HDFS Fault-Tolerance

Thursday, January 26, 2023   3:22 PM

- DFS Failures
  - Common Failures
    - Entire Node (HW)   **K-safety**
    - Partial HW (Non-Storage)   **K-safety, node-down**
    - Storage HW   **RAID, hot-swap**
    - Memory/Data Corruption   **ECC, checksums**
    - OS Error   **monitoring, disy/restart, node-down**
    - Middleware Error   **hdfs/hadoop, log/debug**

  - HDFS Failures
    - Data Node
      - Heartbeat
      - Mark as "dead" on lost heartbeat
          ⤷ re-replicate affected blocks

    - Name Node
      - Persist state of Meta Data
        - fsimage   } **RDBMS**
        - edit log   }

      - Manual Failover (HA)
        - Primary (Hot)   ✗
        - Secondary (Cold)   ✓   **run Failover**

        ✳ Hot-Hot / Hot-Warm (Multi-Master)

      ✳ ==Microservice:== state replicated across (mm 2)
                     Name Nodes via HTTP/REST

      ✳ ==Zookeeper:== Maintains Primary/Secondary
                     failover

  ✳ Usage: CLI

      hdfs dfs   ← old
      hadoop fs   ← new (+object store)
                          Azure, S3 ...