

Everything about HPe MSA 2040/2042/1020 Storage Device. (Part I)

پس از عرضه دستگاه ذخیره سازی MSA 1040/2040 Series در سال 2013، شرکت HPe با توسعه ی سیستم عامل این دستگاه توانست قایت های جدیدی را به آن اضافه کند. در این نوشتار سعی شده تا با بررسی ویژگی های فنی این دستگاه، خوانندگان این مقاله با کلیه قابلیت های جدید دستگاه آشنا شوند.

بطور کلی خانواده HPe MSA Series، جزء رده ی اولیه (Entry level) از تجهیزات ذخیره سازی با عملکرد بهینه می باشند. این تجهیزات بر خلاف باور های عمومی که به درست و یا غلط در حوزه ی فناوری اطلاعات رواج یافته، در صورتی که به درستی پیکربندی شوند، برای بسیاری از مقاصد کاری مانند مجازی سازی، VDI، پایگاه داده و پشتیبان گیری می توانند بکار روند. یکی از نکات برجسته در خصوص این خانواده از محصولات ذخیره سازی شرکت HPe، هزینه تامین و عملیاتی سازی یا CAPX/OPEX بسیار مناسب آن در مقایسه با محصولات ارائه شده توسط دیگر رقبا مانند EMC, IBM, NetAPP می باشد. لازم به یاد آوری می باشد که در طرح زیر ساخت ذخیره سازی، ایجاد تناسب میان نیاز ها و هزینه حرف اول را می زند. چه بسا بکارگیری دستگاه گران تر به مفهوم بهره وری و کارایی بیشتر نمی باشد. در بخش اول این نوشتار به توضیح ویژگی های ظاهری، فنی و دستگاه ذخیره سازی MSA 2040/1040 می پردازیم و در بخش دوم به اصول و مبانی اتصال دستگاه به شبکه ارتباطی SAN و پیکربندی استاندارد اشاره می شود. با عرضه ی خانواده MSA1040/2040/2042، دو رویکرد کلی در ساختار کاری (Mode) این دستگاه بوجود آمده است. این دو رویکرد یا فناوری به نام **Linear** و **Virtual** معروف می باشند. در شرایط کاری **Linear** همان اصول کارکردی دستگاه های قدیمی MSA P2000 GX اجرا می شود. داده های ذخیره شده بصورت **پیوسته و از پیش مشخص** روی بخش های کوچکی بنام **Page** بر روی دیسک های عضو **Disk Group** نوشته می شوند. در شرایط کاری **Virtual** الزامی به رعایت این پیوستگی در **page** ها وجود ندارد بنابراین داده های خیره شده یا بهتر است بگوییم **Page** ها می تواند در دیسک های عضو چندین **Disk Group** که در **Virtual Pool** قرار دارند بصورت **نا پیوسته** نوشته شوند. هر دو روش دارای مزایا و ایراداتی می باشند که با توجه به شرایط و نیازهای محیط عملیاتی می توان مد کاری را برگزید. برخی از مزایای روش **Virtual** را می توان به شرح زیر دانست

- افزایش کارایی با افزایش تعداد دیسک
- از بین بردن محدودیت ایجاد Volume روی 16 دیسک
- امکان استفاده از منابع دیسک بصورت مشترک میان Volume ها بصورت موثر
- امکان اجرای راه کار Storage Tiering، Read Cache، Quick Rebuild، Thin provision و ...

مشخصات فنی دستگاه MSA 1040/2040:

در جدول زیر مشخصات فنی دستگاه MSA 1040/2040 ارائه شده است.

Item	MSA 1040	MSA 2040	MSA 2042
CACHE/ Controller	6 G.B	6 G.B	6 G.B
# of SFF/LFF Support	48/99	96/199	96/199
6/12 Gbps Backend Connection	yes	yes	Yes
Host Port/ Controller- Type	2 * 8 Gb	4* 16 Gb	4* 16 Gb
iSCSI Connection	1/10 Gb	1/10 Gb	1/10 Gb
# of Enclosure	3	7	7
Thin Provision	Yes	yes	Yes
Sub-LUN Tiering	Yes	yes	Yes
Read Cache	Yes	yes	Yes
Performance Tier	Yes	yes	Yes
Wide Striping	yes	yes	Yes
iSCSI/FC in Single Controller	No	yes	Yes
Full Drive Encryption	No	yes	Yes
Build-in Hybrid Cache	No	No	Yes/ 800 G.B
real-time Tiering	No	No	Yes
Remote Snapshots	Yes/ LV	Yes /LV	Yes / LV
Remote Replication	No	No	Yes
SED Drive support	No	Yes	Yes

نمای ظاهری دستگاه ذخیره ساز HPe MSA 1040/2040

در شکل های زیر نمای از پشت تجهیزات به تفکیک ارائه شده است .



شکل 2: MSA 1040



شکل 1: MSA 2040

نمای ظاهری دستگاه از جلو بصورت زیر می باشد.



شکل 4: MSA 1040/2040 12 LFF



شکل 3: MSA 1040/2040 24 SFF

با توجه شکل ظاهری دستگاه می توان به این نتیجه رسید که خانواده ی MSA 2040 دارای کارایی بیشتر می باشند زیرا دارای پورت های Front-End پر سرعت و بیشتری می باشند.

اصطلاحات و مفاهیم فنی

همانطور که پیشتر بیان شد، ابتدا به شرح مختصر و مفید مفاهیم و اصطلاحات رایج در تجهیزات ذخیره سازی MSA 1040/2040 می پردازیم.

Disk Group: به تعدادی دیسک البته بر اساس قواعد مشخصی از یک خانواده با Raid Type یکسان گفته می شود. Raid type های مجاز در دستگاه MSA 1040/2040 شامل موارد روبرو می شوند (0, 1, 10, 5 & 6).
Disk Group یا DG به سه نوع تقسیم می شوند که هر یک دارای مزایا و کارکرد ها خاص خود می باشد که در ادامه به شرح ان می پردازیم.

Linear DG: تعدادی دیسک هم خانواده با Raid Type یکسان که امکان افزودن دیسک در آینده را ندارد. البته امکان آمیختن دیسک های نا خانواده در DG وجود دارد اما به هیچ وجه پیشنهاد نمی شود. Linear DG به LDG و یا Vdisk نیز شناخته می شوند. دستگاه MSA 1020/2040 از LDG 64 پشتیبانی کی کند. Raid type های قابل اعمال به DG شامل موارد زیر می باشد. **RAID 1, RAID 5, RAID 6, RAID 10, and RAID 50**

Virtual DG: مجموعه ای از چندین دیسک با Raid Type یکسان، که پس از ایجاد می بایست به عضویت Virtual Pool در آیند. VDG می تواند شامل SSD باشد. VDG نام مختصر Virtual DG می باشد. به منظور ایجاد VDG مشکل از دیسک های SSD به عنوان لایه ی Performance نیاز به وجود لایسنس Performance Tier می باشد.

Read-cache DG: این نوع DG به نوعی Virtual Disk Group بوده که با هدف ایجاد فضای Cache برای فقط Virtual Pool بکاربرده می شود. فضای ایجاد شده توسط این نوع DG به فضای کلی Virtual Pool افزوده نمی شود،

امکان افزودن این نوع DG به Virtual Pool در هر زمان ممکن می باشد. به منظور ایجاد Read-cache DG می بایست از دیسک های SSD با ویژگی $1 < DWPD$ استفاده کرد. تنها یک Read-Cache DG در یک Virtual Pool می توان ایجاد کرد. برای فعال سازی Read-Cache DG نیازی به لایسنس نمی باشد. از Read-Cache DG برای افزایش کارایی Virtual Pool استفاده می شود. افزایش فضای Read-Cache DG با حذف Read-Cache DG پیشین و ایجاد Read-Cache DG جدید ممکن می باشد.

Pool: فضایی شامل یک یا چند DG. Pool از نظر کارکرد به دو نوع تقسیم بندی می شود.

Linear Pool: در این خانواده از Pool تنها می توان یک LDG وجود داشته باشد. امکان افزودن LDG پس از ایجاد Linear Pool وجود ندارد. بلکه تنها می توان تعداد دیسک های LDG را افزایش داد.

Virtual Pool: در این خانواده از Pool می توان، چندین VDG را در بر گیرد. نام دیگر آن Storage Pool می باشد. دستگاه MSA 1040/2040 تنها از دو Virtual Pool پشتیبانی می کند. بیشینه تعداد VDG در یک Pool، شانزده VDG می باشد. امکان افزایش کارایی در Pool با استفاده از لایسنس tier ممکن می باشد. در واقع در دستگاه MSA 104/2040 فناوری Storage Tiering وجود دارد. امکان حذف DG از Virtual Pool در صورتی که فضای خالی معادل در کل Virtual Pool وجود داشته باشد، ممکن می باشد. در صورت حذف یک یا چند DG از Pool با شرایط فوق، هیچ داد ای از دست نمی رود بلکه تنها داده ها مورد نظر در VDG های عضو در Virtual Pool جابجا می شود. هر virtual Pool می تواند از بخش یا در اصطلاح Tier زیر تشکیل شود.

- Performance Tier (SSD)
- Standard Tier (10K, 15K SAS)
- Archive Tier (7.2K Midline SAS)

Chunk: کوچکترین بخش از داده که بصورت پیوسته رو یک دیسک در Disk Group نوشته می شود. اندازه ی Chunk می بایست بر اساس Application block Size هماهنگ باشد.

Volume: بخشی یا تمام فضای Pool که به عنوان دیسک به initiator تخصیص می یابد. داده های Volume بر روی دیسک های Linear/Virtual Pool توزیع می شود. در واقع هر دیسک در LDG یا VDG به بخش های ثابتی نام page با سایز 4 M.B تقسیم می شود. امکان ایجاد volume تا 300 T.B بصورت پیش فرض وجود دارد. بیشینه فضای Volume حدود 512 ترابایت می باشد. نام دیگر Volume در حوزه ذخیره سازی، LUN می باشد. Volume یا LUN قابل ارائه به میزبان می باشد. در دستگاه MSA 1020/2040 هر Volume دارای مالک یا Owner می باشد. مالک Volume در هر لحظه می تواند یکی از Controller های می باشد و دیگر Controller به عنوان جانشین در صورت

خرابی مالک قابل ارائه سرویس می باشد. کمترین اندازه ی Volume معادل اندازه یک Page یعنی 4 M.B می باشد. سیستم عامل Volume یا LUN را بصورت دیسک شناسایی می کند.

Virtual Volume: روشی برای مدیریت فضای اختصاص داده شده برای حجم Volume در دیسک های عضو DG به گونه ای که فضای Virtual Volume بر روی هر بخشی از دیسک های عضو VDG در Virtual Pool می تواند ذخیره می شود و هیچ رابطه ی خطی وجود ندارد.

Linear volume: روشی برای مدیریت فضای اختصاص داده شده برای حجم Volume در دیسک های عضو DG به گونه ای از پیش تعیین شده. بنابراین محل داده های ذخیره شده بر روی LDG عضو Linear Pool از پیش تعیین می شود.

Volume Mapping: به منظور ایجاد دسترسی میزبان به Volume ایجاد شده، مفهومی بنام mapping وجود دارد. در واقع Volume Mapping، مانند Storage Access List عمل می کند. با پیکربندی Mapping می توان گفت که چه LUN یا volume توسط چه پورت های کنترلر قابل دسترس برای میزبان می باشد. لازم به یاد آوری می باشد که دستگاه ذخیره ساز MSA 1040/2040 از مد کاری ALUA پشتیبانی می کند. در این مد کاری هر دو کنترلر دستگاه به عنوان مالک خود را معرفی کرده و دستگاه ذخیره ساز به گونه ای بصورت Active/ Active به ایفای نقش می پردازد. در صورت از کار افتادن کنترلر مالک، دیگر کنترلر به ارائه سرویس می پردازد و هیچ وقفه ای در کار بوجود نخواهد آمد.

Write-back VS Write-through Caching: دستگاه MSA 1040/2040 هنگام نوشتن داده از مکانیزم ذخیره ی اطلاعات در Buffer استفاده می کند. هدف از این کار افزایش کارایی و ارسال تایید به میزبان از نوشته شدن اطلاعات می باشد. در حین نوشته شده اطلاعات در Buffer دستگاه، پروسه ای داده های Buffer را به دیسک سخت انتقال می دهد. در صورت غیر فعال کردن فرآیند write-Back، فرآیند write-through یا ارسال مستقیم داده های میزبان به دیسک اجرا می شود. این روش در برابر قطع و خرابی برق دستگاه امن تر می باشد ولی به دلیل نبود نسخه ی آینه ای از داده های در حال نوشته شدن در حافظه ی دیگر Controller، خرابی کنترلر به هیچ وجه قابل تحمل نمی باشد. این گزینه در سطح volume قابل تعریف می باشد.

Cache Optimization: دستگاه MSA 1040/2040 دارای دو مد کاری برای کارکرد حافظه ی Cache خود می باشد.

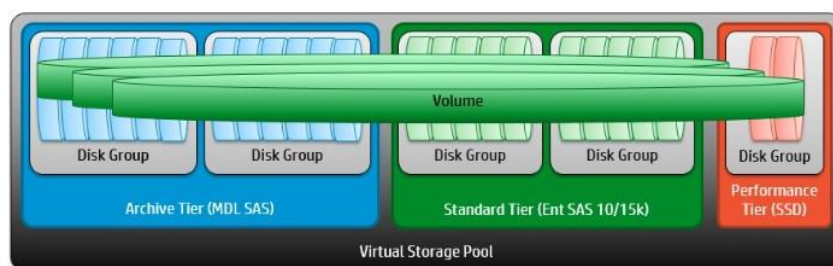
- 1- Standard: در این حالت هر Controller نسخه ای آینه ای از داده های یکدیگر را در خود نگهداری میکند تا هنگام اختلال در ارتباط با میزبان، دیگر Controller بتواند کارهای جاری را ادامه دهد.
- 2- Non-Mirror: در این حالت حافظه Cache دو دستگاه Controller، بصورت جدا از هم کار می کند و هیچ حالت آینه ای وجود نخواهد داشت. در صورت از کار افتادن کنترلر، امکان ادامه کار توسط دیگر کنترلر نمی باشد ولی کارایی سیستم بهبود پیدا می کند.

Optimizing read-ahead caching: به منظور کاهش زمان دسترسی به داده های دیسک، نسخه ای از داده های خوانده شده در حافظه Cache کنترلر ذخیره می شود تا در مراجعه بعدی از آن استفاده شود. دستگاه MSA 1040/2040 برای تعیین فضای Read-Cache به سه روش پیکربندی می شود.

- 1- تناسبی: در این روش کاری، controller دستگاه با محاسبه ی پویا اندازه ی بار کاری، نسبت به تعیین اندازه حافظه ی Cache اقدام می کند
- 2- ردیفی: در این روش کنترلر اندازه حافظه Cache را برابر با اندازه ی Stripe تنظیم می کند.
- 3- دستی: در این روش راهبر سیستم مقدار حافظه ی Cache را تعیین می کند.

Read Cache: امکان افزایش حافظه ی Cache کنترلر دستگاه MSA 1040/2040 با استفاده از دیسک های SSD

Sub-LUN Tiering: Volume های که در Virtual Pool های متشکل از دیسک SSD و دیسک سخت تعریف می شوند، می توان داده های پر دسترس را با استفاده از فناوری Sub-LUN Tiering به بخش های پر سرعت مانند SSD ارسال و داده های کم دسترس را به بخش های کم سرعت مانند دیسک های 7.2 K منتقل کرد. این فناوری باعث بهبود کارکرد دستگاه می شود.

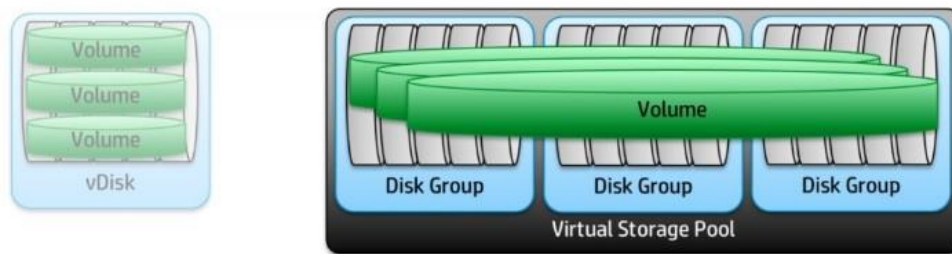


شکل 5 Storage Tiering

T10 Unmap for Thin Reclaim

دستگاه ذخیره ساز MSA 1040/2040 دارای قابلیت آزاد سازی فضای ذخیره سازی بدون استفاده، بطور خودکار میباشد. هدف از این کار آزاد سازی فضای اشغال شده با ساختار Thin بعد از حذف یا جابجایی ماشین مجازی در دستگاه ذخیره ساز می باشد. ویژگی Thin provision برای virtual Storage فعال می باشد. این قابلیت در سطح pool قابل تعریف می باشد.

Wide striping: با پیاده سازی Virtual Pool در تجهیزات MSA 1040/2040، پخش و توزیع کردن page های مربوط به virtual Volume بر روی کلیه ی Virtual DG های عضو Virtual pool ممکن می باشد. البته این ویژگی در مد کاری Virtual storage قابل اجرا می باشد.



شکل 6: wide Striping

Snapshot امکان تهیه وضعیت جاری اطلاعات Volume در لحظه و استفاده از آن برای بازیابی داده ها در صورت بروز اشکال در Volume. دستگاه MSA 1040/2040 امکان تهیه 512 نسخه ی snapshot را دارد. بصورت پیش فرض دستگاه از 64 وضعیت Snapshot پشتیبانی می کند برای افزودن به این عدد می بایست لایسنس و مجوز مربوطه خریداری شود. Snapshot به دلیل وابستگی به مبدا بر خلاف Volume copy نمی توانند بصورت مستقل بکار روند.

Remote snap: ویژگی و گواهینامه ای برای ارسال (Replication) داد های Volume بصورت نا هم زمان از یک دستگاه MSA به دستگاه MSA دیگر. برای مواقعی که ممکن دستگاه MSA مبدا دچار اختلال در کار کرد شود.

Drive spin down: امکان غیر فعال کردن دیسک های بیکار به منظور کاهش توان و برق مصرفی

Volume copy: امکان تهیه نسخه ی مشابه از Volume در دستگاه دیگر. نسخه ی تهیه شده هیچ وابستگی به نسخه مبدا ندارد. در واقع کپی کامل از مبدا در مقصد توسط دو دستگاه MSA 1040/2040 انجام می شود.

Scrub Procedure: فرایند درونی دستگاه MSA 104/2040 به منظور شناسایی خطا در سطح دیسک و رفع آن بصورت خودکار. این فرایند در سطح دیسک اجرا می شود و کلیه Disk errors و Raid errors ها را برطرف می نماید. فعال سازی این پروسه پیشنهاد می شود.

Full Disk Encryption (FDE): به منظور تامین امنیت و محرمانگی اطلاعات ذخیره شده در دستگاه MSA 2040(only)، این ویژگی امکان رمزنگاری داده های ذخیره بر روی دیسک های SED را فراهم می کند.

Repurposing: فرایند آزاد سازی سیستم به منظور پاکسازی دیسک های دستگاه در حالتی که ویژگی FDE فعال شده است.

SSD wear gauge: به دلیل محدود بودن تعداد دفعات نوشتن و خواندن بر روی دیسک های SSD، دیسک های SSD عرضه شده برای دستگاه MSA 1040/2040 دارای شاخص wear gauge می باشند. از این پارامتر برای پیش بینی وضعیت SSD استفاده می شود. هنگام رسیدن دیسک SSD به 5% این شاخص، می بایست نسبت به تعویض دیسک SSD اقدام کرد.

لایسنس های دستگاه MSA 1040/2040

کلیه ویژگی های تجهیزات ذخیره سازی MSA 1040/2040 بطور پیش فرض فعال نمی باشد و می بایست با خرید و فعال سازی لایسنس مربوطه عملیاتی گردد. به منظور آشنایی بیشتر به شرح لایسنس های دستگاه می پردازیم.

Advanced Data Service software suite

مجموعه ای از سه لایسنس شامل

1. HPE MSA 2040 Performance Automated Tiering LTU
2. HPE MSA 2040 512 Snapshot Software LTU
3. HPE MSA Remote Snap Software LTU

با استفاده از لایسنس های فوق می توان

HPE MSA 1040 Advanced Virtualized Features

با استفاده از این لایسنس می توان دستگاه را در مد کاری Virtual Storage علاوه بر linear پیکربندی نمود.

MSA Performance Tiering

با استفاده از این لایسنس می توان ساختار Sub-LUN Tiering را در سه سطح ایجاد نمود.

1. A Performance tier with SSDs
2. A Standard SAS tier with Enterprise SAS HDDs
3. An Archive tier utilizing Midline SAS HDDs.

پس از ایجاد لایه های فوق با توجه به میزان دسترسی، می توان داده های پر کاربرد را در لایه ی مناسب قرار داد این عمل بطور خودکار انجام می شود. به منظور ایجاد لایه ی Performance با استفاده از SSD می بایست لایسنس Advanced Virtualized و لایسنس Performance Tiering خریداری شود.

MSA 1040 Snapshot and volume copy software

بصورت پیش فرض امکان تهیه snapshot و volume copy تا 64 وضعیت وجود دارد و نیازی به لایسنس نیست. با استفاده از این لایسنس می توان تعداد Snapshot ها و یا Volume copy را به 512 وضعیت افزایش داد.

HPE MSA Remote Snap Software

این لایسنس برای هر دو مد کاری Linear و Virtual وجود دارد. با استفاده از این ویژگی می توان در سطح دستگاه ذخیره ساز (block level) نسبت به ارسال و کپی اطلاعات به دستگاه دیگر اقدام کرد. فرایند رونوشت برداری بصورت نا همزمان (asynchronous) بوده، و از فناوری Snapshot استفاده می شود. روش کار به این گونه می باشد که داده های volume از دستگاه مبدا به volume ای در دستگاه مقصد کپی می شود. امکان انتقال اطلاعات از بستر پورت های FC و iSCSI مکن می باشد. امکان ایجاد رونوشت برداری بصورت 4 ورودی به 1 ممکن می باشد. یعنی از چهار دستگاه به یک دستگاه می توان رونوشت ارسال کرد. مثلا از چهار سایت اقماری می توان داده های مهم را به سایت مرکزی ارسال کرد. با استفاده از یک کنترلر می توان فرایند رونوشت برداری را نیز انجام داد. امکان رونوشت برداری بین دستگاه MSA 1040/2040 با MSA P2000 G3 در مد linear نیز ممکن می باشد. هدف اصلی از تهیه این لایسنس راه اندازی راه کار پدافند غیر عامل و سایت های پشتیبان می باشد.

ملاحظات کلی مربوط به کارایی:

در این بخش از مقاله به روش های کلی در خصوص بهبود کارکرد دستگاه MSA 1040/2040 ارائه می شود. عدم رعایت موارد زیر باعث اختلال در کارکرد کلی دستگاه نمی شود ولی قویا پیشنهاد می شود به روش های زیر در کلیه مراحل پیکربندی دستگاه در نظر گرفته شود.

- تا حد امکان از Virtual Pool بجای Linear Pool استفاده شود. مفهوم این عبارت پیکربندی دستگاه در حالت Virtual Storage می باشد.
- Disk Group ها جدایی از ماهیت linear یا Virtual می بایست دارای پیکربندی Raid یکسان باشند تا بهترین کارایی بدست آید. در واقع Disk Group های هر tier بهتر است دارای پیکربندی یکسان Raid باشند.
- تعداد دیسک در Raid 5/6 می بایست ضریبی از دو باشد. (جدول زیر)

Raid Type	Total Drive/ DG	Data Drive	Parity Drive
Raid 5	3	2	1
Raid 5	5	4	1
Raid 5	9	8	1
Raid 6	4	2	2
Raid 6	6	4	2

Raid 6	10	8	2
--------	----	---	---

- دیسک های هر Tier می بایست از همه نظر یکسان باشند.
- بخش Performance tier که از دیسک های SSD تشکیل شده است می بایست حدود 10 تا 20 درصد بخش Standard Tier در نظر گرفته شود. تا بیشترین کارایی حاصل شود. در صورت نیاز به Performance بیشتر، می توان تا 40 درصد، فضای این لایه را افزایش داد.
- از هر نوع دیسک حداقل یک نوع به عنوان یدک (spare) استفاده شود.
- در صورت نیاز به افزایش ظرفیت Pool، معادل DG نسبت به اضافه کردن دیسک اقدام شود. اگر DG در Pool شامل 10 دیسک 300 G.B 15 K در حال پر شدن می باشد. دقیقاً به همان تعداد و نوع دیسک به Pool اضافه شود.
- همیشه از آخرین نگارش firmware استفاده شود.