**C Program that generates a input file(input.txt) with four random words**

**PROGRAM:**

#include<stdio.h>

#include<string.h>

#include<limits.h>

int main()

{

char \*inp[15];

int i;

FILE \*fd;

fd=fopen("input.txt","w+");

inp[0]="apple";

inp[1]="orange";

inp[2]="banana";

inp[3]="pineapple";

for(i=0;i<USHRT\_MAX\*300;i++)

{

fprintf(fd,"%s ",inp[i%4]);

if((i+1)%10==0)

fprintf(fd,"\n");

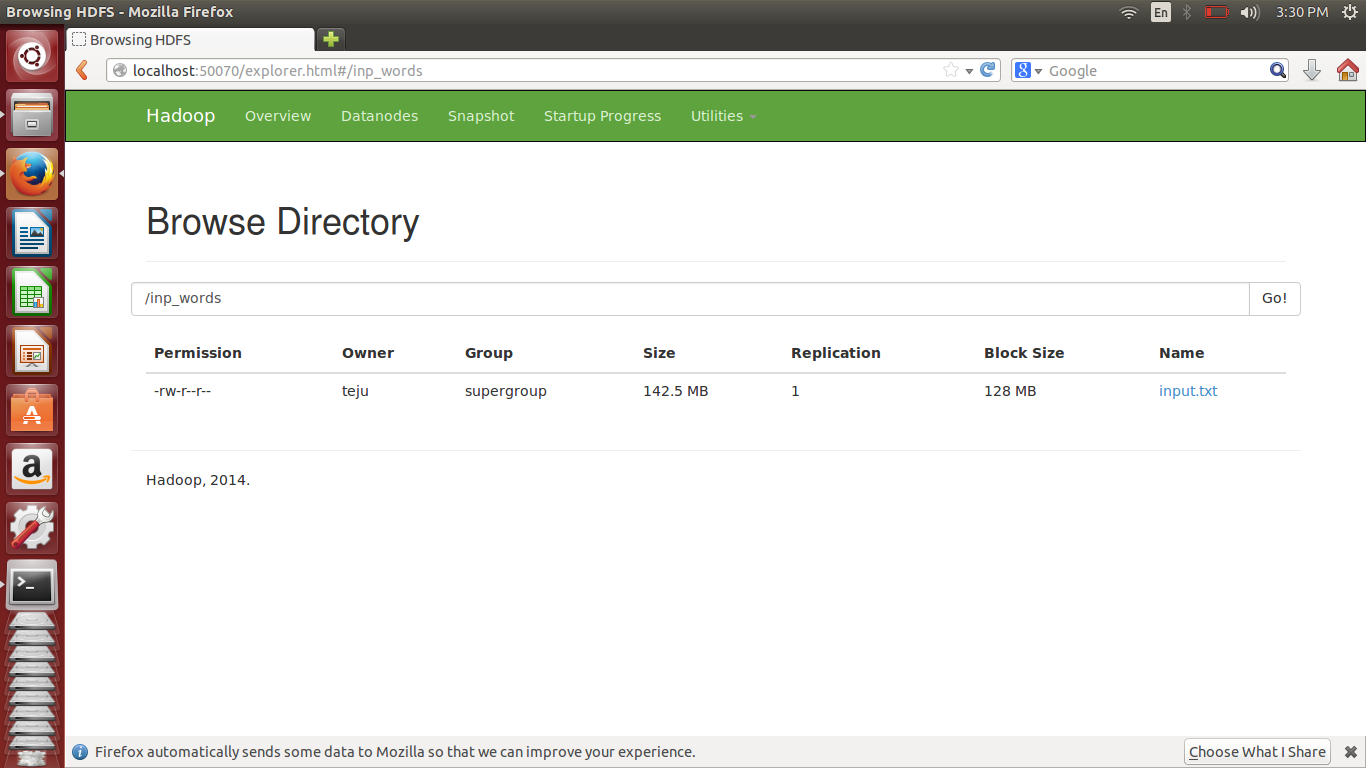
}

fclose(fd);

return 0;

}

This creates a input file of 143MB which results in **TWO** splits as block size in Hadoop is 128MB.That “input.txt” is uploaded to HDFS.



The “input.txt” file look like:

