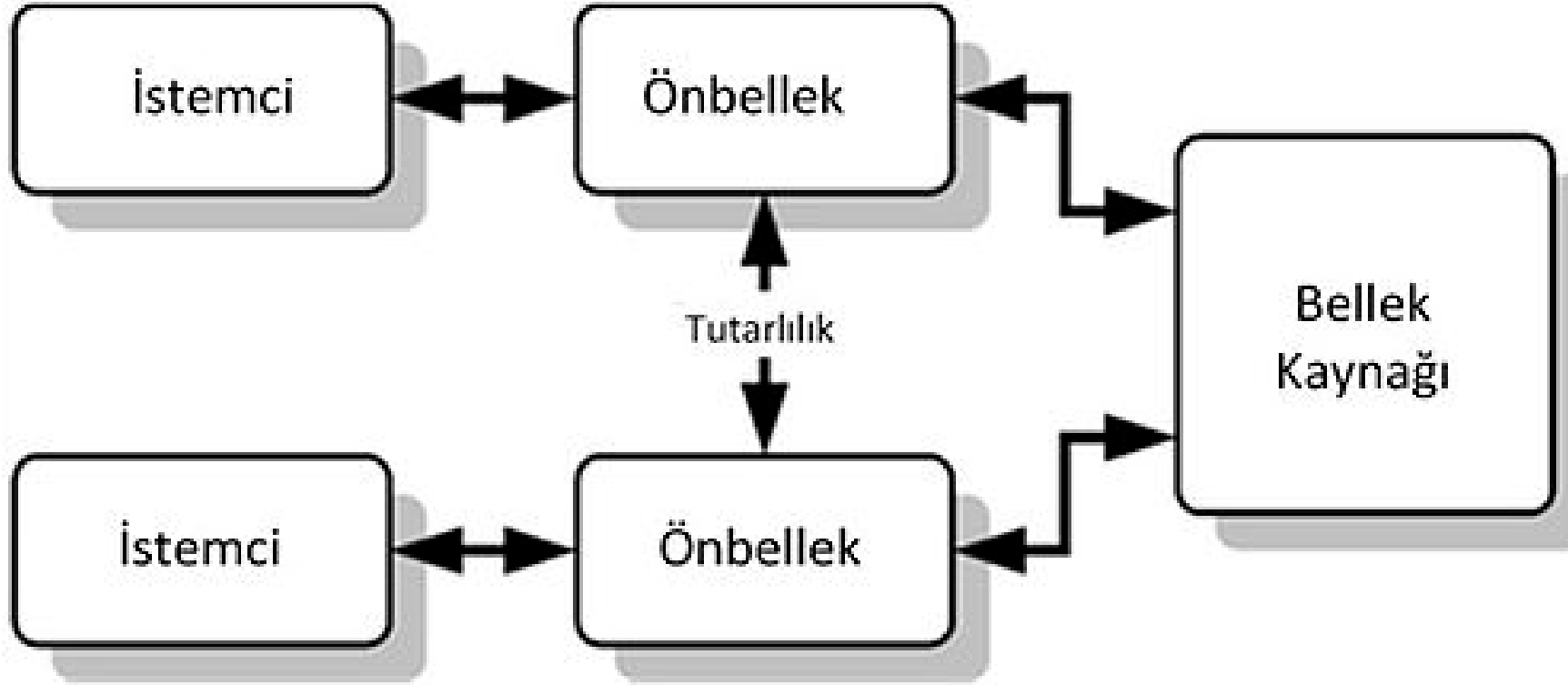


BİLGİSAYAR MİMARİSİ ve ORGANİZASYONU



10. BÖLÜM - BELLEK ORGANİZASYONU

Bellek Hiyerarşisi



1



- ❖ Bilgisayarlar enformasyon (bilgi) üzerinde çeşitli işlemler yaparak yeni enformasyon oluşturmaktadır. Gittikçe üzerinde işlem yapılacak ve bu işlemleri yapacak yardımcı enformasyon miktarları artmaktadır. Bunları kendinde depolayan ve gerektiğinde işlemcinin kullanımına sunan çeşitli bellek birimleridir. Temel olarak işlemci ile direk olarak iletişim kurabilen ve adreslenebilen bellek birimine ana bellek (main memory) denmektedir. Hem ana bellek, hem diğer bellekler erişim hızı ve enformasyon depolama kapasitesi ile birbirinden ayrılmaktadır.
- ❖ Erişim hızı, bir bellek ünitesine bir enformasyonun yazılması/okunması süresine denilir. Erişim hızı birimi genelde ns'dir.
- ❖ Bellek kapasitesi, aynı anda bellekte tutulabilen maksimum enformasyon miktarıdır.



2



❖ Bellek kapasitesinde aşağıdaki birimler kullanılırlar:

1 bit = 0;1

4 bit = 1 nibble

8 bit = 1 Byte

16 bit = yarım kelime (sözcük)

32 bit = bir kelime (sözcük)

64 bit = iki kat kelime (sözcük)

$1024 = 2^{10} = 1k$ (kilo)

1 milyon = $2^{20} = 1m$ (mega)

1 milyar = $2^{30} = 1g$ (giga)

1 trilyon = $2^{40} = 1t$ (tera)



❖ Bu arada sanal bellek terimi de kullanılmaktadır. Fiziksel belleğin kapasitesinin yeterli olmadığı durumlarda ki modern belleklerin temel organizasyon tipi olan sanal bellekten kullanılır. Yani sanal bellek, az miktarda bellek kapasitesine sahip olan bilgisayarda büyük kapasiteli bellek yeri isteyen programları çalıştırmak için geliştirilmiş olan bir yöntem olarak görülür. Örneğin, ana belleğimizde 5 mb'lık boş yer var ise veya onun kapasitesi 5 mb ise ve 10 mb'lık belleğe ihtiyacı olan bir programı çalıştırmak isterseniz, olağan durumda elinizdeki program belleğe sığmadığı için çalıştırmak mümkün değildir, fakat sanal bellek organizasyonu kullanan bir bilgisayarda bunu yapmak mümkün olmaktadır. Bunun nedeni programınızın bütün parçalarının aynı anda çalışacak durumda olmamasıdır. Bunu kullanarak o an ihtiyaç duyulmayan parçaların yardımcı bellek olarak da adlandırılan sabit diske atılması ve gerektiğinde ana belleğe çağırılması gibi işlemler yapılarak bu darboğazdan çıkılmaktadır.



❖ Bu örnekten işlemcide çalışan programın aktif olmayan kısımlarının ana bellekte tutulmasına gerek olmadığı, böylece bir yardımcı belleğe ihtiyaç olduğu gözükmemektedir. Böyle yardımcı bellekler gibi sabit manyetik disk (hard disk), esnek manyetik disk (floppy disk), optik disk (cd-rom) vs. kullanılmaktadır.

❖ Diğer taraftan günümüzde işlemci saat frekansının, dolayısı ile de onun hızının çok artması daha yavaş erişim hızlı ana bellek birimleriyle ile işlemci arasında hız bir tezat teşkil etmektedir. Bu sistemin performansını düşürmekte, sistem hızını daha yavaş olan ana bellek erişim hızına düşürmektedir. Bunu engellemek için işlemci ile aynı saat frekansında çalışan bellekler kullanılmaktadır. Bu türlü belleklere önbellek (ara-bellek, tampon bellek, cache) denmektedir.



5



❖ Ön-bellek kullanan bir işlemci bir veriyi ana bellekten okumak (veya yazmak) istediğinde önce ön-belleğe bakar, eğer istenen veri oradaysa ve ön-belleğe yazıldığı andan sonra içeriği değiştirilmemiş ise (bu, özel tanımla belirlenir) bu veriyi okur. Yazma sürecinde ise önce ön-bellekteki, sonra ise uygun bir zaman bulunduğu ana bellekteki bir adresin içerikleri değiştirilir. Dolayısıyla belleğe yazma, sisteme belli bir yük getirir ve ön-belleğin kapasitesinin fazla artırılması sistemin performansını artıracakları yerde düşürmeye neden olur.



❖ Böylece, görüldüğü gibi bellek birimleri erişim hızı ve kapasitelerine göre değişik tiplere bölünmektedirler ve bir hiyerarşi oluşturmaktadırlar. Bu hiyerarşinin en yukarı kısmında hız itibari ile işlemciye yakın olan bellek birimleri durmaktadırlar. Bu birimler işlemcinin kaydedicilerinden oluşan işlemcinin yerel belleği ve ön-bellektir. Ön-bellekler son zamanlarda çok seviyeli olarak kullanılmaya başlanmıştır. L1 (level – seviye) adlanan ön-bellek direk işlemcinin içerisinde yer almaktadır. İkinci seviye ön-bellek (L2) çoğu zaman işlemci kartına yerleştirilmektedir. Bundan dolayı onun erişim hızı L1'e göre biraz daha düşük, kapasitesi ise biraz daha büyüktür.



7



Şekil 3.1: Bellek hiyerarşisi

Bilgisayardaki bellek birimlerinin hiyerarşik yapısı gösterilmektedir. Hiyerarşinin üst kısımlarındaki bellek birimleri daha hızlı, fakat küçük kapasiteli alt kısımlarındaki ise daha büyük kapasiteli fakat yavaşlardır.



- ❖ İşlemci yalnız hızlı belleklerle direk temasta bulunabilir. Genelde hiyerarşide komşu olmayan bellek birimleri birbirleri ile direk temasta bulunmamakta, aradaki birim yardımıyla temasa geçebilmektedirler. Üç ayrı bellek alt sistemi oluşturmaktadırlar. Bunlar;
 - ✓ L1 ve L2 ön-belleklerini içeren ön-bellek alt sistemi,
 - ✓ Ana bellek ve yardımcı belleği içeren sanal bellek alt sistemi,
 - ✓ Yardımcı bellek ve çeşitli çıkarılıp takılabilen manyetik bellek birimlerini de içeren bellek alt sistemidir.
- ❖ Belleğin böyle hiyerarşik birimlere bölünerek kullanılması tabii ki, bilgisayar organizasyonunu karmaşık hale getiren bir yöntemdir. Bunun nedeni de bellek biriminin pahalı olmasıdır. Özellikle, yüksek hızlı belleklerin üretimi çok pahalıya mal olmaktadır.

Önbellek Organizasyonu



10



❖ Önbellek (Cache memory), işlemci ile her zaman direkt temasta olan tampon (Buffer) bellek türüdür. Önbellek, sürdürülmekte olan program ve verilerin en yoğun olarak kullanılan kısımlarını tutmak ve bu yolla bilgisayarın işlem sürdürme hızını (verimliliğini, performansını) yükseltmek için kullanılır. Bu yüzden, bir bilgisayarda kullanılacak olan önbellek, statik RAM'dan 4 - 10 kat hızlı olacak ve veri transferinin hızı açısından işlemciden geri kalmayacak şekilde tasarlanır, DRAM dan sonra kullanan çağdaş kişisel bilgisayar teknolojisinde ise bu oran 25 kata kadar çıkartılmıştır.

Önbellek Alt Sisteminde Zaman Değerlendirmeleri



11



❖ İşlemcinin Belleği Yöneten Birimi BYB (Memory Management Unit), gereken bilgiyi Okumak/Yazmak için önce önbelleğe başvurur. Eğer gereken bilgiyi Verecek/Alacak, olan esas bellek bloğu önbellekte yedeklenmiş durumda ise, işlemci esas belleğe başvurmaksızın yalnız önbellekle çalışır. Aksi durumda ise:

- ✓ Sürdürülmekte olan program durdurulur (kesilir).
- ✓ Gereken bilgiyi içinde bulundurmakta olan veya yazılacak olan bilgiyi tutacak olan esas bellek bloğunun kopyası önbelleğe getirilir.
- ✓ Durdurulmuş olan program, durmaya sebep olmuş komutun tekrar sürdürülmesi ile hiçbir şey olmamış gibi devam ettirilir.



- ❖ Önbelleğin kullanılması neticesinde elde edilebilecek verimlilik (performans) yükselişi aşağıdaki gibi değerlendirilebilir.
- ❖ BYB bellekten okumak veya belleğe yazmak için önbelleğe başvurur, aradığı bloğu önbellekte bulamadığında ise esas belleğe müracaat eder. Eğer gereken bloğu esas bellekte bulursa onu önbelleğe getirir ve kullanır. Bulamadığında ise işlemcinin yardımı ile işletim sistemini çağırır. İşletim sistemi, gereken bloğu içinde bulunduran kütüğü (segment, sayfa) sabit diskten esas belleğe getirir. Veri yokluğu nedeni ile durdurulmuş olan komutun sürdürülmesi tekrar edilir. Bu defa BYB, yeni gelen kütüğün gereken bilgiyi içeren bloğunu önbelleğe koyar ve işini devam ettirir. Böylece, istisnasız olarak bütün müracaatlar önbellekten geçer.



❖ Esas belleğe ise yalnız, önbellekte arama başarısız olduğunda (Cache miss) başvurulur. Bu yüzden, önbelleğe başvuru olasılığı her zamanı 1'dir. Esas belleğe başvuru olasılığı ise önbellekte arama başarısızlığı olasılığı ile aynıdır. Böylece, aranan bilgiyi önbellekte bulma (Cache hit) olasılığı P ise, esas belleğe başvuru olasılığı

$$Q = 1 - P \text{ 'dir.}$$

❖ Önbellekten ve esas bellekten birim bilgi (bayt, söz vs.) alma suresi uygun olarak T_k ve T_e olsun. Bu durumda, işlemcinin birim bilgi alma suresinin ortalama değeri **$T_o = T_k + (1 - P)T_e = T_k + Q \times T_e$**

gibi, önbellek sayesinde bellekle çalışma hızının yükselme katsayısı ise

$K = T_e/T_o = T_e/(T_k + Q \times T_e)$ gibi hesaplanabilir. Böylece, esas ve önbellekten oluşan önbellek alt sistemi ile çalışma hızı, yalnız esas bellekle çalışma hızından K kat yüksek olabilir.



❖ **Örnek 1:** Aranan bilginin önbellekte bulunma olasılığı $P=0.90$, önbelleğin hızı ise esas belleğin hızından 20 kat yüksek olduğu durum için önbellek alt sisteminin esas belleğe nispeten kaç kat hızlı çalışacağını belirleyelim.

$$Q = 1 - P = 1 - 0.90 = 0.1;$$

$$T_k = (1/20)T_e = 0.05T_e;$$

$$K = T_e / (0.05xT_e + 0.1xT_e) \cong 6.7' \text{dir}$$

❖ Önbellek, hız yükseltme amacı ile kullanılsa da bazı durumlarda bunun tersini de yapabilir. Yukarıdaki formüllerden görüldüğü gibi, P 'nin belli bir değerinin altında $T_o > T_e$ olacaktır ki, bu da hız düşürme demektir. Yani bu durumlarda, önbelleğin kullanılması bilgisayarın performansının düşmesine sebep olacaktır.



14



❖ Bu olay, matematik olarak aşağıdaki gibi ifade edilebilir. Önbelleğin negatif etki yapıcı duruma düşmesi için aşağıdaki şartın oluşması gerekir: **$T_o > T_e$** Bu formülde, **$T_o = T_k + Q \times T_e$** olduğu dikkate alınırsa

$T_k + Q \times T_e \geq T_e$; $Q \times T_e \geq T_e - T_k$; $Q \geq 1 - (T_k / T_e)$ $Q = 1 - P$
olduğu dikkate alındığında ise **$1 - P \geq 1 - (T_k / T_e)$; $P \leq T_k / T_e$**

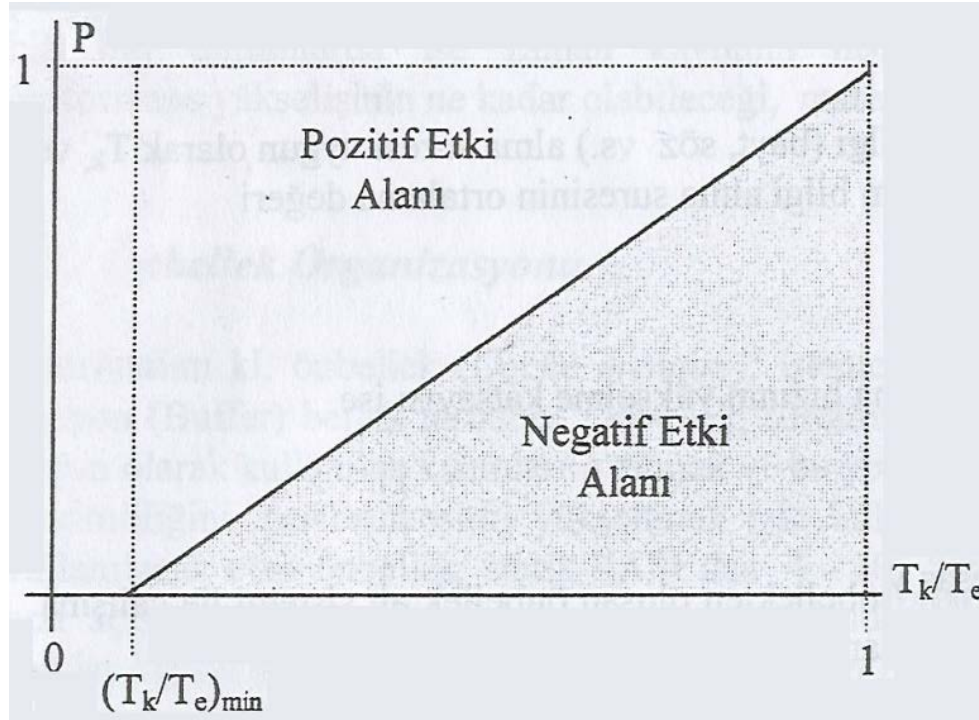
Görüldüğü gibi, $P \leq T_k / T_e$ durumunda önbellek, hız yükseltme yerine hız düşüşüne sebep olacaktır. Yani, pozitif etki yerine negatif etki yapacaktır. Grafik olarak Şekil 1'de gösterilen bu olaydan dolayı, ileri mikroişlemcilerin sistem yönetici komutları arasında önbelleği Etkili/Etkisiz yapan komutlar da bulunmaktadır.



15



❖ Şekilde gösterilen bu olaydan dolayı, ileri mikroişlemcilerin sistem yönetici komutları arasında ön belleği Etkili/Etkisiz yapan komutlar da bulunmaktadır



Şekil :Önbelleğin pozitif ve negatif etki yaptığı durumlar