+ YARN’s ApplicationMasters provides tasks to Containers and each task status (i.e. MapReduce, Spark etc) is submitted by container to respective Application Master.

+ Containers executes their allocated tasks using CPU, Memory and FileSystems etc which are underlying within it. Also Containers may step on each other’s feet by stealing CPU, Meemory or File Resources.

+ YARN allows HADOOP system to execute various data processing programing paradigms on existing cluster by interacting between “Resource Manager” as a Master and each “Node Manager” as a slaves in entire cluster. Currently this needs each and every required supported files need to be installed based on the specific processing programing paradigm’s Job. This puts entire Hadoop System cluster in complex to manage as well as Administrating.

+ When tasks are run in Docker containers, they can have their own file system, CPU and Memory limitations (Which is user configurable) and therefore users can install whichever version of software that they need for their tasks in a container. They can then ship the containers to the NodeManagers. As a result, their tasks are guaranteed to run without interference.

+ We can manage User access levels too as privileged user using DefaultContainerExecutor (as a YARN user) for running tasks and Non-privileged user by LinuxContainerExecutor.