BIG DATA ON CLOUD
SOCIAL BEHAVIOUR OF
PEOPLE ON SOCIAL MEDIA
SITES

GAUTAM KUMAR
PRITIKA SACHDEVA
IRSHAD QURESHI

GUIDE NAME: MS. NAVNEET POPLI
MANAGEMENT EDUCATION AND
RESEARCH INSTITUTE, NEW DELHI

BACKGROUND

• CLOUD COMPUTING
• CHALLENGES - DATA
  FRAGMENTATION
• SOCIAL MEDIA
• SOLUTION - HADOOP
  SYSTEM

PROBLEM STATEMENT

• To determine the behaviour of people
  that helps retail industry to predict
  what kind of product can be sold to
  people.
• To manipulate the data generating on
  social media site at 3V.
• Segregations of business value data
  from big data.

SOLUTION STRATEGY
**USING HADOOP SYSTEM**
- HDFS (Hadoop Data Files System)
- Apache Flume – Data Ingestion Utility
- API streaming (Twitter)
- Endpoint
- Apache Oozie - Loads the data into Hive
- Hive

**APPLICATION ARCHITECTURE**

**Development Environment**
- **Tools used for development**
  - IDE: Net beans IDE 7.3.1
  - Webserver: Glass Fish server 4.0
- **Development Language**
  - J2EE
- **Development OS**
  - Windows 8
Future scope

- LINKING CLOUD INFORMATION TO SMART DUST - “a tiny micro electromechanical sensor which is a bundle of source of energy.”