

優先制御待ち行列のクラス別バッファ使用量の推定

01307040 NTT サービスインテグレーション基盤研究所 *土屋利明 TSUCHIYA Toshiaki
NTT ネットワークサービスシステム研究所 宮坂昌宏 MIYASAKA Masahiro

1 はじめに

インターネットのサービスは、従来のベストエフォートサービスから品質保証型のサービスへとシフトしつつある。品質保証の対象は最優先クラスであることが多いが、DiffServ [1] 技術などにより、複数クラスの優先制御が行われる場合、他の優先クラスについてもある程度の保証が必要となる。そこで本検討では、複数クラスの優先制御を実現するノードシステム上での各優先クラスのバッファ使用量を推定する方法について検討する。

2 モデル

本検討では、品質保証複数の優先クラスに対応したノードシステムを、サービス速度 C 、クラス n に対応したバッファ容量を B_n とする非割り込み優先待ち行列によりモデル化 (図 1) する。個々のデータフロー k は、

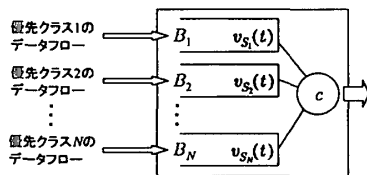


図 1: 優先待ち行列システムモデル

ピークレート P_k 、トークンレート ρ_k 、バーストサイズ σ_k で規定されるトークンバケットモデルに従うものとする。優先クラスが単一の場合、各フローへの割り当て帯域とバッファ使用量の関係が検討されており ([2], [3])、特にバッファ使用量の最大値は

$$\min \left\{ \sum \sigma_k \frac{P_k - c_k}{P_k - \rho_k} \right\}, \quad (1)$$

$$\text{s. t. } \sum c_k \leq C, \quad \rho_k \leq c_k \leq P_k$$

で与えられることが知られている [3]。従って、式 (1) で求められた値とバッファ容量とを比較することで、パケット損失の無い、あるいは極めて低い損失率を保証する受付判定が可能となる。一方、複数の優先クラス

がある場合、最も優先度の高いクラス 1 に対しては式 (1) が適用可能となるが、他のクラスに対しては直接適用できないため、複数の品質クラスに対応するための手順が必要となる。

3 低優先クラスのバッファ使用量

簡単のため、クラス数が 2 の場合のみを考える (ただし、より多くのクラスに対しても、自然な拡張が可能である)。ノードに加わるフロー全体の集合を S とし、優先クラス n のフローからなる集合を S_n ($n = 1, 2$) とする。時刻 t における優先クラス n のバッファ使用量を $v_{S_n}(t)$ 、全フローを単一バッファに収容した仮想システムでのバッファ使用量を $v_{S_1, S_2}(t)$ とおくと、パケット損失がないという条件の元では、

$$v_{S_1, S_2}(t) = v_{S_1}(t) + v_{S_2}(t)$$

という関係が成り立つ。従って、優先クラス 2 の最大バッファ使用量は、この仮想システムの最大バッファ使用量

$$b_{S_1, S_2} = \min \left\{ \sum_{k \in S} \sigma_k \frac{P_k - c_k}{P_k - \rho_k} \right\}, \quad (2)$$

$$\text{s. t. } \sum_{k \in S} c_k \leq C, \quad \rho_k \leq c_k \leq P_k$$

で上から抑えられる。

式 (2) で求められる b_{S_1, S_2} は優先クラス 2 の最大バッファ使用量のタイトな上限となる場合もあるが、優先クラス 1 のバッファ使用率を無視しているため、条件によっては効率が上がらない。そこで、以下に述べる手順を更に加える必要がある。

フロー k が最大速度でデータ転送を行う場合、時刻 0 から $T_k = \sigma_k / (P_k - \rho_k)$ まではピークレート P_k 、それ以降ではトークンレート ρ_k でのデータ転送となる。したがって、時刻 t の転送レート $r_k(t)$ および期間 $(0, t)$ での転送データ量 $R_k(t)$ は

$$r_k(t) = \rho_k + 1_{t \leq T_k} (P_k - \rho_k)$$

$$R_k(t) = \rho_k t + \min\{t, T_k\} (P_k - \rho_k)$$

で表される。優先クラス 1 のフローが時刻 0 から最大速度でデータ転送を行い、優先クラス 1 のバッファに

データが溜まり始めてから当該バッファが空になるまでの時間を $T^{(1)}$ とする. また, T_1, T_2, \dots を長い順に並び替えたものを $T_{(1)}, T_{(2)}, \dots$ とあらわす. このとき,

$$\sum_{k \in S_1} R_k(T_{(1)}) \geq CT_{(1)} \quad (3)$$

ならば, 時刻 $T^{(1)} \geq T_{(1)}$ が成り立ち,

$$T^{(1)} = \frac{\sum_{k \in S_1} \sigma_k}{C - \sum_{k \in S_1} \rho_k} \quad (4)$$

が得られる. それ以外の場合には,

$$\sum_{k \in S_1} R_k(T_{(n)}) < CT_{(n)}, \quad \text{かつ}$$

$$\sum_{k \in S_1} R_k(T_{(n+1)}) \geq CT_{(n+1)}$$

を満たす n に対し,

$$T^{(1)} = \frac{\sum_{k \in S_1} [R_k(T_{(n+1)}) - R_k(T_{(n)})]}{\sum_{k \in S_1} [T_{(n)} R_k(T_{(n+1)}) - T_{(n+1)} R_k(T_{(n)})]}$$

が成り立つ.

今, 優先クラス2のフローが時刻0から最大速度でデータ転送を行ったと仮定すると, 時刻0から時刻 $T^{(1)}$ までの間は全ての帯域が優先クラス1のデータを処理するために使用されているので, 優先クラス2のバッファ使用量 $v_{S_2}(t)$ は増加する. 式(4)で求めた時刻 $T^{(1)}$ における優先クラス2のフロー全体の転送レート

$$r^{(2)}(T^{(1)}) = \sum_{k \in S_2} r_k(T^{(1)})$$

$$= \sum_{k \in S_2} \{ \rho_k + 1_{\{T^{(1)} \leq T_k\}} (P_k - \rho_k) \}$$

が, この時点で優先クラス2が利用可能な帯域より小さい場合, すなわち

$$r^{(2)}(T^{(1)}) \leq C - \sum_{k \in S_1} r_k(T^{(1)}) \quad (5)$$

であれば, 時刻 $T^{(1)}$ 以降 $v_{S_2}(t)$ は減少する. 従って, 優先クラス2のバッファ使用量は $t = T^{(1)}$ の時点で最大となり,

$$b_{S_2} = \sum_{k \in S_2} \max\{P_k T^{(1)}, \sigma_k + \rho_k T^{(1)}\}$$

$$= \sum_{k \in S_2} \{ \rho_k T^{(1)} + (P_k - \rho_k) \min\{T^{(1)}, T_k\} \} \quad (6)$$

で与えられる (図2). 逆に,

$$r^{(2)}(T^{(1)}) > C - \sum_{k \in S_1} r_k(T^{(1)})$$

が成り立つ場合, 時刻 $T^{(1)}$ 以降も $v_{S_2}(t)$ は増加する. この時, $T^{(1)}$ 以降は $v_{S_2}(t) = v_{S_1, S_2}(t)$ であること, $\sum_{k \in S} r_k(t)$ の転送レートが加わった場合, サーバ稼働期間中のバッファ使用量は上に凸となることから, $b_{S_2} = b_{S_1, S_2}$ となる (図3). 式(5)が成り立つ場合は式(6)の値, それ以外ならば式(2)の値が優先クラス2の最大バッファ使用量 b_{S_2} となるので, この値を実装されたバッファ容量 B_2 と比較することで品質管理が可能となる.

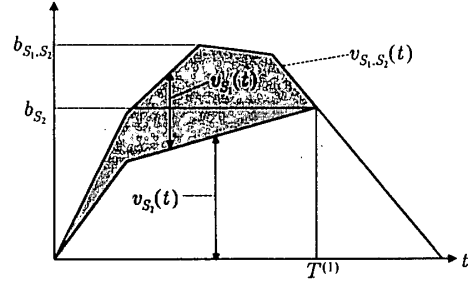


図2: バッファ使用量の推移 (a)

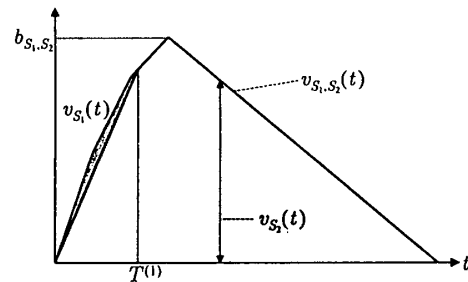


図3: バッファ使用量の推移 (b)

参考文献

- [1] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang and W. Weiss, "An Architecture for Differentiated Service," Internet Engineering Task Force, RFC2475, Dec. 1998.
- [2] A. Elwalid, D. Mitra and R. H. Wentworth, "A New Approach for Allocating Buffers and Bandwidth to Heterogeneous, Regulated Traffic in an ATM node," IEEE Journal on Selected Areas in Communications, vol. 13, no. 6, pp. 1115-1127, Aug. 1995.
- [3] F. L. Preste, Z.-L. Zhang, J. Kurose and D. Towsley, "Source Time Scale and Optimal Buffer/Bandwidth Tradeoff for Heterogeneous Regulated Traffic in a Network Node," IEEE/ACM Transactions on Networking, vol. 7, no. 4, pp. 490-501, Aug. 1999.