observations:

- **Capped CRF** - On this mostly low-motion clip, capped CRF delivered very significant savings while keeping the VMAF value above 93 (though you might consider one lower CRF value to boost quality slightly at the cost of some bitrate savings).
- **VBR** - Very good among VBR variants
- **Capped CBR** - Nothing to distinguish from VBR at these test parameters
- **Constrained VBR**
  - Much higher bitrate than other VBR variants
  - Highest peak bitrate
  - Highest VMAF but at 1 Mbps higher than VBR/Capped VBR
  - Lowest low-frame VMAF of VBR variants
- Let's look at the graphs

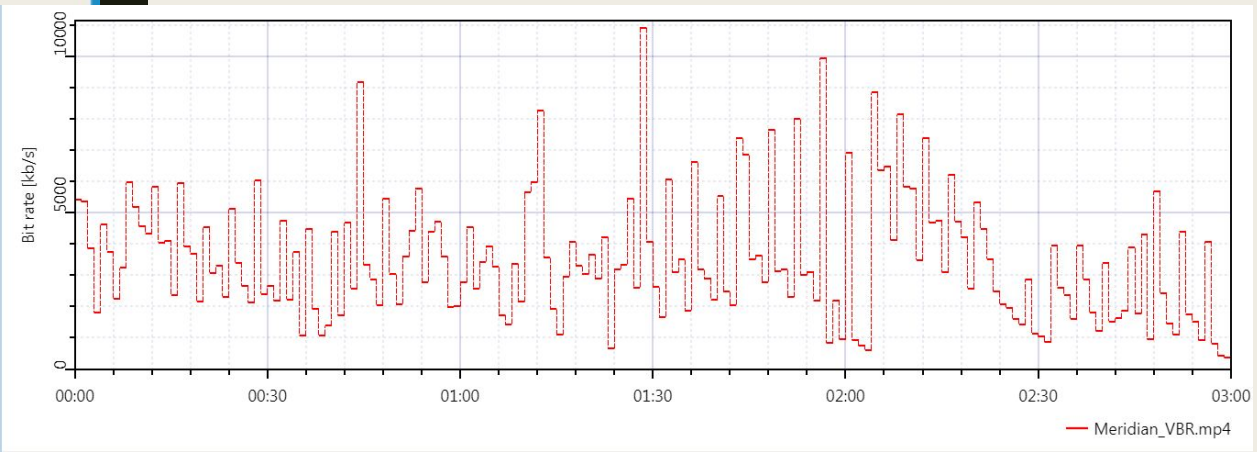
Meridian - Bitrate Graphs



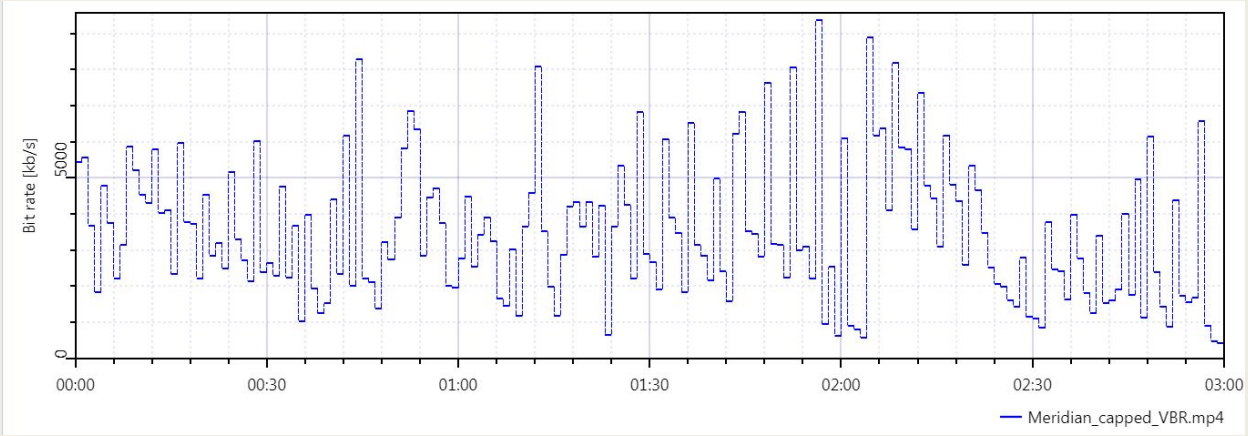
Capped CRF



Constrained VBR

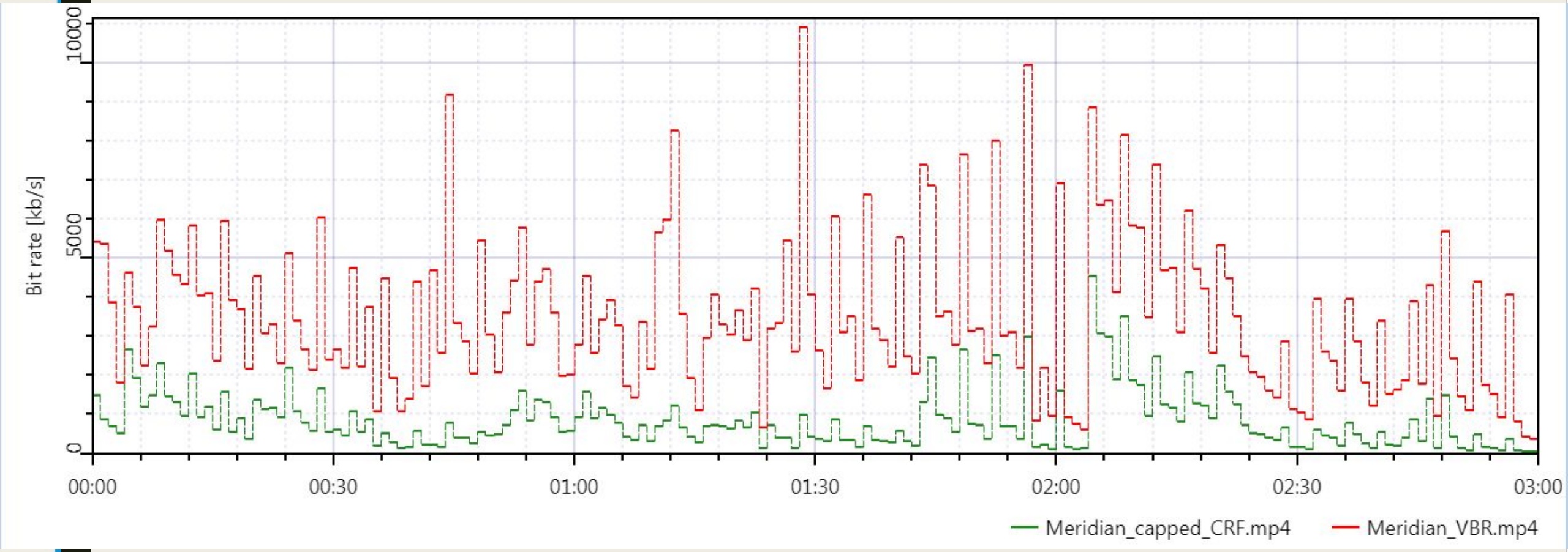


VBR



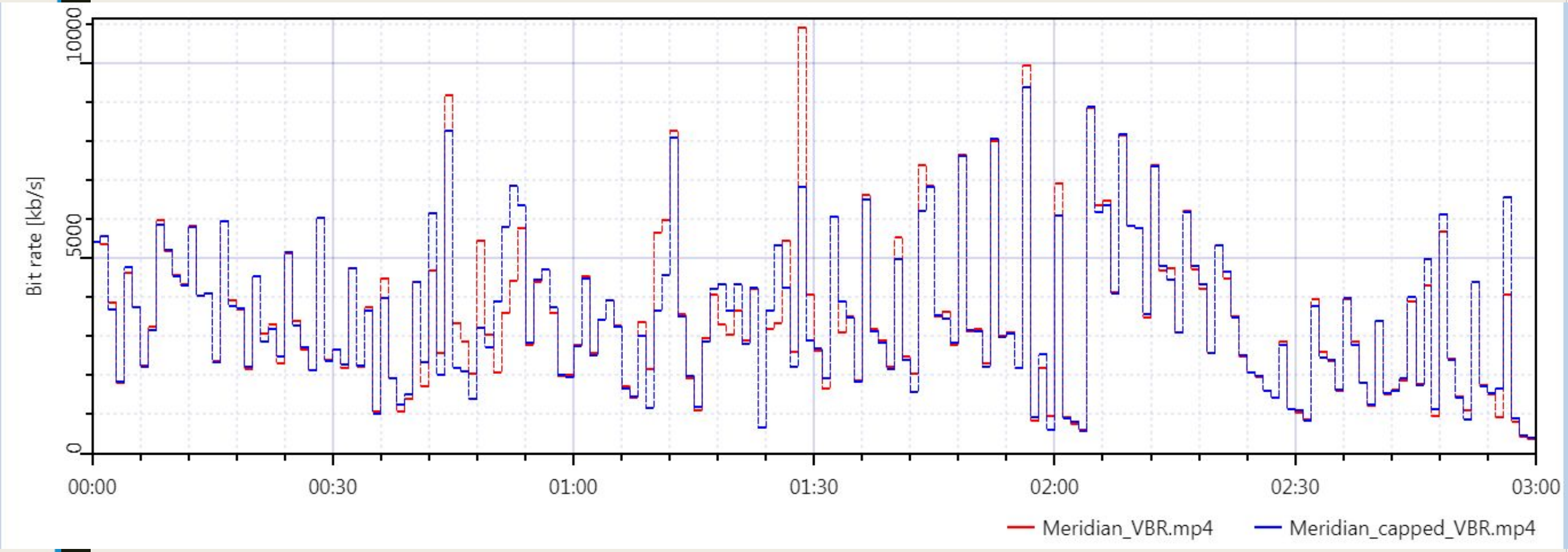
Capped VBR

Meridian - Capped CRF/VBR



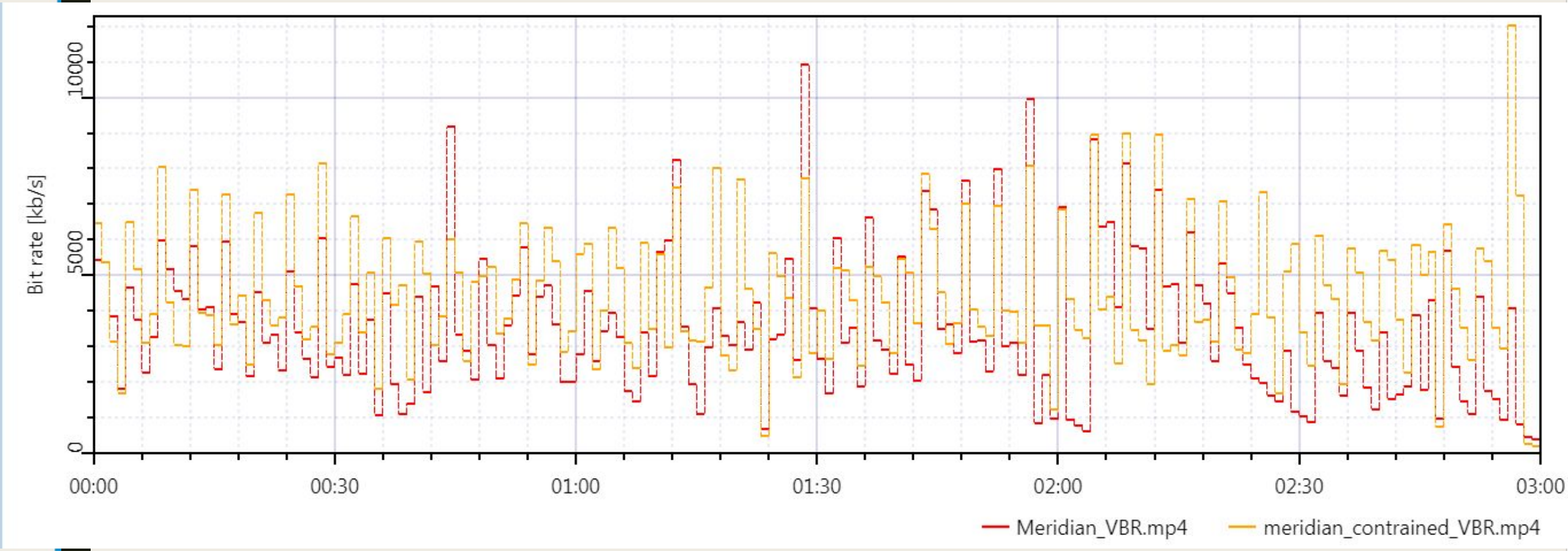
No surprise that capped CRF was ~75% lower bitrate than VBR.

Meridian - VBR/Capped VBR



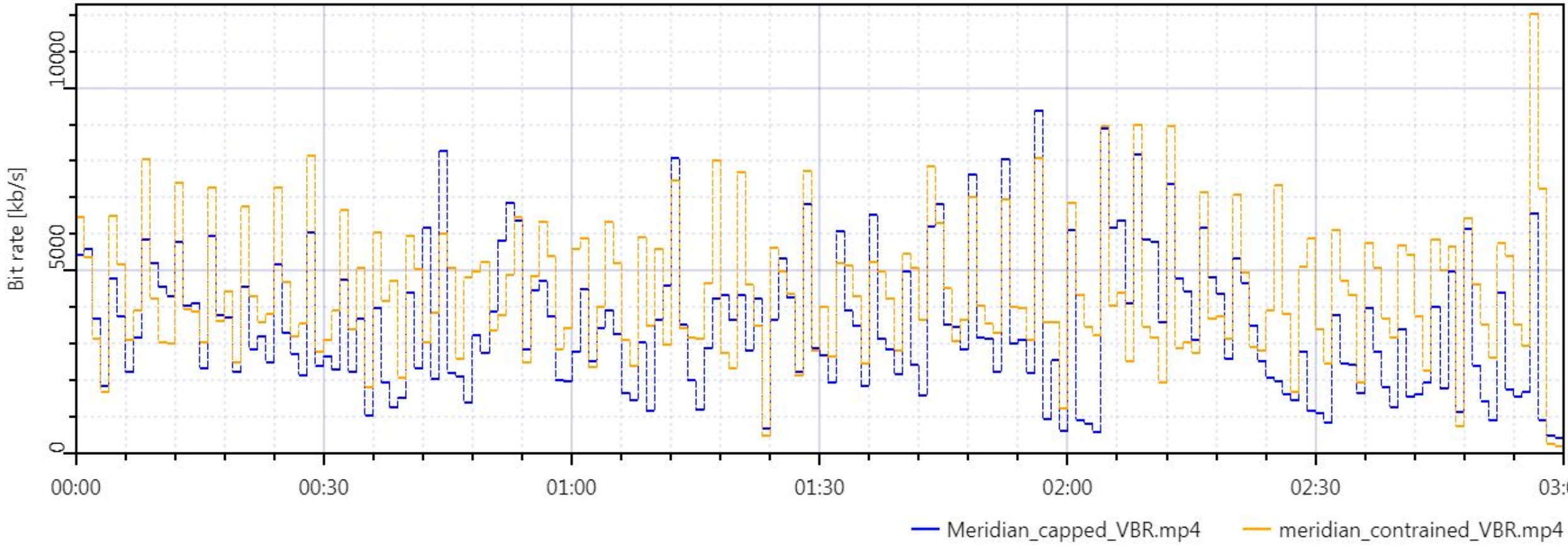
Slight differential between VBR and capped VBR, but not striking.

Meridian - VBR/Constrained VBR

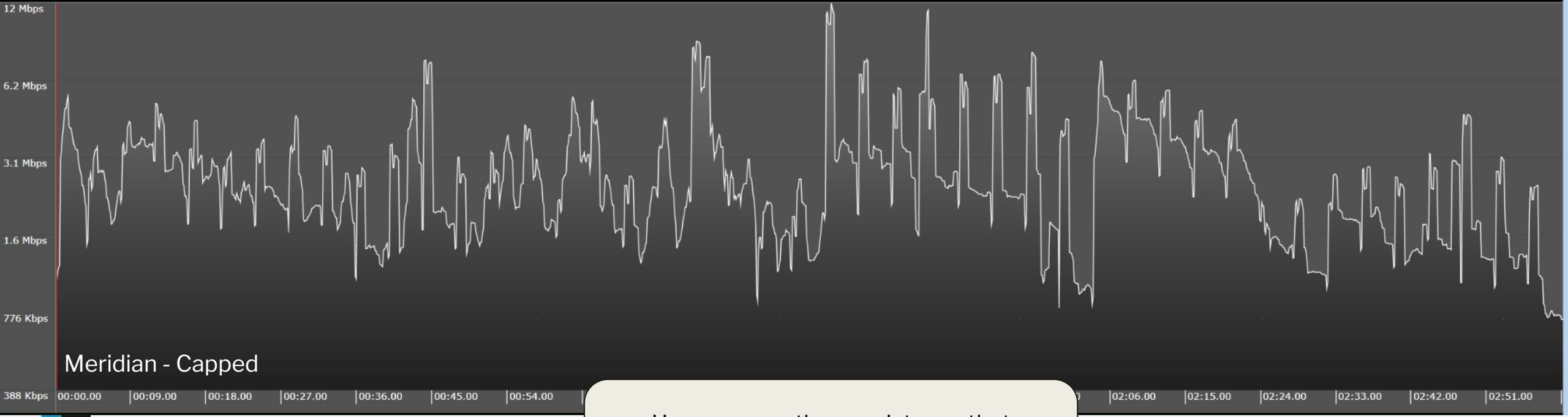


Again, clearly different algorithms at work here.

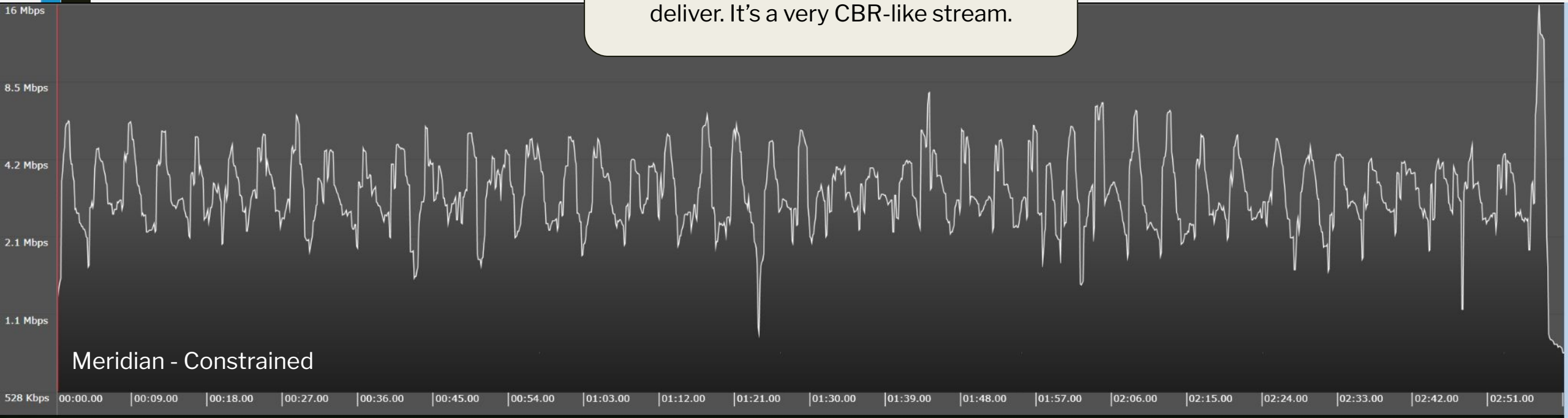
Meridian - Capped VBR/Constrained VBR



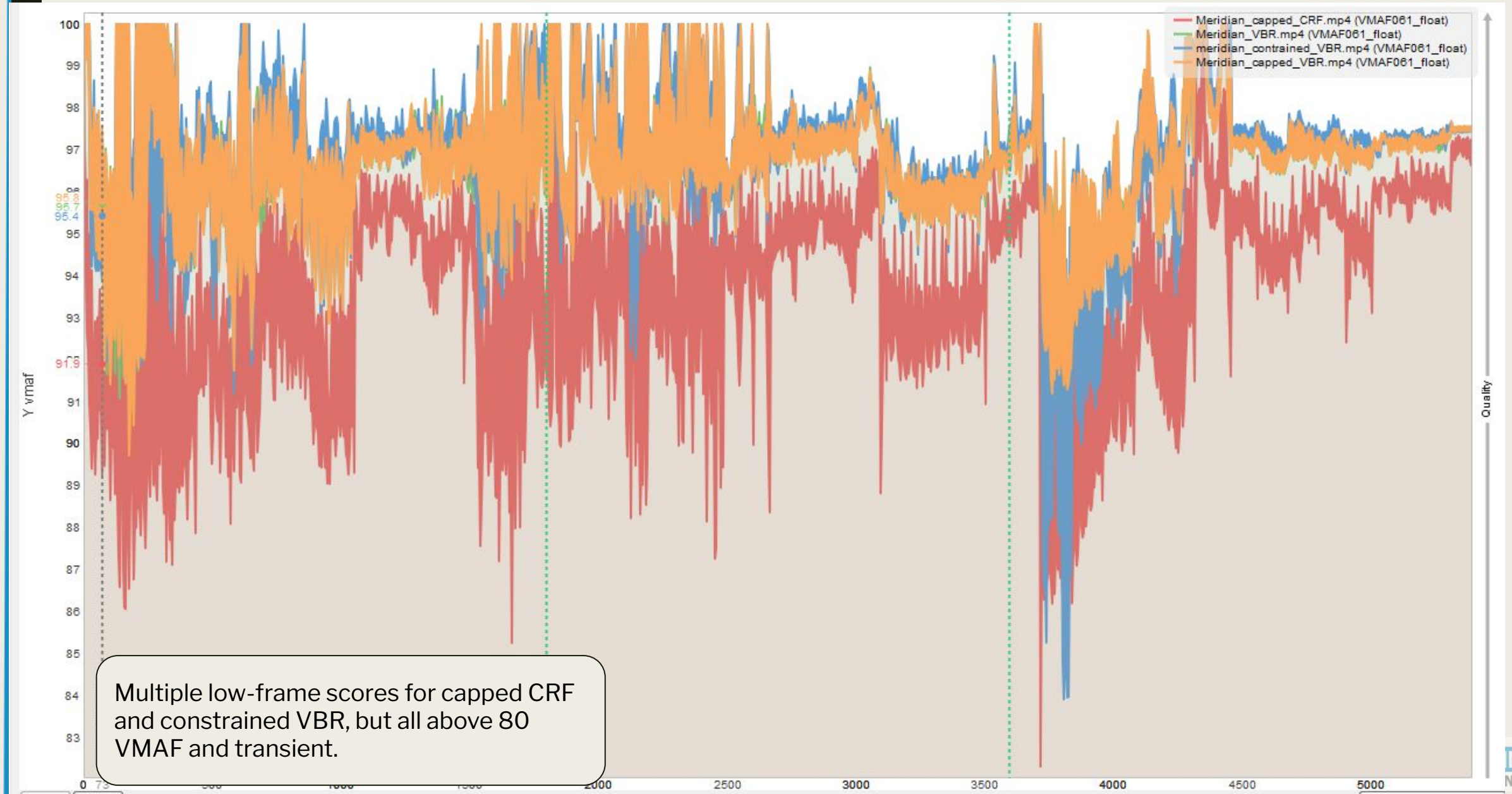
And here.



Here you see the consistency that constrained VBR (below) is supposed to deliver. It's a very CBR-like stream.

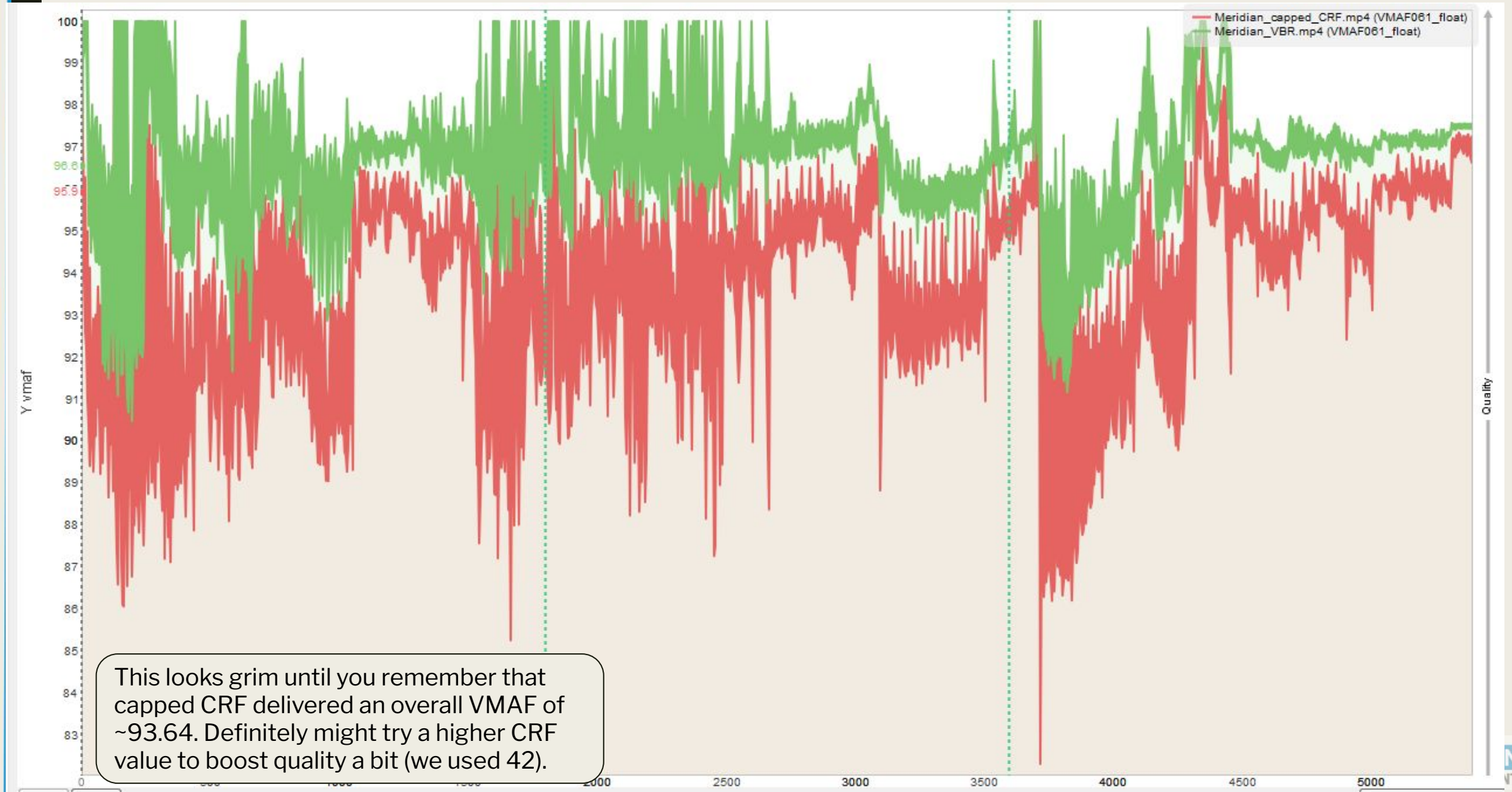


Meridian - All



Multiple low-frame scores for capped CRF and constrained VBR, but all above 80 VMAF and transient.

Meridian - Capped CRF/VBR



This looks grim until you remember that capped CRF delivered an overall VMAF of ~93.64. Definitely might try a higher CRF value to boost quality a bit (we used 42).

Meridian - Constrained / Capped VBR