



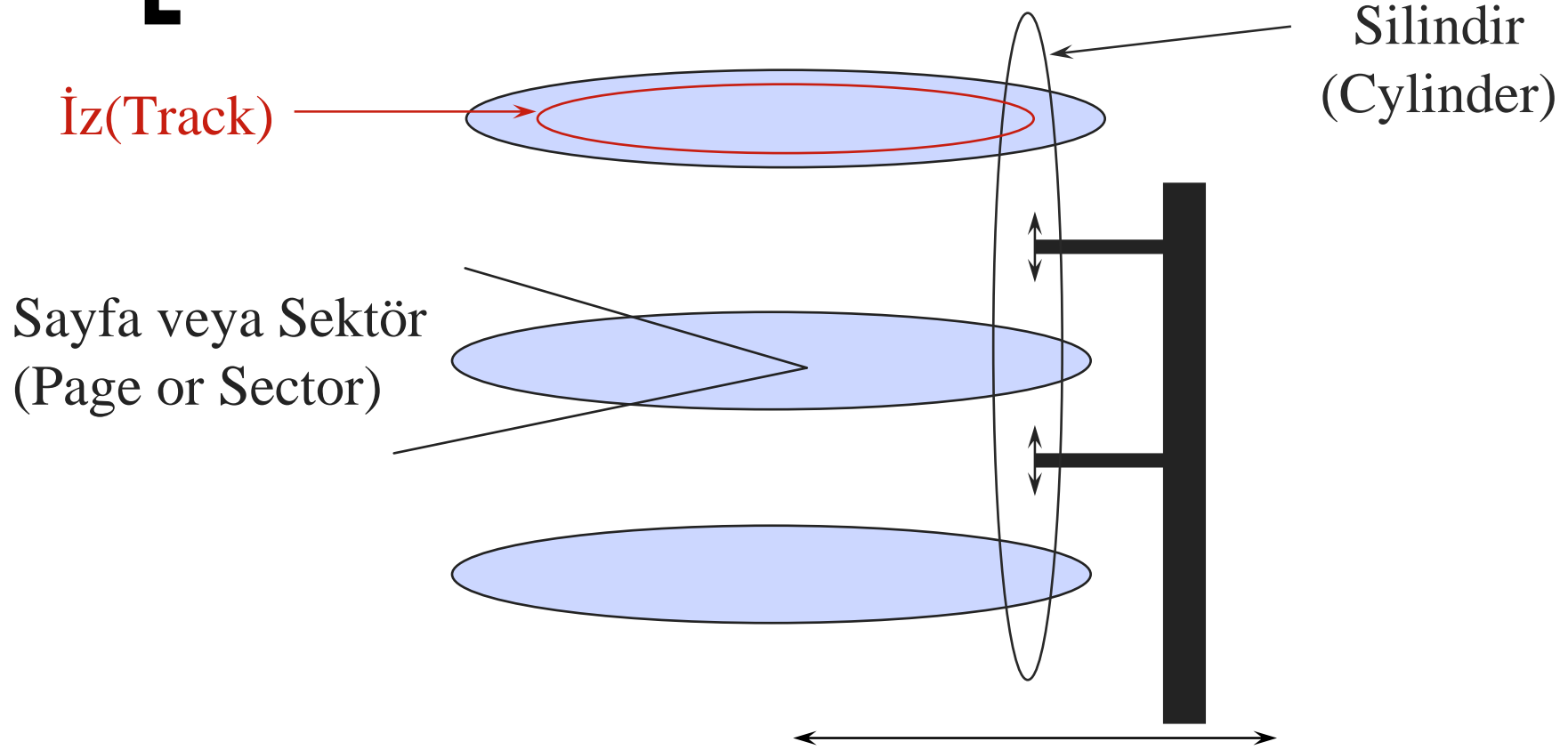
Fiziksel Tasarım Konuları

Dosya Organizasyon Teknikleri
Kayıt Erişim Yöntemleri
Veri Yapıları

[Fiziksel Tasarım]

- İyi performans için
 - Hızlı cevap zamanı
 - Minimum disk erişimi

[Disk Yapısı]



Erişim Zamanı= Arama Zamanı + Dönme Zamanı + Aktarım Zamanı

[Tanımlar]

- Arama zamanı
 - Okuma yazma kafasının hareketi için ortalama süre
- Dönme zamanı
 - Okuma yazma kafasının ilgili sektöre ulaşabilmesi için geçen ortalama süre
- Transfer
 - Bir sektörden okuyup belleğe aktarmak için geçen süre

[Ek tanımlar]

- Mantıksal kayıt
 - Bir varlık hakkındaki veri (tabloda bir kayıt)
- Fiziksel kayıt
 - Saklama ortamı üzerinde; bir sektör, sayfa veya blok
- Tipik olarak birkaç mantıksal kayıt bir fiziksel kayıta depolanır.

Dosya Organizasyon Teknikleri

- Üç teknik
 - Heap (sirasız)
 - Sıralı
 - Sequential (SAM-Sıralı erişimli bellek)
 - Indexed Sequential (ISAM-İndeksli sıralı erişimli bellek)
 - Hash edilmiş veya Direct (doğrudan)

[Heap File

Tony Lama en önce eklendi,
Digital en son eklendi.

<i>ID</i>	Şirket	Endüstri	Sembol	Fiyat	Kazanç	Dividnd.
1767	Tony Lama	Apparel	TONY	45.00	1.50	0.25
1152	Lockheed	Aero	LCH	112.00	1.25	0.50
1175	Ford	Auto	F	88.00	1.70	0.20
1122	Exxon	Oil	XON	46.00	2.50	0.75
1231	Intel	Comp.	INTL	30.00	2.00	0.00
1323	GM	Auto	GM	158.00	2.10	0.30
1378	Texaco	Oil	TX	230.00	2.80	1.00
1245	Digital	Comp.	DEC	120.00	1.80	0.10

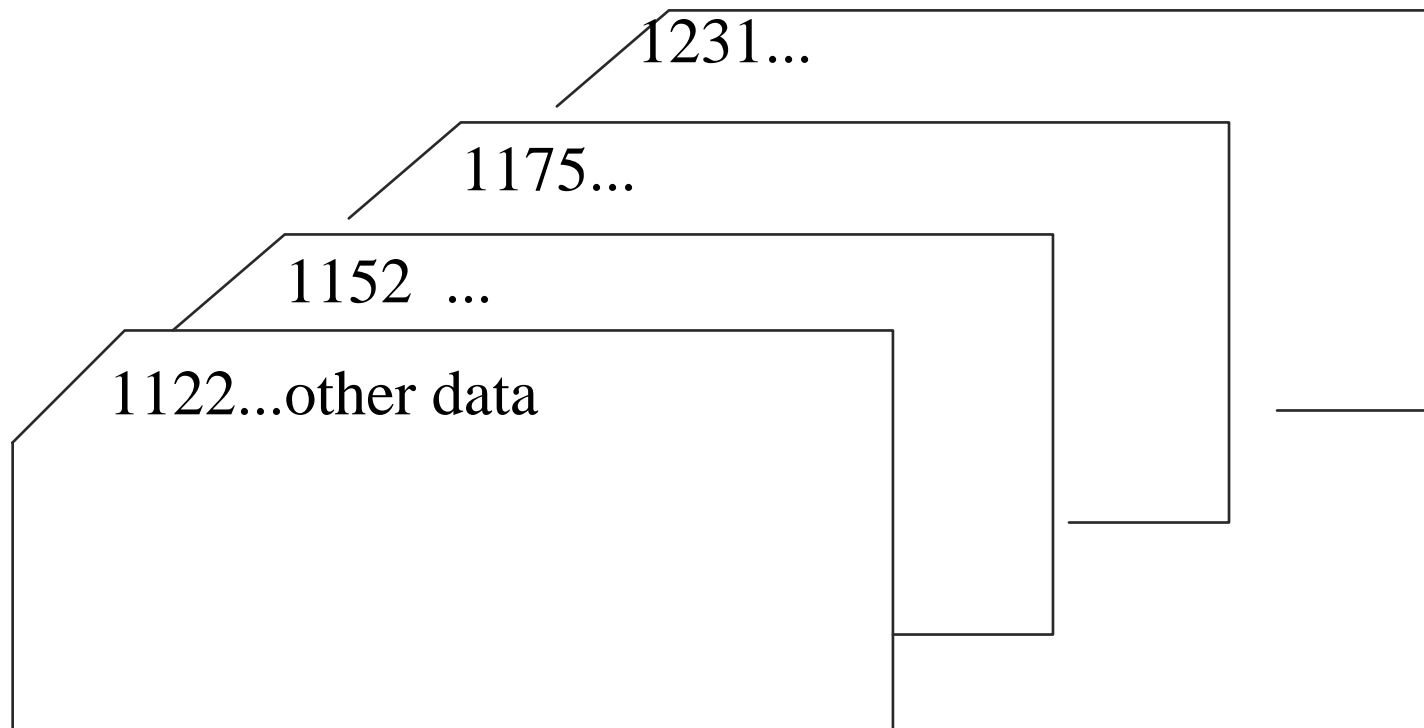
[Heap File Özellikleri]

- Ekleme
 - Hızlı: Yeni kayıtlar dosyanın en sonuna ekleniyor
- Geri alma
 - Yavaş: Sıralı erişim ihtiyacı
- Güncelleme - Silme
 - Yavaş:
 - Sayfayı bulmak için sıralı tarama lazım
 - Silme için işaretleme veya güncelleme yapılır
 - Sayfa yeniden yazılır

[Sequential (Sıralı) File]

<i>ID</i>	Company	Industry	SymbL.	Price	Earns.	Dividnd.
1122	Exxon	Oil	XON	46.00	2.50	0.75
1152	Lockheed	Aero	LCH	112.00	1.25	0.50
1175	Ford	Auto	F	88.00	1.70	0.20
1231	Intel	Comp.	INTL	30.00	2.00	0.00
1245	Digital	Comp.	DEC	120.00	1.80	0.10
1323	GM	Auto	GM	158.00	2.10	0.30
1378	Texaco	Oil	TX	230.00	2.80	1.00
1480	Conoco	Oil	CON	150.00	2.00	0.50
1767	Tony Lama	Apparel	TONY	45.00	1.50	0.25

[Sequential Access]



[Sequential File Özellikleri]

- Daha eski ortamlar (kartlar, teyp üniteleri)
- Kayıtlar birincil anahtara göre sıralıdır
- Bireysel kayıtlara doğrudan erişim kullanıldığında ihtiyaç duyulmaz
- Kayıtlara erişim
 - Kayıt bulunana kadar sıralı arama
- İkili arama hızı artırır
 - Dosya boyutu ve orta noktadaki eleman bilinmelidir

[SAM dosyalara kayıt girişi]

- Ekleme

- Yavaş:

- Kaydın nerede olduğunu bulmak gerektiği için
- Eğer yeterli yer yoksa üstüne yazma yapılır
- Eğer yeterli yer yoksa bazı kayıtlar diğer sayfaya taşınır
- Eğer orda da yer yoksa boşluk bulunana kadar tutulur

- Zamanı azaltmak için “overflow” dosyası kullanılabilir

SAM dosyalarında silme ve güncelleme yavaştır

■ Silme

○ Yavaş:

- Kaydı bulma
- Hem silme için işaretleme hem de belleği serbest bırakma
- Üstüne yazma

■ Güncelleme

○ Yavaş:

- Kaydı bulma
- Değişiklik yapma
- Üstüne yazma

GM (1323) kaydını bulmak için ikili arama

<i>ID</i>	Company	Industry	SymbL.	Price	Earns.	Dividnd.
1122	Exxon	Oil	XON	46.00	2.50	0.75
1152	Lockheed	Aero	LCH	112.00	1.25	0.50
1175	Ford	Auto	F	88.00	1.70	0.20
1231	Intel	Comp.	INTL	30.00	2.00	0.00
1. 1245	Digital	Comp.	DEC	120.00	1.80	0.10
3. 1323	GM	Auto	GM	158.00	2.10	0.30
2. 1378	Texaco	Oil	TX	230.00	2.80	1.00
1480	Conoco	Oil	CON	150.00	2.00	0.50
1767	Tony Lama	Apparel	TONY	45.00	1.50	0.25

- Doğrusal arama ile 6 seferde bulunacak kayıt 3 aramada bulunmaktadır

[Indexed Sequential]

- Disk (genellikle)
- Kayıtlar birincil anahtara göre fiziksel olarak sıralıdır
- Index her bir kaydın fiziksel yerleşimini verir
- Kayıtlar indeks ile sıralı olarak veya doğrudan erişilebilir
- İndeks bir dosyada tutulur ve dosya açıldığında belleğe bilgiler alınır.
- İndekslerin bakımı yapılmalıdır

[Indexed Sequential Access]

- Anahtar için verilen bir değer
 - Kayıt adresleri için indeks aranır
 - Adresler için bir komut okunur
 - Hızlı : bazen tek bir disk erişimi ile

Indexed Sequential Access: Fast

Find record with key 777-13-1212

<u>Key</u>	<u>Cyl.</u>	<u>Trck</u>	<u>Sect.</u>
279-66-7549	3	10	2
452-75-6301	3	10	3
789-12-3456	3	10	4

777-13-1212 < 789-12-3456
Search Cyl. 3, Trck 10, Sect. 4
sequentially.

222-66-7634
255-75-5531
279-66-7549
333-88-9876
382-32-0658
452-75-6301
701-43-5634
777-13-1212
789-12-3456

[Inserting into ISAM files]

- Çok etkin değil
 - İndeksler güncellenmeli
 - Kaydın gideceği yere konumlanılmalı
 - Eğer yer varsa yeni kayıt eklenmeli ve üstüne yazılmalı
 - Eğer yer yoksa, bir taşma alanı kullanılmalı
 - Periyodik olarak taşmalar bir dosyada birleştirilmeli

ISAM için Silme ve Güncellemeler

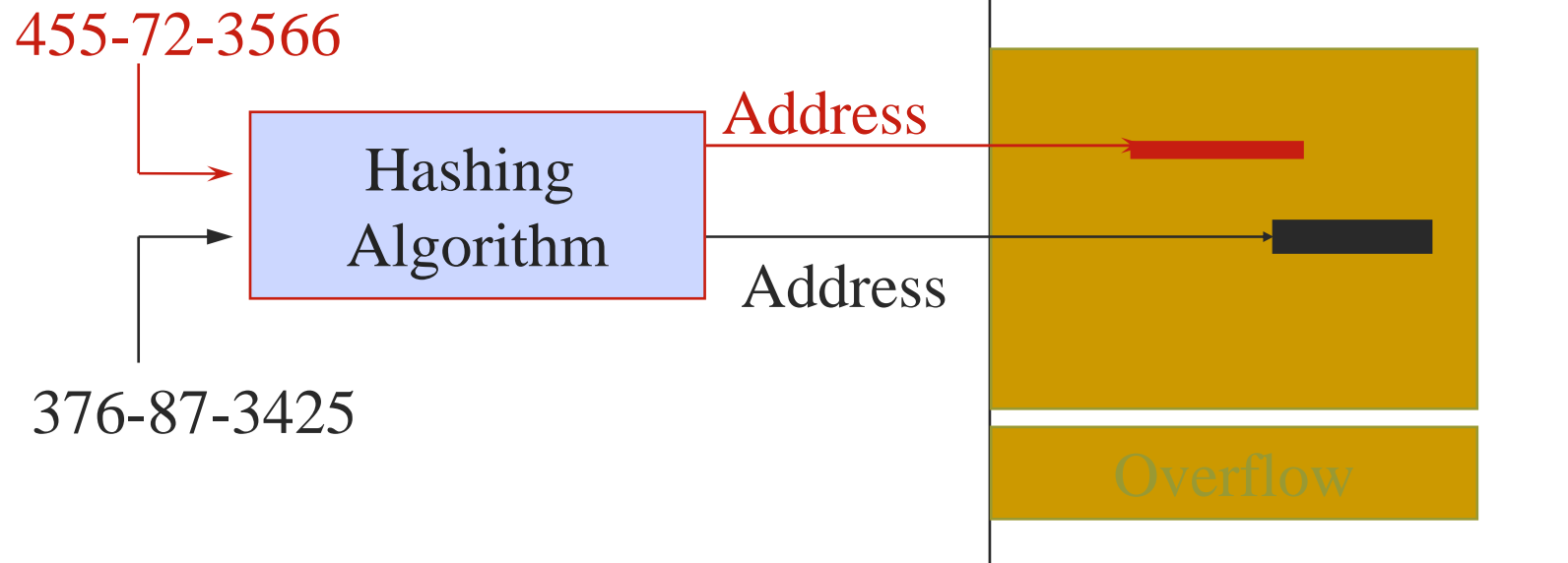
- Oldukça etkin
 - Kaydı bul
 - Silme için işaretle veya değiştirme yap
 - Yeniden yaz
 - Periyodik olarak silme için işaretlenmiş olan kayıtları sil

[ISAM dosyalar kullanılırken:]

- Sıralı ve doğrudan erişim ihtiyacı vardır.
- Müşteri dengelemeleri ve güncellemeler günlük yapılır.
- Toplu güncellemeler için genellikle sıralı erişim daha etkindir.
- Fakat biz genellikle müşteri sorularına cevap için doğrudan erişimi kullanıyoruz.

[Direct veya Hashed Erişim]

- Diskin bir parçası rezerv edilir
- Bir “hashing” algoritması kayıt adreslerini hesap eder



[Hashed Erişim Özellikleri]

- Arama ve bakım için indeksler kullanılmaz
- Doğrudan erişim için çok hızlıdır
- Sıralı erişimler için etkin değildir
- Doğrudan erişim için uygun olup sıralı erişimde kullanılmaz

[İkincil İndeksler]

- Anahtar olmayan öznitelikler üzerinden erişim sağlanır
- Örnek üç veri yapısı:
 - Linked lists: gömülü göstergeler
 - Inverted lists: çapraz-indeks tablolar
 - B-Trees

Basit Linked List: Petrol Endüstrisi

Head: 1122

<u>ID</u>	<u>Company</u>	<u>Industry</u>	<u>Symb.</u>	<u>Price</u>	<u>Earns.</u>	<u>Dividnd.</u>	<u>Next</u>
1122	Exxon	Oil	XON	46.00	2.50	0.75	1378
1152	Lockheed	Aero	LCH	112.00	1.25	0.50	
1175	Ford	Auto	F	88.00	1.70	0.20	
1231	Intel	Comp.	INTL	30.00	2.00	0.00	
1245	Digital	Comp.	DEC	120.00	1.80	0.10	
1323	GM	Auto	GM	158.00	2.10	0.30	
1378	Texaco	Oil	TX	230.00	2.80	1.00	1480
1480	Conoco	Oil	CON	150.00	2.00	0.50	-null-
1767	Tony Lama	Apparel	TONY	45.00	1.50	0.25	

Ring: Oil Industry

Head: 1122

<u>ID</u>	<u>Company</u>	<u>Industry</u>	<u>Symb.</u>	<u>Price</u>	<u>Earns.</u>	<u>Dividnd.</u>	<u>Next Rec.</u>
1122	Exxon	Oil	XON	46.00	2.50	0.75	1378
1152	Lockheed	Aero	LCH	112.00	1.25	0.50	
1175	Ford	Auto	F	88.00	1.70	0.20	
1231	Intel	Comp.	INTL	30.00	2.00	0.00	
1245	Digital	Comp.	DEC	120.00	1.80	0.10	
1323	GM	Auto	GM	158.00	2.10	0.30	
1378	Texaco	Oil	TX	230.00	2.80	1.00	1480
1480	Conoco	Oil	CON	150.00	2.00	0.50	1122
1767	Tony Lama	Apparel	TONY	45.00	1.50	0.25	

[Liste tipleri]

- Non-dense: Birkaç kayıt (Oil stocks)
- Dense: Birçok veya bütün kayıtlar (Symbol)
- Two-way: Sonraki ve Önceki göstergeleri
- Multi-list: aynı kayıt tipi üzerinde birçok liste

[Listeler ile işlem]

Bir petrol endüstrisi listesi ve bir sorgu varsayalım:

```
SELECT * FROM STOCKS  
WHERE INDUSTRY = "Oil" AND EARNINGS > 2.50
```

Endüstrisi listesi başlığını al getir.

“Sonraki Kayıt Göstergesi = Null” olana kadar

Petrol endüstrisi kaydını al getir

Eğer Earnings > 2.50, çıkış listesine dahil et.

Döngüyü Bitir.

[Listeler ile Problemler]

- Bakımı zordur
- Gösterge zincirleri bozulabilir
- Bir veya birkaç kayıt için yoğun listelerde arama yapmak etkin değildir

[Inverted List: Endüstri]

<u>Değerler Tablosu</u>	<u>Tablodaki Tekrarlar</u>
Aero	1152
Auto	1175, 1323
Apparel	1767
Computer	1231, 1245
Oil	1122, 1378, 1480

Earnings (kazançlar) ve Dividends (paylar)

Earnings		Dividends	
<u>Values</u>	<u>Occurrences</u>	<u>Values</u>	
1.25	1152	0.00	1231
1.50	1767	0.10	1245
1.70	1175	0.20	1175
2.00	1231, 1480	0.25	1767
2.10	1323	0.30	1323
2.50	1122	0.50	1152
2.80	1378	0.75	1122
		1.00	1378

[B-Trees]

- Hiyerarşik (ağaç) yapısıdır
- Dengeli
 - Yukarıdan aşağıya ve aşağıdan yukarıya bütün yollar aynı uzunluktadır
- İçerir:
 - Index Set: anahtar alan için indekslenmiş erişim
 - Sequence Set: anahtar alan için sıralı erişim

[Özet]

- Üç dosya organizasyonu vardır
 - Heaps
 - Sorted
 - Sequential Access Method
 - Indexed Sequential Access Method
 - Direct Access Method
- İkincil anahtarlar için üç veri yapısı vardır
 - Linked lists, Inverted lists, B-trees