**Solid State Drives (SSDs)**

**Soyut(Teorik)** Katı hal sürücüler (SSD'ler), kurum depolama performanslarını bir sonraki seviyeye taşımak için uygun faktör olarak kabul edilir. Gerçekten de, Sabit Disk Sürücülerinin (HDD) rotating-storage (dönen depolama) teknolojisi, yanıt süresinin kritik faktör olduğu uygulamaların gerektirdiği erişim zamanını başaramaz. Aksine, SSD'ler NAND Flash bellekleri gibi katı hal belleklerine dayanmaktadır; bu durumda, herhangi bir mekanik parça ve depolanmış verilere rastgele erişim çok daha hızlı olabilir ve bu şekilde yukarıda belirtilen gereksinimleri karşılar. Ancak birçok uygulamada, ana işlemci ve sürücüler arasındaki arabirim, performans darboğazı olarak kalmaktadır. Bu nedenle SSD’nin arayüzü, SSD’nin ana işlemci ile doğrudan bağlantısını sağlayan SAS ve SATA gibi eski depolama arabirimlerinden PCIe’ya dönüşmüştür.

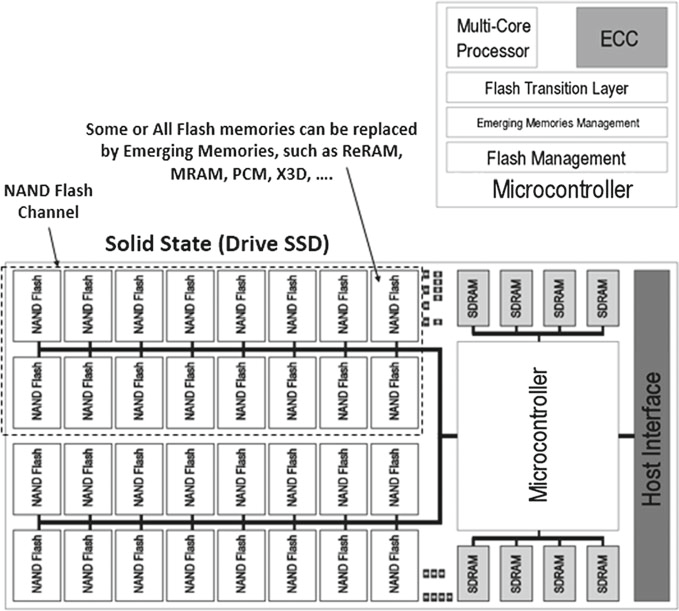
**1.1 Tanıtım**

Katı Hal Sürücüler (SSD'ler), kurumsal depolama performansını büyük ölçüde geliştirmeyi vaat eder. Elektromekanik Sabit Disk Sürücüler (HDD'ler) kapasiteyi sürekli olarak arttırırken, döner depolama teknolojisi, çevrimiçi işlem, veri madenciliği ve bulut bilişim dahil olmak üzere zorlu kurumsal uygulamalarda gerekli olan erişim zamanını veya aktarım performansını sağlamaz. İstemci uygulamalarının, daha hızlı yanıt süreleri sunabilen, daha az güç harcayan ve daha küçük mobil form faktörlerine sığabilen elektromekanik disk sürücülerine bir alternatif ihtiyacı vardır.

Flash bellek tabanlı SSD'ler, verilere daha hızlı rastgele erişim ve daha hızlı aktarım oranları sunabilir. Dahası, SSD'nin veri depolama kapasitesi artık HDD’lerin yerini alabilecek noktada. Ancak birçok uygulamada, ana bilgisayar ve sürücüler arasındaki arabirim, performans darboğazı olarak kalmaktadır. SAS ve SATA gibi eski depolama arabirimlerine sahip SSD'ler yararlı olduğunu kanıtlıyor. PCI-Express (PCIe) SSD'ler, ana işlemci ile doğrudan bağlantılı olarak performansı daha da artıracak ve yanıt vermeyi artıracaktır.

**1.2 SSD’s Mimarisi**

Bir SSD'nin temel blok diyagramı Şekil 1.1'de gösterilmiştir. Belleklere ve bir Flash denetleyicisine ek olarak, genellikle başka bileşenler de vardır. Örneğin, dahili güç kaynağını sürmek için harici bir DC-DC dönüştürücü eklenebilir veya daha iyi bir saat hassasiyeti için bir kuvars kullanılabilir. Tabi ki, güç kaynağının stabilize edilmesi için makul filtre kapasitörler yerleştirilir. Güç yönetimi sebepleri için bir dizi sıcaklık sensörüne sahip olmak da çok yaygındır.

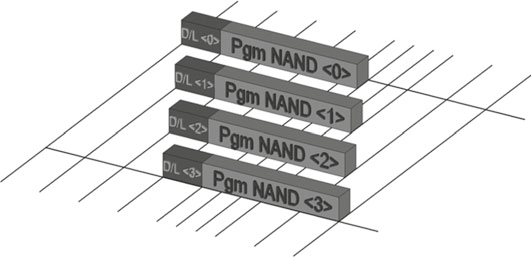


**Fig. 1.1** SSD'nin blok diyagramı

Veri önbellekleme için hızlı bir DDR bellek sıklıkla kullanılır: yazma erişimi sırasında, önbellek, Flash'a aktarılmadan önce verileri depolamak için kullanılır.

Faydası veri güncelleme olup, örn. yönlendirme tabloları daha hızlıdır ve Flaş'ı yıpratmaz. Tipik bir hafıza sistemi, birkaç NAND belleğinden oluşur. Tipik olarak, genellikle “kanal” olarak adlandırılan 8 bitlik bir veri yolu, farklı bellekleri denetleyiciye bağlamak için kullanılır (Şekil 1.1). Bir sistemdeki birden fazla Flash belleğinin hem depolama yoğunluğunu artırmaya hem de okuma / yazma performanslarına yönelik bir araç olduğunun altını çizmek önemlidir.

Bir kanaldaki işlemler aralıklı olabilir, yani birincisi hala meşgulken ikinci bir çipin ele alınabileceği anlamına gelir. Örneğin, çok sayıda yazma işleminin bir dizisi, Şekil 1.2'de gösterildiği gibi farklı NAND'lere hitap eden bir kanala yönlendirilebilir: Bu şekilde, veri yükleme aşaması boru hattı ile kanal kullanımı maksimuma çıkarılır.



**Fig. 1.2** Interleaved operations on one Flash channel

Aslında, program işlemi bir bellek yongası içinde gerçekleşirken, ilgili Flash kanalı serbesttir. Toplam Flash kanalı sayısı, hedef uygulamanın bir fonksiyonudur, ancak onlarca kanal oldukça yaygın hale gelmektedir. Aynı Flaş programlama süresi göz önüne alındığında, dönüşümlü-çalıştırma(interleaving) sayesinde SSD’nin verimi büyük ölçüde artıyor.

Bellek denetleyicileri, bellek kanallarına erişimi zamanlamaktan sorumludur. Denetleyici, Flash ile düşük seviyeli iletişim protokolü için özel motorlar kullanır.

Ayrıca veri yükleme aşamasının, program işlemine kıyasla ihmal edilemediği açıktır (aynı yorum veri çıkışı için geçerlidir): bu nedenle, I / O arabirimi hızını artırmak, performansları iyileştirmenin başka bir akıllı yolu: DDR benzeri arayüzler. Hız arttıkça, kanalı doyurmadan önce daha fazla NAND paralel olarak çalıştırılabilir. Örneğin, 30 MB / sn'lik bir hedef varsayıldığında, yaklaşık 50 MHz'lik bir minimum DDR frekansıyla 2 NAND'a ihtiyaç vardır. 200 µs'lik bir sayfa program süresi verildiğinde, 50 MHz'de dört NAND aralıklı modda çalışabilir ve yazma işinin iki katına çıkar. Tabii ki, güç tüketimi özenle dikkate alınması gereken bir başka ölçüttür.

**1.3 Flash Controller**

Bir bellek denetleyicisinin iki temel görevi vardır:

1. hem ana bilgisayara hem de Flaş belleğine en uygun arabirimi ve protokolü sağlamak;

2. Verilerin verimli bir şekilde ele alınması, aktarım hızının en üst düzeye çıkarılması, veri bütünlüğü ve saklanan bilgilerin saklanması.

Bu tür görevlerin yerine getirilmesi için, zaman-kritik görevleri ele almak üzere özel bir donanımla birlikte standart bir işlemci (genellikle 8–16 bit) yerleştirerek uygulamaya özel bir aygıt tasarlanmıştır.

Genel olarak, bellek denetleyicisi donanımda veya bellenimde uygulanan dört bölüme ayrılabilir (Şekil 1.3).

Ana bilgisayardan Flaş'a geçerken, ilk kısım, gerekli endüstri standardı protokolünü (PCIe, SAS, SATA, vb.) Uygulayan ana bilgisayar arabirimidir. Böylece SSD'ler ve ana bilgisayarlar arasında hem mantıksal hem de elektrik birlikte çalışabilirlik sağlar. Bu blok, ana makine tarafından çağrılan komut dizisini çözen ve Flash belleklerine veri akışını işleyen bir donanım (arabellek, sürücü, vb.) Ve ürün yazılımı karışımıdır.

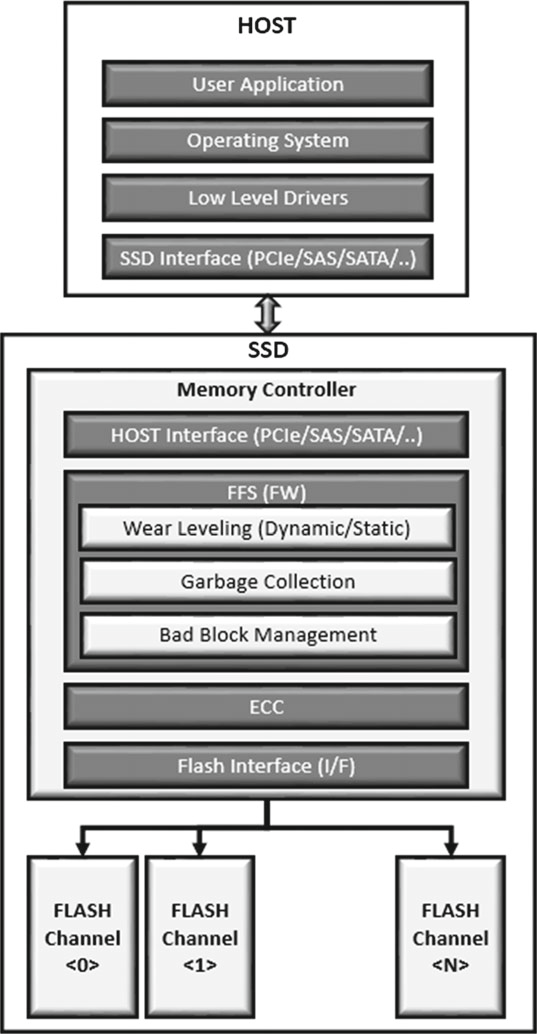
İkinci kısım Flaş Dosya Sistemi (Flash File System) : yani SSD'lerin manyetik diskler gibi kullanılmasını sağlayan dosya sistemi.

Örneğin, bir dosya oluşturan çok sayıda alt sektördeki sıralı bellek erişimi, dosya tarafından Dosya Ayırma Tablosunu (File Allocation Table-FAT) oluşturmak için kullanılan bağlı listeler (SSD'nin kendisinde saklanır) tarafından düzenlenir.

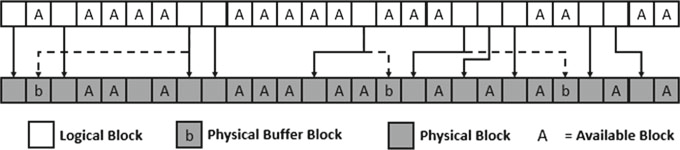
FFS genellikle denetleyicinin içindeki yerleşik yazılım biçiminde gerçekleştirilir; her alt katman belirli bir işlevi yerine getirir.Ana fonksiyonlar şunlardır: Aşınma dengeleme yönetimi, Çöp Toplama ve Kötü Blok Yönetimi.

Tüm bu işlevler için tabloları, Şekil 1.4'te gösterildiği gibi, sektörleri ve sayfaları mantıksal etki alanından fiziksel etki alanına (Flash Çeviri Katmanı veya FTL) eşlemek için yaygın olarak kullanılır. En üstteki satır hafızanın mantıksal görünümüdür, alt satır ise fiziksel olanıdır. Ev sahibi perspektiften, veriler belirli bir mantıksal sektörde şeffaf bir şekilde yazılır ve üzerine yazılır: Flaş sınırlamaları nedeniyle, aynı sayfada yer almak mümkün değildir; bu nedenle, fiziksel bloğa yeni bir sayfa (sektör) atanmalı ve bir önceki geçersiz olarak işaretlenmelidir. Bir noktada, mevcut fiziksel bloğun dolduğunu ve ikincisinin (“tampon” blok havuzundan gelen) bu mantık adresini ele geçirmesi gerektiği açıktır.

Gerekli çeviri tabloları SSD'nin kendisinde depolanır ve böylece toplam depolama kapasitesini azaltır.



**Fig. 1.3** High level view of a Flash controller



**Fig. 1.4** Logical to physical block management

**1.4 Wear Leveling**

Genellikle, aynı hafıza konumunda saklanan tüm veriler aynı sıklıkta değişmez: bazı veriler genellikle diğerleri çok uzun bir süre değişmezken güncellenir - (en uç durumda, cihazın tüm ömrü boyunca). Sıklıkla güncellenen bilgileri içeren blokların daha fazla sayıda yazma / silme döngüsü ile vurgulandığı, çok nadiren güncellenen bilgileri içeren blokların çok daha az stresli(kullanılmış) olduğu açıktır.

Rahatsızlıkları gidermek için, her bir sayfanın / bloğun yaşlanmasını mümkün olduğunca minimum ve tekdüze olarak tutmak önemlidir: yani, her sayfaya uygulanan hem okuma hem de program döngülerinin sayısı izlenmelidir. Ayrıca, bir blok için izin verilen maksimum program / silme döngüsü sayısı (yani dayanıklılığı) göz önünde bulundurulmalıdır: SLC NAND hafızalarının kullanılması durumunda, bu sayı 20–30 k döngülerindedir, bu da 10'a düşürülür. Sırasıyla MLC ve TLC NAND için 15 k ve 1-5 k.

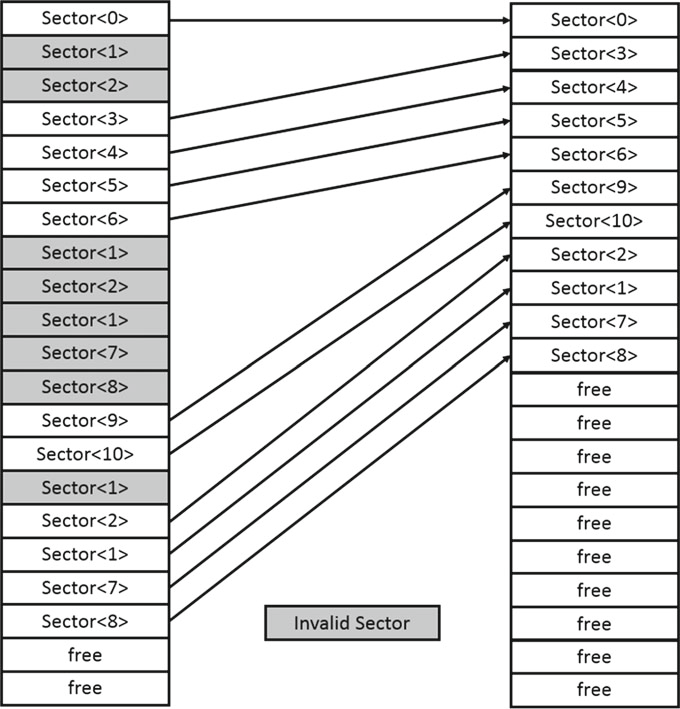
Aşınma Seviyelendirme teknikleri, mantıksal olarak fiziksel çevirme kavramını kullanır: ana bilgisayar uygulamasının her zaman aynı (mantıksal) sektörü güncellemesi gerektiğinde, bellek denetleyicisi bu sektörü dinamik olarak farklı (fiziksel) bir sektörle eşleştirir, tabi ki haritalamayı izler. Sektörün güncel olmayan kopyası geçersiz olarak etiketlendi ve silme için uygun. Bu şekilde, tüm fiziksel bloklar eşit olarak kullanılır, böylece yaşlanmayı makul bir değerde tutar.

İki tür yaklaşım mümkündür: Dinamik Aşınma Seviyelendirme, kullanıcının güncelleme talebini takip etmek için kullanılır; en düşük silme sayısına sahip ilk silinen bloğa yazma; Statik Aşınma Seviyelendirme ile Her blok, en az modifiye edilmiş olsa bile, yaşlanma ortalama değerinden saptığında yeniden haritalama için uygundur.

**1.5 Garbage Collection**

Her Aşınma Seviyelendirme teknikleri, güncellemelerle doldurulabilecek serbest sektörlerin kullanılabilirliğine güvenmektedir: serbest sektörlerin sayısı belirli bir eşiğin altına düştüğünde, sektörler “sıkıştırılmış” ve birden fazla, eski kopya silinir.Bu işlem, geçersiz sektörleri içeren blokları seçen Çöp Toplama modülü tarafından gerçekleştirilir, en son geçerli içeriği serbest sektörlere kopyalar ve ardından bu blokları siler (Şekil 1.5).

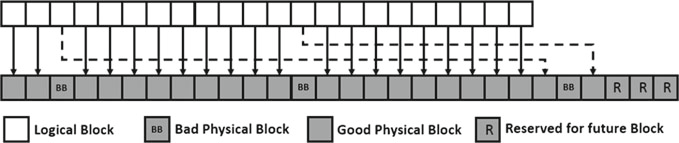
Performans üzerindeki etkileri en aza indirmek için, çöp toplama arka planda gerçekleştirilebilir. Aşınma dengeleme dağılımları tarafından tahrik edilen yaşlanma tekdüzeliği, tek sıcak noktalardan ziyade tüm dizi üzerinde stres yıpranmaktadır. Bu nedenle, belirli bir iş yükü ve kullanım süresi verildiğinde, bellek yoğunluğu arttıkça, hücre başına daha az aşınır.



**Fig. 1.5** Garbage collection

**1.6 Bad Block Management**

Aşınma seviyelendirme algoritması ne kadar akıllı olursa olsun, NAND Flash belleklerinin içsel bir sınırlaması, sözde Kötü Blokların (BB) varlığıyla temsil edilir.Diğer bir deyişle güvenilirliği garanti edilmeyen bir veya daha fazla konum içeren bloklardır. Kötü Blok Yönetimi (BBM) modülü, Şekil 1.6'da gösterildiği gibi, hatalı blokların bir haritasını oluşturur ve korur. Bu harita fabrikada oluşturulur ve SSD'nin ömrü boyunca, bir blok ne zaman kötü olursa olsun güncellenir.



**Fig. 1.6** Bad Block Management (BBM)

**1.7 Error Correction Code (ECC)**

Bu görev genellikle bellek denetleyicisinin içindeki bir donanım hızlandırıcısı tarafından yürütülür. Gömülü ECC ile hafıza örnekleri de rapor edilmiştir. Birden fazla hatayı düzelten en popüler ECC kodları Reed-Solomon ve BCH'dir.

NAND raw BER, üretimin yaklaşmasından sonra, Shannon sınırının bir sonucu olarak, daha da kötüye gidiyor. Sonuç olarak, yumuşak bilgi işlemeye dayalı düzeltme teknikleri giderek daha popüler hale gelmektedir: LDPC (Düşük Yoğunluklu Parite Kontrolü) kodları bu yumuşak bilgi yaklaşımının bir örneğidir.

**1.8 SSD’s Interfaces**

SSD'leri sunucuya ve / veya depolama altyapısına bağlamak için kullanılan 3 ana arayüz protokolü vardır: Seri Bağlı SCSI (SAS), Seri ATA (SATA) ve PCI-Express. PCI-Express tabanlı SSD'ler, en yüksek performansı sunar ve çoğunlukla sunucu tabanlı dağıtımlarda, sunucunun içinde bir eklenti kart olarak kullanılır.

SAS SSD'ler oldukça iyi düzeyde performans sunar ve hem üst düzey sunucularda hem de orta ve üst düzey depolama kasalarında kullanılır. SATA tabanlı SSD'ler çoğunlukla istemci uygulamalarında ve giriş seviyesi ve orta sınıf sunucu ve depolama kasalarında kullanılır.

**1.9 SAS and SATA**

Seri Bağlı SCSI (SAS), depolama aygıtları ve ana bilgisayar arasındaki verileri taşımak için geleneksel olarak kullanılan bir iletişim protokolüdür. SAS, bir seri noktadan noktaya fiziksel bağlantıya dayanmaktadır. Cihaz iletişimini sürmek için standart bir SCSI komut seti kullanır. Bugün, SAS tabanlı cihazlar en çok 6 Gbps hızında çalışmaktadır, ancak 12 Gbps SAS da kullanılabilir. Diğer tarafta, SAS arayüzü ayrıca eski sistemleri desteklemek için daha düşük hızlarda (1.5 Gbps ve / veya 3 Gbps) çalıştırılabilir.

SAS, ikinci nesil SATA sürücülerle geriye dönük uyumluluk sunar. Uluslararası Bilgi Teknolojileri Komitesi (INCITS) T10 teknik komitesi, SAS protokolünü geliştirir ve sürdürür; SCSI Ticaret Birliği (SCSITA) teknolojiyi desteklemektedir.

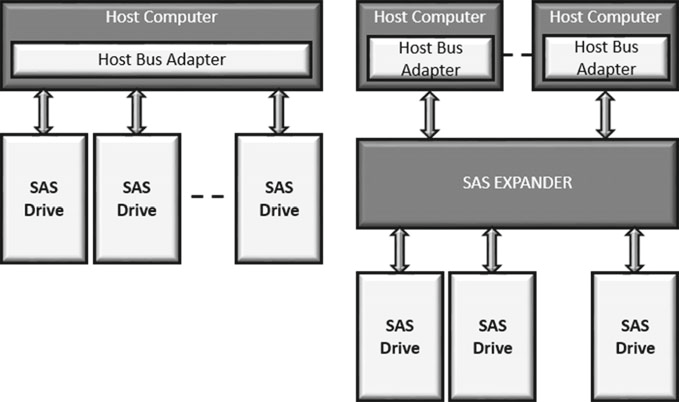
Seri ATA (SATA veya Serial Advanced Technology Attachment), ana bilgisayar veri yolu adaptörlerini, sabit disk sürücüleri ve katı hal sürücüleri gibi yığın depolama aygıtlarına bağlamak için kullanılan başka bir arabirim protokolüdür. Seri ATA, eski paralel ATA / IDE protokolünün yerini alacak şekilde tasarlanmıştır. SATA ayrıca bir noktadan noktaya bağlantıya dayanmaktadır. Cihaz iletişimini yürütmek için ATA ve ATAPI komut setlerini kullanır. Bugün, SATA tabanlı cihazlar 3 veya 6 Gbps'de çalışır.

Seri ATA endüstri uyumluluk özellikleri, Seri ATA Uluslararası Organizasyonundan (aka. SATA-IO) kaynaklanmaktadır.

Tipik bir SAS eko-sistemi, bir SAS arka paneline ya da bir noktadan noktaya bağlantı yoluyla bir ana bilgisayar veri yolu adaptörüne takılan SAS SSD'lerden oluşur; bu da, ya bir genişletici ya da doğrudan yoluyla ana mikroişlemciye bağlanır. Şekil 1.7'de gösterildiği gibi.

Her genişletici, toplam 65535 (64k) SAS bağlantısı sağlamak için 255'e kadar bağlantıyı destekleyebilir. Aslında, SAS tabanlı dağıtımlar, paylaşılan bir depolama ortamında çok sayıda SAS SSD'nin kullanılmasını sağlar.

SAS SSD'ler iki bağlantı noktası ile üretilmiştir. Bu çift bağlantı noktası işlevi, ana sistemlerin SAS SSD'lerine yedek bağlantı kurmasına izin verir. SSD'ye olan bağlantılardan birinin kırılması veya düzgün çalışma olmaması durumunda, ana sistemlerin hala SAS SSD'ye sürekli erişimi sağlamak için kullanılabilen ikinci bağlantı noktası vardır. Yüksek kullanılabilirliğin mutlak bir gereklilik olduğu kurumsal uygulamalarda bu özellik önemlidir.



**Fig. 1.7** SAS connectivity

SAS SSD'ler ayrıca hot-plug'ı da destekler: Bu özellik SAS SSD'lerinin sistem çalışırken dinamik olarak kaldırılmasını veya takılmasını sağlar; yeni takılan SAS SSD'lerin otomatik olarak algılanmasını sağlar. Aslında, bir sunucu veya depolama sistemi çalışırken, yeni takılan SAS SSD'leri dinamik olarak yapılandırılabilir ve kullanıma sokulabilir. Ayrıca, SAS SSD'leri çalışan bir sistemden çekilirse, ana sisteme zaten bağlı olan tüm uçuş verileri SAS sürücüsünde saklanır ve SSD tekrar açıldığında daha sonraki bir noktadan erişilebilir.

SAS'tan farklı olarak tipik bir SATA altyapısı, ana bilgisayar mikroişlemcisi tarafından çalıştırılan bir ana bilgisayar veri yolu adaptörüne bağlanan SATA SSD'lerden nokta-noktadan oluşur. Ayrıca SATA diskleri, SAS SSD'lerden farklı olarak, tek bir portla üretilmiştir. Bu iki ana farklılık, SATA tabanlı SSD'leri giriş veya orta ölçekli dağıtımlar ve tüketici uygulamaları için iyi bir uyum haline getirmektedir.

SATA protokolü hot-plug'ı destekler; Ancak, tüm SATA sürücüleri bunun için tasarlanmamıştır. Aslında, hot-plug, gerçekte hala in-flight(havalı\*) olan işlenmiş verilerin bir güç düşüşü sırasında güvenli bir şekilde saklanmasını garantilemek için özel bir donanım (yani ek maliyet) gerektirir.

SATA sürücülerinin SAS arka planlarına bağlanabileceğini vurgulamak gerekir, ancak SAS sürücüleri SATA arka planlarına bağlanamaz. Tabi ki, piyasada geniş SATA etkisi için bu başka bir sebep.

SAS ve SATA arasındaki benzerlikler şunlardır:

• Her iki tip de SAS arka paneline takılır;

• Sürücüler bir SAS sürücü bölmesi modülünde değiştirilebilir;

• Her ikisi de dünya çapında kabul görmüş uzun süredir kanıtlanmış teknolojilerdir;

• Her ikisi de noktadan noktaya(point-to-point) mimariyi kullanır;

• Her ikisi de takılabilir.

SAS ve SATA arasındaki farklar şunlardır:

• SATA cihazları daha ucuzdur;

• SATA cihazları, ATA komut setini, SAS'ı SCSI komut setini kullanır;

• SAS sürücülerinde çift bağlantı noktası kapasitesi ve daha düşük gecikmeler vardır;

• Her iki tip de SAS arka paneline takılıyken, bir SATA arka paneli SAS sürücülerini barındıramaz;

• SAS sürücüleri çok daha katı özelliklere karşı test edilir;

• SAS sürücüleri daha hızlıdır ve değişken sektör boyutu, LED göstergeleri, çift bağlantı noktası ve veri bütünlüğü gibi ek özellikler sunar;

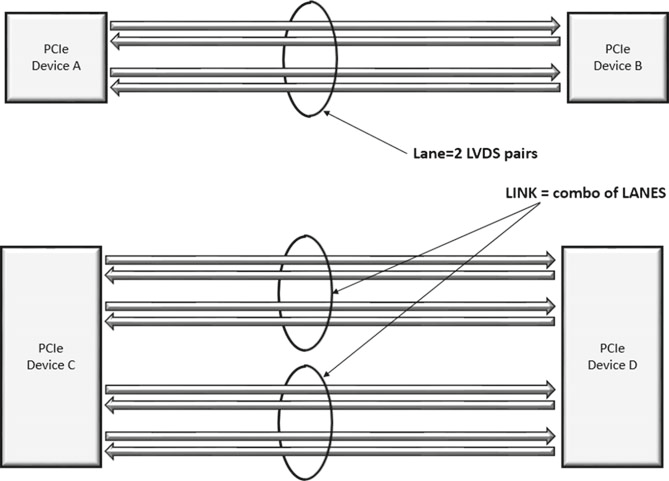
• SAS bağlantı kümelemeyi (geniş bağlantı noktası) destekler.

**1.10 PCI-Express**

PCI-Express (Peripheral Component Interconnect Express) veya PCIe, PCI ve PCI-X'in yerini alan bir veriyolu standardıdır.

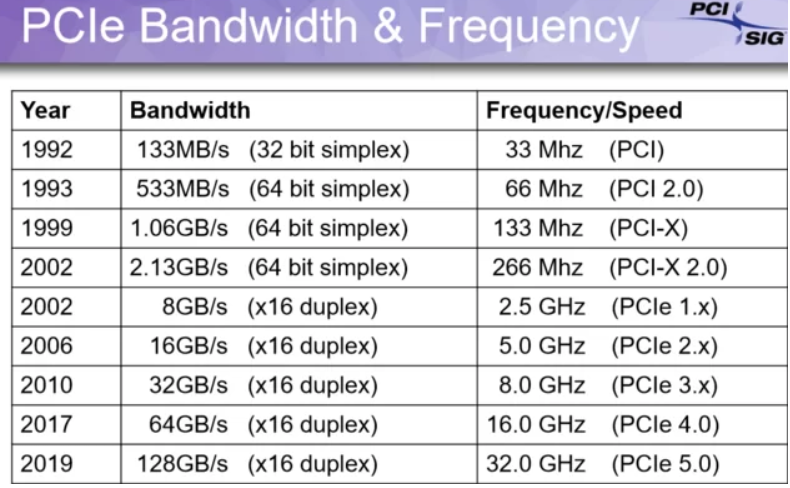
PCI-SIG (PCI Special Interest Group) PCIe tanımını oluşturur ve sürdürür.

PCIe, kurumsal sunucular, tüketici kişisel bilgisayarlar (PC), iletişim sistemleri ve endüstriyel uygulamalar dahil olmak üzere tüm bilgisayar uygulamalarında kullanılır.



**Fig. 1.8** PCI Express lane and link. In Gen2, 1 lane runs at 5Gbps/direction; a 2-lane link runs at 10 Gbps/direction

Paylaşılan paralel veri yolu mimarisini kullanan eski PCI veri yolu topolojisinin aksine, PCIe, her aygıtı kök kompleksine (ana bilgisayar) bağlayan ayrı seri bağlantılarla noktadan noktaya topolojiye dayanır.Ayrıca, bir PCIe bağlantısı iki uç nokta arasında tam çift yönlü iletişimi destekler. Veri aynı anda yukarı yönde (UP) ve aşağı yönde (DP) akabilir. Bu özel tek yönlü seri noktadan noktaya bağlantıların her bir çiftine, Şekil 1.8'de gösterildiği gibi bir şerit denir. PCIe standardı sürekli olarak geliştirilmekte olup, PCIe 4.0 piyasadaki standartların en son versiyonudur (Tablo 1.1). Standardizasyon kuruluşu şu anda Gen5'ü tanımlamak için çalışıyor. PCIe'nin diğer önemli özellikleri arasında güç yönetimi, çalışırken değiştirilebilir cihazlar ve eşler arası veri aktarımlarını işleme yeteneği (ana bilgisayar üzerinden yönlendirmeden iki uç nokta arasında veri gönderme) bulunur.



Buna ek olarak, PCIe, paralel veriyolu mimarilerine göre kablo sayısını önemli ölçüde azaltan bir seri teknolojiden yararlanarak pano tasarımını basitleştirir.

İki cihaz arasındaki PCIe bağlantısı 1–32 şeritli olabilir. Paket verileri şerit boyunca şeritlidir ve aygıt başlatması sırasında şerit sayısı otomatik olarak görüşülür.

PCIe standardı, çoklu genişlikler için yuvaları ve konektörleri tanımlar: × 1, × 4, × 8,× 16, × 32. Bu, PCIe'nin düşük verimlilik, düşük maliyetli uygulamalar ve performans açısından kritik uygulamalar sunmasına izin verir.

PCIe, bir işlem katmanı, bir veri bağlantı katmanı ve bir fiziksel katmandan oluşan paket tabanlı katmanlı bir protokol kullanır. (Şekil 1.9'da gösterildiği gibi.) İşlem katmanı, verilerin ve durum mesajı trafiğinin paketlenmesini ve \*\*\*paketten çıkartmasını(de-packetizing)\*\*\* sağlar. Veri bağlantı katmanı, bu İşlem Katmanı Paketlerini (TLP'ler) diziler ve iki uç nokta arasında güvenilir bir şekilde teslim edilmelerini sağlar.Bir verici cihaz bir uzak alıcı cihaza bir TLP gönderir ve bir CRC hatası tespit edilirse, verici cihaz bir bildirim alır.Verici cihaz TLP'yi otomatik olarak tekrar eder. Hata kontrolü ve arızalı paketlerin otomatik tekrarlanması ile PCIe çok düşük Bit Hata Oranı (Bit Error Rate) sağlar.

Fiziksel Katman iki bölümden oluşmaktadır: Mantıksal Fiziksel Katman ve Elektriksel Fiziksel Katman.

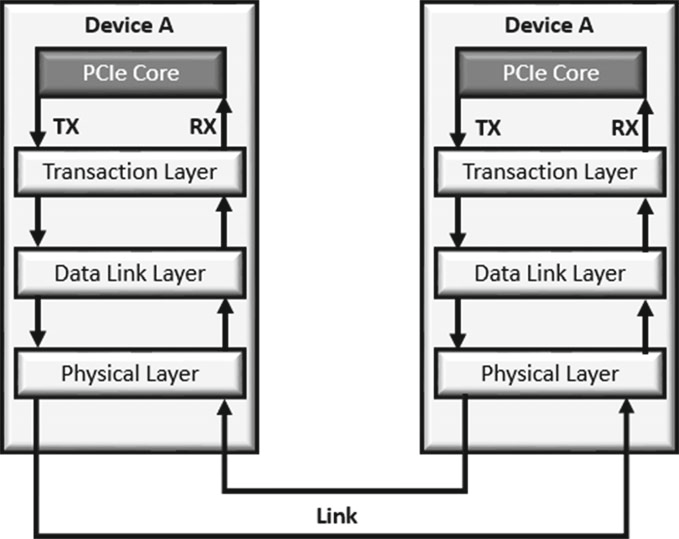
Mantıksal Fiziksel Katman: Link üzerinde iletim yapmadan önce paketlerin işlenmesi için mantık kapılarını ve Bağlantıdan Veri Bağlantısı Katmanına giden paketleri işlemektedir.

Elektriksel Fiziksel Katman: Fiziksel Katman'ın analog arabirimidir: her şerit için farklı sürücü ve alıcılardan oluşur.

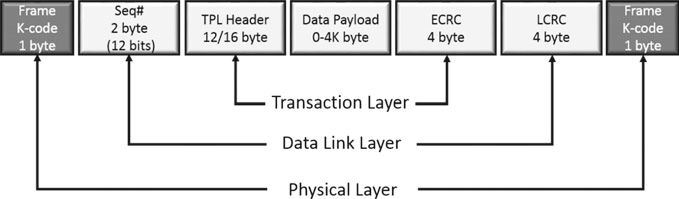
TLP montajı Şekil 1.10'da gösterilmiştir. Başlık ve Veri Yükü TLP’nin temel bilgileridir: İşlem Katmanı, bu bölümü uygulama yazılımı katmanından alınan verilere dayanarak toplar.

Pakete isteğe bağlı bir Uçtan Uca CRC (End-to-End CRC- ECRC) alanı eklenebilir. ECRC, Header ve Data Payload içindeki CRC hatalarını kontrol etmek için bu paketin son(nihai\*) hedeflenen cihazı tarafından kullanılır. Bu noktada Veri Bağlantısı Katmanı, kimliği korumak için bir sıra kimliği ve yerel CRC (local CRC-LCRC) alanı ekler. Ortaya çıkan TLP, her biri pakete 1 baytlık bir Başlangıç ve Bitiş çerçeveleme karakterlerini birleştiren Fiziksel Katmana iletilir. Son olarak, paket mevcut şeritleri kullanılarak Bağlantıda kodlanır ve farklı şekilde iletilir.

Bugün, PCIe, ana bilgisayar CPU'larını yerleşik IC'lerle ve ek çevresel genişletme kartlarıyla birbirine bağlayan ana anakart teknolojisi olarak, tüketici dizüstü bilgisayarlarından kurumsal sunuculara kadar hemen hemen tüm bilgisayarlarda kullanılan yüksek hacimli bir emtia(satılacak şeyler,mallar) ara bağlantıdır.



**Fig. 1.9** PCIe Layered architecture



**Fig. 1.10** Transaction Layer Packet (TLP) assembly

**1.11 The Need for High Speed Interfaces**

İşlemci satıcıları, bireysel işlemci çekirdeğinin performansını artırmaya, birden çok çekirdeği tek bir çipte birleştirmeye ve çoklu işlemci sistemlerinde birden fazla çipi birbirine yakınlaştırabilecek teknolojiler geliştirmeye devam etti. Sonuçta, böyle bir senaryodaki tüm çekirdeklerin aynı depolama alt sistemine erişmesi gerekir.

Kurumsal BT (bilgi teknolojisi) yöneticileri, çok işlemcili sistemleri kullanmaya isteklidir çünkü bir sistemin işlem yapabileceği saniye başına I / O işlem sayısını (IOPS) ve ayrıca Watt başına IOPS sayısını artırma potansiyeline sahiptir.

Bu çok işlemcili bilgi işlem yeteneği, maliyet ve güç tüketimine göre daha iyi IOPS sunar - işlem öğelerinin verilere zamanında erişebildiğini varsayar. Verileri bekleyen aktif işlemciler zaman ve para harcar. Elbette, sistemde her işlemci çekirdeğine kod ve veri besleyen çoklu depolama teknolojisi seviyeleri vardır. Genellikle, her çekirdek çekirdek hızında çalışan yerel önbellek içerir. Bir çipte birden çok çekirdek ikinci düzey ve bazen de üçüncü düzey bir önbellek paylaşır. Ve DRAM önbellekleri besler. DRAM ve önbellek erişim süreleri, veri aktarım hızıyla birlikte işlemcinin performansını eşleştirecek şekilde ölçeklenmiştir.

Sorun, erişim süresi ve veri hızı açısından DRAM ve HDD arasındaki performans boşluğu. Disk / sürücü satıcıları, daha yüksek kapasiteli, daha düşük-Gbyte diskleri / diskleri tasarlamak ve üretmek için harika bir iş çıkardılar; Ancak sürücüler, ne kadar hızlı bir şekilde veriye erişebilecekleri konusunda sınırlamalara sahiptir ve bu verileri DRAM'e ne kadar hızlı aktarabilirler.

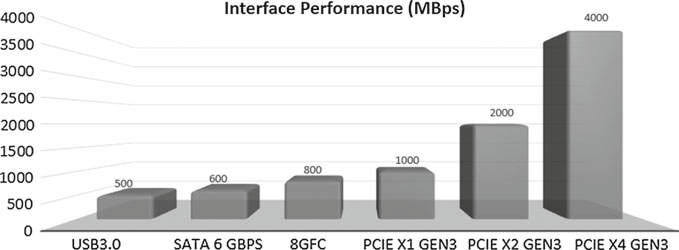
Erişim süresi, bir sabit diskin ne kadar hızlı bir şekilde okuma kafasını bir diskin üzerindeki gerekli veri yolu üzerinde hareket ettirebileceğine ve adreslenen sektörün dönme gecikmesinin kafanın altına hareket etmesine bağlıdır. Maksimum aktarım hızı, diskin dönme hızı ve veri kodlama şeması tarafından belirlenir: birlikte, diskten okunabilecek saniye başına bayt sayısını belirlerler.

Sabit diskler, sıralı verileri okumak ve aktarmak için nispeten iyi performans gösterir. Ancak rastgele arama işlemleri gecikme katıyor ve sıralı okuma işlemleri bile en son işlemcilerin veri iştahı ile eşleşemez.

Bu arada, finansal işlemler ve veri madenciliği (örneğin müşteri ilişkileri yönetimi için başvurular) gibi çevrimiçi işlem gerçekleştiren kurumsal sistemler, verilere oldukça rastgele erişim gerektirir. Ayrıca bulut bilişim, özellikle sanallaştırmaya bakıldığında, tek bir sistemin tek seferde aktif olduğu farklı uygulamaların kapsamını genişleten güçlü rastgele gereksinimlere sahiptir. Her bir mikro saniyelik gecikme süresi doğrudan para, işlemcilerin kullanımı ve sistem gücü ile ilgilidir.

Neyse ki, Flash bellekler DRAM ve HDD arasındaki performans farkını azaltmaya yardımcı olabilir. Flash, DRAM'den daha yavaştır, ancak Gbyte depolama alanı başına daha düşük bir maliyet sunar. Bu maliyet, disk depolamadan daha pahalıdır, ancak işletmeler, Premium'u, Mbyte / s ve rasgele verilere daha hızlı erişim açısından çok daha iyi bir çıktı sunarak, döner depolamaya kıyasla daha iyi bir IOPS maliyeti sağladığından, prim ödemesini memnuniyetle karşılayacaktır.

Ne eski disk sürücüsü form faktörü ne de arabirim Flash tabanlı depolama için idealdir. SSD üreticileri, disk sürücüleri için geliştirilmiş güç profilini kolayca aşacak şekilde 2,5 inç form faktöründe yeterli Flash cihazı paketleyebilir. Ve Flash, en yeni nesil disk arabirimlerinden bile daha yüksek veri aktarım hızlarını destekleyebilir.

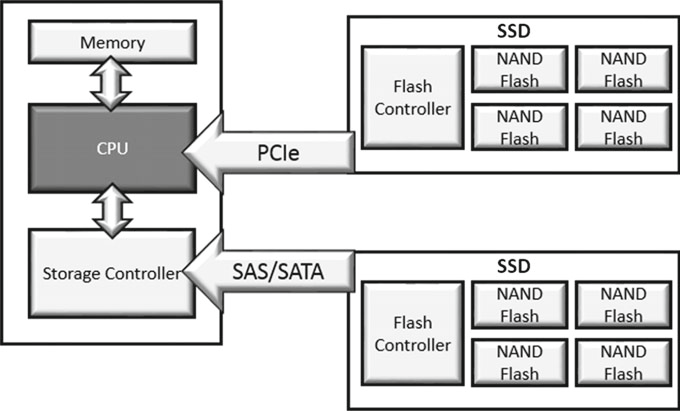


**Fig. 1.11** Arabirim performansı. PCIe, gecikmeyi azaltarak ve işlem hacmini artırarak genel sistem performansını artırır

Disk arayüzlerini performansı (Şekil 1.11). Üçüncü nesil SATA ve SAS 600 Mbyte / s verimini destekler ve bu arabirimlere dayalı sürücüler, kurumsal sistemlerde zaten kullanım bulmuşlardır. Bu veri hızları en hızlı elektromekanik sürücüleri desteklerken, yeni NAND Flash mimarileri ve çoklu kalıplı Flash paketleme, SATA ve SAS arabirimlerinin çıktı kapasitesini aşan toplu Flash bant genişliği sağlar. Kısacası, SSD performans darboğazı depolama ortamından ana bilgisayar arayüzüne kaymıştır. Bu nedenle, birçok uygulamanın Flash depolamadan tam olarak yararlanmak için daha hızlı bir ana bilgisayar bağlantısı gerekir.

PCIe ana bilgisayar arayüzü, bu depolama performansı darboğazının üstesinden gelebilir ve SSD'yi doğrudan PCIe ana bilgisayar veriyoluna bağlayarak benzersiz bir performans sunar. Örneğin, 4 şeritli (x4) PCIe Generation 3 (Gen3) bağlantısı 4 GB / sn veri hızı sağlayabilir. Basitçe söylemek gerekirse, PCIe istenen depolama bant genişliğini karşılar. Ayrıca, doğrudan PCIe bağlantısı sistem gücünü azaltabilir ve eski depolama altyapısına atfedilen gecikmeyi kesebilir.

Açıkçası, PCIe gibi bir arabirim çok kanallı bir Flash depolama alt sisteminin bant genişliğini ele alabilir ve ek performans avantajları sunabilir. Disk arabirimi kullanan SSD'ler, disk I / O işlemlerini gerçekleştiren bir depolama denetleyicisi IC tarafından eklenmiş gecikme de yaşar. PCIe cihazları doğrudan ana veri yoluna bağlanır, böylece eski depolama altyapısı ile ilişkili mimari katman ortadan kaldırılır. PCIe SSD'lerin zorlayıcı performansı, sistem üreticilerinin PCIe diskleri sunuculara ve depolama dizilerine yerleştirerek, IOPS başına maliyeti artırırken uygulamaları hızlandıran katmanlı depolama sistemleri oluşturmasıyla sonuçlandı.



**Fig. 1.12** PCIe SSD vs SAS/SATA SSD

PCIe'yi bir depolama ara bağlantısı olarak kullanmanın yararları açıktır. SATA veya SAS'a kıyasla 6 kattan fazla veri çıkışı elde edebilirsiniz. SATA ve SAS arayüzlerinde ana bilgisayar veri yolu adaptörleri ve SerDes IC'leri gibi bileşenleri ortadan kaldırabilirsiniz; sistem düzeyinde para ve güç tasarrufu sağlar ve Şekil 1.12'de gösterildiği gibi PCIe depolamayı ana CPU'ya yaklaştırır ve gecikmeyi azaltır.

Gecikme, IOPS, bant genişliği, güç, arabirim hızı, kanal sayısı, NAND türü (SLC, MLC, TLC, QLC), SSD tasarımcılarının hedef özelliklerini en düşük maliyetle karşılamak için dikkate almaları gereken parametrelerdir. Gelecekte, ortaya çıkan bellekler de oyunun bir parçası olacak. Bu değişkenler göz önünde bulundurulduğunda, özellikle prototip yapımına dayanan basit bir yaklaşımın, özellikle pazara çıkış zamanına bakıldığında, takip edilmesi zordur. Bu nedenle, SSD simülatörleri olması zorunludur.