**מפגש פורום HPC**

24.2.2022

11:00-12:00

**אג'נדה**

11:00-11:15 – אלי שמואלי – הדרכות מוכנות מ- PRACE

11:15-11:30 דיון פתוח בנושאים:

1. סקירה לגבי תצורת העבודה על גבי הקלאסטרים הקיימים באוניברסיטאות השונות , וכן לגבי מנגנוני חישוב עלות שימוש
2. האם קיימת מדיניות לגבי גיל מקסימלי למכונות בצביר (מכונות בכלל ומכונות שנרכשו ע"י חוקרים)
3. האם קיימת מדיניות לגבי שיתוף מכונות
4. העדר אנשי מקצוע טובים לסביבת HPC
5. אינטגרציה בין HPC ומחשוב ענן ציבורי כגון AWS, Azure, GCP.
6. ניהול capacity של חדרי מחשב.

**משתתפים** (15):

* אלי שמואלי, מחב"א
* הנק נוסבכר, מחב"א
* ואדים מלקין, מכון ויצמן
* אמיחי סילברמן, טכניון
* אנטוני וולר, בר-אילן
* בני פרידמן, או"פ
* דני שקד, תל-אביב
* דרור פריגוזין, בר-אילן
* חיים אבודרהם, אב"ג
* אינגה פסטר, אב"ג
* נועם נודלמן, בר-אילן
* סמיון צרנין, אב"ג
* שי בנישו, בר-אילן
* יאיר גלבוע, חיפה
* יהבי בורבין, העברית

**פרוטוקול**

* אלי הציג הדרכות דרך PRACE:
	+ PRACE Training Portal
	+ <https://training.prace-ri.eu/>
	+ קורסי ה- MOOCS של PRACE
	+ <https://www.futurelearn.com/partners/prace>
* אב"ג: יש להם קלסטר ישן, מכרז בקרוב, לא מצאו מערכת ניהול טורים טובה, בסוף החליטו על Run:AI
* בר-אילן: כ- 400 שרתים, כל חוקר בונה לעצמו קלסטר, אין קלסטרים בבעלות מרכז החישובים, יש קלסטר בבעלות Data Science Instituteשנותן לחוקרים להוסיף שרתים משלהם
* ויצמן: הקלסטרים הם בבעלות מרכז החישובים עם ניצול לשל 60% עבור CPU, וניצול של 50% עבור GPU. כ-550 שרתים, תשתית IB מבוססת מתגי EDR/HDR, מערך אחסון 3PB איסילון ועוד 3.5PB GPFS. שאיפה לא לקנות שרתים על שם של מדען אלה לתת למדען x מיליוני שעות בשנה. זה שלNvidia ישראל אין ניסיון בבניה של קלאסטרים היה ברור, ניסינו יחד עם אמת מחשוב לבנות מערכת מבוסס פרויקט deepops
	+ <https://github.com/NVIDIA/deepops>

בזמנו לא היה בשל, היו בעיות חמורות של אבטחת מידע, בסופו של דבר עברנו לLSF של IBM.

* טכניון: 2 קלסטרים, כ- 160 נודים, חוקרים קונים נודים שמתווספים לקלסטר, משתמשים ב- PBS עבור ניהול טורים, ניסיון רע כשהרכש נעשתה לא דרכם ולא באופן מנוהל נכון, לקח: תכנון לפני רכישה
* חיפה: כ- 2000 קורים, קונים עוד 6000 קורים בעתיד הקרוב, בעיות חום בשרתים, 11 ארונות, יתירות נמוכה, בעיה בחשמל –רק 100KV, אין פתרון ליציאה לענן, מי שמשקיע בקלסטר מקבל עדיפות
* או"פ: קלסטר בת 4 שנים, רק קלסטר אחד מנוהל ע"י מרכז החישובים, 2000 קורים, כמה כרטיסי NVIDIA, ניהול טורים ע"י SLURM, לא גובים כסף – מימון מרכזי. מי שמשקיע בקלסטר מקבל עדיפות על החומרה שרכש.
* העברית: הוחלט על מרכז שבהקמה, 1.2MV(נכון?!), ניהול טורים ע"י SLURM, חוקר קונה אשראי כל חודש
* תל-אביב: יש אשכול פרטי ויש אשקול מרכזי, כ- 12,000 קורים, בצעו PoC על Run:AI , משתמשים ב- PBS חינמי – ללא תמיכה, כ- 300KV ב- 30 ארונות
* ניסיון מכון ויצמן בענן: מצריך כ"א יותר יקר מאשר כ"א HPC רגיל, אי אפשר להחזיק כ"א לכל שלושת סביבה ענן – רק לאחד אפשר, מימון עבור שרותי ענן מגיע מרכזי מהמכון, 8 מעבדות כבר עברו לסביבה ענן, אין ל- AWS מספיק GPUים באירלנד (200) ולכן חייבים לבצע העבודה בארה"ב
* ניסיון הטכניון בענן: עלות GPU יקר לעומת עלות CPU שהיא זולה, מקבלים 37% הנחה ממיקרוסופט עבור Azure, איש מיקרוסופט, בשם רועי בן חיים עוזר לחוקרים להסב את התוכנות שלהם לעבור בסביבת ענן Azure

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **שם מוסד** | **כ"א HPC קיים** | **כ"א HPC נדרש** | **כמות ארונות** | **צריכת חשמל** |  **סה"כ קורים** |
| בן גוריון | 3 | 5 | 9 |  | 8352 |
| בר אילן | 4 |  | 20 |  | 8076 |
| מכון ויצמן | 8 | 10 | 19 | 400KW | 21000 |
| טכניון | 3 | 4 | 10 | 120KW | 13000 |
| חיפה | 2 | 5 | 11 | 360KW | 2000 |
| הפתוחה | .5 | 1.5 | 3 |  | 1712 |
| העברית | 1.5 | 4 | 5 יגדל ל- 38 | תכנון ל- 1.2MW |  |
| תל אביב | 5 | 7 | 30 | 300KW | 12000 |