Cloud Scalability is the ability to scale on-demand the facilities and services as and when they are required by the user.

Cloud Fault Tolerance is tolerating the faults by the cloud that are done by mistake by the user.
Here the scaling is beyond the limits, it means we can't even imagine what will be the limit. loud middleware is designed on the principle of scalability along different dimensions in mind e.g.:- performance, size and load.

The cloud middleware manages a huge number of resources and users which depends on the cloud to obtain that they can't obtain within the premises without affording the administrative and maintenance costs.

So in this overall scenario the ability to tolerate failure is normal but sometimes it becomes more important than providing an efficient & optimized system.

The overall conclusion says that "it is a challenging task for the cloud providers to develop such high scalable and fault tolerance systems who can get managed and at the same time they will provide a competitive performance.