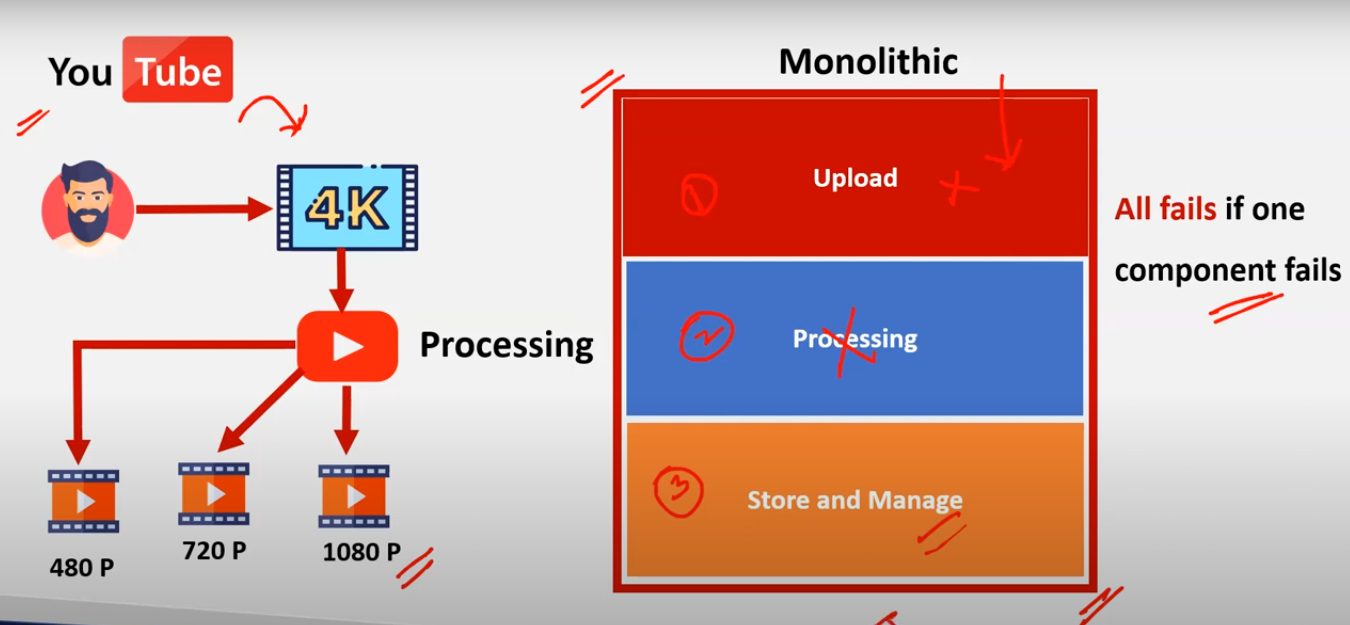
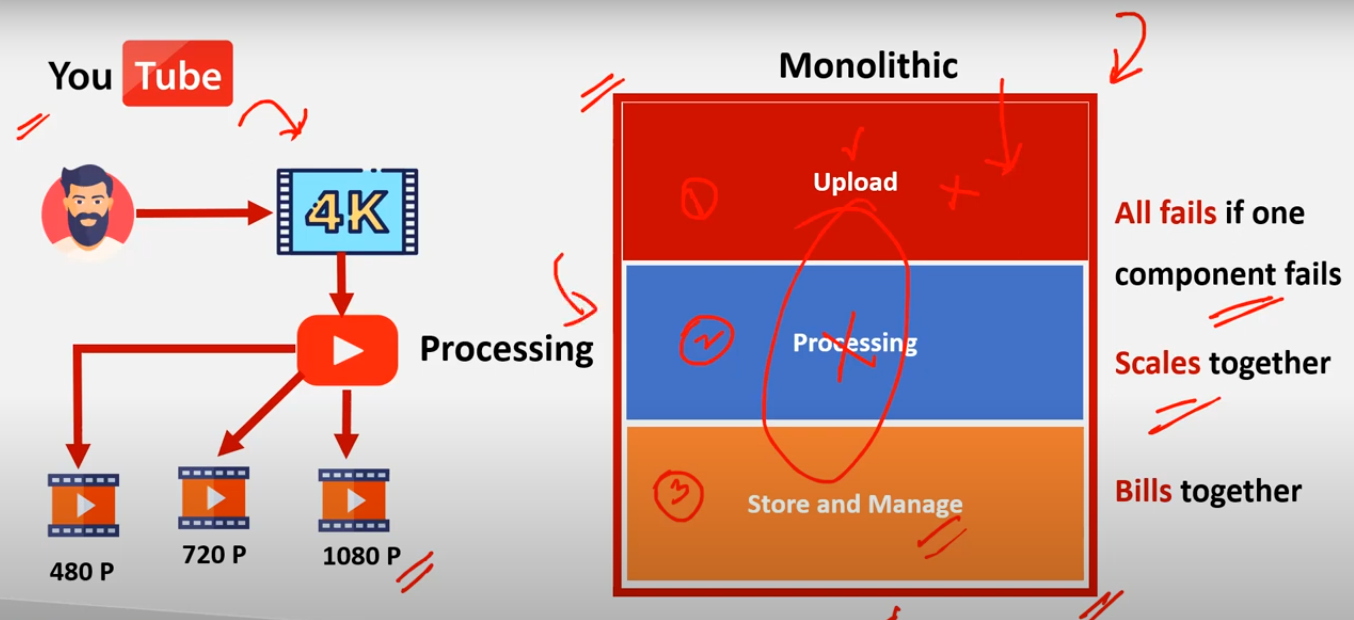


When u r uploading vedio in youtube next process is processing and then store and manage comes into picture

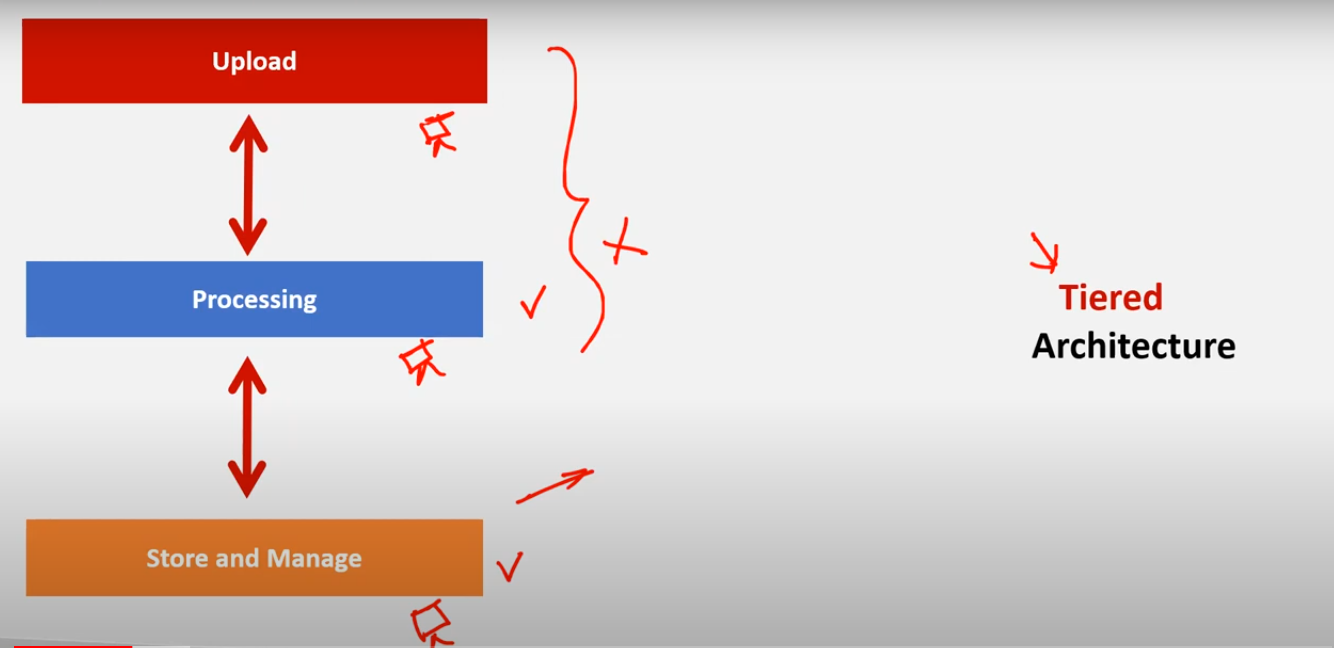
So there are 3 modules in when we keep 3 modules in single server its called monolithic



If one module fails all fails

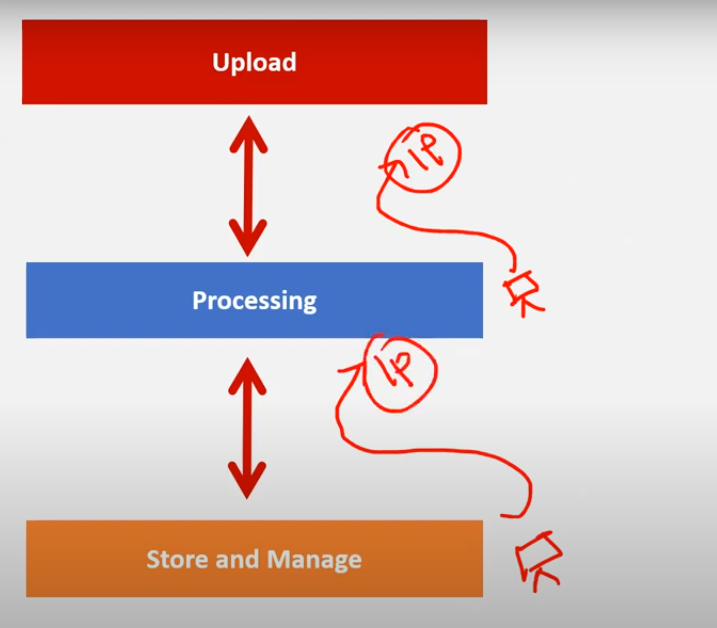


To avoid monolithic architecture we have implemented tired arhitecture

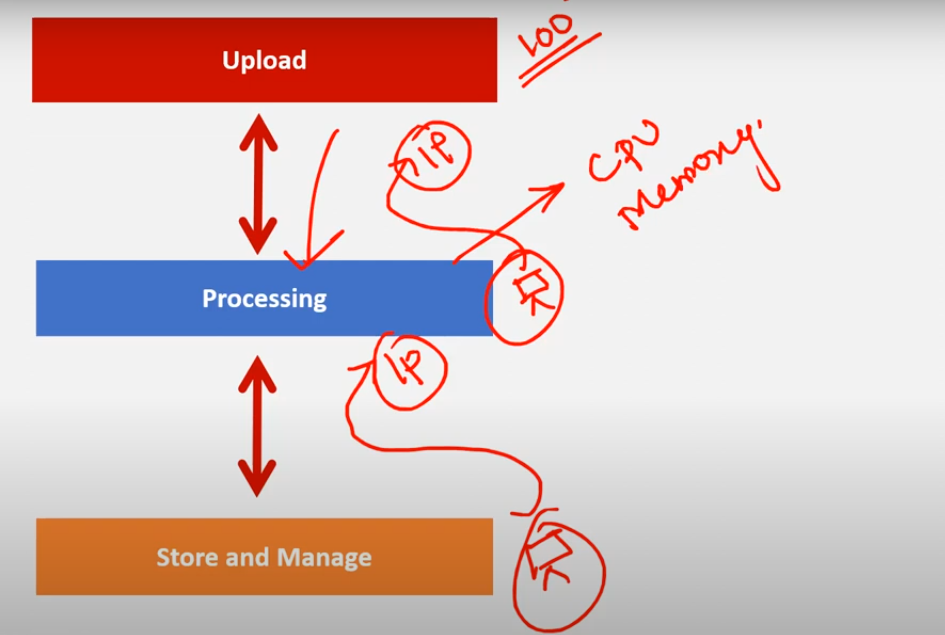


To avoid this issue we have implemented tiered architecture

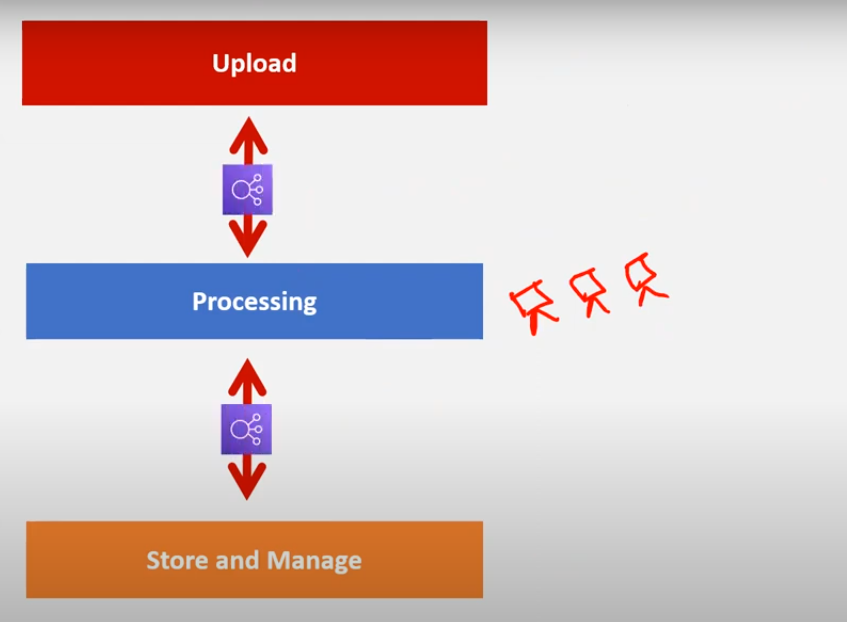
* If one server fails other two will be running



If you want to connect to the processing server u need to mention that processing server ip in upload server

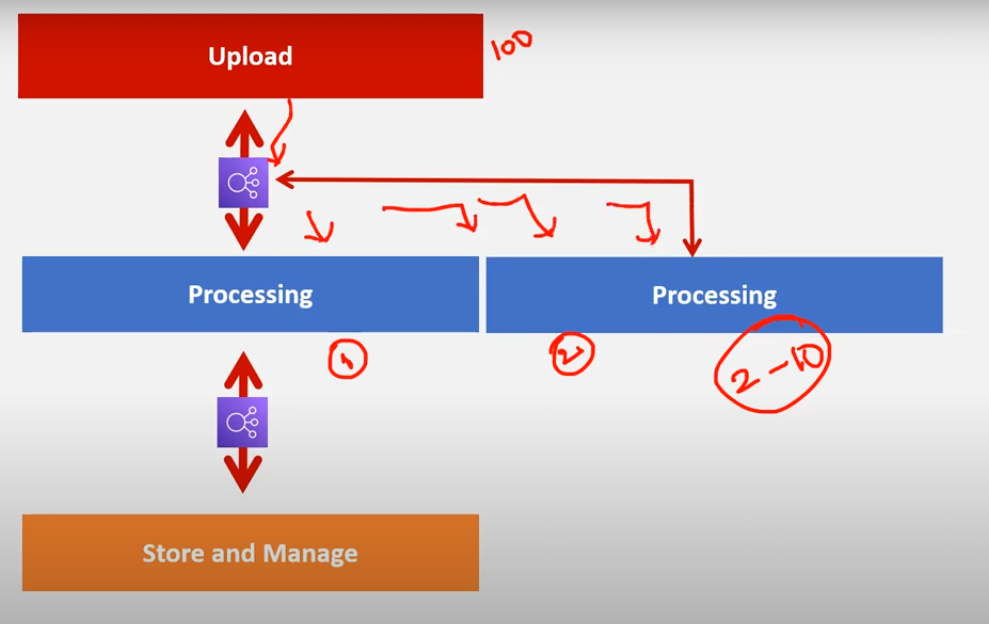


If you keep 100 uploaded vedioes and hit gets hit to processing section at once so cpu and memory of processing server will be overloaded



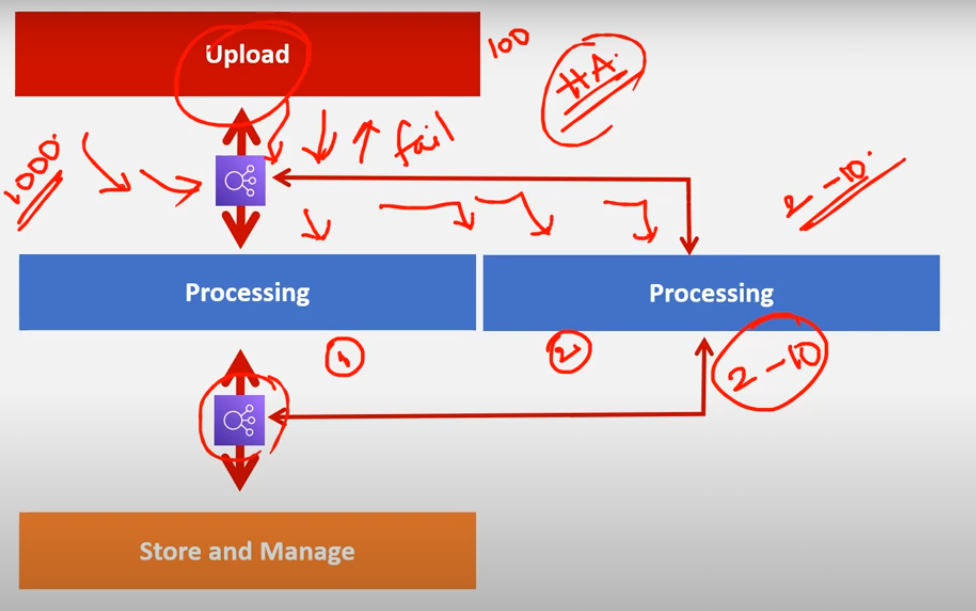
To overcome that issue we have created loadbalancer and will keep that load balancer url in upload section but still problem arises as there is only 1 server

So if we create autoscaling the problem gets resolved

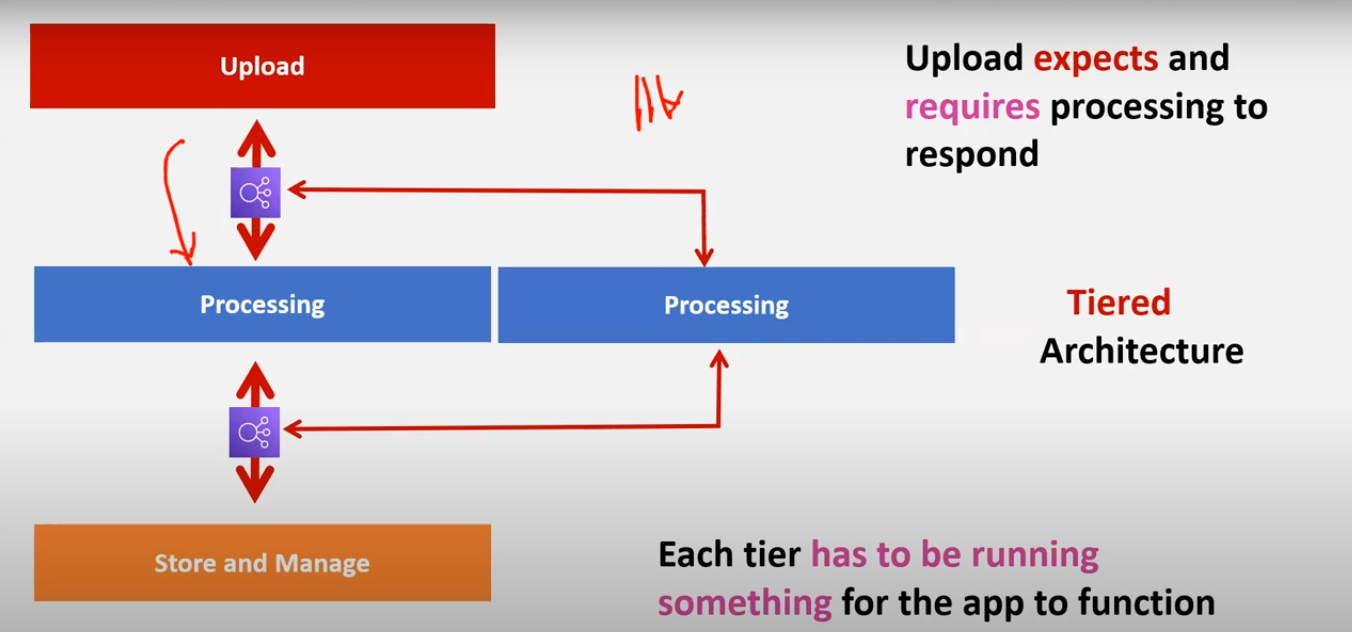


If load gets increased thn autoscaling servers manages the servers according to the load

Keeping high availability with the help of load balancer and auto scaling

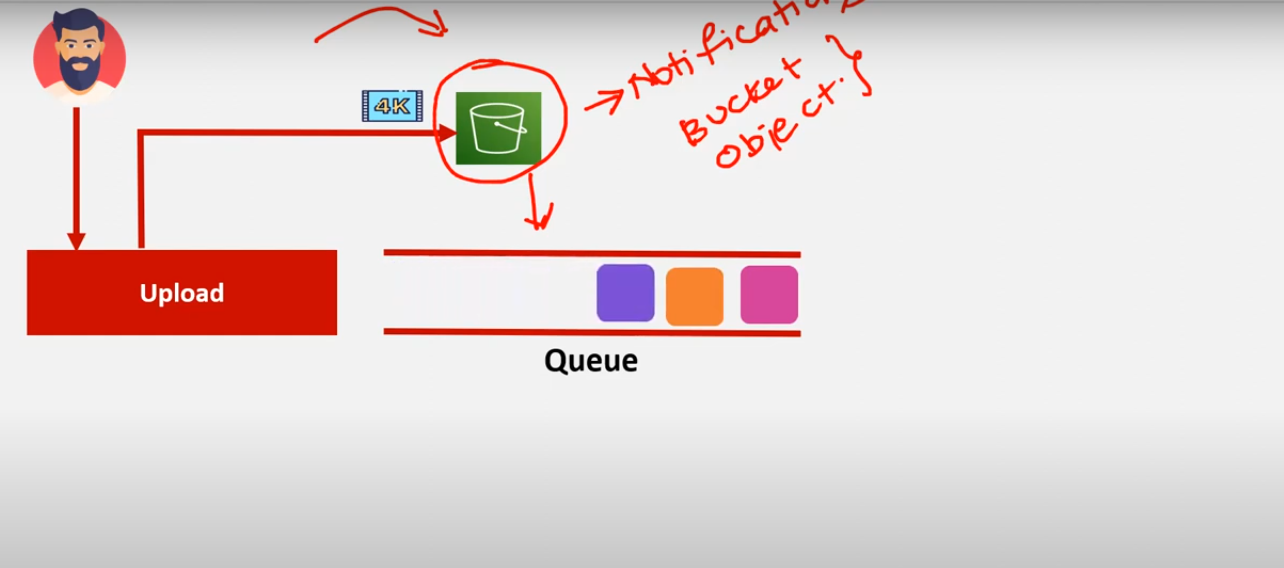


But still keeping high availability if there are 1000 vedioes to upload in autoscaling we declared 2-10 servers to launch if still that gets overloaded again our upload system gets fails



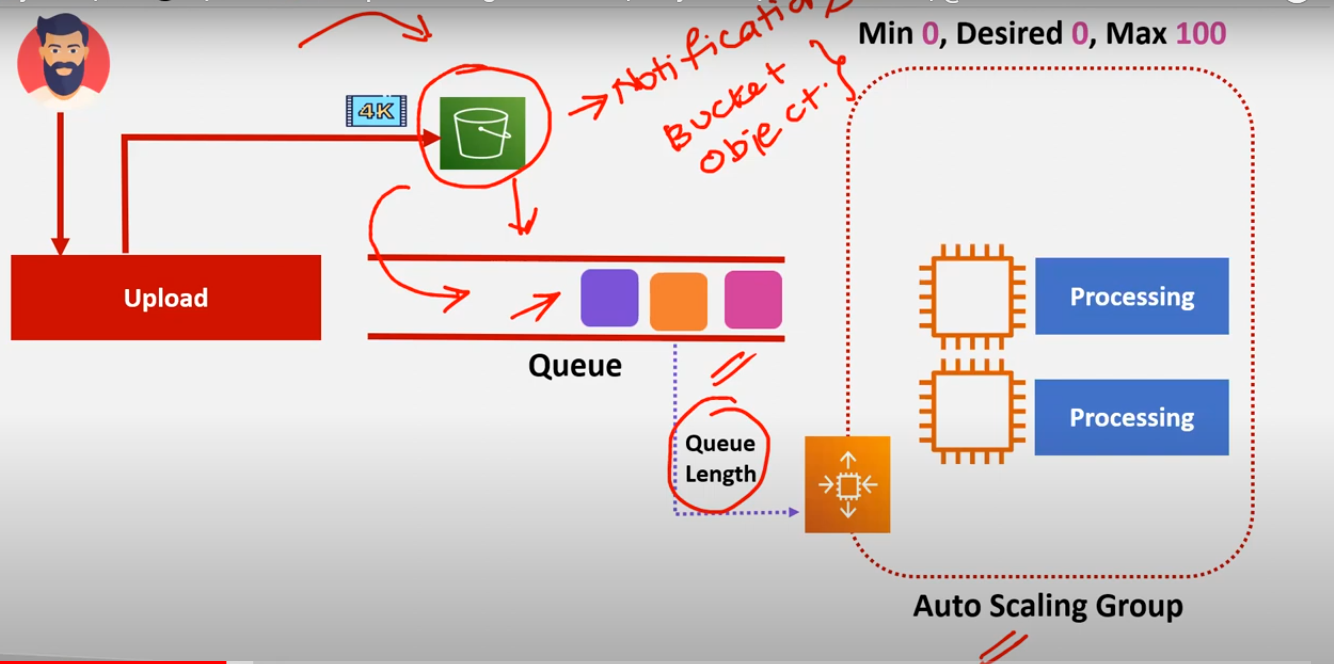
There is no control overthis

To overcome this issue we have Event driven architecture



In event driven architecture we wont directly upload our files to servers instead of that we upload those vedioes in s3 bucket and will integrate s3 to queue

* And by using s3 notification service we send metadata to queue metadata is not entire file but information about the file/object



Integrate queue to auto scaling group so based on the queue length servers will be launched