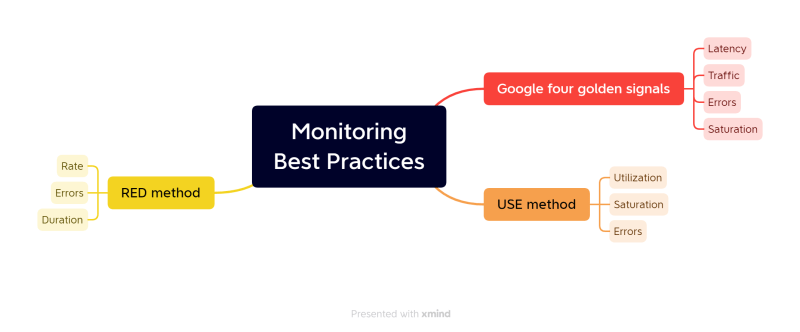
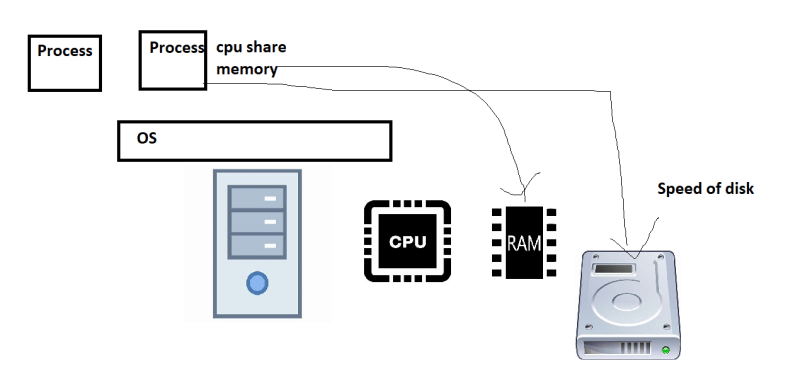
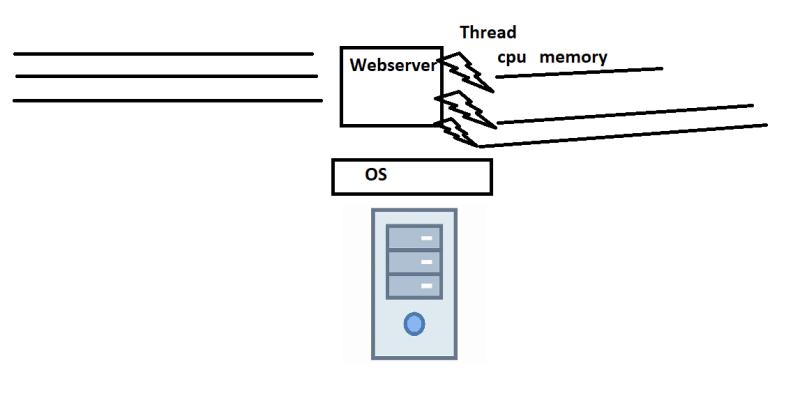
**What has to be monitored?**

* There are organizations and individuals who have published the best practices on implementing a monitoring solution
  + Google Four Golden signals <https://sre.google/sre-book/monitoring-distributed-systems/>
  + USE method <https://www.brendangregg.com/usemethod.html>
  + RED method <https://www.weave.works/blog/the-red-method-key-metrics-for-microservices-architecture/>   
    

**Terms in Monitoring**

* Latency
* Traffic
* Errors
* Saturation = [(Total-request-success-request)/Total-request] \*100}

**Some basic Stuff**

* Impact of CPU, Memory and DISK on your applications  
  
* Webserver: When requests are sent, threads are created which will have its own cpu and memory share. So as number of requests increase the load on cpu and memory increases.  
  
* Generally, to figure out the saturation points, organizations stress/load the systems with the help of performance test engineers.