

# サーバ用高性能ハードディスク

## High-performance Hard Disk Drives for Servers

### あらまし

富士通は、エンタープライズ型サーバ、ワークステーションの高性能、高信頼性に応える高性能ハードディスク(HDD)として、3.5インチAllegroシリーズを開発している。1994年に、世界に先がけてMR(磁気抵抗効果型)ヘッドを搭載したAL-1(Allegro-1)を出荷し、その後大容量化、高性能化を進めAL-2、AL-3、AL-4と開発し、今回AL-5シリーズを開発した。本稿で紹介するAL-5シリーズは、3.5インチのフォームファクタで41 mmの高さに3インチディスク媒体10枚を納め、10,000 rpmで36 Gバイトの大容量を達成したAL-5E、および25 mmの薄さで5枚のディスク媒体を使用し18 Gバイトを達成したAL-5LEである。これらの性能は、高性能なG-MR(Giant-MR)ヘッド、EPR4ML方式の高速リードチャネルなどの採用により実現した。インタフェースについては、ULTRA-2 SCSIによる80 Mバイト/秒の高速転送、ファイバチャネルによる100 Mバイト/秒の高速リアル転送をサポートした。

### Abstract

Fujitsu has developed the Allegro series of 3.5-inch high-performance hard disk drives. These drives are suitable for use in high-performance, high-reliability enterprise servers and workstations. In 1994, Fujitsu shipped the Allegro-1 (AL-1) which was the first drive in the world to have a magneto resistive (MR) head. Fujitsu then developed three series of larger-capacity and higher-performance products: the AL-2, AL-3, and AL-4. Now, Fujitsu has finished development of the AL-5 series of products.

This paper introduces AL-5 series products AL-5E and AL-5LE. The 10,000 rpm rotational speed of these drives is achieved by using 3-inch disks in a 3.5-inch form factor. The AL-5E is 41 mm thick and accommodates 10 disks for a capacity of as high as 36 gigabytes. The AL-5LE is 25 mm thick and uses five disks to provide a capacity of 18 gigabytes. The performance of these products is achieved by using a high-performance Giant-MR (G-MR) head and the EPR4ML high-speed read channel. For the interfaces, high-speed transfer at 80 megabytes per second using the ULTRA-2 SCSI and high-speed serial transfer at 100 megabytes per second using a fiber channel are supported.



関野光久(せきの みつひさ)

1977年慶応大学大学院機械工学修士課程了。同年富士通入社。以来磁気ディスク装置の開発に従事。磁気ディスク事業部第一技術部

ま え が き

ハードディスク(以下、HDD)の年間伸長率は、台数ベースでここ数年20~25%の驚異的成長を続けてきた。ネットワーク時代の到来により、PCサーバやエンタープライズサーバに使用されるSCSI(Small Computer System Interface)高性能HDD市場は2000年以降更なる成長が続くものと期待される。

富士通では高性能・高信頼性3.5インチHDDとしてAllegroシリーズを提供してきた。大容量化・高性能化の要求を受けて、この度Allegro-5シリーズを開発した。これは、新たに3.0インチサイズの小径ディスクを採用したことにより、10,000 rpmの高速回転と高速アクセスを実現したものである。また、G-MR(Giant-MR)ヘッドなど最新の記録技術を採用したことにより、ディスクを小径化したにもかかわらず、最大36 Gバイトの大容量を達成した。

以下、Allegro-5シリーズの技術と特長を紹介する。

サーバ用ハードディスクの動向

SCSI機は出荷台数ではWorld Wideベースで全HDD出荷数の約1割を占めるに過ぎない。しかしエンタープライズサーバやネットワークサーバなど主要なシステムでは高性能HDDに対する期待が大きく、重要なマーケットである。

HDDはMR(磁気抵抗効果型)ヘッドなど新技術の出現により、従来にも増して大容量化が進んでいる。このトレンドはATA(AT-Attachment)インタフェースのデスクトップ用HDDも例外ではなく、性能面ではSCSI機に肩を並べてきた。このため、大容量化だけでは今後SCSI機の成長は望めない状況となってきた。SCSI機がハイエンド機の座を維持・発展するためには、さらなる高性能化と高機能化が必要である。

インタフェースの高速化については、高速プロトコルチップ、電波ノイズ、伝送系の改善、低消費電力化といった技術開発が必須となる。サーバやハイエンドシステムでは数台~数十台のHDDが単一ロッカに実装される場合が多く、SCA-2などの活電保守タイプのコネクタが多用されている。システムの小型化に伴って、HDDの相互干渉やケーブル実装技術、空冷技術の向上も並行して必要となる技術である。

これらの要求に対応するため、HDDは数々の機能拡張・高性能化が図られている。最近のHDDの技術革新としては以下のものが挙げられる。

高速回転

図-1はSCSI機HDDの回転数別トレンドを示す。3.5イ

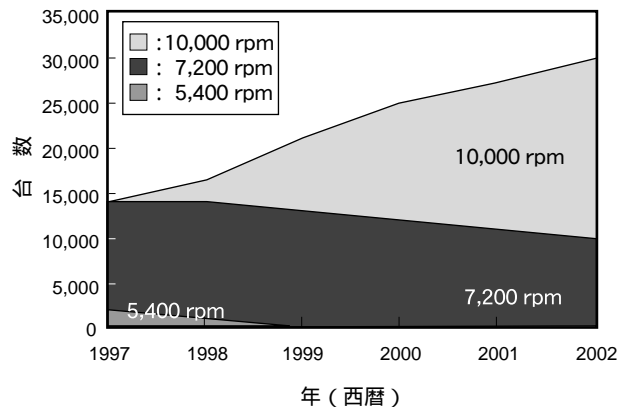


図-1 SCSI HDD出荷予想(W.W.)  
Fig.1-Shipment of SCSI HDD(World Wide)

ンチHDDの回転速度は高速化の一途である。現在は7,200 rpmが主流であるが、1999年以降は10,000 rpm以上が主流になる。

回転数の高速化には、スピンドルの振動・ディスクフラッタ・騒音の増大、消費電力・発熱の増大といった問題を解決する必要がある。ディスクのダウンサイジングにより、これまでは回転に伴う振動や発熱はそれほど深刻な問題ではなかった。ディスク回転に伴う風損およびフラッタはそれぞれ、

$$\begin{aligned} \text{風 損} &= 2.8 D^{4.6}, \\ \text{フラッタ} &= 2 D^5 \end{aligned}$$

で表される。ここで  $\omega$  は回転数、Dはディスク径である。14インチや10インチといった昔のHDDにおいて課題であった発熱やディスクフラッタの問題が10,000 rpm化によって顕在化した。このため、高性能機ではディスク径を小さくしなければならないという気運が高まった。

ディスクの小径化はヘッド・アクセス・ストロークを小さくできることから、アクセス性能の改善にもつながる。しかし、小径化は記憶容量の減少を伴うことから、最適サイズをどこに置くかがポイントとなる。世界のすう勢は現在3.0インチに収束しているが、アクセス性能を重視して2.5インチを選択したメーカーもある。

大容量化

サーバ系のシステムでは大容量に対するニーズは依然として大きい。図-2はサーバ用HDDの容量推移を示す。大容量化には記録密度技術の向上が必須であるが、MRヘッドの改良やG-MRヘッド、低ノイズ媒体、リードチャネルなどのハイテク技術の進展により、面密度は着実に増加している。実際の装置では、搭載ディスク枚数、トラック密度(TPI)と線密度(BPI)の配分比などの違いが各社の技術やねらいの違いとして見えてくる。

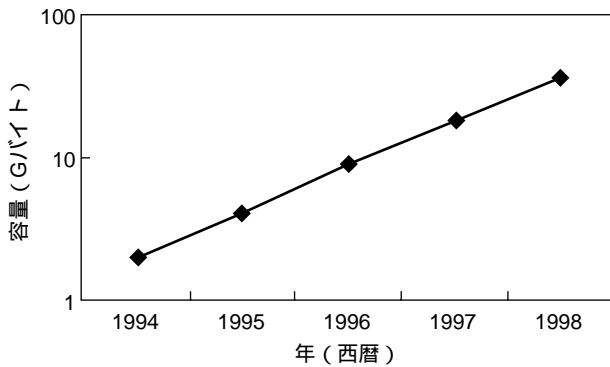


図-2 HDD容量推移  
Fig.2-Improvements in HDD capacity.

### 高速転送

記録密度の向上、ディスク回転数の増加は必然的に内部データ転送速度の高速化を押し進めている。現状の勢いでは1999年内にも50 Mバイト/秒を超えることになろうが、記録ヘッドの高速記録性能とリードチャネルLSIの高速性能がこのトレンドを律則することになろう。

### 高速インターフェース

表-1はSCSI系HDDインターフェース高性能化の推移を示す。ここで、Narrowはバス幅8ビット、Wideは16ビットである。

高性能化を達成するにはドライブとホスト間のデータ転送をいかに速くするかがキーとなる。SCSIインターフェースは、当初の5 Mバイト/秒の転送速度から現在はWideタイプで80 Mバイト/秒まで高速化が進んできた。富士通は、SCSIの標準化を推進するとともに、ULTRA-2 SCSIを含めたHDDを業界に先がけて提供している。

一方、従来のパラレル転送に加えて、高速シリアル転送をサポートするFCAL(Fibre Channel Arbitrated Loop)インターフェースが製品化された。FC(Fibre Channel)は1988年からNCITS(旧ANSIのX3T10)で検討されてきた。SUN Micro Systems社では1994年にRAIDのホストとの接続に使用した。HDDにFCが採用されたのは1995年(Seagate, Barracuda-4)である。富士通では1998年からハイエンド向けHDD(AL-4)でサポートを開始した。今後は急速にFCALが一般化するものと予想される。

### 富士通におけるディスク装置の進展

ハイエンドSCSI機として、富士通は1994年からAllegroシリーズを投入してきた。“Allegro”は3.5インチの高性能磁気ディスク装置の開発コード名であり、とくにディスクアレイ、高性能サーバ、ワークステーション用として開発された装置である。最近では年率約100%の容量増

表-1 インタフェース高性能化の推移

パラレル転送	SCSI	Fast10	Ultra-1	Ultra-2 (LVD)	Ultra-3 (将来)
Narrow (Mバイト/秒)	5	10	20	40	80
Wide (Mバイト/秒)	10	20	40	80	160
シリアル転送				FCAL	FCAL2
転送速度 (Mバイト/秒)				100	200

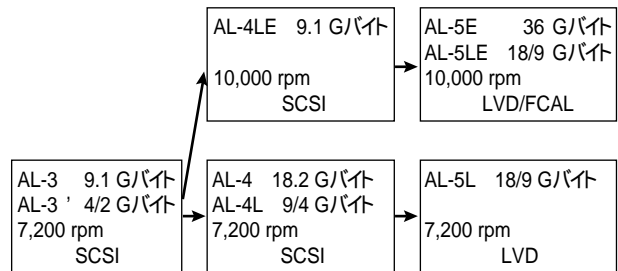


図-3 Allegroシリーズの製品推移  
Fig.3-Allegro series products.

加を達成している。1997年には18 Gバイト機Allegro-4を市場に投入したが、1999年初頭には36 Gバイト機Allegro-5を投入することになった。

最近のAllegro系列の流れは図-3に示すとおりである。Allegro-4LEから3.5インチディスクによる高速回転(10,000 rpm)を投入したが、消費電力の低減やより高いアクセス性能を実現するために、Allegro-5Eでは3.0インチディスクを採用した。

### Allegro-5シリーズの概要

ここでは、富士通の最新鋭のハイパフォーマンスHDD Allegro-5(以下、AL-5)シリーズについて、その性能と特長を概説する。AL-5LEの機構部を図-4に、製品の主要な仕様を表-2に示す。

#### 特長

AL-5シリーズは最新技術により高速大容量を達成した装置であり、以下の特長を持っている。

- (1) 36 Gバイト(AL-5E)、18 Gバイト(AL-5LE/5L)の大容量
- (2) 3インチ媒体採用で低振動、低消費電力の10,000 rpm
- (3) G-MRの全面採用による高密度記録
- (4) 高速アクセスタイム(リード平均シークタイム 5 ms)
- (5) 2 Mバイトのキャッシュを標準装備
- (6) 45 Mバイト/秒の高速転送速度

最新技術であるSpin-Valve G-MRヘッドとEPR4MLのリードチャネルの採用により、3.6 Gビット/平方インチの高記録密度を実現し36 Gバイトの大容量を達成した。

ハイパフォーマンス機(10,000 rpm)においては、従来の3.5インチ径の円板から3インチ径に変更し、風損を低減したことによって7,200 rpmと同等レベルの消費電力に抑えることができた。25 mm厚(AL-5LE/AL-4L)と、41 mm厚(AL-5E/AL-4)の各々について7,200/10,000 rpm機の消費電力比較を図-5に示す。

新型ヘッドスライダと高剛性サスペンションおよび高性能アクチュエータを採用し、シーク時の高加速度にも十分耐えられる構造とした。平均アクセスタイムは従来機に比べ2 ms以上と大幅に短縮し、5 msの高速化を達成した。

キャッシュサイズも従来の512 Kバイトから2 Mバイトに拡張し、トランザクション性能を向上させている。内部転送速度45 Mバイト/秒は世界最高速度であり、AVディスクなどで要求されるサステイン速度を大幅に向上させる。



図-4 Allegro-5LE機構部  
Fig.4-Allegro-5LE mechanism.

AL-5Lはロープロファイルの普及モデル(7,200 rpm機)であり、従来からの3.5インチ径ディスクを使用したことにより、4枚ディスクで18 Gバイトを実現した。

インタフェース

従来ハイエンドHDDはパラレルインタフェースSCSIが主流であったが、近年シリアルインタフェースFCALを使用した高速転送機も各社から製品化されるようになった。富士通ではAL-5E/5LEシリーズから本格的にFCALをサポートする。

本節では、AL-5シリーズでサポートしている、SCSI系装置の主要機能を述べる。また、FCALについては、まだ比較的新しい仕様であることから、その内容と富士通の取組みについて紹介する。

【SCSIインタフェース】

Allegro-5シリーズでサポートしている主要なSCSIコマンドを表-3に示す。

Allegro-5でサポートしている主要な機能、および性能は次のとおりである。

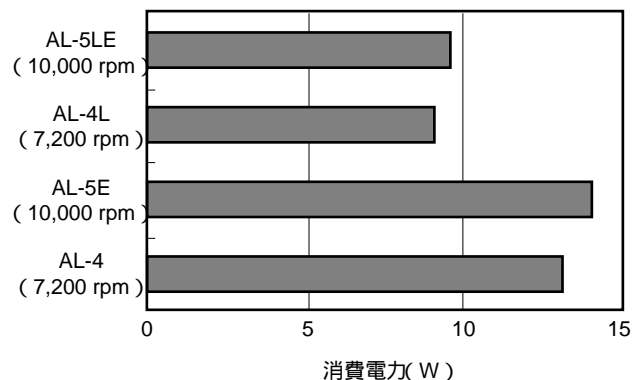


図-5 7,200/10,000 rpm消費電力比較  
Fig.5-Power dissipation of 7,200 rpm and 10,000 rpm.

表-2 Allegroシリーズの主要仕様

主要仕様	AL-3	AL-3'	AL-4	AL-4L	AL-5E	AL-5LE	AL-5L
装置高さ (mm)	41	25	41	25	41	25	25
容量 (Gバイト)	9	4/2	18	9/4	36	18/9	18/9
媒体数	10	5/3	10	5/3	10	5/3	4/2
媒体サイズ (インチ)	3.5	3.5	3.5	3.5	3.0	3.0	3.5
ヘッド	MR		MR		G-MR		
回転数 (rpm)	7,200		7,200		10,000		7,200
アクセスタイム (ms)	10	8	8	7.5	5.5	5	7
TPI (トラック密度)	6,438		9,620		13,500		13,250
BPI (ビット密度)	119,000		159,000		275,000		270,000
面密度 (Gビット/平方インチ)	0.77		1.5		3.7		3.6
転送 (Mバイト/秒)	14	14	19.5		45		32.8
インタフェース	Ultra-1		Ultra-1		LVD, FCAL		LVD

表-3 SCSI コマンド

CHANGE DEFINITION	RECEIVE DIAGNOSTIC
FORMAT UNIT	RESULTS
INQUIRY	RESERVE
LOG SELECT	REZERO UNIT
MODE SELECT (6)	SEEK (6)
MODE SELECT (10)	SEEK (10)
MODE SENSE (6)	SEND DIAGNOSTIC
MODE SENSE (10)	START STOP UNIT
READ (6)	SYNCHRONIZE CACHE
READ (10)	TEST UNIT READY
READ BUFFER	VERIFY
READ CAPACITY	WRITE (6)
READ DEFECT DATA	WRITE (10)
READ LONG	WRITE AND VERIFY
REASSIGN BLOCKS	WRITE BUFFER
RELEASE	WRITE LONG
REQUEST SENSE	WRITE SAME

## (1) 高性能

### (i) 高速データ転送速度

AL-5シリーズでは、ULTRA-2 SCSIとして標準化された新しい低電圧差動信号伝送方式LVDC (Low Voltage Differential)を採用することにより、各モードごとに以下のような高速転送が可能となっている。

- ・ ULTRA-2 SCSI LVD：最大80 Mバイト/秒(16ビットバス)
- ・ ULTRA-1 SCSI SEモード：最大40 Mバイト/秒(16ビットバス)

これにより、最大転送速度で従来方式の2倍の性能を達成している。

### (ii) リオーダーリング機能

コマンドキューイングを使用することにより、ホストシステムはHDDに対し複数個のコマンドを連続に発行することができる。複数コマンドを受け取ると、HDDは実行待ちの各コマンドごとのアクセス時間(シーク、回転待ち時間)を予測し、アクセス時間が最短となるコマンドを選択し、順次処理を行う。

このリオーダーリング機能により16コマンドのキューイング環境では、アクセス時間の約50%程度の短縮を達成している。

### (iii) キャッシュ機能

キャッシュメモリは、従来512 Kバイトが標準であった。AL-5シリーズより2 Mバイトの大容量メモリを標準搭載し、高性能化を達成した。キャッシュメモリはセグメント分割し、先読み(read-ahead)、ライトバックキャッシュをサポートしている。ライトバックキャッシュ機能では、従来はシークンシャルアクセス時のみに効果を発揮していた。ランダムアクセスにおいても、アルゴリズムの最適化を図り、キャッシュセグメントに複数コマンドのライトデータの先取りを行うことなどにより、コマンドレスポンスタイムの向上を図っている。

LBA	ユーザデータ	CRC	ECC
LBA：論理ブロックアドレスであり、IDレスフォーマットでは、従来のIDに代わり、アドレス位置情報として、読み取り位置を保証する。			
ECC：ディスク媒体欠陥のエラー検出・訂正を行う。4インタリーブ、5バーストReed-Solomon符号を採用し、最大20バイトまでの訂正が可能。			
CRC：ECC誤訂正の検出、データバッファの正当性を保証。			

図-6 データ形式

Fig.6-Data format.

## (2) 高信頼性

### (i) ユーザデータ保証

ディスク媒体上のデータは、LBA(Logical Block Address)アドレス位置情報(論理ブロックアドレス)、CRC(Cyclic Redundancy Check)、ECCの付加情報により保護されている。ディスク媒体上のデータ形式を図-6に示す。

### (ii) エラー回復機能

シークエラーやデータチェックなどのエラーが検出された場合には、最適化された内部エラー回復手順を自動的に行う。

### (iii) 故障予測通知機能(SMART)

HDD内部で、リードエラー、シークエラー、稼働時間などを監視し、徐々に進行するような障害を予測し、ホストシステムに事前に通知する機能を有している。

## (3) その他の機能(オプション)

### (i) XORコマンド

ディスクアレイ(RAID)などのアプリケーション用として、HDD側でパリティの生成を可能とするためのXORコマンド機能を備えている。

### (ii) BUS FAIRNESS

複数HDDを接続する環境では、バス使用権の優先順位の低いデバイスが頻繁にバスを使用するため、優先順位の低いデバイスのスループットが低下する。これを回避する手段としてBUS FAIRNESS機能を備えている。

【ファイバチャネル】

### (1) FCの特徴

FCはシリアルインタフェースなのでパラレルで問題になるスキューがなく、高速で長距離の通信が可能である。また信号線(光ファイバ2本、または電気的対線2組、計4本)が少ないので、コネクタが小型化でき高密度実装に適している。

装置の接続方法は、1対1接続、ループ接続、スイッチ(Fabric)を使ったネットワーク的な接続がある。これは従来のインタフェースにはない特徴である。HDDがサポートしているのはFC-AL接続で、一つのループ上には

最大126台の装置(イニシエータやターゲット)が接続可能である。さらにこのループがスイッチにつながる構成(Fabric Loop Attachment : FLA)がある。

このようなFCの特徴は、ホストに15台以上のドライブを接続するときや、マルチホストによるファイル共用に適しており、この分野ではSCSIからFCへの切り換えが始まっている。さらにこれを発展させた形として、ストレージデバイスを集積して複数のシステムで共用する方式が、Storage Area Network(SAN)として注目され始めている。SCSI, ATAとFCの主要機能の比較を表-4に示す。

### (2) インタフェースプロトコル

FCのインタフェースプロトコルは階層構造であり、様々なプロトコルを使用することができる。HDDではSCSIのプロトコルを使用している。データ転送はフレーム単位で行う。

ループ接続における実際のフレーム転送は、一定のシリアルパターン(プリミティブシグナル)をループ上に出して排他制御を行い、ループの専有権を得た装置が相手装置を指定して一時的に1対1の論理的接続を確立して、フレームの送受を行う。

### (3) 富士通の取組み

非常に広範な機能を持つFC仕様のうち、HDD用の機能を抽出したプロファイル Private Loop Direct Attach (PLDA)が定義されており、富士通も他社と同様にこのPLDAにのっとっている。

新しいインタフェースでは、互換性が問題となることが多い。とくにFCはネットワーク的な接続をするので、多数のベンダの装置が一つのネットワークにつながり、相互に通信することになる。このため、たとえ標準に準拠していても、規定の解釈などの細かい部分で矛盾が起きて動作できなくなることが懸念される。この問題を未

表-4 FC, SCSIとATAインタフェースの主要機能比較

	ATA	SCSI Ultra-3	FC
プロトコル	単純	やや複雑	複雑
コントローラ構成	単純	やや複雑	複雑
データ転送速度	66 Mバイト/秒	160 Mバイト/秒	100 Mバイト/秒 × 2
ケーブル長	46 cm	25 m (全長)	30 m (装置間)
デバイス台数	2 × 2	15	125
接続方法	1対1	ディジーチェーン	ループ

然に防ぐため、米国ではFCLC(Fibre Channel Loop Consortium)が主催する相互運用性確認のために接続試験を行う集まり(プラグフェスタ)が3~4か月ごとに開かれ、FC製品を供給する各メーカーが自社装置を持ち寄って試験を行っている。富士通は1998年6月からこれに参加しており、標準の準拠と相互運用性を保証している。

富士通は、このインタフェースの互換性を損ねずに機能強化を図っている。例えば、デュアルポートについては、従来品が一方のポートが動作中は他方のポートはビジー状態になるが、富士通では一方のポートが動作中でも他のポートはコマンドの受け付けができるようにして、ループ間の負荷のバランスを取りやすくしている。

## む す び

パソコン、ワークステーション、サーバなどの進歩は止まるところを知らず、これに使用されるHDDの重要性はますます高まることはまちがいない。これらの要求を満たすHDD開発を進めるには、高性能化・大容量化・高機能化をいかに早く実現するかがキーポイントである。富士通は、G-MRヘッド、低ノイズ媒体、高速リード/ライトチャネル、高精度位置決め技術などを開発してきた。これからもリーディング企業として、世界に先がけた技術開発と製品の提供を続ける所存である。