

## インターネット・アクセス・トラフィック制御技術 TelePolling

## 3P-3 — アクセス履歴情報を用いた制御の精密化の検討 —

丸山剛一 岸田克己 渡部智樹 田中一男

NTT ヒューマンインタフェース研究所

## 1. はじめに

近年、インターネット上で WWW ブラウザを用いた実時間での参加型イベントが数多く行われるようになった。WWW ブラウザベースでインタラクションを行う際には、サーバ側の CGI プログラム起動などを伴うことも多いため、参加者が非常に多数となる場合には、サーバやネットワークに一時的な負荷集中が発生し、システムの安定動作が困難となる。

このようなイベント開催に際しては、イベントサーバ側で、ネットワーク帯域やサーバ処理能力などのリソースを増強するのみでなく、トラフィックを時間的に平準化し、限られたリソースを有効に活用することが必要である。我々はこのようなイベントにおいて、サーバ上データの定期的な更新（疑似プッシュ）や、参加者からのデータ集約のためのアクセス・トラフィック制御技術 TelePolling [1]を開発し、テレビやラジオなどの放送メディアと同時進行するインターネットイベントにおける実証実験によってその効果を実証してきた[2]。

本稿では、その TelePolling におけるアクセス制御をさらに精密化する手法について検討したものを述べる。

## 2. アクセス・トラフィックの平準化技術 -TelePolling-

まず、TelePolling アクセス・トラフィック制御技術[1]の概要を説明する。TelePolling はアクセス数がサーバの受付能力を超えないように時間的にアクセスを分散させる制御である（図1(b)）。この制御のために、予約数と呼ばれる各参加者端末からのアクセスを受け付けた時点でのサーバ混雑指標に応じて、その端末に対し再接続を行うまでの待機時間  $T_w$  を算出し、指示する。予約数  $R$  は次のように管理される（初期値は0）。

- 1) 端末からのアクセスを検出する度にインクリメント。
- 2) 単位受付間隔  $T_0$  が経過する度に0を下限としてデクリメント

ただし、単位受付間隔  $T_0$  とは、サーバが1アクセスを受け付けるために確保する時間であり、サーバ側の能力に

応じて設定する。例えば1秒間に  $N_0$  個のアクセスを受け付け可能であるサーバでは、余裕を見て  $1/N_0$  より少し大きい値を設定する。イベント主催者サーバ側ではあらかじめ、この  $T_0$  と、端末にとって無理のない基本アクセス周期  $T_a$  を設定し、端末ブラウザからアクセスがあった時点での予約数  $R$  を基に、以下の式で各端末の待機時間  $T_w$  を算出する。

$$T_w = T_a + (T_0 \times R) \quad (1)$$

サーバが混雑している時には予約数  $R$  が増加し、自動的に各端末ブラウザの再接続時間が大きくなるため、サーバ側の能力を越えないように、アクセスの平準化が行われる。イベントによっては、ある一定の周期で情報を更新すれば十分である場合も多く、そのような場合には、必要以上のアクセスを行わないように、イベントで許容できる情報更新アクセス間隔を考慮して基本アクセス間隔  $T_a$  を適度に大きめの値に設定する。

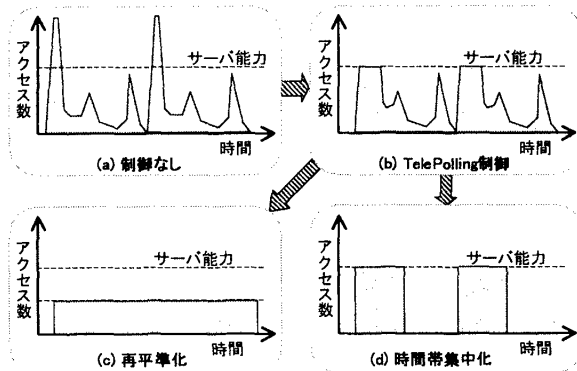


図1: アクセス・トラフィック制御

## 3. 制御の精密化

前節で述べた TelePolling 制御において、基本アクセス間隔  $T_a$  とサーバの単位受付間隔  $T_0$  から求められる1アクセス周期当たりのアクセス可能数  $T_a/T_0$  に比べて実際の参加者数が少ない場合には、さらに次のいずれかの制御が可能となる。

(i) アクセスを再平準化する（図1(c)）

(ii) 特定の時間帯にアクセスを集中させる（図1(d)）

(i)の制御は基本アクセス周期中のアクセスのばらつきをさらに平準化し、サーバの負荷を定常的に低く保てるため、サーバの安定度が増し、他の通信への影響も少なくできる。(ii)の制御では、アクセス時間帯のスケジュー

A Study on More Precise Control Method of Internet Access Traffic using Access Log

Koichi MARUYAMA, Tomoki WATANABE,  
Katsumi KISHIDA, Kazuo TANAKA  
NTT Human Interface Laboratories

ーリングが可能になり、例えばデータの集約と、その集計を繰り返すようなイベントにおいて、集計処理時間を確保できる。また複数イベントにアクセス時間帯を重ならないように割り当てて、同一サーバ上で複数イベントを並列進行させることもできる(図2)。本節では、サーバへのアクセス状況を周期的に把握することによって、これらの制御を自動的に行う手法について説明する。

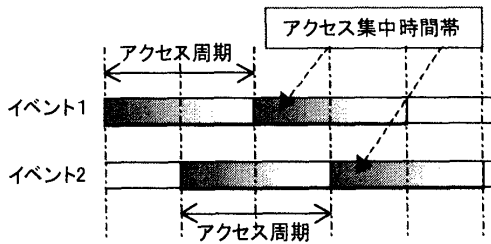


図2: 複数イベントの並列進行例

3-1 アクセスの動的再平準化

アクセスを再平準化するために、各基本アクセス周期ごとにアクセス数  $N$  を計測し、その値にみあったサーバの単位受付間隔  $T_0$  を再設定する。基本アクセス周期内で最も効率よくトラフィック分散が行われるための、単位受付間隔  $T_0$  の値は、

$$T_0 = T_a / N \quad (2)$$

であり、この値を用いて再平準化を行う。ただし  $T_0$  は2節で述べた下限値  $1/N_0$  以上にしか設定できない。この制御によりアクセスが再平準化する例を図3に示す。まず、基本アクセス周期1 ( $T_0 = T_a/4$ ) のアクセス数を計測し、基本アクセス周期の切り替わるタイミングで、その結果 ( $N=2$ ) と(2)式により、新たな単位受付間隔  $T_0$  を設定する ( $T_0 = T_a/2$ )。基本アクセス周期2の区間では各端末に新たな単位受付間隔と(1)式による待機時間の指示が行われ、結果として基本アクセス周期3で所望のアクセスタイミングに分散される。

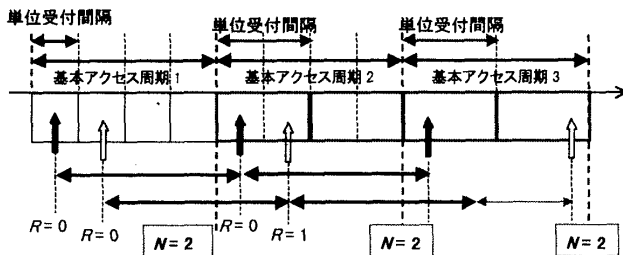


図3: アクセスの再平準化

3-2 アクセスの特定時間帯集中化

参加者端末からのアクセスを、基本アクセス周期毎の早い時間帯に集中させる制御を行う。制御における基本アクセス周期開始のタイミングを調整すれば任意の時間帯にアクセスを集中させることが出来る。この制御を

行うために、新たにサーバ側で、超過数  $Q$  という混雑指標を管理する。超過数  $Q$  の管理は初期値を0として、

- 1) 端末からのアクセスを検出する度にインクリメント
- 2) 単位受付間隔  $T_0$  が経過する度にデクリメント
- 3) 基本アクセス周期を経過したら0に戻す

という手順で行うが、予約数の管理と異なる点はデクリメント時に下限がない(負の数までデクリメントすること、周期的に0に戻すことである。この超過数は基本アクセス周期内のある時点までにおける、実際のアクセス数と受付可能アクセス数との差を表す。したがって、(1)式において、予約数  $R$  の代わりに超過数  $Q$  を用い、

$$T_w = T_a + (T_0 \times Q) \quad (3)$$

によって、再接続時間を計算すると、図4の例のように、各端末のアクセスを順に基本アクセス周期内の早い時間帯側にシフトさせることができる。

例えば、参加者数  $N$  人の場合に、 $T_0 = T_a / (kN)$  と設定し(3)式によるアクセス制御を行なうことで、基本アクセス周期の前半  $1/k$  の時間帯にアクセスを集中させることができる。

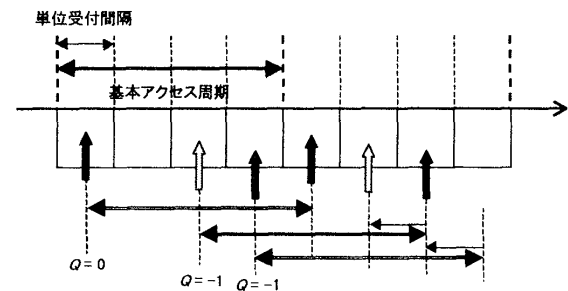


図4: アクセスの特定時間帯集中化制御

4. まとめ

本稿では、TelePolling アクセス・トラフィック制御において、周期的に混雑状況を把握する指標を導入することによって、リソースの利用効率をさらに上げる制御を可能にする方法について述べた。今後は、サーバの単位受付間隔が小さくなった場合に問題となる各クライアントに対するサーバの応答遅延のばらつきを考慮した制御や、予期できない外乱などの影響でサーバが過負荷状態になった場合の検出方法と、その制御方法について検討し、TelePolling アクセス制御のさらなる精密化を図る。

【参考文献】

[1] 酒井, 丸山, 渡部, 岸田, Web における情報更新アクセスの適応的トラフィック制御, 情処学会第56回全国大会 3F-3, 1998  
 [2] 岸田, 酒井, 渡部, 丸山, JoiNet マス・イベント・システムのサービス・アーキテクチャ, 情処学会第56回全国大会 4J-4, 1998