Blazar Just-In-Time-Transcoder (JITT)

Optimized Cloud DVR Solution with Lowest Cost

Save Millions on Storage and CDN Distribution Costs

Overview

The surge in video streaming services and the increasing reliance on video in all web-related content necessitates the need for affordable, high-quality streaming technology. The varying usage scenarios and deployment situations greatly influence the streaming system costs. For instance, cloud DVR systems can place a substantial financial strain on storage, but the expenses associated with the CDN/edge distribution architecture shouldn't be overlooked either. Factors such as the application (DVR, VoD, Live), the quantity of channels or number of media assets stored, the user count, and the CDN or distribution network's architecture/bandwidth all contribute to shaping the system's design and cost centers.

While advancements in various technologies have boosted streaming performance and cost-efficiency, the deployment of streaming systems remains a considerable investment. Content-aware transcoding and just-in-time packaging (JITP) are two technologies that have positively impacted streaming deployments.

In this paper, we will introduce and discuss "Just-in-Time Transcoding" (JITT). We will explain how igolgi's Blazar JITT product sets a new standard for affordable video streaming deployment, eases generational codec transitions and enhances video quality and performance for the end-users in the process.

In addition, Blazar JITT reduces energy consumption per stream helping reduce carbon emissions. A separate white paper will be available soon on this topic.

ABR Streaming Background

A video streaming system employs a technique known as adaptive bitrate (ABR) encoding. The audio and video streams are first chopped up into small time chunks known as segments. These segments may last anywhere from 2 to 10 seconds(this is only a guideline, not a requirement). The file segments are then encoded into a variety of resolutions and bitrates, referred to as a "bitrate ladder". The audio signal is similarly divided into several audio channels and bitrates. The goal of ABR is to deliver the appropriate version of video and audio segments based on the capabilities of the receiving device and the available bandwidth. The entire streaming operation is thus converted to a series of file transfers(that can be supported by any Internet server) from the headend to the client devices. Rate selection is done by each individual client based on a manifest of rates provided by the ABR encoding head end.

Figure 1 shows a high-level view of an ABR streaming system with a cloud DVR. The video/audio content is transformed into various profiles and kept in storage. When a device asks for this content, the Origin Server initially sends a manifest to the requesting device detailing the available profiles. The device, in turn, selects the profile that best fits its characteristics (such as 4K TV, HDTV, smartphone, laptop, and so on) and the bandwidth it has at its disposal, then requests that specific profile. If there are fluctuations in the available bandwidth, the device has the flexibility to request different profiles as needed.

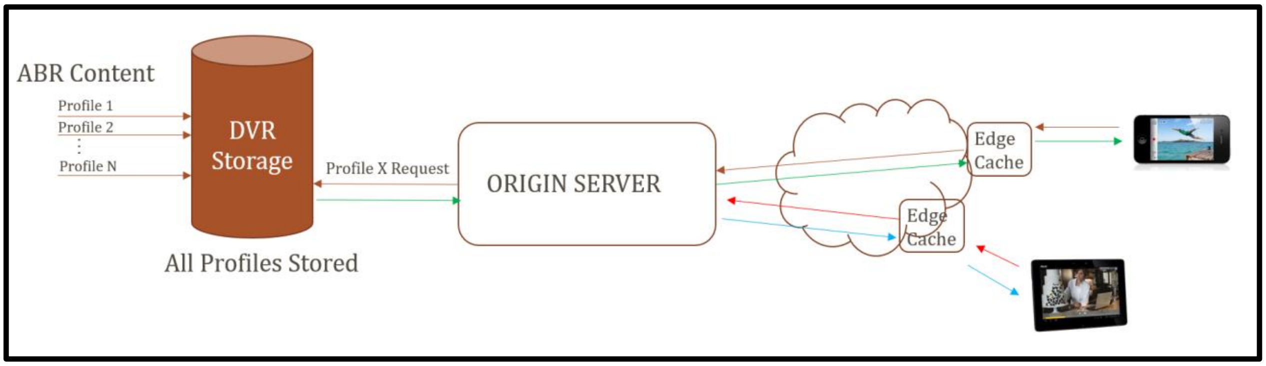

Figure 1: ABR video streaming system with a cloud DVR

As alluded to earlier, another key feature of this system is the division of content into brief time segments. The advantage of this is that the device receiving the content can easily switch between profiles by just requesting the next small segment from a different profile. Common packaging formats in use are HLS and DASH. The process of packaging these segments can happen at various stages in the network: before the storage, after storage in the Origin Server, or at the edge of the network. Each of these packaging architecture options has tradeoffs with storage size, core network bandwidth, and edge cache complexity and edge storage.

Just-In-Time Packaging (JITP) is an enhanced version of packaging that's particularly efficient when various digital rights management (DRM) or packaging formats are required for different types of devices. In this setup, the original video and audio signals are stored just once, but they are packaged in real time to satisfy the specific needs of the receiving device. This method significantly cuts down on storage and bandwidth costs.

Bitrate Ladders

Streaming systems typically have many different resolutions and bitrate profiles available so that each client can stream the content efficiently. In Figure 2, we show two example bitrate ladders. In the first Static Ladder, 8 different profiles are provided in typical resolution and bitrate combinations. Static Ladders are often used in live systems. With a large number of profiles, the endpoint has more choices to choose a profile that dynamically matches its bandwidth and device type requirements. The downside of this many profiles in the ladder is that the DVR storage requirement is high. This also may reduce the efficiency of the edge caches since more profiles may be stored at the edge.

	Static Ladder		Content Aware Ladder	
	resolution	bitrate (kbps)	resolution	bitrate (kbps)
	1080p	3900	1080p	2285
	1080p	3300		
	720p	2600	720p	1960
	540p	1800		
	540p	1200		
	360p	800	480p	860
	360p	650		
	270p	450	320p	510
Bandwidth Required (kbps)		14,700		5,615
Reduction in Bandwidth(%)			vs static	62%

Figure 2: Example of a Static ABR Ladder and a Content Aware Ladder

In the Content Aware Ladder system, a smaller number of profiles are generated using "content-aware" encoding, which creates unique profiles based on the specific requirements of the content. For instance, a high-action movie may need different combinations of resolutions and bitrates than a news program featuring people simply talking. Video-on-Demand (VoD) systems are good candidates to employ a content-aware ladder. While Digital Video Recorder (DVR) systems might also use content-aware encoding, they may require re-encoding from the live signals.

The advantage of an Automated (content-aware) ladder is that it uses fewer profiles. This means a DVR system would use less storage, and the bandwidth needed by the Content Delivery Network (CDN) to deliver a certain video quality to a specific client could also be marginally reduced compared to the static ladder. In the provided example, the content-aware ladder results in a 60% reduction in storage requirements.

In both the content aware ladder or static ladder, neither provides an optimal profile based on the end-client screen size or dynamically available bandwidth. Since both systems require transcoding to the ladder profiles and storing ahead of time in the DVR, the end client can only choose from a limited set of operating points.   The static ladder provides more operating points so gives more choices for bandwidth options but requires more storage. The content aware ladder uses less storage but does not give as many options for bandwidth which creates less efficient streaming for certain end point devices.

For example, if an end client had available 1400 Kbps, the static ladder could provide the 540p profile at 1200 Kbps, while the content aware could only provide 480p at 811 Kbps, so in this case the content aware system actually produced worse quality than the static system.

The result is that neither of these solutions are ideal.

Blazar JITT

Just-in-Time Transcoding (JITT) is a method where compressed video and audio signals are converted from one format to another at the precise moment when a request for the signal is made by a video or audio player (client). This necessitates that the transcoding system functions at a speed much faster than real time. Blazar, a JITT transcoder, possesses several unique characteristics, primarily its ability to transcode at a speed up to 100 times faster than real time. Blazar JITT is also dynamically scalable which allows for an array of new applications and use cases.

Blazar JITT fully leverages the latest developments in CPU and GPU technologies and makes optimal utilization of the various resources in both computational platforms. While JITT has been contemplated in the past, neither the computational power nor the SW architecture could achieve the types of consistent gains in speed for all the hybrid operations that Blazar is able to accomplish.

Blazar JITT Ladder

The most ideal bitrate ladder would be an "infinite" ladder, offering every possible profile. This would deliver the exact resolution and bitrate that aligns with the client device, the available bandwidth from the Content Delivery Network (CDN), and the type of content, thereby ensuring the highest video quality. This is precisely what the igolgi Blazar solution delivers.

In the Blazar system, only the topmost profile — the one with the highest resolution and bitrate — needs to be stored on the DVR storage system. When a client device requests the content, Blazar's Just-In-Time Transcoding (JITT) is capable of immediately converting the content to the best bitrate/resolution profile that suits the client's device type, available last-mile bandwidth, CDN bandwidth, and content type. The system architecture is shown in figure 3.

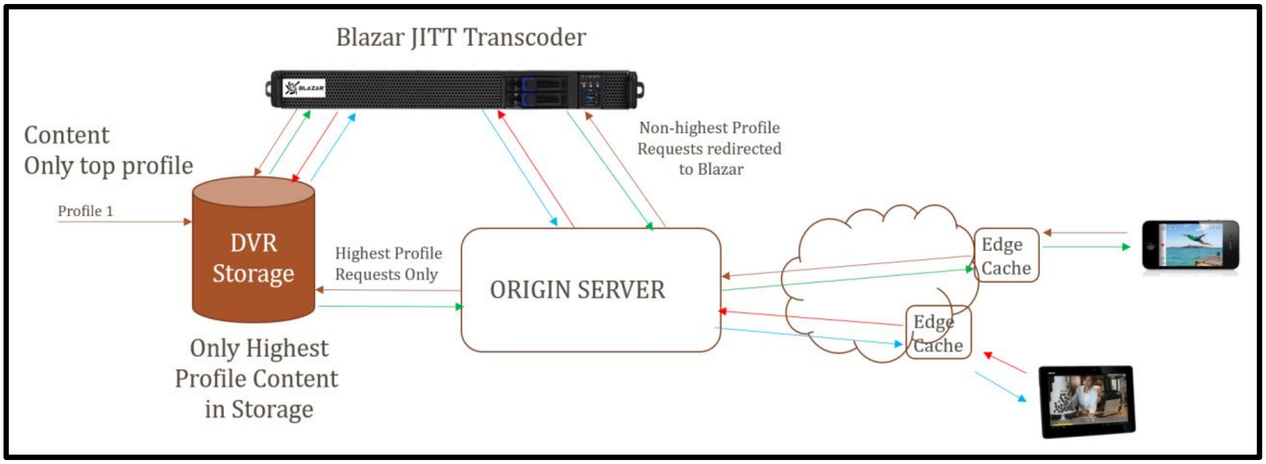

Figure 3: ABR system with Blazar JITT to provide lower profiles

When the end device has sufficient bandwidth for the top profile, then the Origin can serve that profile directly from the DVR storage. When the endpoint needs to use less bandwidth, then the Origin re-directs the request to Blazar JITT which retrieves the top profile from storage and instantly transcodes it to the lower profile needed.

In practice an infinite ladder would lead to an infinitely large manifest file that the client would need to parse. Therefore, a practical approximation of an infinite ladder is to construct a ladder with 20-50 profiles, yet only the top profile media is preserved in the DVR storage. The client can then select one from these 20 or more profiles and request it from the Origin server. Blazar is able to create the specific profile faster than real time. To the end device it appears as if the file was served from the storage server in the same amount of latency.

Figure 4 shows a typical example where the new Blazar infinite ladder saves **84%** storage compared to a static ladder, and **59%** compared to a content aware ladder. It should be noted that independent of the ladder type, Blazar can save the same amount of DVR storage since only the top profile is in storage.

	Static Ladder		Content Aware Ladder		Blazar Ladder	
	resolution	bitrate (kbps)	resolution	bitrate (kbps)	resolution	bitrate (kbps)
	1080p	3900	1080p	2285	1080p	2285
	1080p	3300			Any profile to match system needs	
	720p	2600	720p	1960		
	540p	1800				
	540p	1200				
	360p	800	480p	860		
	360p	650				
	270p	450	320p	510		
Bandwidth Required (kbps)		14,700		5,615		2,285
Reduction in Bandwidth(%)			vs static	62%	vs static	**84%**
					vs content aware	59%

Figure 4 Blazar ladder compared to standard static or content aware ladders

To put this into perspective for some use cases. A 500 channel cloud HD DVR storing all the channels for 1 year would require nearly 30 PetaBytes with the static ladder, 11.3 PetaBytes with the Content Aware Ladder, and 7.7 PetaBytes with the Blazar system that only stores the top profile.

Figure 5 shows an example of a 41 profile ladder with 50 Kbps increments between profiles. Using the example given earlier, where an endpoint device has 1400 Kbps available bandwidth, the Blazar ladder provides a near match at 1385 Kbps with 540p resolution. Whereas the content aware ladder has nothing near 1400 Kbps and must provide a lower resolution 480p at 811 Kbps. The static ladder has 1200 Kbps available at 540p. Worse quality than Blazar.

This example demonstrates both advantages of Blazar JITT:

1) Lower Storage and bandwidth cost: Blazar JITT only stores the top profile
2) Better video quality and optimum for the system available bandwidth.

Static Ladder		Content Aware Ladder			Blazar Ladder	
resolution	bitrate (kbps)	resolution	bitrate (kbps)		resolution	bitrate (kbps)
1080p	3900	1080p	2285	1	1080p	2285
1080p	3300			2	1080p	2235
720p	2600			3	1080p	2185
				4	1080p	2135
				5	1080p	2085
				6	1080p	2035
		720p	1960	7	720p	1985
				8	720p	1935
				9	720p	1885
540p	1800			10	720p	1835
				11	540p	1785
				12	540p	1735
				13	540p	1685
				14	540p	1635
				15	540p	1585
				16	540p	1535
				17	540p	1485
				18	540p	1435
				19	540p	1385
				20	540p	1335
				21	540p	1285
540p	1200			22	540p	1235
				23	540p	1185
				24	540p	1135
				25	540p	1085
				26	540p	1035
				27	540p	985
				28	540p	935
		480p	860	29	480p	885
360p	800			30	480p	835
				31	480p	785
				32	480p	735
360p	650			33	480p	685
				34	480p	635
				35	480p	585
		320p	510	36	320p	535
				37	320p	485
270p	450			38	270p	450
				39	270p	400
				40	270p	350
				41	270p	300

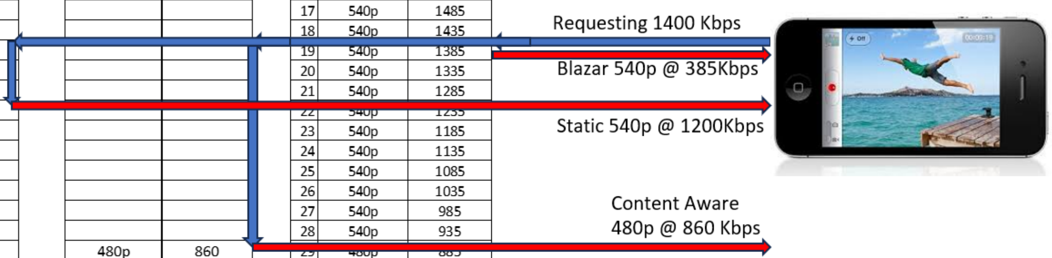

Figure 5: Example of a Blazar JITT virtual ladder vs traditional static and content aware ladders.

Blazar JITT Codecs

Another benefit Blazar JITT provides is an adaptable solution for the ever-evolving landscape of audio and video compression standards. During the codec transition phase, which typically lasts several years, both codec ladders need to be supported by the DVR for streaming to legacy and newer devices. In the past, this has always slowed the introduction of newer codecs in the marketplace. With Blazar JITT, the newer codec can be supported by Blazar JITT. When the number of client devices supporting the newer standard(which is usually more bandwidth efficient) reaches a critical mass, then the entire DVR library can be turned over to the newer codec and legacy devices are now supported with the same Blazar JITT. For instance, today H.264 is the most widely adopted video compression standard in client devices and set-top boxes, with newer devices increasingly integrating H.265 and AV1. The transcode conversion by Blazar would initially be from H.264 to HEVC/AV1. When the client population is more mature with HEVC or AV1, the library could be converted to that technology and all other devices would be supported by the Blazar JITT. This way, there is no requirement to pick technology winners early in the game as well.

H.265 and AV1 demonstrate bitrate efficiency surpassing H.264 by 50% or more. Consequently, employing these more efficient codecs when the end device supports them can result in significant savings in network bandwidth and edge cache storage.

Blazar broadens the concept of a bitrate ladder to encompass the video or audio compression format, effectively multiplying the effective ladder size by the number of codecs supported. A full Blazar ladder for every codec that is needed can be supported with still just a single codec and one top profile stored on the DVR storage.

This is not feasible with static or content aware ladders unless all the codec versions and associated ladders are stored which is an expensive proposition.

Blazar's dynamic ability to switch video codec formats offers several substantial benefits in terms of architecture and cost:

- When end devices support more efficient codecs, Blazar can perform Just-In-Time Transcoding (JITT) to these lower bitrate codecs. This can lead to bandwidth savings or an enhancement in video quality for the end device.
- The DVR storage can now maintain the highest profile in the most bitrate-efficient codec, with Blazar only transcoding to less efficient legacy codecs when an end device cannot support the newer ones. For instance, H.265 could be used for storing the highest profile in the cloud DVR. If an end device requires H.264, Blazar's JITT can transcode to H.264 for those specific devices, while transcoding to H.265 for devices that support it.

Blazar and Edge Caches

The storage and core network bandwidth savings are even higher when the edge caches are considered. In an Adaptive Bitrate (ABR) video streaming system, a Content Delivery Network (CDN) and edge caches work together to deliver a smooth, high-quality viewing experience.

The edge caches store the different video profiles (different bitrates and resolutions) that make up the ABR ladder for a particular video. When a client device requests a video, the edge cache can quickly provide the version of the video that's most appropriate for the device's current network conditions, thus minimizing buffering and maximizing quality.

The device doesn't just request and stick with one profile but dynamically adapts to changing network conditions. When network conditions change, the client device may switch to a different profile (either higher or lower quality), which the edge cache should also have available.

This is why it is beneficial for edge caches to store all the profiles of the ABR ladder rather than just the one that the client device initially requests. By storing all profiles, the edge cache can quickly respond to changes in the client's network conditions and provide the most appropriate video profile.

Of course, this might not be the case for all systems. In some systems with severe storage limitations, the edge cache might store only the most requested profiles. But such a setup could lead to more cache misses and poorer performance if the network conditions change, and the client device needs to switch to a profile that the cache doesn't have.

Blazar JITT at the edge can greatly optimize such a system. Figure 5 shows an ABR system with Blazar JITT used at the edge caches. In this case, the core network only has to store and forward the top profile to the edge caches for any video content. When an endpoint device requests a lesser profile, the blazar JITT can provide that on the fly, reducing the edge cache requirement to store all the profiles.

Therefore, the same storage savings as described before can be achieved at each edge point. Depending on the number of edge cache devices this can be a significant multiplying effect of edge cache storage savings. In addition, the network bandwidth from the origin to the edge has the equivalent percent savings since only the top profile needs to be pushed to the edge.

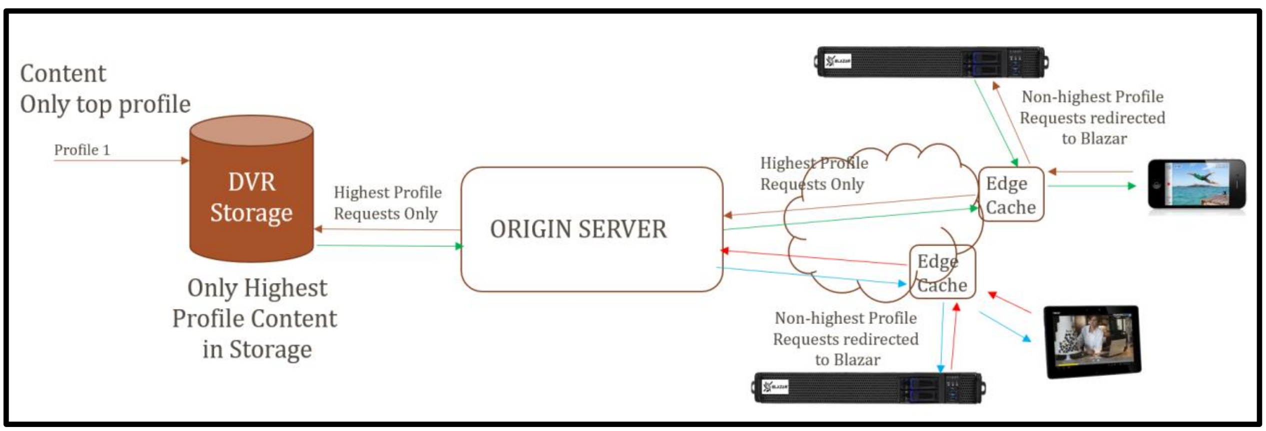

Figure 6 Blazar JITT architecture for edge caches

Figure 6 shows the ABR system with a Blazar JITT implemented at the edge cache. When an endpoint device requests the top profile, then the origin provides that top profile from the DVR storage and the edge cache can cache it for future use. When an endpoint client requests a lesser profile then the edge cache redirects the request to Blazar JITT which retrieves the top profile from the Origin and then transcodes it to the desired profile the endpoint device requests.

The benefits of this JITT edge approach are multifold

- Cache Hit Ratio will be twice or better with the same cache size. This is due to the cache only storing the top profile. (Alternatively the edge cache storage size can be reduced by twice or more with the same cache hit ratio).

- Network bandwidth from the edge cache back to the origin will be reduced by the effect of the edge cache hit ratio increasing. This reduces network capacity requirements on multiple components including switches, origin, and storage system.

Blazar JITT Design

igolgi's Blazar JITT transcoder is a software solution that runs on standard Linux servers with GPU compute accelerator hardware. Blazar is available as a software component that can run on any cloud provider such as Google Cloud, AWS, Azure, OCI etc., or is available as an on premise solution(sold as a software license or full turnkey server appliance). The core patent pending technology in Blazar are very efficient algorithms that manage compute resources in the CPUs and GPUs to perform a myriad of audio and video processing functions for JITT.

The Blazar Software stack is controlled from a simple but powerful API that can be integrated into any workflow. igolgi can customize the API and features for certain use cases or workflows if needed. The hardware can also be selected for optimum cost/performance tradeoffs depending on the system workload.

In some use cases, it makes sense to have Blazar JITT nodes be dynamically spun up or down depending on workloads (peak viewing times for example). This can most easily be done in cloud deployments, but Blazar nodes can also be used as VoD transcoding resources in an on premises deployment when not needed for Blazar functions.

For JITT operation, Blazar takes segmented top profile media files as input (usually fMP4 or MPEG-TS segments), and transcodes them fast enough for the end device playback to not be disturbed any more than if that end device was retrieving content from a storage server. A typical use case is 2 second media segments transcoded 10 times faster than real time which achieves this goal.

In some use cases more than 10x faster than real-time or greater transcode speed are needed and in those cases the Blazar hardware architecture must be scaled to support that.   This can be done statically with the hardware deployment up front, or dynamically by combining Blazar nodes together.

Cost and Benefit Analysis

There are many system aspects that impact the cost benefit of Blazar JITT solution compared to traditional solutions. For any given customer application and use case, it is important to do a cost benefit analysis for that specific circumstance. For example, the following variables have significant impact on the cost benefit of JITT in the use case of cloud DVR system:

=> Number of Channels (or Programs) to be recorded
=> Allocated number of DVR hours per user

=> Private Copy per User requirements depending on channel
=> Number of Users
=> Peak Number of Users and Average Number of Users of DVR channels
=> Percent of programs streamed that need a profile lower than the top profile
=> Codec and bitrate ladder design options
=> Edge cache and intermediate cache architecture sizes

These are just some of the parameters of importance. We present below a few example configurations which give a sense of the cost savings potential.

In the chart of Figure 7, we compare 2 architectures compared in terms of cost per number of subscribers:

1) DVR with 20 hours per user of private copy capability without Blazar JITT
2) DVR with 20 hours per user of private copy capability with Blazar JITT DVR

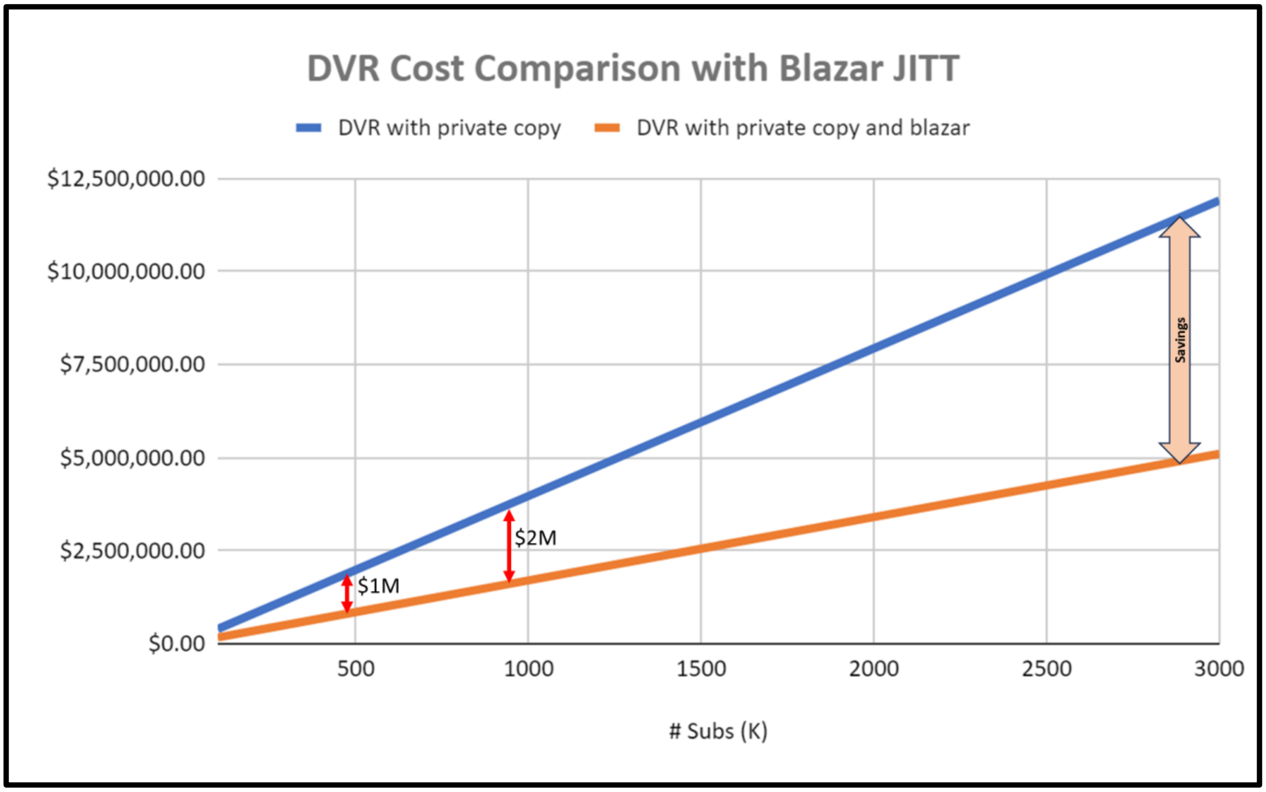

Figure 7 : DVR Cost Comparison with Blazar JITT

Assumptions

- Bit rate ladder as described in this document: 14.7 Mbps ladder, 2.3 Mbps top profile
- 60% peak viewers during prime time for 3 hours
- 5% of subscribers using a lower profile

- 30% of subscribers using the DVR
- $ 0.005 per GB/month (storage capital cost amortized)

The Blue line is the total cost for DVR storage needed for a static bitrate ladder with no Blazar JITT transcoding as a function of number of end users. The system supports 20 hours of private copy per end user. Note that most cloud DVR services offer 100s of hours of “private copy” per user but since many channels do not require the private copy requirement, the actual amount of individual private copy ends up being much lower, hence we used a 20 hour estimate.

As the number of users increases the amount of storage grows and the cost grows with it. Note we did not include other cost factors such as increased network bandwidth (switching/routing), power and cooling which would impact the total cost of ownership as the storage grows also.

The orange line represents total cost for the same DVR 20 hour private copy service, but only needing the top profile saved in storage, with an appropriate number of statically provisioned Blazar JIT transcoders added to provide the lower profiles on demand to end devices. As the number of users increases the cost advantage compared without Blazar JITT gets very significant. In this example, the Blazar JITT system annually saves over $1M at 450K subscribers and grows from there. If the Blazar JITT systems were dynamically spun up and down based on demand, the savings could be even higher.

In a second comparison illustrated in Figure 8, a cloud DVR service records all 500 channels continuously for either a year or two years. However, in this scenario, there's no need for a private copy. The storage size is fixed in this case to about $1.8M for 1 year or $3.5M for 2 years. The storage system may have to grow continuously but also many items can be removed over time as well. The light blue line is the yearly cost of such a DVR system which would be almost 30 Petabytes in size.

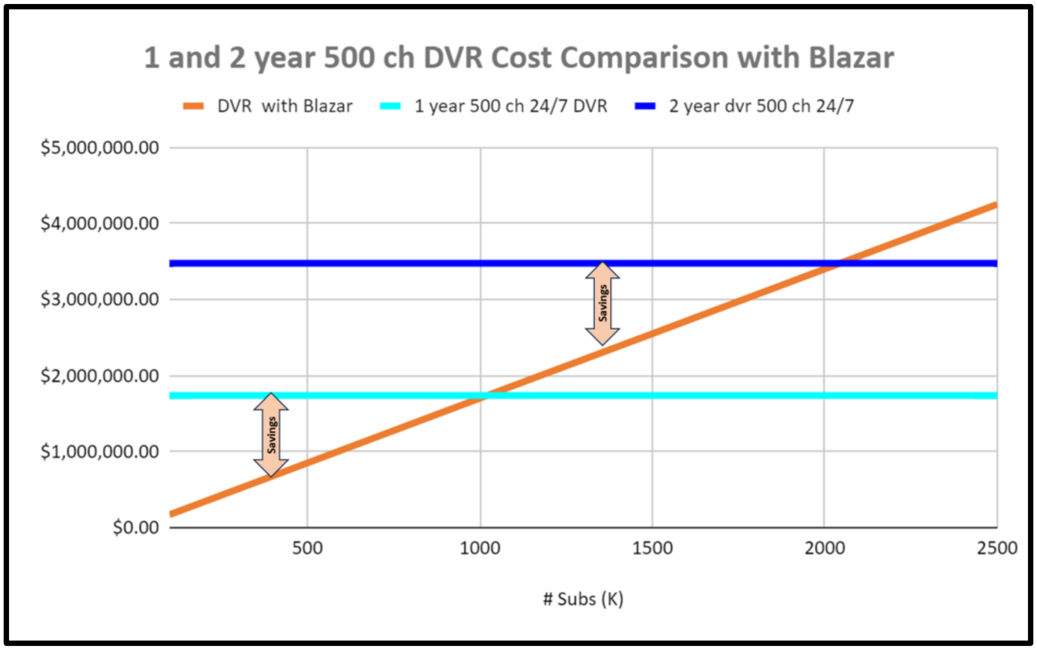

Figure 8 : One and Two Year 500 channel cloud DVR Cost Comparison with Blazar JITT

In this example, if private copies are not mandated, then storing all available 500 channels in ABR format will still be more expensive than a Blazar JITT solution for under a million users (for 1 year DVR) or for under two million users (for 2 year DVR). Smaller operators who get a private copy waiver with less than a million DVR subscribers thus save with the Blazar solution.

The savings with Blazar JITT go beyond just storage costs. The fixed DVR cost we have assumed does not include the extra networking, routing, power, and cooling required for such large storage systems. A one-year HD DVR storing all ABR profiles requires nearly 30 PetaBytes or twice that for a 2 year storage. These are huge storage systems that require a lot of networking and power. For example, a 30 Petabyte system is roughly 60U or more than a full rack in size. Compared to this, the Blazar JITT system occupies just 15RU including storage and JITT processing.

JITT and Green Initiatives:

ABR content distribution, while convenient, is very wasteful from a power perspective. All content is transcoded to many profiles and stored for long periods of time and may never be used. So, in addition to the transcode cost of creating all the profiles, there is also the added storage cost and distribution cost of maintaining these profiles all over the distribution chain. The concept of “create only what you consume” is a better paradigm from a power perspective and is a much greener option. A more detailed analysis of the power performance of JITT (with various hardware acceleration options) will be in a forthcoming white paper.

Conclusions:

Just in Time Transcoding for Adaptive Bit Rate distribution of content is now feasible at scale thanks to the availability of hardware acceleration options and Blazar software stack that efficiently uses the available resources. It is this combination that makes Just in Time Transcoding a reality. While there are a wide spectrum of use cases for JITT, we have detailed the DVR use case for operators in this white paper. We have demonstrated that Blazar JITT can be a cost-saving, power efficient solution for small operators who do not have private copy requirements. It is also a tremendous tool for large operators who do have a private copy restriction for DVR. Finally, we have also made an argument for Blazar JITT being a key enabler for new codec technology transitions which are usually impeded by the legacy problem. Blazar JITT can be efficiently deployed at the ORIGIN of a headend or at the edge depending on the operator's network architecture.

Other application use cases of Blazar JITT will extend to video CDNs and everywhere ABR traffic is generated and handled. Our future work will center around the economics of these use cases.