

国立国語研究所学術情報リポジトリ

Discribed language data processing

メタデータ	言語: jpn 出版者: 公開日: 2017-06-13 キーワード (Ja): キーワード (En): 作成者: 斎藤, 秀紀, SAITO, Hidenori メールアドレス: 所属:
URL	https://doi.org/10.15084/00001305

分散処理システムへの試み

齋藤 秀紀

1. はじめに

国立国語研究所にコンピュータが導入され、14年が経過した。現在使用中の装置は昭和49年度に第2回のリプレース機種として設置され、以後バッチ処理による用語用字調査など各種の研究調査に使用されてきた。しかし、この数年間、利用者の増加と多様化により、装置能力に限界が生じ、より大型装置の導入が要求されるようになった。現在、3回目のリプレースとして新機種の導入計画が進められているが、本稿は、これらの装置に関する基本的な考え方について述べたものである。新機種 HITAC—M150 は、中央処理装置、磁気ディスクなどの記憶容量の大幅増設また各種ターミナル装置として L—320 インテリジェント・ターミナル、漢字ディスプレイ装置など端末系の充実が図られており、特に L—320 についてはオフィス・コンピュータによるインハウス・オンライン形式での分散処理を指向したものである。

新システムでのフロッピーディスクの採用は、従来の紙テープ、カードに対して高記録密度反復利用可能媒体であるため、保管スペースの縮小、データ管理の容易性、消耗品の無駄を防ぐことが可能となる。

また、L—320 とフロッピーディスクの結合により、データエントリ修正の効率化、ローカル・マスタファイル方式の採用、ターンアラウンド時間の短縮、L—320 によるフロッピーディスク専用処理システムなど従来のバッチ処理方式に対し多くの機能が追加される。これらの諸機能は今後本格的分散処理導入のための足がかりとなるものであり、組織内での適用業務の把握と運用面での種々の確認が必要となろう。

2. 現行装置の問題点

現行の HITAC—8250の稼働時間は、月平均使用時間で160時間を越えており、特にピーク月では 200 時間まで増加する。これは処理時間、勤務体制またコンピュータ処理能力、いずれも限界に達していると見てよい状態である。しかし、今後コンピュータ利用者の増加と多様化する利用者に対処するためには、既存コンピュータに対する機能アップ、また新機種導入など、いずれかの対策が必要となる。

以下、現行コンピュータの問題点について概要を述べ、新機種に対する基本的な考え方を示す。なお、現在使用中のコンピュータ機器構成は表 1 の通りである。

表 1 HITAC—8250 機器一覧

内部記憶装置	98K B	2 領域
磁気ディスク装置	29M B	2 台
磁気テープ装置	40K B	6 台
ラインプリンタ装置	1250行/分	1 台
カード読み取り装置	1000枚/分	1 台
カードさん孔装置	160字/秒	1 台
紙テープ読み取り装置	500字/秒	1 台
紙テープさん孔装置	110字/秒	1 台

上記の性能一覧に示した通り、装置規模は小型に属するものである。総内部記憶容量は98K Bであるが、OS部に28K B、またユーザ使用領域に40K B、30K Bと2領域にそれぞれ分割配当している。処理傾向としてはファイル処理を主体とした入出力処理が多いが、利用効率を現状以上に上げるためには、マルチ処理で並行して処理できるジョブの本数を増加させることが必要である。しかし、現状規模のまま単純に多重度を上げることには問題が多く、コンピュータ性能がプログラム処理に十分対応できなければマルチ処理の効果は期待できないことになる。つまり、マルチ処理の多重度を上げるためには、内部処理速度と十分な記憶容量が用意されていなければ、領域当りに配当される記憶容量が減少し、逆にプログラムサイズに制限を受けることになる。また、単なる

記憶容量の増設は内部処理速度、記憶装置の費用の点からも最良であるとは言いがたい。以上の点は現行装置でのプログラムの分割処理を極力少なくし、プログラム作成本数の増加によるプログラム処理時間の増加を極力避けるべきであることを示している。

この問題に対処するためには、中央処理装置の演算速度の高速化と内部記憶容量の増加が必須となるが、さらに実際的なプログラム処理からは、個々のプログラムサイズに拘束されない記憶区画の動的配当が望ましいことになる。

この機能は仮想記憶方式の導入によって解決されると思われるが、現在発表されている中型以上のコンピュータでは、記憶管理に仮想記憶方式が採用されているため、新機種導入によるプログラム分割問題、そのほか物理的記憶装置上の記憶域の断片化は一応解決できると思われる。

しかし、物理的には有限の内部記憶容量とバックにある拡張記憶装置に磁気ディスクを使用するため、記憶装置と磁気ディスク間のプログラムのロールイン、ロールアウト回数が処理効率に大きなインパクトを与えることになる。

これは、いわゆるページング・アルゴリズムの問題であるが、実用限界として、仮想対実記憶の比は1.3～2倍程度と実測がなされていることから、これも、無制限に仮想記憶領域の拡大は行うべきではないと思われる⁴⁾。しかし、一般的に入出力処理 (i/o bound) の多いジョブに対しては、前者の内部処理 (cpu bound) に対し、この条件はゆるめられる⁵⁾。

また、現行の HITAC—8250での使用可能記憶区画は40KB、30KBであることから、今後数年間は、極端な記憶容量の増加はないものと思われる。M—150では、ユーザ区画4～5程度とし、平均的実領域サイズは80KB、また仮想領域は一部のユーティリティが400KBを要求しているため（このユーティリティは使用頻度はあまり多くないが）この条件を満たす方向での区画配当が必要となろう。ユーザ領域区画は主に、現行の区画と増設される入出力の装置数によって決定されるものと思われるがM—150でのOSはVOS2常駐部は推定150KB、ユーザー利用可能領域362KBである。

次に磁気ディスク上の問題点として、現行ディスク装置は29MB（データ読み取り速度60ms）2台であるが、各ユーザ領域で使用できる台数はシステム

用を省き1台である。また、物理的記憶容量から見た場合、現行ディスクでは実際に記憶可能な容量としてレコード長320B, 100Bでそれぞれ約3.6万件, 8万件とレコード長が少なくなるにつれ実用容量は減少する。また、漢字データの場合その量はさらに半分となり、最多使用レコード長100Bではディスク最大容量29MBに対し17MBと約2分の1の容量になる。この容量では、ユーザ領域2区画に対し、マルチ・ファイル構成は問題なしとは言えない。新装置では、この点を考慮し、ウェンチェスタ型70MB4台とし、システム以外にユーザ領域3区画に対し各々1台の配当を、さらに容量面でマルチ・ファイルの利用も十分可能となるように配慮した。

これは、駆動装置の導入コストの他に、複数台のディスク装置設置によってバックの取りはずしができること、オフラインシークまたダウン時の装置の切り換え、その他ファイルの個人別管理に重点をおいたためである。これによって、1ディスク当りのレコード長320B, 100Bで各々8.6万件, 20万件と処理可能件数は2.5倍程度増加し、3バックのマルチ・ボリュームでは、さらに25.8万件, 60万件となる。これで、現在蓄積されている約400万件のデータ中、一部であるが、共同利用のためのデータベース化が可能となる。また、従来のSAMのほか、新たにVSAMの利用によって、データ長、ブロッキング・ファクタの自動処理化ができ、ディスク利用のさいにも仮想記憶方式と同様、論理データのみを考慮すればよいことになる。以下の5点は新ディスクの導入によって得られる利点である。

- 1) 各ユーザー領域単位で独立したディスクの配当が可能
- 2) マルチ・ファイル、マルチ・ボリュームの利用が可能
- 3) 大量の漢字データを共同利用するためのデータベース化が可能
- 4) VSAM形式のファイル利用が可能
- 5) 処理速度の高速化が可能

次に消耗品関係としては、プログラム作用及びデータ作成に紙カードまたは、漢字などの記憶媒体に紙テープを使用してきた。しかし、データ修正の容易性、データ保管場所、媒体強度の問題、記憶密度、反復作用の可否、その他データ管理面で問題があった。

フロッピーディスクの採用は（以下FD）、クリーン・データ作成また、プログラム・デバック時における反復修正時のデータサーチの容易さ、媒体のコンパクト化、媒体の反復使用が可能なことから、大量のデータを扱う場合、消耗品の無駄を防ぐことができる点で有利であると思われる。またFDの記憶容量は紙テープやカードに比べ、物理的に英数字約243KBまたは約1MB記憶可能であり、243KBの場合でカード約3000枚、紙テープ620メートル分に相当する記憶容量を持つことができる。

また、オフライン装置によるデータ作成、修正、編集処理が可能なところから、ホスト側の負荷を軽減することができ、併せてコンピュータ組織外でのデータ管理、作成処理が非常に容易にできることが利点となる。そのほかのFDの重要な点として、その初期の目的が単にカードの代替品としてサポートされていたにもかかわらず、利用方法としてはローカル・マスタファイルの考え方を導入できる簡易型ディスクとしての位置付けが可能になる。

特に、オンライン・インテリジェント機能を持った分散処理装置と併用されることによって、その位置付けが強力になり、将来、クリーン・データ作成のための前処理として、階層的組織構造とマッチしたコンピュータシステムを組織できることになる。しかし、現在FD用にサポートされている専用入出力装置については、前述のカードイメージ処理を前提としているため、例えば異った記録媒体のFD使用にはそれぞれの専用装置が必要となる。また、バッチ処理形態を指向しているため、FDによるホスト側データのリアルタイム処理では、ターンアラウンド時間の点で問題なしとは言いがたい。この種の処理は将来man-machineの相補的処理を必要とする場合に重要な形態になると思われ、無視できない機能の一つとなろう。ただし、FDに対するデータフォーマットに対しては端末側のインテリジェント機能を利用して、ロジカル構造に特殊なファイル形式をとった場合、ホスト側処理でデータの再構成を必要とする。このオーバーヘッドに注意する必要がある。

そのほか、プリンタなど出力装置上の問題として、長時間の英数字出力処理を行った場合、他のジョブでのプリンタ使用が不可となることから、出力装置の複数化の必要が生じてくる。

特に今後使用者の増加と共にこの様な状態が多くなると思われる。しかし、実際に長時間使用の頻度は相対的に低いため、そのために特に、副プリンタを導入することは無駄が多い。この種の処理で直接影響を受けるのは大部分プログラム・デバック処理及びラン用メッセージ処理であり、前項目で述べたFD処理用インテリジェント・ターミナル側での低速プリンタによって、また端末側のリモート処理で、これらの弊害は大部分回避できることになる。

また、プリンタ装置については将来漢字プリンタと現在の英数字プリンタの共用が進められると思われるので、これらの使用形態はバッチによる大量印刷時のシステムストップを防ぐ手段として有効であると考えられる。ただし現在市販されているノンインパクト方式の高速プリンタは印字用紙及びトナーなどの消耗品を含めた場合11×15インチ1枚4～5円程度となり、通常インパクト式に比べ用紙価格が高いという問題がある。

最後に漢字入力及びプリンタ装置上の問題として、現行装置の構成は、ホスト側に通常の汎用コンピュータを、また漢字処理装置はスタンド・アローン形式で漢字プリンタを使用している。当初このシステムは、HITAC-8250の性能上の限界から、処理の分散化を図り、負荷をホスト側から少しでも、はずす目的で計画されたものである⁶⁾。これはまた、漢字プリンタが高速バッチ印刷を対象として開発されているため、デバックなどモニター用としての利用は不便なことが多い。しかし、プログラム・デバック時などには、低速用モニター装置が必要になることが多く、漢字キーボード付きディスプレイ装置、またL-320付属プリンタを副プリンタとして使用することができれば、これらの問題は解決する。当然、漢字ディスプレイ装置などのオンライン端末は、他のインテリジェント・ターミナルL-320からの直接コントロールが可能である。

以上現行コンピュータの問題点を概観し新機種に対する基本的な要求事項を簡単に示してきた。これらの事項は、次の9項目にまとめることが可能であり、これらの項目を満足することが新機種に対する基本的な要求事項と一致する。

- 1) ユーザー利用領域の拡張について
- 2) 磁気ディスク、磁気テープなどのファイル容量と使用形式について

- 3) カード、紙テープなどの消耗品の節約について
- 4) プリンタ装置などの出力装置の複合化について
- 5) コンソール処理を含むオペレート運用について
- 6) ホストコンピュータでの漢字入出力機能と漢字データ処理について
- 7) ローカル・ファイルの採用と分散処理について
- 8) データ保存場所の節約について（カード、紙テープ、磁気テープなど）
- 9) L-320内でのデータ修正機能について

3. 国語研究所における処理の特長

コンピュータ導入の目的が大量の用語用字調査を短期間に行うために計画されたため、利用範囲も言語情報処理に関する方法論、また調査の集計、統計処理、用例集の作成などに重点がおかれていた。取扱うデータについても、一部統計計算用の英数字データを省き、大部分は漢字データである。当然、コンピュータ処理も大量データを指向したファイル処理が中心となっている。まず、通常（定型）業務として調査関係の特長を述べ、次に研究的（非定型）業務について、同様の説明を行う。定型業務の特長は、次の3点が主なものである。

- 1) 同一プログラムの反復使用が多くかつ長期使用である。
- 2) 入出力のデータ量が多く、大部分は漢字データによるファイル処理である。
- 3) オペレートは、クローズ方式である。

定型業務上の問題点として、1) については、プログラム保守と管理上の問題、特に装置の増設、またリプレースに伴うプログラムの拡張と移植性などのプログラム維持について 2) については、ファイル保存スペースと管理上の問題、また、多種多様化する媒体管理が問題となる。

これに対して、非定型業務は、プログラム開発に関して、試行錯誤が多く、中小規模のデータを使用した反復テストが大部分となる。当然オペレートについても、作業の性格上、研究者自身の処理が中心となろう。

以上が、研究所におけるコンピュータ処理上の特長であったが、以下これを従来のコンピュータ処理方式であるバッチ形式で実行した場合の問題点と利点

について述べ、分散処理へ移行していくことがバッチ処理形態の問題点の解決になることを示す。次の4項目はバッチ処理における長所となるものである。

- 1) 導入初期においては運用管理が容易である。
- 2) オペレーション、プログラム作成を集中化でき、かつ専門要員を効果的に使用できる。
- 3) 装置システム構成が一元的であり、複雑なシステム構成を必要としない。
- 4) 大量一括集中発生 of データ処理に適している。

以上の4点については、いずれもコンピュータ利用初期段階では、必須の条件であり、全てこの条件を最初に満足することが運用上重要である。しかし、前述の現行コンピュータの問題点で述べたように、集中化処理方式には、問題点も少なくない。

- 1) 装置が専有化され、運用面で自由度が失われる。
- 2) オープン処理利用ができにくい面がある。
- 3) 端末装置の多様化により、ホスト側サポートに限界が生じた。
- 4) データ修正など処理効率化を妨げる要因がホスト側で無視できなくなった。(例外処理に対する効率)
- 5) 組織からの多様な要求を一セクションではカバー不可能となってきた。
- 6) ターンアラウンド時間がかかりすぎる。
- 7) ホスト側CPU負荷が多い。
- 8) システムの大型化が価格対性能比の向上に結びつかなくなった。

そのほかの問題として、例えばオープン処理利用時に非専門要員によるオペレートミス、これは教育期間、平均的勤務期間にも直接関係するが多重度が増加するにつれ、同一コンソール上に複数の要員が集中してくるため、他領域のキャンセルミス、またコンソールメッセージの読み違いなどの基本的操作上の問題が多く発生する。同時に磁気テープ装置、磁気ディスク装置の割りあて、出力装置の長時間専有による他ジョブの出力停止の弊害がこれに加わる。これはデバッグ処理のさい、出力装置が使用できないことは他領域のプログラムに大きな影響を与えることになる。

以上が、集中処理における長短の主なものであったが、これらの問題を解決するために次に述べる分散処理の考え方の導入が有効であると考えられる。

4. 分散処理への試み

分散処理については、かならずしも新しい考え方ではなく、従来から現場で、経験的に行われてきた部分が少なくない。しかし、従来の装置では価格の面で装置、人などの二重投資を極力避ける傾向があったため、あまり積極的に行われているとはいえなかった。これについては最近の素子のLSI化やマイクロコンピュータの普及により経済性的問題は徐々に解決されつつあり、分散処理導入と普及の気運もここにあると言ってよい。また分散処理について種々の概念や定義が発表されている。しかし、本稿で述べる分散処理機能に対する対象は、データエントリを含むFD処理、専用装置としての位置付けとファイル管理面でのローカル・ファイル機能またリモートバッチ処理に伴うシステム概念に範囲を限定する。この方式は、データエントリ装置として必要なコンソール、プリンタ、補助記憶装置などをリモートターミナル付属の各装置と重複させることにより分散処理を可能にしている。次に、システム面から見た場合、コンピュータ処理形態と人的作業組織への対応付けが、このシステムをささえる重要なポイントとなるが通常はデータ処理の対象になる問題に対し処理部門単位でグループ化され、それらが有機的に結合されていることが多く、運用上はシステム構成をどの人的組織にどのように対応させるかが重要な問題となる。

我々が分散処理を導入する動機は前述の通りであるが、定型業務からの利用としては、いかにクリーンなデータを迅速に作成するかという点がコンピュータ利用効率を上げる上で重要となる。通常データ作成の場合、漢字または英数字に限らず第一次さん孔の後、校正用ゲラシートを作成、校正作業の後コンピュータまたはオフライン装置を使って修正が行われる。この間、作業の各セッションは独立して処理されることが多い。この点、人的組織面から見た場合このシステムは従来の分散処理方式で定義されている階層型分散処理システムに近いと言えよう。このシステムは人的組織が論理的に結合されているのと同様、かならずしも装置間が物理的に接続されている必要はなく、装置の論理的

結合を許すものとする。以上の点を基本にし、実際の装置ではどのようなシステムが実現可能か、また分散処理で得られる利点はなにかについて述べる。

集中処理については、HITAC-8250周辺問題として述べたが、分散処理はこれらの障害を排除する目的で導入計画されたものである。この形式はシステムを構成する装置構成とシステムの運用面に分けられる。運用面では、人的組織に対応しているため各部門の目的にあった専用装置としての機能が要求される。これは、オペレート、教育、管理など運用面でも独自の方法が許されることを意味する。また各部門ごとの使用目的の変化に応じ、専用機としての機能は失わず他部門への影響も最少限におさえた処置が可能である。オフィス・コンピュータによる専用装置としての位置付けは、データ入力、修正機能を持つFD処理専用装置が対応する。

そのほか、現場で発生した一次データをホスト側に転送することなく、大部分の処理を端末側で処理できるため、プログラムやデータファイルなどは部門単位で自主管理できることになる。また、ターンアラウンド時間の短縮を含んだ全体として小回りのきく処理が可能となる。これらの端末側の機能はローカル・ファイルが組織内で閉じている場合に特に有効となる。同時に、使用目的と部門固有の操作、プログラムなど独自のものが作成でき自由に運用できることになる。

当然、この処理形態を満足させるためには、ローカル・ファイルを維持するための補助道具として、ローカルでのプログラム開発、オペレートの容易性、またローカルがゆるされる環境として維持費用が安いこと、データ保存媒体の安定性また、他システムとのデータの互換性が保障されなければならない。そのほか、ホスト、ローカル装置及びソフト共に一般的な対象から任意に選べるものであることが必要である。

以上の点をまとめた場合、我々が分散処理システムとして構成した装置は、現行のHITAC-8250の問題を次の装置で、それぞれ複合して解決できるものと思われる。

- 1) コンソール機能の複合化
- 2) プリンタ装置の複合化

3) インテリジェント・ターミナルの導入とリモート機能の導入

4) F D処理専用装置とローカル・ファイルの採用

新規導入の対象となるコンピュータシステムは、インテリジェント機能を持ったオフィス・コンピュータを端末装置(L-320)としてホストと回線で結んで相互伝送を可能とする形態をとる。L-320の機器構成は表2に示したが、形式的には従来のリモートバッチ処理方式に類似している。しかし、前述の通り L-320付属端末装置の利用により、端末側で個々に独立した処理が可能であり、この機能によってローカル処理及びローカル・ファイル処理に重点をおいたシステムと言えよう。

表2 L-320 機器一覧

内部記憶装置	48K B
ディスプレイ装置	40字×20行
F D装置	2台 243K B / 1 M B 併用可
シリアルプリンタ	120字/秒

L-320は、オンライン及びオフライン形式、いずれの処理も可能であるが、それぞれ代表的な使用方法としては、次のような使い方が考えられる。

オフライン形態（スタンドアローン形態）

- 1) データエントリ処理
- 2) インテリジェント機能を使った各種F Dユーティリティ処理
- 3) 簡易言語R F DまたはR P Gを使った各種バッチ処理

オンライン形態（フロントエンド形態）

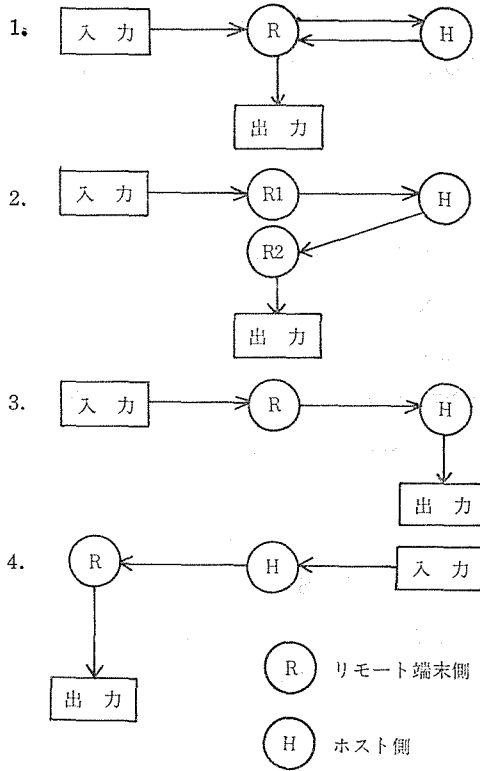
- 1) リモートジョブによるバッチ処理
- 2) ホストに対するF Dの入出力処理
- 3) ホストに対するオンライン処理

L-320ではそのほかの機能として、フォグランド、バックグラウンド処理機能があるため、これを適当にジョブに割り合てることによって、異なった複数のジ

ジョブを同時に実行させることが可能である。これらの機能は、端末側でのデータ入力または出力など装置専有時間の比較的長い処理によって、他のジョブが停止することなくホストに対するFD入出力処理を並行して実行できることを示す運用上重要な機能である。また、各種FD処理用ユーティリティの利用は、現行のホスト側FD入出力専用装置の処理機能に満足できにくい点が多いため、FDをベースとしたオフィス・コンピュータのFD機能の利用が必要となる。これによって、より自由度の高いFD処理とホストに対するFD入出力装置としての併用機能また、手足を持った専用デバックマシン、そのほかローカルで非専門要員による簡単なプログラム開発とホスト側で処理される汎用プログラムの開発と実行が可能となる。この機能は、定型、非定型業務に限らずコンピュータ非専門要員の教育期間などを考えた場合、オペレータの容易性と共に重要である。また、将来ローカルで作成されたデータに対し、ホスト側マスターファイルからの情報の付加、半自動的処理結果に対する人間の目視チェックなど man-machine 相互の連帯処理が増加するものと思われ、小なくても、この処理を実行する場合、ターンアラウンド処理における時間を無視できないものとなるはずである。この点、従来紙カードの代替として開発されてきたFDに対し、小中型のホスト・コンピュータに対する位置付けが明確でなく、むしろ、オフィス・コンピュータレベルに小回りのきく機能を持った装置が多い。我々の導入対象となったM-150についても、FD入出力装置にL-320インテリジェント・ターミナルを接続した意味の大部分はここにある。当然、L-320はFDを記憶媒体の基本としているため、FD処理のためのユーティリティが完備しており、ホスト側で開発しなければならないFD処理用基本プログラム作成労力から大部分開放されることになる。

また、FD処理用ユーティリティの積極的利用によって、定型業務におけるデータ形式の標準化を進めることができ、併わせて、前述のFD内のデータ伝送機能、データエントリ装置の利用も可能となる。また、L-320でのFD入出力装置は1セクタ128Bまたは256BのFDの併用ができるため、英数字データ、プログラム、漢字データなどデータ量の大小に応じた使いわけと、FDデータエントリ専用装置とのデータ互換性も確保できる。

図1 ホスト、リモート間データ伝送経路



そのほか、1枚のFDにプログラム及びデータの混在したファイル記録が可能であり、デバッグ終了までのあいだ磁気テープ、磁気ディスク装置への配当、割りあてから開放されることになる。

次にリモート端末について、リモート端末及びホスト間のデータ伝送の方向は次の4種の形式が可能である。1の流れは、リモート側からのデータ投入後、ホストによる処理結果をリモートに返す形態である。これは、リモートからの直接ホスト側プログラムの呼び出し実行も含まれる。2の形態は、1とほぼ同じ処理であるが、出力結果を他のリモート端末に転送する例である。3については、出力結果を直接ホストに出力する場合、また磁気ディスク、磁気テープそのほか漢字ビデオ装置への情報格納の場合に相当する。

最後の形は、逆にホスト側で処理されたジョブ結果を、リモート端末に出力す

る場合で、ホストでの長時間にわたるラインプリンタ使用時のサブプリンタとして使用、またFDへの出力に使用する。なお、本装置でリモート側に直接出力可能な媒体は、FD及びプリンタである。L-320付属ディスプレイへの出力は全てFDを介して間接的に行われる。

5. おわりに

本稿では、新規に導入される HITAC-M150 について、主に現行のコンピュータの持つ問題点と業務の特性から、必要機能を検討してきた。

導入決定までは、種々の議論がなされ、またリモート方式以外に例えばTSS (Time Sharing System) の導入の可否についても検討された。

しかし、一般定型業務への移行問題と予算レベルの制約により、将来の問題として残された。我々の要求する必要機能としては、TSSでのプログラム開発とローカル処理によるデータ作成またローカル・ファイル形式などインテリジェント機能と手足を持ったTSSターミナルがメーカーによってサポートされることを望むが、現行の中型以下では、TSS機能の拡張など、ハード、ソフト上の問題など無視できない問題も少なくない。しかし、オフコンのインテリジェント機能を積極的に利用することによって、会話型データ修正処理をローカルで実行でき、一部であるがシステムにTSSエディタ能力を持たせることができるようになる。

本稿で提案したオフコンによるシステムは、従来の集中処理方式から分散処理方式への移行への足がかりとなるものであり、本システムの導入によって、従来の利用形態と業務組織についてもゆるやかな変革が必要となると思われる（逆にシステムを研究所組織にあわせたと言うこともできるが）。

分散処理の利点としては、各部門単位での独立業務と運用、オペレート、教育などについても使用環境に順応したシステムの構成が可能である。また、データ管理面のソフトウェア上の問題を人間側での作業に置き換えることによって逆に各担当者の責任を明確にし、これらの形態は、これによってはじめて各自の自由裁量と自主的運用が許されることになる。しかし、オペレート及び運用上で非専門員に対するガイダンスやコンピュータによる会話型処理方法など

従来のバッチ処理と異なった面も少なくなく、実際面でさらに細部にわたって検討が必要となろう。

日本語の言語情報処理について、ほぼ14～15年が経過している。当初5～6の機関で開始されていたシステムも、現在では汎用機器装置上で日本語処理パッケージが利用できるところまで一般化している。また、入出力装置についても、ほぼ実用レベルに適しており今後これらの手足を効果的に利用していく上で、データ蓄積の方法が重要になると思われる。また、データ入力後の修正方法と校正効率の改善なども同様に重要な課題の一つとなろう。本稿で述べてきた分散处理的システム導入のきっかけとなったのも、データ修訂正処理の効率化と、よりクリーンなデータ作成に重点をおいた、システム構成を必要としたからにほかならない。

なお本稿のデータ修訂正処理に関するハードウェア構成上の基本的な考え方は、文部省科学研究費「言語解析を応用した日本語文修正の効率化に関する研究」（一般研究C研究代表者斎藤秀紀）を受けて行ったものの一部である。

参考文献

- 1) 野口正雄：「データエントリシステム入門」1978
- 2) 半沢孝雄：「磁気ディスク技法集」1978
- 3) P. J. DENNING：「Virtual Memory」Computing Surveys Vol. 2 No. 3 1970
- 4) D. J. Hatfield and J. Gerald：「Program restructuring for virtual memory」IBM System Journal No. 3 1971
- 5) J. E. Morrison：「User program performance in virtual storage systems」IBM System Journal No. 3 1973
- 6) 斎藤秀紀：「言語処理におけるターンアラウンド・システム」国語研究所報告 59, 63—112 1976
- 7) 斎藤秀紀：「多目的漢字入力システムの試案」国語研究所報告61, 41—56 1978
- 8) 戸田正直：「人間の行動における分散と統合の概念形成」システム工学会資料 19—34 1978
- 9) 「HITAC—L320システム概説書」日立マニアル

図2 システム構成図

