第4回の課題について

到着順、処理時間順でそれぞれタイムチャートを作成すると以下のようになります。 ここで、 以下を仮定しています:

- 同じプロセスでは CPU 処理が終了した後に出力処理が実行できる
- 異なるプロセスの CPU 処理と出力処理は同時に実行できる

非プリエンプティブなスケジューリングではプロセス実行中の CPU の横取りはありませんが、入出力処理で自ら CPU を解放することは横取り(プリエンプション)には当たらず、あるプロセスが入出力を行っている間に他のプロセスの CPU 処理を実行することができます。

到着順									
	0			35	55 60				
CPU		P1		P2	P3				
出力				P1		P2		P3	
				45			75		100
処理時間順									
	0 5		25		60	70			
CPU	Р3	P2		P1					
出力		P3		P2		P1			
			30		50				

到着順の場合、まず P1 の CPU 処理が実行され 35 秒間かかります。それが終わった後に、このプロセスの出力処理が 10 秒間実行されますが、これと同時に 2 番目に到着した P2 の CPU 処理が実行でき 20 秒要します。これから P1 の実行完了には 45 秒かかることになります。

P2 の出力は P2 の CPU 処理が終了した後に時刻 55 秒から開始され(45 秒から 55 秒までは 出力は動作していない状態となります)、これと同時に P3 の CPU 処理が実行され、5 秒要して 時刻 60 秒で終了します。P2 の出力処理は 20 秒要しますので、時刻 75 秒に終了します。P2 の 実行には(システム開始(時刻 0)から)75 秒かかることになります。P3 の CPU 処理は終わっていますが、P2 の出力処理が続いていたので P3 の出力処理が行えず、時刻 75 秒から P3 の出力処理が開始して 25 秒で終了します。P3 はシステム開始から 100 秒で実行が完了します。したがって、

P1:45 秒、P2:75 秒、P3:100 秒 (なお、平均の実行時間は 73.3 秒になります)

CPU の利用率は、全体の実行時間 100 秒の中で実際に CPU が動作してプロセスを処理していた時間 60 秒から、60% となります。

同様に出力の利用率は、55/100 で 55% となります。

「利用効率」は問題中にも示していますが、実際に処理を行っていた時間が全体の所要時間に占める割合で、上図のタイムチャートでいいますと灰色部分(ここが処理をしていない – 遊んでいる – ところです)を除いたものの全体に対する比率です。(教科書では p.210 に「資源の利用率」として、以下の記述があります)

■2. 資源の利用率

ある資源が利用されている割合を利用率という.

① 順次使用資源の場合

プロセッサや入出力装置の場合、利用率は全時間に対する使用中の時間の割合である。図 15.2 に示すように個々の使用中の時間を T_i . 空き時間を t_i とすると、利用率は次で表される.

利用率 =
$$\frac{\sum T_i}{\sum (T_i + t_i)}$$

$$T_1$$
 t_1 T_2 t_2 T_3 t_3 \longrightarrow t_3 使用中 空き

図 15.2 順次使用資源の使用状況の例

② 空間資源の場合

メモリなどの場合、利用率は次のようになる.

全体として、CPU や入出力がどれだけ遊ばずに仕事をしているかの目安になります。OS の役割として「効率性の提供」をあげました(第1回講義資料の p.11、p.21)が、この遊んでいる時間をできるだけ減らそうということがその方策の一つです。

この問題のように CPU 処理と別プロセスの出力処理を並列実行して遊び時間を減らすことや、スケジューリングによって遊びの部分を減らすようにできます。処理時間順によるスケジューリングでは、タイムチャートで灰色の部分が少ないことがよく分かります。

処理時間順では、CPU 処理時間が一番短い P3 が最初に実行されます。それが 5 秒で完了した後、その時点で処理時間が一番短い P2 の CPU 処理が実行され、同時に P3 の出力処理が実行されます。P2 の CPU 処理が終了すると P1 の CPU 処理が実行されます。P3 の出力処理が終了すると、その時点で P2 の CPU 処理が終了しているので、P1 の CPU 処理が実行できます。P2 の出力処理が終了しても P1 の CPU 処理が終了するまで P1 の出力処理は実行されず、時刻 60 秒から実行されて 10 秒要します。

これより、P3 の実行完了には 30 秒、P2 の実行完了には 50 秒、P1 の実行完了には 70 秒要することになります。したがって、

P1:70 秒、P2:50 秒、P3:30 秒 (なお、 平均の実行時間は 50 秒になります)

CPU の利用率は、全体の実行時間 70 秒の中で実際に CPU がプロセスを処理していた時間 60 秒から、85.7% となります。

同様に出力の利用率は、55/70で78.6%となります。

それぞれの方式のメリットとデメリットは、教科書やスライド (p.24~25 など) の記述を確認してください。 実行完了時間の平均を比べると、処理時間順の方が短いことが分かります。(プロセスの処理時間はあらかじめ分からないということが問題ですが)

【提出いただいた解答へのコメント】

「P1:45 秒、P2:75 秒、P3:100 秒」などといった数値だけを解答とされる方も多くいらっしゃいますが、それに至る考え方、たとえば、「P2 は P1 の CPU 処理後に 20 秒間 CPU 処理され、その後 20 秒出力処理されて完了するので」・・・ などといった説明を記述していただくのが望ましいと考えています。

上図のようなタイムチャートを書いていただくのも良いと思います。

スケジューリング方式を比較する際には、 問題の具体的な値と結びつけて解答いただくのが良いと思います。たとえば、「到着順は CPU を独占するプロセスが 1 つでもあればそのあとのプロセスは長時間待たされてしまうという欠点がある」という一般的な記述について、 P1 のために P3 が実行できずに待ち時間が長くなるなど具体的な点を示していただくとさらに良いのではと思います。