

Horizontal Vs Vertical Scaling

We have a system that serves for the incoming requests and give response, it can be a server or a cloud system that can take in request and respond. If we have case when the incoming requests bump up and we are not in a stage of responding to all the incoming requests, here our system will beak. To over come this outage, we can have add more machines to our system or we can increase the size of the existing machines.

The ability to handle more requests is called **Scalability**. Two Scalability Techniques are :

- **Vertical Scaling** : The technique in which we increase the size of the existing machines.
- **Horizontal Scaling** : The technique in which we increase the number of machines to handle the incoming requests.

| Horizontal Scaling | Vertical Scaling |
|--------------------------------------------------------------------------------------------------------------------------|------------------------------|
| Machines : 1 2 3 4 5 | Single Machine |
| Need a Load Balancer | Doesn't need a Load Balancer |
| Resilient [If one goes down another takes it's place] | Single Point of Failure |
| Network calls between the each servers of the network. (RPC: Remote Procedure calls) | Inter Process Communications |
| Data Inconsistance. [In case of transaction where we sends data from one server to other, we have to lock the servers] | Data is consistent. |
| Scales well as user increases | Hardware Limit |