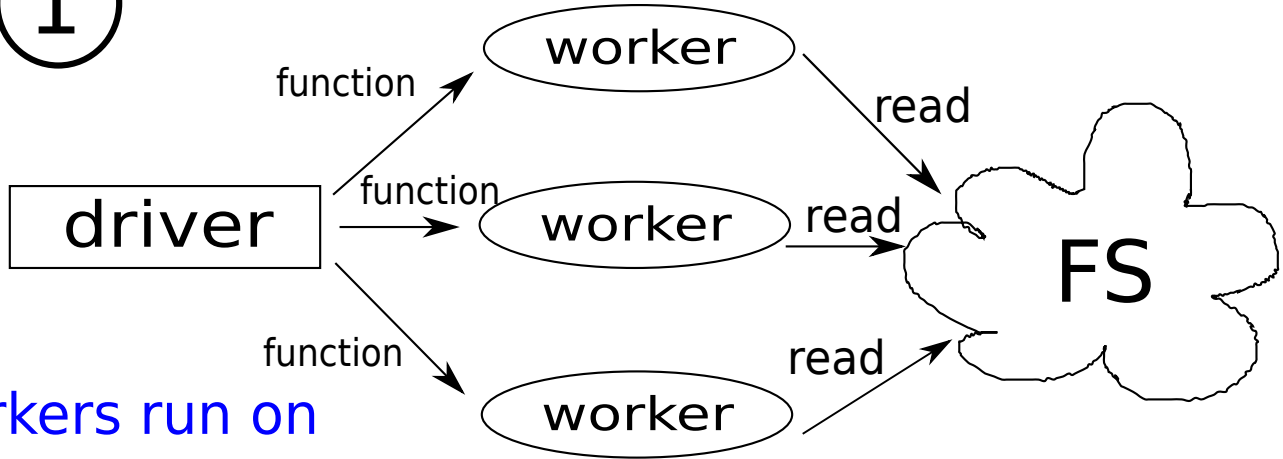


RDDs are distributed
across workers

①



workers run on
cluster nodes

②

