

大規模分散システムにおけるリソースサーバ構成法

2P-7

中村 元紀

NTT 光ネットワークシステム研究所

1 はじめに

大規模分散システムにおいて分散リソースの捕捉を高速に行うための、リソースサーバの構成法について述べる。分散された複数のリソースをそれぞれ排他的な部分集合(以下部分リソース)に分割し、複数のリソースサーバ(Resources Managing Agent)(以下 RMA)が各々全てのリソースについて排他的に部分リソースを管理することにより、リソースを高速に捕捉する分散リソース管理方式、HDRM(Horizontally Divided Resources Management)を既に提案している[2]。HDRMでは、RMAが管理するリソースに関する情報(以下管理権)を排他的に分割するため、分割損により可用性が低下する可能性がある。本稿では可用性の低下の防止に関して、RMAでの対処の必要性とその対処法について述べる。特にHDRMの特徴である高速なリソースの捕捉を維持しつつ可用性を高めるためのRMAの再捕捉方式について述べる。

2 分散リソース管理方式

2.1 対象とする分散リソース

HDRMの管理対象は次のような性質を持つリソースとする。(i)日本全国のように広範囲に分散されたノードに存在し、(ii)ノードに特有で移動が不可能であり、(iii)部分リソースに分割して使用され、(iv)リソースの利用者(以下クライアントと呼ぶ)が特定のいくつかのノードの部分リソースを同時に使用する。

特に本稿では、将来の適応型通信サービス[1]で上記の性質を持つ通信リソースを利用する場合を想定する。適応型通信サービスでは、ユーザの置かれた環境に応じて最適なサービスを提供するため、ユーザが利用する可能性のあるリソースを全て捕捉し、その後ユーザが選んだ環境を構築するのに必要なリソースのみを有効化する。

2.2 HDRMの基本方式

筆者は2.1で述べた性質を持つ分散リソースの管理の際に生じる捕捉遅延を削減する方式HDRMを提案してきた[2]。HDRMでは1節で述べたように各RMAが全てのリソースについて、それらの部分リソースを排他的に管理する。すなわち、あるRMAが管理する部分リ

ソースは他のRMAが捕捉することはできない。これにより、各RMAは複数の部分リソースの捕捉・解放をそのRMAのローカルな情報のみで実行可能となり、管理のためのRMA間通信が不要となるのでリソースの捕捉遅延の削減が可能となる。

3 分散リソースの可用性

リソースを排他的に分割して管理することにより、HDRMでは分割損が生じる可能性がある。例えば、あるRMA(R1とする)へ捕捉要求が集中して、R1の管理する空きリソース量が少なくなると、他のRMAでは捕捉可能なリソースがR1では捕捉不可能という状況が起こり得るため、リソースの可用性が低下する。

これを避けるための方式については、クライアントで対処する方式とRMAで対処する方式が考えられる。本稿ではそれぞれのHDRMへの適用可能性について述べる。

3.1 クライアントでの対処

クライアントでの対処により可用性を高める方式としては、大きく分けて以下の二つが考えられる。

- (α) 捕捉要求を全てのRMAへブロードキャストする。
- (β) 一つのRMAへ要求して捕捉失敗の通知を受けたクライアントは、別のRMAへ再度要求する。これを成功するまで順次繰り返す。

HDRMの適応対象である大規模システムにおいて(α)は全てのクライアントが捕捉要求のたびにブロードキャストを行うので通信量の増大を招く。(β)では最初のRMAで捕捉が失敗した場合、順次別のRMAへ捕捉要求を行うため、捕捉が成功するまでの処理遅延が増大する。また、リソースの捕捉は各クライアントが共通的に利用する一種のサービスであるので、クライアントの側で複雑な処理を実装するのは好ましくない。よってHDRMでは、クライアントで対処するのは適していない。

3.2 RMAでの対処

RMAでの対処としては、RMA間の協調制御によって対処する方式と、捕捉要求を受けたRMA単体で対処する方式が考えられる。

3.2.1 RMA間の協調制御による対処

RMA間の協調制御による対処として、あるRMAがリソースの捕捉要求を受けても捕捉不可能な場合に、そ

のRMAが3.1の(α)や(β)と同様に他のRMAへの捕捉要求を行う方式が考えられる。しかし大規模システムでは転送遅延が無視できない上に、システム全体でリソースが不足する際には要求の転送が氾濫する可能性があることなどを考えるとHDRMでは適していない。

そこで空きリソースが少なくなった場合、RMA間でリソースの管理権の譲渡を行うものとする[2]。すなわち、RMAは次の捕捉要求を処理するのに十分な空きリソースがないと予測した場合、事前に他のRMAから空きリソースの管理権を譲り受ける。また、譲り受けられない場合はリソースの捕捉サービスを停止し、空きリソースの管理権を他のRMAへ譲渡することにより、システム全体でのリソースの不足に対応する。このようにHDRMでは、リソース捕捉の成功確率(リソースの可用性)の向上のためにRMAが協調制御を行う。

3.2.2 RMA単体での対処

2.1で述べたように、適応型通信サービスでは、捕捉されたリソースが比較的短時間で解放されるケースが増える。よって捕捉に失敗しても、待機後に再捕捉を行うことにより、RMAが協調を行わなくても単体で捕捉失敗数を削減することが可能となる場合が増える。

捕捉要求の集中や、クライアントからの大量のリソースの捕捉要求などにより捕捉が成功しない場合、3.2.1で述べた協調制御だけでは対処できない。よってRMA間の協調による対処だけでなく、RMA単体での再捕捉による対処も重要となる。

3.2.3 分散リソースの再捕捉方式

前述のような再捕捉の処理が他の要求の処理遅延に影響を与えて、HDRMの特徴である捕捉遅延の削減効果を弱めることは好ましくない。ここでは捕捉遅延の削減を維持しつつ可用性を高めるための再捕捉方式について述べる。

RMAでの再捕捉方式については、再捕捉前に他の要求を処理するか、また再捕捉の契機の設定方法により、以下のような方式が考えられる(図1)。

- (1) 他の要求は処理せずに一定時間後再捕捉を試みる。
- (2) 他の要求の処理を続行し、一定時間後に割り込みをかけて再捕捉を試みる。
- (3) 捕捉が失敗した要求を処理待ちキューの後方に挿入することにより、他の要求をいくつか処理した後再捕捉を試みる。

単一リソース管理の場合であって、クライアントが常に一定量のリソースを要求するのであれば、(1)は要求を到着順に処理すべき場合には適している。しかし、HDRMでは複数ノードのリソースを組み合わせた捕捉要求を受けるため、(1)では本来捕捉が失敗した要求とは

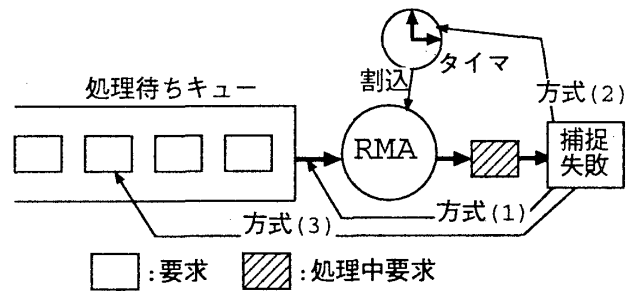


図1: 再捕捉方式

独立に捕捉可能であった要求の処理が待たされることとなり、全体として可用性が低下する可能性がある。よって(1)は適していない。

(2)は、一定時間を適当な値に設定すれば、再捕捉処理待ちによる遅延を抑えつつ、再捕捉が失敗する確率もある程度に抑えることが可能となる。しかし厳密に実行するために割り込みなど複雑な処理が必要となる。このような処理は本来RMAが使用可能なプロセッサパワーを消費し、他の要求の処理時間に影響を与えることとなる。HDRMでは高速なリソースの捕捉のためにRMAの処理を軽量化することが重要であるため、また最適な値を決定するのも難しいため、(2)は適していない。

一方(3)はHDRMにおいてRMAが受ける複数ノードのリソースを組み合わせた捕捉要求に対しても、捕捉が失敗した要求が他の要求の処理遅延に対して影響を与えにくく、特にキューの最後に挿入するようにすれば、その時点で到着していた要求の処理遅延にはほとんど影響を与えない。またRMAの処理も失敗した要求をキューに挿入するだけであり、軽量化が可能である。よってHDRMにおいては方式(3)が適している。

4 まとめ

分散リソース管理方式HDRMについて、リソースの可用性低下の防止に関して、RMA間の協調及びRMA単体での対処の必要性を示し、また一旦捕捉に失敗した要求を、RMA内で再捕捉を試みることによって可用性を高める方式について述べた。今後は今回示した方式により、捕捉の遅延を抑えつつ可用性を高めることが可能なことを、シミュレーションにより定量的に示す予定である。

参考文献

- [1] ITAO, T. AND MATSUO, M.: DANSE: Dynamically Adaptive Networking Service Environment, *Globecom 98*(Nov. 1998).
- [2] 中村元紀: 分散オブジェクト指向ネットワークアーキテクチャにおけるペアラリソースの制御方法, *信学技法 SSE96-177*(Feb. 1997), 89-94.