TEE Distributed Provisioning Relay-00

IETF 117
TEEP Meeting
Big data computing framework like Spark, MESOS, YARN always uses distributed architecture to arrange applications and computing resource.

Master node runs the main part of application and dispatches tasks into Worker nodes.

Worker node runs tasks in Executor.

Tasks are the real place that big data is processed. And tasks could change dynamically during the process stage of application.
Question: When using TEE cluster to run big data frameworks like Spark, how to attest this distributed system during runtime.

- Executor in Worker node runs the tasks that distributed by application.main
- Tasks could be updated during the running stage of this application.

=> Remote attestation of distributed framework is a dynamic action, different stage has different reference value.
Resolution

- Attest Master node first. Master runs the application.main, reference value of which is stable during runtime.

- Using Master as the “transfer station” to relay tasks information of Workers back to Verifier/Relying Party.

- Verifier/Relying Party using this relay information to generate reference value of different stage.
Relay message Type

There are three types of Relay messages generated by Master

- Task Serialized
  Tasks that generated by application.main in Master, and will be serialized by Relay message

- Worker Hardware Manifest
  Hardware information of Worker node, like TEE architecture, memory setting, etc.

- Worker Software Manifest
  System and software information of Worker node and Executor, like OS version, system software list, etc.
Next Step

There are some further steps of this idea:

- Verify the hypothesis of the design
- Specify the message context
- Align with SUIT and other documents
Thanks