תרגיל בית 2 – מאמר

סער גולן

205717077

**שם המאמר:**

Provenance as First Class Cloud Data

**מטרת המאמר:**

מטרת המאמר היא להראות את התועלת והיתרונות בשימוש של provenance (שהוא מטא-דאטה שמתאר מקור של אובייקט מסויים בענן) בספקי אחסון בענן, להראות את אי השימוש בו כיום ולהציג דרישות שבעזרתן provenance יהיה נגיש ושימושי לספקי ענן.

תיאור טכנולוגיות\מחקרים במאמר:

* Simple Storage Service (S3) – שירות אחסון של אמזון שמאפשר לשמור אובייקטים שגודלם עד 5GB, בעזרתו ניתן לאחסן את הprovenance. קיים שירות ייעודי המאפשר לאחסן provenance שהוקלט (Amazon’s SimpleDB).
* Microsoft Azure Blob – שירות דומה לS3 עם הבדלים קטנים – קיימת הגבלה על מטא-דאטה, ועל גודל של PUT (לעומת S3). גם פה קיים שירות דומה לאחסון provenance – Azure Table service.
* Nirvanix Internet Media File System (IMFS) – מערכת לניהול קבצים מבוזרת שניתן לגשת אליה דרך האינטרנט. כמו Azure Blob וS3, מערכת זו נועדה לטפל בקבצים גדולים שנכתבים לעתים רחוקות ונקראים לעתים קרובות. בניגוד לשתי הטכנולוגיות שתוארו למעלה, זו מספקת ממשק לניהול הקבצים.

**סיכום המאמר:**

המאמר מציע להשתמש בprovenance ובמידע שהוא מספק בצורה טובה יותר.

Provenance מספק הצגה מפורטת של מידע ומקורו של אובייקט בענן. המאמר מציג תחילה את השימושיות של הprovenance ומדוע צריך ספקי ענן צריכים להתייחס לprovenance כמידע חשוב.

המידע שמגיע מהprovenance מספק את ההתנהגות ה"נורמלית" של בענן של המשתמש, לכן הוא יכול לעזור ולזהות אנומליות, בצורה דומה לחברות אשראי שמזהות אנומליות בחיובי כרטיסי אשראי ומתריעות על כך. בנוסף, ניתן להשתמש במידע זה לחיפוש שיותאם למשתמש, ככל שיותר ויותר מידע ישותף ויאוחסן בענן.

המאמר מציג את שירותי הענן הקיימים היום -

* Simple Storage Service (S3) – שירות אחסון של אמזון המאפשר לשמור אובייקטים שגודלם עד GB5 (בעזרתו ניתן לשמור את הprovenance).
* Microsoft Azure Blob – שירות דומה לS3 (של מייקרוסופט) אך עם הבדלים קטנים.
* Nirvanix Internet Media File System (IMFS) – מערכת לניהול קבצים מבוזרת שניתן לגשת אליה דרך האינטרנט.

אך גם עם כל אלה אף אחד מהשירותים הקיימים כיום לא מיועד בצורה מפורשת לתמיכה ואחסון של provenance. במקרה הטוב משתמשים יכולים להתייחס לprovenance כמטא-דאטה ולאחסן אותו בנפרד בdb, דבר שלא יאפשר לנצל את התועלת שלו.

כדי לאפשר להתייחס לprovenance בצורה בה המשתמשים יוכלו לנצל את המידע ממנו באופן השימושי ביותר, עדיין הprovenance צריך לקיים כמה דרישות:

* שיתוף פעולה בין יכולות אחסון ומחשוב בענן (למשל אמזון מציעה שירות מכונה וירטואלית שמספק למשתמשים כוח חישוב לפי דרישה – כך המכונה הווירטואלית יכולה לייצר provenance שיועבר לאחסון ביחד עם המידע)
* לאפשר ללקוחות להקליט provenance באופן מקומי.
* לאפשר להשתמש בprovenance באופן חיצוני, כלומר לאפשר למשתמשים לאחסן את זה.
* שמירה על ביטחון ופרטיות – provenance יכול להכיל מידע רגיש.

לסיכום, כותבי המאמר טענו שצריך להתייחס לprovenance בצורה טובה יותר שתאפשר לייעל את השימוש בו. דבר זה יטיב עם המשתמשים שיוכלו להפיק תועלת ממידע זה , ויטיב גם עם ספקי הענן שיוכלו לנצל את המידע העשיר שעובר דרך הענן.

**האם מטרת המאמר הושגה:**

מטרת המאמר הושגה. כותבי המאמר מנסים להעלות את חשיבות המידע המגיע מprovenance ומציעים לפתח שירותים ייעודיים לו כדי לייעל את השימוש לו. במאמר מציגים את השירותים הקיימים כיום שלדברי כותבי המאמר לא מספקים, ומציעים דרכים ורעיונות כיצד ניתן להשתמש במידע בצורה שתטיב עם ספקי הענן ועם המשתמשים וכך שני הצדדים יתייחסו יותר למידע זה וחשיבותו תעלה.

**הצעות נוספות שלי:**

פיתוח מנגנון שיאפשר המרה של provenance ונגישות שלו בין ספקי ענן שונים. מחד, מנגנון כזה עדיין צריך לעמוד בדרישות שהוצגו במאמר (לרבות שמירה על ביטחון המידע והפרטיות) ומאידך יאפשר גמישות רבה יותר למשתמשים ויעודד שיתוף פעולה בין ספקי הענן השונים.