

# 分散高速情報ストレージ実装リアルタイムプッシュ型 システムの開発

Development and Trial of Real-Time Push Interface Application System  
Used Distributed High-Speed Access Information Storage

小川 浩二<sup>1)</sup>      衛藤 敏夫<sup>2)</sup>      中村 洋一<sup>3)</sup>      花崎 雄二<sup>4)</sup>  
Koji OGAWA      Toshio ETO      Youichi NAKAMURA      Yuji HANASAKI

- 1) 株式会社数理技研 ソリューション第3部  
(〒106-0022 東京都新宿区新宿四丁目1番9号 新宿コースビル E-mail: ogawa@suri.co.jp)
- 2) 株式会社数理技研 ソリューション第3部  
(〒106-0022 東京都新宿区新宿四丁目1番9号 新宿コースビル E-mail: eto@suri.co.jp)
- 3) 株式会社数理技研 オープンシステム研究所  
(〒400-0111 山梨県中巨摩郡竜王町竜王新町1901-1 E-mail: youichi@suri.co.jp)
- 4) アルビス株式会社 リテールサポート部  
(〒939-0402 富山県射水郡大門町流通センター水戸田三丁目4番 E-mail: yhanasa@albis.co.jp)

**ABSTRACT.** This paper presents “Development and Trial of Real-Time Push Interface Application System Used Distributed High-Speed Access Information Storage”. This system applies to Retail-Support, Super Market Store Operation. On the assumption High-Speed access to Business-Data, this system realized push type user interface in addition to pull type user interface.

## 1. 背景

今日企業においては、様々なビジネスの意思決定を迅速に行うために、企業活動の結果として生まれるビジネス・データを、効率的に情報をロストすることなく蓄積・管理でき、また蓄積・管理されたデータに対して効果的にアクセス、分析できる情報システムが求められている。

我が国における流通業においても同様のことが言える。

特にスーパーマーケットをはじめとする店舗活動を行う関係者にとっては、POS(Point Of Sales)情報等の大量のデータを効率的に蓄積し、迅速かつ効果的に分析活用することで、発注・店内加工・販売促進活動等の業務精度の向上に繋げ、結果としてロス(機会ロス、廃棄ロスなど)を軽減させることが急務である。

しかし流通業における、システム導入・活用の現状は、以下に掲げる「流通業における情報システムの問題」を抱え、こうした問題を改善し、効率的なデータ利用を行うためには、

- ・ 安価な費用投資で効率的にビジネス・データの情報蓄積ができること
- ・ 蓄積した情報に対して効果的な分析ができること
- ・ 利用者に対して有益な情報を簡易な操作で短時間で得られる仕組みを持つこと

が課題となっている。

### 「流通業における情報システムの問題」

大量のビジネス・データを蓄積するデータウェアハウスを構築して、様々な分析ツールを用意し、バスケット分析などで著名なデータマイニングを実施して、ビジネスに生かす必要性が喧伝されているが、現場での積極的な活用事例は少ない。

一般に市販RDBMSやERPパッケージを基にしてデータウェアハウス構築し、データを有効活用するためには多大な投資コストがかかり、このため情報をまるめることなく蓄積してデータ分析を行っている企業は、ごく一部の大手企業だけに限られる。

大半の企業は費用対効果等の関係から、流通ビジネス・データ(POS、発注情報など)を日単位などで、まるめて管理することとなり、日常業務現場で意思決定を行うための、時系列単位等のきめ細かい分析ができていない。

また流通現場では、ビジネス・データの分析・活用の意識を持っていても、利用者がコンピュータ操作などに精通しているレベルにないことが多く、また現場業務に追われコンピュータと対峙する時間も満足にとれないため、大量の情報から有益な情報を引き出すことが、かなりの困難な作業となっているのが現状である。

## 2. 目的

本システム開発の目的は、「1. 背景」に述べた流通業における課題を解決するために、

- ・ 安価な費用投資で効率的にビジネス・データの情報蓄積ができること
- ・ 蓄積した情報に対して効果的な分析ができること
- ・ 利用者に対して有益な情報を簡易な操作で短時間で得られる仕組みを持つこと

を満たす、アプリケーションシステムを構築することである。

ここで、情報の蓄積と、蓄積した情報への高速アクセスを安価に実現する分散型で高速アクセスを可能とする情報ストレージ実装技術、データマイニングアプローチによる分析技術を基盤として、流通業におけるリアルタイムプッシュ型インタフェースを搭載する店舗活動支援システムを開発し、実地検証などを通して、現場業務での有効性を検証する。

## 3. システム構成

下図1は、システム構成を表したものである。

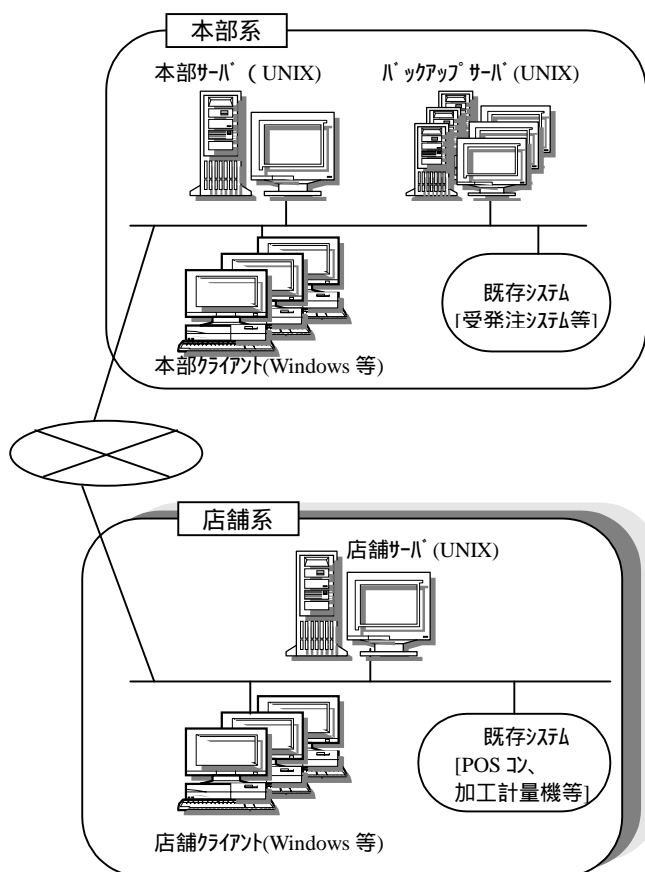


図1 システム構成図

本システムでは、店舗で発生するトランザクションデータは店舗サーバで、本部で発生するトランザクションデータは本部サーバで分散管理蓄積する。

また、数店舗に1台の割合でバックアップサーバを設け、データ二重化を行い、店舗サーバの故障時には代替機として稼働する。

## 4. 開発システム概要

本開発システムは、中堅流通業（スーパーマーケット）の販売計画、発注、加工、販売、分析からなる店舗活動を支援するために、

- ・ POS コン、受発注、経理、人事などの他システムから取得する、売上・加工・入荷の情報や、商品マスター、人事マスターなどのビジネス・データ
- ・ 本システムが提供するアプリケーション画面から投入される、販促プロモーション、コーザル情報、プッシュメッセージ出力のためのパラメータ値等の入力データを、店舗、本部にそれぞれ分散配置されたサーバマシンの主記憶メモリ上に、情報を保証しながら蓄積し、また分散型で管理される情報に対するシームレスで、且つメモリアクセス性能を活かした高速アクセスを実現する。

この蓄積された情報への高速アクセスを前提に、

- ・ ニューラルネットワーク理論、統計解析理論に基づいたデータマイニングツールを利用した、時系列予測、相関分析、クラス分類などの分析処理
- ・ 蓄積された過去の情報と、刻々と取り込む入力データに対応して、発注、販売、加工などのE値（下記（注）を参照）を算出するシミュレーション
- ・ アプリケーション画面から指示された要求に従った、ビジネス・データへの検索処理と、その検索結果に対するデータ加工

などの処理を行う。さらに、

- ・ アプリケーション画面からの指示要求に従った、ビジネス・データの検索処理結果の画面出力
- ・ 分析処理や、E値シミュレーションと、プッシュメッセージ出力のためのパラメータ値などによって導き出される、アラーム・ガイダンスメッセージの画面への出力
- ・ 本システムで蓄積される発注量、売上、就業情報データの、受発注、経理、人事などの他システムへの出力

を行う。

（注）E値

Estimate 値の略。本システムで扱う過去データを元に、各種コーザル情報、スケール情報で補正を施した予測値。販売E値、発注E値などがあり、特に、1日の時間帯別に算出する予測値をE曲線と呼ぶ。

## 5. 分散高速情報ストレージ

本システムが実現するリアルタイムプッシュ型システムのコアとなる、「分散高速情報ストレージ機能」について解説する。

昨今の情報システムにおけるデータ（ビジネス・データ）は日々の業務処理の結果（単なる副産物）ではなく、それ自体価値のあるものとみなされることが圧倒的である。

つまり、今日のビジネス・データは、ビジネスの現況そのものを表現・反映しているのであり、これを過去のデータと組み合わせれば、現在のビジネスの位置と今後の方向が判明することが多い。

流通業で考えると、典型的にはPOSデータがあり、POSデータはアイテム（商品）単位にトランザ

クシオン量を考えると、中堅食品系スーパーマーケット1店舗（～年商20億円）では、年間約10,000,000レコードがトランザクションとして発生する。

この大量のトランザクションデータを、データベースシステムに蓄積保存し、ビジネスのための意思決定などのために利用する場合、スーパーマーケットの発注業務や、店内加工、販売値引のタイミングなどの意思決定に利用されることが考えられ、利用者は当然オンラインでの情報引き出しを要求する。

市販のRDBMSを使ってこの大量の情報蓄積（データウェアハウス）を構築する場合、大容量のデータ蓄積を前提に考えると多大なハードウェア投資は必至である。

また市販RDBMSは一般性を持たした関係データベースとして実装され、格納する情報へのアクセス効率を向上させるために、様々なアーキテクチャ上の工夫がなされてたが、一般性を追求するために、データ格納効率と、そのアクセス性能は、ある面では犠牲にならざるを得ない。

従って性能面と、費用面から考え、現行の大半のシステムは、発生したトランザクションをある単位で括り（集計データとして）データ保存するか、数日分のデータしか情報蓄積せず、情報システムを構築している。

このようなデータ蓄積の現状は、貴重なデータそのものが欠損・欠如しているために、効果的な分析が成されないことになり、大企業と比較すれば明らかに投資腕力の劣る中小の企業は、規模のみならず情報システムでも格差を縮めることが困難となる。

ここで、本システムでは、情報を発生地点でそのまま収集し、ネットワークを介して情報共有し、コンピュータのアクセス性能が最も高速なメモリに情報蓄積することによるアクセス性能の劇的な向上をもたらす、分散型情報ストレージを実装した。

メモリ上では、いわゆる「ダイナミックデータ構造」が可能となるため、より柔軟性の高いデータ格納方法が実現でき、変更や進化に強いシステムの構築が可能となる。

一般に、今日安価に大量に入手可能なメモリであるDRAMは、揮発性メモリであり、データベースやデータウェアハウス用には不適格だとする考え方が多々あるが、本システムでは、メモリのアクセス性を活かしながら情報の復旧性を保証すること及び分散されたメモリ上の情報をシームレスにアクセス可能とし、必要なものについては同期を保証する機能を具備した。

メモリデータベースのアーキテクチャを実現するためには、

- (1) 対象となる業務で発生・管理する情報量を正しく推定する
- (2) 管理する情報を完全に正規化してコンパクトにモデル化する
- (3) メモリ格納の利点を最大に生かし、データ構造に柔軟性を保つ
- (4) 処理（アクセス）時間コストの正しく推定するといったノウハウを持って、対象・適用業務毎にモデル設計する必要がある。

流通業スーパーマーケットにおける店舗活動支援に関する情報をシミュレーションすると、日別取り

扱い商品数2万点、年商20億円店舗の、販売活動支援のための情報は、3年分を保持するのに、1GBバイト未満のメモリで足りる。

これは、国内外超大手事業体のデータウェアハウスの持ち方の同じ程度の情報を1/10のサイズで持つことになる。

さてこのような分散高速情報情報ストレージを前提にアプリケーションシステムを考えると、様々なパラダイムチェンジが現実となる。

データベース処理が超高速処理を可能となることから、加工・販売の刻々の変化に完全にリアルタイムに対応した評価が可能であり、システムが高速データアクセスを利用したシミュレーションを行い、利用者に対してアクション指示（online push）を出力することが可能となる。

これは従来システムを利用者が条件をシステム投入して、結果をデータベースから引き出すPull型システムとすれば、システムがある条件下で刻々の全データにアクセスし、シミュレーションを行い、利用者に対して結果などを出力するPush型システムと名づけることができる。

また、データベースアクセスの限界を超えた情報システムの業務への展開がある。

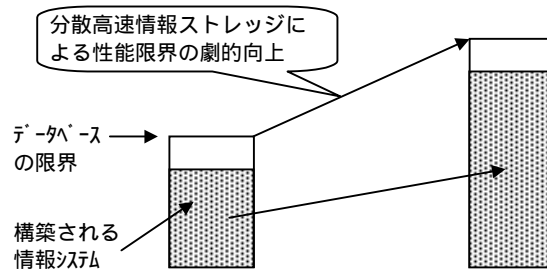


図2 システム限界の突破

上記図2で示すように現状システムは、データベース限界（資源・性能・費用）の範囲の中でそれを100%活用すべくシステム構築されているわけであるが、本開発システムのアプローチではこの限界を突破して、アクセス性能の観点で見れば数千～10000倍の向上が、従来型のハードウェア投資より確実に安価に提供できる。

ここから従来型では、業務モデルをコンピュータシステムで構築しようとも

「無理だ、できない、時間がかかって使えない」といった案件が、現実にシステム構築できる可能性が広がる。

端的な例が、1店舗あたり年間約10,000,000アイテム(10M)がトランザクションレコードとして発生するスーパーマーケットのビジネスデータをすべてそのままサーチして、異常な動向を示しているアイテムを検出するような発想が従来型システムではありえない。

10Mレコードをアクセスすることを考えると、1000TPSとしてもレコードサーチに10000秒(=2.8時間)かかり、システムの実現という意味で除外されてきた。（本システムのアプローチでは、同様の処理は1秒～10秒である）

この性能を持って、これまで「無理だ、できない、時間がかかって使えない」と、SEが説明し、コンピュータの性能をこんな物かと勝手に線引きした業務

SE が、本当にやりたかったことを実現できる可能性が持てるのである。

## 6 . 開発機能

本開発システムの開発機能を、下図 3 に示す。

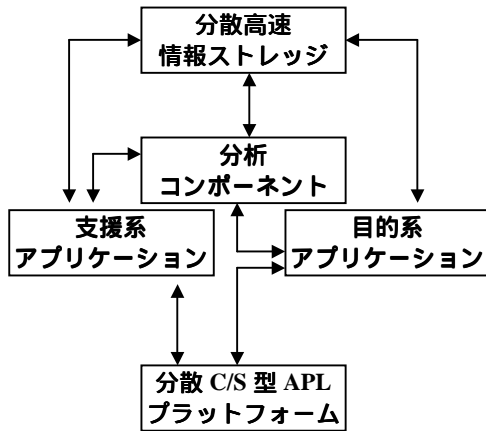


図 3 開発機能関連図

### 6 . 1 分散高速情報ストレージ

- (1) **情報ストレージ情報管理エンジン機能**  
情報ストレージに格納する情報をトレーサブル実時間タグ付き情報として、サーバマシンのメモリ上に展開し、蓄積管理する機能。
- (2) **情報ストレージ情報分散管理機能**  
ストレージ情報の分散管理 [ 店舗毎に発生する情報は自店舗に設置されるサーバマシンで管理し、その集約情報を本部に設置されるサーバマシンで管理を行い、分散管理されたメモリ上の情報のデータ同期を保証する機能。]
- (3) **情報ストレージ情報保証機能**  
2 サイトシステム (サーバが 2ヶ所に置かれ、一方のサーバが何らかの要因で機能停止しても、もう一方のサーバが最新情報を保証しつつ業務を引き継ぐシステム) を保証する機能。
- (4) **情報ストレージアクセス API 機能**  
アプリケーション層が、ストレージ情報に対して更新・検索アクセスするインターフェースを提供する機能。
- (5) **情報ストレージ運用ツール機能**  
障害発生時等の、バックアップサイトへの切替えや、障害解析を行う情報ダンプツール機能。

### 6 . 2 分析コンポーネント

- (1) **ニューラルネット系データマイニングツール機能**  
ニューラルネットワーク理論を用いて本機能において作成するニューラルネットワークモデルを基にデータの分析を行ないデータマイニングを実現する機能。
- (2) **統計的データマイニングツール機能**  
統計的データ解析手法を用いて、ストレージ情報より知識発見 (マイニング) をビジュアル的に行う機能。

### 6 . 3 分散 C/S 型 APL プラットフォーム

- (1) **アプリケーション通信制御機能**  
アプリケーション層で稼動する C/S 型プログラムの通信制御を行うライブラリを提供する機能。
- (2) **クライアントアプリケーションミドルウェア機能**  
C/S 型プログラムのうち、クライアント側で稼動するプログラムが利用する GUI 制御を行うライブラリの提供とクライアントプログラムを管理するミドルウェア機能。
- (3) **サーバアプリケーションミドルウェア機能**  
C/S 型プログラムのうち、サーバ側で稼動するプログラムが利用するユーティリティ関数の提供及びサーバプログラムを管理するミドルウェア機能。

### 6 . 4 支援系アプリケーション

- (1) **ストレージ情報 情報取込インタフェース機能**  
既存システムの情報を情報ストレージへストレージ情報として追加更新する機能と、既存システム以外からの情報をユーザ自身が追加更新する機能。
- (2) **学習インタフェース機能**  
プッシュ型アプリケーションのためのパラメータ値 (ルール情報) のメンテナンス処理を行う機能。
- (3) **マスターメンテナンス機能**  
ユーザが情報ストレージのマスター情報を参照、更新する機能。

### 6 . 5 目的系アプリケーション

- (1) **発注支援サブシステム機能**  
店舗、本部での商品発注を処理、管理するアプリケーション機能。
- (2) **加工支援サブシステム機能**  
店舗の店内加工量の予測シミュレーションと、モニタリングを行うアプリケーション機能。
- (3) **販売支援サブシステム機能**  
店舗の販売量の予測シミュレーションと、モニタリングを行うアプリケーション機能。
- (4) **販売分析支援サブシステム機能**  
販売分析 (マーチャングデザイン) を行うアプリケーション機能。
- (5) **販売計画支援サブシステム機能**  
本部、店舗による販売計画 (予算計画、販促計画: プロモーション、フェイスング) を行うアプリケーション機能。
- (6) **作業管理支援サブシステム機能**  
店舗における作業計画、作業実績管理を行うアプリケーション機能。
- (7) **オンラインアプリケーション共通機能**  
ユーザ管理、ログインメニュー画面処理、プッシュ画面処理、相乗積、粗利計算の流通用計算ツールからなるオンライン用の共通機能

## 7 . 実地検証項目とその結果

### 7 . 1 検証項目概要

実際のスーパーマーケットの店舗、本部で、流通小売現場で生の POS データ、発注、入荷、加工情報、マスター情報を取り込み、本システムを稼働させ、現場業務の中で流通業関係者とシステム開発者が一体のチームを編成し、2ヶ月に渡り、以下の検証項目について実施した。

### (1) ビジネスデータに対するデータマイニングアプローチ

流通現場ノウハウや通説を、実際のデータにより数値的検証する。またデータマイニングアプローチにより未発見の意味のある情報を検出するトライアルを行う。

### (2) 情報蓄積コストの評価

生のビジネスデータが、分散環境でデータ間の同期を保証しながら、本来業務を疎外（情報収集および蓄積の運用時に関して、人的な新規コストが本システムを導入する事で発生する）することなく情報蓄積することが可能であるかの検証。

### (3) アプリケーションの精度の検証

分散高速情報ストレージを実装する本システムを利用することで、種々の業務解析テーマが実践でき、その結果として、利用者現場において、システム導入による効果（流通業における典型例が、機会ロス・廃棄ロスの軽減による粗利益率の向上）が得られることの検証。

### (4) プッシュ型アプリケーションシステムの有効性の検証

従来のアプリケーションの実現形態であるプル型アプリケーションに加えて、本システムで提案実現するプッシュ型アプリケーションシステムが、利用者業務現場で歓迎されるか、また有効な手段として活用し得るかの検証。

上記検証ツール機能により、実験期間で以下の結果を得た。

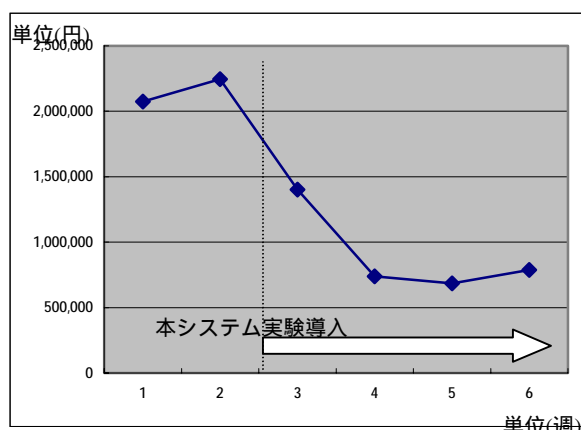


図4 推定機会ロスの推移

#### ・棚効率の分析

本システムに投入された棚割情報から、商品毎に販売情報、粗利情報、E値などをABC分析できる機能を搭載し、視覚的に判断できるGUI画面を具備した。

本機能により、現場認識と売れ筋商品のズレを簡単な操作で確認できるものが提供できた。

商品名	値	商品名	値	商品名	値
ヤクルト ヤクルト 200 65m 13	288	クロレラ アップル 65...	590	クロレラ メロン 65...	0
生活応援クロレラジモンボ袋 100m 19	3,892	クロレラ3Pサワー	0	ヤクルト 5本入	4,900

図5 棚効率の分析画面

## 7.2 検証結果

### (1) ビジネスデータに対するデータマイニングアプローチ

流通現場から要望の高かった、機会ロスの推定、棚効率の分析等を中心にシステムを使って、実験を行った。

#### ・機会ロス推定

機会ロスは、以下のアルゴリズムをシステムに実装し、オンライン等で分析できる機能を準備した。

#### 品切れ商品の検出：

在庫量が0となった時間

販売レシートから販売が一定時間以上0となった時間

商品によって、在庫量が正しく把握できるもの（入荷商品＝販売商品）とできないもの（店内加工品など）により、の算出方法を変えた。

#### 機会ロス数量・金額

商品毎に販売E値、E曲線を算出して、品切れ時間以降、商品が陳列されていれば売れたであろう数量、金額を商品毎に推定算出する。

### (2) 情報蓄積コストの評価

本システムとのデータインタフェースにかかる既存システムとの環境整備、ソフトウェア変更等のコストは、実地検証の結果より、初期導入時で1～5人月程度の稼働と算出された。

また一旦インタフェースが確立すると、ほぼ完全自動化が実現できる。

これは、もちろん既存システムのインタフェースに依存するが、本システムは、UNIX、TCP/IPを前提にした完全特注のオープン系で実装されていることによる。

但し大手メーカーの提供するPOSシステム、加工計量機などから情報インタフェースする際の、メーカーサイトの対応の遅さ、非公開性には閉口する。

他山の石として、本システムを基盤とした事業化の際には、インタフェース公開を前提にするものとする。

また本システムのオンライン機能を使った情報投入コストは、概ね現行のコストと比較して同等の評価は確保できた。

また特に店舗現場では、本部から送られる販促テーマ、販促企画などの冊子、資料が年間を通すと膨

大(月平均 50cm)になり、狭いバックヤードに保管することが困難で、現実問題として、資料は一定期間を過ぎると廃棄している事実がある。

本システムの利用は、上記課題を解決するひとつの手段となり(電子化して保存蓄積するため、保管場所問題の解決、情報の引き出しやすさに繋がる)、最低限現行業務と同等であれば、有効性を実証できることがわかった。

### (3) アプリケーションの精度の検証

下記表 1 に本システムが算出する販売予測値(E 値)と販売実績値(1ヶ月間)の比較を示す。

	E 値総額	実績金額	比率
豆腐類	3,756,083 円	3,906,002 円	103.99%
牛乳類	4,530,340 円	4,668,604 円	103.05%
牛肉類	7,478,477 円	8,068,040 円	107.88%
刺身類	8,109,253 円	9,254,075 円	114.12%
嗜好品	10,212,237 円	9,98,1532 円	97.74%

表 1 E 値と販売実績値の比較

表 1 からわかるように、生鮮系のカテゴリを除けば、ほぼ本システムの算出する E 値は現実に則していると判断できる。

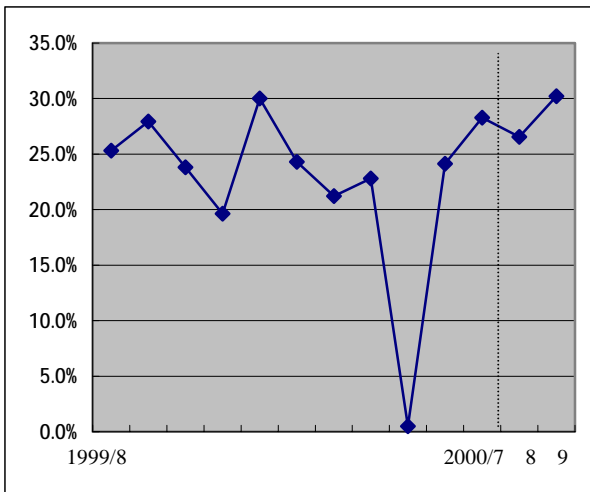


図 6 粗利率の推移

また本システムの導入(実地検証)における粗利率の数値を上記図 6 に示す。

### (4) プッシュ型アプリケーションシステムの有効性の検証

実際の店舗現場にてシステム稼働させ、プッシュ型アプリケーションによる、アラーム・ガイダンスメッセージ出力の内容、タイミングの履歴をとり、その履歴を検証、分析してシステムパラメータのチューニングを行う形式で実験を行った。

現場ヒアリングなどから以下の結果を得られた。

- ・ 実験当初は販売量 E 値の算出アルゴリズムの実際現場の動向とのずれがあり、実際にそぐわないアラームの出力が多数あった。
- ・ 実験中にチューニングした結果、実験 2 ~ 3

週間以降は、日配、グロッサリー系では、実態に近づく結果を得られた。

- ・ 生鮮系をはじめとする店内加工商品に関する販売 E 値は未だ満足した結果を得られていない。

これは、同じ商品でも、たとえば 100g あたり 100 円で、商品によっては 300g (300 円)、500g (500 円) と、同一商品コードで売価が異なる現象が発生するためである。

このような商品については、実験途中から販売 E 値は、販売総金額の評価として予測値を算出したが、まだ今後課題を残している。

## 8. 全体考察

本システム機能では、機会ロス検出機能、棚割画面によるマーチャライジングと E 値シミュレーション機能、作業要員計画管理機能など、既存システム、メーカーが提供するシステムでは、国内ではまだ存在しない機能を中心に、流通現場から多大な評価を頂いた。

さらに、流通業関係者が、本システムが提供する機能、情報を見ることで、

- ・ こんな情報は見られないのか
  - ・ こんな風に表示すると、もっとわかりやすい
  - ・ ここまでできるのなら、こんな機能もつけてくれ
- 等の意見を頂くまでほど意欲的に検証作業に取り組んでいた頂き、開発者と密にシステム企画・修正・機能追加構想を行える関係となった。

開発全体を通じては、

- ・ ソフトウェアハウスという事情から、国内では、ソフトウェア大企業によって抑えられ、なかなか接することができない現場の生のビジネスデータに接する機会が作れ、開発機能へのイメージが膨らみ、今後のためにまたとない経験をつむことができたこと。
- ・ 処理高速化のため、揮発性メモリにデータベースを構築し、ネットワークを介した分散 DB が、企業の情報システムに十分対応できることが実証できたこと。
- ・ E 値演算など、高速アクセスを前提にしたデータベースをもとに、既存システムでは不可能と考えられた演算(データマイニングなど)が検証できたこと。

などの成果が挙げられる。

## 9. 参加企業及び機関

アルビス株式会社  
株式会社 カスミ  
株式会社 数理システム  
株式会社 エム研  
株式会社 数理技研

## 10. 参考文献

特になし。