



Moshav Bnei Zion P.O.Box 151, 60910 Israel Tel. 972-9-7907000 Fax. 972-97442444



סיכום מפגש שולחן-עגול

טכנולוגיות ומתודולוגיות בעולם האחסון

מנחה
פיני כהן

לקוחות נכבדים שלום,

תודה על השתתפותכם במפגש שולחן עגול Round Table בנושא טכנולוגיות ומתודולוגיות בעולם האחסון.

מצ"ב סיכום עקרי הדברים שעלו במהלך המפגש. במפגש עלו נושאים מהותיים שתומצתו בסיכום כפי שעלו. אין בסיכום זה המלצה גורפת ללקוחות אלא מתן פרספרטיבה והצגה של ההתלבטויות שעלו במפגש כלומר "מהשטח".

שוק התשתיות ואתו שוק האחסון נמצא בתקופת ביניים – תקופת שתמשך זמן רב. מצד אחד העתיד של האחסון ברור. בעתיד, לא נרכוש חומרת אחסון ייעודית אלא נתקין "תוכנת אחסון" על השרתים שמממשים את הענן הפנימי שלנו. כלומר software defined storage. התוכנה תגלה איזה דיסקים יש בתוך השרתים HDD SSD, מה מידת המהירות שבה השרתים מדברים אחד עם השני (לצרכי רפליקציה של האחסון) ותטפל במידע בהתאם כולל כל התכונות שאנחנו מכירים – רפליקציה, DR, TIERS, SNAPS וכד'. ולכן בתחום יש "ירידה מסוימת" כי מבינים שעולם האחסון כפי שאנחנו מכירים אותו כעת – יעלם. עם זאת, נכון להיום, ה- software defined storage עדיין לא בשל כמו האחסון הרגיל זאת בעיקר בגלל שספקי האחסון אומרים (באופן ציורי): "אם תרצה שאתחייב לביצועים – אני חייב לעשות certification לחומרה" ולכן בסופו של דבר עדיין מוכרים חומרה + תוכנה ביחד ברובם המכריע של המקרים.

המלצת STKI בהקשר זה היא להתחיל ולבדוק את עולם ה- software defined storage כאשר ארגונים המחזיקים כמויות גדולות של קבצים יצליחו להפיק תועלת יחסית מהירה מיישום SDS ו-object storage.

טכנולוגיה אשר הבטיחה רבות וגם קיימה היא טכנולוגיית ה- FLASH כאשר כיום הסטנדרט למערכות קריטיות הדורשות את הביצועים הגבוהים ביותר הוא להשתמש בפתרונות FLASH

זאת במקביל לעליה בקיבולת והירידה במחיר ניתן לעיתים (באמצעות compression-deduplication) להוכיח ROI של FLASH לעומת מערכות HDD.

בדיון גם עלתה המלצה חשובה (ובסיסית) – יש לדאוג לרשת ניהולית נפרדת שתוכל לעבוד באופן נפרד לגמרי מהרשת הראשית ואשר תאפשר לגשת לציודים (אחסון, תקשורת ושרתים) גם במקרה של תקלה/נפילה ברכיב ותקשורת עצמה. כי ללא "דרך חלופית כזו" לא ניתן להגיע לציווד בכדי לראות מה הבעיה בזמן תקלה.

STKI ממליצה לארגונים גם לעקוב אחרי טכנולוגיה חדשה שמתפתחת באופן מהיר - RAM DISK. מדובר על זיכרון במהירות RAM אשר אינו נדיף. דוגמה לטכנולוגיה זו הנה https://he.wikipedia.org/wiki/3D_XPoint

טכנולוגיה מתפתחת נוספת הנה ממשק המחבר אחסון ישירות לתוך ה-BUS של המחשב. עוד פרטים ב- https://en.wikipedia.org/wiki/NVM_Express כאשר בתחום זה כבר יש מוצרים לדוגמה הסטרטאפ הישראלי <https://e8storage.com>.

בברכה,

פיני כהן

תוכן

- 4..... האמנם אחסון הוא סטנדרטי?
- 5..... מה מידת הבשלות הנדרשת מספק אחסון?
- 5..... התחייבויות של ספק לביצועים או פונקציונליות ותכנון עסקה
- 6..... ספק אחסון אחד או שניים.....
- 6..... active-active -I Storage virtualization
- 7..... Converged infrastructure (monolithic)
- 7..... hyperconverged -I Software defined
- 8..... הגדרת דרישות של מערכות אחסון ואורח חיים של מערכת

האמנם אחסון הוא סטנדרטי?

לאחרונה דובר רבות על כך שפתרונות האחסון הפכו להיות סטנדרטים עד כדי כך שישנם גורמים הדורשים להתייחס לכל הפתרונות באופן אחיד – ולקבוע ספק מועדף לפי מחיר בלבד. לדעת המשתתפים בדיון עדיין ישנם הבדלים חשובים בין הפתרונות – עם זאת צריך להיות יותר מקצועי בכדי להפנים הבדלים אלו.

להלן מספר דוגמאות לתכונות פונקציונליות אשר לא נמצאות או ממומשות באופן שונה במארזים השונים:

- יכולת של יותר מבקר לעבוד Active-Active בו זמנית מול אותו LUN (או שבו זמנית על LUN אחד עובד רק בקר אחד).
- באיזה מצבים ניתן להפעיל compression, thin provisioning, deduplication והאם online או רק לאחר שמירה.
- האם dedup לפי גודל בלוק קבוע, משתנה. האם ניתן לבטל dedop כשלא צריך?
- HA בין 2 בקרים בלבד או יותר מ-2 בקרים.
- Global dedup/compression או ברמת volume/LUN
- יכולת Active-Active בין אתרים
- Aggregate ברמת בקר או ברמת מספר בקרים (cluster).
- תחום שבו ישנה שונות יחסית רבה בין הפתרונות הנו אופן יישום active-active. עד כמה רגיש ל-latency. האם מדובר על "קופסא נפרדת" או שמדובר על פתרון שמוטמע בתוך המכונות, כיצד נמנע split-brain (האם מחייב מכונה נפרדת ואיזה תקשורת צריך, מה קורה אם יש נפילת תקשורת ל-witness) ועוד ועוד.
- הבדלים באופן שחרור מקום ב-SNAPS – האם משתחרר מקום אם משחררים חלק מ-SNAPS או רק כאשר משחררים SNAP מלא.

מה מידת הבשלות הנדרשת מספק אחסון?

בדיון עלתה השאלה "מה מידת הבשלות הנדרשת מספק אחסון" כאשר שוקלים אם להכניס אותו לארגון. גם אם הטכנולוגיה טובה והמחיר אטרקטיבי – הדעה השלטת בקרב המשתתפים היא שלפתרון חייב להיות "קילומטרז" מספיק עם התקנות מוכחות בישראל – גם בשביל לוודא שהתמיכה בישראל טובה ומקצועית מספיק.

לא רק "שאותן מכונות עובדות כבר בישראל" אלא שמדובר גם על שימוש ברמת פונקציונליות, קיבולת וביצועים שנדרשת.

עם זאת לא הייתה תמימות דעים לגבי "מה נחשב קילומטרז' מספיק". האם 5 התקנות, האם 10? ארגונים רבים יסתפקו ב"קילומטרז" נמוך יותר אך יפעילו את המכונות לבדיקה (או לסביבות נמוכות\מערכות שאינן קריטיות) חודשים רבים עד שנה ויותר.

התחייבויות של ספק לביצועים או פונקציונליות ותכנון עסקה

גם כאשר יש הסכם שמחייב את הספק להוסיף ציוד על חשבוננו או לשלם קנס (במידה ולא עומדים בביצועים או פונקציונליות לא כפי שהובטחה) לקוחות נוטים לא להפעיל אופציות אלו מכיוון שמדובר על פעולה ארגונית מורכבת – יש להפעיל את הגורמים המשפטיים והמנהלים הבכירים ולכן מדובר עדיף לבחור ספק\ציוד מתאים מלחתחילה ולא לבחור ספק\ציוד זול – עם חוזה שמגן על הלקוח.

יש לציין שלאחרונה ישנה תופעה של "תמחור חסר של עסקאות" ואנחנו רואים יותר ויותר מצבים שבהם ספקים לא עומדים בהתחייבויות שלהם ולכן עליהם לשפר את המערכת (הוספת דיסקים מהירים CACHE וכד').

גישה מעניינת לטיפול בנושא זה הנה לחתום על הסכם עם ספק שבמידה ואינו עומד בביצועים (או פונקציונליות) יהיה עליו לשלם קנס לארגון (ולא רק לשפר את המערכת). תשלום פיצוי כספי הנו הרבה יותר בעייתי מבחינת הספק מאשר תוספת ציוד ולכן גישה זו עשויה לגרום לספק להבטיח בצורה טובה יותר שהמערכת שמספק אכן תעמוד בדרישות.

סוגיה נוספת שעלתה היא תכנון וניהול עסקת אחסון. כאשר מבצעים עסקה נשאלת השאלה כמה "שלייקס" לקחת בהקשר של ביצועים ונפחים. מצד אחד עלויות האחסון קטנות כל הזמן (סדר גודל של 25% לשנה במונחי עלות ליחידת נפח) ולכן עדיף לרכוש תוספת אחסון כמה שיותר מאוחר. מצד שני תוספת נפחים כאשר המכונה כבר מותקנת ועובדת עלולה לגרום לבעיות של השבתה ולא תמיד ניתן לרכוש בדיוק את אותם הדיסקים שנרכשו בתחילה והדבר

עלול לגרום למגבלות בהגדרות LUNS או POOLS שלעיתים מחייבים שימוש בדיסקים זהים. ולכן, ישנם מנהלי אחסון שמעדיפים לרכוש את כל האחסון מראש ל-3 שנים.

ספק אחסון אחד או שניים

בדיון עלתה הדעה שעדיף להפעיל מדיניות "שני ספקים" אפילו אם הספק השני מקבל נישה קטנה. זאת לדברי מספר לקוחות שהשתמשו בזמנו עם ספק אחד, ולמרות שהספק פעל באופן מייטבי והלקוח הרגיש שהוא מקבל שירות ומחירים טובים, כאשר הכניסו את הספק הנוסף, אפילו אם בסביבת נישה, היה שיפור במחירים ובשירות.

active-active ו- Storage virtualization

לקוחות ציינו את היתרונות התפעוליים הרבים של שימוש ב-storage virtualization – אפשרות לעבוד מול מספר ספקים באופן שקוף, מיגרציה נוחה בין מכונות ועוד ועוד. אולם עדיין החסרונות של תוספת latency, הוספת גורם כשל נוסף, הפרויקט הראשוני (ברוב המקרים יש ליישם טופולוגיה חדשה של רשת האחסון) ובמידה מסוימת גם עלות הטכנולוגיה עדיין עוצרים לקוחות רבים מליישם טכנולוגיה זו.

בדיון עלה נושא של active active. בתצורה זו נמצאים שרתי ה-VMWARE (נדון בסיטואציה פופולארית זו) על אותו LAN (למרות שנמצא בין שני אתרים שונים) והם כותבים למה שנראה להם כאחסון אחד (שהוא "מחביא" בתוכו שתי מכונות אחסון שנמצאות בשני האתרים). טכנולוגיה זו מייטרת את השימוש ב-SRM. וגם שקופה ליישום במצב של נפילת אחסון באחד האתרים כמו גם נפילת אתר שלם. עם זאת יש לציין שחייבים לשים באתר נוסף שרת quorum אשר "מחליט" מי האתר הפעיל במידה ויש נתק בין שני האתרים (וזאת בכדי לא לקבל מצב של split brain – שבו כאשר הקשר בין האתרים מנותק ושני האתרים ממשיכים לפעול במקביל – ללא ידיעה שהאתר השני עובד).

למרות שההמלצה הכללית לעבודה ב- stretched cluster היא של 30-50 קילומטר המתורגמים ל-5 millisecond, לקוחות ציינו שמגיעים בטכנולוגית dwdm מתקדמת לפחות מ-1 millisecond latency ב-100 קילומטר!!

יש לציין שיישום active-active אינו סטנדרטי ומחייב טיפול במצבים אפליקטיביים מסוימים. לדוגמה אפליקציה שנמצאת על שרת VMWARE ומשתמשת בנתונים שנמצאים על מערכת

NAS (לדוגמה מסמכי ארכיון). כאשר האתר הראשי נופל והאפליקציה מתחילה לפעול באתר המשני (HA של VMWARE), המערכת עדיין מצפה לאתר את ה-NAS באתר הראשי ולא במשני.

Converged infrastructure (monolithic)

בדיון עלה הנושא של מערכות converged הבנויות מרכיבים "גדולים" (exadata, vblock) (וכד'). ללקוחות כבר יש ניסיון ביישום פתרונות בתחום זה בעיקר אורקל (exadata ו-ODA) וגם VBLOCK.

התחושה היא פתרונות אלו "מספקים את הסחורה". כלומר מספקים ביצועים ואמינות. וגם התמיכה והתפעול יותר נוחים מהסטנדט של בניית סביבה המורכת מספק אחסון, ספק תקשורת, ספק מחשוב וכד'.

אך מצד שני התחושה היא של "חתונה קתולית" ולכן ישנה מן הצד השני רתיעה מפתרונות אלו וגם לא בטוח שהפרמיה (כי פתרונות אלו יקרים יותר) משתלמת.

hyperconverged - Software defined

לקוחות נמצאים בשלבים ראשוניים בלבד של יישום software defined storage ובשלבים מתקדמים מעט יותר של יישום converged infrastructure.

לגבי use case – converged הוא ראשוני הוא לסניפים מרוחקים שבהם מותקנים היום מספר שרתים ואחסון קטן. במקרים אלו הפידבק שמערכות אלו עובדות באופן מיטבי.

אחד הלקוחות תיאר מצב שבו בצעו בחינה של מערכות hyperconverged בסביבה ארגונית תובענית של vmware. על פי תוצאות הבחינה לא הגיעו לאותם הביצועים של מערכות ה-high end storage "הרגילות". למרות שאותם דיסקים SSD נמצאים בשימוש הן במערכות hypervonverged והן במערכות high end (סדר גודל של 70 מיקרו-שניות) אבל במערכות hyperconverged העברת הכתיבות לשרת השני מחייבת יציאה לשכבת התקשורת וכך מגיעים לכ-3 מילי-שניות, לעומת מערכות highend storage ששם מגיעים לסדר גודל של 1 מילי-שניות. כלומר לפחות בשלב הנוכחי, hyperconverged אינם מגיעים לאותם ביצועים של highend storage.

אולם גם בתחום זה צפויים להיות שינויים. זאת מכיוון שאת התחום של hyperconverged אפשר לחלק לשניים.

Hyperconverged שבהם הספק נותן שרתים והלקוח מחבר אותם בתקשורת שלו. זה המקרה מהדובר והמוכר יותר כיום.
Hyperconverged שבהם הספק נותן את השרתים ואת התקשורת- כלומר נותן ארון שלם. במקרה הזה הספק יכול להתחייב לביצועים טובים יותר עד לאפשרות שהמערכת תגיעה ל-sub millisecond (אם משתמש בתשתית G10 או אפילו G40 ב-Ethernet).

לגבי software defined storage לקוחות ציינו את Microsoft storage spaces בתור מוצר מעניין וחינמי.

הגדרת דרישות של מערכות אחסון ואורח חיים של מערכת

לקוחות התייחסות להגדרת דרישות בסיטואציה של שדרוג מערכת אחסון. מצד אחד אפשר לדרוש את אותם הדרישות הנוכחיות (במידה שמרוצים מהביצועים הנוכחיים) אבל בנפחים גבוהים יותר. אולם לקוחות ציינו שהם מעדיפים לא להתפשר על איכות וביצועים. לקוח תיאר מצב שבו הוא מסתכל 3 שנים קדימה ותמיד דורש פי 3 ממה שיש היום במכונה החדשה מיד (כלומר הלקוח מעדיף לא להוסיף נפחים במהלך חיי העסקה אלא רוכש את כל הנפחים העתידיים מיד). זאת מן הטעם שאינו רוצה לבצע מהלכים של הוספת דיסקים שמחייבים reorg במערכות והיו כבר מצבים שבהם רצו לקנות דיסקים ספציפיים וכבר לא היו במלאי\בייצור. מערכת האחסון נרכשת עם אחריות ל- 8 שנים כי לאחר השימוש במערכת בייצור המערכת עוברת ל- DR וכך מגיעים ל- 8 שנים.

הערת STKI- מבחינה תחרותית עדיף שאורח החיים של מערכת הייצור וה- DR יסתיימו באותו זמן כן שניתן יהיה לבצע תחרות בין ספקים באופן נוח.
ספציפית לגבי תחום האחסון, בפתרונות אורקל, נראה אנשי האחסון פחות נדרשים כי מערכות exadata ו- ODA מגיעות עם יכולות אחסון כחלק מהאפליקציה.