第6章　QNXを用いた分散環境の構築

100CPUを用いた大規模マルチプロセッシング・システムの構築

岡澤 幸一

8CPU程度のマルチプロセッサであれば，従来のマルチタスク技術の延長でシステムを構築できた．しかし，それより大規模な，例えば100CPUなどのシステムになると，通常の構成とは異なるアプローチが必要になる．本稿では，マイクロカーネル・アーキテクチャを採用したOS「QNX」を例に，分散環境とは何かについて解説する．また，大規模分散環境の構築事例も紹介する．（編集部）

CPUのクロック速度の向上が限界に達しつつあるにもかかわらず，ユーザからは高性能化や低消費電力化の要求が高まっています．それに対する処方せんとして，ハードウェア・ベンダにより提供されている技術がマルチコアであると筆者は考えます．

マルチコア化したシステムに対して，従来のアプリケーションがそのまま動作することが要求されます．また，マルチコアにより動作するCPUコンテキストが複数あるため，同期制御や排他制御というクリティカルな問題が浮上してきます．マルチコア化とは，ソフトウェアの課題と言っても過言ではありません．

本稿では，マルチコアを始めとした分散システムの分類について解説します．ハードウェア・システムは，それぞれのCPUが独立したメモリとI/O資源を持ち，高速通信路で接続された並列プロセッサを用いた構成とします．ターゲットとなるプロセッサの数は数十個～数百個を想定しています．

1．並列と分散

そもそも「並列」と「分散」という用語は同じでしょうか．普段は何とはなしに使っているかもしれません．まず，背景などを整理してみます．

簡単に二つの言葉が含まれているものを挙げてみます．

「並列」コンピュータ，「並列」処理，超「並列」，…

「分散」ファイル・システム，「分散」システム，…

これらの例から言うと，走行するハードウェア側から，複数の環境で同時に走行することに対して語られる言葉が「並列」であるといえます．一方，ソフトウェアやシステムの立場から実際のソフトウェア・コードを「分散」する方法や，その結果できあがるシステムに対して「分散」システムと名付けられます．

つまり「並列」化したハードウェア・システム環境で，「分散」化したシステムを実現すること，これが本稿のテーマとなります．

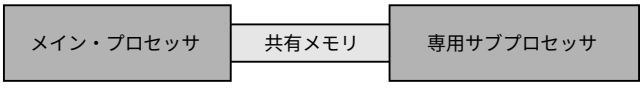
2．ハードウェアの視点――「並列」

ソフトウェアの観点から疑似「並列」を実現するOSについて考えてみる必要がありますが，ここではまず，実際に並列性を実現するためのハードウェア・システムの技法から見ていきます．

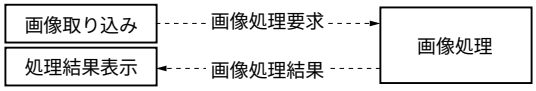
CPUを複数使い，要求される処理を分割してシステムとして構築する手法は，要求される処理の分割のタイプによって分類されます．

● 専用サブプロセッサ（ジョブ実行型）

専用サブプロセッサは，データ処理の視点が中心となるアーキテクチャといえます．画像処理や信号処理などを行



（a）ハードウェア構成



（b）ソフトウェア構成

図1　専用サブプロセッサの構成

KEYWORD――並列，分散，アクタ，オブジェクト指向，メッセージ・パッシング，QNX，マイクロカーネル

う場合，データ処理の一部である数値計算の部分を別の専用サブプロセッサで実行するようなシステムを構築します．専用サブプロセッサは，切り出された数値計算（例えば画像の特徴を抽出したり，画像の微分を取るなどのデータ処理）を行うハードウェアとなります．

メインのプロセッサからはコマンドとして処理を起動し，結果を受け取ります．一般には，**図1**のように専用サブプロセッサとメイン・プロセッサの間は共有メモリで結合され，データそのものは共有メモリを介して受け渡されます．与えられた処理を行う処理要求元（メイン・プロセッサ）から処理先（専用サブプロセッサ）へ下請けに出すタイプの構成となります．

二つのプロセッサを接続する方法としては，例えばメイン・プロセッサ上のバス（VME，PCI，CompactPCIなど）に接続するサブボードが利用されます．このボードに専用サブプロセッサを搭載し，CPUからアクセスできるバス空間上に共有メモリ領域を置き，専用サブプロセッサを制御します．

● クラスタリング（並列計算）

クラスタリングは，複数のコンピュータ・システム（プロセッサとOSを持つ）をネットワークで接続して並列コンピューティング環境を作成する手法です．低コストなコンピュータおよびネットワーク・ハードウェアにより実現されます．

この場合のハードウェア構成を**図2**に示します．

特徴としては，並列計算に特化していることです．行列演算を分割して各コンピュータに処理を振り分け，最終的に結果を集めて全体の行列演算の結果を得るといった使い方が考えられます．

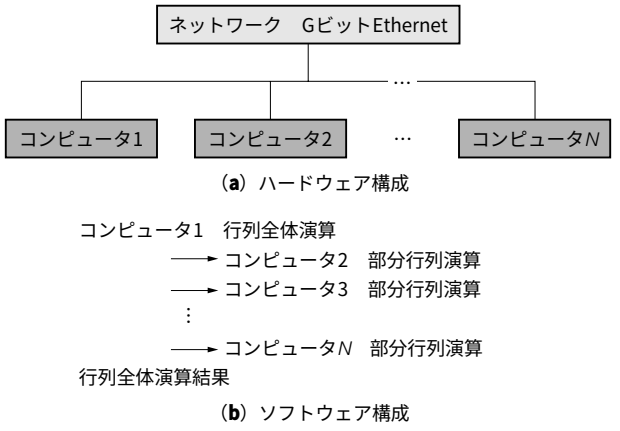


図2　クラスタリングの構成

この流れでは，現在グリッド・コンピューティングとして，数百，数千のコンピュータに並列計算させる技術が登場しています．低コストのパソコン・ベースのコンピュータをGビットEthernetで接続することにより，誰でも利用できる技術になっています．また，ネットワークのスイッチ技術を利用して，独立したルート（回線）で接続することが可能となっています．

● CPUセル形式の並列コンピューティング

CPUそのものにメモリを持たせ，CPU間をハードウェアで接続したセル形式の並列コンピュータが存在します．これは「セル・コンピュータ」と呼ばれています．

1980年代に生まれた「トランスピュータ」がその一例です．商業的には成功しませんでしたが，その発想には目を見張るものがありました．

トランスピュータの構成を**図3**に示します．ハードウェア的にはCPU間をセル形式で接続します．また，ソフトウェアの視点から見ると，仮想チャネルで接続されたマルチプロセスの構成となります．「プロセスがどこのCPUにいるか」，それは割り付けの問題となります．マルチプロセスが仮想チャネルを介してプロセス間通信を行いながら全体のアプリケーションを構成します．トランスピュータの場合，それと対となる並列処理記述言語OCCAMが作られ，コンパイラが用意されました．

トランスピュータのCPUは四つの高速通信チャネルを持ち，4方向に接続していくことで，大規模な並列コン

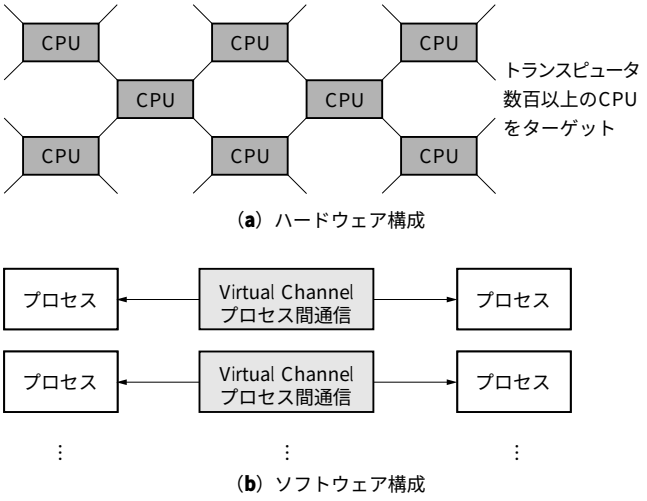


図3　トランスピュータの構成