Cost Efficient Disaster Recovery

Sebastian Manemann, December 2021

"The best insurance is the one you'll never need"

Seeing your OTT Service falling apart is the worst thing that can happen to every provider, lucky those who implemented proper disaster recovery schemes at the right time. But to what extent does the recovery work? for how long are your services going to be affected? Are all channels covered equally?

And what will be the cost of your recovery scheme?

In this Whitepaper, you will see if and how the cloud can address some or all of the needs and how traditional software systems with license based business models might not be the best option to provide reasonable DR at a low cost without wasting a lot of idle CPU resources.

There is no other way to do that? - well, buckle up and continue reading...

What is a Disaster anyway?

What is the purpose of a disaster recovery system? We need to define the wording we want to use here first, in order not to mix up things. When talking to people about this topic, there is often a slight confusion between DR and redundancy mechanisms. A redundant system will make sure your service remains uninterrupted in the event one piece of your workflow goes crazy (e.g.

encoder outage). In that event, a secondary encoder (instance) will take over in the best case, and the service will not be affected.

What are the primary cause(s) of your Organizations largest outages?

■ 2018 ■ 2019
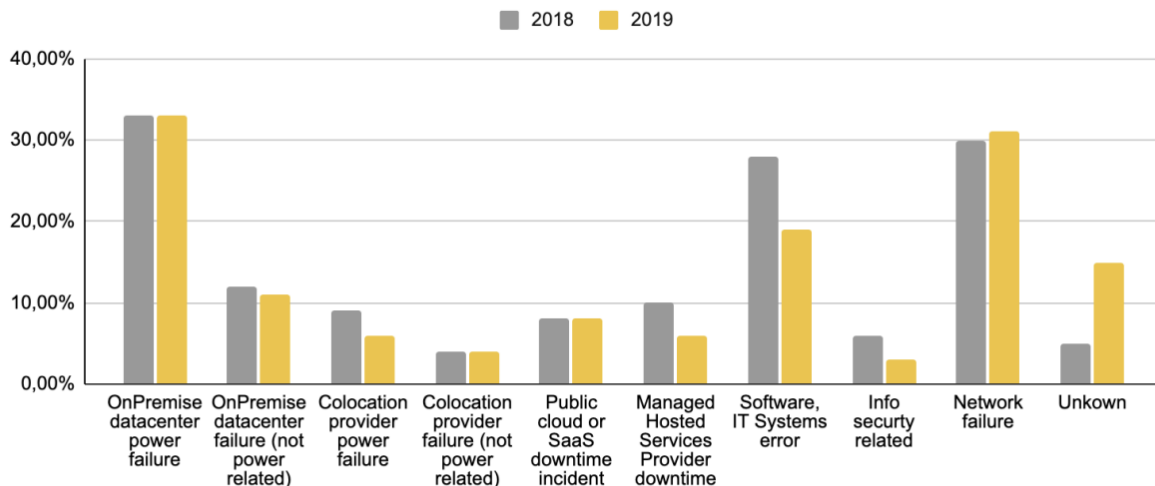
Source: Uptime Institute Global Survey of IT and Data Center Managers 2019

When talking about a Disaster, we are really talking about the complete fallout of a physical location , like a natural disaster or your favorite employee Homer Simpson setting it up for an explosion. That would mean that your entire workflow is gone in a single moment, there are no onsite redundancy mechanisms that can cover up for this.

Moreover, people have different understanding of what a typical recovery duration should be. While redundancy mechanisms (e.g. instance failure) are usually seamless, there are quite a few hurdles to ensure a zero downtime when switching from a facility to another (just think about synchronization for instance). Some will find 10mn to be acceptable, some will consider 10s as being too long, while others expect something seamless. Let's just say that everyone agrees this should be as short as possible.

Traditional Disaster Recovery Scheme
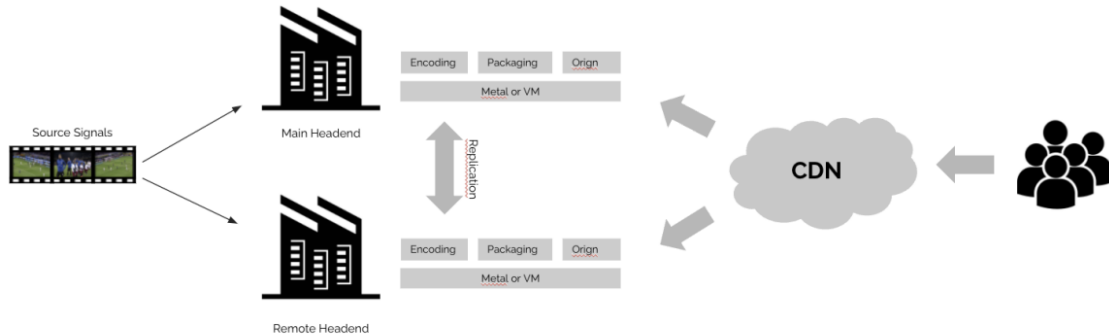
Going to the cloud to have a secondary physical location that is independent is the obvious choice for this scenario. One does not need to maintain the hardware actively or take care of any SLAs that need to be addressed (Other than the one with your cloud provider of choice obviously). Also, since cloud computing is meant to be pay as you go, theoretically it could be an affordable alternative. The Problem with this idea is that traditional Software implementations tend to fail in realizing this promise.

Disaster Recovery Scheme - Cloud based
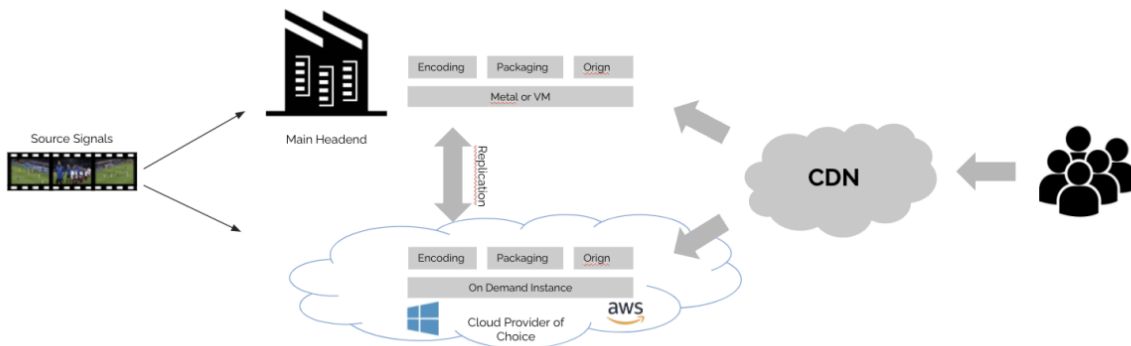
Software vs. Software in the Cloud

It is no secret that our beloved streaming industry is a little late to really embrace the usage of cloud capabilities. You have seen similar reluctance before when the move from hardware towards software on COTS Hardware was knocking at our doors. For sure, people are constantly talking about how they want to or have already moved their workloads towards cloud computing, but most of the time in the same sentence they will tell you how terribly expensive the cloud is. Looking at other industries, cloud computing is widely adopted with very little concern about the cost, as a matter of fact, in most cases there are huge savings associated with moving towards cloud.

Use of Cloud Computing in Enterprises as of 2020

Source: https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20210121-1
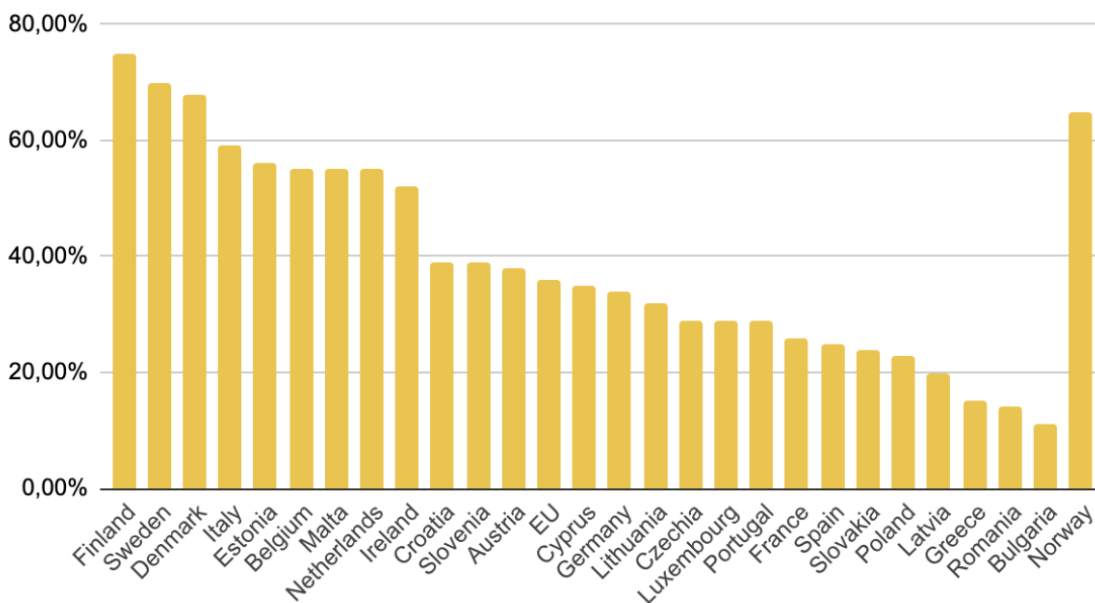
The underlying misconception here is that cloud is not about taking a piece of software and letting it run as a Virtual Machine on a remote hardware which belongs to someone else that happens to be a cloud provider. To do such "traditional" deployments, you are faced with two options:

1. The first one is to basically duplicate your complete infrastructure and likely duplicate the bill as well,

2. Or you can deploy upon requirement, which is likely to take hours if not days, not to mention possible license issues with changing IP/MAC or CPU identifiers.

I like to think of cloud computing as a platform that enables us to break up monolithic workloads into small services that do exactly what they are good at - not more, and not less. If you want to learn more about the microservice approach, I can recommend having a look at microservices.io.

Someone hand me a tool

There are tons of cloud services that are made to optimize the usage and efficiency of each and every of those microservices. One of them is called Kubernetes (or K8s - where the 8 stands for the amount of letters in between K and s - pretty nerdy stuff, huh?). Kubernetes is a management foundation that helps to automate the scheduling, provisioning, deprovisioning and maintenance of docker containers on a flexible set of resources. K8s is widely adopted throughout industries and cloud providers offer their own K8s services in their respective environments. Taking Elastic Kubernetes Service on AWS for Instance, you don't need to worry about setting up a Kubernetes cluster from scratch, you can just go to AWS and let them spin up a cluster according to your needs. The beauty of EKS is that it not only manages the Containers within the cluster, but also provisions and de-provisions resources according to the given workload. We published another article on that which describes the benefits of K8s for OTT very well.

Another great tool to use is Terraform, which enables its user to define their infrastructure as code, leaving the "heavy lifting" part of actually provisioning this infrastructure to the cloud provider. So what?

While taking monolithic software to cloud providers and hoping for the best seems to be a bad choice to make, using software that is built upon the principles of being "cloud-native" can help to scale and react in unprecedented short timeframes. Remembering the two options for traditional deployments ?

Should you choose to use Microservices, K8s and Terraform instead, hours can be turned into minutes,while the automatic scaling of Infrastructure can prevent costs for unused instances and spin up complete workflows on demand. Doesn't that sound like the perfect approach for a DR?

Alongside the technological disruption, there are also changing business rules when talking about the cloud. Providers offer computing resources without enforcing commitments and there is a spot market evolving around selling unused

resources for short timeframes. While monolithic software is relying on license based business models most of the time, the paradigm change in how software is built and operated is also changing the ways vendors engage with their customers. Highly flexible subscription/consumption based models are there all over the place and those models are a perfect fit for unpredictable events like our disaster recovery.

Just in Time Everything (JiTE) to the Rescue

Now that you know what options are on the table to build our DR site, who can help you leverage that? The good news: we at Quortex got you covered.

Quortex JiTE technology - Pull Mode architecture

Our patented Just-in-Time Everything technology is completely cloud native. All functions are operating as distinct stateless microservices managed by Kubernetes. We don't have any adherence between the functions and the underlying hardware, meaning that a function can do its job on any available piece of hardware and be moved to another piece instantly and seamlessly.

Therefore we are not only able to reverse the OTT workflow to be truly pull mode and only encode a chunk of video or audio when it is actually requested, we can also provision resources and spin up new channels in a snap. Leveraging this technology, the Quortex.io SaaS provides the matching business model and flexibility to this approach, since customers only get charged for what is actually processed and requested by the viewers, without any commitment or upfront fees. Even the CDN is included. Just configure your channel lineup, connect the SRT sources and rest assured that our technology will do the rest in the case of disaster.
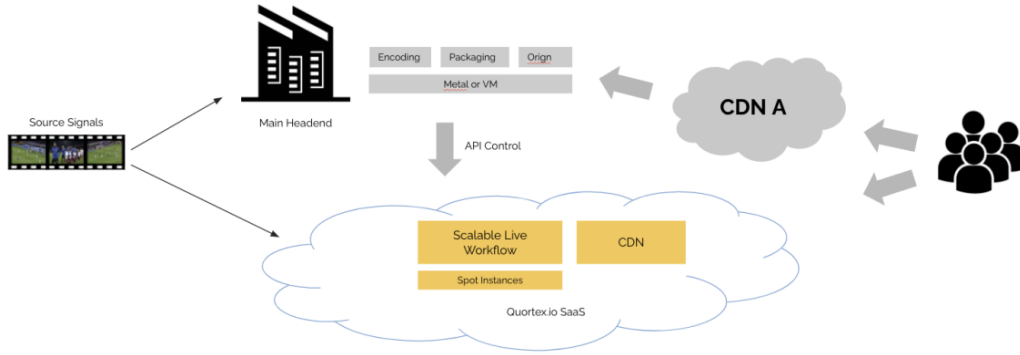
Disaster Recovery Scheme - Quortex.io based

"Come on down!" - The Price is Right

Lets say your OTT Platform runs 20 Channels and has a subscriber base of 20k Users. A Disaster Recovery site based on our Quortex.io SaaS will cost 971€/month when Idle, based on our Input Tier pricing.

Input			Processing			Delivery		
Total (GB)	Tiers (GB)	Price	Total (GB)	Tiers (GB)	Price	Total (GB)	Tiers (GB)	Price
100	100	0,050 €	100	100	0,587 €	10	10	0,035 €
500	400	0,040 €	500	400	0,505 €	100	90	0,030 €
2000	1500	0,030 €	2000	1500	0,385 €	1000	900	0,028 €
6000	4000	0,020 €	6000	4000	0,354 €	10000	9000	0,025 €
20000	14000	0,014 €	20000	14000	0,275 €	100000	90000	0,022 €
100000	80000	0,012 €	100000	80000	0,231 €	1000000	900000	0,019 €
200000	100000	0,011 €	200000	100000	0,195 €	5000000	4000000	0,018 €
	Above	0,010 €		Above	0,165 €		Above	0,015 €

Quortex.io consumption based pricing

Assuming each of your sources comes at 12 mbit/s, you'll ingest a total of 72,419,64 GB/month.

Your calculation should be:
(100*0.05€)+(400*0,04€)+(1500*0,03€)+(4000*0,02€)+(14000*0,014€)+ (52419,64*0,012€)= 971,04€

Remember that when the Platform remains unused, there is not only very limited cost, it also allocates almost no resources. Thanks to this fact, the carbon footprint of your DR systems remains minimal. If you are interested in reading more about how to reduce your services carbon footprint in general, have a look at the other White Paper called Mission: Emission.

To estimate the cost for the usage, we need to make some assumptions here, so imagine you have 50k subscribers and each profile is pulled at least one time (worst case scenario) with an average bitrate of 3.2mbit/s per user.

Let's say now the Disaster Event is happening and you need to run your channels for 6 hours on the DR, using Quortex.io the cost will be 7878 € for the whole 6 hours in the worst case, based on the three price tiers.

| Input	24,10€	Processing	341,59€	Delivery	7.536,79€

You can find some example calculations and how you can play around with the usage here.

One step beyond

Thinking about the JiTE approach and the previously mentioned cloud principles, why wouldn't you move your complete workload to a technology that is in itself redundant and self healing? Furthermore the system will scale according to your business needs, without any headache when adding new channels or trying popup channels for only a few days or weeks. A DR site might not be needed at all and you'll still be able to operate at a maximum level of security.

Every cloud provider operates their infrastructure on a topology of Zones and Regions. A (Geo-)Region can have multiple Zones in it , they can be utilized to spare workloads between physical data centers. If you want to take it one step further,

you can distribute your workload over different Regions, which will incur a higher cost, since traffic between regions will be charged. AWS for instance provides around 20 regions with 61 availability zones currently. If you want to go absolutely bulletproof, you can even build multi cloud clusters that will leverage resources from different cloud providers in different regions in different zones. I truly doubt that any private data center on earth will be that resilient.
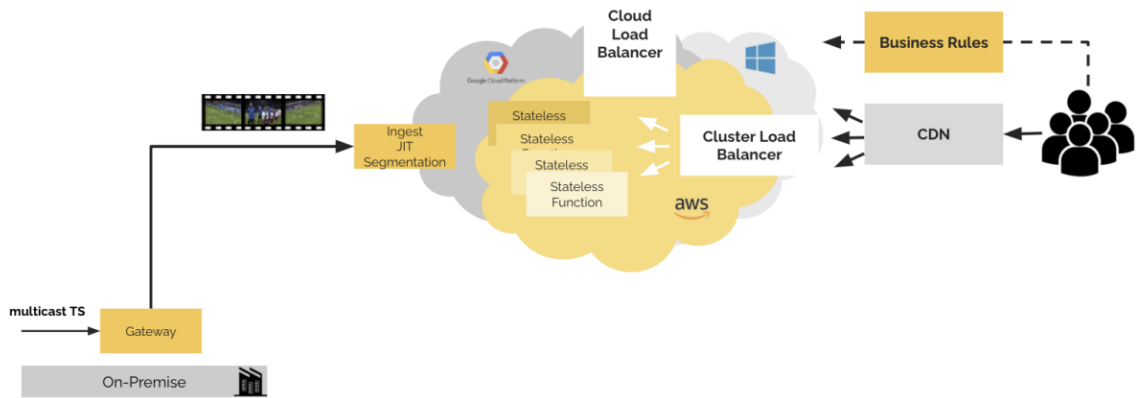
OTT Workflow using Quortex tech stack - multi cloud/multi AZ

Clearly, there is some work to be done to deploy and maintain such a cluster, and inter-cloud traffic is not the cheapest thing on earth, but still it might be worth giving it some thought since permanently provisioning instances and licenses that are doing literally nothing is no solid business case at all and is especially unhelpful when thinking about the environmental impact of the increased carbon footprint.

At Quortex we are able to provide all levels of engagement for the disaster recovery or the full headend. Starting with our self sign up SaaS Platform , over managed service JiTE technology stack all the way to customer managed deployments, we are there to help you build the most robust OTT Environments.