

# ウィンドウ方式を用いたフロー制御の評価

## An Evaluation of Window Flow Control

松尾一紀, 安藤純利, 小野欽司

Kazunori Matsuo, Sumitoshi Ando, Kinji Ono

国際電信電話株式会社 研究所

### 1. まえがき

公衆パケット交換網は、既に数か国でサービスが行なわれており、さらにサービスを準備あるいは計画中の所は多数にのぼっている。こうして公衆パケット交換網は最も有力なデータ交換網として定着しつつあるといえる。

このため公衆パケット交換網の標準化の努力がCCITTにおいて長年にわたって続けられており、その結果、勧告X.25に象徴されるものが公衆パケット交換網の特徴を現わしている。つまり、交換方式としては、ヴァーチャル・ゴールの概念に基づくパケット交換サービスであり、技術的なポイントは、誤り制御とフロー制御にある。このうちフロー制御の問題は、パケット交換網で新しく出てきた問題であって、まだまだ研究が不足しているため、システム設計に役立てられるような技術的手法としては確立されていない。

フロー制御の目的は ① パケット網および端末のリソースの有効利用をはかりつつ、② 伝送スループットを上げることである。前者に関しては、例えばあるフロー制御方式と網あるいは交換機のリソース（ここではバッファメモリ）の割当法との関係について包括的な解答は全く得られていない。しかしながら、本論文は後者に関する

ものである。CCITTパケット網のフロー制御は、ウィンドウ方式を採用しているので、このスループットとウィンドウ数の関係を明らかにした。パケット交換網は端末が必要とするスループットを確保すべきであるから、スループット値が与えられた時、必要なウィンドウ数をセッティングすることになる。

本論文では、このための計算式と具体例を示す。

### 2. CCITTパケット網の特徴

公衆パケット交換網は、ARPANETやTSSサービス用の専用網を技術的な先駆者とし、各国通信業者の実験網の経験をベースにして標準化が進められてきた。その技術的根幹をなすのは、言うまでもなく統計的多重化（*Statistical Multiplexing*）、言い換えれば伝送路のダイナミックな割当て（*Dynamic Allocation*）である。これは計算機におけるメッセージ転送と親和性の強い技術であり、また、公衆パケット網が主要なユーザとして見込んだものがTSSサービスあるいはコンピュータアクセスサービスであったので当然の帰結である。それ故、現在のパケット交換網は、データ通信の大部分に適した網ではあるが、デジタルのデータ交換の全領域に適合する網である

とは言えない。良く知られているように、大量のデータを送る呼に対して、パケット交換機は呼処理の負荷が大きく、回線交換に比して不利になる。しかしながら、パケット交換は伝送路を有効利用できる交換方式であるし、現在の情勢では、世界的に広く利用可能となるデータ交換網はパケット交換網であることから、これの適切な利用を考えることが大切である。例えば、低速端末から計算機へアクセスする呼の場合に限定すれば、伝送のスループットを重視する必要はないということになるが、もっと一般的に高速端末による通信も考慮して、パケット網におけるスループットの問題を研究する必要がある。

CCITTで標準化されつつある公衆パケット交換網の特徴の主なものを以下に述べる。

### (1) ヴァーチャル・コール

パケット交換の方式として、大別してヴァーチャル・コールとデータグラムの2種があり、両者の比較に関して昔から論じられてきたが、CCITTパケット網の主流はヴァーチャル・コールである。

### (2) フレームレベルとパケットレベルの分離

パケット交換の機能を、1リンク上で誤り制御を主体とするフレームレベルとヴァーチャル・コールの処理やフロー制御を主とするパケットレベルとに分けて構成する。

### (3) ウィンドウ方式によるフロー制御

フロー制御に関しても多くの手法が

提案されているが、CCITTパケット網では、ウィンドウ制御による方式が採用されている。パケットレベルに送信 $P(S)$ 、受信 $P(R)$ のシーケンス番号を設け、ウィンドウ数を $W$ とすると送信側は受信側から $P(R)$ で示されたウィンドウ下端の番号から上端 $P(R)+W-1$ の番号のパケットまでしか送出できない。受信側はバッファが空くと、 $P(R)$ を更新して、さらにパケットが受信できることを示す。

### (4) *link-by-link* と *end-to-end*

勧告 X.25 は交換機と端末の間のインタフェース条件を規定するものであり、網内をいかに作るかは通信業者の自由となっている。上記のウィンドウ制御も厳密に言えば、端末とパケット交換機間でのパケット送受信を規定しているだけということになる。しかし送信端末と受信端末とはヴァーチャル・コールで対応づけられており、一方の側でのバッファ塞り状態が長く続くと次第に網内に伝わって、最後には相手端末に至り、影響を及ぼすことになるので、実効的には端末一端間みのフロー制御をしていることになる。

X.25でのウィンドウと網内のフロー制御との関連は本来自由であるとしても、 $P(R)$ を返してウィンドウ更新をするその方法については、従来から *link-by-link* と *end-to-end* の2つの方法の対立があり、勧告 X.25はその両者の妥協で文章化されている。

*link-by-link* とは、端末→交換機→-----→交換機→端末といった接続の各リンク毎に独立にウィンドウ制御を行なう、つまり各ノード毎に自分のバッファの状態に応じてウィンドウ更新の $P(R)$ を返す方式である。

*end-to-end* とは、せっかく $P(R)$ という番号があるのだから、これをフロ

一制御だけでなく、受信端末での受信確認を送信端末へ伝える手段に使おうというものである。P(R)は途中の交換機で勝手に返送されることはなく、必ず受信端末が出したものが順次前位のリンクへ渡される。これは、端末-端末間の通信という観点を強くした方法といえる。

現在での問題点は、① 公衆パケット網として、link-by-linkを採用した網とend-to-end型の網とが現実に存在していることである。両者を相互接続した場合、end-to-endをユーザに保証出来ない。② 端末-交換機のインタフェース上でX.25はパケット長に複数種類を認めており、パケットの分割・併合があるとしていることである。これではP(R)をend-to-endに戻すといっても、パケット長変換があるとそのロードで番号を適当に変換して記憶しておかねばならない。特にウィンドウ数が大きい時では変換処理が面倒になる。

現在CCITTでは、勧告X.25を修正して全ての網がlink-by-linkとend-to-endの両方をサービスするという方向にある。これにより上記①の問題は解決するにしても②の問題は残るし、また、修正案がend-to-endの場合にスループットを十分に保証できるまでの改修を考えているわけではない。

#### (5) スループット選択かウィンドウ選択か

ヴァーチャル・コールの伝送スループットを決めるパラメータについて詳しくは3章に述べるが、大まかにいって

- ① パケット長
- ② ウィンドウ数
- ③ 往復遅延時間
- ④ 回線速度

である。

端末の立場からは、必要なスループットを、パケット交換網が常に保証してヴァーチャル・コールをサービスしてくれるのが望ましい。

しかし、パケット交換網の立場からすると、現実的に可変なパラメータはウィンドウ数だけであって精密なコントロールはできないし、輻輳になるとスループットは落ちてしまうということである。それでも出来る範囲でユーザの要求に応えるために、オペショナル・ファシリテイとして呼設定時に選択させることが考えられている。

ただし、これも網によって提供条件が異なっている。

- (i) スループット選択主体でゆく
- (ii) ウィンドウ選択で代行する

要求方法としては、

- A. 呼毎に選択する
- B. 回線固定として、加入時に設定する。

(i)は要求されたスループットを確保するようにパラメータ(現実には網内のウィンドウ数)を設定するのは網の責任という考え方である。網側では、その遅延時間やパケット長から判断して適当なウィンドウ数を決めなければならない。

(ii)は実際に設定できるパラメータがウィンドウ数だけならば、それを直接指定しようという考えである。もしここでend-to-endの方式が取られているならば、スループット確保の責任は端末にある。

本論文は、この(i)、(ii)のウィンドウ数設定に役立つ計算式を与えるものである。

### 3. スループットとウィンドウ数

ウィンドウ数とスループットとは密接な関係を持ち、ウィンドウ数をいくつに設定するかということは、端末間の伝送スループットを大きく左右する。しかしながらこの関係は、データパケットの受信確認をリンクごとに独立に返すか (*link-by-link*)、相手端末から返すか (*end-to-end*) によって異なる。

ここでは、スループットとウィンドウ数との関係を *link-by-link*, *end-to-end* それぞれについて計算する方法について述べる。

スループットとウィンドウ数の関係は次の式で表わされる。

$$T = \frac{W \cdot L}{D}$$

ここで、 $T$  はスループット (*bits/sec*)、 $W$  はウィンドウ数、 $L$  はパケット長 (*bits*)、 $D$  は往復遅延時間 (*sec*) である。

回線の伝播遅延などが大きいと、送出的データパケットの受信確認が返ってくるまでに時間がかかり、回線使用率がおさえてスループットが低くなってしまいが、ウィンドウ数  $W$  を適当に選ぶ、その遅延時間に見合った個数のパケットを連続して送出できるようにしておけば、遅延によるスループットの低下をある程度防止することが可能となる。

また、パケット長が短い場合にも、ウィンドウ数  $W$  をふやし、連続して送出できるパケットの個数をふやすことにより、スループットの低下を防ぐことが出来る。

(1) *end-to-end* の場合のスループットとウィンドウ数の関係

*end-to-end* の場合のシーケンス例

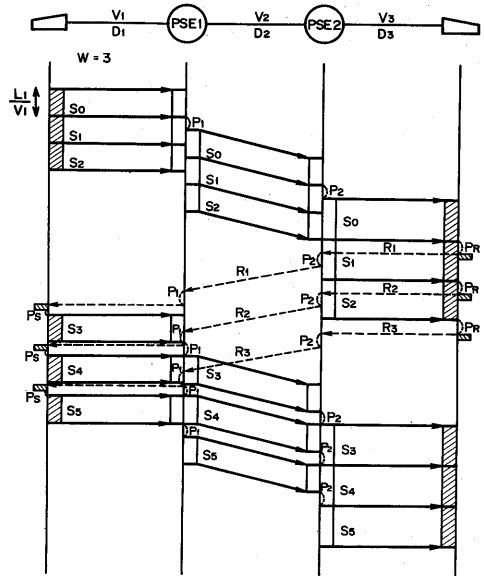


図1. *end-to-end*におけるシーケンス

を図1.に示す。この例では2つの網が大きな遅延を持つ中継リンクで接続されている網構成を図示した。

この図からわかるように、*end-to-end* の通信では、いくつかのリンクのすべての遅延要素の和がスループットに影響を与える。このことはスループットとウィンドウ数の関係式において、 $D$ の項に端末間のすべての遅延要素の和をあてはめればよいことを示している。

この例の場合

$$T = \frac{W \cdot L}{D}$$

$$D = 2(D_1 + D_2 + D_3) + 2(P_1 + P_2) + (P_1 + P_3) + (L_1 + L_2) \left( \frac{1}{V_1} + \frac{1}{V_2} + \frac{1}{V_3} \right)$$

ここで

$D_1, D_2, D_3$  ; それぞれのリンクの伝播遅延 (*sec*)

$V_1, V_2, V_3$  ; それぞれのリンクの回線速度 (*bits/sec*)

$P_1, P_2$  ; それぞれ PSE1, PSE2での処理時間 (*sec*)

$P_S, P_R$  ; それぞれ送信端末, 受信端末での処理時間 (sec)

$L_1, L_2$  ; それぞれデータパケット長, 確認パケット長 (RRパケット長またはデータパケット長)

で与えられる。

データパケットに対する確認は, 反対方向へのデータパケットの受信シーケンス番号  $P(R)$  を使用する場合は, RRパケットの受信シーケンス番号を使用する場合とがあるが, それぞれの場合について確認パケット長  $L_2$  は異なる。

他の網構成の場合についても, この式を応用して関係式を求めることが出来る。

ところで, スループット値はいくつかあるリンクのうち一番低い回線速度を越えて要求することは不可能である。そのような例を図2に示す。図では, 受信端末の速度が遅いため, 受信端末側の回線使用率が100%となり, スループット値は上限値の  $V_3 \text{ bits/sec}$  に等しくなっている。

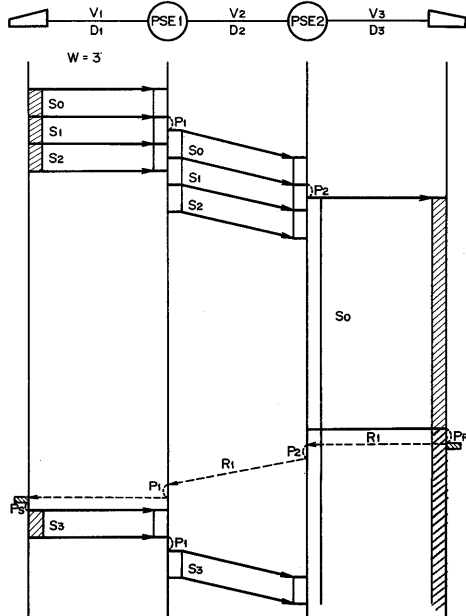


図2. end-to-endの場合におけるシーケンス(受信端末が遅速)

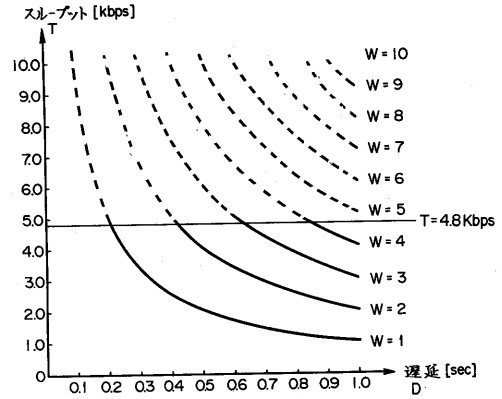


図3 スループットと遅延の関係(パケット長  $L=128$  バイト)

ゆえに, 要求できるスループットには上限があり, これはいくつかのリンクの回線速度の最小値  $\min\{V_1, V_2, \dots, V_n\}$  と一致する。この関係を図3に示す。図3の例では, 使用する回線速度の最小値が  $4800 \text{ bits/sec}$  であり, このとき点線の部分のスループットを要求することは出来ない。

## (2) link-by-linkの場合のスループットとウィンドウ数の関係

link-by-linkの場合のシーケンス例を図4に示す。この例の網構成は(1)の場合と同じである。

パケット交換機は入側回線からデータパケットを受信すると, 出側回線のウィンドウ塞がりをチェックし, ウィンドウがあいている場合には即データパケットを出側回線に送出する。ウィンドウがすでに塞がっている場合には, このデータパケットを待ち  $Q_{max}$  につなぎ, 出側回線から RR パケットを受信してウィンドウがあき次第出側回線に送出される。

さて, 図4の例では, 中継線の遅延のためこのリンクのスループットが低くなり, このため全体のスループットもそのスループットの値でおさえられ

ているが、大きな遅延をもつ中継リンクのウィンドウ数を大きくすることで全体のスループットを高くすることが可能である。

link-by-link の場合のスループットは、それぞれのリンクを独立させて考えたスループットのうち、いちばん低い値のスループットでおさえられることになる。

この例の場合、それぞれのリンクのスループットは

$$T_{e1} = \frac{W_{e1} \cdot L}{D_{e1}}$$

$$D_{e1} = 2D_1 + 2R + P_3 + (L_1 + L_2) \cdot \frac{1}{V_1},$$

$$T_{e2} = \frac{W_{e2} \cdot L}{D_{e2}}$$

$$D_{e2} = 2D_2 + 2P_2 + P_1 + (L_1 + L_2) \cdot \frac{1}{V_2},$$

$$T_{e3} = \frac{W_{e3} \cdot L}{D_{e3}}$$

$$D_{e3} = 2D_3 + P_R + P_2 + (L_1 + L_2) \cdot \frac{1}{V_3}$$

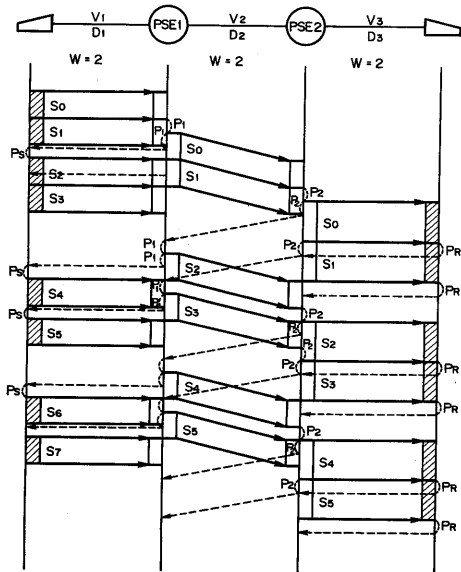


図4. link-by-linkにおけるシーケンス

で表われ、これらのリンクを経由した全体のスループットは、

$$T = \min\{T_{e1}, T_{e2}, T_{e3}\}$$

で与えられる。

link-by-link の場合特に問題となるのは、入側回線へRR パケットを返すタイミングであるが、図4の例では、出側回線のウィンドウがあり、待ち Queue につながれているデータパッケージが送出のために、level 3 のバッファから level 2 のバッファに引きつがれた時点としている。

他の方法として、出側回線のウィンドウがあり、待ち Queue につながれているデータパッケージを送出後、このデータパッケージに対するフレームレベルの確認をもらった時点で、入側回線にRR パケットを返す方法も考えられる。計算式においてはこの場合、データパッケージを出側の回線に送出後、RR パッケージを作って入側回線に送出するまでの処理時間  $P_1, P_2$  が大きくなったとして扱えばよい。

### (3) 最大データパッケージ長が異なる場合の問題

ところで、X.25 インタフェース上で許される最大データパッケージ長が、入側リンクと出側リンクとで異なる場合、パケット交換機においてモアデータマークを使ったデータパッケージの分割および併合がおこなわれる。

パケット交換機はモアデータマーク付きのデータパッケージを受信しても、そのデータパッケージ長が、出側回線で許される最大データパッケージ長より小さければ、つぎに来るデータパッケージの到着を待ち、到着後に組み合わせて1個のデータパッケージとして送出する。

この際、RR パッケージの返送が上記

のように出側のフレームとして編集した後とすると、最初に到着したモデママーク付きデータパケットに対するRRパケット送達が、つぎのデータパケット到着まで待たされるために、入側回線のスループットが低下することになる。

パケット長変換のある場合のP(R)の変換とスループットに関しては慎重な配慮が必要である。

#### 4. ウィンドウ数設定例

あるスループットを要求された時、それを満たすために必要となるウィンドウ数の設定例をend-to-end, link-by-linkそれぞれについて以下に示す。

##### (1) end-to-end の場合のウィンドウ数設定例

図2の網構成において、要求されたスループット値を満足するように設定すべきウィンドウ数を表1に示す。

この例では、国内網における回線速度  $V_1, V_2$  は  $9600 \text{ bits/sec}$ , 国内網の伝播遅延  $D_1, D_3$  はなし, 交換機1,2の処理時間  $P_1, P_2$  は  $50 \text{ ms}$ , 送信端末, 受信端末での処理時間  $P_s, P_r$  は  $50 \text{ ms}$ , データパケット, 確認パケット長  $L_1, L_2$  は  $138 \text{ bytes}$  とした。

また、中継リンクとしては衛星回線経由と海底ケーブル経由とを考慮、衛星回線経由の場合の回線速度  $V_2$  は  $56 \text{ Kbits/sec}$ , 伝播遅延  $D_2$  は  $300 \text{ ms}$ 。海底ケーブル経由の場合の回線速度  $V_2$  は  $19.2 \text{ Kbits/sec}$ , 伝播遅延  $D_2$  は  $100 \text{ ms}$  とした。

end-to-end であるから、送信側端末が自分の必要とするスループットを得るために、このようなウィンドウ数を設定して網に要求するわけである。

表1. end-to-end の場合、要求するスループットとウィンドウ数

中継リンク	往復遅延時間	要求するスループット <sup>*</sup> (bps)	ウィンドウ <sup>*</sup>	
衛星リンク	$D_1=D_3=0 \text{ ms}$ $D_2=300 \text{ ms}$ $P_1=P_2=50 \text{ ms}$ $P_s=P_r=50 \text{ ms}$ $V_1=V_s=9600 \text{ bps}$ $V_2=56 \text{ Kbps}$ $L_1=L_2=138 \text{ byte}$	1399 MS	300	1
		1200	2	
		2400	4	
		4800	7	
		9600	13	
海底ケーブル	$D_1=D_3=0 \text{ ms}$ $D_2=100 \text{ ms}$ $P_1=P_2=50 \text{ ms}$ $P_s=P_r=50 \text{ ms}$ $V_1=V_s=9600 \text{ bps}$ $V_2=56 \text{ Kbps}$ $L_1=L_2=138 \text{ byte}$	808 MS	300	1
		1200	2	
		2400	3	
		4800	5	
		9600	10	

\* end-to-end のスループットおよびウィンドウ数

表2. link-by-link の場合、要求するスループットとウィンドウ数

中継リンク	往復遅延時間 <sup>***</sup>	要求するスループット <sup>**</sup> (bps)	ウィンドウ <sup>**</sup>	
衛星リンク	$D_2=300 \text{ ms}$ $P_1=P_2=50 \text{ ms}$ $V_2=56 \text{ Kbps}$ $L_1=L_2=138 \text{ byte}$	770 MS	300	1
		1200	1	
		2400	3	
		4800	4	
		9600	8	
海底ケーブル	$P_2=100 \text{ ms}$ $P_1=P_s=50 \text{ ms}$ $V_2=19.2 \text{ Kbps}$ $L_1=L_2=138 \text{ byte}$	465 MS	300	1
		1200	1	
		2400	2	
		4800	3	
		9600	5	

\*\* 中継リンクのスループットおよびウィンドウ数

\*\*\* 中継リンクの往復遅延時間

## (2) link-by-link の場合のウィンドウ数設定例

図4の網構成において要求されたスループット値に対し、設定すべきウィンドウ数を表2に示す。各種の数値は上の end-to-end の場合と同じとする。ただし、国内網のリンクでは往復遅延時間が小さく、ウィンドウ数を2~3に設定すると  $9600 \text{ kbit/sec}$  のスループットを保証できるので、これらの国内網のスループット  $T_{e1}, T_{e3}$  は保証されたものとして、中継リンクのウィンドウ数と全体のスループットを求めた。

link-by-link の場合、端末が要求するスループットに応じて、パケット網内の各リンクでここに示すようなウィンドウ数を設定してサービスするということになる。

## 5. おわりに

パケット交換網プロトコル X.25, X.75 の標準化により、回線上の手順については統一されつつあるが、一方、パケット交換機のつくり方および処理方式についてはいろいろな方法が考えられ、フロー制御に関しては問題が多い。

そのうちのスループットとウィンドウ数の関係について、end-to-end, link-by-link のそれぞれに対する考察を行った。

日頃御指導頂く中込所長、鍛冶副所長、寺村次長ならびに関係各位に感謝する。

## 参考文献

- (1) V. Cerf and R. Kahn, "A protocol for packet network intercommunication," IEEE Trans. on Commun., vol. 22, no. 5, pp. 637-648, May 1974.

- (2) L. Pouzin, "Flow control in data networks - models and tools." in Proc. 3rd ICC, pp. 467-474, Aug. 1976.

- (3) M. C. Penzotti and M. Schwartz, "Congestion control in store and forward tandem links," IEEE Trans. Commun., vol. COM-23, no. 12, pp. 1434-1444, Dec. 1975.

- (4) 竹中, 他, "パケット制御方式のシミュレーションによる評価," 信学論(B), J61-B, 6, pp. 327 (PB53-06)