

CPU 스케줄링 요약

1. 스케줄링 개요

CPU 스케줄링은 여러 작업(Job)이나 프로세스가 CPU를 어떻게 배분받는지 결정하는 방식입니다.

핵심 질문: "어떤 스케줄링 정책이 가장 효과적인가?" 이 질문에 대한 답을 찾기 위해 다양한 스케줄링 기법이 개발되었습니다.

2. 스케줄링 평가 항목

반환 시간 (Turnaround Time): 작업이 시스템에 들어와서 완료될 때까지 걸리는 시간. 반환 시간은 시스템 성능을 평가하는 중요한 지표입니다.

응답 시간 (Response Time): 작업이 도착한 후 처음 실행되기까지 걸리는 시간으로, 상호작용이 중요한 시스템에서 중요한 평가 기준입니다.

3. 스케줄링 알고리즘

선입선출 (FIFO, FCFS): 먼저 도착한 작업을 먼저 처리하는 방식. 간단하지만, 작업의 길이에 따라 짧은 작업이 긴 작업 뒤에서 대기해야 하는 문제(convoy effect)가 발생할 수 있습니다.

최단 작업 우선 (SJF): 짧은 작업을 우선 처리하여 평균 반환 시간을 줄이는 방식. 그러나 작업 길이를 사전에 알아야 하는 한계가 있습니다.

최소 잔여시간 우선 (STCF): SJF에 선점을 추가한 방식으로, 남은 실행 시간이 가장 적은 작업을 우선 처리합니다. 주로 대화형 시스템에서 성능이 뛰어납니다.

라운드 로빈 (RR): 일정한 시간 슬라이스마다 작업을 교체하면서 처리하는 방식. 응답 시간 측면에서 매우 효과적이지만, 반환 시간은 길어질 수 있습니다.

4. 입출력 작업 고려

입출력(I/O) 작업이 포함되면 CPU를 효율적으로 사용하기 위해 다른 작업으로 전환하는 방식이 필요합니다. STCF와 같은 스케줄러는 입출력 작업이 끝나기를 기다리지 않고 CPU 집약적 작업을 실행시켜 시스템 자원을 최대한 활용합니다.