

LANとパソコンからなる分散型データベース システム LIS について

(財)日本情報処理開発協会 滝沢 誠(Makoto Takizawa)

(財)日本情報処理開発協会 横塚 実(Minoru Yokotsuka)

1. はじめに

VLSIを中心としたハードウェア技術の進歩により、従来の
小中規模並みの性能を有した超小型計算機が、事務所、工場
等の各作業者に普及してきている。これにより、各作業者が
単に計算能力を持てるだけでなく、必要なデータを保有出来
る様になり、てきている。一方、Ethernet [METCR76] 等のローカル
エリア網(LAN)の発展により、一つの組織体内の種々の計算
機の相互結合が可能となり、てきている。この様なシステムで
は、超小型計算機は、情報の送受信端末としてではなく、デ
ータベースシステム(DBS)を中心とした作業用のワークステーションと考
えられ、他の計算機のDBSから、自分の計算機にデータを導
出し、蓄積して操作する為に用いられてくる。

既に、従来の大型の異種DBSを統合化する為の分散型デ
ータベースシステム(DDBS)としては、JDDBS [TAKIN78-84]が実現されて

いる。本論文では、一つの組織内の種々の計算機内のDBSをLANを用いて統合化する為のDBSとしてのローカル情報システムLISについて論じる。

2. LISのアーキテクチャ

LISは、LANで結合された複数の仮想情報システム(VIS)から構成される(図2.1)。VISには、超小型機上に実現された個人情報システム(PIS)と、従来の大型機上に実現された中央情報システム(CIS)との二種がある。PISは、各個人の作業用の口

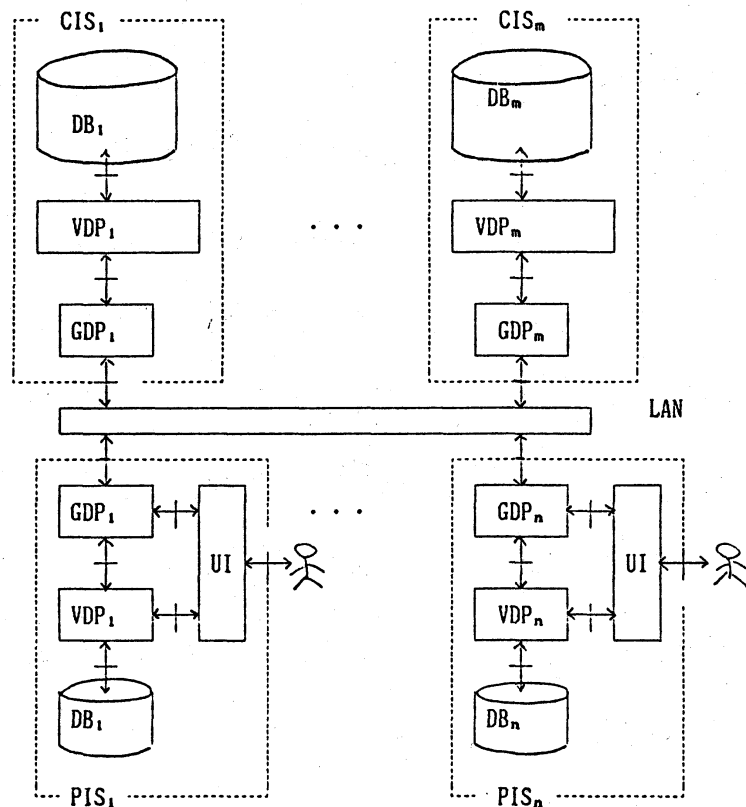


図2.1 LISの構成

- カルなデータが格納され、GISには組織体全体の共有データが格納される。各VISは、仮想DBS (VDS) と全体データベースプロセッサ (GDP) とから成る。VDSは、ローカルなDBSであり、データの集合としてのDBと、DB内のデータ操作を行う仮想DBプロセッサ (VDP) とから成る。VDSは、関係型のDBSとする。即ち、VDPは関係型のDBMSと考える。DBSが関係型でない場合には、関係インタフェースLDP (ローカルデータベースプロセッサ) [AKI80, 82, 83] により、関係型DBSに仮想化されているものとする。GDPは、VIS間での通信処理の為に制御システムであり、(i)分散検索処理、(ii)分散トランザクション処理、(iii)セキュリティ管理、(iv)DTPD管理を行う。UIは利用者インタフェースである。

3. 放送網としてのLAN

EDDBSを実現するうえで、各VIS間の通信を行う通信網は、重要である。LANは、従来の広域パケット交換網と比較して、高速(10~100Mbps)、高信頼、狭域(数km)、及び放送通信性を特徴としている。特に、LANの放送通信性は、EDDBSの通信処理を簡単で有効なものとする[AKI81, 84b]。

一つの点から、 n ($n \geq 1$) 個の点に同一情報 X [情報量 x bits] を通信することを放送通信とする。特に $n=1$ のとき、一

対一通信とする。点 i から、点 j_1, \dots, j_n に情報 X を通信する時、通信網に送信される情報量 (= パケット量) を通信量 $T_{ij_1, \dots, j_n}(\alpha)$ 、又通信を始めてから終了するまでの時間を通信時間 $R_{ij_1, \dots, j_n}(\alpha)$ とする。LAN では、両者は次式で与えられる [AKI 81, 84b]。

$$T_{ij_1, \dots, j_n}(\alpha) = \alpha + \beta \alpha \text{ [bit]} \quad R_{ij_1, \dots, j_n}(\alpha) = (\alpha + \beta \alpha) / \omega \text{ [sec]} \quad \dots \quad (1)$$

ここで、 α と β は定数である。以上より、LAN は、放送通信を行うとき、従来の広域パケット交換網に対して以下の利点を持つ。又、 ω は、LAN の伝送速度 [bit/sec] である。

① 通信量と通信時間の減少。①式より、一パケット送信すると全点にパケットを通信出来る事から明らかである。

② 同時到着性。送信されたパケットは同時に全点に届く。

本論文では、以上の LAN の特徴を生かした DBMS 処理手法を与える。

4. 全体き-タバースプロセス (GDP)

GDP は、VIS 間の通信処理を行う為の制御システムである。

4.1 GDP の階層構造

GDP は、図 3.1 に示す様な三層の階層構造を有している。GDP は、LAN の提供する Ethernet 接続 (EDL) サービスを用いて実現されている。IPC 層は、EDL サービスを用いて、VIS 間のプロセス間通信 サービスをシステム サービス (SS) 層に提

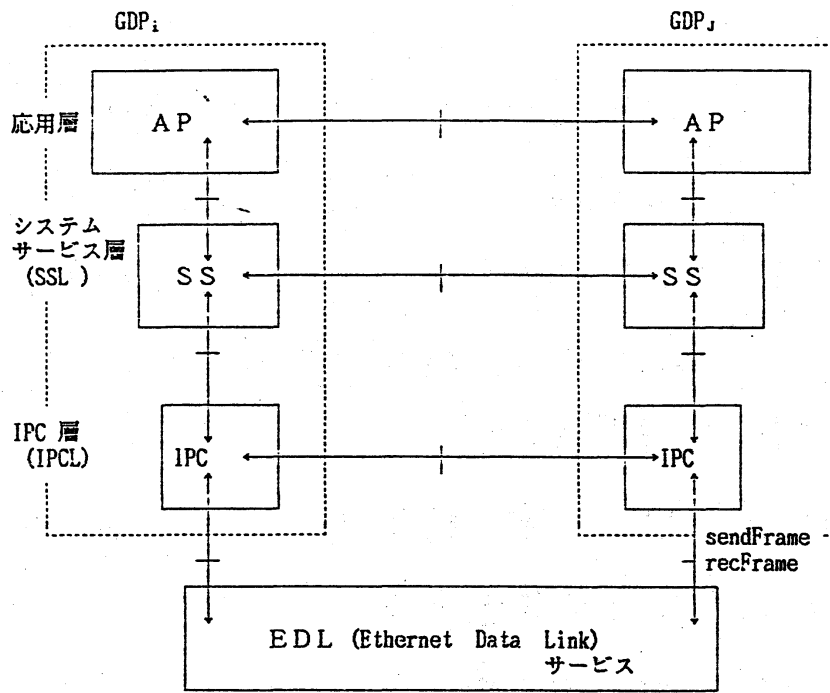


図4.1 GDPの階層構造

供する。SS層は、DDBSの基本サービス(検索、トランザクション、セキュリティ、DDI)を提供する。応用(AP)層は、DDBSの応用サービスを提供する。

4.2 IPC (interprocess communication) 層

VIS間のプロセス間通信を、メッセージパッシングにより行う。IPCは、EDL層の送受信サービスを基本として、以下のサービスを提供する。

群サービス 複数のVIS内のプロセス間で群(cluster)を形成する。N2Nの1対1のプロセス間の仮想回線サービスをm(m>2)個のプロセス間に拡張したものである。群内のあるプロセスから送信されたパケットは、群内の全点に同時に届けられる。群サービスには、更に次のモードがある。

- (1) { 動的群 必要に応じて、プロセス間で群を設定する。
 永続群 VISの親への 長変時に群が設定される。
- (2) { 保障群 群内で、あるプロセスから送信されたパ
 ケットは、必ず全宛に到着することを保障。
 非保障群 保障されない。

(3) データグループサービス 群の設定をせずに、パケットを任意
 の(複数の)プロセスに送信出来る。

4.2 システムサービス(SS)層

SS層では、IPサービスを用いて、以下の基本DBSサービ
 スを提供する。

(1) 分散検索演算処理サービス ^(非手動的な) 複数のVIS内の関係に対する複集演
 算の条件を満たす目標集合を、VIS間の通信処理によ
 り導出する。利用者に、関係の所在を意識する必要のない
 検索演算を提供する。

(2) 分散トランザクション処理サービス 複数のVIS内のデータの
 更新を行うトランザクションに対して、VIS間のデータの
 インテグリティを保ち、かつDBS全体の性能向上をはかる為
 の通信処理を行う。トランザクションの複数のVIS内のデータに
 対する更新の原子性を保障するコミットメント制御、各VIS
 での複数のトランザクションからの演算の実行順序を一定に保
 つ為の同時実行制御、障害かたVIS間のインテグリティを保つ為

の障害管理とを行う。

③ セキュリティ管理サービス 各VLSへの不正アクセスと、情報の流出に対する制御を行う。

④ DB管理サービス システム内のデータの存在、操作手法についてのメタデータの管理を行う。

4.3 応用(AP)層

AP層は、S層のサービスを用いて、EDBSの応用サービスを提供する。現在、応用サービスとしては、視野サービスがある。の利用者が、LIS内のVLSから、自分のPISにデータを導出したとき、このデータと、VLS内のデータとの一貫性を保つことと、⁽¹⁾導出されたデータを視野としたときの視野更新とがある。後者は、視野を抽象データ型として、許される演算のみを利用者に使用させ、各演算の実現としてトランザクションを与えておくことにより、解決する。

5. 分散検索演算処理方式

LIS内の各VLS内の関係⁽²⁾のもの⁽³⁾の集合を全体データ構造とする。全体データ構造の全体検索演算は、QUEL[HELD67]で表せられるとする。QUELで表現された全体検索演算を全体問合せとする。単独問合せについては、[CAMP81, 83c, 84b]に論じがあるので、複数属性問合せについて考える。

5.1 複数属性問合せ

全体問合せ Q の条件式 σ が、複数の属性の等結合条件式($x_i.v_i = x_j.v_j$)の積であるとき、 Q を複数属性問合せとする。ここで、 Q の参照する関係の集合を $\Sigma = \{X_1, \dots, X_n\}$ とし、 X_i は VIS_i にあるとする。 A を Q の結合属性 v_i の集合とし、 A_i を Q の目標属性集合とし、 A_i を X_i の目標属性集合とする。 A_i を X_i の結合属性集合とする。 $A = \bigcup_{i=1}^n A_i$, $A = \bigcup_{i=1}^n A_i$ 。又、各 VIS_i 毎の初期ローカル問合せ処理(IQP) [HEVNA78]は終了しているものとする。例として、図5.1を考える。四個の関係 X, Y, Z, W

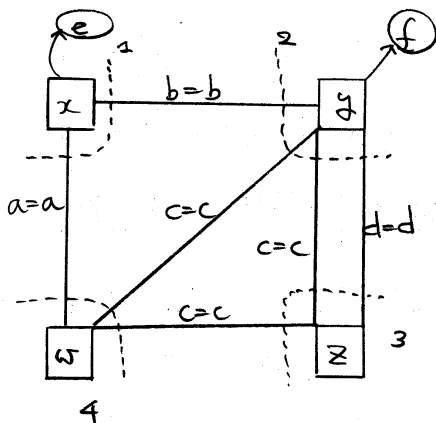


図5.1 問合せ例

が、各 $VIS_{1,2,3,4}$ に存在するとする。これに対する全体問合せの例を以下に与える。

```
var (x, X) (y, Y) (z, Z) (w, W);
get into Temp [x.e, y.f] where
x.a = w.a and x.b = y.b and y.c = w.c and
w.c = z.c and z.c = y.c and y.d = z.d ; -- (1)
```

(1)のグラフ表現を図5.1に与える。

各 VIS での操作実行と、 VIS 間での通信実行の系列をトランザクションとする。全体問合せ Q の目標集合を求めるトランザクション T を求め、各 VIS で T の基本実行を実行させることを、 Q のEDBS処理とする。

5.2 基本定義

全体閉合せのRDBS処理手順を説明するうえで必要となる基本操作算の定義を与える。

A. ビットマップ生成算 //

X を全順序集合 $\langle a_1, \dots, a_n \rangle$ (ここで $a_1 < \dots < a_n$) とし、 Y を X の部分集合とする。 X に対する Y のビットマップ生成算 $Y // X$ を次の様に定める。

$$Y // X = \text{系列} \langle b_1, \dots, b_n \rangle \quad \text{ここで } b_i = \begin{cases} 1 & \text{if } a_i \in Y \\ 0 & \text{otherwise} \end{cases} \quad \text{--- (2)}$$

$Y // X$ を、 Y の X に対するビットマップとする。

B. ビットマップ制限算 *

A. の X と Y について考える。 $\langle X \rangle_i$ を a_i とし、 $BM = Y // X$ とする。 $R(A, B)$ を関係とし、 A と B は各々属性集合とし、 $\text{dom}(A) = \{a_1, \dots, a_n\}$ とする。このとき、以下を定義する。

$$X * BM = \langle \langle X \rangle_{e_1}, \dots, \langle X \rangle_{e_2} \rangle \quad \text{--- (3)}$$

ここで、 $\langle BM \rangle_{e_i} = 1$ かつ $e_i < e_j$ ($i < j$)。即ち、ビットマップ

J の BM のビットオン (=1) に対応した系列 X の値の系列である。

$$R * BM = \{ t \mid t = \langle a, b \rangle \in R \wedge a \in X * BM \} \quad \text{--- (4)}$$

C. コード代算 ^

B. の X と R を考える。 σ を置換 $\{0/a_1, 1/a_2, \dots, (m-1)/a_n\}$ とする。このとき、以下を定義する。

$$R \wedge X = \{ \langle v, b \rangle \mid \langle a, b \rangle \in R \wedge v = \sigma a \} \quad \text{--- (4)}$$

R^X を関係 $S(A, B)$ とする (ここで, $\text{dom}(A) = \{0, 1, \dots, m-1\}$).
 $S(A, B) (= R^X)$ を, $R(A, B)$ の X によるコード化関係とする. 又,
 A を, 属性 A の X によるコード化属性とする.

D. 放送通信計算

- (i) broadcast (X); X は, ある順序集合とする. X を全点
 に, 値順に放送する.
- (ii) $Z \in \text{rec}(U)$; 放送された情報を, 送信順に受信し順
 序集合 Z とする.

5.3 DBMS処理の為の通信手順

複製属性問合せ Q のDBMS処理する為の通信手順を以下に示
 える. これを B^2 と [broadcast, bit-map responding, and coding] 手
 順とする.

[B^2 通信手順]

- (1) Q を受けた VIS_i は, Q の参照する関係を有している全点 x_i に
 Q を放送する.
- (2) Q を受けた VIS_i は, 初期D-カル問合せ処理(ILOP)を行い,
 目録集合 X_i を求め, X_i の性能情報(組数, 各属性値の数, X_i
 の物理構造) PI_i を放送する.
- (3) 各点 VIS_i は, 全点から PI を受信したならば, 次の乗に初期
 化を行う.

$$\Sigma' \leftarrow \{X_1, \dots, X_n\}; \quad \Sigma'' \leftarrow \{\emptyset\}; \quad \alpha \leftarrow A \text{ (全結合属性集合);}$$

各 $X_R \in \Sigma$ に対して、 $\mathcal{A}_R \leftarrow A_R$ (X_R の結合属性集合);

(4) [送信 VIS_s と関係 X_s の決定]

全点 VIS_s は、 Σ' のなかから、最優先 X_s を選ぶ。但し、 \mathcal{A} が Γ -ド化属性 \mathcal{A} を含むならば、 $\mathcal{A} \in \mathcal{A}_s$ なる X_s のなかから選ぶ。

(5) [X_s の放送]

VIS_s は、 $X_s[\mathcal{A}_s]$ を読み出して、一定順序で放送する。

$$Y_s \leftarrow X_s[\mathcal{A}_s]; \quad \text{broadcast}(Y_s); \quad BM_{ss} \leftarrow Y_s // Y_s;$$

$$X'_j \leftarrow X_s; \quad j \in S; \quad \text{右左} \quad (7);$$

(6) [X_s の受信]

(i) $\mathcal{A}_i \cap \mathcal{A}_s \neq \emptyset$ なる関係 $X_j \in (\Sigma' \cup \Sigma'')$ を有する全点 VIS_j は、 VIS_s からの Y_s を受信し、 Z_j とする。ここで、 $\mathcal{A}_{sj} = \mathcal{A}_s \cap \mathcal{A}_j$ 。

$$Z_j \leftarrow \text{rec}(Y_s);$$

Z_j と、 X_j とを \mathcal{A}_{sj} について結合し、 $X'_j(\mathcal{A}_s, \mathcal{A}'_j, A_j)$ を得る。

$$X'_j \leftarrow X_j[\mathcal{A}_{sj} = \mathcal{A}_{sj}] Z_j; \quad (\mathcal{A}'_j = \mathcal{A}_j - \mathcal{A}_{sj})$$

$$BM_{bj} \leftarrow X'_j[\mathcal{A}_s] // Z_j; \quad [\text{セトニアの生成}]$$

(ii) if $X_j \in \Sigma'$ then broadcast(BM_{sj});

(7) [セトニアの受信]

全 VIS_j (ただし、 $X_j \in (\Sigma' \cup \Sigma'')$) は、全セトニアを受信する。

while (全セトニアを受信してない)

$$\{ BM \leftarrow \text{rec}(Y); \quad BM_{bj} \leftarrow BM_{bj} \cap BM; \}$$

全ビットマップを受信し終えたならば、ビットマップにより X_j を制限し、コード化する [図 5.2].

$$X_j'' \leftarrow X_j' * B\Gamma_j; \quad [\text{ビットマップ制限}]$$

$$X_j \leftarrow X_j'' \wedge (z_j * B\Gamma_j); \quad [\text{コード化, } X_j(\tilde{\mathcal{A}}, \mathcal{A}_j, \Lambda_j)]$$

$\mathcal{A}_R \cap \mathcal{A}_S \neq \emptyset$ なる全 X_R に対して、

$$\mathcal{A}_R \leftarrow (\mathcal{A}_R - \{\mathcal{A}_j\}) \cup \tilde{\mathcal{A}};$$

$$\mathcal{A} \leftarrow (\mathcal{A} - \{\mathcal{A}_S\}) \cup \tilde{\mathcal{A}};$$

$$\Sigma' \leftarrow \Sigma' - \{X_S\};$$

(5) [不要関係の除去]

$$\text{if } (X_j \in \Sigma' \vee X_j = X_S)$$

$$\text{then if } \mathcal{A}_j = \tilde{\mathcal{A}} \text{ (全コード化された)}]$$

$$\text{then } \{ \Sigma' \leftarrow \Sigma' - \{X_j\};$$

$$\text{if } \Lambda_j \neq \emptyset \text{ then } \Sigma' \leftarrow \Sigma' \cup \{X_j\} \};$$

各 $\text{VIS}_j \in (X_j \notin \Sigma' \cup \Sigma')$ に対して、 VIS_j は Q の処理を終了する。

(6) [目標集合処理]

$$\text{if } |\Sigma'| \geq 2 \text{ then go to (4); } [|\Sigma'| = \Sigma \text{ 内の関係数}]$$

[終了]

$\Lambda_R \neq \emptyset$ なる全 VIS_R は、 $X_R(\tilde{\mathcal{A}}, \Lambda_R)$ を VIS_0 に送信する。 VIS_0 は、

X_R を受信する毎に $G \in G[\tilde{\mathcal{A}} = \tilde{\mathcal{A}}] X_R$ を行い、 $G[\Lambda]$ を目標集合とする。□

以上の手順で、 Σ' を候補関係の集合、 Σ'' を非候補目標関係の

集合とし、 \mathcal{A} をコード化属性とする。又 \mathcal{Y} を送信関係とする。

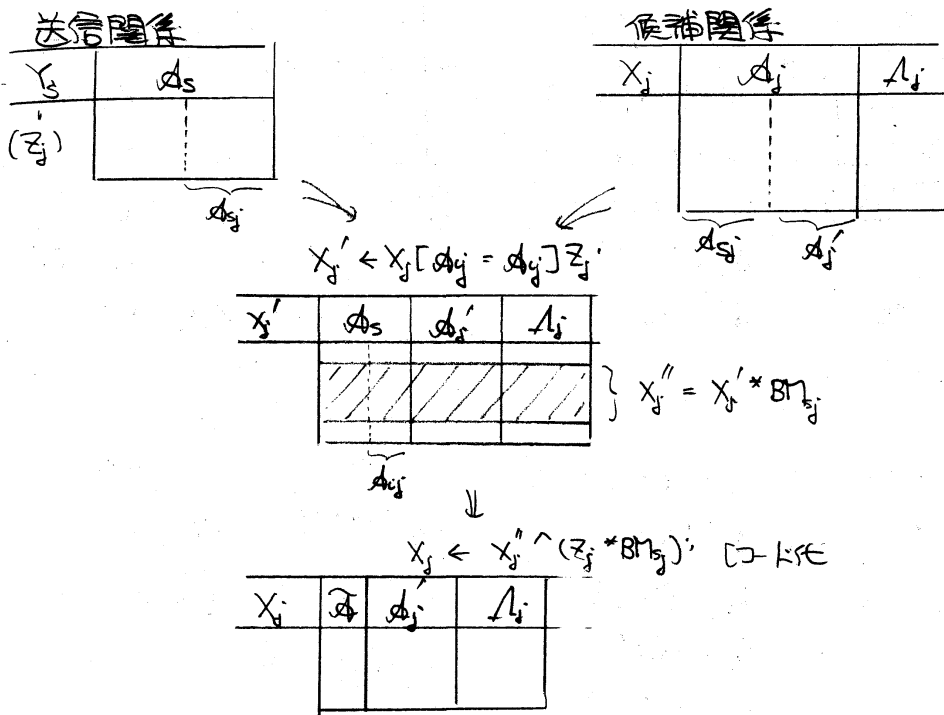
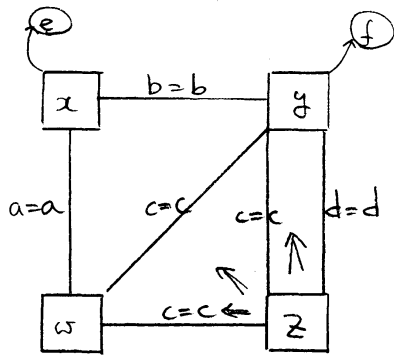


図5.2 結合, ビットマップ制限, コード化

図5.3には、図5.1の問合せに対する B^2 通信手順の適用例を示す。コード化された属性 \mathcal{A} は、 $\log_2 |Z_j * BM_{sj}|$ ビット長あればよいので、1~2ビットあれば十分である。

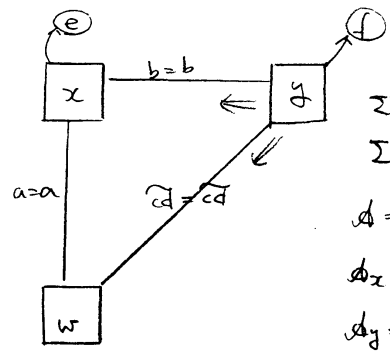
5.4 B^2 通信手順の全通信量と通信時間

B^2 通信手順での全通信量 T と通信時間 R とを考える。 B^2 通信手順で、関係 X_s の $Y_s (= X_s[\mathcal{A}_s])$ が放送されたとする。このとき、 $\|Y_s\|$ と $|Y_s|$ を各 R 、 Y_s の情報量と組数とする。又、 m_s を、 \mathcal{A}_s の $\mathcal{A}_j \neq \emptyset$ なる候補関係 X_j の数とする。 Y_s の放送と、ビットマップの放送の為の全通信量 T_s と通信時間 R_s とは次式で与えられる。



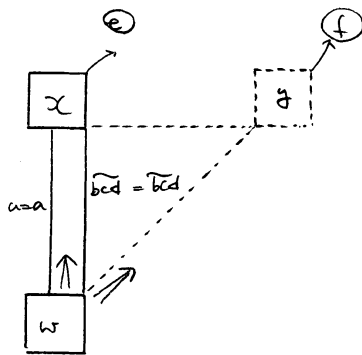
$\Sigma' = \{X, Y, Z, W\}$
 $\Sigma'' = \{\Phi\}$
 $\mathcal{A} = \{a, b, c, d\}$
 $\mathcal{A}_x = \{a, b\}$
 $\mathcal{A}_y = \{a, c, d\}$
 $\mathcal{A}_z = \{c, d\}$
 $\mathcal{A}_w = \{a, c\}$

(1) Z[c, d] の送信



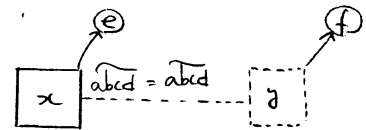
$\Sigma' = \{X, Y, W\}$
 $\Sigma'' = \{\Phi\}$
 $\mathcal{A} = \{\bar{c}\bar{d}, a, b\}$
 $\mathcal{A}_x = \{a, b\}$
 $\mathcal{A}_y = \{\bar{c}\bar{d}, b\}$
 $\mathcal{A}_w = \{\bar{c}\bar{d}, a\}$
 $\bar{x} = \bar{c}\bar{d}$

(2) Y[$\bar{c}\bar{d}, b$] の送信



$\Sigma' = \{X, W\}$
 $\Sigma'' = \{Y\}$
 $\mathcal{A}_x = \{\widehat{bcd}, a\}$
 $\mathcal{A}_y = \{\widehat{bcd}\}$
 $\mathcal{A}_w = \{\widehat{bcd}, a\}$
 $\bar{x} = \widehat{bcd}$

(3) W[\widehat{bcd}, a] の送信



$\Sigma' = \{X\}$ $\Sigma'' = \{Y\}$
 $\mathcal{A}_x = \{\widehat{abcd}\}$
 $\mathcal{A}_y = \{\widehat{abcd}\}$
 $\bar{x} = \widehat{abcd}$

(4)

図5.3 B²C通信手順の例)

$$T_s = p \cdot \{ |X_s[\mathcal{A}_s]| + (m_s - 1) \cdot |X_s[\mathcal{A}_s]| \} \text{ [bit]} \quad \dots (5)$$

(5)式で第一項はY_sの送信と、第二項はビットマップの送信の為の全通信量である。3章の(1)式でα=0とする。従って、このときの通信時間R_sは、次式で与えられる。

$$R_s = T_s / \omega \quad \text{[sec]} \quad \dots (6)$$

ωは、通信網の伝送速度[bit/sec]である。全体問合せQのDDBS処理の為のTとRは、B²C通信手順に従って、各々T_sとR_sを計算したものとなる。

5.5 VIS での処理手法

2 番手順で、最初の差信関係 $X_i (= X_i[d_i])$ の読み出しと、 X_i と X_j との結合は、[AKIM84b] により、各関係 X_i と X_j の物理構造に基づいて行われる。

コード化関係について考える。 $X_i(Q_i)$ を VIS_i の関係とする。 X_i から、2 番手順で得られたコード化関係を $\tilde{X}_i(\tilde{d}_i, d'_i, \Lambda_i)$ とする。 \tilde{X}_i は、図 5.4 の様に物理的に実現される。主記憶内の表 $\tilde{\pi}_i(\tilde{d}_i, d'_i)$

$\tilde{\pi}_i$	\tilde{d}_i	d'_i	Λ_i

} 主記憶

は、 $\tilde{X}_i[d]$ の値 \tilde{t} と、 $\tilde{t}[d]$ ($\tilde{t} \in \tilde{X}_i$) に対応した組 $t \in X_i$ の識別子 d との対応表 $\{(t, d)\}$ である。主記憶が十分大きいときには、 d'_i と Λ_i の値も

X_i	d'_i	Λ_i

(二次記憶)

$(\Lambda_i = d'_i \cup d'_i)$

いときには、 d'_i と Λ_i の値も $\tilde{\pi}_i(\tilde{d}_i, d'_i, d'_i)$ 、又は $\tilde{\pi}_i(\tilde{d}_i, d'_i, \Lambda_i) (= \tilde{X}_i)$ として主記憶内にもたせる。

$\tilde{\pi}_i(\tilde{d}_i, d'_i)$ で、 \tilde{d}_i は $2B$ 、 d'_i は $4B$ とすると、 \tilde{X}_i が 10000 組ある、たととして

も $60kB$ であるので $\tilde{\pi}_i$ は主記憶に保持出来る。

$\tilde{X}_i(\tilde{d}_i, d'_i)$ は、各 VIS の差受信バースに格納出来る JDR の単位に $4B$ から放送される。JDR を受信する毎に $\tilde{\pi}_i$ をサーチして、 \tilde{d}_i が等しい \tilde{X}_i の組識別子を求め、読み出すべきページを見つけ、

同-ページ内の組を一回の入出力で読む様にする。その後、 α_i の値のチェンジを行う。

5.6 送信関係の決定法

Bと手順の(4)で、そのなかから送信関係 X_k は、 X_k の發送、結合、ビットマップ応答の發送、及びコード化の為の全通信量 T_k と、全処理時間量 P_k との和を最小化する様に決定される。 T_k は、(5)式で与えられた。処理時間量は、 α を索引とした時(但し入出力時間は0)の[TAKI784b]で与えられた。

5.7 まとめ

本論文で示したBと通信手順は、(i) LANの發送通信の利用、(ii) 各関係毎に全結合属性値の組の發送、(iii) 發送された結合属性値のコード化による格納量と通信量の減少、(iv) 各VISでの分散制御と動的な送信関係決定とを特徴としている。

6. 分散トランザクション処理方式

LANの發送通信を用いると、従来の一対一網を用いることにより生じるDBSのコミットメント制御、同時実行制御を容易に出来る。

6.1 コミットメント制御

DBSでは、 n 個のプロセス間で、全員一致の場合のみ合意(コミット)が得られたとするコミットメント制御が、DBS処理の基

本と仮定している。従来の一対一網では、[GAT81]に示される様に、(i)ある点 i から、 m 個の点一つ一つに同一の request を送信し、(ii) request の応答を受信し、(iii) 全点が合意したかしないかの結果を m 点に送信せねばならない。この為には、 $3m$ 回の通信が必要になる。LAN の放送通信を用いると、[AKI82] に示す様に、 $m+1$ 回の通信で合意を取れる。

(i) ある点 i から、全点に request を放送する。

(ii) request を受けた各点 j は、OK または Nb の応答を放送する。

(iii) 全点 $(i \neq j)$ は、他の点からの応答を待ち、全応答が OK であれば合意が得られ、一つでも Nb のときは合意が得られなかったとする。□

以上の手順を、トランザクションが m 個の VIS 内のデータを更新する場合に適用すると、従来の二相コミット手順 [GAT81] が $4m$ 回の通信が必要となるのに対して、 $2m+1$ 回の通信で行える [AKI81, 84]。

[B2P 通信手順]

(i) トランザクションを受けた VIS₀ に、更新データを B² の通信手順で求める。求めた更新データを precommit メッセージとして、更新対象の全 VIS_{1, 2, ..., m} に放送する。

(ii) 各 VIS _{i} は、precommit を受信したならば、ログ log_i に更新データを記録させる。完了したならば、precommitted を放送する。

(iii) 各 VIS_i は、他の全 VIS_j からの precommitted を受信する。全て受信したならば、 log_i を用いて DB_i 内のデータの物理更新を行う。完了したならば、acknowledgement を返送する。

(iv) 各 VIS_i は、全 ack を受信したならば、トランザクションを終了させる。

(v) (iii) で、全 precommitted を受信出来ない時、トランザクションをアボートさせる。□

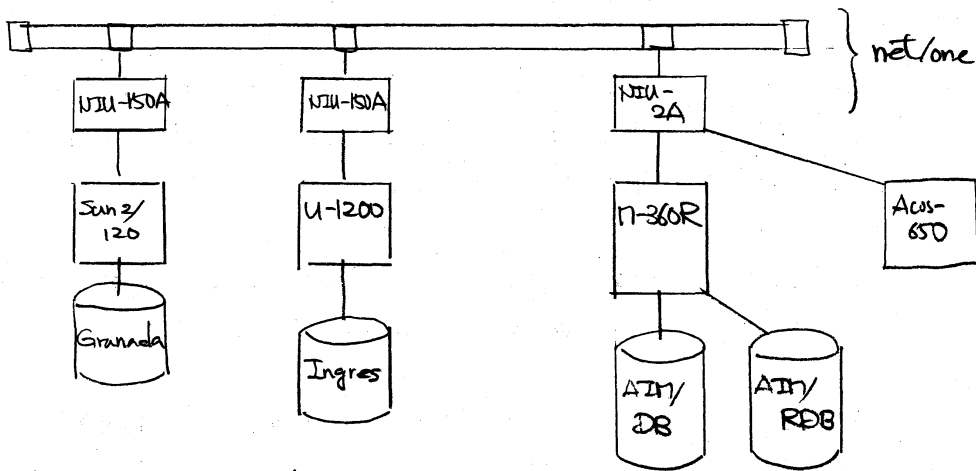
6.2 同時実行制御

同時実行制御とは、複数の トランザクションの競合する実行を一定順序に従うことにより、データベースの整合性を保ち、多くの トランザクション実行の多重度を高めて、システム全体のスループットを向上させる為の制御である。一つの DBS に対して、RDBS では、各 DBS で実行される 競合実行 (同一データの read/write, write/write) を一定順序に従う必要がある。広域ネットワーク交換機では、各点に 情報を送信する時、各通信リンク毎に通信遅延が異なる為に、各点に 複数の トランザクションからの操作実行が一定順序で到着しない点が主要な問題となり、時刻印をトランザクションに設ける オメガ [Omega] 等が考えられている。この為の制御は、通信と処理の負荷をもたしている。一応、LAN は、ある点で送信された情報は、全点に同時に到着する性質を有している[第3章]。即ち、各 VIS_i には、通信網に

送信された順にトランザクションの処理が到着することになり、
 一対一網での通信遅延の相違による問題が生じなくなる。

7. おおりに

本論文では、ローカルエリア網(LAN)と種々の計算機から
 成るDBMSとしてのローカル情報システム(LIS)の構成と、分
 散検索処理と分散トランザクション処理の為の通信手順を示し
 た。LISは、現在図7.1に示す機器構成のもとで、1983年より
 1985年までの3カ年で開発を進めている。



LAN	Net/one	同軸 CSMA/CD	バス 10M bps
Sun 2/120	CPU 68010 MEM 2MB Disk 140MB OS Unix 4.2bsd		
			U-1200 CPU 16 bits MEM 1MB Disk 130MB OS Unix II + 4.1bsd
N360-R	DBMS ATN/DB (CODASYL型) ATN/RDB (待待型)		DBMS Ingres

図7.1 LISの構成

参考文献

- [BERN81] Bernstein, P.A. and Goodman, N., "Concurrency Control in Distributed Database Systems," ACM Computing Surveys, Vol.13, No.2, June 1981, pp. 185-221.
- [HELDG 76] Held, G., et al., "INGRES - A Relational Database Systems," AFIPS Conf. Proc., 1976, pp.409-416.
- [HEVNA 78] Hevner, A., et al., "Query Processing on a Distributed Databases," Proc. of the 3rd Berkeley Workshop, 1978, pp.91-107.
- [METCR 76] Metcalfe, R.M., et al., "ETHERNET Distributed Packet Switching for Local Computer Networks," CACM, Vol.9, 1976, pp.395-404.
- [TAKIM 78] Takizawa, M., et al., "Resource Integration and Sharing on Heterogeneous Resource Sharing Systems," Proc. of the ICC'78, 1978, pp.253-258.
- [TAKIM 79] Takizawa, M., and Hamanaka, E., "The Four-schema Concept as Gross Architecture of Distributed Databases and Heterogeneous Problems," JIP (IPSJ), Vol.2, No.3, 1979, pp.134-142.
- [TAKIM 80] Takizawa, M., and Hamanaka, E., "Query Translation in Distributed Databases," Proc. of the IFIP '80, 1980, pp.451-456.
- [TAKIM 81] Takizawa, M., "分散型データベースシステムJDDBS-IIと通信処理," 電子通信学会 AL81-22, 1981.
- [TAKIM 82a] Takizawa, M., "Distribution Problems in Distributed Databases," JIP (IPSJ), Vol.5, No.3, 1982, pp.139-147.
- [TAKIM 82b] Takizawa, M., Yokotsuka, M., et al., "CODASYL データベースシステムに対する関係インタフェースシステム(LDP-V 1.5)の設計と実現," 情報処理学会 論文誌, Vol.23, No.3, 1982, pp.665-675.

- (TAKIM 82c) Takizawa, M., "放送通信機能を用いた分散型データベースシステムの通信処理の実現について," 情報処理学会 DBMS研31, 1982.
- (TAKIM 82d) Takizawa, M. and Yokotsuka, M., "ローカルデータベースプロセッサ (LDP) の評価," データセマンティクスの理論と実際に関する研究, 京大数理解析研講究録461, 1982, pp.127-150.
- (TAKIM 83a) Takizawa, M. and Noguchi, S., "CODASYL データベースシステムに対する非手続的更新インタフェースの基本概念," 情報処理学会 論文誌, Vol.24, No.6, 1983, pp.
- (TAKIM 83b) Takizawa, M., "Distributed Database System-JDDBS," JARECT Vol.7, Computer Science and Technologies (Kitagawa, T., ed.), Ohmsha and North-Holland, 1983, pp.262-283.
- (TAKIM 83c) Takizawa, M., Yokotsuka, M., and Noguchi, S., "分散型データベースシステムに於ける放送通信を用いた通信処理方式," 情報処理学会「ローカルエリアネットワーク」シンポジウム報告書, 1983, pp.211-218.
- (TAKIM 84a) Takizawa, M. and Noguchi, S., "CODASYL DML 上の非手続グラフ問合せの設計と実現," 投稿中, 1983.
- (TAKIM 84b) Takizawa, M. and Noguchi, S., "放送通信を用いた分散型データベースシステムの統合化手法," 投稿中, 1983.
- (GRAYJ 81) Gray, J., "Notes on Database Operating Systems," Operating Systems: Advanced Course, Springer-Verlag, 1981.
- (YAMAH 82) Yamazaki, H., et al., "A Proposal for Broadcast Architecture Network (BANET)," Proc. of the ICC '82, 1982.