

hw6

Cpu iş yükü

Her iş aynı süre çalışır: Bu varsayım, sistemdeki tüm işlerin tamamlanması için geçen sürenin aynı olduğunu ifade eder.

Tüm işler aynı anda gelir: Bu varsayım, sistemdeki tüm işlerin aynı anda sisteme yüklendiğini belirtir.

Tüm işler sadece işlemciyi kullanır: Bu varsayım, sistemdeki işlerin sadece işlemci (CPU) kaynaklarını kullandığını ve başka herhangi bir donanım (örneğin, disk, ağ) kaynağı kullanmadığını ifade eder.

Her işin çalışma süresi bilinir: Bu varsayım, her bir işin tamamlanması için gereken sürenin önceden bilindiğini ifade eder.

İş sıralama (scheduling)

- Performance metric: Turnaround time (Bir işin sistemde tamamlanması için geçen toplam süredir.) , **Bir iş sıralama algoritmasının ne kadar iyi çalıştığını ölçmek için kullanılan değerlerdir.**
- **Döngü Süresi = İşin Tamamlandığı Zaman - İşin Sisteme Girdiği Zaman**
- **Adillik (Fairness):** Tüm işlerin eşit şekilde ele alınması ve hiçbir işin diğerlerine göre daha fazla bekletilmemesi ilkesidir.
- Genellikle, bir iş sıralama algoritmasında performansı artırmak için yapılan değişiklikler, adilliği azaltabilir. Yani, bazı işler diğerlerine göre daha hızlı tamamlanırken, bazıları daha uzun süre bekleyebilir.

İlk Gelen İlk Çıkar (FIFO) İş Sıralama Algoritması

Sisteme en önce gelen iş, en önce tamamlanır. Yani, işler bir kuyrukta bekler ve kuyruğa ilk giren iş, kuyruktan ilk çıkar.

Ortalama Döngü Süresi: işlerin tamamlanması için geçen toplam sürenin (30 saniye) iş sayısına (3) bölünerek ortalama döngü süresinin 20 saniye olarak hesaplandığı gösterilmiştir. Döngü süresi, bir işin sistemde ne kadar süre kaldığını ifade eder.

Konvoy Etkisi Nedir?

FIFO algoritmasında, işler sisteme geldikleri sıraya göre işlenir. Ancak, bu durum bazı senaryolarda "konvoy etkisi" olarak adlandırılan bir duruma yol açabilir. Konvoy etkisi, kısa işlerin uzun bir iş nedeniyle gereğinden fazla bekletilmesi durumudur.

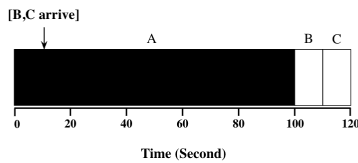
En Kısa İş İlk (SJF) İş Sıralama Algoritması

Sisteme gelen işler, tamamlanma sürelerine göre sıralanır ve en kısa sürede tamamlanacak iş öncelikli olarak işlenir.

İşlerin sonra gelmedi durumu :

A işi sistemde zaten var ve 100 saniye çalışması gerekiyor. B ve C işleri ise A işinden 10 saniye sonra sisteme geliyor ve her biri 10 saniye çalışacak.

- A işi zaten çalışmaya başladığı için durdurulamaz (SJF algoritması genellikle önleyici değildir, yani bir kez başlayan bir işin ortasında durdurulmaz).
- B ve C işleri geldiğinde, A işi zaten çalıştığı için sıraya girerler.



$$\text{Average turnaround time} = \frac{100 + (110 - 10) + (120 - 10)}{3} = 103.33 \text{ sec}$$

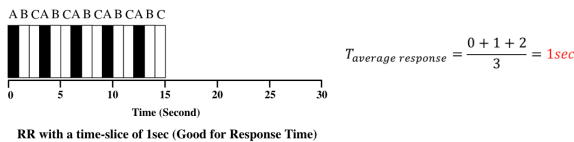
En Kısa Sürede Tamamlama İlk (STCF) Algoritması

SJF'nin aksine, bir iş bir kez başlatıldıktan sonra tamamlanıncaya kadar devam etmek zorunda değildir. STCF algoritmasında, sistemde yeni bir iş geldiğinde, kalan tüm işlerin (yeni gelen iş dahil) kalan çalışma süreleri yeniden değerlendirilir. En kısa sürede tamamlanacak iş, işlemciyi kullanmaya devam ederken, diğer işler bekler.

- **Yanıt Süresi = İlk Çalıştırılma Zamanı - İşin Sisteme Giriş Zamanı**

Round Robin (RR) İş Sıralama Algoritması

Bu algoritmada, her bir iş belirli bir zaman dilimi (time slice) için işlemciyi kullanır. Bu zaman diliminin sonunda, işlemci bir sonraki işe geçer. Bu döngü, tüm işler tamamlanana kadar devam eder.



Kısa Zaman Dilimi:

- **Avantajları:**
 - **Daha iyi yanıt süresi:** Kısa işler daha hızlı tamamlanır, çünkü uzun süre beklemek zorunda kalmazlar.
 -

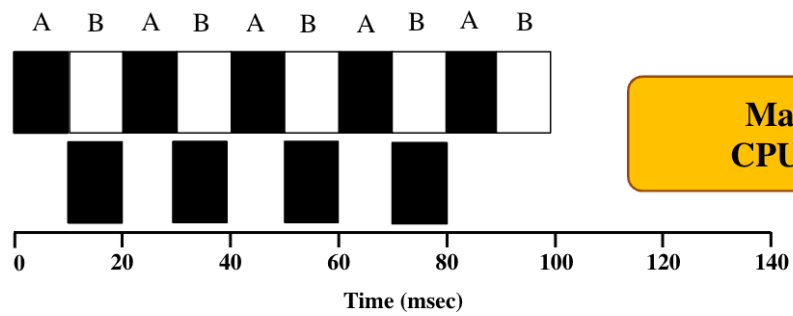
Dezavantajları:

- **Yüksek bağlam değiştirme maliyeti:** İşler arasında sık sık geçiş yapıldığı için, bağlam değiştirme (context switching) işlemi daha sık gerçekleşir.

Uzun Zaman Dilimi:

- **Avantajları:**
 - **Bağlam değiştirme maliyetinin azalması:** İşler arasında daha az geçiş yapıldığı için, bağlam değiştirme maliyeti düşer.

(İşlemci Kullanımını Maksimize Et): İkinci örnekteki durum. İşlemci sürekli olarak bir iş üzerinde çalıştırılıyor.



**Maximize the
CPU utilization**

Overlap Allows Better Use of Resources

Çok Seviyeli Geri Bildirim Kuyruğu

Multi-Level Feedback Queue (MLFQ)

İşler, sistemde geçirdikleri süreye veya diğer kriterlere göre farklı kuyruklara taşınırlar. Bu sayede, hem kısa süreli işlere hızlı yanıt verilmesi hem de uzun süreli işlerin de adil bir şekilde çalıştırılması amaçlanır. **İşlerin ne kadar süreceği önceden bilinmese bile, sistemin kendini adapte etmesini sağlamak.**

İşlerin Taşınması: Yeni gelen işler en yüksek öncelikli kuyruğa eklenir. Bir iş belirli bir süre kuyrukta kaldıktan veya işlemciyi belirli bir süre kullandıktan sonra daha düşük öncelikli bir kuyruğa taşınır.