BROADBAND WIRELESS NETWORKING IN THE ERA OF BIG DATA

Presented by: Dr. Tamer Omar

Colleage of Enfineering & Technology

Technology Systems Departmet

East Carolina University

INTRODUCTION

- Organizations accumulate huge amounts of data from various systems, however more often the data is stored but not organized or analyzed by these organizations.
- Mobile service providers (MSPs) in their efforts to provide more efficient networks deal daily with huge amount of signaling data characterized by the same features of big data.
- The successful implementation of a big data system (BDS) involves having the required infrastructure in place to process the data.
- The four main characteristics of Big data are
  - High Velocity
  - High Volume
  - High Variety
  - High Veracity

THIS RESEARCH PROPOSES THE UTILIZATION OF BDS TO DETERMINE IF IT BRINGS

A **VALUE** TO MSPS AND THEIR CUSTOMERS.

INTRODUCTION

- This research aims at
  - Designing
  - Implementing
  - Operating

both a wireless 4G het-net and a big data system (BDS) in a testbed platform.

- The purpose of this research is to implement a 4G heterogeneous network (het-net) that depends on a big data infrastructure in order to:
  - Collect
  - Organize
  - Analyze

the network performance.

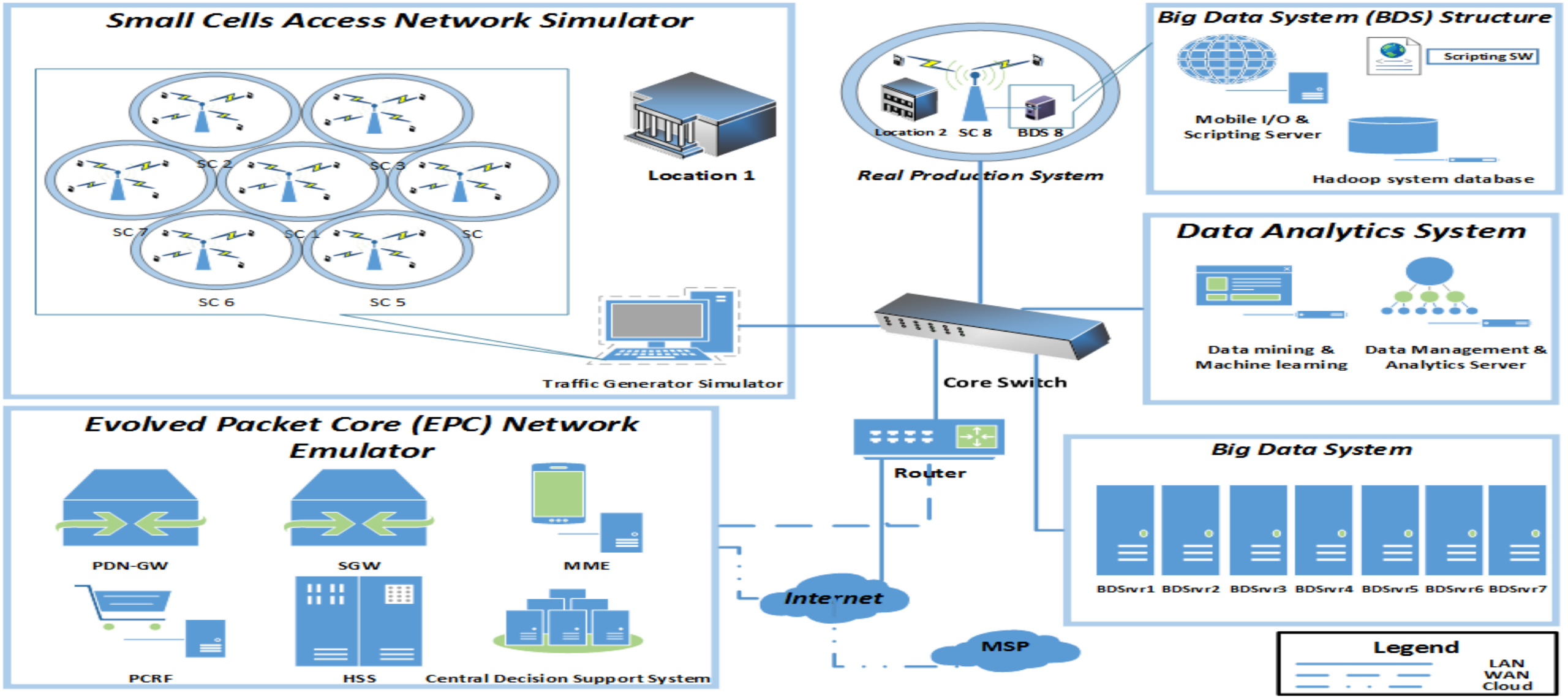
- **Objective:** The analysis results will be used to tune the network parameters in order to enhance the self-organized network (SON) self-healing functions.

DESIGN

SYSTEM ARCHITECTURE

- The proposed architectures is composed of the following four layers:
  1. Small Cells Access Network Simulator
  2. EPC Network Emulator
  3. Big Data System
  4. Data Analytics System

SYSTEM MODEL

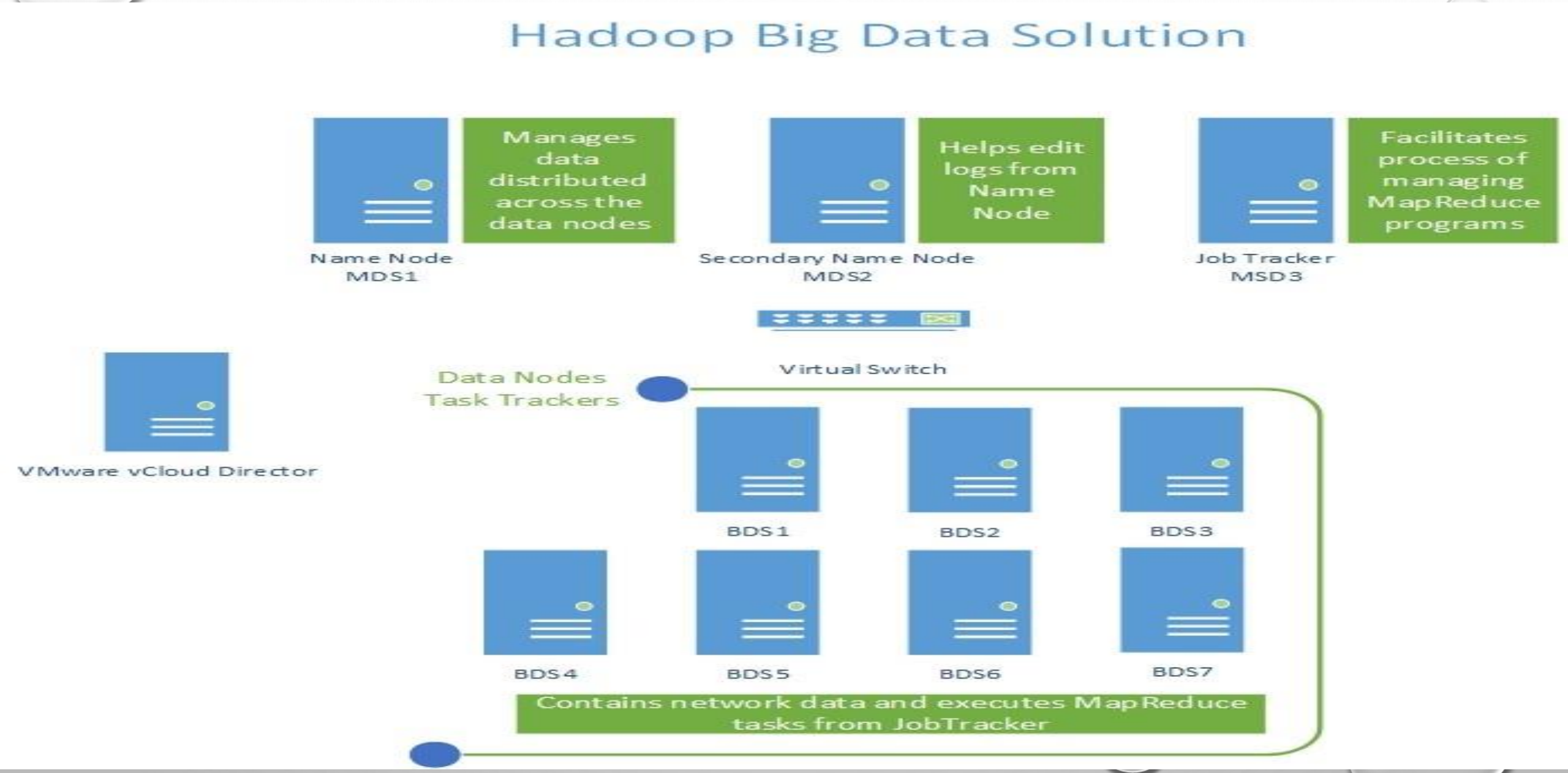


IMPLEMENTATION

IMPLEMENTATION PHASES

- PHASE 1: INDIVIDUAL SYSTEMS
  - Small cells access network simulator
  - EPC network emulator
  - Big data servers
  - Data analytics system
- PHASE 2: COMPATIBILITY TESTING
  - The integration of all the systems in the testbed will be performed through the core switch. Part of network traffic will be routed through the internet to test the cloud-based networking performance for the traffic generated by the simulated access network and the EPC
- PHASE 3: FULL IMPLEMENTATION
  - All the systems will go into production to start generating the daily traffic and perform the needed analysis for the research purposes. A lab manual will be developed to document the network scenarios and create the different lab procedures

BIG DATA SYSTEM

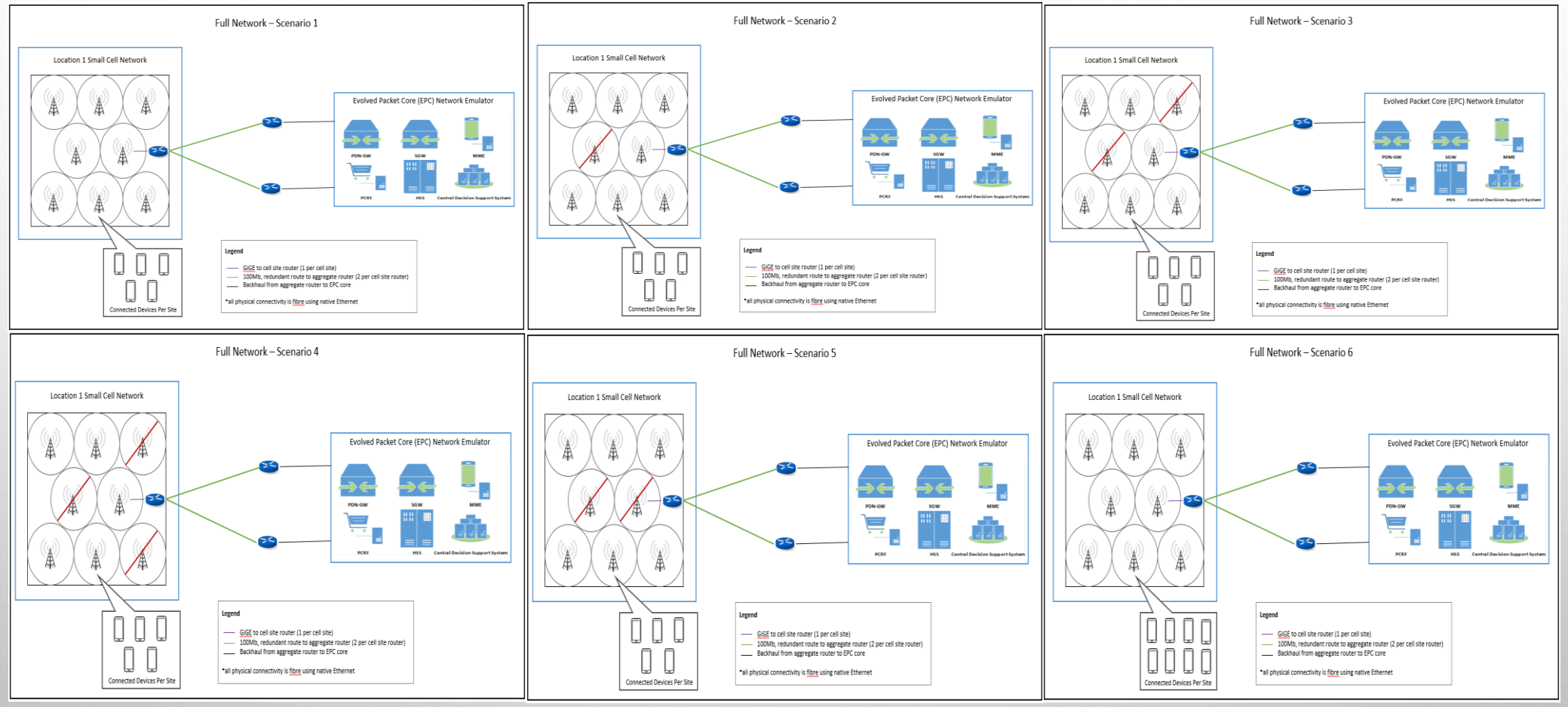


MODEL COMPONENTS

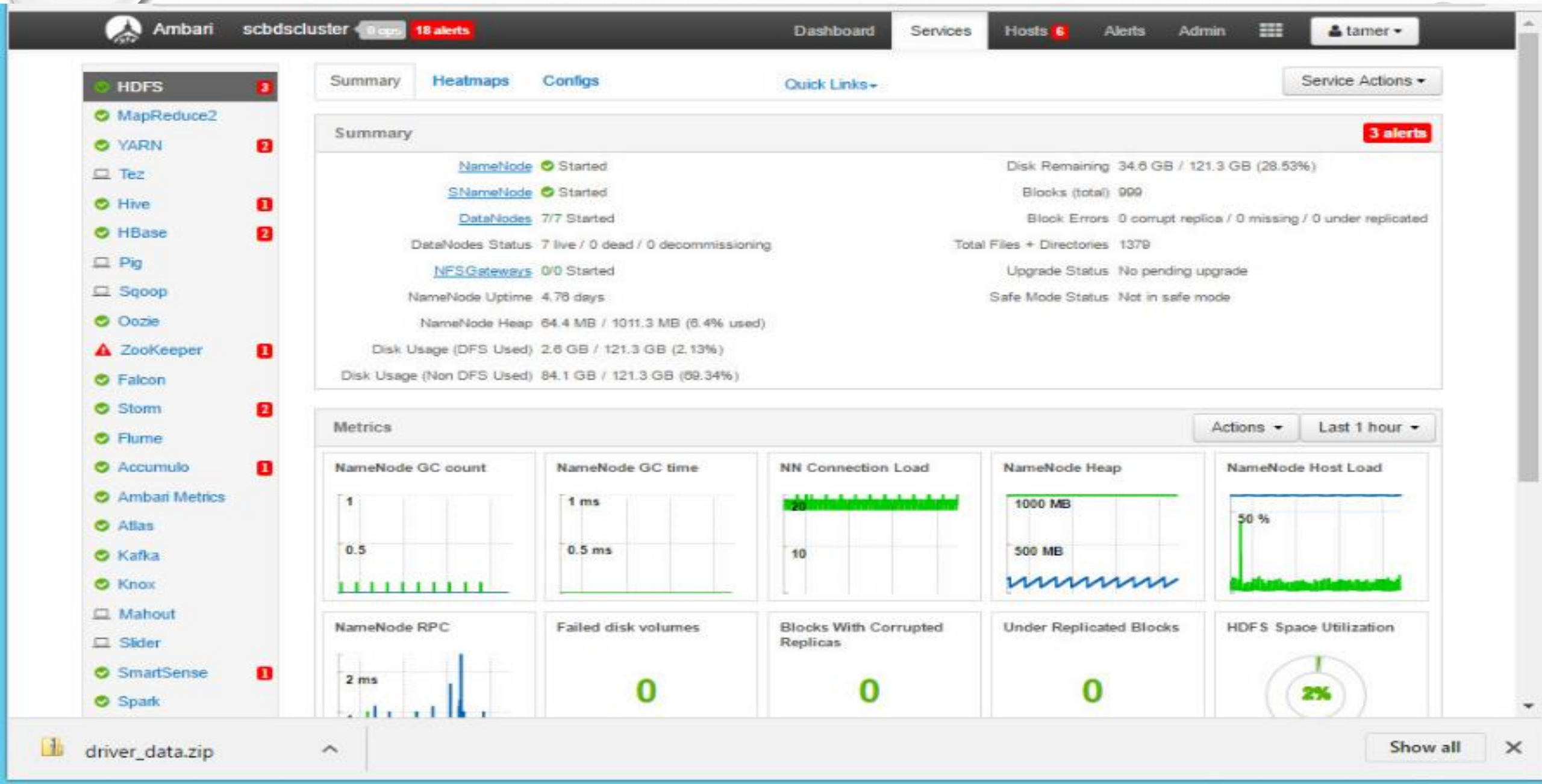
- VIRTUALIZATION:
  - Application server for cloud services: VSphere is used as the hypervisor and the model application server.
  - Virtual machines (nodes, small cells)
- VIRTUAL NETWORKING:
  - Virtual switch's (simulate LAN connectivity)
  - Virtual routers (simulate WAN connectivity)
- BIG DATA:
  - Data nodes
  - Name nodes
  - Job trackers
- APPLICATIONS
  - Hive used to store and analyze the data.
  - Pig is a scripting language used for creating Map-Reduce programs

OPERATIONS

WIRELESS NETWORK SCENARIOS

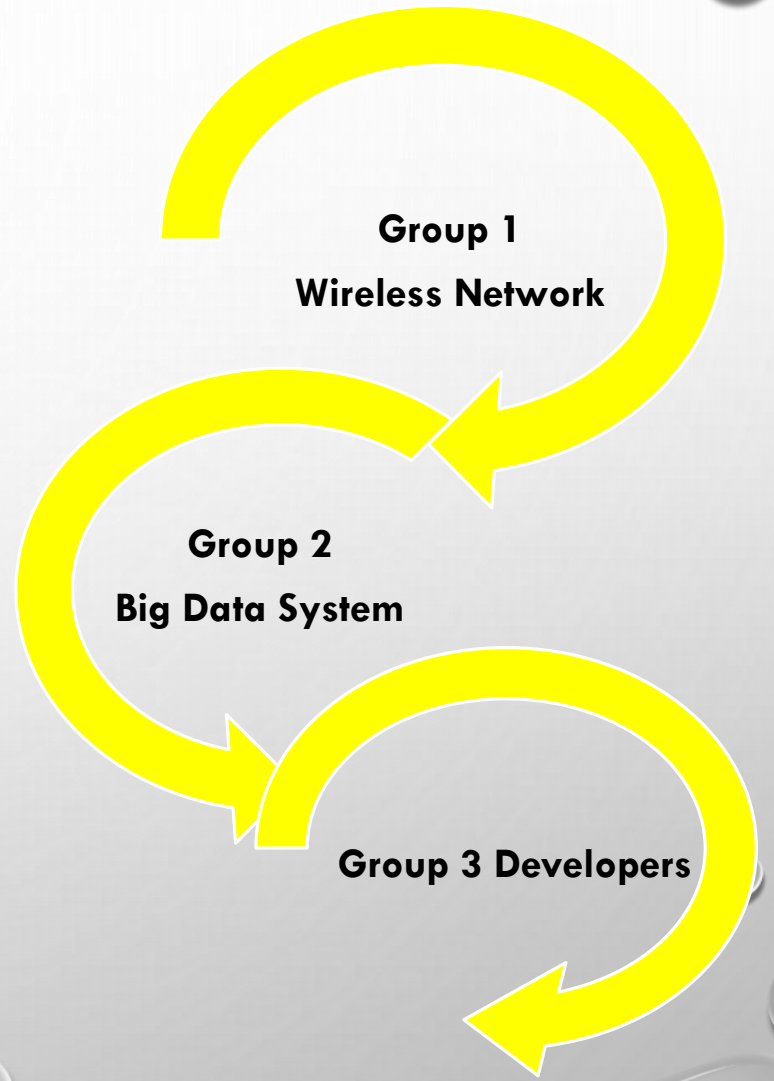


BIG DATA SYSTEM



TEAM

- Group 1 responsible for:
- Implementing the access network using the simulator.
- Sampling the network to generate the traffic needed for simulating a real operating access network.
- Group 2 are responsible for:
- Implementing, maintaining, and administering the Hadoop system HDFS.
- Maintain the virtual platform and ensure the reliability of the system.
- Group 3 responsible for:
  - Importing the datasets created by group two into the analytic tools.
  - Conducting the analysis using different algorithms and optimization techniques.
  - Creating the scripts needed for collecting the empirical data from the access network and exporting it to the BDS
  - Importing the optimized network parameters recommended by the SON module to the access network to enhance the network performance.

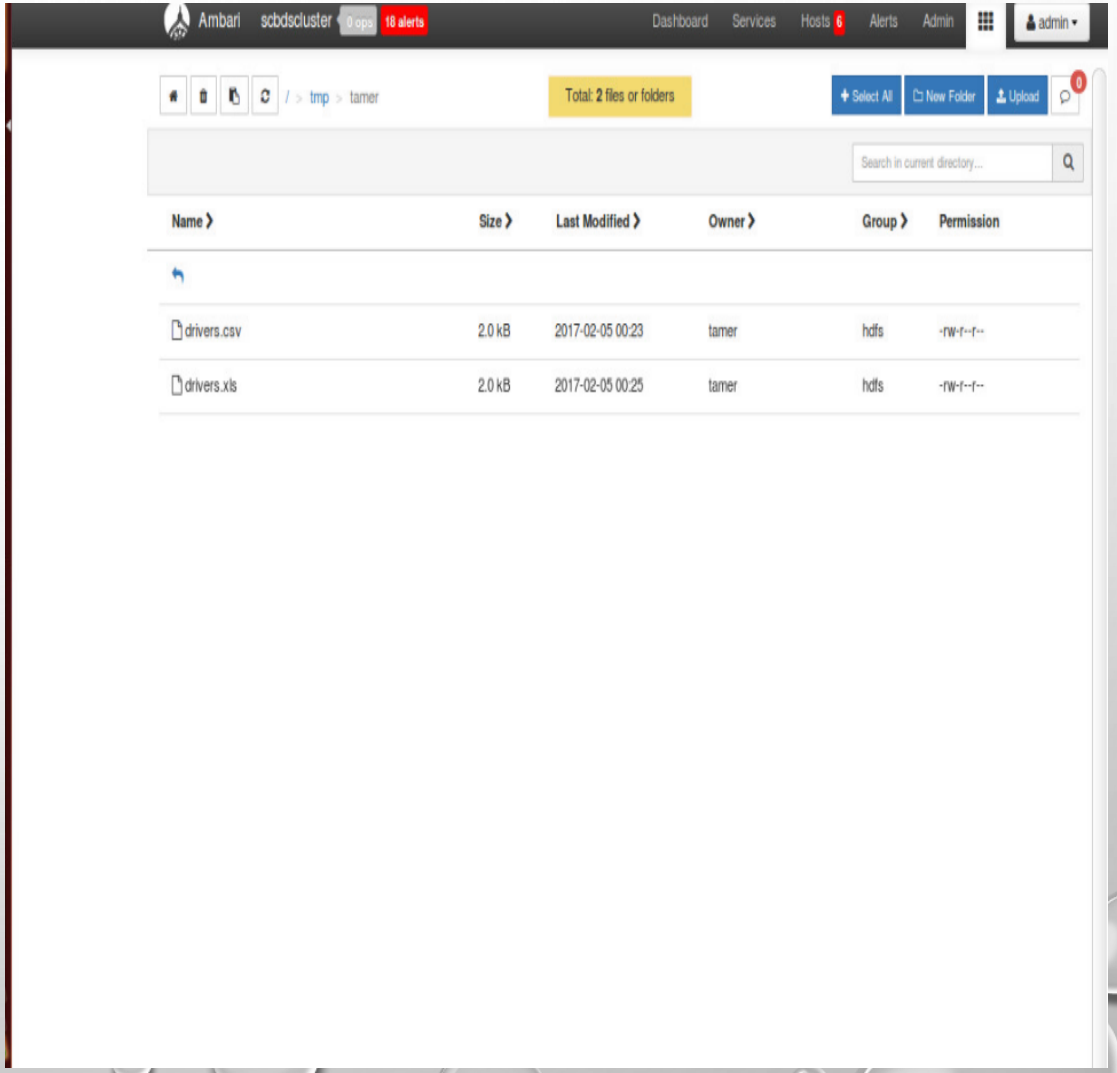


DATA COLLECTION

DATA EVERYWHERE

- Data collected from the networks have different characteristics that incur different challenges:
  - Unstructured Vs Structured data
  - Distributed Vs Centralized system
  - Local vs Global accessibility
  - Cost Effectiveness (servers Vs nodes)

Big Data Systems with cloud accessibility are considered are great repository solution that tackle these challenges

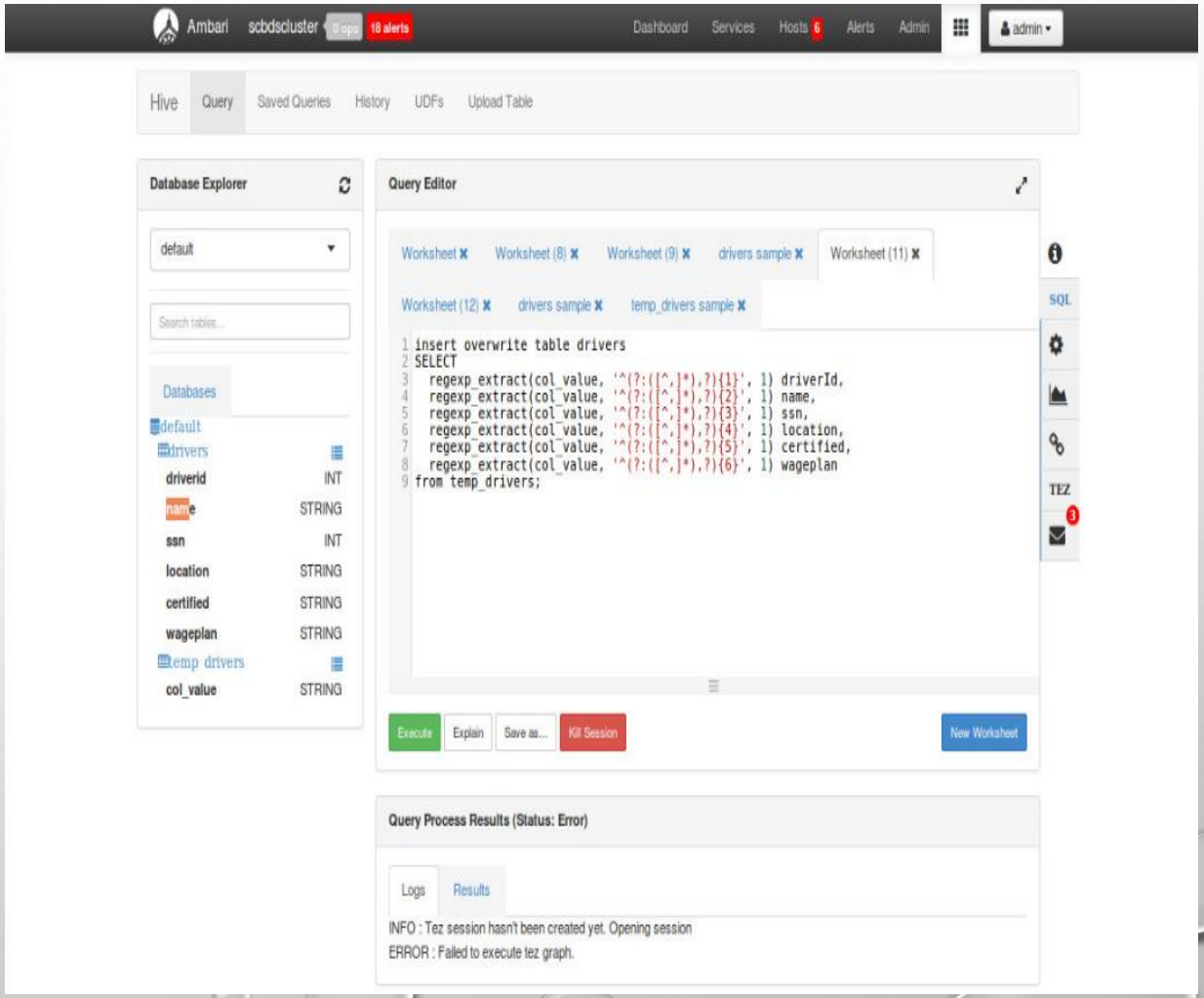


DATA ORGANIZATION

CENTRAL & DISTRIBUTED DECISION SUPPORT SYSTEMS

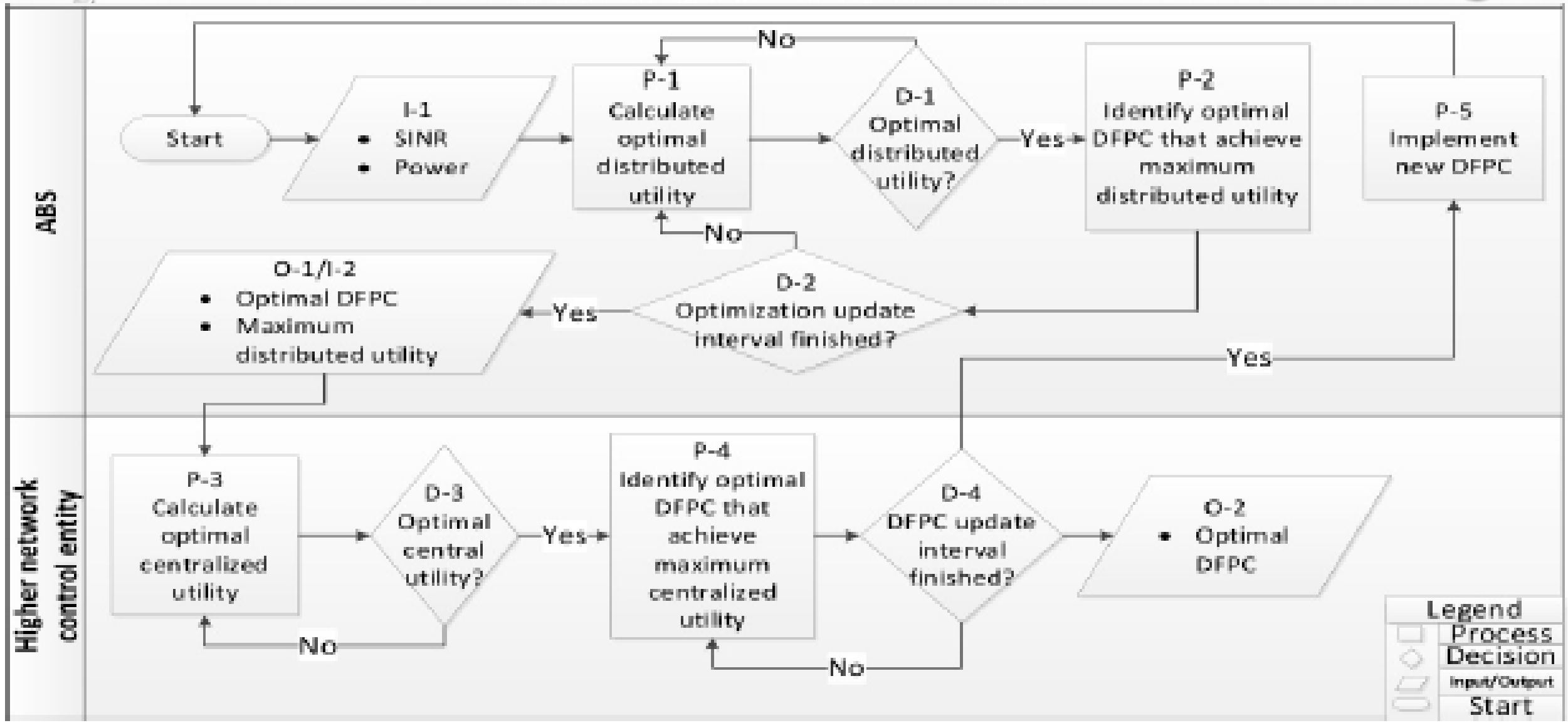
- BIG DATA BURDENS:
  - Data collected in distributed nodes need to be organized and integrated for decision making.
  - Small data sets from big data
  - Traffic considerations

HADOOP DISTRIBUTIONS PROVIDE MULTIPLE SERVICES TO AID IN ORGANIZING THE DATA AND REPRESENT IT IN A STRUCTURED FORM SUITABLE FOR ANALYSIS



DATA ANALYSIS

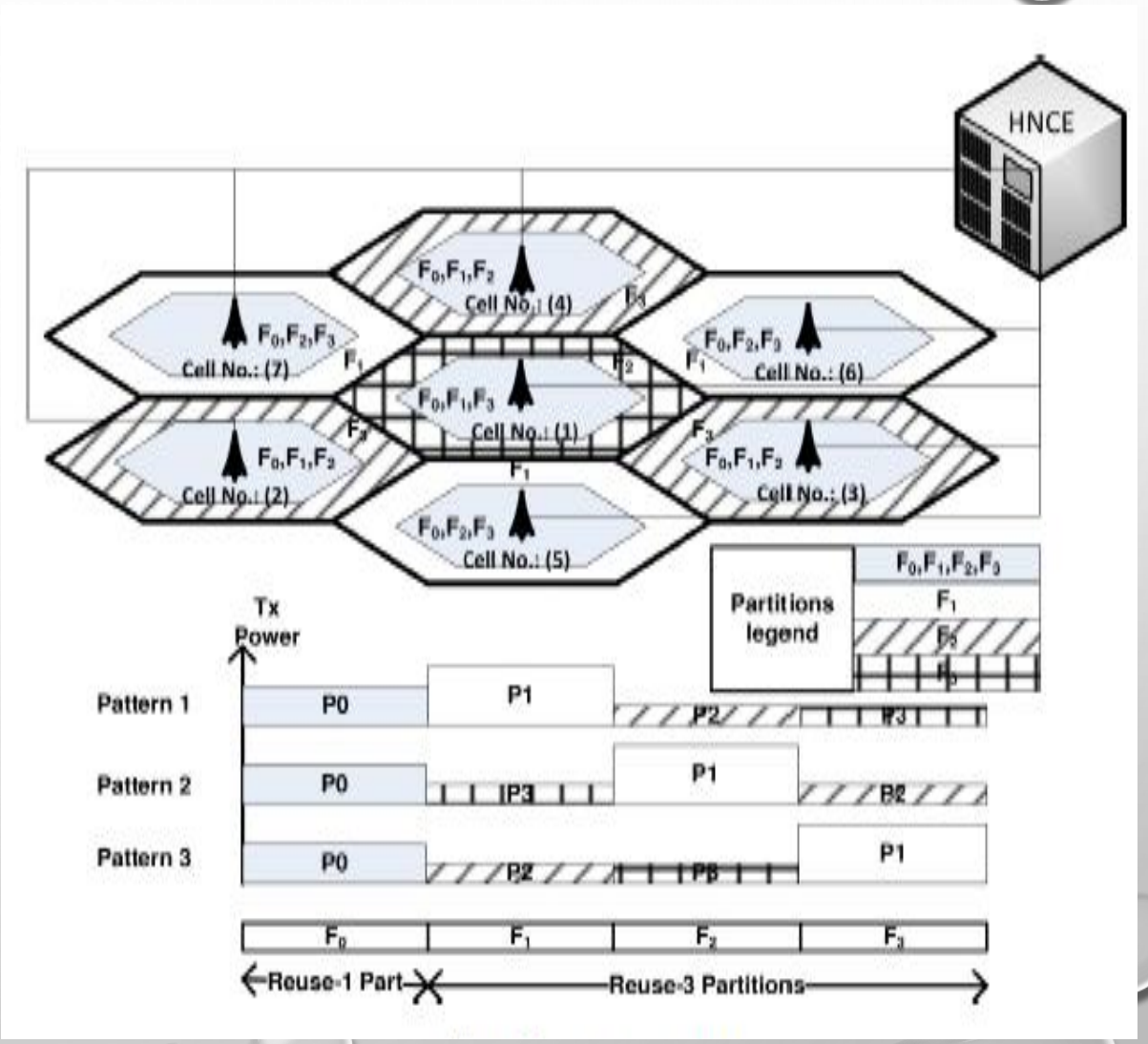
RADIO RESOURCE MANAGEMENT MODEL FLOW DIAGRAM



BROAD-BAND WIRELESS NETWORKS

FREQUENCY PARTITIONING

- Adaptive fractional frequency reuse is utilized in the system to partition the frequency in each cell.
- The number of physical resources and power levels in each partition are determined according to selected reuse factor in the network
- The figure shows a reuse 3 example used in the system



PARAMETERS OPTIMIZATION

- The distributed phase joint optimization problem is used to calculate the optimal network utilization $U_{B;MAX}$

- Each SC reports maximum distributed utilities & their corresponding frequency partitions to big data system.

$$
\begin{aligned}
& max && \sum_{i=1}^{M}\sum_{j=1}^{S}\sum_{k=1}^{N} U_{ijk}X_{ijk} \\
s.t \quad & && \sum_{i=1}^{M}\sum_{j=1}^{S}\sum_{k=1}^{N} P_{ijk}x_{ijk} \leq P_{max} \\
& && P_{ijk}x_{ijk} \leq P_{i,max} \\
& && \sum_{k=1}^{N} r_{ijk} \geq r_{ijk,min} \\
& && \forall (j = 1, 2, ..., S) \\
& && \sum_{i=1}^{M}\sum_{j=1}^{S}\sum_{k=1}^{N} x_{ijk} = 1 \\
& && x_{ijk} = \{0, 1\}
\end{aligned}
$$

FINALLY A VALUE CAN BE ACHIEVED

Q&A