

מדריך מעבדה קורס Big Data

במעבדות נשתמש בשרת של האוניברסיטה. בשרת הותקנו עבורנו MongoDB, Jupyter Python, Redis, Hadoop, Hive, Impala ו-Spark (נלמד בהמשך הקורס!).

מטרת המדריך ללמד אתכם להתחבר לשרת באמצעות MobaXterm ולהפעיל jupyter python בשרת.

לפני שנתחיל צריך לדאוג שיהיה ברשותכם:

(1) כתובת IP של השרת: `bdl1.eng.tau.ac.il`

(2) שם משתמש: **באתר הקורס**

(3) סיסמה: **1qaz!QAZ**

תוכנה MobaXterm מותקנת במחשב (במעבדה 424 תוכנה מותקנת על כל המחשבים)

שם משתמש להתחברות לשרת ניתן למצוא באתר הקורס (moodle), כל סטודנט יקבל שם משתמש.

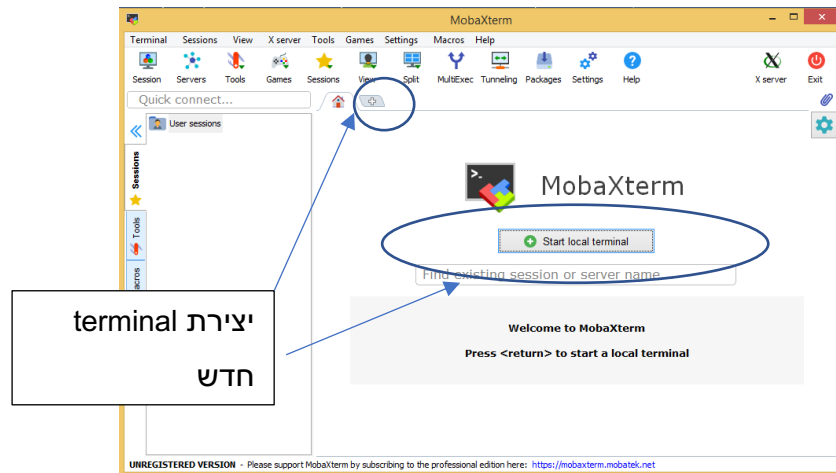
תוכנה MobaXterm

תוכנה (צד לקוח) חינוכית המשתמשת בפרוטוקול SSH (Secure Shell) המאפשר התחברות לשרת מרוחק בצורה מאובטחת. התחברות מאובטחת נדרשת מכיוון שאנו משתמשים בשם משתמש וסיסמה (כמובן שיש תוכנות נוספות שניתן להשתמש...)

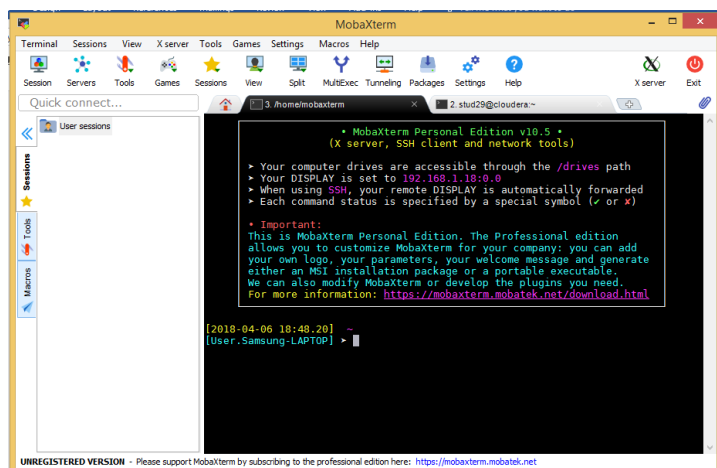
התקנת התוכנה: הורידו את [קובץ ההתקנה מהאתר](#), בחרו Installer edition (צבע ירוק). תחלצו את קובץ הzip ותבצעו את ההתקנה (שימו לב ששני הקבצים נדרשים להתקנה!). תפתחו את MobaXterm_installer_20.2 להתקנת התוכנה.

התחברות לשרת

פתחו את התוכנה MobaXterm :



באיור למעלה מסומן כיצד לפתוח טרמינל חדש ליצירת קשר עם השרת (יש עוד מספר דרכים).
לאחר לחיצה נקבל את המסך הבא :

התחברות מהבית

נדרש להתחבר ל-VPN ואז אפשר להתחבר לשרת.

התחברות ל-VPN :

<http://www.tau.ac.il/cc/helpdesk/communication/sslvpn/sslvpn.html>

במחשבים באוניברסיטה (מעבדה 424) אין צורך להתחבר ב-VPN.

התחברות לשרת תתבצע באמצעות הפקודה הבאה :

ssh <username>@<IP>

לדוגמא :

ssh stud1@bdl1.eng.tau.ac.il

לאחר מכן תתבקשו להזין סיסמה, תזינו את הסיסמה הראשונית שתקבלו באתר הקורס.

ברכות, כעת את מחוברים לשרת!

קבצי תרגול בשרת

כתבו את הפקודה ls. כעת ניתן לראות את התיקיות הבאות :

- תיקיית bigdata_lab מכילה סקריפטים של ipython שנשתמש בשיעורי המעבדה
- תיקיית test מכילה קבצי נתונים שנשתמש במעבדה
- שאר הקבצים לא נצטרך לתרגול

פקודות בסיסיות ב-LINUX

כעת אתם מחוברים לשרת LINUX שנמצא בענן, ונכיר פקודות בסיסיות:

פקודה	הסבר
pwd	מציג את הניתוב לתיקייה הנוכחית
ls	הצגת תוכן התיקייה
cd /directorypath	מעבר לתיקייה
mkdir	יצירת תיקייה
rm -r directoryname	הסרת קובץ או תיקייה
cp filelocation copylocation	העתקת קבצים
cp -R filelocation copylocation	העתקת תיקיות
mv filelocation newlocation	העברת תיקייה
touch filename	יצירת קובץ חדש
cat filename	מציג תוכן של קובץ
echo hello world>> new.txt	הוספת טקסט לקובץ

העברת קבצים ממחשב המקומי לשרת:

נפתח טרמינל חדש ללא התחברות לשרת, ונריץ את הפקודה הבאה על המחשב המקומי:

```
scp -r <fromlocation> <tolocation>
```

כאשר -r נועד להעתיק תיקיות

לדוגמה: העברת קובץ text.txt ממחשב מקומי לשרת (כתובת במחשב המקומי הינה: C:\Users\User\Documents\MobaXterm\home).

```
scp -r text.txt stud1@bd11.eng.tau.ac.il: /home/stud1/text.txt
```

העברת קבצים משרת למחשב מקומי:

בטרמינל המקומי ולא בשרת נריץ את אותה פקודה אבל הסדר של מיקומים שונה.

לדוגמה: העברת קובץ text.txt משרת למחשב מקומי

```
scp -r stud1@bd11.eng.tau.ac.il: /home/stud1/text.txt .
```

נקודה (.) מסמן מיקום של תיקייה home.

תרגיל

- (1) צרו תיקייה בשרת שלכם בשם myhello
- (2) הציגו את התיקיה
- (3) צרו קובץ txt ריק בשם hello במחשב המקומי
- (4) העבירו את hello לתיקייה שיצרתם בשרת
- (5) בשרת הוסיפו את הטקסט "hello world"
- (6) בדקו והציגו שהטקסט נוסף לקובץ
- (7) שנו את שם הקובץ ל-helloworld
- (8) הציגו את הקבצים הנמצאים בתיקייה myhello
- (9) הורידו את הקובץ מהשרת למחשב המקומי
- (10) מחקו את תיקייה myhello מהשרת

פתרון

תחילה מתחברים לשרת (נדרש להחליף לשם משתמש שלכם!):

```
ssh stud1@bdl1.eng.tau.ac.il
```

ניצור תיקייה:

```
mkdir myhello
```

```
ls
```

יוצרים קובץ hello.txt בתיקייה C:\Users\User\Documents\MobaXterm\home

פותחים terminal חדש וכותבים את הפקודה:

```
scp hello.txt stud1@bdl1.eng.tau.ac.il: /home/stud1/myhello/
```

או

```
scp hello.txt stud1@bdl1.eng.tau.ac.il: ~stud1/myhello/
```

חוזרים לטרמינל של השרת:

```
ls myhello
```

```
cat myhello/hello.txt
```

הוספת טקסט:

```
echo hello world >> myhello/hello.txt
```

```
cat myhello/hello.txt
```

שינוי שם הקובץ:

```
mv myhello/hello.txt myhello/helloworld.txt
```

```
ls myhello
```

חוזרים לטרמינל המקומי :

```
scp stud1@bd11.eng.tau.ac.il: ~stud1/myhello/helloworld.txt .
```

הולכים לתיקיה C:\Users\User\Documents\MobaXterm\home ובעת ניתן לראות את הקובץ helloworld.txt.

לבסוף מחיקת התיקיה בשרת :

```
rm -r myhello
```

הרצת Jupyter Notebook

לאחר התחברות לשרת ניתן להריץ את הפקודה הבאה :

```
jupyter notebook
```

לאחר מכן, לפתוח דפדפן (אם לא נפתח אוטומטית) ולכתוב את כתובת URL הבאה :

```
bd11.eng.tau.ac.il:####
```

מספר הפורט #### ניתן למצוא בטרמינל חפשו כתובת http://bd11.eng.tau.ac.il:####/.

הזינו את הסיסמה הראשונית שקיבלתם.

שימו לב: לאחר סיום שימוש חובה לסגור דרך הטרמינל את ה-jupyter ע"י CTRL+c פעמיים.

סיכום

השלימו את הפרטים :

ssh <username>@<IP> ssh stud##@bd11.eng.tau.ac.il	התחברות לשרת
stud__	שם משתמש
1qaz!QAZ	סיסמה ראשונית
jupyter notebook	jupyter

המלצה: לאחר כל שעת עבודה בשרת הורידו את הקבצים למחשב המקומי!