Recap:

OLTP --->ETL -DW --OLAP

RDB ⇒ TB

PB /ZB/YB →

Bigdata Intro:

1 kg rice /paddy ----> cook --->200gb /400tb →RDB/ Dw

1 tonne rice → chef → PB →

Bigdata ===> 3 v

**Volume**

**variety**  → Any kind /type

Human --Structured / semi-structured (json/XML)/ unstructured data

Machine created data => sensor/GPS / logs

**velocity**

CPU GPU TPU

Volume ---> Any size

Multinode cluster /data center :

Scalability :

Horizontal → adding new nodes/machines to the cluster at anytime without disturbing existing nodes

Vertical -> scaling in a single node

Client --server ⇒

Distributed storage 👍

Post office →

C1 (stamps/speed post) c2 (speed post/stamp/banking c3 speed post/stamp/banking

M1 m5 m100

Sales.txt sales.txt sales.txt

Replication

Fault tolerance

Data availability

Single point of failure (SPoF)

========================================

Vmware ⇒ edit virtual machine

Memory => increase 4gb to 8gb

Processor => 1 to 4

Virtual machine password : hadoop

day 2

cd

cd /

cd ..

cd ../..

cd -

ls

mkdir dta

mkdir -p 2024/feb/20

cd

1745 pwd

1746 mkdir dta

1747 mkdir /2024/feb/20

1748 mkdir

1749 mkdir --help

1750 mkdir 2024/feb/20

1751 mkdir -p 2024/feb/20

1752 cd 2024/feb/20

1753 cd ..

1754 cd ../..

1755 cd -

1756 cd 20

1757 cat >sample.txt

1758 cat sample.txt

1759 cat >>sample.txt

1760 cat sample.txt

1761 ls

1762 ls -lrt

1763 ls --help

1764 cd

1765 ls -lrtt

1766 ls -lrt

1767 cd 2024

1768 ls -lrt

1769 cd

1770 cd dta

1771 ls -l

1772 cd

1773 cd datasets

1774 ls -l

1775 cd

1776 ls

1777 cd 2024/feb/20/

1778 cp sample.txt demo.txt

1779 mv demo.txt 2024/feb

1780 mv demo.txt /home/hduser/2024/feb

1781 ls

1782 cp sample.txt demo.txt

1783 ls

1784 rm sample.txt

1785 cd

1786 rm 2024

1787 rmdir 2024

1788 rm -rf 2024

1789 history

Step1: Download Vmware workstation player using link

below:

https://customerconnect.vmware.com/en/downloads/detail

s?downloadGroup=WKST-PLAYER-

1700&amp;productId=1377&amp;rPId=97014

Step2: Download Jigsaw vm using link below:

https://drive.google.com/file/d/1wCa43KSpPlZSi22KVS3jcnTy

3iLEx\_6C/view

https://unextlearning-my.sharepoint.com/:u:/g/personal/krishna\_shetty\_u-next\_com/EafqCbLFx29Et8MPqKqNdlcBb6Q43MFslSkwDsMGbRLgZg?e=X82vhr