

## 第8回 データをより速く、より確実に、安全に

ディスク編の2回目は、ハードディスク(HDD) インターフェースと RAID のお話です。

HDD インターフェースは、他の PC パーツと同様に、時代の要請に応じて高速化、安定化、標準化が図られてきました。今回は HDD インターフェースの変遷を追いながら、現在のインターフェースの仕組みやはたらきに触れます。

また、HDD の容量が大型化するにつれ、ディスククラッシュなどの事故によるデータ消失などのダメージも大型化します。こうしたリスクの軽減とシステムの高速化を目指して RAID が注目されています。個人レベルでの導入も進んでいます。後半ではこの RAID を説明します。



### ■ HDD の誕生とインターフェース

インターフェースのお話の前に、HDD の誕生から黎明期について簡単に触れます。

世界初のハードディスクは、1956年にIBMが発売した「IBM 305 RAMAC」です。

- ・記憶容量 約 5MB
- ・転送速度 60KB/sec (8800 字/秒)
- ・ディスク回転数 1200rpm
- ・ディスク直径 24 インチ
- ・ディスク枚数 50
- ・ヘッド数 2 (独立した二つのヘッドがディスク間を上下した)
- ・サイズ 縦 75cm x 横 152cm x 高 172cm(大きめの自動販売機なみ)
- ・重量 1 トン



IBM RAMAC

記憶容量 5MB に驚かされますが、当時主流だったパンチカード 6 万枚分に相当する画期的な容量だったといえます。このディスクの開発者の中に、Alan Shugart 氏がいました。この後、彼は Shugart 社を設立しました。ここで生まれたインターフェースが、SASI (Shugart Associates System Interface の頭文字) です。本来はミニコン用のインターフェースですが、我が国では国民機といわれた PC98 シリーズのドライブに搭載されていました。1979 年には Alan Shugart 氏は Seagate 社を設立しています。翌 1980 年、SASI を規格化する動きが始まり、曲折があつて、1986 年、SCSI が ANSI で標準と認められました。SCSI の名称は ANSI が規格名に企業名を入れることを許さない方針で、意味を変えずに最小限の変更ですませた結果です。SASI と SCSI、確かによく似ているはずですが。

SASI の標準化が動き始めた当時、Shugart 社の開発に Larry Boucher 氏がいました。彼は後に、SCSI ボードで有名な Adaptec 社を設立します。

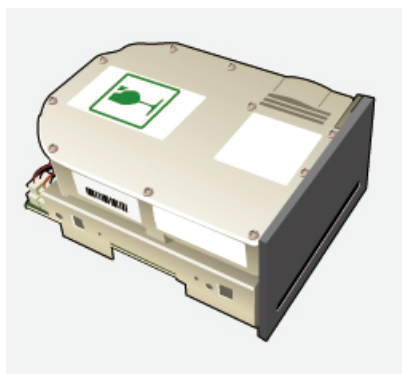
IBM は別にして、1970 年から 1980 年の間に 5 大メジャーといわれる HDD メーカーが次々に設立されました。IBM、Western Digital、Seagate Technology、Maxtor、Quantum の 5 社ですが、現在は Quantum 社が Maxtor 社に買収され、その Maxtor 社が Seagate Technology に買収され、IBM の HDD 部門が日立に買収されるなど、さま変わりしています。

## ■HDD インターフェースの変遷

RAMAC の時代から HDD と CPU を結ぶインターフェースは存在しますが、特定の機種に依存しないインターフェースが登場するのは 5 インチディスクの頃からと考えて良いようです。ちなみに、RAMAC の 24 インチは別格としてもミニコン用のディスクは 14 インチ、8 インチといった大型のディスクが採用されていました。5.25 インチディスクが登場する頃からインターフェースのあり方も少しずつ変化してきます。

### ・ST-506 と ESDI

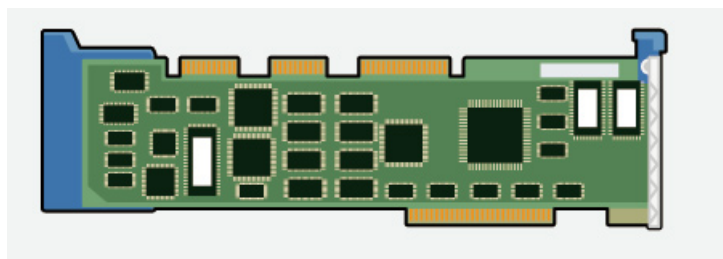
5.25 インチディスクは、1980 年 Seagate 社が発売した ST-506 が最初です。IBM-PC にも採用されました。このディスクに搭載されたインターフェースが、ST-506 です。ディスクの製品名がインターフェースの名前になっています。



Seagate ST-506

5.25 インチディスク、容量 5MB、転送速度 625KB/sec

ST-506は転送速度が625KB/secと遅く、設計上140MBの容量制限を超えることができません。この容量と速度の改善を目指して1983年にMaxtor社が中心となって規格化を進めたのがESDI(Enhanced Small Device Interface)です。



IBM ESDI アダプター 対象ディスク 5.25 インチ

転送速度 5Mbps~3MB/sec(一般的な仕様で、上記アダプターのものではありません)

ESDIが今日のATAやSCSIに大きな影響を与えているのが、コマンドによるドライブ制御を可能にしたことです。また、7台までのデバイス(うち1台はホスト)を接続することができ、デバイスのアドレスが設定できるので、コントローラーは個別にアクセスします。この3年後、1986年SCSIがANSIによって標準に認められます。さらにその2年後には、IDEへの規格作りが始まり、新しいHDDインターフェースの時代が始まります。

## ・SCSI

SCSI(Small Computer System Interface)はSASIを元に開発されたインターフェースです。元になったSASIはミニコン用のバスインターフェースでした。ANSIによる標準化を通じて、SCSIはコンピュータやドライブを含む周辺装置を接続する汎用インターフェースになりました。接続された機器でケーブルを共有するバス型インターフェースで、接続できる機器は合計7台、ケーブルの総延長は6mです。データバスは8ビット+パリティ1本、50ピンのケーブルで機器間を接続します。転送速度は5MB/secです。

しかし、SCSI-1と呼ばれる最初の規格は、内容に曖昧な点が多く、機器への組み込みの自由度が高く、問題が生じやすかったといわれます。1990年にANSIはデータ転送の高速化や機能拡張を入れるとともに、標準的なコマンドセットCCSを取り込んで規格案を作成し、1994年に正式に承認されました。これがSCSI-2です。

その後も規格は拡張が続き、SCSI-3に至って従来のSCSI-2はパラレル伝送部分を残すのみで、ファイバーチャネルやシリアルバスなどが入っています。

今日、SCSIはPCに使われなくなってきています。安価なATAインターフェース機器が十分な性能を発揮しており、PC内蔵機器にはATAが使われるようになっていきます。外付け機器の接続にはUSB 2.0やIEEE 1394といった新しいシリアルインターフェースが使われるようになっており、しだいに番が失われているのです。

しかし、一方で最新の SCSI はターゲットをハイエンドに向けていることも事実です。PC の世界に見切りをつける方向で進化しているように見えるのです。

多様な SCSI 規格

転送モード	規格	SCSI-3	バス幅	転送レート
SCSI	SCSI-1		8	5
Wide SCSI	SCSI-2		16	10
Fast SCSI	SCSI-2		8	10
Fast Wide SCSI	SCSI-2		16	20
Ultra SCSI	SCSI-3	SPI	8	20
Ultra Wide SCSI	SCSI-3	SPI	16	40
Ultra2 SCSI	SCSI-3	SPI-2	8	40
Ultra2 Wide SCSI	SCSI-3	SPI-2	16	80
Ultra3 SCSI	SCSI-3	SPI-3	16	160
Ultra160 SCSI	SCSI-3	SPI-3	16	160
Ultra320 SCSI	SCSI-3	SPI-4	16	320

ハイエンド傾向が強く感じられるようになったのが、Ultra3 SCSI からです。Ultra3 SCSI の特徴的な機能にその傾向が現れています。

**CRC (Cyclic Redundancy Check)**

Ultra ATA や Ethernet、ファイバーチャネルなどで採用されているエラーチェックです。SCSI の高帯域化に伴い、より高速な信号にも耐えるエラーチェックが必要とされています。さらに、SCSI は長いケーブルを使用する機会が多く、強力なエラーチェックで確実にデータを保護しなくてはなりません。

**ドメインバリデーション**

SCSI ホストアダプターと機器間のデータ幅の不一致やケーブルの断線、ターミネーションの異常などをチェックする基本テストと、実際の読み書きテストを通じて SCSI バスの電気特性をチェックする拡張テスト行って、確実に動作する安全性の高い SCSI システムを提供する機能です。

**パケット化 SCSI (Packetized SCSI)**

データの転送効率を大きく高める改良型の SCSI プロトコルです。非同期で非効率的な転送を行っていたコマンドやメッセージ、ステータスといった非データを、データとともに同期転送ですべてやり取りできるようにしています。非データやデータの内容を SPI 情報ユニットという情報単位にまとめ、この SPI 情報ユニットを最大ネゴシエーション速度 (Ultra320 SCSI ならば最高 320MB/sec) でやり取りします。

### **QAS (Quick Arbitration and Selection)**

どのデバイスがバスを占有するかを調停する時間を短縮し、バスの使用効率を高めます。バスフリーフェーズ(SCSI バスが使用されていないことを示す状態)を介することなく高速にバス調停を行います。パケットサイズド SCSI では、バスフリーを除くすべてのバスフェーズが最大ネゴシエーション速度で遷移します。そこで、QAS を利用して残りのバスフリーフェーズもさらに切りつめ、SCSI バスの使用効率を極限まで引き上げます。

以上、Ultra3 SCSI の主な機能を紹介しましたが、いずれも、高速な HDD を多数束ねて使う用途でこそ効果が発揮される機能です。SCSI ホストアダプターに HDD を 1 台や 2 台接続するような使い方では、効果どころかほとんど意味をなさないといいでしょう。どうやら SCSI は個人用 PC の世界からは遠ざかって行くようです。

### **•IDE(Integrated Drive Electronics)**

1986 年 Compaq PC の製品化のための企画が始まりといわれています。Western Digital 社や Seagate 社、Conner Peripherals 社が関わりました。従来、ホスト側にあったディスクコントローラーを、ドライブ側に移し、ドライブ側のロジック回路と一体化しました。ドライブとインターフェース回路を統合したのです。その結果、ドライブはディスクコントローラーを搭載したアダプターを介することなく(実際にはアダプターはありましたが、従来のものよりもずっと簡略化されました)、マザーボードに接続され、転送速度が向上しました。

しかし、ロジックが簡単ならえに標準規格が存在しなかったことで、互換性の問題が生じるようになりました。これらの問題の解決を目指して、1988 年にドライブメーカーが集まって IDE の仕様を整備し、ANSI に提案して 1994 年に承認されます。これが ATA(AT Attachment) インターフェースです。この仕様が標準となり、ドライブメーカー、ホストメーカーがこの ATA 仕様に準拠するようになります。

ケーブル 1 本あたり、最大 2 台の機器が接続可能(マスタ/スレーブ接続)で、40 芯または 80 芯のリボンケーブルを用いて接続し、ケーブル長は最大 18 インチ(45.7cm)と規定されています。ところが、IDE HDD には、504M バイトを超える容量が認識されないという問題がありました。HDD 側のパラメーターと PC/AT の BIOS のパラメーターのミスマッチが原因です。そうしたこともあって、1996 年には Enhanced IDE として拡張規格を統一した ATA 規格が、ATA-2 として ANSI に承認されています。その後も SCSI 同様に ATA も規格の拡張が続いており、現在 ATA-8 が ANSI で審議中です。

規格	承認	内容
ATA-1	1994年	IDEの規格化
ATA-2	1996年	PIO 3,4 Multiword DMA 1,2 追加による高速化
ATA-3	1997年	Singleword DMA の削除、リムーバブルメディアのサポート、S.M.A.R.T 対応
ATA/ATAPI-4	1998年	ATAPIの統合。UltraDMA 0,1,2 スキャナ、プリンタ、メディアチェンジャー等 SCSI 準拠の多種デバイスのサポート
ATA/ATAPI-5	2000年	UltraDMA 3,4のサポート。
ATA/ATAPI-6	2002年	UltraDMA 5、BigDriveのサポート
ATA/ATAPI-7		UltraDMA 6のサポート。1.8、2.5インチHDDの3.3V規格定義

IDE HDDについて知っておきたい用語と、データ転送モードについて解説しておきましょう。

### Ultra ATA (Ultra AT Attachment)

1996年に米 Quantum 社や米 Seagate Technology 社、米 Intel 社などにより提唱された IDE の拡張仕様です。業界標準(デファクトスタンダード)で IDE の公式な規格ではありませんが、その仕様は、ATA/ATAPI-4 や ATA/ATAPI-5 などに受け継がれ、規格化されています。データ転送速度と信頼性が強化されており、Ultra DMA という新しい転送方式で、最大転送速度を 33.3Mbytes/s にまで高め、CRC によるエラーチェック機能が追加され、信頼性を向上させています。

### ATAPI (ATA Packet Interface)

IDE コントローラーから SCSI 互換のパケット形式のコマンドを発行することにより、HDD 以外の機器を接続できるようにした規格。CD-ROM 等がサポートされました。

### BigDrive (ビッグドライブ)

28 ビット LBA を 48 ビットに拡張し、128PB (ペタバイト、128 ペタバイト=134,217,728GB) までの容量を扱えるようにした規格。ATA/ATAPI-6 で採用されました。LBA (Logical Block Addressing) は HDD のすべてのセクタに通し番号を振って、その通し番号によってセクタを指定する方式のことで、28 ビットの場合の上限は 128GB です。ただし、ハードウェアの規格上は 128PB まで可能とはいっても、Windows XP など 32 ビット OS では、OS の構造上 2TB (テラバイト、2TB=2,048GB) が限界です。

### PIO (Programmed input/output) 転送モード

CPU が直接 IDE コントローラー I/O ポートを経由してデータの送受信を行います。

ATA 機器はすべて機器転送速度、転送モードのネゴシエートのために PIO Mode 0 をサポートしています。Mode 0 (3.3MB/sec) ~ Mode 4 (16.7MB/sec)



### Singleword DMA 転送モード

PC 本体に搭載されている 8bit の DMA 転送が可能な DMAC を用いて転送を行います。ATA/ATAPI-3 規格において廃止されています。Mode 0(2.1MB/sec) ~Mode 2(8.3MB/sec)

### Multiword DMA 転送モード

PC/AT で拡張された 16bit の DMA 転送が可能な DMAC を用いて転送を行います。  
Mode 0(4.16MB/sec) ~Mode 2(16.6MB/sec)

### Ultra DMA 転送モード

ATA/ATAPI-4 以降で追加されたチップセットや UIDE コントローラーカードに搭載された、専用の高速な DMAC を使用して転送を行います。  
Mode 0(16.7MB/sec) ~Mode 6(133.3MB/sec)

## ・シリアル ATA

パラレル転送方式による速度向上が技術的に困難になってきたため(詳しくは第4回「バスの基礎知識」をご覧ください)、ATA インターフェースの転送速度を向上させるための規格として登場しました。

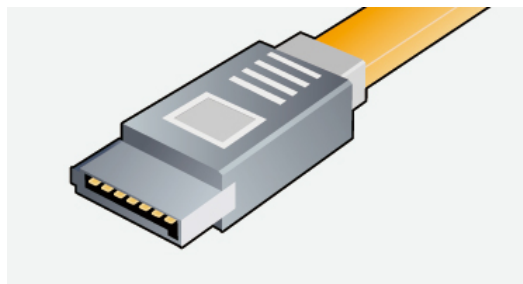
### シリアル ATA 1.0

2003 年に規格化され、150MB/sec のデータ転送を行います。

ホットスワップに対応し、信号の伝送に使用する電圧はパラレル ATA の 5V から 0.5V に低下しており、消費電力と信号の干渉の低減、ケーブル長の延長を実現しています。

ケーブルは 7pin で、最大長は 1m、1 本のケーブルに 1 台のデバイスを接続します。

従来のパラレル ATA とはプロトコルレベルでの互換性を持っていますが、ATA のマスタ/スレーブ接続の概念は廃止されています。



シリアル ATA ケーブル

### シリアル ATAII

転送速度 300MB/sec を実現しています。シリアル ATA II は、Serial ATA 1.0a を基にして、2004 年ごろまでの技術的な拡張全体のことをいい、サーバやワークステーション、ネットワークストレージ向けの拡張機能が主となっています。NCQ(Native Command Queuing) やマルチポートなどの概念を取り入れています。

### **NCQ (Native Command Queuing)**

シリアル ATA II で盛り込まれた機能です。HDD のコマンド実行の完了を待たずに複数のコマンドを連続して発行し、そのままキュー(待ち行列)にためていきます。これにより、長い時間を要する HDD の機械的動作の完了をコマンドごとに待つことなく、次々と HDD にコマンドを送り込むことができます。しかも、ためて順番に実行するのではなく、キューに格納された複数のコマンドを、HDD の動作効率の良い順番に並べ替えてから実行に移します。これによって、HDD の機械的動作によるオーバーヘッドを最小化します。

### **シリアル ATA III**

2007 年に規格化が予定されており、転送速度 600MB/sec の実現が期待されています。

### **eSATA (External Serial ATA)**

Serial ATA 1.0a の拡張規格で外付けドライブ向けに定義されています。一部で対応製品は出ていますが、同様の機能を実現する USB や FireWire ほど普及していません。

誤接続を防ぐため、eSATA のコネクタ形状はシリアル ATA のコネクタ形状とは違うものになっています。接続ケーブルの長さは最大 2m。パソコンの電源を入れたまま、接続ケーブルを抜き差しできるホットプラグに対応し、USB2.0 接続の 2 倍以上の速度で通信可能です。



## ■ RAID(Redundant Arrays of Independent(Inexpensive) Disks)

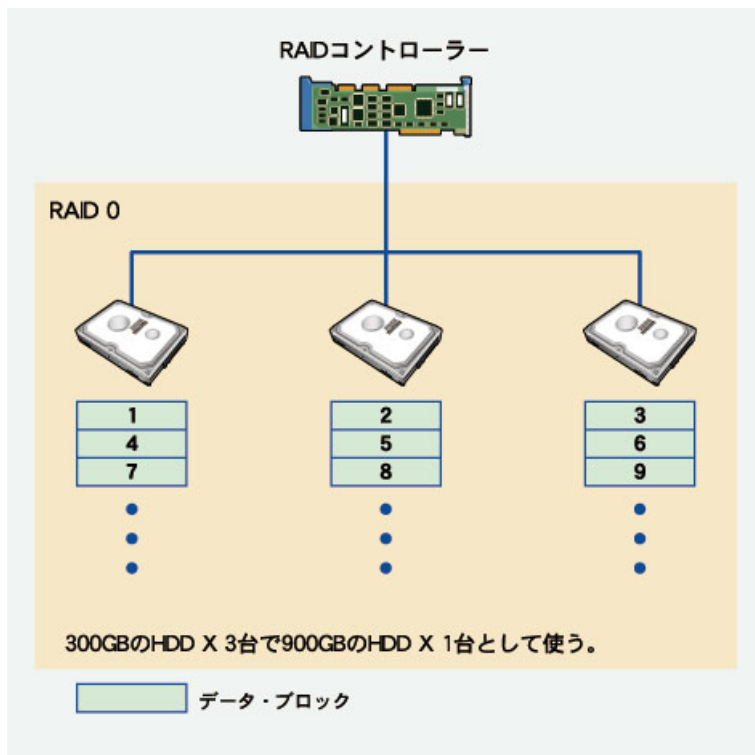
RAID は、ディスククラッシュなど不意のディスクの故障などからデータを守ったり、あるいはディスクシステムを高速化したりする技術です。

IDE RAID の登場でハードウェア RAID 製品の初期導入コストが下がっており個人レベルの PC でも RAID を導入しやすい環境になっています。

しかし、RAID の仕組みは比較的複雑です。今回は、「RAID レベル」を中心に RAID の原理を概説します。RAID レベルは、数字の大小によって性能やグレードを表すものではなく、RAID の実現方式を表すものです。用途に合った最適な RAID レベルを選ぶための基礎知識として、RAID レベルを整理しながら理解するようにしましょう。

### ・RAID 0(ストライピング)

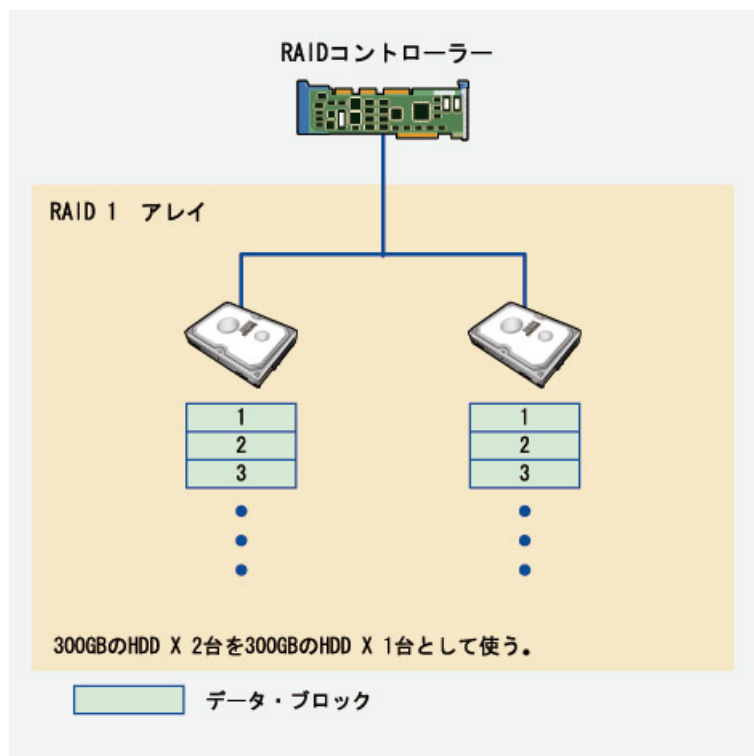
RAID 0 では、2 台以上のディスクを組み合わせ、ディスクに対する読み書きの処理を複数のディスクに対して同時並行的に実行することで、アクセス速度を高速化します。



ディスクが 3 台あれば、データを 3 つに分けてそれぞれのディスクに対して同時に書き込むことにより、理論的には 1 台の場合の 3 分の 1 の時間で書き込むことができます。複数のディスクをあたかも 1 台のディスクであるかのように扱うことで大容量の単一ストレージを実現します。この方式は、ディスクのうち、1 台でも壊れてしまうと全体へアクセスできなくなります。つまり、冗長性(Redundancy)をまったく備えていないので、ディスク台数が増えると、かえって耐障害性は悪化します。

### ・RAID 1(ミラーリング)

RAID レベルの中で、最も単純な手法でディスクの耐障害性を高めています。その手法とは、同一のデータを複数のディスクに書き込み、一方のディスクが故障しても、他方で処理を続行できるようにします。同じデータを格納したディスクの「コピー」を用意することで、故障に備えています。通常は2台のディスクを使って実現します。



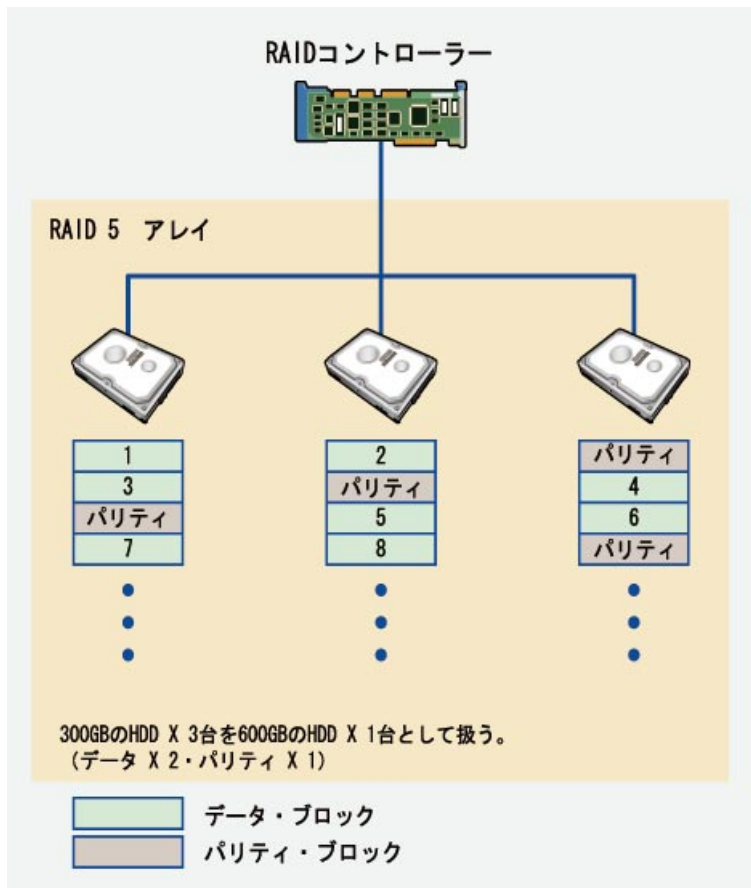
データはアレイ内の各ディスクにまったく同一のデータが書き込まれます。1台のディスクが故障しても、残りのディスクが稼働し続けることで、データは失われません。

ただし、同一のデータを2台以上のディスクに書き込むため、ディスク容量の利用効率は50%以下になってしまうというデメリットがあります。また、2台のディスクに同一のデータを書き込むため、ディスク1台の場合と比べ、書き込み時のオーバーヘッドが大きくなり、速度も遅くなります。

### ・RAID 5(分散データ・ガーディング)

耐障害性の向上と高速化、大容量化のすべてを実現できます。

ディスクの故障時に記録データを修復するために「パリティ」と呼ばれる冗長コードを、全ディスクに分散して保存するのが特徴です。



RAID 0と同様にデータを分割して各ディスクに格納しますが、データ・ブロックの組(上図で例えば1/2や3/4、5/6)ごとにパリティが生成される点がRAID 0と違います。1台のディスクが壊れても、残りのディスクに格納されたデータとパリティから、失われたデータを復活させることができます。

複数のディスクにデータを分散して書き込むと同時に、パリティも計算・生成して同時にディスクに書き込みます。負荷を分散するためパリティ用ディスクは特に決まっておらず、全ディスクに分散して書き込まれます。どれか1台のディスクが故障しても、それ以外のディスクのデータとパリティ情報から、元の完全なデータを生成して回復できます。ただし、2台以上のディスクが故障すると回復できません。

ディスク台数が多いほど容量の利用効率も向上します。RAID 1(ミラーリング)と比較した場合、この利用効率の高さがRAID 5のメリットの1つです。

## ・余り使われない RAID レベル

RAID レベルが 0、1、5 と番号がとんでいます。実は抜けている 2、3、4、6 という RAID レベルも存在します。しかし、あまり使われていません。概要を説明しましょう。

### ・ RAID 2

元データとともにエラーを修復するための冗長コードを、複数のディスクに記録します。この冗長コードは、主記憶装置などで使用される ECC (Error Correcting Code : 誤り訂正符号) で、パリティよりも複雑で生成に時間がかかります。しかも ECC による冗長コード自体のサイズが、元のデータに対してさらに大きくなりがちで、容量面でもデメリットが目立ち、市販の RAID 製品にもほとんど実装されていません。

### ・ RAID 3

RAID 3 も元のデータに冗長コードを加えて複数のディスクに記録します。冗長コードには RAID 5 と同じパリティを使用します。

データは、ビットまたはバイト単位に分割されて、専用の複数のディスクへ同時に書き込まれます。そのデータからパリティが生成され、パリティ専用のディスクに書き込まれます。常にすべてのデータディスクに並行してアクセスし、データを一括して転送する方式なので、転送速度は高速ですが、そのために、ほかのディスクと同期して回転させたり、データ転送したりするための特殊な仕組みが必要となるなど、コスト面でほかの RAID レベルより不利です。このため、RAID 3 を実装している製品は多くありません。

### ・ RAID 4

RAID 0 のストライピングに、パリティ専用ディスクを追加してデータを再生成する機能を持たせたものです。ブロック単位で複数のディスクに記録しますが、パリティは特定のディスクだけに格納されます。データとパリティそれぞれを格納するディスクが分かれています。そのため、データの書き込みが発生すると、パリティ用ディスクへの書き込みが集中して、そこがボトルネックとなって書き込み性能が下がります。これをパリティの分散書き込みで改善しているのが RAID 5 です。つまり、RAID 4 は、RAID 5 で代替されるため、サポートしている製品はほとんどありません

### ・ RAID 6

RAID 5 の改良版で、1つのデータ・ブロックにつき 2つのパリティを生成し、別々の HDD に記録することで、同時に 2 台のハードディスクが故障しても、元のデータを修復可能とします。しかし、パリティが増えた分だけ、計算や書き込みのオーバーヘッドも増加するため、書き込みの性能は高くありません。パリティ用に 2 台分のディスク容量を必要とするため、ディスクの利用効率は RAID 5 より低下します。現在のところ RAID 6 を実装している製品は多くありません。

・RAID レベルの組み合わせ

RAID 0/1/5 はそれぞれ異なるメリット/デメリットを持っています。

そこで複数の RAID レベルを組み合わせ、ある RAID レベルが不得意とする機能は別の RAID レベルでカバーして、総合的にデメリットの少ないシステムを実現するというものです。

RAID レベルの特徴と組み合わせによる特徴の変化を表にしました。

		RAID 0	RAID 1	RAID 5	RAID 10 <sup>*1</sup>	RAID 50 <sup>*2</sup>
耐障害性		なし	◎	○	◎	◎
性能	シーケンシャル読み出し	◎	△	◎	◎	◎
	シーケンシャル書き込み	◎	○	△	◎	◎
	ランダム読み出し	◎	○	◎	◎	◎
	ランダム書き込み	◎	○	△		○
	リビルド時*1		◎	△	◎	○
必要なディスク台数		2 台以上	2 台	3 台以上	4 台以上	6 台以上
ディスク容量の利用効率 (n : ディスク台数)		100%	50%	100× (n-1)/n%	50%	100× (n-1)/n%
実装にかかるコスト		小	中	中	大	大

\*1 RAID 10 : RAID 1 と RAID 0 を組み合わせた RAID レベルです。RAID 0+1、RAID 1+0 と呼ぶこともあります。

\*2 RAID 50 : RAID 5 と RAID 0 を組み合わせた RAID レベルです。RAID 0+5、RAID 5+0 と呼ぶこともあります。

単一の RAID レベルに比べ、ディスク台数が多くなりがちですが、複数の RAID レベルの組み合わせは、耐障害性の向上と高速化、大容量化を同時に実現する手段としてよく利用されます。実際、多くの製品が、ハードウェア RAID コントローラーには、RAID 0/1/5 という単体の RAID レベルだけではなく、これらを組み合わせた RAID レベルをサポートしています。

ただし、RAID レベルを組み合わせた RAID には、呼称からしくみまでベンダーによって異なっているのが実情で、これが標準というものがありません。呼称が同じでも原理が異なっている例さえあります。選択にあたっては注意が必要です。

RAID の構築は簡単ではありませんが、ソフトウェアも充実してきており、ハードウェアも扱いやすくなっています。何よりも、今回のテーマである HDD の低価格化は魅力です。

次回は、ネットワークと HDD、そして PC 以外の用途で活躍する HDD のお話しです。