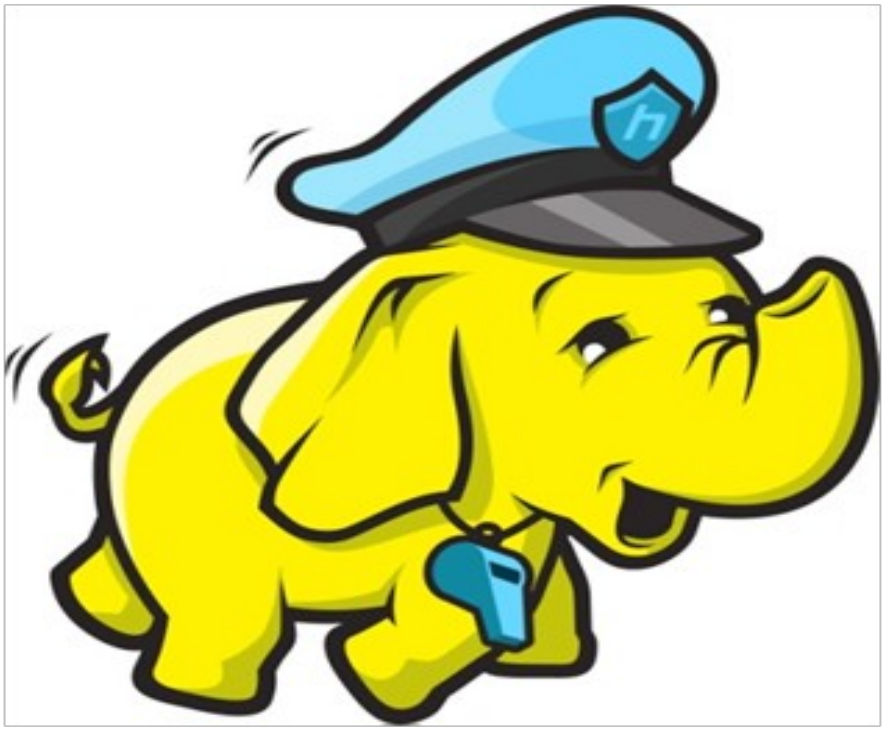
Recover from Namenode failure



Agenda

What is Namenode
Responsibility of Namenode
Single point of failure
Causes of Namenode failures.
Namenode recovery
Role of Secondary Namenode
FsImage & Edits files.
Checkpoints in Hadoop
Creating checkpoints
Recovery with the help of checkpoint

What is Namenode

Namenode is a process which runs on master machine of Hadoop cluster.

We need to contact Namenode for any read/write operation in HDFS.

Namenode keeps metadata of the data which is stored in HDFS.

Namenode cordinates with datanodes to read/write data in HDFS.

Responsibility of a Namenode

Namenode keeps a block map of all the files in HDFS.

It contacts each datanode & ask for block report.

It creates bigger block report from all datanode block reports.

It keeps list of live nodes & dead nodes.

It balances the storage of Hadoop cluster.

Single point of failure

Namenode is single point of failure in Hadoop cluster.

Hadoop cluster is not accessible if Namenode is down.

We can't do any read/write operation, even datanodes have all the data.

Hot backup is not yet supported in Hadoop.

Causes of Namenode failure

Master machine can stop working due to hardware problem.

Namenode metadata can get corrupt.

Without metadata, Namenode is not capable of finding the data in HDFS.

We can't contact Datanode directly for data.

Checkpoint can be used to recover metadata.

Namenode recovery

Namenode must be recovered in order to access HDFS.

Hadoop cluster will remain offline, untill we recover Namenode.

Secondary Namenode can help Namenode to recover.

We can only recover till the last checkpoint saved.

Stale data is far better than no data.

Role of Secondary Namenode

Secondary Namenode must be on separate machine in Hadoop production cluster.

Add the following entry in **hdfs-site.xml** to run Secondary Namnode on another machine.

```
<property>
    <name>dfs.secondary.http.address</name>
    <value>192.168.1.2:50090</value>
    <description>The secondary Namenode address and port </description>
</property>
```

Checkpoints, which are stored on Secondary Namenode, helps in Namenode recovery.

FsImage & Edits files

FSImage contains snapshot of HDFS metadata.

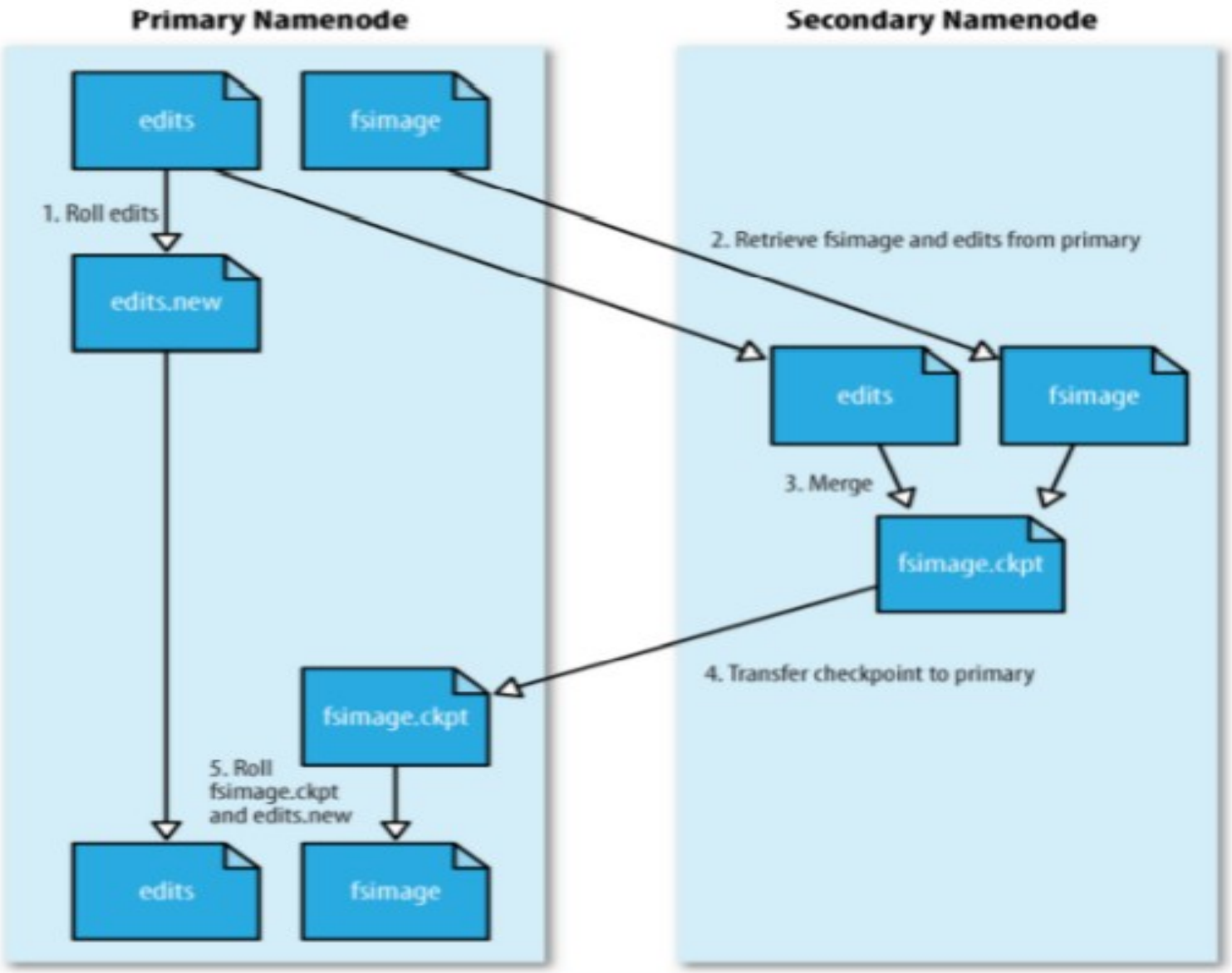
Namenode loads FSImage at it's startup.

After every read/write operation, FSImage is not updated.

Instead, all the changes are recorded in edits file.

Later, a new FSImage can be created by merging old FSImage & edits file.

FsImage & Edits files

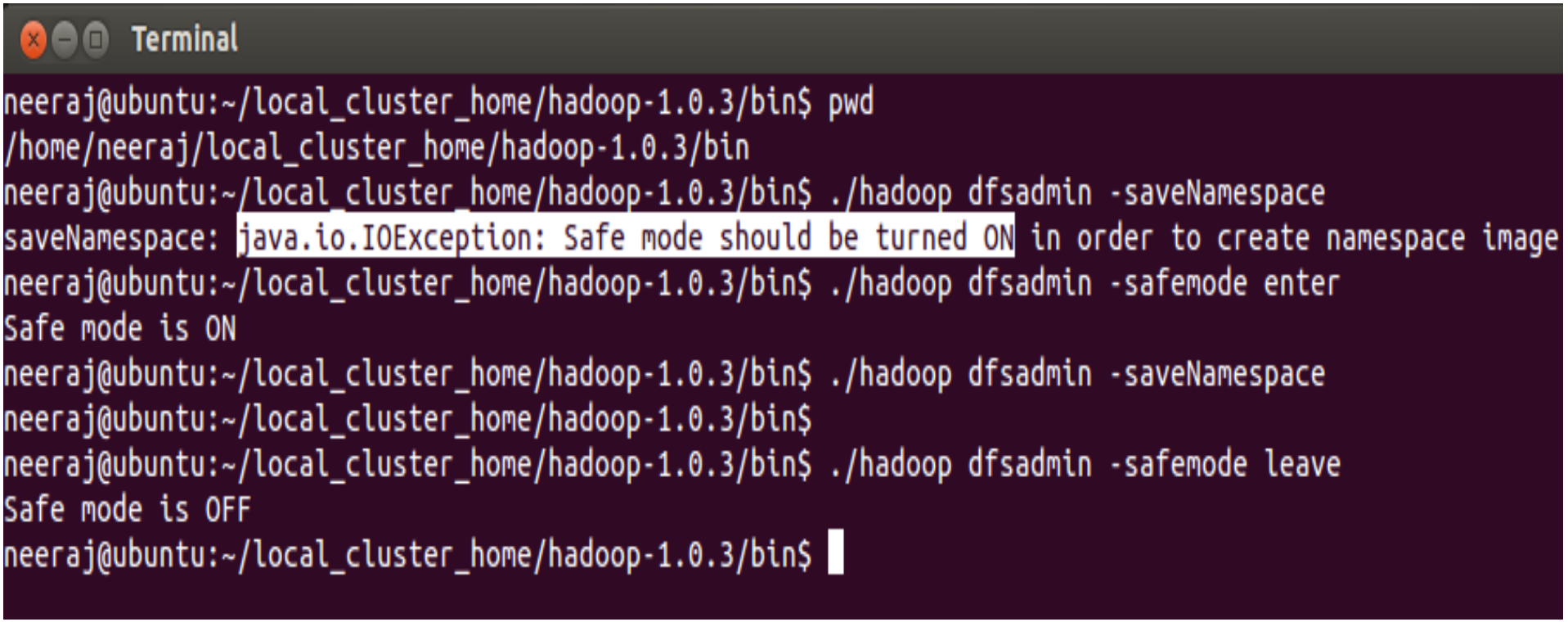


Creating checkpoints

Checkpoints are taken after every 1 hour (by default)

Checkpoint are useful for recovering from failure.

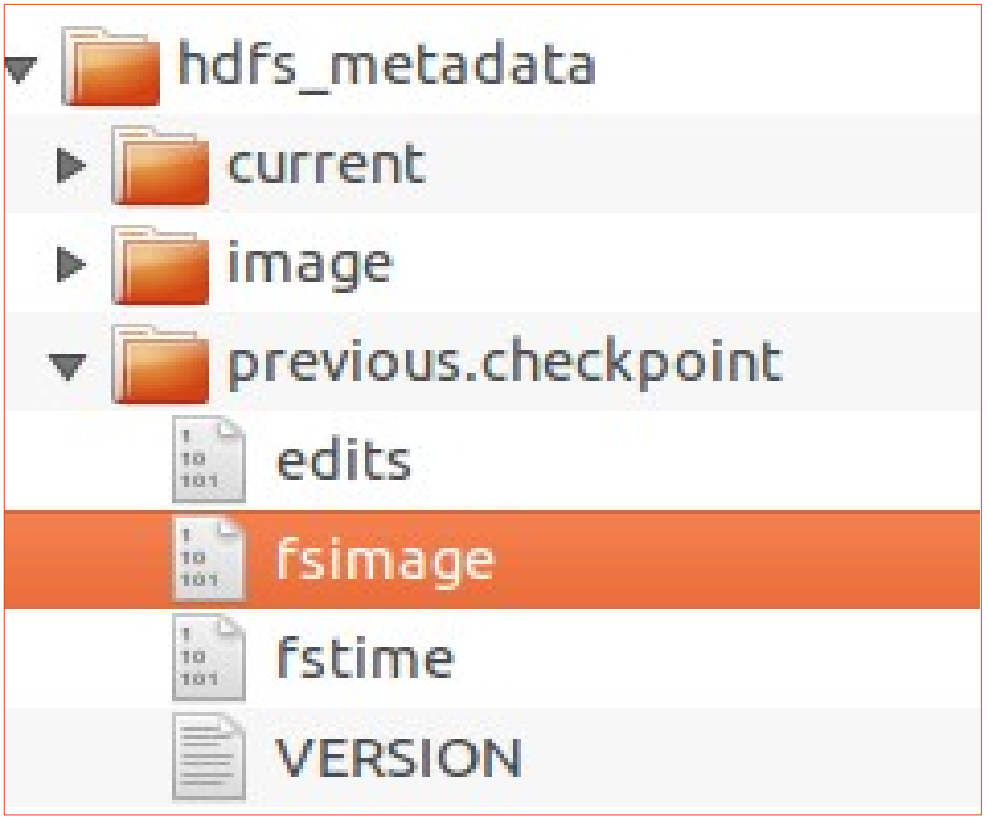
We can create checkpoints, by running the below command



Checkpoints in Hadoop

Checkpoint are stored in below shown directory

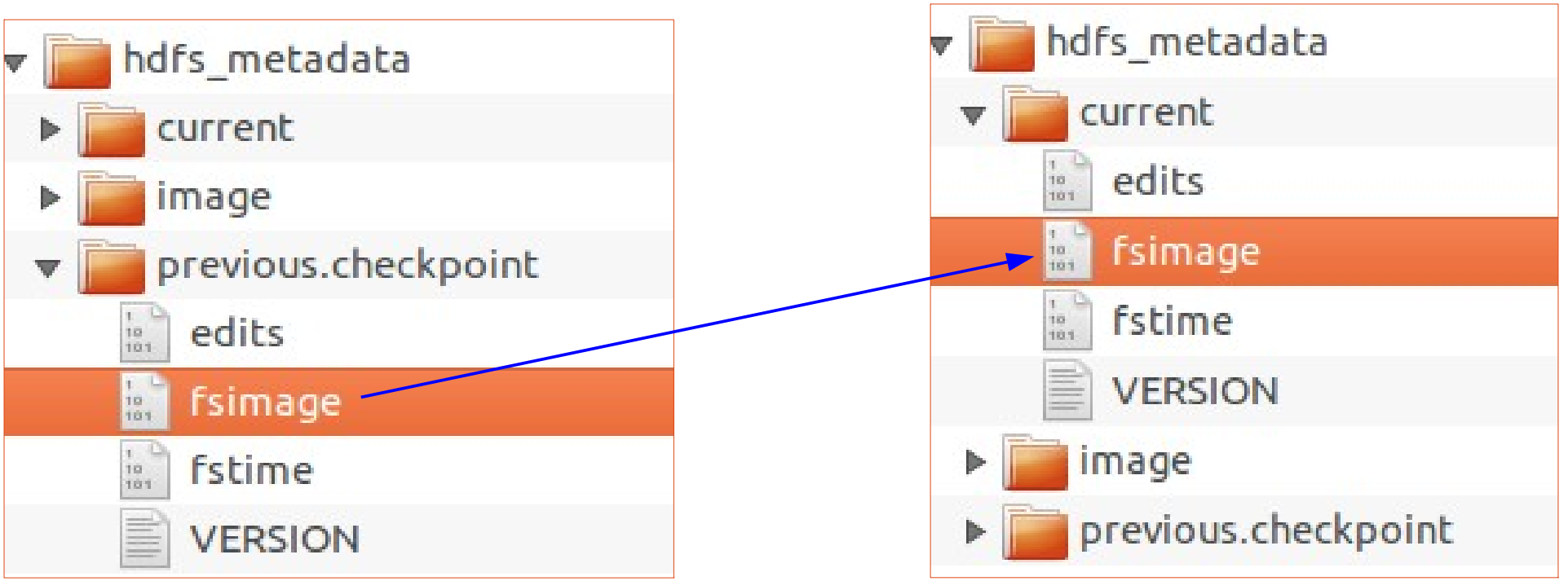
Fsimage is used to recover the namenode



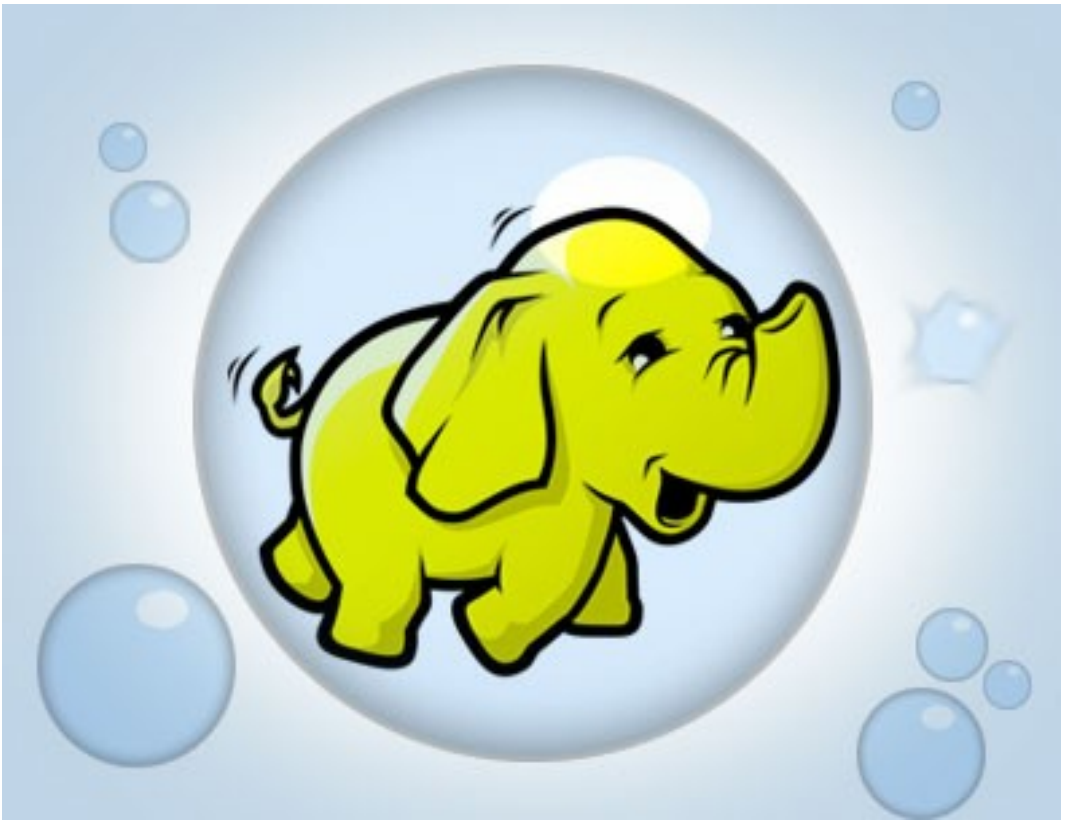
Recovery with the help of checkpoint

Follow below mentioned steps to recover namenode

1. Stop Hadoop by runnning ./stop-all.sh command
2. Copy **fsimage** file from checkpoint directory to current directory.
3. Start hadoop by runnning ./start-all.sh command



...Thanks...



For online Hadoop training, send mail to neeraj.ymca.2k6@gmail.com