

# 新 キャッシュ付きディスク制御装置“H-6581-C3”

## New Cache Disk Control Unit H-6581-C3

磁気ディスクサブシステムの性能向上・信頼性向上を可能とする新形キャッシュ付きディスク制御装置H-6581-C3(以下、H-6581-C3と略す。)を開発した。

H-6581-C3は、12式のマイクロプロセッサによるマルチマイクロプロセッサ構造をとり、大容量キャッシュ、高速データ転送路、不揮発メモリなどを搭載している。機能的には、高速書込み機能、学習機能応用によるキャッシュ制御、ディスク二重書きなどをサポートするインテリジェント制御装置である。さらに、ハードウェア冗長度を高め、SVP(Service Processor)を装備し遠隔保守を可能にするなど、信頼性・保守性についても向上させている。

宮崎道生\* *Michio Miyazaki*  
 久野 潔\* *Kiyoshi Hisano*  
 北嶋弘行\*\* *Hiroyuki Kitajima*  
 加茂善久\*\*\* *Yoshihisa Kamo*  
 川村哲士\*\*\*\* *Satoshi Kawamura*

### 1 緒 言

磁気ディスクサブシステムは、コンピュータシステムで外部記憶装置の中心として幅広く使用されている。磁気ディスク装置(以下、DKUと略す。)の装置当たり記憶容量は年約30%で増加しているが、データ転送速度、シーク速度、回転速度などによって決まる処理速度は年数パーセントの向上にとどまっている。DKUとCPUとの処理速度の差(I/Oボトルネック)が広がっており、これをいかに小さくするかがコンピュータシステムの大きな課題の一つとなっている。

このI/Oボトルネックを解消し、より高性能のディスクサブシステムを実現するために、ディスク制御装置(以下、DKCと略す。)の高機能化が重要になっている。例えば、キャッシュと呼ばれる半導体メモリをDKCに搭載し、使用頻度の高いデータはキャッシュに格納する。CPUからの読取り要求で、要求されたデータがキャッシュにあれば(この場合をヒットと呼ぶ)、DKUにアクセスすることなくキャッシュから直接CPUへ転送する方式がある。これにより、高速アクセスが可能となり、ディスクサブシステムの高性能化が実現できる(図1)。

新形キャッシュ付きディスク制御装置H-6581-C3(以下、H-6581-C3と略す。)(図2)は、マルチマイクロプロセッサ構造をとり、大容量キャッシュ、高速書込み機能、ディスク二重書き機能のサポートなどの特長を持つ高性能・高信頼のDKCである。H-6581-C3により、コンピュータシステムの大幅な性能向上・信頼性向上が可能となった。

本稿では、H-6581-C3の開発思想、構成および新機能について述べる。

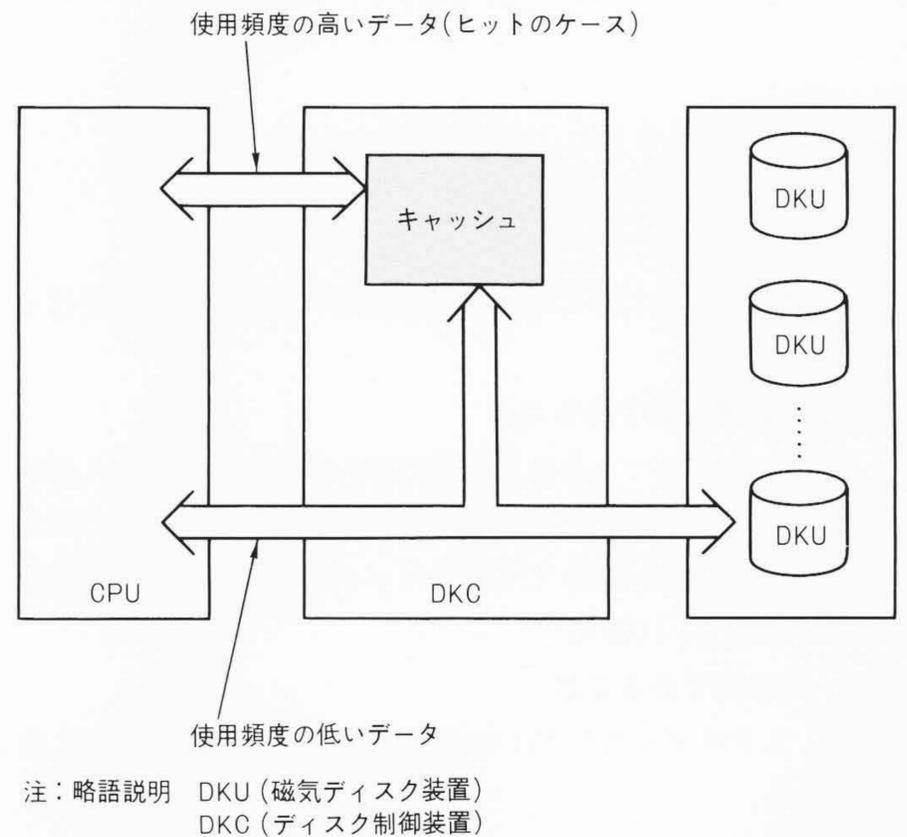


図1 キャッシュ付きDKCの概念 キャッシュ付きDKCの基本的動作を示す。

### 2 H-6581-C3の開発思想

外部記憶装置としてのディスクサブシステムに対する要求仕様を整理すると、下記の4項目に要約できる。

- (1) 大記憶容量であること。

\* 日立製作所 小田原工場 \*\* 日立製作所 システム研究所 \*\*\* 日立製作所 中央研究所  
 \*\*\*\* 日立製作所 マイクロエレクトロニクス機器開発研究所 工学博士

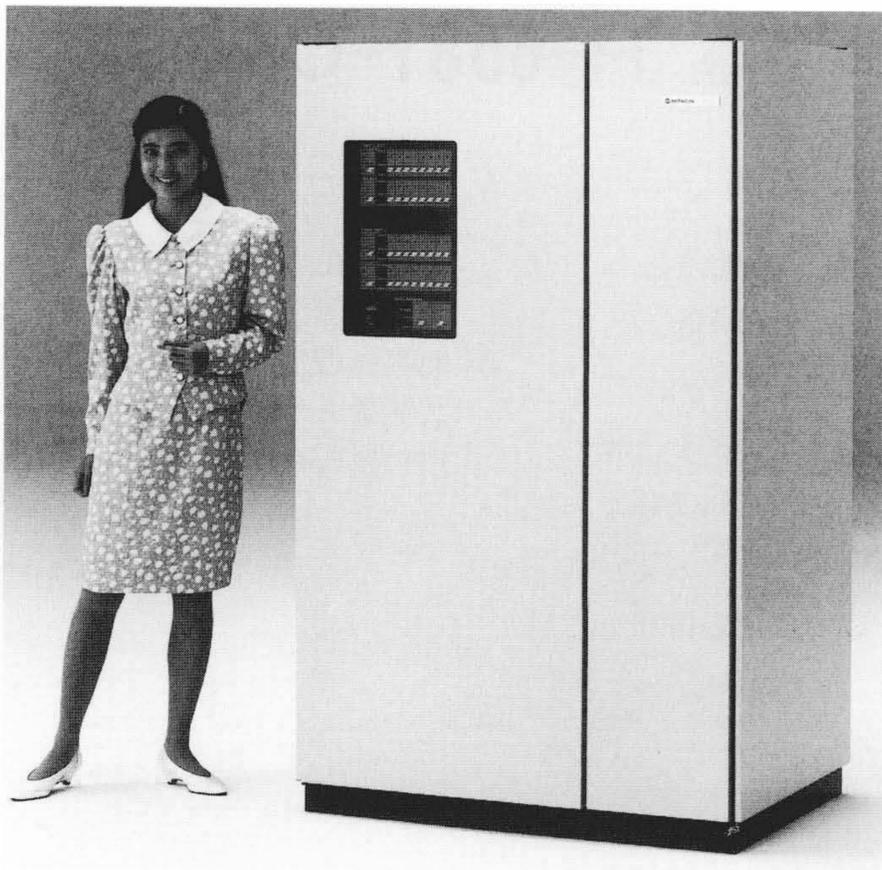


図2 H-6581-C 3の外観 H-6581-C 3の外観を示す。

- (2) 高性能であること。
- (3) 高信頼であること。
- (4) 使いやすいこと。

H-6581-C 3は以下に述べる基本思想に基づいて開発を行った。

### 2.1 大記憶容量であること

DKUのボリューム当たり容量の増大に伴い、アクセス要求も増大する。ボリューム当たりのアクセス要求の増大に耐えられるように、H-6581-C 3ではキャッシュ機能の拡大を図り対処した(2.2 (1)参照)。

### 2.2 高性能であること

ディスクサブシステムに対する性能要求は下記2項目に要約できる。

- (1) 高頻度のアクセス(DKUに対する読み、書き)要求に耐えられること。

この性能は特にオンライントランザクション処理に重要であり、次に示すキャッシュ機能の拡張によって性能向上を図った。

- (a) キャッシュ容量を増大させる。
- (b) 書込み要求に対しても、キャッシュ機能を適用(高速書込み機能)する。
- (c) 高度なマルチマイクロプロセッサ方式の採用により、高頻度、高多重の処理を可能とする。
- (2) データ転送能力の強化

この性能要求は、バッチ処理をはじめ大量データ処理で重要であり、データ転送路の能力の拡大によって対処した。DKCとDKU間のパス数、およびDKCとCPU間のパス数を増大させ、

DKCとCPU間のデータ転送については、より高速化を図った。

さらに、DKCとDKU間のデータ転送路の性能向上のために、ALC(Actuator Level Cache)<sup>1)</sup>を開発した。

### 2.3 高信頼であること

この要求項目は、(1) 装置としての信頼性、(2) データ保証、(3) 保守性、に分けることができる。

- (1) 装置としての信頼性

装置に障害が発生してもこれを確実に検出し、かつ障害部位を切り離してディスクサブシステムとして動作を継続することが重要である。このため以下の機能を開発した。

- (a) 障害検出機能・ハードウェア冗長度の向上
- (b) 自動縮退・回復機能

- (2) データ保証

2.2で述べた新たにサポートする高速書込み機能では、書込みデータがキャッシュにいったん格納された後DKCの制御によってDKUに書き込まれる。このため次の2項目により、キャッシュ部の高信頼化を図った。

- (a) 書込みデータを保持するキャッシュを二重化する。
- (b) 二重化したメモリの一方を不揮発化する(不揮発化機構)。

さらに、DKUについても二重書き機能をサポートする。

- (3) 保守性

より高度な保守・診断を可能にするためSVP(Service Processor)を開発した。

### 2.4 使い勝手

使いやすいディスクサブシステムを提供するために下記機能を開発した。

- (1) キャッシュ使用効率を自動的に最適化することにより、人手によるチューニングを不要とし、導入を容易にする(学習ローディング)。
- (2) 大容量DKUのデータを磁気テープなどにバックアップするダンプ処理の負担が大きくなっている。これを解決するためダンプ機能の高速化を図る(アイドルダンプ)。

## 3 H-6581-C 3の構成

H-6581-C 3では以下の構成を採用した(図3)。

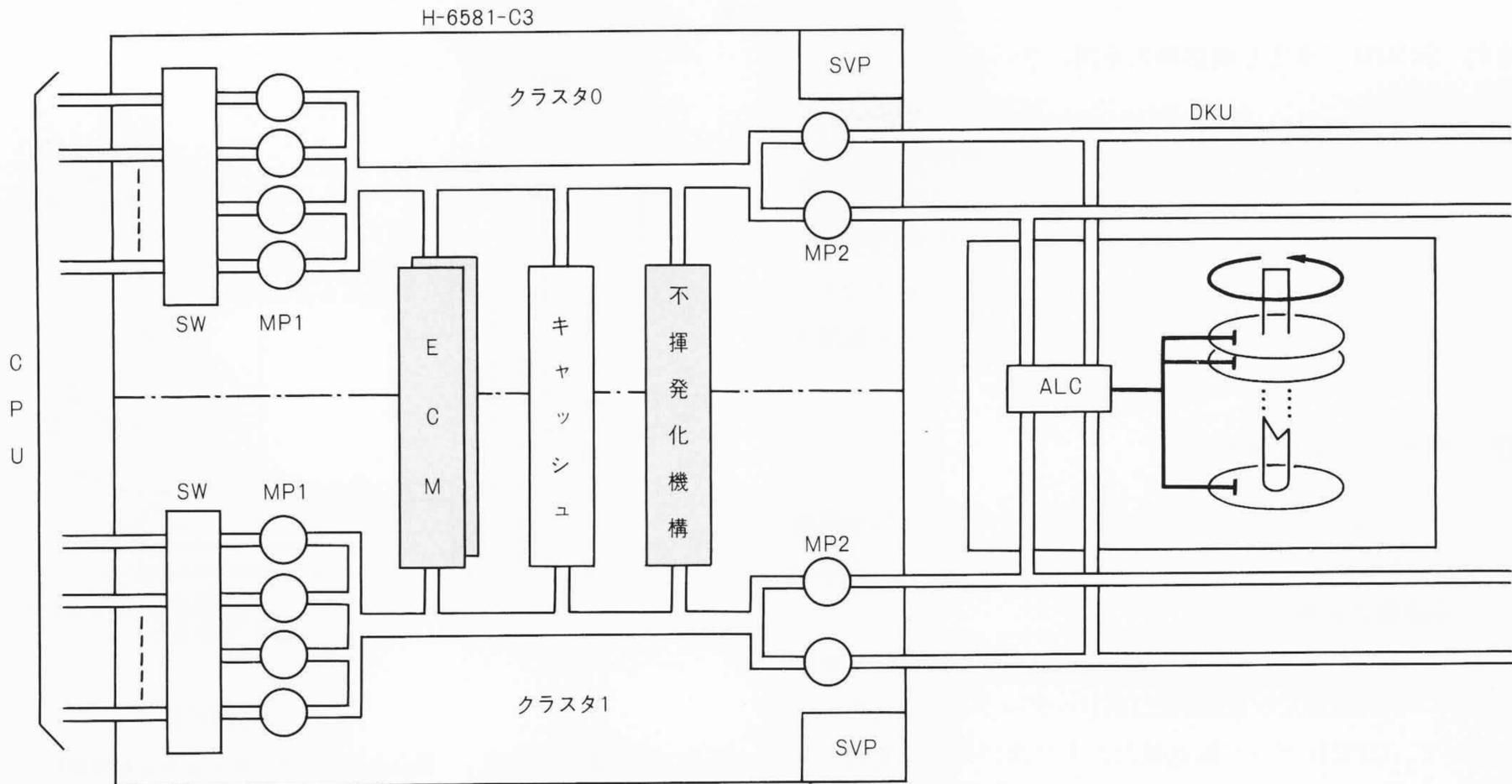
また、主な仕様を表1に示す。

- (1) 大容量キャッシュおよび不揮発化機構

H-6581-C 3は最大512 Mバイトのキャッシュを搭載でき、キャッシュのヒット率をより高くすることが可能である。また、高速書込み機能をサポートするために、バッテリーバックアップによって48時間にわたってデータ保存が可能な8 Mバイトから成る不揮発化機構を開発した。

- (2) 高速データ転送

大容量キャッシュによる高ヒット率の効果を、より有効に発揮させるために、H-6581-C 3ではキャッシュヒットの場合、



注：略語説明ほか MP1 (マイクロプロセッサ-1), MP2 (マイクロプロセッサ-2), SW (チャンネルスイッチ), SVP (Service Processor)  
ALC (Actuator Level Cache), ECM (Extended Control Memory), ○ (マイクロプロセッサ), □ (不揮発メモリ)

図3 H-6581-C 3 の構成図 H-6581-C 3 のハードウェア構成とDKUの接続を示す。

CPUとの間で最高9 Mバイト/sのデータ転送を可能としている。また、対CPUとのデータ転送路を、従来DKC(H-8538-C 3)の4倍の8パス(最大構成)とした。さらに、DKUとのデータ転送能力を高めるため、従来の2パスから倍の4パスをサポートし、データ転送能力を向上させた。

表1 H-6581-C 3 の仕様 H-6581-C 3 の仕様を従来機との比較で示す。

項目	H-6581-C 3	H-8538-C 3 (従来機)	
制御論理 (ディレクタ)	4 (2クラスタ)	2	
最大キャッシュ容量 (Mバイト)	512	64	
不揮発化機構 (Mバイト)	8 (48時間データ保持)	なし	
データ転送速度 (Mバイト/s)	CPU ↳ DKC間	9	3
	DKC ↳ DKU間	4.2	3
データ転送経路数	CPU ↳ DKC間	8	2
	DKC ↳ DKU間	4	2
マイクロプロセッサ数	12	2	

対CPUパスと対DKUパスとの合計12パスのデータ転送路はキャッシュにも接続しているため、キャッシュは合計90 Mバイト/s以上のデータ転送能力を持つようにした。

(3) マルチマイクロプロセッサ構造

3章の(2)で示したように、H-6581-C 3は、合計最大12パスをサポートしている。各データ転送路が負荷を均等に処理し、その転送能力を十分に発揮することが重要である。このため、各データ転送路ごとに、互いに独立なマイクロプロセッサ(最大12式)を設けた。このマルチマイクロプロセッサ構造により、対CPU側処理と、対DKU側処理の独立性を高め、空いたデータ転送路を有効に動作させることが可能となった。

12式によるマルチマイクロプロセッサを高性能・高信頼に動作させるため、マスタプロセッサに相当する部位を持たない分散協調形マルチマイクロプロセッサ方式を開発した<sup>2),3)</sup>。この方式の特徴は、ECM(Extended Control Memory)と呼ぶ共用の制御メモリ上に一元管理されたシステム状態情報に基づいて、各プロセッサが独立に処理を進める点にある。この方式によれば、特定のプロセッサに負荷が集中することがなく、システム性能を向上できる。また、マスタプロセッサを持たないため、プロセッサの障害時には影響範囲を当該プロセッサに局所化でき、ディスクサブシステム全体への影響を回避することができる。

なお、マイクロプロセッサは、H-6581-C 3用に新たに開発したもので、1.3 μmプロセスのHi-BiCMOS<sup>4)</sup>を使用している。

また、ECMは二重化し高信頼化を図っている。

(4) 電源構造

H-6581-C 3の主要構成要素である2式の制御部(クラスタ)・キャッシュ・不揮発化機構・ECMはおのおの独立のDC電源から給電する。これにより、各構成要素に障害が発生した場合、DKCとしての動作続行中でも障害・回復作業が可能となる。さらに、AC入力・制御部を二重化し、AC給・受電系統障害に関しても信頼性・可用性を高めている。

4 キャッシュ拡張機能

本章では、H-6581-C 3で開発した主なキャッシュ拡張機能について述べる。

4.1 高速書込み機能

高速書込み機能は、まずCPUからキャッシュへデータを書き込み、その後DKCが自動的にDKUへデータを書き込む。したがって、CPUはデータ転送後ただちに次の処理を実行できるため、システムの性能向上が可能となる。H-6581-C 3では新たに不揮発化機構を開発し、高いデータ保証性を確保することで、高速書込み機能を実現した。すなわち、CPUからの書込みデータはキャッシュと不揮発化機構に二重に書き込まれる。通常は、その後キャッシュからDKUへデータが書き込まれるが、DKUへ書き込む前に停電などの障害が発生した場合には、不揮発化機構からDKUへ書き込む(図4)。H-6581-C 3では、高速書込み機能として次の2種類をサポートしている。

(1) ディスク高速書込み機能

書き替え対象のレコードがすでにキャッシュ内にある場合(ヒット)に高速書込み処理を行う。書き替え対象のレコードがキャッシュ内にない場合(ミス)には、当該レコードのトラック上での位置を確認する必要があるため、通常の手続きと同様に、DKUへ直接書き込む。

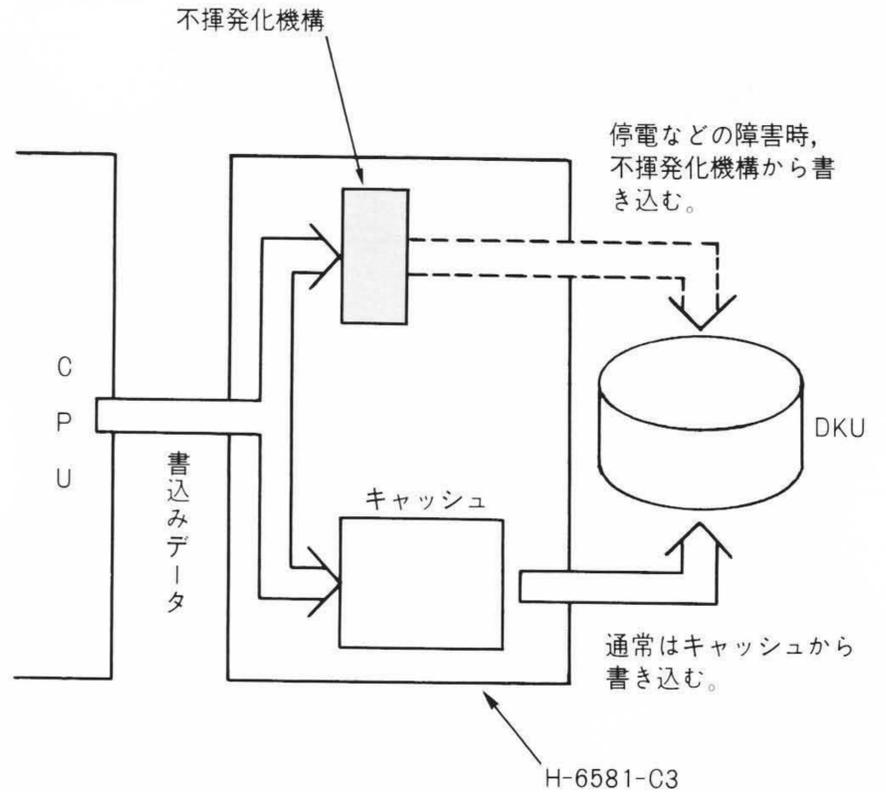


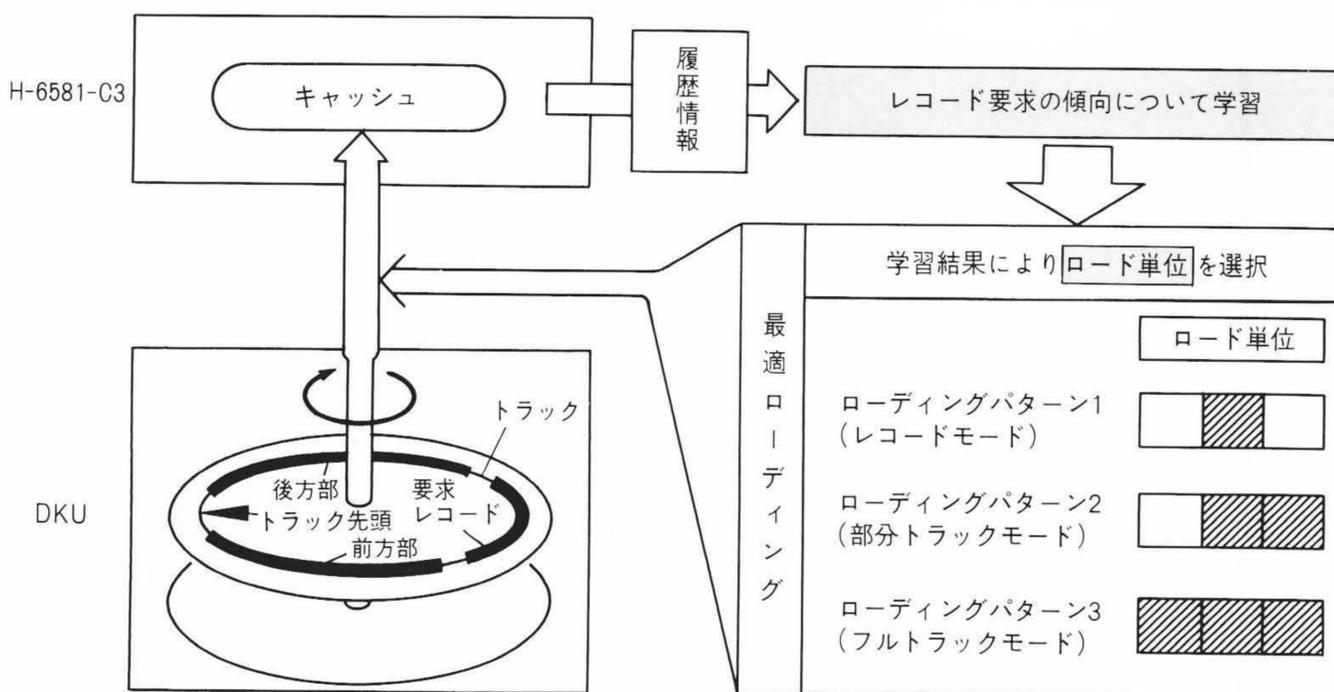
図4 高速書込み機能 書込みデータは、キャッシュと不揮発化機構に二重に保持される。CPUは書込みデータ転送後、次の処理を行うことができる。

(2) 拡張高速書込み機能

等長ブロックの属性を持つファイルの更新では、キャッシュミスでもトラック上での位置は計算可能であることに着目し、高速書込み機能の適用を可能とした。オペレーティングシステム(VOS 3 /AS : Virtual-storage Operating System/Advanced System Product)がファイルの属性を確認のうえ本機能を使用する。

4.2 学習ローディング機能

読取り動作でのキャッシュミスの場合、要求されたレコードとその同一トラック内の周辺レコードを、DKUからキャッシュに読み込む(ローディングと呼ぶ)。ところがヒット率が



注: [斜線] はロード単位を示す。

図5 学習ローディング機能 CPUからのレコード要求の傾向について学習する。その結果により、3種のパターンの中から最適のローディングパターンを選択する。要求レコードだけをアクセスする傾向が強い。…パターン1 要求レコードと、それ以後をアクセスする傾向が強い。…パターン2 要求レコードと、その前後をアクセスする傾向が強い。…パターン3

低いと、ローディングが高頻度で発生し、性能向上が十分達成できない場合があった。このため導入時に、キャッシュを適用するDKU内記憶領域を選択するというチューニング作業が必要であった。

この問題を解決するために、学習ローディング機能を開発した。本方式では、CPUからあるレコードが要求された後に発行される周辺レコードに対する要求の傾向について学習する。ローディング対象として、要求レコードに加えて、今後要求される確率が高い周辺レコードを学習結果を用いて選択するため、ローディングの最適化が可能となった(図5)。

#### 4.3 ディスク二重書き機能

従来HITACシステムでのディスク二重書きは、オペレーティングシステムが二重に書き込み要求を発行する、ソフト二重書き方式を採用してきた。本方式では、CPUのチャンネルとDKCを含めて二重化が可能である。

今回、上記方式に加え、CPUからの1回の書き込み要求に対し、DKCが自動的に異なる2台のDKUへ同一のデータを書き込むハード二重書き方式もサポートすることにした。

#### 4.4 ダンプ機能の高性能化

ダンプ時間の短縮と、ダンプ用磁気テープの本数削減を目的とした差分ダンプ機能を、従来機(H-6581-1)からサポートしている。差分ダンプ機能は、前回ダンプ処理を行った時点以後に更新された部分(トラック)だけをダンプする機能である。H-6581-C3では、これに加え、アイドルダンプ機能を開発した。

従来、オンライン稼動下でのダンプ処理は、主な業務であるオンライン処理に悪影響を与えないよう、ダンプ対象を細分化し、かつ一定時間間隔で実行していた。このため、ダン

プ処理の高速化には限界があった。アイドルダンプ機能は、オンライン処理の合間にDKUからデータをキャッシュへ先読みし、同じくオンラインの合間にキャッシュからCPUへダンプデータを高速で転送する方式をとる。これにより、オンライン業務への影響を最小限にとどめ、ダンプ処理の高速化が可能となった。

### 5 評価

#### 5.1 性能

一般に、ディスクサブシステムの性能の尺度は、ディスクサブシステムに対する読取り・書き込み要求(I/O要求)を、CPUが発行してから必要なデータ転送が完了するまでの時間(応答時間)とされている。キャッシュ付きディスクサブシステムの性能は、次のような要因で大きく変動する。

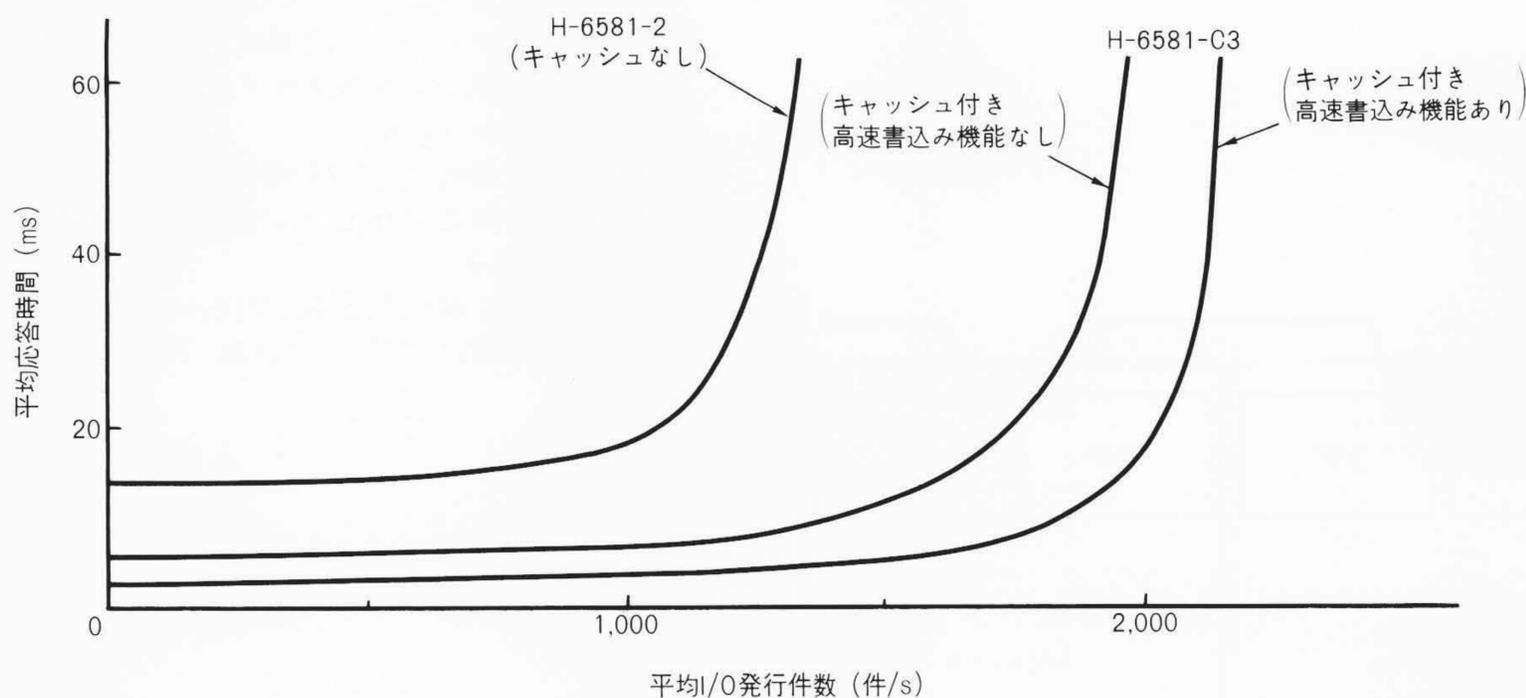
- (1) I/O要求の発生頻度
- (2) 必要とするデータの長さ
- (3) 読取り要求と書き込み要求の比(リード・ライト比)
- (4) キャッシュヒット率

さらに、このキャッシュヒット率も、キャッシュ容量、必要とするレコードのDKU内位置などによって大きく変動する。

H-6581-C3による性能改善例を図6に示す。同図により、キャッシュを装備することによって、応答時間、スループット(単位時間当たりのI/O要求件数)ともに改善され、さらに高速書き込みをサポートした場合は、応答時間をほぼ半分に短くできることがわかる。

#### 5.2 信頼性・可用性・保守性

信頼性・可用性・保守性の向上のため、H-6581-C3では次のような技術を採用した。



注：シミュレーション条件 ●DKU=64台(H-6587-3), ●キャッシュ~CPU間=9 Mバイト/s, ●DKC~DKU間=4.2 Mバイト/s  
●ブロック長=4 kバイト, ●リード・ライト比=4:1, ●リードヒット率=80%, ●ライトヒット率=100%

図6 性能シミュレーション H-6581-C3の性能をシミュレーションした一例を示す。

(1) 信頼性・可用性

H-6581-C 3では、ハードウェア論理のレベルからDKCの機

表2 高信頼性・可用性のための機能・構造 ハードウェア論理レベルから機能レベルまで、各種の機能・構造を採用している。

レベル	対象	機能・構造	
ハードウェア論理レベル	内部バス	制御系	パリティチェック
		データ系	パリティチェック + サムチェック
	メモリ用ECC	キャッシュ, 不揮発化機構	1, 2, ビットエラー修正 3ビットエラー検出
		マイクロプロセッサ用メモリ, ECM	1ビットエラー修正 2ビットエラー検出
構成要素レベル	ECM	二重化	
	基本クロック*	二重化	
	AC入力部	二重化(オプション)	
機能レベル	不揮発化機構	書込みデータの二重保持, 不揮発化	
	障害発生時	該当構成要素を切り離して処理を継続する(自動縮退)。	
	回復時	回復した構成要素を自動的に組み込み使用する(自動回復)。	

注：\* 各クラスタ, キャッシュ, 不揮発化機構, ECMはおのおの独自の基本クロック部を持っている。

略語説明 ECC(Error Correcting Code), ECM(Extended Control Memory)

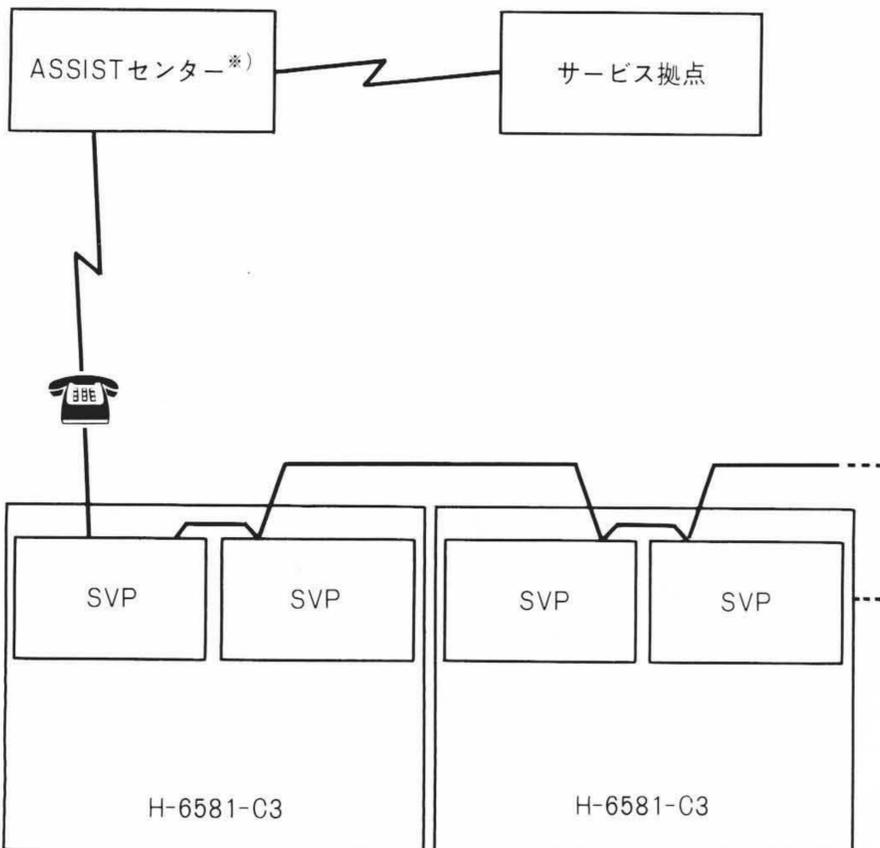


図7 遠隔保守システム H-6581-C 3のSVPによる遠隔保守システムを示す。複数台のDKCを接続して、ASSISTセンター\*と通信が可能である。

能レベルに至るまで、おのこのレベルに適合した信頼性・可用性向上のための機能・構造を備えている(表2)。ディスクサブシステム, ひいてはコンピュータシステムの信頼性・可用性を向上させることが可能となった。

(2) 保守性

H-6581-C 3では、ディスクサブシステム全体の保守性を向上させるためSVPを搭載している。SVPはディスクサブシステム内の各種情報を収集・記録(ロギング)する。保守員は、ロギング情報を用いて各DKC, DKUの品質状況を容易に知ることができ、障害の早期修復だけでなく的確な予防保守が可能となった。

さらにSVPは通信機能を持っており、通信回線を経由してH-6581-C 3をASSISTセンター\*に接続でき、サービス拠点からのディスクサブシステムに対する遠隔保守を実現することができた(図7)。

6 結 言

ディスクサブシステムに対する高性能化・高信頼化などの幅広いニーズに対応するため、新形キャッシュ付きディスク制御装置H-6581-C 3を開発した。H-6581-C 3は、マルチマイクロプロセッサ構造をとり、高速書込み機能, 高速データ転送機能などによる高性能化を実現している。また、ハードウェアの冗長度を高め、自動縮退・回復機能のサポートによって信頼性・可用性を高めることができた。

今後も、ディスクサブシステムの高性能化・高信頼化のための技術開発に取り組み、顧客要求にこたえていく考えである。

参考文献

- 1) 大山, 外: 磁気ディスクドライブ内蔵キャッシュメモリの開発, 電子情報通信学会秋季全国大会論文集, D-106(平2-10)
- 2) 佐藤, 外: キャッシュ付き磁気ディスク制御装置のマルチプロセッサ制御方式, 情報処理学会第41回全国大会講演論文集(6), 65~66(平2-9)
- 3) 横畑, 外: 磁気ディスク制御装置用マルチプロセッサ対応実時間OS, 情報処理学会第41回全国大会講演論文集(6), 67~68(平2-9)
- 4) 上遠野, 外: 超高速Hi-BiCMOSゲートアレー “HG29M100/HG29A32/HG21T30”, 日立評論, 71, 12, 1219~1222(平1-12)

\* ASSISTセンター: 日立電子サービス株式会社の遠隔保守センターの略称である。