Big Data Processing System Architecture

Problem Statement :

Design a Big Data System Architecture for the given use case.

Use Case :

Digital marketing is the component of marketing that uses the Internet and online based digital technologies such as desktop computers/web browsers, mobile phones and other digital media and platforms to promote products and services. An Advertisers wants to build its own campaigns analytics platform for its campaigns. In order to fetch and explore campaign's insights to tune the campaign further. User activities are recorded continuously and being send/dumped into advertisers log server in real-time. Data dumped from campaigns is in form of text file of **JSON rows**. On an average each campaign produces 1 GB of data per hour. Which can reach up to 32 GB during peak seasons. At a particular time multiple campaigns runs (between 0-50). Clearly, data generated from campaigns are large enough and suitable to be processed using Big Data technologies. You have to design Big Data log processing system for the Advertisers to help them generate the insights of the individual campaigns. Along with the specific analytics for example location or device based analytics, advertisers need two more kind of information first, total number of users or hits on the advertisement and second, total number of unique users or unique visitors.

Architecture Components Considerations :

Your Architecture design should be able depict the following.

- Overall architecture type. For example, Cloud based or inhouse system, Realtime or Batch System ?
- Data Storage system such as, homogeneous or heterogeneous, single or hybrid, NoSQL or SQL.
- Storage structure such as, folder structure or primary key etc.
- The data processing system and reasons for choosing the system.
- Data flow i.e. how the data is transferred from one module to another.
- Data format if any and the reason for using that format.
- Data Security features if any.

Submission Deliverables :

A **presentation deck (.ppt or .pdf)** to show and provide details about architecture design & its components/modules and their selection criteria and advantages of the proposed architecture.