

תפקיד ה-DBA במחשוב ענן



מאת: אביאל אילוז (בסיוע של דוד יהלום).

תפיסת מיחשוב הענן רואה את החומרה ואת התוכנה (מערכות הפעלה, אפליקציות, שפות תכנות) כשירותים המסופקים מתוך הענן לכל דורש, ומסדי נתונים אינם יוצאים דופן; הם מספקים מצע איחסון (בתוך הענן) ומנוע ניהול נתונים, כשירות לכל דבר (DaaS = Database-as-service). עם זאת, תחום מסדי הנתונים במיחשוב הענן מחייב התייחסות מיוחדת, משום שהוא מציב אתגרים מיוחדים שאינם קיימים לגבי שירותי ענן אחרים.

מהם היתרונות העיקריים במעבר מסדי נתונים אל הענן?

היתרונות של מעבר מסדי נתונים אל הענן הינם בכמה תחומים. בתחום הכלכלי ישנה הפחתה בעלויות. ההוצאות למשאבי חומרה ותוכנה ולהעסקת צוותי IT יורדות באופן משמעותי. ספציפית לגבי נפחי איחסון, המחיר ליחידת נפח בענן נמוך בהשוואה לאיחסון מקומי. חברות המספקות שירותי ענן מציעות מודל של "שלם כשתצטרך" (Pay-As-You-Go) וכך אין צורך ברכישת נפחים מעבר לדרוש לנפח המידע שבמסד הנתונים. הצורך בצוות IT מיומן שעוסק בתחזוקה, בשידרוגים ובמעקב אחר ביצועי החומרה פוחת, ומאפשר לחסוך בעלות זאת, או לנצל למשימות אחרות.

אך לדעתי, יתרון זה איננו הראשון במעלה לגבי מסדי נתונים, אלא דווקא יתרון יכולת ההתרחבות הדינמית או סקלביליות דינמית. אחת הדרישות העיקריות ממסדי נתונים, במיוחד בנפחים גדולים, היא האפשרות להתרחב או לבזר את הנתונים כדי לשפר ביצועים או כדי לתת מענה לצורך בגדילה, וזאת במינימום או בהעדר מוחלט של פגיעה בפעילות השוטפת (no downtime). הענן מקנה את היתרון הזה באופן מוחלט. לא זו בלבד שניתן להתרחב ע"פ הצורך באופן דינמי וללא השפעה על הפעילות, אלא שתשתיות החומרה עצמן, למשל אלו של פלטפורמת Azure של מייקרוסופט (SQL Server ו-Windows) ממוקמות למעשה ברחבי העולם (אמריקה אירופה ואסיה). כך שתשתית הנתונים מתרחבת ומבזרזת ללא מאמץ של ניהול או חשש מקריסה ולאיזורים גאוגרפיים שונים, ע"פ הדרישות העסקיות. סקלביליות זו כמובן תורמת רבות לביצועי מסד הנתונים כשיש צורך לנהל נתונים בנפחים גדולים.

יתרון זה מביא על גביו יתרון נוסף למסדי נתונים בענן והוא הזמינות היתרה (High Availability). במערך איחסון מבזר, סקלבילי וענק כמו הענן, יורדת למינימום האפשרות לאי-זמינות של מסד הנתונים.

ובכל זאת, למרות היתרונות האלו, מהם האתגרים במעבר מסדי נתונים לענן?

אכן, למרות יתרונות אלו, ישנם כמה אתגרים. ראשית, זהו האתגר הכללי של המעבר לענן, שהוא תהליך מורכב המחייב היערכות נרחבת של האירגון להעביר את תשתיותיו, והעיקרי שבהן הוא מסדי הנתונים ומאגרי המידע, אל הענן. בכנס של פורום C3 של "אנשים ומחשבים" ברודוס לפני כחודש, **אבי שלום**, מנהל תוכנית EXP בגרנטר המזרח התיכון, מיתן את הציפיות ואמר שהתהליך יהיה ארוך ויקר. לדעתי, גם אם התהליך יארך, ברור שהדבר הבא הוא "מיחשוב ענן" וארגונים רבים יעשו את המעבר.

אלא שכאן יש אתגר גדול יותר לגבי מסדי הנתונים עצמם. ארגונים רבים לא מרגישים נוח עם העובדה שמאגרי הנתונים העסקיים שלהם לא יושבים קרוב אצלם, אלא אי שם באיזשהו ענן אמורפי. קשה לדמיין ארגון ביטחוני, או כל ארגון אחר המחזיק בנתונים חסויים, שיסכים להעביר את מסדי הנתונים שלו אל הענן. משום כך, מבחינים בין "ענן ציבורי" ל"ענן פרטי". ארגונים בעלי מידע חסוי, יממשו פתרונות ענן במסגרת פרטית, שבו ענן אחד מספק את כל שירותי הארגון, ובכלל זה שירותי מסדי נתונים, ומקנה במידה מספקת את היתרונות

הכללים של הענן. ארגונים בעלי מידע ציבורי לגמרי, כמו מכוני מחקר מדעיים ואחרים, ימצאו את המעבר לענן, עם מסדי הנתונים, יותר קל ופחות בעייתי. עם זאת, גם אלו שיעבירו את מסדי הנתונים לענן יצטרכו להתמודד עם אתגרי אבטחה חדשים ולהפעיל, למשל, הצפנות נתונים ברמה מתקדמת יותר, ואבטחת גישה קשיחה יותר.

[במפגש הראשון של פורום DBA](#) של "אנשים ומחשבים" נדגים את ההתמודדות בנושא האבטחה בעולם ענן בעזרת כלי של חברת Sentrigo.

איך ישפיע מעבר מסדי נתונים לענן על תפקיד ה-DBA?

נראה שאכן יש חשש מסוים בקהילת ה-DBA לגבי נחיצותם של DBA בסביבת הענן, כפי שעולה ממאמרים ומפורומים מסוימים. אבל, לדעתי, החשש הזה מוגזם ומיותר; לא נראה לי שמסדי נתונים בעולם הענן מיייתרים את תפקיד ה-DBA, אלא משנים את הדגשים בעבודתם. נכון הוא שבפלטפורמת ענן, DBA לא יעסקו יותר בניטור תשתיות ובניהול משאבי חומרה, וכל השיקולים התשתיתיים הקשורים לאופטימיזציה (CPU, RAM, IO), יהפכו ללא-רלוונטים, שהרי הכל מטופל באופן שקוף בענן. אבל זהו דווקא יתרון, כי מעתה, ישקיעו DBA את זמנם במשימות אחרות שמטרתן להבטיח מסד נתונים אפטימלי. ראשית, על ה-DBA מוטלת המשימה לטעון נתונים אל תוך המסד הענני ולסנכרן אותם במקרה הצורך. כמו כן, לא יעלם הצורך בעיצוב נכון ויעיל של מסד הנתונים, ניסוח נכון של שאילתות, אופטימיזציה של שאילתות, אסטרטגיית אינדוקס, תחזוקת אינדקסים ומשימות דומות נוספות, שהן בסופו של דבר ה"לחם והחמאה" של ה-DBA ובלעדיהן ביצועי מסד הנתונים יהיו ירודים, על אף הענן המרענן.

על זה כמובן נוסף אתגר האבטחה ומשימות ה-DBA בתחום זה יתרחבו לצורך מימוש טכנולוגיות הצפנה והקשחת גישה, דבר שבלי הענן, היה אולי נדרש פחות. גם בנושאי תשתית ותחזוקה יש ל-DBA תפקיד; עצם העובדה שמסד הנתונים יושב בענן מרוחק ואמין, איננה ערובה מלאה לזמינות מוחלטת, ועל ה-DBA מוטלת החובה לדאוג לזמינות ע"י שמירת גיבויים לצורך שחזור בעת פקודה. חוץ מזה, אל לנו לשכוח שבקרב ספקי שירותי הענן עצמם, יש צורך ב-DBA מקצועיים ומנוסים כדי לטפל בתשתיות הענק שלהם ולהבטיח את היתרונות המוצהרים - סקלביליות, זמינות ואמינות - ואין ספק שצורך זה יגדל בעתיד.

מייקרוסופט מציעה היום את SQL Azure כשירות מסדי נתונים בענן, מה תוכל לספר על המוצר?

SQL Azure הוא למעשה SQL Server שהותאם לסביבת הענן. כל הפעולות העיקריות המוכרות ל-DBA בעיצוב מסדי נתונים ניתנות לביצוע ב-SQL Azure: הקמת מסדי נתונים, טבלאות, אובייקטי קוד, אינדוקס, אבטחה וכו'. אך בניגוד לגרסת SQL Sever הרגילה המקומית, גירסת Azure לא מאפשרת, ובצדק – כמתחייב מתפיסת הענן, לנהל או להגדיר הגדרות הקשורות לתשתית מסד הנתונים כמו מיקום, Partitioning, ניצול משאבים, מנגנוני High Availability וכו'.

החיבור אל מסד נתונים כזה מתבצע ע"י הכלים הרגילים המוכרים של SQL Server, אם כי עדיין במגבלות מסוימות. גיבויים ותהליכי תחזוקה שונים כרגע לא ניתנים לביצוע, אך יתכן שבגרסאות עתידיות תתאפשר גם פעולה זאת. כצעד ראשון, זוהי התחלה טובה המאפשרת התנסות ראשונית וקלה יחסית במעבר מסדי נתונים אל הענן.

[במפגש הראשון של פורום DBA](#) של "אנשים ומחשבים" נדגים הקמה וניהול בסיסי של מסד נתונים ב-SQL Azure

אילו טכנולוגיות מציעה אורקל למסדי נתונים בסביבת הענן?

אורקל מאפשרת לארגונים לייצר "עננים" במצעות טכנולוגיות ה-Grid Computing הן ברמת האפליקציה והן ברמת מסדי הנתונים. בתשתית ענן של מסדי הנתונים אורקל מציעה שלל טכנולוגיות יחודיות (כמו RAC, In Memory Database Cache-i Automatic Storage Management) המאפשרות לארגונים להקים פתרונות המבטיחים סקלביליות דינמית, זמינות יתרה (High Availability) וביצועים גבוהים, כמצופה מפלטפורמות הממוקמות בענן. זאת בנוסף לכלים המאפשרים הקמה מהירה של פתרון ענן שלם הכולל חומרה ותוכנה (כמו ExaLogic-i ExaData).

לאורקל גם טכנולוגיה ייחודית לניטור הפעילות בענן המשמשת את ה-DBA וצוות ה-IT. באמצעות **Oracle Grid Control** ניתן לנטר ולבצע מעקב פרו-אקטיבי על כל מערכות המידע של הארגון, גם אלו שבענן, ממקום מרכזי אחד.

ומה בנוגע לאינטגרציה בין מערכות המידע בענן לאלו הנשארות "קרוב"?

האפשרות לסינכרון המידע בין מסדי הנתונים המקומיים לאלו שבענן, הוא הרכיב החשוב המאפשר קודם כל להקים תשתית נתונים בענן, ולאחר מכן לשמור על העדכניות שלה.

לשם כך מציעה מייקרוסופט כחלק משירותי **SQL Azure** את השירות **SQL Azure Data Sync** המאפשר לסנכרן נתונים בין הענן למערכות פיזיות בארגון או בתוך הענן עצמו. כמו כן ניתן להשתמש בכלים המוכרים הרגילים כמו **SSIS** ו**BCP**.

אורקל מצידה מציעה את טכנולוגיית **GoldenGate** היחודית המאפשרת לארגונים קטנים כגדולים, לבצע רפליקציה הטרורגנית וגמישה בין מסדי נתונים מפלטפורמות שונות בזמן אמת. רפליקציות **GoldenGate** יכולות להיות דו כיוונית ובין מסדי נתונים פיזיים בארגון למסדי נתונים הממוקמים בענן.

השימושים בטכנולוגיית **GoldenGate** מגוונים מאוד: הזרמה סלקטיבית וראשונית של המידע מהמערכות המקומיות למערכות בענן, רפליקציה לצורך "השקה רכה" של הפלטפורמה הוירטואלית (שמירת עותק מקומי מסונכרן בזמן אמת מול עותק "ענני"), הקמת מערך **DRP** מסונכרן בענן ופרויקטים בתחום ה-**Load Balancing** ו-**RealTime BI** המאפשרים עבודה במקביל בשתי הסביבות, ועוד קונפיגורציות רבות נוספות – כיד הדמיון.

[במפגש הראשון של פורום DBA](#) של "אנשים ומחשבים" נדגים שימוש ברפליקציה הטרורגנית של **GoldenGate** בין **SQL Azure** לאורקל.