Team Chaitan

Chaitan Baru, Sanjeev Kumar, Venkat L., Anil Majumdar, Florian Stegmaier, Vinayak Borkar

Topic 1: Process

- Open source approach of benchmark.
- MOB -Massively Open benchmark.
- Controlling group: Core group could have advisory groups (possibly vertical oriented like healthcare).
- Four different communities:
  - academics,
  - vendors,
  - system integrators
  - customers

Topic 2: Benchmark Availability

- Every six months, 2 things happen
  - New version of benchmark released
  - Official BigData 100 list is updated
- Adoption cycle can be different
  - When a new release of the product is out
  - Adopter needs to implement the latest version of the spec to be on the list
- One idea: Use the W3C Working Group Model for releases
  - Working Draft -> Recommendation

Topic 3: Reporting Requirements

- Aggregate Systemic properties (compute, latency, bandwidth, IOPS) -- CAPEX 2
  - As reported by well defined queries in the benchmark
- Scale Factor for Data
  - in turn could determine the hardware configuration (# of nodes in the cluster)
- Workload
  - Real time to insert and extract data
  - "Uber" batch to execute a job (different workloads and response to run) for structured, semi structured and unstructured)

Topic 4: Audit?

- YES!
- Who pays?: Vendor if it is for profit, free for open source.