

## Smart Scatter：インターネット／イントラネットの負荷分散機構 －全体構成と制御部－

4 T-1

細井 聡\* 高橋 英一\* 山田 拓也\*\* 高場 浩一\*\* 安達 基光\*  
\* (株) 富士通研究所 \*\* (株) 富士通

### 1. はじめに

近年の Internet/Intranet の急速な普及に伴い、ネットワークを安定かつ効率よく利用することが重要になってきている。特に、WWW などのサーバへのアクセス集中を回避したり、故障を遮蔽したりする環境が求められている。Smart Scatter は、以上のような機能を提供するネットワークミドルウェアである。

本稿では、まず、Smart Scatter とは何か、その特徴と必要性について述べる。次に、Smart Scatter の全体構成を説明し、最後に、その制御部について述べる。

### 2. Smart Scatter とは

#### 2.1 Smart Scatter の必要性と特徴

近年の大規模な Web server は、高容量処理と高信頼化のために複数ノードによるクラスタ構成を取ることが多い。多くの大規模サイトは、round robin DNS でノードにリクエストを分配している。round robin DNS はひとつのホスト名に複数の IP アドレスリストを対応させ、解釈毎に循環的に IP アドレスを選択させる方法である。

しかし DNS は広域ネットワークでの問い合わせトラフィックを抑制するために、ネットワーク上にキャッシュが存在し、ノード故障などのクラスタ構成変更の IP アドレスリストへの伝搬に時間がかかり、ノード故障によるサービス不全状況が起こってしまう。また、負荷分配を IP アドレス単位に行なうため、ノードに性能差がある場合のリクエスト分配の粒度が大きくなってしまふ。

Smart Scatter はクラスタと外部ネットワーク接続するルーターであり、クラスタ全体を表す代表 IP アドレス宛の TCP,UDP/IP データグラムを、クラスタ管理者が定義する割合でノードに分配する。さらに通過データグラムの統計情報とノードの運転状況を監視し、故障時にノードへの分配割り当てを

自動的に調整する。また、WWW サーバの負荷不均衡時には、分配割り当てを自動的に調整する。

#### 2.2 Smart Scatter の全体構成

Smart Scatter は、実際に振り分けを行う「振り分け部」と、「振り分け部」に対しどのような振り分けを行うかの制御を行う「制御部」の2つの部分から構成される（図1）。

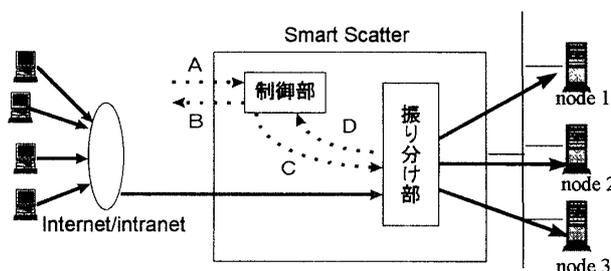


図1 Smart Scatter の構成

「制御部」は、「振り分け部」に対して、各ノードにどのような割合でパケットを振り分けるかの指示を出す（図1の矢印 C）。

「振り分け部」は、各ノードの状況を定期的に監視して、種々の統計情報を収集し、それを「制御部」に伝える（図1の矢印 D）。

「制御部」は、Smart Scatter の状態をユーザに伝える（図1の矢印 B）。また、ユーザからの指示を受け付け（図1の矢印 A）、必要に応じて、「振り分け部」に対して指示を出す（図1の矢印 C）。

（なお、「振り分け部」の詳細は、4T-02 で述べる。）

### 3. Smart Scatter 制御部分について

「制御部」は、「振り分け部」に対して、振り分け方法を指示し、ノード間の負荷分散や、故障時に他ノードへの割り当て変更を行う。「制御部」が提供する機能には以下の3つがある。

#### 3.1 負荷分散

Smart Scatter では、入力パケットをハッシュして、複数ノード間に振り分けることにより、負荷分散を行う。

Smart Scatter

Akira Hosoi

Fujitsu Laboratories, 4-1-1 Kamiodanaka,  
Nakahara, Kawasaki, Japan

具体的には、まず、クライアント IP アドレスやサービスポート番号などをキーとしたハッシュ関数により、パケットをある範囲の値（たとえば、0～255）にマッピングさせる。そして、次に、それぞれの値をどれかのノードに割り当てる。

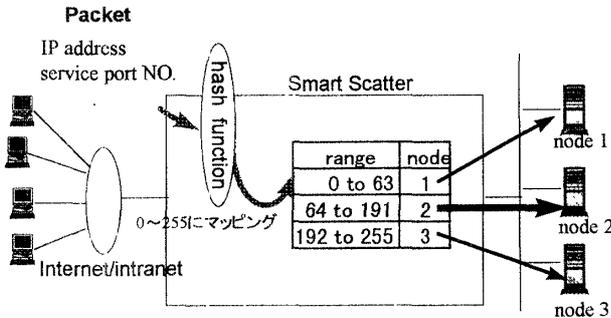


図2 Smart Scatter による負荷分散

図2は、0～63をノード1、64～191をノード2、192～255をノード3に割り当て、3つのノードの負荷を1:2:1の割合に分散させた場合を示している。このようにして、Smart Scatterでは負荷分散を行う。

### 3.2 WWW サーバの自動調整

「振り分け部」は、定期的に各ノードの性能情報を収集している。（この性能情報の収集サイクルを「監視サイクル」と呼ぶ。）この情報を元に、「制御部」は、負荷分散の自動調整を行う。たとえば、図2において、ノード2が高負荷であると判断した場合、64～68をノード1へ、69～73をノード3へ割り当てるように変更（ノード2自身は、74～191の範囲の値に対して割り当てられるように変更）することにより、ノード2の負荷を軽減させる。次の監視サイクル後もやはり高負荷であったなら、同様の操作を行う。このように、ハッシュ値に対するノードの割り当てを動的に変更することにより、負荷の不均衡を自動調整する。

Smart Scatter では、一定の監視サイクルごとに、各ノードの response 時間（VCの接続から終了までの時間）、connect 時間（SYN～SYN/ACKに要する時間）、同時 active VC 数、connection rate（監視サイクルでの fin の通った総数を計測し、1秒間に fin の通過する数に換算した値）、connect 再送回数（VCの接続要求が再送される回数）などの性能情報を収集する。これらのうちのどれか1つまたは、複数を組み合わせて、高負荷判定を行う。デフォルトでは、connect 時間と connection rate を組み合わせた指標を元に、高負荷かどうかの判定を行う。

### 3.3 ノード故障対応

Smart Scatter は、ノードの故障が起こった場合、故障ノードに割り当てられていた分を他のノードに割り振ることにより、故障を Smart Scatter の外部に対して隠蔽させることができる。

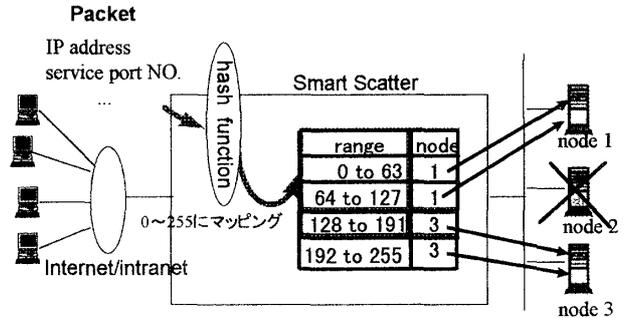


図3 ノード故障の場合の割り当て変更

図3は、ノード2が故障し、ノード2の割り当て分を、ノード1と3に割り当てるように変更した場合である。これにより、ノード2が行っていた処理を他ノードが引き継ぐ。（これはちょうど、自動調整で、ノード2の割り当てを0にしたことと同様である。）

#### 「故障からの復帰について」

「制御部」は、どのノードも故障していない場合の、最新のノードの割り当て表を常にファイルに保持している。よって、ノードが復旧した場合には、このファイルを元に、ノード故障が起きる直前の割り当て状態に復帰させることができる。

## 4. まとめ

Smart Scatter の特徴、必要性、全体構成およびその制御部について解説した。

今後、さらなる評価を行っていく予定である。