

計算機システムの構成設計支援技法 ISCP/S: Synthesis

北嶋弘行, 大町一孝, 本山博司
(日工製作所 システム開発研究所)

1. はじめに

システム建設の初期段階で、ユーザ要求を満す経済的なシステム構成を導くこと(以下、これをシステム構成設計と呼ぶ)は、システム建設の要諦である。しかし、仮想化、分散化に伴う計算機システムの大規模・複雑化により、このシステム構成設計は困難かつ多くの工数を要するようになってきた。したがって、これを計算機を用いたツールで支援することが有効と思われる。

従来からシステム建設の初期段階に適した性能予測ツールはいくつか提案されてきている。しかし、分散システムのように大規模なシステムの構成設計にこれらを適用しようとするると、次の問題がある。(1)モデルも大がかりになりがちで、モデル作成の工数や計算時間が大きくなる。(2)性能だけでなく、同時に信頼性の予測が必要なものが多い。(3)上記の評価結果、性能や信頼性の目標水準に比へ過不足を生じた場合、システム構成案の改善が必要である。しかし、モデルのパラメータを人間が試行錯誤的に調整する方法では、適正なシステム構成案を得るまでに長時間を要し、しかもこれに至る保証もない。

上記の問題に対処するために、報告者らは分散型を含む計算機システム構成を、性能だけでなく信頼性の面から評価し、さらに改善案を自動作成する技法 ISCP/S (Integrated System for Configuration Planning / Synthesis) を提案し、この TSS 型ツールにも行っている。

本報告の以下では、先ず第2章で提案技法の枠組を示した後に、性能面の支援機能に限って述べる。信頼性面の支援機能は別途報告の予定である。第3章で性能評価機能、第4章で改善案の自動作成機能も、各々述べ、これらの適用例も第5章に述べる。

2. ISCP/S の枠組

ISCP/S は構成設計の多段階を支援する。すなわち、各設計段階で得られる処理費や仕様とシステム構成案第一次を入力とし、

(1) 先ず、性能指標(各処理費の平均応答時間、各資源の利用効率)と信頼性指標(各処理費のアップタイムや障害回復時間、および各サブシステム故障時のシステム処理能力)を見積る。

(2) 上記の結果、目標水準に対し過不足がある場合、この改善案を機能配分と機器構成の観点から自動作成する。

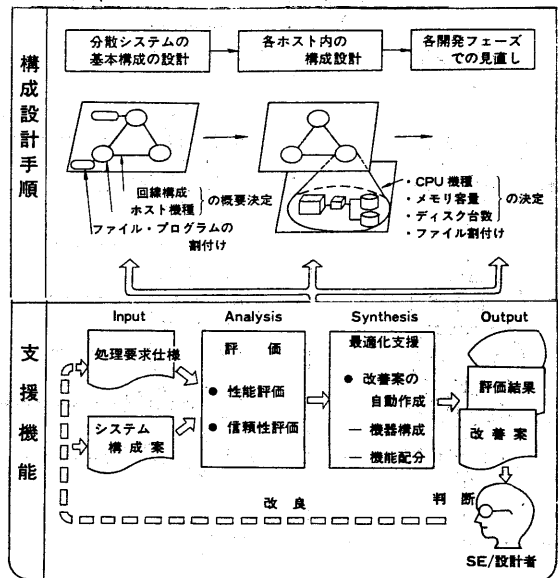


図1 ISCP/S の枠組

3. 性能評価モデル

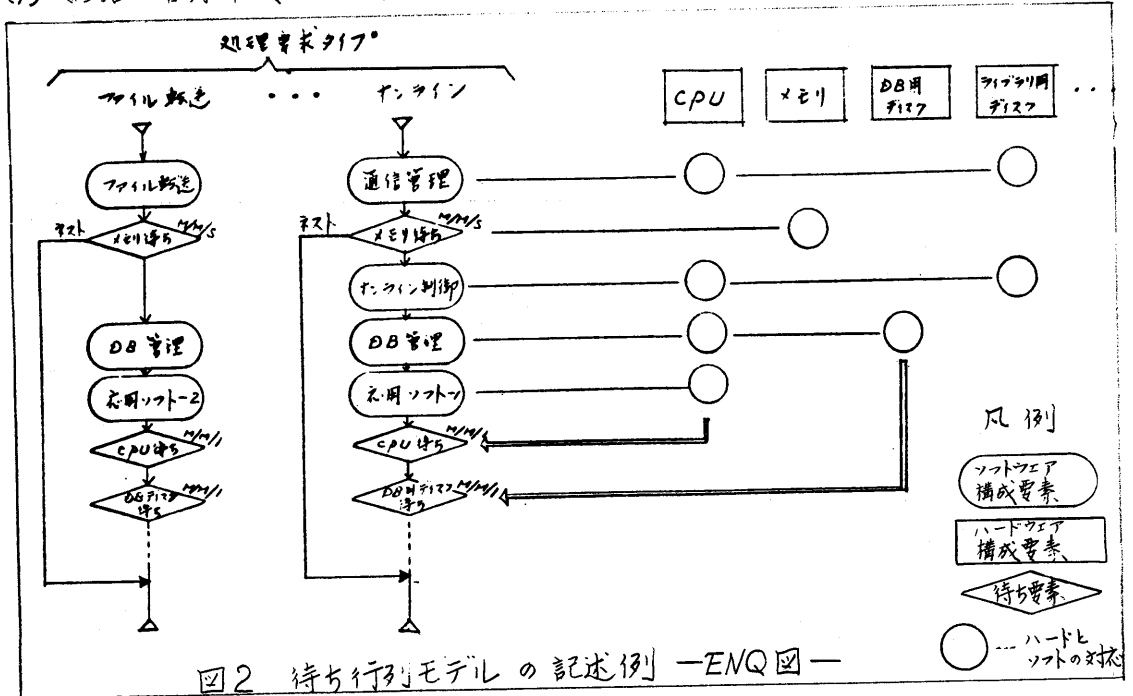
性能評価モデルは、分散システムの場合でも、① TSS対話形式で使える程に計算時間が短く、② 設計仕様に近い形で簡単にモデル化できること、の2点を意識した。このために下記の手段をとった。

3.1 計算時間の短縮

まず、シミュレーション法と解析的手法とも比較し、計算時間の莫から後者を採ることにした。さらに、この中から、モデル化の柔軟性という理由で、分解型の待ち行列網手法を採用した。この手法は、対象システムをM/M/S/N型の解析容易な部分に分解し解析した後、これをネットワークとして再統合する方法である。本手法を計算機ネットワークに適用した例があるが²⁾シミュレーションによる結果と良く一致したと述べられている。

3.2 モデル作成の容易化

- (1) 階層的モデル化 ... 分散システムの構成を、計算機などの各ノード内でのシステム構成を表す下位モデルと、ノード間の回線結合関係を表す上位モデルの2階層で定義する。これによって、任意の分散システム構成の場合でも、下位モデルも多様に組合せることにより、容易に表現できる。
- (2) モデルベース機能 ... 上記のように階層的にモデル化することによって、さらに、ノード内のモデルは、例えば、オンライン+TSS+バッチの混在システムなどのように標準化できることが期待される。そのため、ISC/P/Sではノード内に関する標準モデルを登録・編集できるものとした。
- (3) ハード構成/ソフト構成/処理要求/機能の分離 ... システム内の要素を上記の3つに明示的に分離して定義できるようにした。これによって、処理要求仕様やシステム構成仕様などに近い形でデータ化できるようとした。
- (4) 経路の自動作成 ... 分散システムのモデル作成で面倒なことの1つは、各



理要求が同じる経路の定義である。しかも、この経路は機能の配置を変更することにより全く異なったものとなる。ISCP/Sでは、機能配分および回線構成の情報から、各処理要求の経路を最短経路アルゴリズムで自動作成している。

(5) 待ちモデルの記述方式の提案 ... 待ち行列網手法に基づいてモデルを作成する場合のサポートとなるように、モデル記述方式を新たに提案した。従来、シミュレーションモデルのためには、GPSS用ダイアグラム、E-net³⁾などの記述方式が用いられている。一方、解析的な待ち行列モデルに対しては明確な記述方式は見あたらなかった。提案した記述方式を図2の例に沿って概説する。

① 処理要求を処理要求タイプ〔注1〕に分類し、処理要求タイプごとに、これを使用するソフト構成要素、および待ちの発生箇所(待ち要素と呼ぶ)の列を作成する。各待ち要素には待ちモデルの特性(例: M/M/1, 処理要求間の優先スケジューリングを行うか否か、等)を区別する。また、各ソフト構成要素については資源使用量(CPU実行ステップ数、I/O回数等)を記す。

〔注1〕処理要求タイプ... アプリケーションに依存する部分(DB, AP等)を除くと、処理内容で待ち要素を同一とみなせる処理要求群を指すものとする。例えば、オンライン、TSS、バッチ等の区別が相当する。

② 2次資源に関する待ち要素については、その占有時間に対応するソフト構成要素や待ち要素をネスト記号で括る。

③ ハード構成要素の行も作成し、各々の処理速度を明らかにする。

④ 上記①③より得た行列にて、各ソフト構成要素が使用するハード構成要素、および各待ち要素が対するハード構成要素、を分るようネットワークを記す。以上の提案した記述方式の概要であり、これをENQ図(Extended Network Queues 図)と呼ぶこととする。

4. 最善手作成機能

4.1 機能配分モデル

分散システムの計画・設計において、ファイルやプログラムの各々をどのサイトに割り付けるか(すなわち機能配分)は重要な課題である。機能配分の仕方によって通信コスト、蓄積コスト、および障害に対する処理の続行可能性(可用性)が異なる。従って機能のコピー数の増加は、必要な機能がより近くで得られるため参照の通信コストを減少し、かつ、計算機や通信回線の障害時に、別計算機のコピーで代替できる可能性が増すため可用性の向上が期待できる。反面、コピー数の増加は、蓄積コストとコピー間の奔合をとるための更新コストの増大を招く。そこで、機能配分モデルでは、これらの評価尺度間のトレードオフを最適にする機能配分案を計算し提示する。

(1) 機能配分モデルの定式化 機能配分モデルは従来から幾つか提案されてきているが、提案手法に比べ、提案手法は次の特長をもつ。

① 従来、各機能は独立と仮定してきた。しかし、実際の分散システムでは、機能間にデータや制御が受け渡されるため、独立でない。提案手法は、機能間の関連を考慮できる。

② 可用性の制約を扱える。

本モデルの定式化の概要も下記する。(詳細は紙面の制約で省略する)

〔決定変数〕 機能 i を サイト j に配置するの否 $x_{ij} = 0$ または 1 。
 〔条件〕 処理要求仕様 (各処理要求の頻度, 必要とする機能の系列, 機能間のデータ転送量), 各機能の蓄積容量, 蓄積コスト単価, 通信コスト単価 など。
 〔目的関数〕 参照通信コスト + 更新コスト + 蓄積コスト $\rightarrow \min$ 。
 〔制約条件〕 各処理要求の不可用率 \leq 許容値。
 ここで, 各処理要求の不可用率は, 任意の機能配分 $X = \{x_{ij}\}$ について, 各システム構成要素のアベイラビリティに関する信頼度関数で表す。
 制約条件として, 上記以外に, 各サイトの処理・蓄積容量, 各機能の配置可能なサイト集合, 各処理要求の通信遅れ時間などを制約とすることができる。

(2) 機能配分モデルの解法

上記の定式化内容は, 2値変数 $X = \{x_{ij}\}$ に関する非線形0-1計画問題となる。しかし, 本定式化では機能間の関連を許しているため, X の規模は [サイト数] \times [機能数] と大きい。したがって, 各機能も独立と仮定する従来の最適解法をそのまま適用したのでは, 計算時間が膨大となる。そのため, 新たに次のように着目した近似解法を提案した。

- ① 機能集合のうち, 任意の1機能を着目し他機能の配置は既知と仮定すると, 従来の各機能の独立の場合の最適解法が適用できる。
- ② 全機能を一定の順に並べ, ある実行可能解 X の各 x_{ij} を, 各機能の順に, 上記①の実施とこの結果による x_{ij} の変更を行う。ここで, 機能の順序づけは, 該機能の配置が他機能の配置に及ぼす影響度によって行うものとし, これによって, 解への収束の効率化を図った。この指標として, 該機能と他機能間での総データ転送量を用いた。

4.2 機器構成最適化モデル

本モデルでは, 性能目標値を最も経済的に達成するための, 各一価収要素の処理容量を決める。本モデルでは, 機器配分・ソフト構成は既知と仮定している。性能目標値 T は, 参照通信回線網 T と T の容量選択と T との関係で表される。しかし, 本モデルと通信回線網を統合して定式化する場合, ソフトウェア構成最適化モデルは, 他に別なモデルが必要。

(1) 機器構成最適化モデルの定式化

〔決定変数〕 y_{kl} (ハード構成要素 k に割り当てる機器 l の台数 (非負の整数))
 〔条件〕 処理要求仕様, 機能配分案, ソフト構成, 各一価収要素の性能単価など。
 〔目的関数〕 機器コストの総和 $C = \sum_{k,l} y_{kl} C_{kl}$ $\rightarrow \min$ 。
 ここで, C_{kl} 機器 l の価格 (¥/台/年)。
 〔制約条件〕 ① 各処理要求の平均応答時間 \leq 許容値
 ② 各回路の利用率 \leq 許容値
 ③ 各一価収要素 k に対する割り当て可能な機器の制約
 ④ 各一価収要素 k の配置台数に対する制約

上記の制約条件中, 平均応答時間および利用率は前述の性能評価モデルを用いて計算する。

(2) 機器構成最適化モデルの解法

上記の定式化内容は, 整数 $Y = \{y_{kl}\}$ に関する非線形整数計画問題となる。こ

の最適解法としては分岐限乗法など知られている。しかし、本問題の場合、変数 Y の個数は〔ハード構成要素数〕 \times 〔各構成要素に対する機種の数〕の規模となるため、分岐システムのように大規模なシステムの場合には、最適解法では計算時間が膨大になる恐れがある。そこで、新たに近似解法を考案した。この近似解法は最急傾斜法の考え方に基くもので、解の改良幅が最大の方角に次々し機種や台数を置き換えていく方法である。この解法は大別して次の4ステップからなる。

〔ステップ1〕2次資源以外の資源について、利用率が許容値を超えるものがある場合、これを許容値以下とするように機種や台数を上げる。

〔ステップ2〕2次資源について同様の操作を行う。

〔ステップ3〕平均応答時間が許容値を超える処理要求がある場合には、次の操作を行う。すなわち、各ハード構成要素について、この機種や台数を上げた場合のコスト増当り制約不満足の処理要求などの程度判約に近づくかを表す指標を計算し、これが最大となるハード構成要素の機種や台数を上げる。この操作を、全ての処理要求が平均応答時間の判約を充すまで繰り返す。

〔ステップ4〕平均応答時間の許容値に対する余裕が一定量を超える場合には、次の操作を行う。すなわち、各ハード構成要素について、この機種や台数を下げた場合に、平均応答時間 \times 利用率の判約に対する余裕もできるだけ残したうえで、その程度コストを削減できるかを表す指標を計算し、これが最大となるハード構成要素の機種や台数を下げる。この操作を、判約に対する余裕量が一一定値以上の間繰り返す。

5. 適用例

以上のモデルは、以下の各サブモデルを仮想的な分散システムに適用した。

5.1 性能評価モデルの適用例

- (1) 適用対象 ある販売会社では現在の集中型から分散型に移行する計画である。これに対してシステム設計者が作成したシステム構成第一次案は次の通り。
 - ① オンライン業務は在庫管理や出庫計画を中心に地域性が強い。そのため、全国を6分割し、各々に分散機を配置する。
 - ② 上記の各分散機から統計情報を収集・加工し、マクロ的計画を立案するために、マクロホッチ処理用計算機を配置する。
 - (2) 入力データ 回線構成、CPU機種は図4に示す通りである。また、各サイトのシステム構成と待ちモデル、処理要求は図3の各々(2),(3),(7)に示すごとく定義する。オンライン処理要求に対しては、平均応答時間を1秒以下にするという判約を課した。この他に、各ハード機種の処理速度、価格、許容利用率、および各ソフト機種の資源使用量の定義が必要であるが省略する。
 - (3) 計算結果 性能評価モデルの計算結果を図4に示す。(各サイト内についても同時に計算しているが紙面の都合で省略する。)利用率や平均応答時間の目標を達成できていない箇所は**印で指摘している。
- 本例の性能評価に要した時間は約1.5秒(M180, CPU時間)であった。

5.2 機器構成最適化モデルの適用例

- (1) 適用対象 上記5.1と同一の例を対象とする。
- (2) 入力データ 性能評価モデルの入力データがそのまま使用できる。

(3) 計算結果 計算結果を図5中に示す。図中で、機種や台数の変更が起った箇所も →印で指摘している。同図中には、機器構成を変更した場合の利用率や平均応答時間も併せて出力されている。この場合、これらの目標値が全て達成されていることは言うまでもない。

ここで、システム全体を一挙に最適化の対象とすると、構成要素数が多いため計算時間の莫大の問題がある。そこで段階的最適化の接近方法をとった。すなわち、

- ① 先ず、分散システム全体としての基本構成(CPU 機種、回線容量など)のみを最適化の対象とする。この時点では、各サイト内の構成要素の利用率は、各々の許容値に等しいものと仮定して平均応答時間を計算する。
- ② 次に、①の結果を受けて、分散システムの基本構成は固定、任意のサイト以外のサイト内構成要素の利用率は許容値に等しいという条件で、該サイト内構成を最適化する。これを、各サイトについて次々と繰り返す。

(1) 処理要求

SERVICE REQUEST	TYPE	TRAFFIC (/HR)	MAX. TAT (SEC)	REQUIRED FILE & PROGRAM
TOKY-OL1	ON-LINE	2908.7	5.0	TOKYO TCE2 >> 0.0 BIT >> TOKY-DB1 >> 4060.0 BIT >> TOKYO TCE2
TOKY-OL2	ON-LINE	1649.2	5.0	TOKYO TCE3 >> 0.0 BIT >> TOKY-DB1 >> 4060.0 BIT >> TOKYO TCE3
TOKY-OL3	ON-LINE	1978.9	5.0	TOKYO TCE4 >> 0.0 BIT >> TOKY-DB1 >> 4060.0 BIT >> TOKYO TCE4
TOKY-B	BATCH	1978.9	9999.0	TOKYO TCE1 >> 0.0 BIT >> TOKY-DB2 >> 4060.0 BIT >> TOKYO TCE1

(2) ハード構成

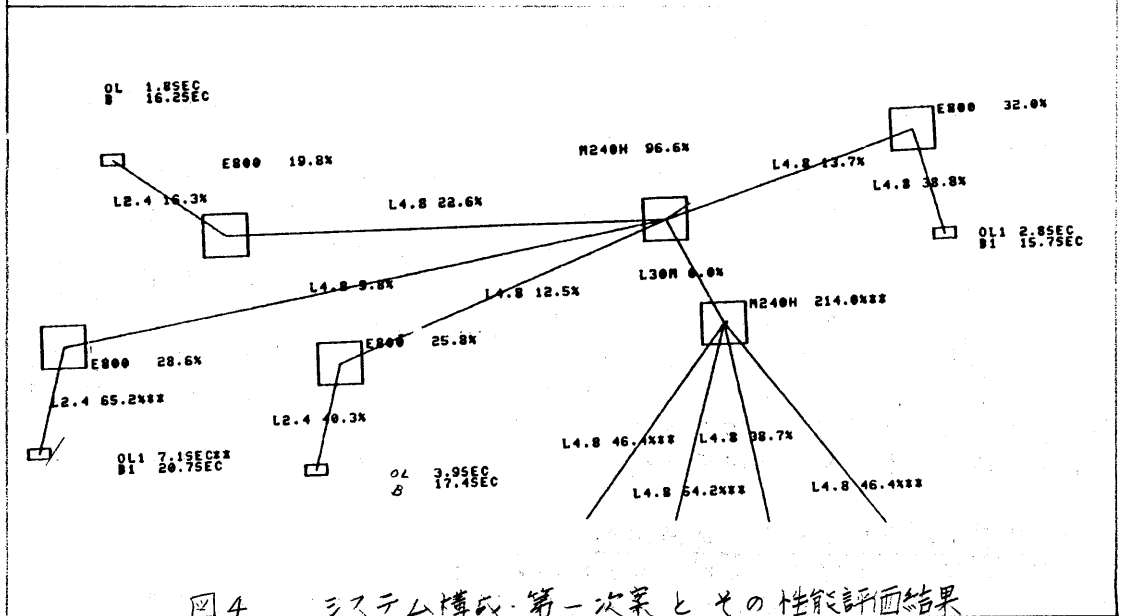
RESOURCE TYPE	MODEL OF DEVICE	NUMBER OF DEVICES
ICPU	M240H	1
IMEMORY	IMEMORY	8
ICHANNEL	ICH-3.0	1
IDISK	HB576	4
IDISK	HB576	4
ICCP	ICCP	1

(3) 待ち行列

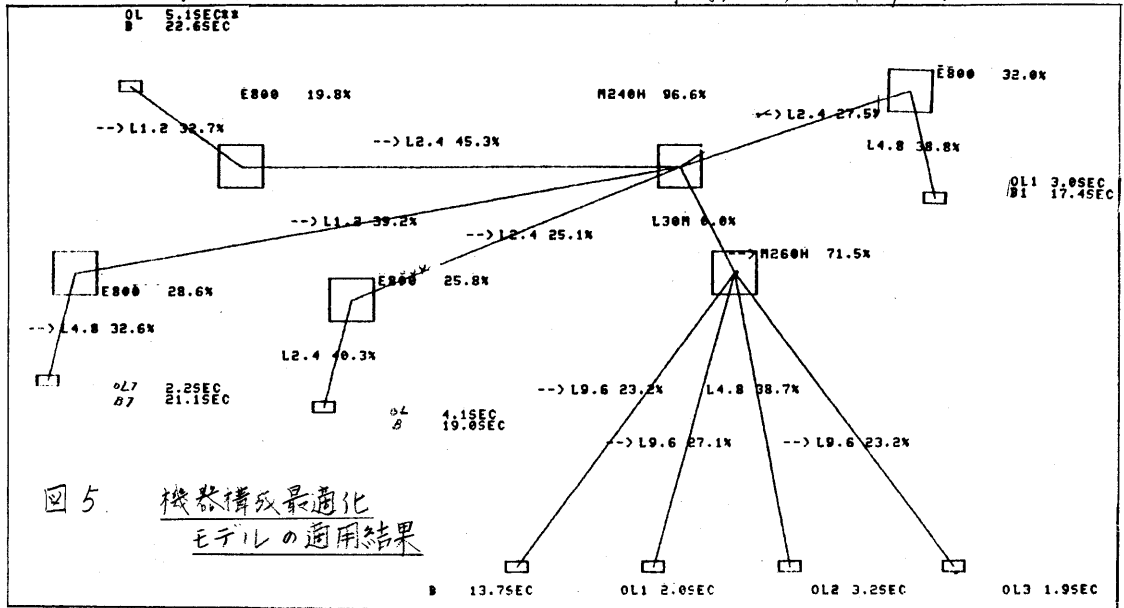
QUEUE ID.	MODEL	S	N	NEST LVL	CORRESP. HARD ID.
WMEM	M/H	S	9999	2	MEM
WCPU	M/H	S	9999	1	CPU
WDISKLB	M/H	S	9999	1	DISKLB
WDISKDB	M/H	S	9999	1	DISKDB
WCH	M/H	S	9999	1	CH

TRANSACTION TYPE : ONLINE

図3 入力データ



段階的最適化による計算時間は、本例の場合、①の段階で約4分、②の段階で1サイト当り約2秒(ともに、M180のCPU時間)で、合計約5分であった。



5.3 機能配分モデルの適用例

- (1) 適用対象 (a) トポロジー … 対象システムのトポロジーを図6に示す。サイト間の通信コスト単価は図6中のごとくとする。蓄積コスト単価は各サイト一様に9(10/KB/月)とする。
- (b) 処理要求 … 処理要求の種類は全部で20だが、その一部を表7に示す。この表中で、機能間の転送データも太線で示しているが、後にこの量も変化させ感度分析を行う。
- (2) 計算結果 機能配分結果を図7(2)に示す。図中で「7」は配置されるサイトを表す。本例の場合、計算時間は約8秒(M180 CPU時間)であった。

表7 機能配分対象の処理要求

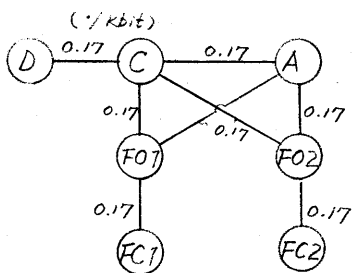


図6 機能配分対象のトポロジー

処理要求	頻度 (件/時)	シーケンス および 転送データ長(単位:106bit)
SR1	300.	D ¹⁰⁰ → FORD → U ²⁰⁰⁰ → D
SR2	300.	C ¹⁰⁰ → FMRK → U ²⁰⁰⁰ → C
SR3	300.	C ¹⁰⁰ → FMRK → Q ²⁰⁰⁰ → FORD → Q ²⁰⁰⁰ → FPRD → U ²⁰⁰⁰ → C
SR4	600.	C ⁰ → FPRD → Q ⁴⁰⁰ → FSCH → U ⁷ → C
SR5	200.	C ⁰ → FPRD → Q ⁴⁰⁰ → FALL → Q ⁴⁰⁰ → FFP → Q ⁴⁰⁰ → F01

U: 更新処理, Q: 参照処理

