Experiment 1

Aim:

Install, configure and run Hadoop and HDFS.

Theory:

Hadoop is a framework for processing big data across a cluster of machines.

HDFS is its storage system.

Installing and Configuring Hadoop entails preparing a cluster with various roles to ensure data distribution and redundancy.

Running Hadoop and HDFS involves executing data replication and processing using map reduce and efficiently managing big data.

Flowchart:

```
  ┌─────────┐              ┌──────────────────┐
  │  Start  │─────────────▶│   Install VMWare │
  └─────────┘              └──────────────────┘
                                    │
                                    ▼
                           ┌──────────────────┐
                           │  Install Ubuntu  │
                           └──────────────────┘
                                    │
                                    ▼
                           ┌──────────────────┐
                           │  Install Hadoop  │
                           └──────────────────┘
                                    │
                        Download Hadoop distribution
                            Configure Hadoop
                                    │
                                    ▼
  ┌─────────┐              ┌──────────────────────┐
  │   End   │◀─────────────│ Start Hadoop Services│
  └─────────┘              └──────────────────────┘
```