DB Project Design

תיאור דרך המימוש של הDB

אופי הDB:relational, row oriented רלציוני מונחה שורות.

בחרנו דווקא בסוג זה כי ברצוננו לתמוך ברשומות בעלות סכמה קבועה בדגש על טבלאות דטרמיניסטיות. הקו המנחה בבחירות השונות של מימוש הDB הינו יעילות קריאה וכתיבה לדיסק – O/I שאלו הפעולות היקרות והמשמעותיות בעבודה עם מסדי נתונים.

היתרונות של DB מונחה שורות: תמיכה ביעלות ב- O/I בפעולות הנעשות פר שורה. חסרונות: אגריגציות כבדות וחישובים יעילים פחות על הדטה.

הDB מהיר במיוחד בפעולות אלו ופחות מתמקד ביעילות של פעולות אגריגציה ופעולות פר עמודה, אם כי כמובן תומך גם בפעולות אלו. כפי שיתואר בהרחבה בהמשך הדוקומנטציה.

אופי הדטה המומלץ לעבודה עם הDB הינו מספרים, תווים, מחרוזות בעלות אורך קבוע וכדומה.

**implementation properties:**

אחסון הטבלאות בדיסק: כל טבלה תשמר בנפרד בקובץ csv ששמו יהיה שם הטבלה, קבצים אלו מהירים וזולים בקריאה וכתיבה לדיסק וכן אופיים מותאם לגישה טבלאית מונחית שורת, לפיכך בחרנו בהם.

פירוט הפעולות הנתמכות ודרך מימושם:

* **הגדרת טבלה**

יצירת קובץ csv כשם הטבלה ואתחול השורה הראשונה לשמות השדות. עלות: (1)O (גישה יחדה לדיסק)

* **מחיקת טבלה**

מחיקת הקובץ שנוצר בעת הגדרת הטבלה. עלות: (1)O (גישה יחידה לדיסק)

* הכנסת רשומה לטבלה

כתיבת השורה על פי הסדר הסכמתי של השדות לסוף הקובץ בו שמורה הטבלה. עלות: (1)O (גישה יחידה לדיסק)

* **מחיקת רשומה מטבלה**

קריאת הקובץ בו שמורה הטבלה, איתור הרשומה הרלוונטית בה וכתיבת ערכי deleted לכל האטריבוטים ברשומה. נעדכן deleted\_counter במידה והמונה עומד על 40% מהקובץ, נקרא את הקובץ ונכתוב ותו מחדש ללא השורות המחוקות. עלות: (n)O איתור הרשומה המבוקשת, במקרה הגרוע נדרשת כתיבה של הקובץ כולו לדיסק.

* **מחיקת מספר רשומות לפי תנאי/ים מסוימים מטבלה**

קריאת הקובץ בו שמורה הטבלה, איתור הרשומות המבוקשות ומחיקתן כפי המוסבר בסעיף מחיקת רשומה בודדת. עלות: (n)O איתור הרשומות המבוקשות, במקרה הגרוע נדרשת כתיבה של הקובץ כולו לדיסק.

* **עדכון רשומה בטבלה**

קריאת הקובץ בו שמורה הטבלה, איתור הרשומה הרלוונטית בה ועדכון השדות הנדרשים. עלות: (n)O – עלות איתור הרשומה המבוקשת, קריאה וכתיבה לדיסק.

* **חיפוש רשומות העונות על תנאי/ים מסוימים בטבלה**

קריאת הקובץ בו שמורה הטבלה ואיתור הרשומות העונות על התנאים. עלות: (n)O איתור הרשומות המבוקשות, קריאה וכתיבה לדיסק.

**4 main consideration**  
**·   Performance**

יחסנו משקל רב ליעילות פעולות קריאה וכתיבה מהדיסק (O/I) לפיכך השתמשנו בקבצי csv שהינו זול בפעולות אלו. בנוסף בtrade-off בין מהירות פעולת ההכנסה למהירות חיפושים וחישובים על הדטה הכרענו לטובת פעולת ההכנסה.  
**·   Space Utilization**

בפעולת המחיקה יצרנו איזון בין חיסכון במקום לחיסכון במשאבי זמן.  
**·   Complexity of DB**

קל לשימוש אך קשה למימוש ☺   
**·  Flexibility**

הDB יחסית קשיח מאופיו כrelational DB . גמישות אינה הצד החזק שלו. אם רצונך ב DB עקבי ודטרמיניסטי, נשמח לשרת אותך באמינות, במידה ואתה מחפש גמישות – חפש בגוגל אחר המתחרים הלא רלציונים ;)

**Pros & cons:**

|  |  |
| --- | --- |
| **Cons** | **Pros** |
| חיפוש ומחיקה עלולים לארוך זמן | מסד נתונים דטרמניסטי ואמין, מהיר בהכנסה |