

4J-7 分散データベース用ファイル管理システム

堀口真志、小田正美、田中英彦、元岡達
 (東大・工) (富士通) (東大・工) (東大・工)

1. はじめに

計算機網の目的の一つであるデータの共有を実現するための分散ファイルシステムの一方式を提案する。これは、ユーザが網を意識することなく、網の上のどのファイルをもアクセスできることを目的としたものであり、研究用計算機網TECNETに実装するために設計されたものである。TECNETの網向きOS(NOS)は、プロセス同士が網を意識することなく情報を交換する手段としてプロセス間通信を提供していることを、その特徴としている。ここでは、NOSの上に実装予定の分散ファイルシステムの構成を述べ、分散データベースを実現するために工夫した点について触れる。

2. 分散ファイルシステムの構成

この分散ファイルシステムの構成を図1に示す。

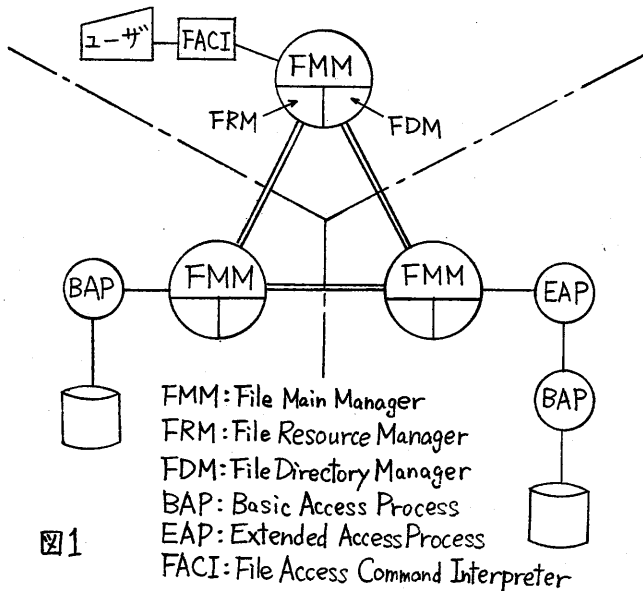
FMMは各ホストに1つあるプロセスであり、そのホストにあるファイルを一括管理し、また他のホストのFM

Mと通信することによって、網全体としてのファイル管理が行われるようになっていた。FMMの下にはFRM(ファイルの確保・解放・生成・削除を行う)、FDM(ファイルのディレクトリの管理を行う)の2個のサブシステムがあって、FMMの依頼によってそれぞれの仕事を行う。

BAPは、ファイルの編成ごとに設けられ、ユーザプロセスと通信を行ってその要求に従って実際にファイルにアクセスするプロセスである。

BAPは、ファイルのRead、Write等の基本的な機能のみを行うものであるが、これだけでは通信のオーバーヘッドが大きい場合、EAPと呼ばれるプロセスに拡張コマンドを送ることもできる。EAPが処理する拡張コマンドとしては、ファイルの複写・比較の他に、データの検索・更新等を用意して、分散データベースシステムの便宜を図る。

FACIは、ユーザのためのファイルアクセスルーチンである。



2. ディレクトリ

この分散ファイルシステムでは、ファイルディレクトリは、検索の効率と信頼性を考慮して、図2に示すような二階層テーブル方式と呼ばれるものを採用した。

UFNは、ファイルのユーザがつけた名前と、ファイルのオーナーがつけた名前とのマッピングを行うための表であり、各ユーザごとに設けられ、使用するファイルごとに1エントリある。

OFNは、ファイルの格納されている装置、その他ファイルに関する詳細

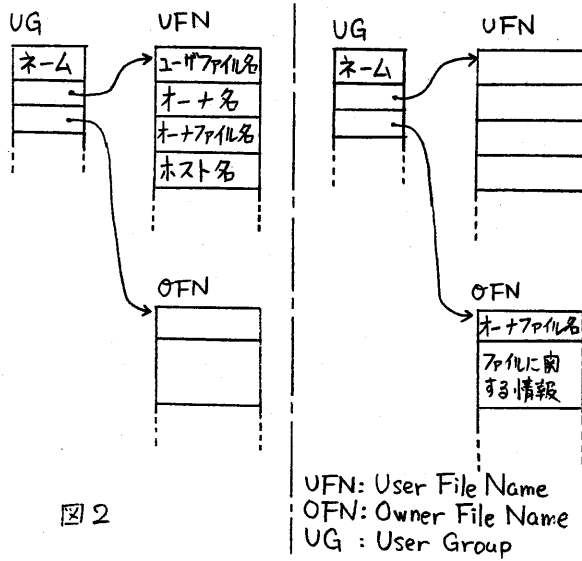
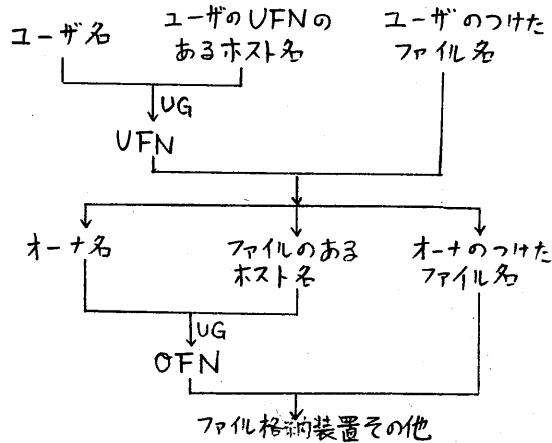


図 2

な情報が記述されているものであって、各オーナー、各ホストごとに設けられ、所有しているファイルごとに1エントリある。UGはUFN、OFNの検索の効率を考慮して設けられたものであり、各ホストに1個ある。

この方式で希望するファイルを探すには以下のような手続きによって行う。



4. ファイル確保の手順

この分散ファイルシステムでは、デッドロック対策としては、実装の容易さからデッドロック回避法のうちの単純法を採用した。そのため、デッドロック対策の必要なファイルはすべて、ジョブステップの最初で確保しておくことになるが、その手順は以下のとおりである。

(1) ユーザ (実際はジョブ管理システム

が代行)が、必要なファイルのリストを作り、自分のホストのFMMに提出する

(2) FMMは、各ファイル名をオーナーのつけた名前に変換するために、必要があれば各ホストのFMM (FDMと通信する。

(3) すべてのファイル名の変換が終了すると、そのファイル名のリストを各ホストのFMMに回覧させる。各FMMは、そのリストのうち、自分のホストにあるファイルを確認する。1つでも確保できないものがある場合は直ちにそのことを要求元のFMMに通知する

(4) すべてのファイルの確保が終わるとFMMはユーザにファイル確保のレスポンスを送る。

5. ファイルアクセスの手順

確保されたファイルをユーザがアクセスするには、次のような手順で行う

(1) ユーザがOPENコマンドを発行すると、システムは、それをファイルの存在するホストのFMMに送る。

(2) FMMはBAPを起動し、ファイルのOPENを指示する。

(3) BAPは、ファイルをOPENし、FMMにレスポンスを返す。

(4) FMMは、ユーザにOPENコマンドのレスポンスを返す。

(5) 以後は、ユーザとBAPとの通信によってファイルのRead, Writeが行われる。

(6) ファイルのCLOSEは、ユーザがFMMにCLOSEコマンドを送ることによって行われる。以下の手順はOPENの場合とほぼ同様である。