

Network Oriented Operating System

Hedehiko TANAKA

Department of Electrical Engineering

An operating system for computer network is designed and some basic functions which the operating system should provide are examined in detail. This network oriented OS is implemented in the experimental computer network TECNET. This system is based on the network-wide inter-process communication facility, and have the network-wide job management system feature and network-wide file management system feature. Each user job is managed by a master-job-manager-process and several local-job-manager-processes which reside on each computer. User can access to any file which is placed at any computer system, through the aid of network-wide file system. The size of OS code is 9kw for kernel part, 6.3 kw for job management, and 5.7 kw for file management each.

1. はじめに

計算機を複数、通信回線等によって結合し相互に利用しあうシステム—計算機網—は今後の情報処理形態として有望であるが、そのようなシステムを有効に働かせる為のオペレーティングシステムの構造については余り研究が為されていず、特に網環境を意識した OS 構造については全く未開拓の分野である。本研究はそのような構造を探るとともに、実験用計算機網の上にそのような OS を実装し評価することを目的としている。この OS ではコンピュータ内の活動はプロセス群からなるとしており、それらプロセスが動く環境を作っているのが核部分と呼ぶ OS の中心部である。ここではこの OS の構造を述べるとともに網環境下でのジョブ管理システムとファイル管理システムについての検討結果を報告する。

2. NOS の構造

NOS はプロセスとシステム核よりなる。プロセスは計算機内におけるプログラム実行の基本単位であり、システム核はそれらプロセスが動く環境を作りあげる部分で次のような仕事を受取っている。

- ・プロセスのスケジューリングとディスパッチング
- ・割込処理
- ・基本操作 (Primitive Operations)

基本操作はプロセスが実行可能なシステム動作で、プロセスの生成/消滅、実行開始/停止、条件待ち状態セット、及びプロセス間通信などがあり、それぞれに対し割込禁

止モードで動くサブルーチンが用意されている。また、プロセスにはシステムプロセスとユーザプロセスとがあり、ユーザプロセスはユーザの仕事を実際に実行するプロセスでありシステムプロセスはその為の OS 動作を実行する。すなわち、ユーザプロセスには、言語プロセッサ、ユーティリティプログラム実行プロセス、ユーザプログラムの実行プロセスなどがあり、システムプロセスには、各入出力装置制御プロセス、ファイル管理プロセス、ジョブ管理プロセス、ロガープロセス、資源管理プロセス等がある。

3. プロセス間通信

NOS のプロセス間通信機構は相手プロセスが存在する場所に依らず容易に他のプロセスと通信できる形になっており、Send, Receive, Sense という3つの基本操作が用意されている。メッセージの授受は双方のプロセスが互いに相手を指定して Send と Receive を出し合った時点で始まり、送信側プロセスの用意したメッセージが、受信側プロセスが用意したバッファ内に転送される。転送が完了すると完了コードが双方のプロセスの状態通知語 (IW) 内にセットされ何時でも Sense 可能となる。この制御は相手プロセスがローカルカリモートかには依らない。即ち、通信プリミティブが出されると通信状態語 (CSW) が作られその通信のパラメータが書込まれるが、リモート通信の場合は相手プロセスが存在する計算機へ問合せのメッセージが送られ相手計算機内にも CSW が作られる。相手プロセスがその後、それに対応した通信プリミティブを出すと制御情報が元の計算機に戻され CSW を埋める。この時点で通信契約が成立して実際のメッセージ転送がおこなわれ、完了すると完了コードが IW にセットされるとともに CSW は消去される。

このような機構の問題は相手プロセスの指定法とデッドロック対策である。相手プロセスの指定法に関しては現在の所、各計算機番号とその中でのプロセス番号との組み合わせを用いているが、これを、インデックスを用いた表引きによって間接的におこなう方法を検討中である。これによって相手を仮想的なネームと呼ぶこともできるし又、表内にプロセス番号を記入する時にチェックをすることで保護対策ともなりうる。又、デッドロックは互いに相手との通信を期待して双方が送信プリミティブを出して待ちの状態に入る場合に生ずる問題で、これに対しては幾つかの解決策が考えられる。1つは、相手からメッセージを受信する可能性がある場合は自分が待ち状態に入る前に Receive を出しておく方法で、2つ目は通信事象が生じれば常に該当プロセスに起動をかける方

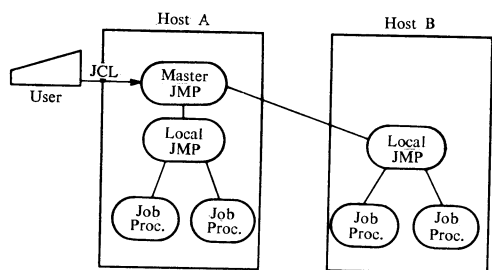


Fig. 1. Network-wide job management system.

法, 3つ目は常時, 待ち状態に入る時はタイマをかけて入る方法である。現在はこれらを適当に併用しているが, 第2版用としては通信事象管理を強化する方向で検討中である。

4. ジョブ管理

NOS における1つのジョブは各所に分散したプロセス群からなる。これを管理する為以下のような網ジョブ管理機構を設けた。まずプロセスが存在する各計算機内に1つずつローカルジョブ管理プロセス (LJMP) を置きその計算機内での統一したジョブ管理をおこなう。更にこれら LJMP を統合する為網内のどこかにマスタとなる主ジョブ管理プロセス (MJMP) を1つ置く。このような管理機構がジョブ毎に作られ, ジョブ制御のコマンドはユーザから MJMP に渡され, MJMP はそれを解釈して適当な LJMP に指令を出す。LJMP は MJMP からの指令に基づいてユーザプロセスを生成したり消去したり, 起動, 停止などの制御をおこなう。この場合 MJMP-LJMP 間のメッセージは形式が網内で統一的に定められており網プロトコルになっているが, LJMP とユーザプロセスとの間はローカルな問題である。ユーザから MJMP に渡されるコマンドはジョブ制御言語 (JCL) の形を取るが, それには次のようなものが用意されている。

- ・指定プログラムのロード
- ・ロードされたコードをプロセスとして生成/消去
- ・指定順序でのプロセスの実行
- ・資源要求/解放
- ・プロセス状態のセンス

5. ファイル管理

NOS が実装されている各計算機は, ディスク, カセット磁気テープなどの二次記憶装置の他に様々な入出力装置をも備えている。これらの装置を各計算機から自由に利用できることを目的とした制御システムが網ファイルシステムである。即ち, 通常のランダムアクセスファイルの他に入出力装置をもそれぞれ1つのファイルみと

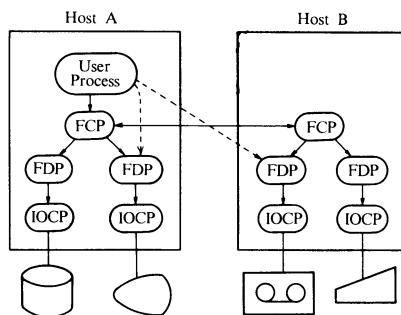


Fig. 2. Distributed file management system.

なして利用規約の統一をおこなっており, 物理的な設置場所に依らぬ一様なアクセスを可能としている。その機構は各入出力装置や二次記憶等の物理的ファイル要素に対して入出力制御プロセスの上にファイルデバイスプロセス (FDP) を置き, 網上ではそれら FDP が仮想的なファイル制御プロセスとなる。又, 各計算機内にはローカルな FDP を統轄するファイル制御プロセス (FCP) を1つ置き計算機間にまたがったファイルアクセスを制御する。即ち, ユーザからのアクセス要求はまずローカルな FCP へ送られ, そこでリソースの位置を決定し, ローカルならばその所要 FDP へアクセス要求が渡されるが, リモートの場合はリモート計算機の FCP へアクセス要求が送られそこから目的とする FDP へとアクセス要求が送られる。要求に対する返事は宛先 FDP からユーザに戻され, 要求が満たされている場合はアクセス名が返される。従ってそれ以降のファイル利用上のコマンドや応答, データ等はユーザプロセスと目的 FDP との間で直接プロセス間通信により授受される。この機構により各要求ファイルの設置場所はユーザにとって透明となる。

6. システムの実装

現在研究室内には複数のコンピュータを回線で結合した実験用計算機網 TECNET が稼動しており, その中の3システムに NOS が実装されている。即ち, 中型ポリプロセッサシステム PPS-1, ミニコンピュータ OKITAC 4300C, FACOM-R である。これらの3システムにはリモートプロセス間通信フィーチャを備えたシステム核が存在し, その上に IOCP, FDP が実装されている。網ジョブ管理システムと網ファイルシステムとは PPS-1 と OKITAC 上に実装されている。OS コードの量は計算機にもよるが, たとえば OKITAC の場合, システム核が 9.5 kW (16ビット/語), ジョブ管理 6.3 kW, ファイル管理 5.7 kW となっている。今後はこのシステムを様々な形で利用しその経験を通して評価をおこなう予定である。