

通信制御装置における回線の優先処理を 考慮した性能解析手法について

太田元助, 岩見彰子, 櫻尾次郎
(日立製作所 システム開発研究所)

1. まえがき

通信制御装置の1種である通信制御処理装置(以下CCPという。Communication Control Processor)の性能解析について、一方法を検討したので報告する。

CCPは、内蔵する通信制御プログラムにより回線に対するデータ送受信の制御を行ない、メッセージ・バッファ方式でもキャラクタ・バッファ方式でも使用できるといふ融通性をもち、ここでは、キャラクタ・バッファ方式で利用した場合を対象としている。

今回の対象モデルの特徴は次のとおりである。まず、CCPでの処理は、回線サービスとチャネル転送サービスとに大別できるが、このうち、回線サービスがより高いプログラム・レベルで処理される。次に、一般に、1台のCCPに複数種の回線速度が混在して接続されるが、その高速側回線からのサービス要求を優先処理するとしている。

CCPの性能解析は、処理部での待ち時間がその許容値を超える確率を求める、という方法がなされるべきものである。回線別の優先処理を行わない場合には、M/M/1のモデルとみなして分布関数を求める解析法が適用できるが、優先処理を行なう今回のようなモデルについての解析の例は見当たらない状況にある。

実際には、この種モデルの性能は、実験テストでの測定によるかあるいは最悪値の積み上げ計算に依らねている場合が多く、後者による場合は値が低めに出るものと予想される。ここでは、2速度の混在モデルという条件を与え、かつ近似式であるが、一つの解析式を得た。

2. 解析対象モデル

今回は、次のモデルを対象とした。

- (a) 2種の回線速度が混在するシステムとする。
- (b) CCPは、1チャネルに複数台接続される。

(図1)

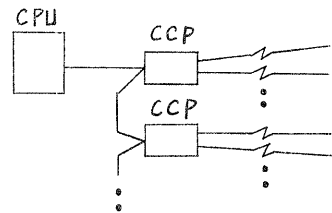


図1. 対象構成

また、伝送制御手順、メッセージ長、メッセージ発生間隔分布等が異なると、収容可能回線数に影響すると思われる。さらに、単純化を考えると、次のようにモデルの条件を設定した。

- (1) 回線に対する1バートの送信または受信ごとにプログラム処理がなされるが、送信と受信とで1回当りの処理時間が異なる。単純化のため、全回線の処理時間につき、負荷の大きい受信処理時間の方を一律に適用するものとした。
- (2) 回線利用率が小さめな程度収容回線数を増せられると予想されるが、安全側を考えると、回線利用率はどの回線も1.0であるとみなした。
- (3) 回線数は半二重回線換算で表わすものとした。すなわち、全二重通信を行なう回線は、2回線であると数える。

- (4) ある回線に着目すると、回線から1バイト受信完了するごとに処理要求(以後、Lサービス要求と称す。)が発生し、また、一定のバイト数単位でチャンネルを回線へ転送を行わねばならぬ。これらの処理時間は、CCP機種、CPU機種、伝送制御手順等に依りた値を用いる。
- (5) 回線から受信したデータは、まず1バイト・バッファに収容され、次に2面バッファへ転送される。2面バッファは、1面当りBバイトであるとする。(図2)
- (6) 高速側回線を優先処理する。(図3)
- (7) 2面バッファにつき、1面の全バイトが収容完了になると、チャンネルサービス用キューに登録し、サービスを待つ。(以後、このサービス要求をQサービス要求と称す。)
- (8) チャンネルを回線へ転送を扱う待ち行列は、図3のごとく、処理の優先レベルごとに行列をつくる。今回は、2種の行列とした。

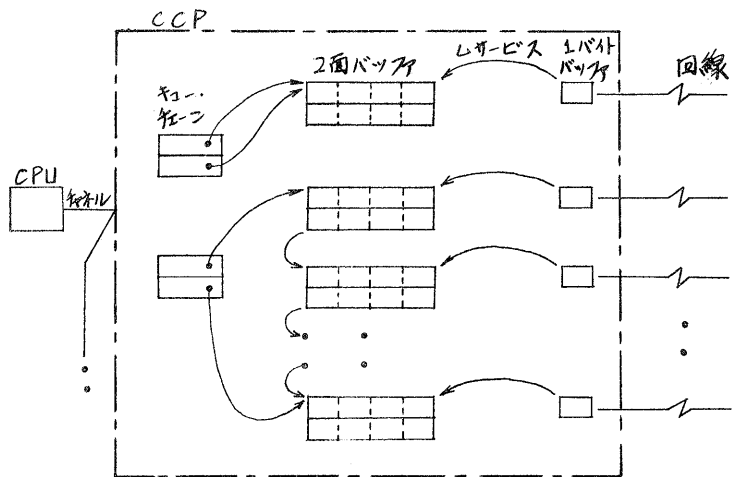


図2. 対象モデル

- (9) CCPのプログラムは、5つのレベルに分けられ、プログラム・レベル1が最も高位で、プログラム・レベル2は回線に対する送受信を扱い、プログラム・レベル3はチャンネルを回線からのデータ転送を扱う。Qサービス中にLサービス要求が発生すると、Qサービスは中断され、Lサービス要求が優先処理される。Lサービス完了後、Qサービスは中断点から再開される。

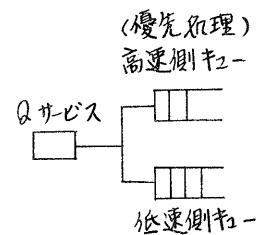


図3. 待ち行列モデル

- (10) メッセージをその先頭バイトから数え始め、Bバイトおつに区切り、Bバイトを単位としてチャンネルを回線へ転送を行なう。

3. 解析

3.1 概要

CCPの性能を決定的に上げるためには、サービス要求が発生したときそれが必要時間内に処理されないと、そのデータの属するテキスト・ブロックの再送が必要となることである。CCPでは、図2に示すLサービスと図3に示すQサービスの部分で処理待ちを生ずる。Lサービスは、要求が発生してから1バイト・タイム以内に終了しないとデータの追突を生じ、追突したデータまたは追突されたデータを破棄することになる。この現象はオーバーランと称される。Qサービスは、要求が発生してからBバイト・タイム以内に終了しないとオーバーランとなる。

いま、CCPに高速側1回線と低速側1回線を接続したとし、そのとき、Lサービス、Qサービスと各処理に十分余裕があったとする。次に、低速側回線数を増していくと、オーバーランを生ずる確率も増す。こゝで、そのシステムとして許容できるオーバーラン確率値を設定すると、オーバーラン発生確率値がその許容値に達したときの接続回線数か、収容可能な限界を示すものとなる。また、LサービスとQサービスとは、プログラム・レベルの関係からQサービスの方が先に限界に達するものと仮定し、こゝでは、Lサービスの解析を省略し、Qサービスの解析についてのみ述べるものとする。

3.2 着目回線について

高速側1回線と低速側 n 回線を接続されているとして、 n を0, 1, 2, ... と増していったとき、高速側と低速側のどちらが先にオーバーランの限界値に達するかであるが、高速、低速の速度比が大きいときは、概略試算の結果、高速側回線の方が先に限界に達することと認められたので、高速側回線のオーバーランに着目して解析するものとする。

3.3 解析式の誘導 ----- 高速側1回線と低速側 n 回線

2面バッファの1面のサイズがB(バイト)であるから、Qサービス要求がオーバーランなしにキューに滞在できる時間は、各回線につき、それぞれの回線速度でのBバイト・タイムである。高速側回線に着目すると、そのBバイト・タイム($\equiv t_b$)中にプログラムが処理すべき時間の合計は、 t_b を超えてはならない。このことから、次式が得られる。

$$t_{Li} \cdot \frac{n_d - (k-1)}{n_d} + t_{hj} + t_{hbj} + t_{Lr} \cdot r + t_{La}(m-r)$$

$$+ w_{ch} < t_b, \quad (k=1, 2, \dots, n_d; i=1, 2, \dots, n_{tL}; j=1, 2, \dots, n_{tL}) \quad \text{--- (1)}$$

$$P_n(t > t_b) = \sum_{i \leq n_{tL}, j \leq n_{tL}} (P_{tLi} \cdot P_{thj}) \sum_{k=1}^{n_d} P_{dk} \left\{ \sum_{n=0}^{\infty} P_{n,m} \left(\sum_{r=0}^m P_{m,r} \right) \right\} \quad \text{--- (2)}$$

$$P_{n,m} = {}_n C_m P_L^m (1-P_L)^{n-m} \quad \text{--- (3)}$$

$$P_L = \frac{t_b}{t_a} \quad \text{--- (4)}$$

$$P_{m,r} = {}_m C_r P_F^r (1-P_F)^{m-r} \quad \text{--- (5)}$$

$$P_F = \frac{N_F}{N_T} \quad \text{--- (6)}$$

ここに、記号の意味は次のとおりである。

t_{Li} : 低速側回線のQサービス時間。 n_{tL} 種類あり。

t_{hj} : 高速側回線のQサービス時間。 n_{tH} 種類あり。

t_{hbj} : t_b 中に発生する高速側回線のLサービス時間の計。

n_d : t_{li} を n_d 区間に等分する。

n : 仮定した低速側回線数。

w_{CH} : 1チャンネルにCIPが複数台接続されているときの、平均のチャンネル使用待ち時間。

t_{er} : テキスト・ブロックの最終バイトのサービスは、通常はバイトの2倍以上の時間がかかる。その時間である。

t_b : キュー滞在滞留時間。高速側回線からQサービス要求が発生したとき、それがキューに滞在できる時間。この時間内に処理されないオーバーラン発生となる。

t_{la} : n 回線から t_b 中に m 個のサービス要求が発生したとして、その m 個のうちの r 個が長時間サービス(t_{er})で、残りの($m-r$)個の平均サービス時間が t_{la} であるとする。

$P_n(t_b > t_b)$: 仮定した n に対するオーバーラン確率。

P_{lei} : 高速側回線のQサービス要求が低速側回線のQサービス進行中に発生したとして、複数種のサービス時間単位のうちの t_{li} のサービス時間単位が進行中であるときに発生する確率。

P_{thj} : 高速側回線のQサービス要求が生じたとき、それが t_{hj} 時間を要するサービスである確率。

P_{dk} : 高速側回線のQサービスが t_{li} 進行中に発生したとして、その中の d_k 区間中に発生する確率。

$P_{n,m}$: n 回線のうち、 t_b 中に m サービス要求を出すものの割合を m 個である確率。ただし、ある回線から t_b 中に生ずるサービス要求の最大個数は定まっている。

$P_{m,r}$: m サービス要求を出した m 個のうち、ちょうど r 個が長時間サービス要求である確率。

r_0 : n を仮定し、それに対して、 $m=0, 1, 2, \dots, n$; $r=0, 1, 2, \dots, m$ の範囲の(m, r)の組合せを考え、式(1)によつてオーバーランするかどうか判定する。そのとき、 r をゼロから1ずつ増していき、初めてオーバーランを起す値を r_0 と置く。

次に算出式の意味につき、主なものを示す。

(a) 式(1)について

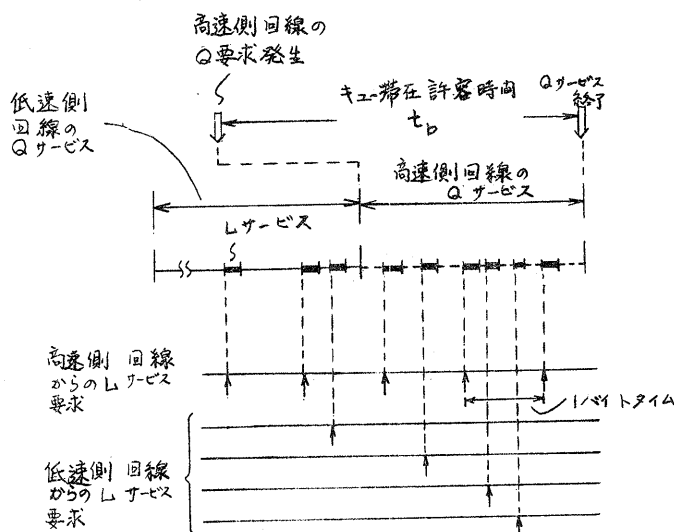
式(1)の全体は、高速側回線が i 回線のみであるとして、そのQサービス要求が発生してから終了するまでの所要時間を、サービス割込み処理等のために終了が遅らされることに着目して表わしたものである。

(b) 式(1)の第1項

この $t_{li} \cdot \frac{n_d - (k-1)}{n_d}$ は、高速側回線のQサービス要求が、低速側回線のQサービス進行中に発生したとして、高速側回線Qサービス開始の遅延時間を表わす。安全側の近似として、高速側のQサービス要求が発生したときに、低速側回線が必ずQサービス中であるとみなすものとする。

低速側回線のQサービス正味時間を t_{li} と置く。 t_{li} のために高速側Qサービス要求の開始が遅らされる状況は図4に示すごとくである。次に、 t_{li} 中のどの位置で発生したかにより開始待ち時間が異なってくるので、それを

n_d 区間に等分することを考える。さうに、その区間のどこで発生したかということであるが、近似として、区間の始めに発生したとみなすものとする。これが、この項の意味である。 n_d の値が小さければ、より安全側に近似したことになる。一方、 n_d の値を増していくと、あるところで解の値(回線収容数)の伸びが飽和することが予想される。



(c) $t_{L \cdot r}$ の項

テキスト・ブロックにつき、最初のバイトから最終バイトまでの各バイトにつき、そのLサービス

図4. 高速側回線のQサービス要求が発生してから、サービス終了するまでの所要時間

所要時間を見ると、最終バイトの処理が他のバイトの2倍以上であるので、各回線からの最終バイトサービス要求が、 t_b 中でどのように集中するかが問題であるとするものである。すなわち、 t_b 中に r 回の最終バイトサービス要求があると、その1回当りの処理に t_{Lr} かかるとして、計 $t_{Lr} \cdot r$ だけの時間がかかる。

(d) 式(1) での判定について

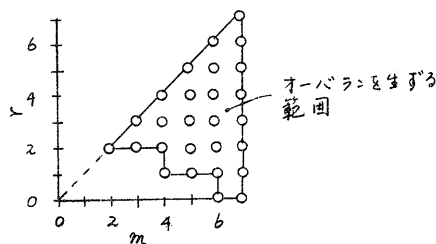
式(1)、式(2) での変数は、次の値ととり得る。

$$\begin{cases} n = 0, 1, 2, \dots, n_{\max}; & m = 0, 1, 2, \dots, n \\ n = 1, 2, \dots, n_{t_k}; & r = 0, 1, 2, \dots, m \\ j = 1, 2, \dots, n_{t_n}; & k = 1, 2, \dots, n_d \end{cases}$$

これらの変数につき、それぞれの値を定めると、式(1)により、その組合せがオーバランとなるか判定できる。

こうして、この変化範囲で示されるすべての組合せに対して、式(1)により、オーバランとなる組合せを求めろ。

例えば、 $n=7$ とし、 n, j, k はそれぞれある値としたとき、オーバランとなる m, r の組合せは図5のごとくなる。



(e) 式(2) について

式(2) は、式(1)においてオーバランする組合せを探し出したとき、その組合せを構成する各

図5. $n=7, n, j, k$ をある値としたとき、式(1)においてオーバランを生ずる m, r 組合せの範囲の例

要素についてそれぞれその要素値となる確率を求め、次にその組合せとなる確率を求め、さらに仮定した n に対して、オーバーランを生ずるすべての組合せの出現確率の総和を求めたものである。これが、そのモデルにおける n に対するオーバーラン確率値となる。

(f) 式(3)、式(4)について

低速側のある1回線に着目し、その1バイト・タイムを t_a とおく。
高速側回線0サービス要求のキュー滞在許容時間は t_b であるから、 $t_a > t_b$ のとき、 t_b 中にその回線からサービス要求の発生する確率は式(4)で与えられる。また、低速側回線が n 回線あって、その n 回線から t_b 中にちょうど m 個のサービス要求の発生する確率は式(3)で与えられる。

(g) 式(5)、式(6)について

どの回線も同様な伝送シーケンスであり、かつあるバイト数を単位として伝送シーケンスが繰り返されるものとする。その繰り返し単位となるバイト数を N_T とおき、そのうち N_T 個が最終バイト・サービス要求であるとする。すると、ある回線からサービス要求が発生したときそれが最終バイト・サービス要求である確率は式(6)で表われ、また、 t_b 中に各回線からたかだか1個ずつしか発生しないとして n 回線から計 m 個のサービス要求が発生したとき、そのうちちょうど r 個が最終バイト・サービス要求である確率は式(5)で与えられる。

3.4 解析式の拡張その1 ----- サービス時間分布

式(1)、式(2)は、サービス時間が2種類のみであるとみなしたが、3種類の場合には次のように表わされる。同様にして任意種類数の場合へ拡張できる。

- (a) 高速側回線は1回線である。
- (b) 低速側回線は1種類の速度で、 n 回線あるとする。
- (c) n 回線のうちから、キュー滞在許容時間 t_b 中に m 個のサービス要求が発生するものとする。
- (d) m 個のうち、ちょうど r 個が t_{lr} 時間のサービス要求である。
- (e) 残りの $(m-r)$ 個のうち、ちょうど u 個が t_{lu} 時間のサービス要求である。
- (f) さらに残りの $(m-r-u)$ 個は、 t_{la} 時間のサービス要求である。

これらを式(1)、式(2)に対応させて表わすと次式が得られる。

$$t_{li} \cdot \frac{n_d - (k-1)}{n_d} + t_{hj} + t_{hbj} + t_{lr} \cdot r + t_{lu} \cdot u + t_{la} (m-r-u) + w_{ch} < t_b \text{ ----- (7)}$$

$$P_n(t > t_b) = \sum_{i \leq n_{t_l}, j \leq n_{t_h}} (P_{t_{li}} \cdot P_{t_{hj}}) \cdot \sum_{k=1}^{n_d} P_{dk} \left[\sum_{m=0}^n P_{n,m} \left\{ \sum_{r=0}^m P_{m,r} \left(\sum_{u=0}^r P_{r,u} \right) \right\} \right] \text{ --- (8)}$$

3.5 解析式の拡張その2----- 高速側複数回線

高速側を複数回線とした場合への拡張について考えてみる。

これには、システムが高速側回線のみから成るものとみなし、そのQサービス時間が低速側回線からのLサービス要求に割込まれて延ばされる値を平均値的に求め、M/M/I のモデルとみてその分布関数により解析する方法も考えられる。

ここでは、これまで述べてきた解析方法の延長の方向で考えてみる。

このモデルで問題となるのは、高速側回線数を n_1 とし、そのうちのある回線からQサービス要求が出たとき、他の (n_1-1) 回線のうちからQサービス要求が既に出ていてサービス待ちになっているものが何回線あるかということである。可能性としては、0, 1, 2, ..., (n_1-1) の各場合があり得る。ここで、安全側への近似を行ない、次のようにモデルを考える。すなわち、高速側のある回線からQサービス要求が出たとき、低速側回線のQサービスが進行中であり、また、そのQサービスが終了しないうちに高速側 n_1 回線のすべてからQサービス要求が出、かつ、いま着目しているのはその n_1 のうち最後に要求を出した回線であるとみる。

Qサービス要求のキュー滞在顧客時間 t_b と高速側速度でのBバイト・タイムにとると、この t_b 中に高速側回線のおのおのから必ず1回ずつのQサービス要求が発生する。

ここで、高速側回線の記号に添字1をつけ、低速側回線に添字2をつけることにして式(1)、式(2)に対応した式と導くと次式が得られる。

$$t_{L_i} \cdot \frac{n_d - (k-1)}{n_d} + t_{Q_{R_1}} \cdot r_1 + t_{Q_{R_{n_1}}} \cdot u_1 + t_{Q_{R_{a_1}}} \cdot (n_1 - r_1 - u_1) \\ + \sum_{\pi_1} t_{L_{R_b}} + t_{L_{R_2}} \cdot r_2 + t_{L_{u_2}} \cdot u_2 + t_{L_{a_2}} (m_2 - r_2 - u_2) + W_{CH} < t_b \quad \text{--- (9)}$$

$$P_{n_1, n_2} (t > t_b)$$

$$= \sum_{i \leq n_{t_k}} P_{t_{L_i}} \sum_{k=1}^{n_d} P_{d_k} \sum_{r_1=0}^{n_1} P_{n_1, r_1} \sum_{u_1=0}^{n_1-r_1} P_{n_1-r_1, u_1} \\ \cdot \sum_{m_2=0}^{n_2} P_{n_2, m_2} \sum_{r_2=0}^{m_2} P_{m_2, r_2} \sum_{u_2=u_{20}}^{r_2} P_{r_2, u_2} \quad \text{--- (10)}$$

ここで、式(9)中の $t_{Q_{R_1}}$ と r_1 は、高速側回線のQサービスで $t_{Q_{R_1}}$ 時間要するものが n_1 個のQサービス要求のうち r_1 個あることを表わし、また式(10)中の P_{n_1, r_1} は、 n_1 回線のうち、 $t_{Q_{R_1}}$ のQサービスを行なうものが r_1 回線となる確率を表わす。また式(9)中の $\sum_{\pi_1} t_{L_{R_b}}$ は、高速側 n_1 回線についての、 t_b 中におけるLサービス時間の計である。

式(9)でオーバーランするか判定し、式(10)により、オーバーランする組合せにつきそれぞれの組合せの出現確率を求め、さらにオーバーランするすべての組合せについてオーバーラン確率の総和を求めるものである。

4. あとがき

式(1)、式(2)をはじめ特に式(9)、式(10)はかなりの近似を行ったものであるが、いずれも安全側(処理能力値が小となる方向)に寄せたものであり、これらの式にて能力の目安値が得られるものとする。

また、各計算式の中で、ある量がいくつかの値をとるとき、その最悪値で代表させて組合せ数を減らし、より簡単化を図ってもよい。

図6に、一つの計算結果を示す。

これらの解析式は今後、実験テスト値等との比較により検証されていくことが望ましい。

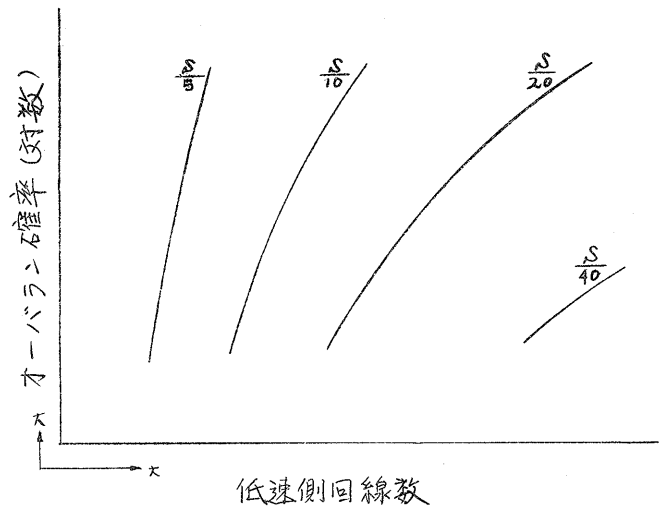


図6. 高速(速度 S)1回線と他の1速度の組合せにおける回線数とオーバーラン確率の関係例
(半対数表示)

5. 謝辞

御討議いただいた東北大学 電気通信研究所野口研究室の白鳥氏をはじめとする方々に感謝いたします。

6. 参考文献

- (1) 森村英典: 応用待ち行列理論, 日科技連, 1975.
- (2) 西田俊夫: 待ち行列の理論と応用, 朝倉書店, 1974.