



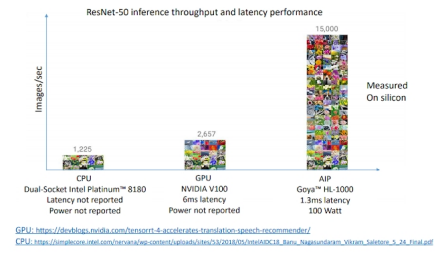
של דוגמאות לשלוח התחילה, שבבים שמתכננת, לאבס הבנה הישראלית ההזנק חברת, Habana Labs, מעבד ההיסק הייעודי שלה ללקוחות נבחרים בספטמבר 2018, יחד עם היציאה של החברה ממצב חשאיות.

איתן מדינה, המנהל העסקי הראשי של Habana, טוען שהשבב HL-1000 הוא עכשיו "מעבד ההיסק עם הביצועים הכי חזקים בתעשייה". הוא מוצע ללקוחות בכרטיס PCIe שנקרא גויה. לפי הניסויים הפנימיים של זה, ResNet50 תמונות זיהוי של הביצועים בבוחן בשנייה תמונות 15,012 להיסק יכול גויה, Habana, בהחלט יזכה אותו בשיא העולמי. והתוצאות נפלטות עם השהיה של רק 1.3ms, שזה יותר ממספיק לאינטראקציה בזמן אמת.

היות שכרטיס הגויה צורך רק כ-100 ואט כשהוא מעבד מודלים של ResNet-50, נצילות האנרגיה בפעולה הזאת היא בסדר גודל של 150 תמונות לשנייה לוואט, וזה נראה עוד שיא בבוחן הביצועים. המעבד הגרפי הטוב מסוגו V100 מספק בערך 18 תמונות לשנייה לוואט, ואילו המעבד Xeon עם שני השקעים מגיע לכ-6 תמונות לשנייה לוואט.

למרות נתוני הביצועים הלא מרשימים של המעבדים המרכזיים, שם מתבצעת רוב ההסקה בימים אלה. זה בעיקר בגלל העובדה שהסקה היא פעילות עם נפח גבוה, שבדרך כלל מסתמכת על ביזור בקנה מידה גדול על פני תשתית בגודל של ענן. בנוסף, מודלים של הסקה באמצעות רשתות עצביות מצריכים הרבה פחות עוצמת חישוב מאשר אימון המודלים האלה, שרובו נעשה במעבדים גרפיים מובחרים זוללי הספק כמו מאופטימליים פחות להיות נוטים אבל בהסקה טובים האלה המתקדמים הגרפיים המעבדים גם V100. לסביבות בגודל עצום. וזאת הסיבה שיצרני שבבים רבים (כולל אנבידיה) מתייחסים להסקה כהזדמנות נפרדת.

ResNet-50 Inference performance 



{loadposition content-related}