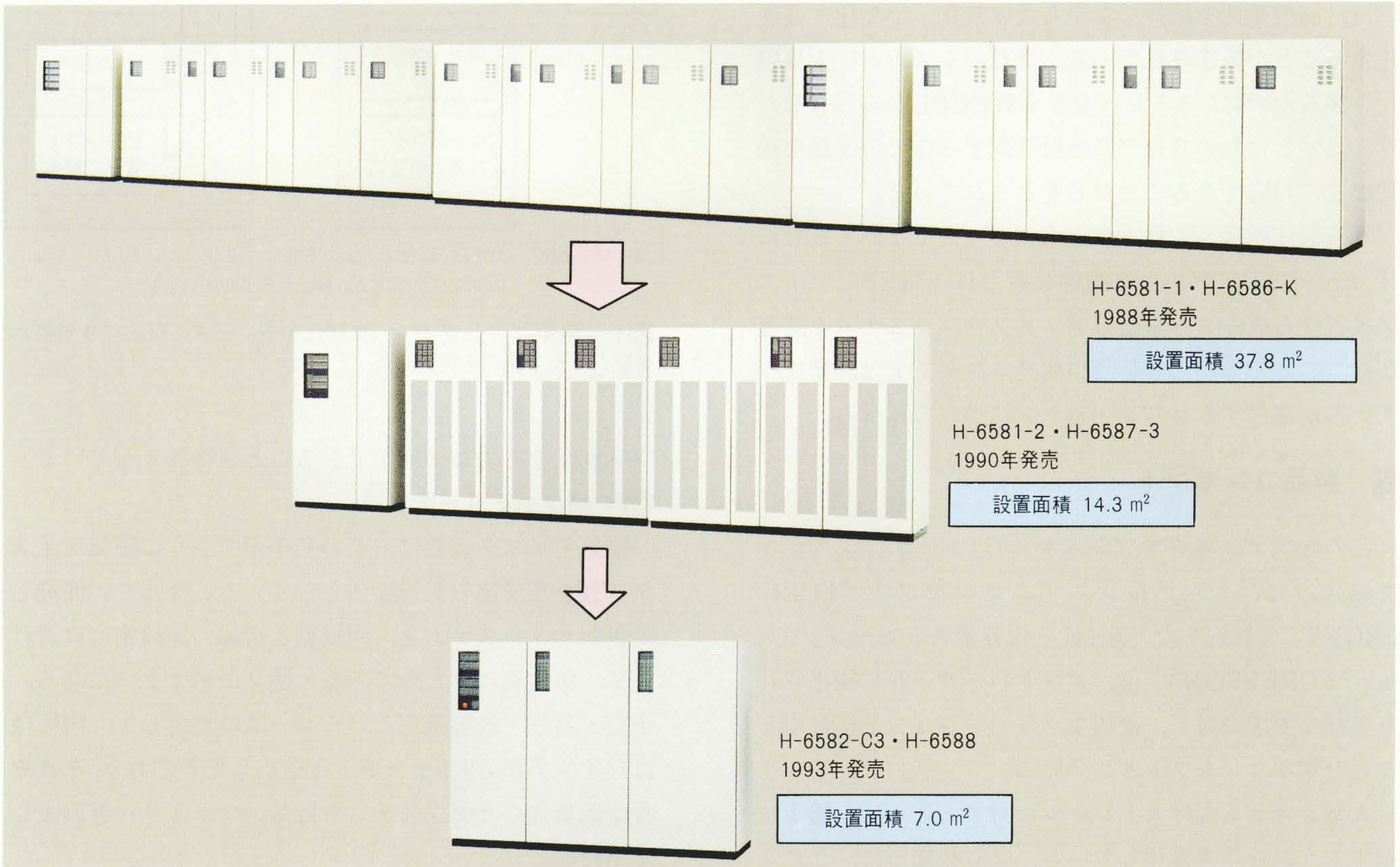
高信頼化・小型化を実現した大容量磁気ディスク装置

Large Capacity Magnetic Disk Storage That Realizes High Reliability and Becomes Smaller

斉藤茂芳*　*Shigeyoshi Saitô*　平島　健*　*Ken Hirashima*
鈴木冨男*　*Tomio Suzuki*　白柳芳朗*　*Yoshirô Shiroyanagi*



180 Gバイト構成磁気ディスクサブシステムの設置面積推移
設置面積を最近の5年間で$\frac{1}{5}$以下に低減(当社従来機比)している。設置面積には保守エリアを含む。

磁気ディスク装置には「小さくて大容量」,「業務処理時間が短い」,「故障しない」,「柔軟な構成ができる」などの使いやすさをはじめ，ランニングコスト削減をも含む幅広い要求への対応が求められている。

日立製作所は，これらの要求に対するソリューションとして大容量磁気ディスクサブシステム(H-6582-C3，H-6588)を開発した。6.5インチ径磁気ディスクなどにより，単位容量当たりの設置面積(保守エリア含む)および消費電力(発熱量)を最近の5年間でそれぞれ$\frac{1}{5}$以下および$\frac{1}{3}$以下に低減し，部品点数の削減，障害兆候検知機能の開発などにより，高信頼性を確保した。また，大容量キャッシュメモリのサポート，光ファイバインタフェース(ACONAR-C：Advanced Connection Architecture)の採用などにより，高速化を達成するとともに，さまざまなエンドユーザー環境での柔軟なシステム構成を可能にした。

* 日立製作所 ストレージシステム事業部

1 はじめに

大型システムでは，トータル マネージメント サーバとしての役割を担うため，高速プロセッサ(CPU)・新システムアーキテクチャの開発などで高性能化が進行している。それに伴い，これらのシステムが取り扱う情報量は急激に増大しており，情報を格納するストレージ製品には，高信頼性を維持しながら大容量化・小型化・高速化することが要求されている。また，システムの大規模化を図るために，ストレージを多数のCPUから共有したり，CPUとは別のフロアに分散設置するなどの多様な接続形態に対応できることが必要とされている。

日立製作所は，これらの要求にこたえて1993年6月にH-6582-C3磁気ディスク制御装置とH-6588磁気ディスク装置から成る大容量磁気ディスクサブシステムを開発した。ここでは，この製品の開発コンセプトと，サブシステムが提供するソリューションについて述べる。

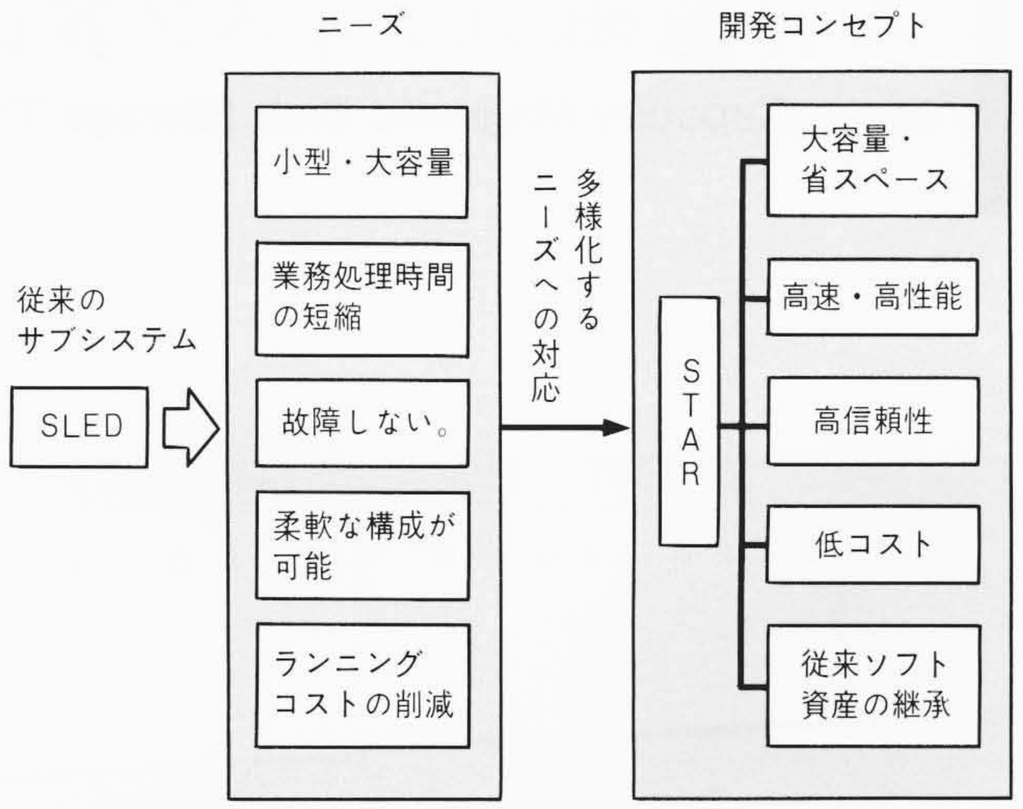
2 製品コンセプト

この磁気ディスクサブシステム(以下，サブシステムと略す。)は，トータルシステムコンセプト“FOREFRONT”に対応した「高性能・大容量ストレージ」である。“FOREFRONT”は，エンドユーザーから提出される多様な要求に対し，適切なソリューションを提供することを基本コンセプトとしている。

大型システム向けストレージに対する要求は，**図1**に示すように近年多様化している。日立製作所は，これらの要求に合致するサブシステムの開発にあたり，“FOREFRONT”に基づいて次に述べる新しいコンセプトを導入した。

2.1 SLEDからSTARへ

従来，大型システム向けストレージの中心的役割を担っている磁気ディスク装置は，**図1**に示すように1ボリュームずつ独立しており，大きく，高価であるという意味でSLED(Single Large Expensive Disks)と呼ばれるようになってきている。しかし，小型・大容量化を中心に多様化するニーズに，今後ともSLEDで対応するのは困難である。日立製作所は，これら多様化するニーズに対応する新しいストレージ製品の開発に際し，STAR(Small, Technologically Advanced Reliable Disks)というコンセプトを導入している。STARは小型であり，かつ従来サブシステムと同等以上の高信頼性を維持すること，およびエンドユーザーが築き上げてきた膨大なソフト資産をそのまま継承できることを特長としている。



注：略語説明 SLED (Single Large Expensive Disks), STAR (Small, Technologically Advanced Reliable Disks)

図1 開発コンセプト 多様化するニーズに対応できる高ポテンシャルの製品を開発した。

2.2 新ファミリーの形成

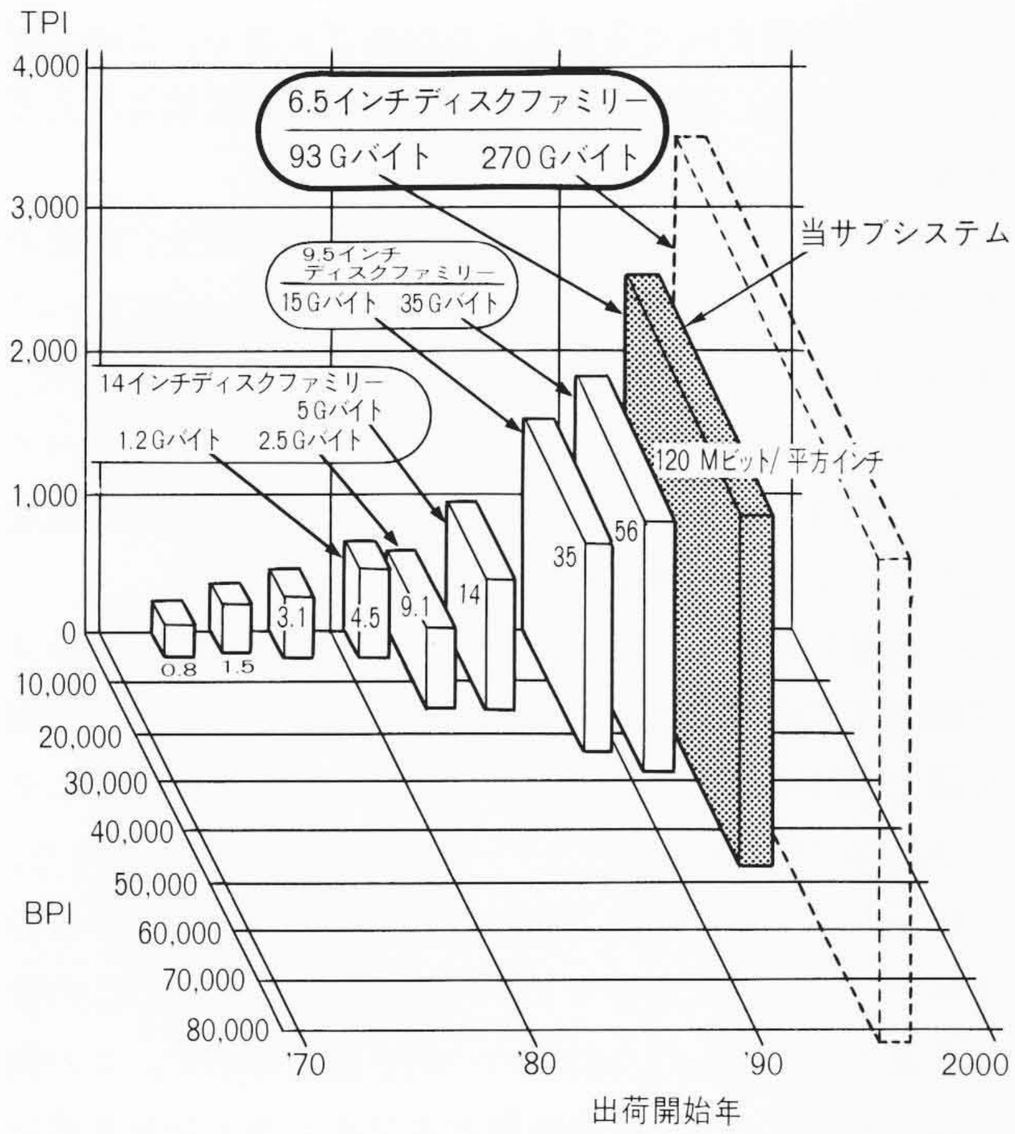
磁気ディスク装置は，その名が示すように磁気ディスク(円板)を情報の記録媒体としている。今までに開発してきたサブシステムを，円板径と情報の記録密度に着目して，年代別にまとめてみると**図2**に示すようになる。小型・高速・大容量化というニーズに対応して，円板径は14インチから9.5インチへと縮小してきており，それぞれが数世代のサブシステムを持ち，ファミリーを形成していた。

STARコンセプトに対応するサブシステムを開発するために，円板径を6.5インチへとさらに縮小した新ファミリーを形成し，サブシステムの小型・軽量化および消費電力の低減を図った。円板径の縮小は，磁気記録再生技術を進展させて記録密度を大幅に向上させることによって達成したものである。6.5インチファミリーでは，高速・高性能を特長とするこのサブシステムを第1世代と位置づけており，今後第2世代でいっそうの大容量化を図る予定である。

2.3 高信頼性の継承

システムの信頼性を支えるために，ストレージ製品に要求されることは，第一に記録された情報が失われず確実に読み出せることであり，第二は，万一の異常発生時でもシステムに及ぼす影響が最小で，正常状態への復帰が早いことである。

サブシステムの大容量化には，その心臓部にあたる磁



注：略語説明　BPI（ビット/インチ），TPI（トラック/インチ）

図2　情報記録密度と円板径の推移　　記録密度は，10年間で約10倍に向上している。

気ヘッド，磁気ディスクおよびこれらを駆動する機構類の小型化が不可欠である。このサブシステムでは，これらのキーコンポーネントの性能を向上させ，より薄く，小さくすることなどにより，小型化を達成している。小型・軽量化は，

(1)　要求機械強度が緩和するなどにより，部品の簡素化や一体化が可能となり部品点数が削減できること

(2)　消費電力が低下するため，発熱が減って各部品の寿命を伸ばすことができること

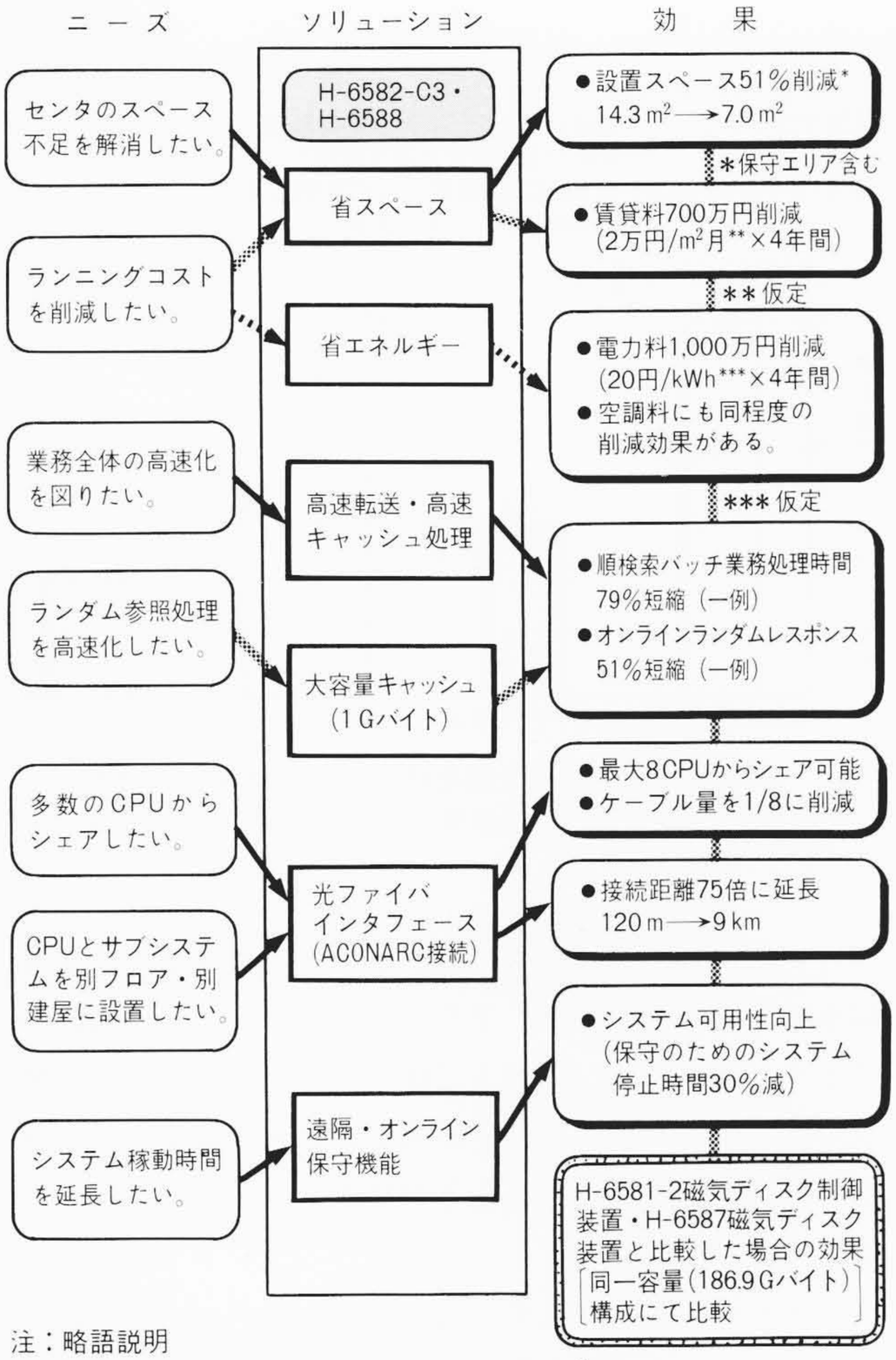
などにより，高信頼性の継承に大きく貢献している。さらに，このサブシステムでは，システム全体として有効な信頼性向上技術を開発し適用している。その代表例については次章で述べる。

3　ソリューションの提供

前章で述べたSTARコンセプトに基づいて開発したこのサブシステムが，エンドユーザーに提供するソリューションは，**図3**に示すようにユーザーニーズに対応している。

3.1　設置スペース不足の解消およびランニングコストの削減

取り扱う情報量の増加に対応してサブシステムを増設し続けてきた結果，多くのエンドユーザーではストレージの占める設置面積の割合が増大し，コンピュータルームが手狭になっている。また，増設に伴う電源・空調設備の見直し，床荷重に対する耐力増強などが必要となってきた。



注：略語説明
ACONARC (Advanced Connection Architecture)

図3　サブシステムが提供するソリューション　　多様なニーズにきめ細かく対応している。

このディスクサブシステムでは，小型・軽量HDA (Head Disk Assembly)，高集積LSIの開発，高効率冷却方式の適用などにより，省スペース・省エネルギー化を推進した。その結果，日立製作所の従来サブシステム（H-6581-2磁気ディスク制御装置・H-6587-3磁気ディスク装置[1]）比で，保守エリアを含む設置面積を$\frac{1}{2}$以下に，消費電力をほぼ$\frac{1}{3}$に低減した。消費電力の低減は，空調用電力の削減にも寄与するため，コンピュータルーム全体の省エネルギー化に大きく貢献している。

3.2　ソフトウェア資産の継承

このサブシステムは，上記3.1節に記載した従来サブシステムと100%の互換性を保持している。したがって，今

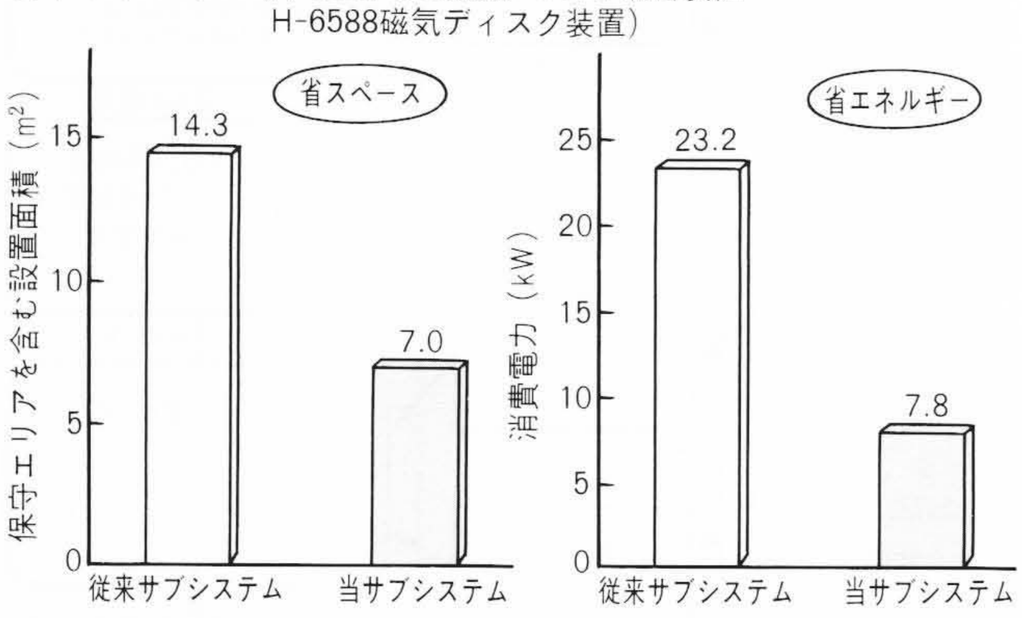


図4 保守エリアを含む設置面積および消費電力の比較
従来サブシステム比で設置面積を$\frac{1}{2}$以下に，消費電力をほぼ$\frac{1}{3}$に低減している。

までに開発してきた膨大なソフトウェア資産をそのまま継承することができる。また，このサブシステムは，ハイトランザクション環境下で使用可能な大容量・高性能ファイルであり，従来サブシステムと同じ環境，同じソフトウェアで稼動させることができる。

3.3 業務効率の向上

ストレージサブシステムに対するエンドユーザーからの要求の中の大項目として，データのアクセスを高速化して業務効率を向上させることが掲げられる。このサブシステムでは，次のような新技術でこの要求にこたえている。

(1) CPUとサブシステム間のデータ転送速度を，従来の4.2 Mバイト/s(金属ケーブルによる転送)から17 Mバイト/s(光ファイバケーブルによる転送)へと引き上げた。

(2) キャッシュメモリを標準でサポート(容量64 Mバイト)し，**図5**に示すようにキャッシュサポート機能を高度化して，より高い性能を引き出した。

(3) **図6**に示すように，磁気ディスク装置と磁気ディスク制御装置の間では4データパスを，磁気ディスク制御装置とCPUの間では8データパスをそれぞれ同時にデータ転送可能な構成とし，データ転送能力を向上させた。

以上のような性能向上策は，いずれもサブシステムのスループット向上に寄与している。実際の業務にこの技術を適用して得られた効果の一例を**図7**に示す。この例では，順検索バッチ処理時間およびオンラインレスポンス時間が，それぞれ79%および51%短縮され，性能が向上している。

3.4 多様な接続形態への対応

ここでは，このサブシステムが採用した光ファイバケーブルによるCPUとサブシステムの間の新インタフェースACONARC(Advanced Connection Architecture)

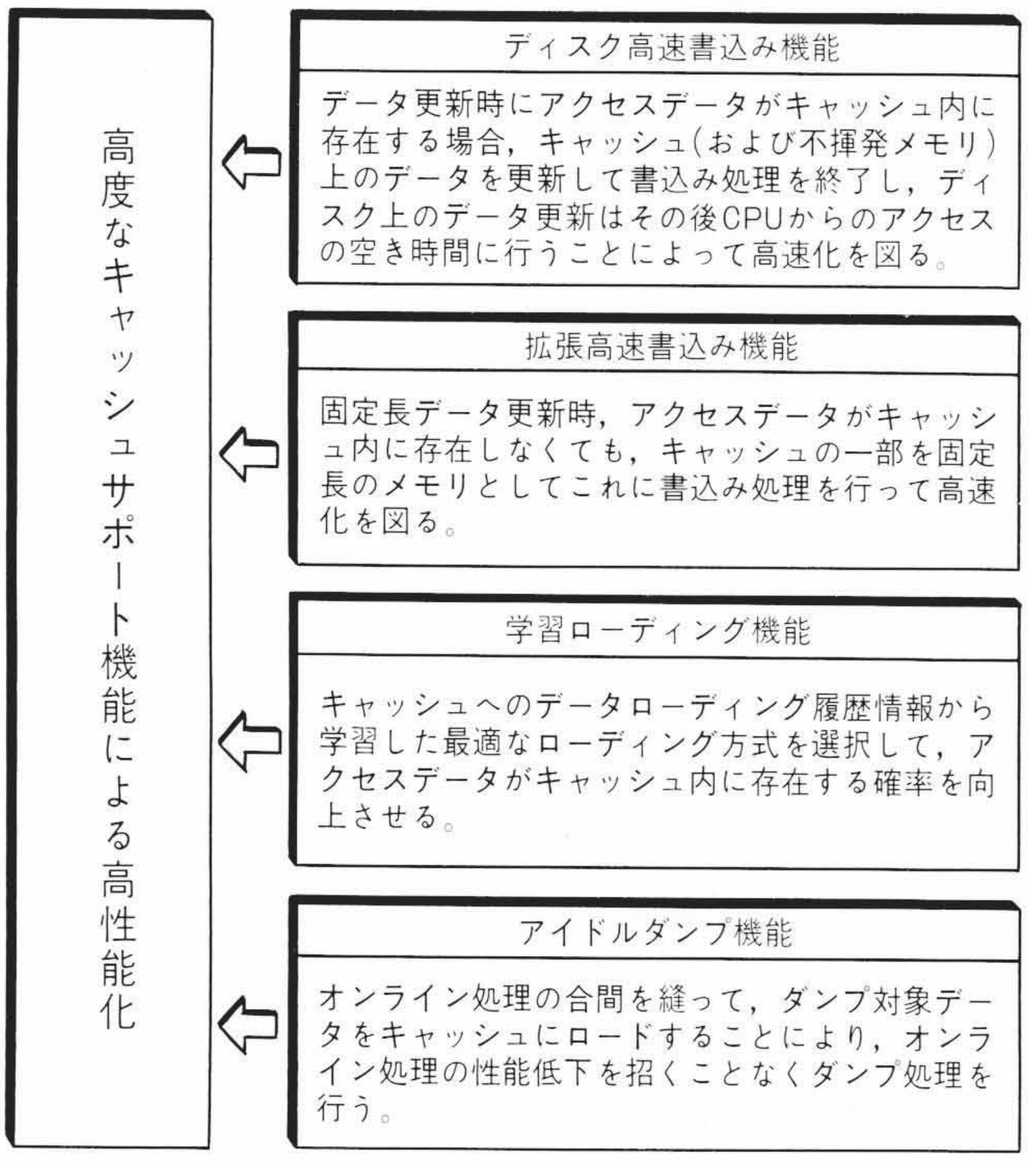
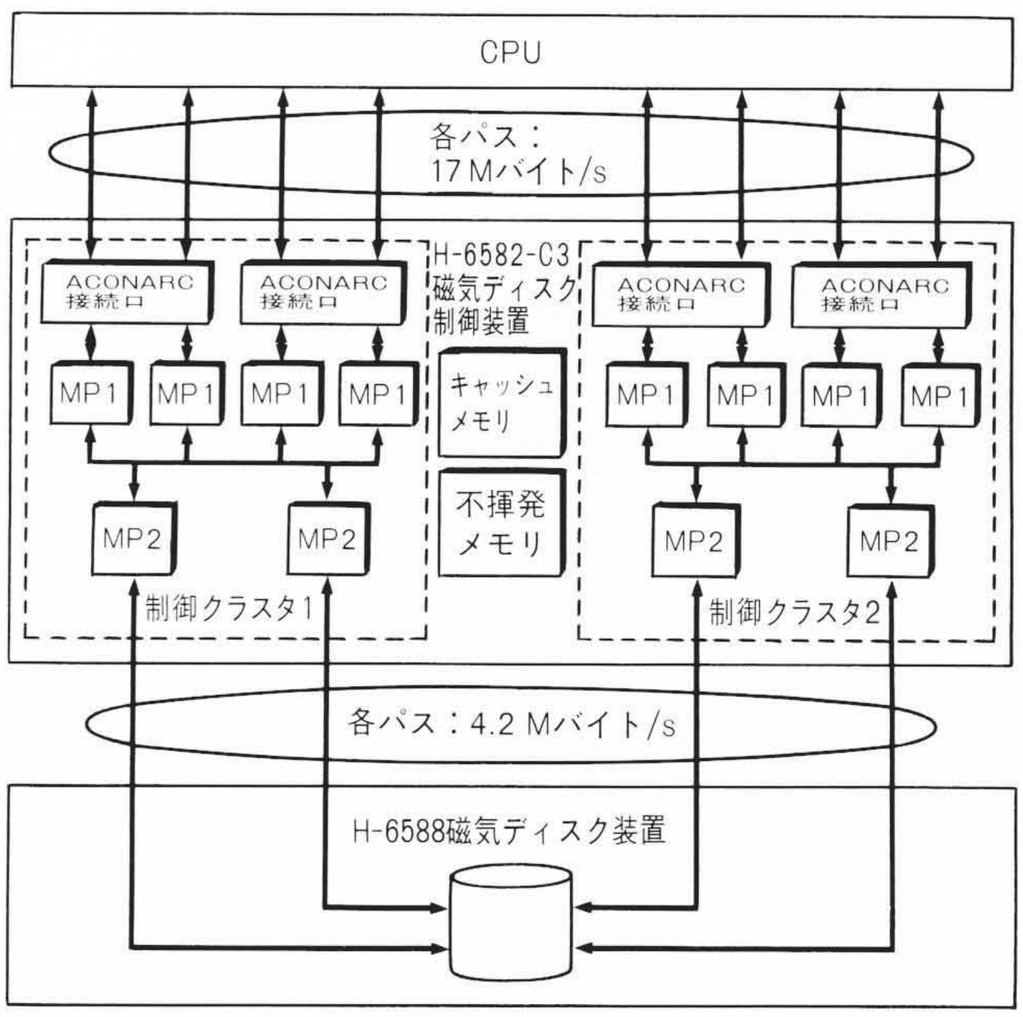
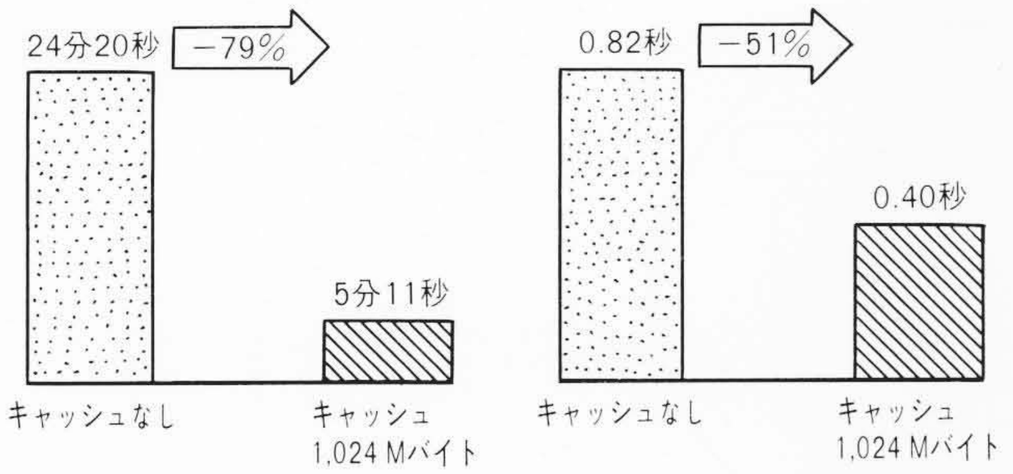


図5 充実したキャッシュサポート機能 各種のキャッシュサポート機能を開発し業務効率を向上させている。



注：略語説明 MP1(マイクロプロセッサ1)，MP2(マイクロプロセッサ2)

図6 データ転送パスの充実 CPUと磁気ディスク制御装置の間を8パスで，磁気ディスク制御装置と磁気ディスク装置の間を4パスで構成し，それぞれのパスが独立かつ同時にデータ転送可能とした。



(a) 順検索バッチ処理時間　　(b) オンラインレスポンス時間

記　号	サブシステムの形名	測　定　条　件
	H-6581-2・H-6587 (従来サブシステム)	キャッシュなし データ転送速度 4.2 Mバイト/s
	H-6582-C3・H-6588 (当サブシステム)	キャッシュ容量 1,024 Mバイト 不揮発メモリ 8 Mバイト データ転送速度 17 Mバイト/s

図 7　サブシステムの性能比較(当社比の例)　　同一の業務をそれぞれのサブシステムを使用して行ったときの処理時間, レスポンス時間を比較した。その結果, この例では当サブシステムが従来サブシステムに比べて50%以上処理時間を短縮しているという結果を得た。

の効果について述べる。

3.4.1　接続距離の延長

ACONARCの適用により, CPUとの間の接続に従来の金属ケーブルだけではなく光ファイバケーブルをも使用できるようにした。これによる効果は次のとおりである。

(1)　CPUとサブシステム間の最大接続距離を, 従来のケーブル長120 mから 9 kmへと延長した(今後いっそうの延長を計画している)。

(2)　軽量で細い光ファイバケーブルを採用してコンピュータルームの床下配線を簡素化することにより, サブシステム小型化の効果である機器の集中設置を可能にするとともに, インストール時間を短縮した。

3.4.2　多数CPUからのシェア

ACONARCディレクタ(**図 8** 参照)を介することにより, このサブシステムは, CPUとの間の論理接続パス数を従来の16本から32本に拡張している。したがって, 一つのCPUからの論理接続パス数を 4 本とすると, 最大 8 台のCPUからサブシステムをシェアすることができる。

以上述べた新インタフェースACONARCの採用により, サブシステムをCPUとは別のビルディングに集中設置して管理したり, いくつかのビルディングに分散設置し, 多数のCPUからシェアするなどの柔軟なシステム構成が組めるようになった。その一例を**図 8** に示す。

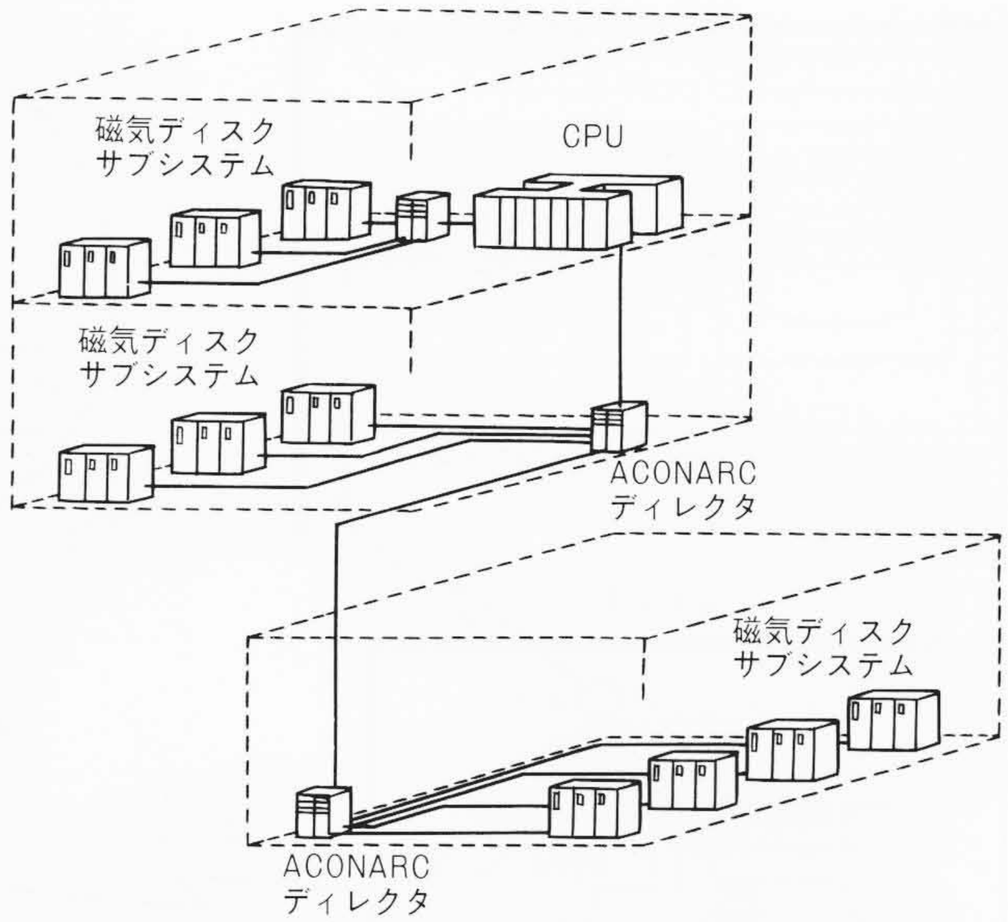


図 8　ACONARCによるシステム構成例　　CPUとサブシステムを別フロア・別建屋に設置するなどの多様なシステム構成が可能となった。

3.5　システム連続稼動の確保

前節で述べたような多様な接続形態への対応に関する要求と並行して,

(1)　容量増設, 接続変更, 障害発生などがあってもシステムを停止させない。

(2)　万一システムを停止させても, 最小時間で復帰させることができる。

ことへの要求は, 従来以上に高まっている。これらの要求に対して,このシステムは次のようなソリューションを提供できる。

3.5.1　オンライン動作中の増設・撤去

このサブシステムは, オンライン動作中の増設・撤去に関して次のように配慮している。

(1)　磁気ディスク装置と磁気ディスク制御装置との間のインタフェースを複数パス(4 パス)で構成しており, 一部のパスを交代で切り離してその経路上の増設・撤去を行うことにより, システムを停止せずに作業を行うことができる。

(2)　一部のパス切り離しによるシステムの性能低下を最小時間にするため, 次のようにして作業時間を短縮している。

(a)　小型化したHDAを, 容量増設単位(4 HDA)ごとに一体化

(b)　磁気ディスク装置と磁気ディスク制御装置をボルト連結とし, 従来使われていた床下インタフェースケーブルを不要にしてケーブル接続作業を簡素化

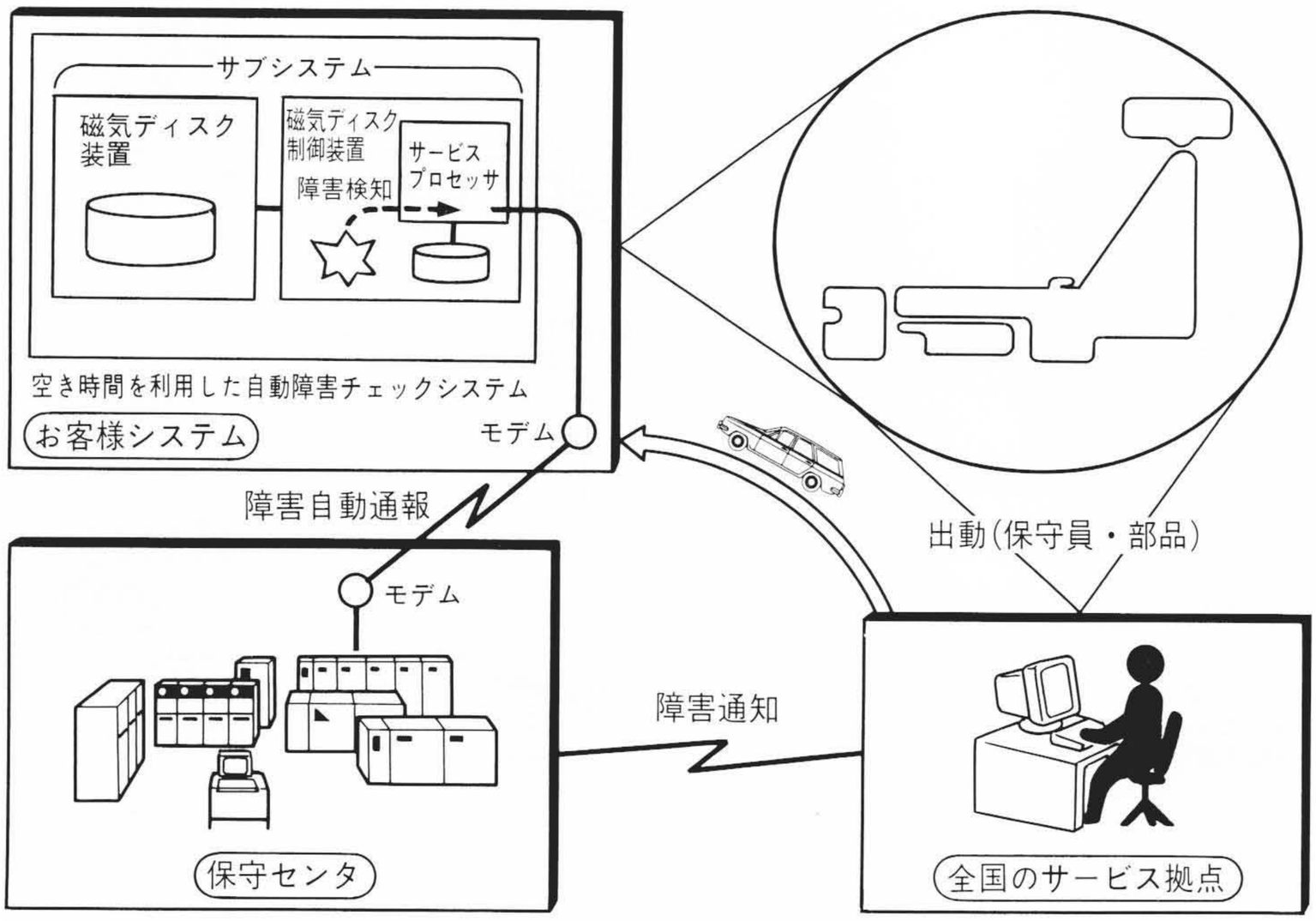


図9 保守サービス体制の充実
エンドユーザーの障害情報，障害検知情報，稼動管理情報を通信回線で保守センタに集めて一括管理することにより，予防保守および万一のときの迅速な対応をさらに充実させた。

3.5.2 オンライン保守

万一の障害発生時に，システムに与える影響を最小にして継続稼動を可能にするために，このサブシステムでは次の機能をサポートしている。

(1) 障害兆候検知機能

(a) 装置内温度，電源電圧，冷却用ファンの回転の異常を検知し，実障害発生に至る前に警告を発する機能

(b) CPUからのアクセスに対し，読出しデータの誤り回数，アクセス動作の異常回数(いずれもリトライ動作によって回復可能)をボリューム単位に管理し，エラーの多いボリュームを検知して，回復不能なエラーに至る前に警告を発する機能

(c) CPUからのアクセスの空き時間を利用して，各HDAに記録されている全データを順次読み出してボリューム単位にエラー数を管理し，エラーの多いボリュームを検知することにより，CPUからのアクセスがある前に書き込まれたデータのヘルスチェックができる機能

(2) 障害発生または障害兆候検知したデータ転送パスを切り離して縮退運転させ，その間に切り離した経路上の不良パーツを交換できる機能

(3) 磁気ディスク制御装置内のマイクロプログラムを，縮退運転中に変更できる機能

障害情報，障害兆候検知情報および稼動管理情報は，**図9**に示すように磁気ディスク制御装置内のサービスプロセッサから通信回線を経由して，保守センタへ自動通報される。この情報により，ユーザーごとの稼動状況を監視するとともに，保守員が交換用のパーツをあらかじめ準備してユーザーサイトに出向くなどの迅速な対応を図ることができる。

4 おわりに

大型システム向け磁気ディスクサブシステムの主力製品について，これが提供しているソリューションを中心に述べた。

近年，RAID(Redundant Array of Inexpensive Disks)技術を使用したディスクアレーが，大型システムにも投入されつつある。今後は従来ソフトウェア資産をそのまま活用できるSTAR型のサブシステムと，高可用性を特長とするRAIDを並行して開発し，ますます多様化する用途・利用形態に的確に対応する製品群を提供していく考えである。

参考文献

1) 高橋，外：35 Gバイト大容量磁気ディスク装置，日立評論，**73**，2，205～210(平3-2)