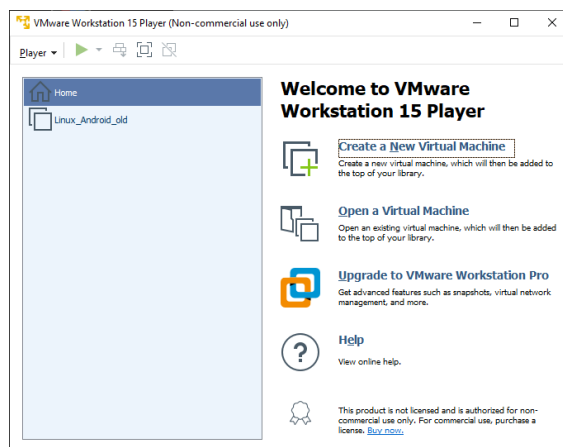


## התקנת לינוקס-אובונטו על מכונה וירטואלית

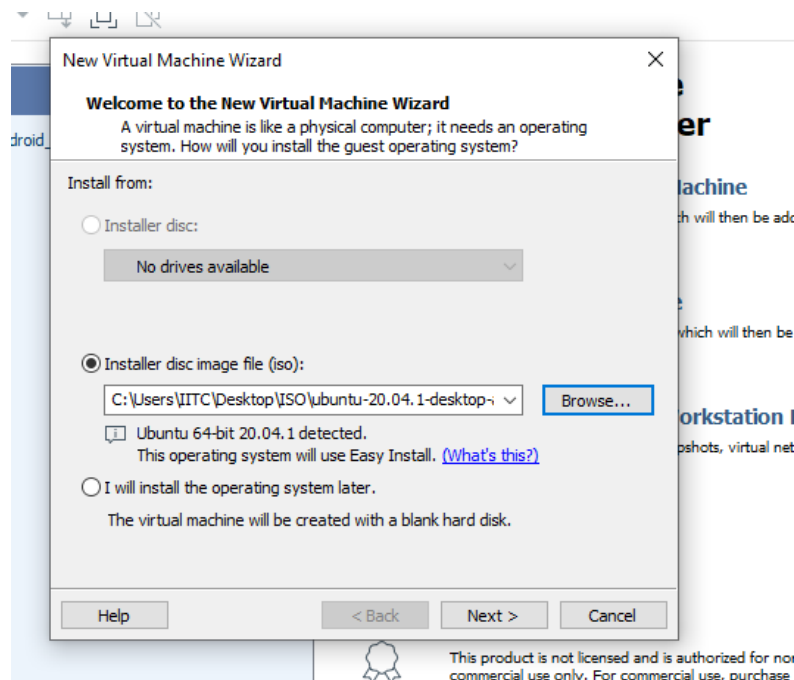
1. הורידו והתקינו VMware Workstation Player  
<https://www.vmware.com/il/products/workstation-player.html>
2. הורידו IMG של אובונטו  
<https://ubuntu.com/download/desktop>

התקנת מע ההפעלה במכונה וירטואלית:  
לפני תחילת עבודה יש לוודא שווירטואליזציה אכן מאפשרת בהגדרות החומרה שלכם שמות אופייניים להגדרה זו Virtualization/VT-x/AMD-V  
הגדרה זו נמצאת בביוס המחשב, ויש לחפש ברשת איך מגדירים זאת על החומרה האישית שלכם.

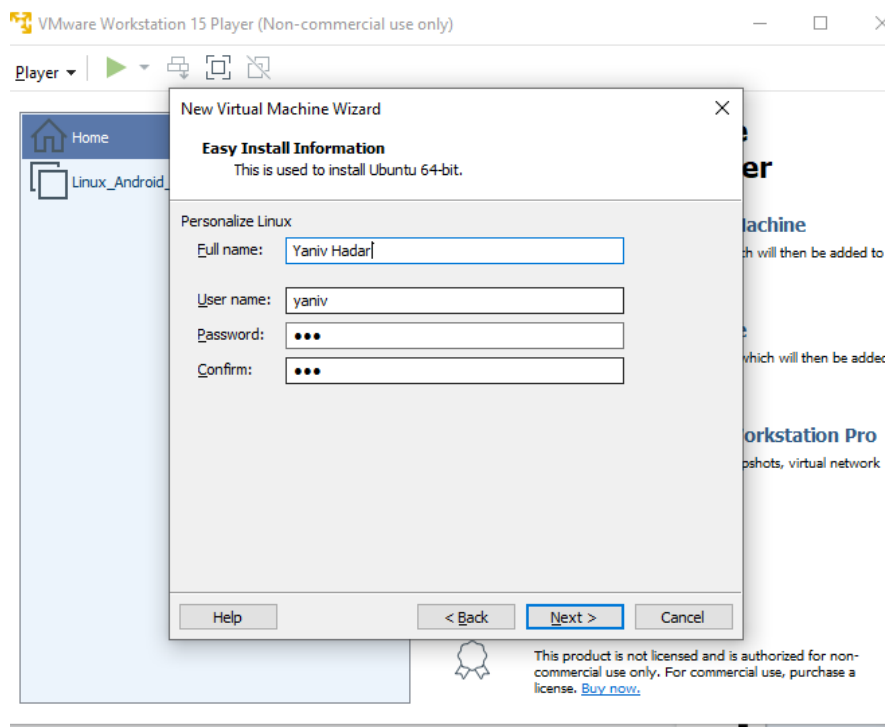
### 1. יצירת מכונה חדשה:



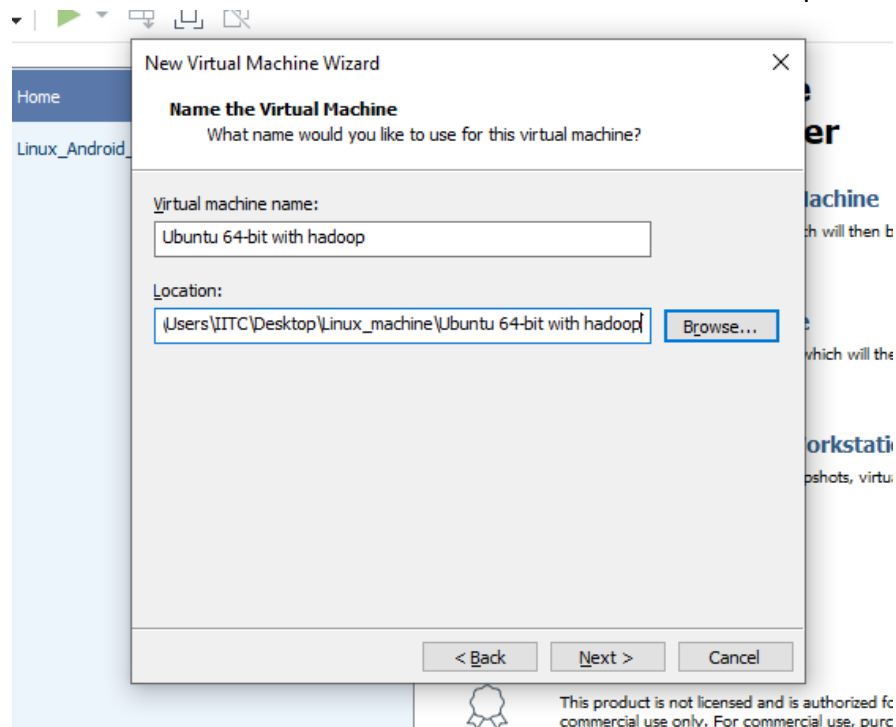
### 2. בחירת נתיב ה IMG



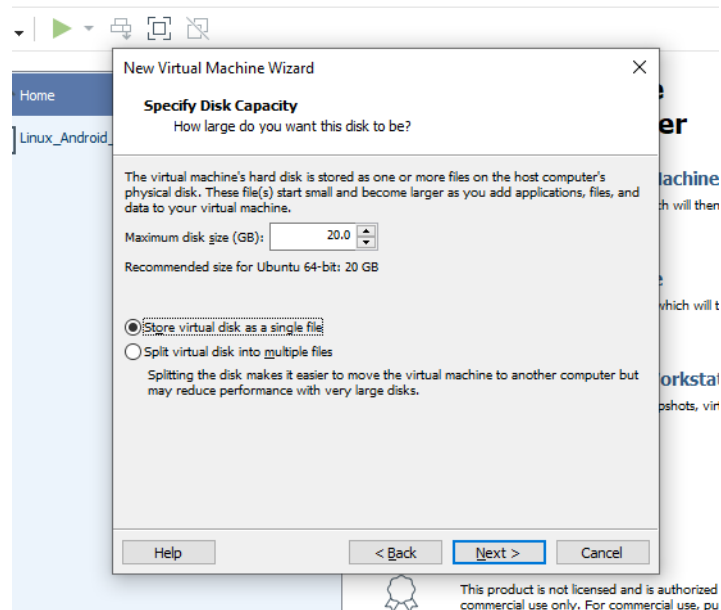
### 3. הגדרת שם מחשב ומשתמש ראשוני



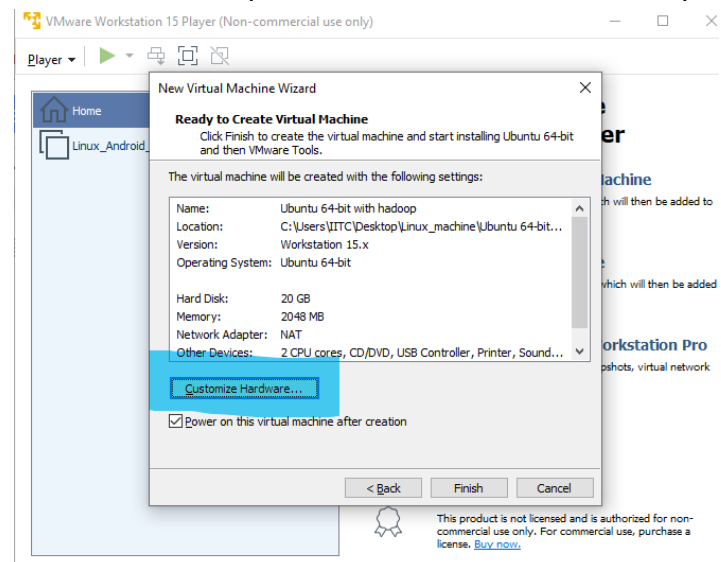
### 4. שם ומיקום המכונה :



## 5. הגדרת אחסון



## 6. מסך סיכום - בשלב זה נשנה את אופי התקשורת מ NAT ל BRIDGED



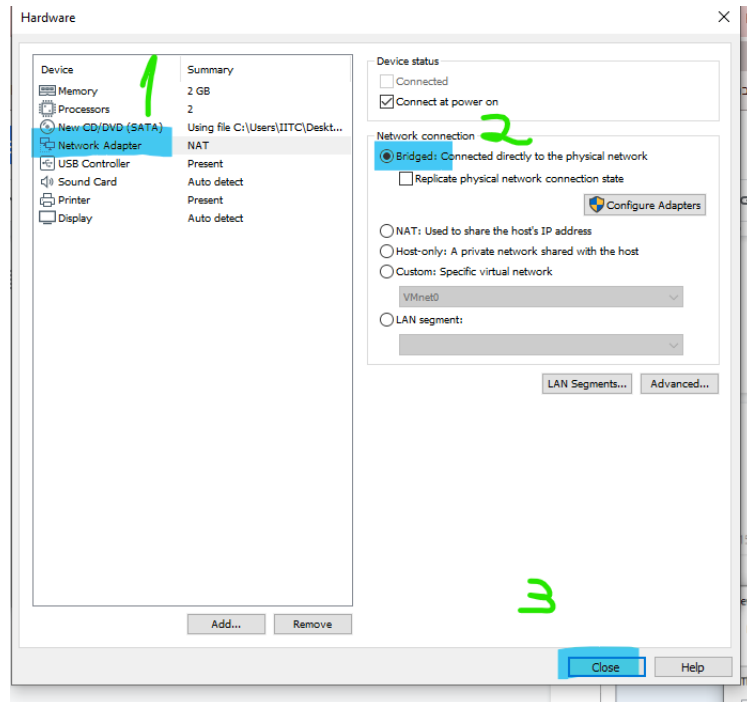
### The difference between NAT / Bridged / Host-Only networking

**Host-Only:** The VM will be assigned one IP, but it's only accessible by the box VM is running on. No other computers can access it.

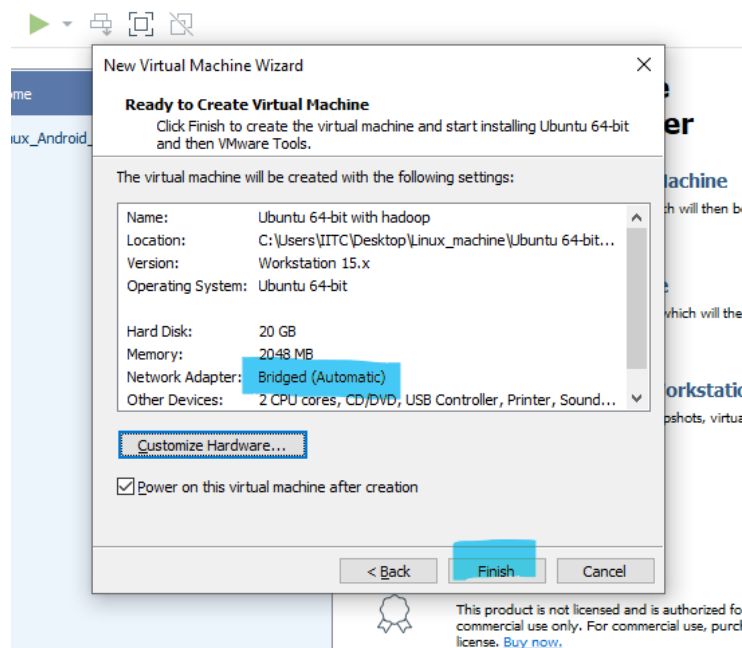
**NAT:** Just like your home network with a wireless router, the VM will be assigned in a separate subnet, like 192.168.6.1 is your host computer, and VM is 192.168.6.3, then your VM can access outside network like your host, but no outside access to your VM directly, it's protected.

**Bridged:** Your VM will be in the same network as your host, if your host IP is 172.16.120.45 then your VM will be like 172.16.120.50. It can be accessed by all computers in your host network.

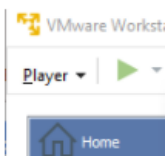
## 7. במסך הגדרות החומרה תחת הגדרות הרשת נבחר ב BRIDGE



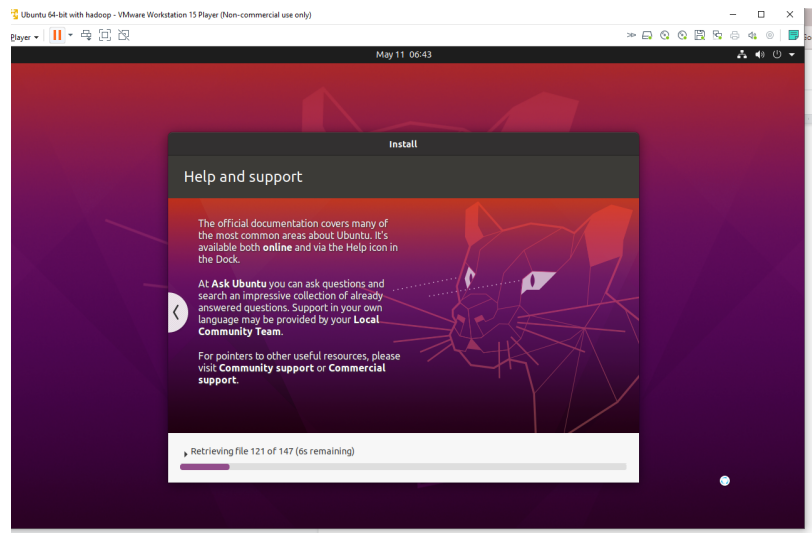
## 8. לאחר מכן נוודא שאכן התעדכן לBRIDGE ונאשר התקנה



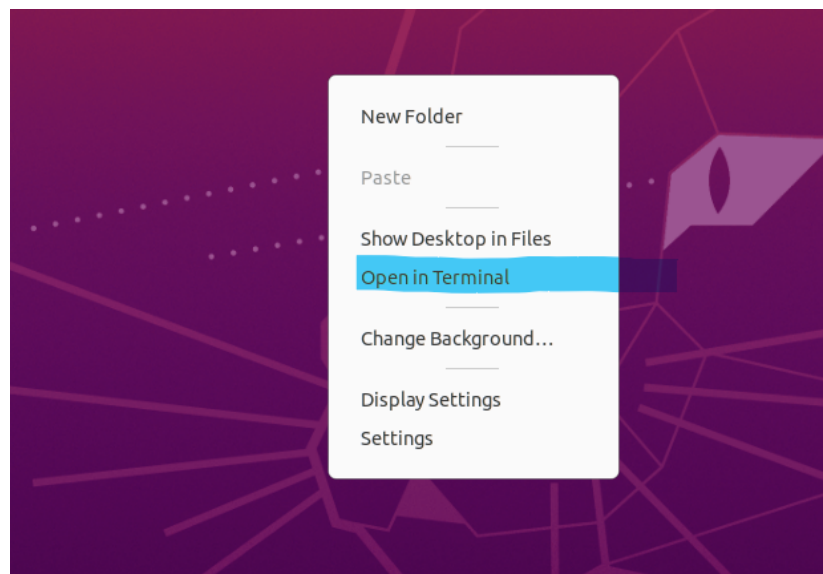
## 9. לאחר הפעלה של המכונה בפעם הראשונה (לחיצה על החץ הירוק)



תתחיל ההתקנה (ללא מעורבות שלכם), ואם נושא הרשת הוגדר כיאות אז גם יותקנו עדכונים באופן אוטומטי



10. פתיחת הטרימינל - ממשק פקודה - CLI כפתור ימני על הדסקטופ, בחירת אופציה בתפריט



11. רענון מאגר העדכונים:  
sudo apt update

## התקנת הדופ

### 1. התקנת ג'אווה

```
sudo apt install default-jdk
```

### 2. נודא שההתקנה עברה

```
java -version; javac -version
```

### 3. נחפש את נתיב ההתקנה בכדי להגדיר את משתנה הסביבה JAVA\_HOME ונשמור בצד:

a. ההתקנה יוצרת לינק סימבולי לקובץ ב bin (כמוסכמה בלינוקס)

b. בעזרת which נוכל לוודא זאת

```
which java
```

```
yaniv@ubuntu:~/Desktop$ which java
/usr/bin/java
```

### c. בעזרת readlink נוכל לראות היכן נמצאת התיקיה הפיזית של קובץ ההפעלה של

ג'אווה, וכך למצוא את נתיב ההתקנה

```
readlink -f /usr/bin/java
```

```
yaniv@ubuntu:~/Desktop$ readlink -f /usr/bin/java
/usr/lib/jvm/java-11-openjdk-amd64/bin/java
```

### d. קבצי הפעלה ימצאו בדכ ב bin - לכן הנתיב לכל ה JDK הוא :

**/usr/lib/jvm/java-11-openjdk-amd64/**

### e. או בפקודה אחת לקבלת הנתיב הרלוונטי

```
readlink -f /usr/bin/java | sed "s:bin/java::"
```

## 4. הורדת הדופ והעתקה למיקום מתאים

```
sudo wget
```

```
https://apache.mivzakim.net/hadoop/common/hadoop-3.2.2/hadoop-3.2.2.tar.gz
```

### 5. פתיחת הארכיון שהורדנו

```
tar -xzvf hadoop-3.2.2.tar.gz
```

### 6. העברת התיקיה שפתחנו למיקום נכון מבחינת מוסכמות מע' ההפעלה

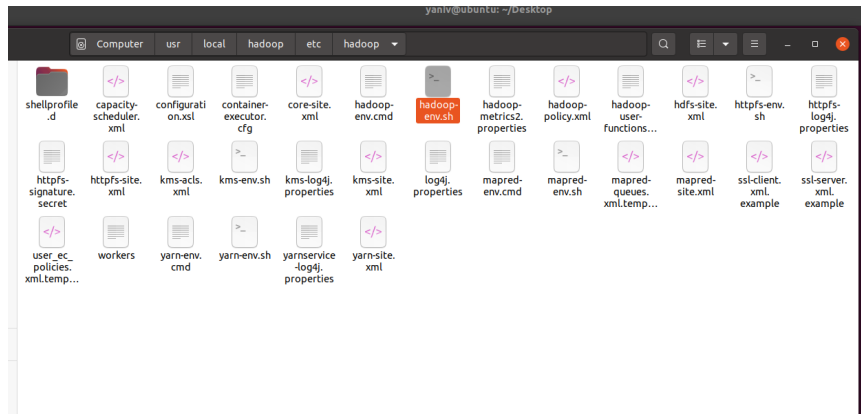
```
sudo mv hadoop-3.2.2 /usr/local/hadoop
```

### 7. הגדרת הדופ

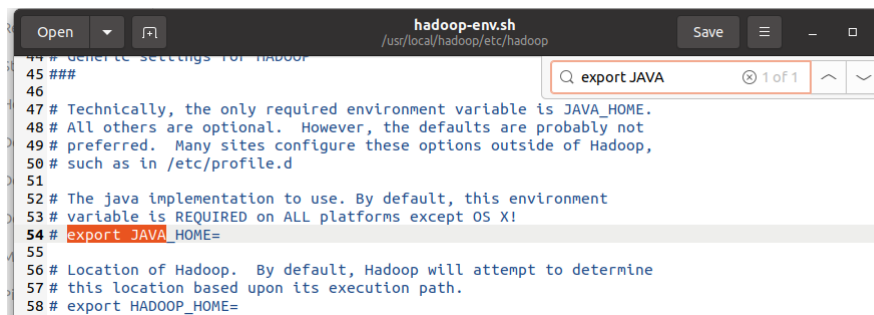
נערוך את הקובץ hadoop-env.sh ונוסיף לו את הנתיב של JAVA

```
sudo nano /usr/local/hadoop/etc/hadoop/hadoop-env.sh
```

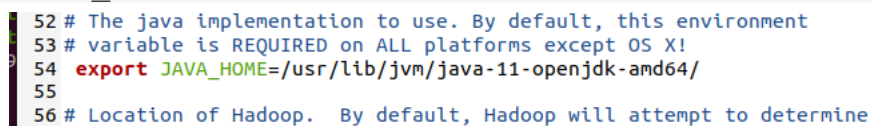
## או לנווט אל מיקום הקובץ בממשק הגרפי ולערוך עם gedit



### 8. יש לחפש בקובץ את ההערה export JAVA\_HOME

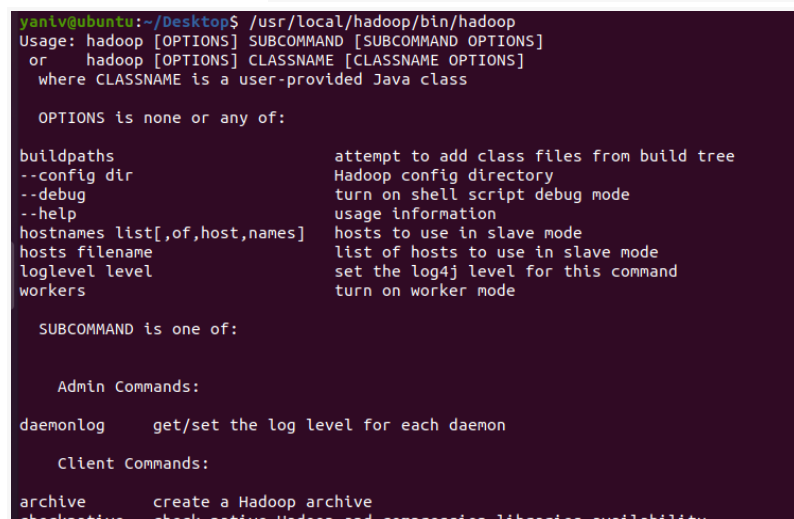


### 9. יש לבטל את ההערה (למחוק את ה # ) והוסיף את הנתבי של ה JDK שמצאנו מוקדם יותר export JAVA\_HOME=/usr/lib/jvm/java-11-openjdk-amd64/



### 10. הרצת הדופ - sanity check

/usr/local/hadoop/bin/hadoop



## 11. הרצת דוגמה מוכנה

a. יש ליצור תיקיית קלט

```
mkdir ~/input
```

b. יש להעתיק לתוכה תוכן

```
cp /usr/local/hadoop/etc/hadoop/*.xml ~/input
```

c.

```
/usr/local/hadoop/bin/hadoop jar
```

```
/usr/local/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-3.2.2.jar
```

```
grep ~/input ~/grep_example 'allowed[.]'
```

d. יש לוודא שתיקיית הפלט שמסומנת בפקודה אינה קיימת - אחרת נקבל שגיאה

```
WARN impl.MetricsSystemImpl: JobTracker metrics system already  
initialized! org.apache.hadoop.mapred.FileAlreadyExistsException:  
Output directory file:/home/yaniv/grep_example already exists
```



## נספח

התחברות מרוחקת למכונה (אין חובה לעשות חלק זה, אולם ניתן להשתמש בו בכדי להתחבר למכונה בענן EC2 של AWS או שרת במכללה לדוג')

1. יש להתקין כלי רשת  
`sudo apt install net-tools`
2. יש להתקין שרת SSH לצורך התחברות מאובטחת מרחוק  
`sudo apt install openssh-server`
3. הכלי **systemctl** מאפשר להרים/להוריד/לדעת מה מצבם של סרוויסים שונים במערכת (ניתן גם להשתמש ב `service` או להשתמש בתיקיית הסרוויס)  
(באופן דומה ניתן להעלות ולבדוק סרוויסים אחרים)  
נבדוק שהשרת למעלה

`sudo systemctl status ssh`

במידה ולא יש להפעילו עם אחת מהפקודות הבאות:

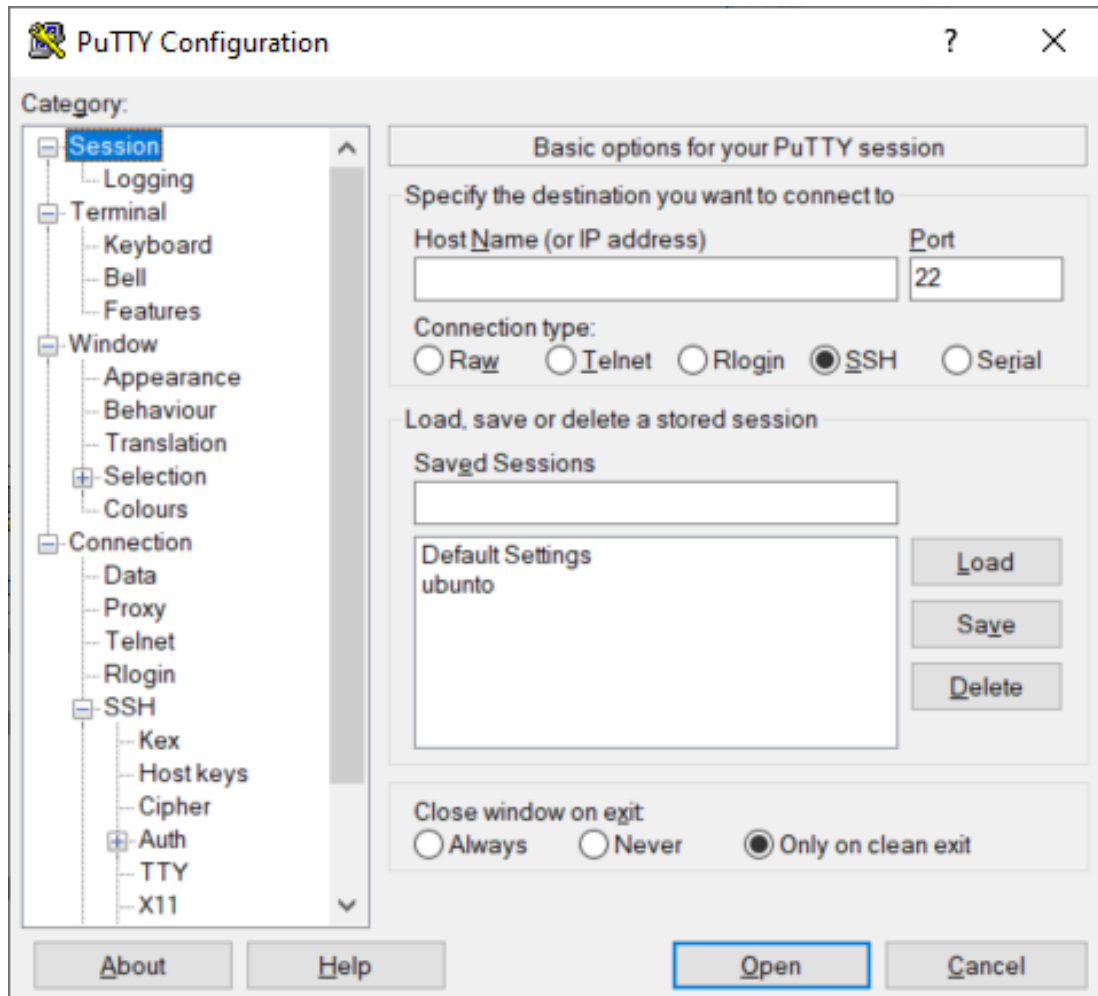
```
sudo systemctl start ssh
sudo service ssh start
sudo /etc/init.d/ssh start
```

4. בעזרת הפקודה `ifconfig` נזהה את ה IP שלנו

```
Processing triggers for initramfs-tools (0.130~deb11.1) ...
yaniv@ubuntu:~/Desktop$ ifconfig
ens33: flags=4163<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500
    inet 192.168.0.165 netmask 255.255.255.0 broadcast 192.168.0.255
    inet6 fe80::67d3:f787:89ad:37a9 prefixlen 64 scopeid 0x20<link>
    ether 00:0c:29:21:01:79 txqueuelen 1000 (Ethernet)
    RX packets 8726 bytes 8992073 (8.9 MB)
    RX errors 0 dropped 0 overruns 0 frame 0
    TX packets 4172 bytes 283966 (283.9 KB)
    TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0

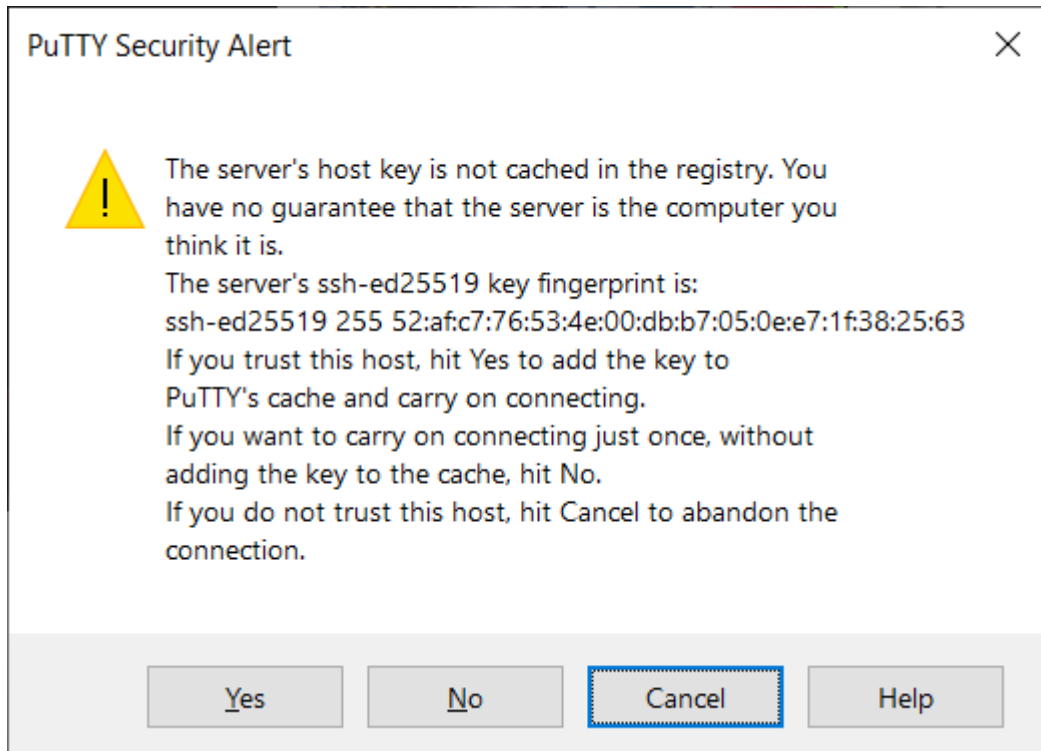
lo: flags=73<UP,LOOPBACK,RUNNING> mtu 65536
    inet 127.0.0.1 netmask 255.0.0.0
    inet6 ::1 prefixlen 128 scopeid 0x10<host>
    loop txqueuelen 1000 (Local Loopback)
    RX packets 248 bytes 21222 (21.2 KB)
    RX errors 0 dropped 0 overruns 0 frame 0
    TX packets 248 bytes 21222 (21.2 KB)
    TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0
```

5. צד הלקוח - נפעיל את putty (או כל טרמינל לבחירתכם)



את הIP של המכונה יש להקליד כ HOST  
(חיבור בענן - אם מתחברים בעזרת קובץ AUTH ל PUTTY זה קובץ PPK יש לתת נתיב עבורו  
בטאב של AUTH)

6. לאחר חיבור ראשון יש לאשר את ההודעה



7. יש לתת שם משתמש וסיסמא (שימו לב בעת הקלדת הסיסמא הסמן לא מתקדם)



8. למי שרוצה להתקין שרת Telnet במקום SSH (חיבור לא מאובטח) במקרה זה יש לשים לב להגדרות החיבור ב PUTTY בהתאם

```
sudo apt-get install telnetd -y
```

```
sudo systemctl status inetd
```