UFS EXPLORER
DATA RECOVERY AND ACCESS SOFTWARE

Knowledge Base

These articles provide basic information and helpful recommendations concerning data access and recovery.

Articles

RAID: Data organization and recovery

Different RAID levels apply different data organization techniques to serve different purposes. Each level of RAID systems has its own advantages and disadvantages in performance and operation.

Data recovery chances from complex RAID systems depend much on the level of RAID array and actual causes of data loss and RAID failure, if applicable. Nevertheless, you may influence the situation and increase data recovery chances in

that you choose efficient and reliable data recovery software working with any RAID level and capable of resolving even complicated tasks.

Content:

RAID array terms

Technical information concerning RAID arrays is generally provided in special terms describing these types of storages. Below you will find most commonly used terms referred to complex RAID storages.

- **RAID** - Redundant Array of Independent Disks. The term denotes a storage scheme where a set of independent storages is combined into one. Depending on actual data organization of the storage, this scheme increases its stability, performance and/or reliability.

- **Hardware RAID** - hardware-driven RAID. Hardware RAID consists of RAID controller chip or board that operates the array and a set of attached disks. Operating system detects complex RAID storage as one solid storage device. Data are operated by hardware controller and can be configured with its BIOS settings.

- **Software RAID** - software-driven RAID. Software RAID doesn't use any hardware and is created on a set of independent storage units by OS. Operating system detects software RAID storage as one solid storage device. Data are operated by OS drivers using CPU time without requiring additional hardware (e.g. NT LDM software RAIDs, Linux, BSD, macOS RAIDs such as LVM/LVM2 or MD).

- **Virtual RAID** - hardware or software RAID virtually reconstructed from its

- **Virtual RAID** - hardware or software RAID virtually reconstructed from its components. This is a virtual storage created by data recovery software to simulate original RAID storage for data recovery purposes.

- **RAID Component (unit)** - disk or disk partition used as a data storage for RAID storage.

- **Mirroring** - data organization technique based on data replication onto separate units. Mirror creates a complete copy of one unit and uses another unit to store this copy. This ensures high fault-tolerance: If one unit fails there remains another copy of data located on another unit(s) of this RAID. Mirroring technique is implemented in RAID level 1.

- **Striping** - data organization technique based on distribution of data fragments among units. Data striping allows users to increase Input/Output (I/O) performance of RAID storage significantly. Data on RAID are divided into small parts (stripes) and distributed across all available units. Striping speeds up storage performance due to parallel reading/writing to all units independently. Striping is implemented in RAID level 0.

- **Parity** - data organization technique based on adding a bit of data from different RAID units to a separate unit. Parity allows increasing fault-tolerance of a storage: in case of any drive failing, the contents of the failed drive can be reconstructed on a replacement drive by subjecting the data from the remaining drives but on condition that no more than one drive fails.

- **Reed-Solomon code** - error-correcting algorithm based on Galois Algebra. Reed-Solomon code allows increasing reliability of RAID storage. This algorithm is used in RAID level 6.

RAID array types (levels)

Levels of RAID storages differ in data organization techniques implemented in it. Below the detailed description of most widely used RAID types with information as to advantages and disadvantages of each level is presented.

RAID, level 0 (RAID0, data striping)

RAID level 0 is the best example of data striping as it is. The term RAID – Redundant Array of Independent Disks – does not actually explain the functionality of this RAID level because it does not imply redundancy at all. This type of storage may consist of two and more units. Stripes are defined as data fragments where each next data stripe is located on *subsequent storage unit*.

0	2	4	...	N-1
1	3	5	...	N

Figure 1. Stripes organization on RAID 0 (2 units)

Figure 1 shows data striping employed in RAID 0. Such data organization scheme allows to speed-up I/O operations up to **U** times (where **U** is the number of **units**in RAID0). This is achieved by sending concurrent or consequent I/O requests to different units (usually different hard disks). *For example*, to read stripes 0..3 (data segment with size of 4 stripes) the controller sends 2 concurrent read requests: to read two first stripes from Unit 1 and two first stripes from Unit 2. Units implement physical reading simultaneously and the controller gets the result two times faster.

Figure 1 shows data striping employed in RAID 0. Such data organization scheme allows to speed-up I/O operations up to **U** times (where **U** is the number of **units**in RAID0). This is achieved by sending concurrent or consequent I/O requests to different units (usually different hard disks). *For example*, to read stripes 0..3 (data segment with size of 4 stripes) the controller sends 2 concurrent read requests: to read two first stripes from Unit 1 and two first stripes from Unit 2. Units implement physical reading simultaneously and the controller gets the result two times faster.

This method of data organization allows using *almost* the whole storage space for data leaving no redundancy on *data area*. However, the capacity of RAID 0 storage is sometimes less than the sum of sizes of individual units because the controller can reserve some storage space for 'own technical needs' or due to different size of storage units making the controller use as much space of each unit *as the smallest unit has*.

unit has.

Advantages of RAID, level 0:

- Extremely high performance in both read and write;
- Simple implementation (even most on-board SATA controllers support RAID 0);
- Up to 100% disk space for data;
- The most cost-efficient RAID solution.

Disadvantages

- Absence of fault-tolerance: Failure of a single unit causes data loss.

Recovery perspectives

- Controller failure/disassembled array: With information about stripe size and units order you can easily recover lost data.
- Damaged unit: As soon as any of the units become unreadable, data recovery for subsequent data segments laying beyond the StripeSize*(UnitsCount-1) is impossible.

RAID, level 1 (RAID1, data mirroring)

RAID level 1 implements the technology of data mirroring. Mirroring creates an exact copy of the data and stores it on a separate disk designated for this. The capacity of RAID1 equals to the size of the smallest storage unit minus the space that can be reserved by the controller. When the controller reads data from RAID 1, it can send

requests to any of the disks to speed-up I/O operation. Writing operation works either in a parallel mode (to both disks simultaneously – faster) or consequently (onto one disk after another – fault-tolerant). RAID 1 doesn't include data segmentation.

Advantages of RAID, level 1:

- Fast read operations;
- Increased fault-tolerance;
- Continuing operation even at least one mirror disk remains (in a 'degraded mode');
- One of the most available solutions supported by most on-board SATA controllers.

Disadvantages

- The most inefficient use of disk space;
- Slow writing operation.

Recovery perspectives

- Controller failure/disassembled array: easy to recover all data from any unit;
- Damaged unit: data can be recovered from any readable unit.

RAID, level 4 (RAID4, stripe set with dedicated parity)

RAID 4 is the first and attempt to compromise between fault-tolerance, speed and cost. The technique implemented in RAID 4 is based on usual Stripe Set (like in RAID 0) extended with one more designated unit to store parity information for error control. This array may consist of 3 or more disks. Actually, this scheme of data storing is also implemented in RAID level 3 with the difference in striping method being byte-level for RAID3 and block (sector) level for RAID 4.

0	2	4	...	N-1
1	3	5	...	N
P	P	P	...	P

Figure 4. Stripe Set with dedicated parity (RAID 4)

Figure 4 displays the method of fault-tolerance working. Stripe set stores actual RAID data. Each 'column' of stripes is summed with XOR to achieve parity.

RAID 4 has such common with RAID 0 features as fast reading of operations and large storage capacity, at the same time this RAID level incorporates its own feature of extended internal error correction. If some stripe becomes unreadable, the controller is capable of reconstructing it basing on information of other stripes and parity. The disk designated for parity is not used for data storage rather as a backup unit.

Advantages of RAID, level 4:

- Ever-faster reading operations;
- High fault-tolerance;
- Operating in a 'degraded mode', if one of the disks fails, basing on parity;
- Efficient solution in the aspect of fault tolerance for money.

Disadvantages

- Remarkably slow write operations: any write/update operations require updates of parity information on one dedicated disk;
- Slow reading of operations in a degraded mode due to a high load on the parity unit.

Recovery perspectives

- Controller failure/disassembled array: easy to recover all data. N-1 disks required, data disks are preferred (to build virtual RAID 0); information about disk order and stripe size is required;
- Damaged unit: recovery chances are close to 100%, if no more than 1 disk fails. If 2 or more disks fail, the same problem as with RAID 0 occurs.

RAID, level 5 (RAID5, stripe set with distributed parity)

Presently, RAID5 is the best compromise between fault-tolerance, speed and cost. The technique used in RAID5 is based on usual Stripe Set (like in RAID 0) that in this level mixes data and parity information. Like RAID 4 it requires at least 3 disks but has no designated disk to store parity information not creating such 'queue' for parity updates during write operations.

Depending on the purpose, implementation, the retailer and other factors RAID level 5 may differ in the methods of parity distribution across the stripe set. Among the most common are: **Left-symmetric** (backward dynamic parity distribution), **Right-symmetric** (forward dynamic parity distribution), **Left-asymmetric** (backward parity distribution) and **Right-asymmetric** (forward parity distribution).

0	3	P	6	9
1	P	4	7	P
P	2	5	P	8

Figure 5. Left Symmetric parity distribution (RAID 5)

0	2	P	6	8
1	P	4	7	P
P	3	5	P	9

Figure 6. Left Asymmetric parity distribution (RAID 5)

P	3	4	P	9
0	P	5	6	P
1	2	P	7	8

Figure 7. Right Symmetric parity distribution (RAID 5)

P	2	4	P	8
0	P	5	6	P
1	3	P	7	9

Figure 8. Right Asymmetric parity distribution (RAID 5)

Fault-tolerance is achieved by the same means as in RAID 4: the stripe set stores actual data and parity information; each column of stripes is summed into a parity stripe of the column.

RAID5 combines features of RAID 0 (fast read operations and large capacity) and RAID 4 (extended internal errors correction). If a stripe becomes unreadable, the controller is capable of reconstructing it basing on other stripes and parity information. Actual capacity of RAID5 storage is (**U**-1) * (**min**(unit size) – Reserved).

Advantages of RAID, level 5:

- Ever-faster reading operations;
- Fast writing depending on data and parity distribution method;
- Fault-tolerance;
- RAID may operate in a 'degraded mode', if one disk fails;
- Efficient solution in the aspect of fault tolerance for money.

Disadvantages

- Slower writing operations in comparison with RAID 0;
- Speed of writing operations depends on the content and parity distribution method.

Recovery perspectives

- Controller fault/disassembled array: easy to recover all data. All undamaged disks preferred, but N-1 required; information about disk order, stripe size and **parity distribution method** is required;
- Damaged unit: recovery chances are close to 100%, if no more than 1 disk fails. If 2 or more disks fail, the same problem as with RAID 0 occurs.

RAID, level 6 (RAID6, stripe set with double distributed parity)

Being reliable and at the same time cost-efficient data storage solution, RAID6 was created with the aim to extend RAID5 with one more stripe for data redundancy. For this purpose Reed-Solomon Code algorithm based on Galois field algebra is used. This technique allows adding one more unit for data redundancy and to efficiently correct disk errors.
Data organization on RAID 6 is similar to RAID5: data and parity (P-stripe) are distributed across storage units. The difference is in additional stripe (Q-stripe) located along with P-stripe and containing GF sum of data.
For more information about RAID6 and Q-stripe algorithms, please, refer to https://www.cs.utk.edu/~plank/plank/papers/CS-96-332.html

Advantages of RAID, level 6:

- Ever-faster read operations;
- Fast writing operations depending on data and parity distribution method;
- High fault-tolerance;
- RAID may operate in a 'degraded mode', if one disk or even 2 disks fail;
- Efficient solution in the aspect of of fault tolerance for money.

Disadvantages

- Slower writing operations in comparison with RAID 0;
- Speed of writing operations depends on the content and parity distribution method.

Recovery perspectives

- Controller fault/disassembled array: easy to recover all data. All undamaged disks preferred, but N-1 or N-2 required; information about disk order, stripe size and **parity distribution method** is required;
- Damaged unit: recovery chances are close to 100%, if no more than two disks fail. If more than 2 disks fail, the same problem as with RAID 0 occurs.

Nested RAID: level 0+1, level 10, level 50, level 51 etc.

ested RAID implementations based on RAID 0, level 5 and level 1 appeared to enhance performance capabilities of RAID systems. RAID level 0+1 uses a mirror of stripe sets to increase fault-tolerance without affecting storage performance. RAID level 10 applies an extension of stripes to mirror the data improving performance capabilities and increased use of storage capacity. RAID level 0+1 and level 10 require at least 4 disks. RAID 50 is a stripe set of RAID5 storage created for performance reasons and RAID51 is a mirror of RAID5 created for fault-tolerance (requiring at least 6 disks to be built). Below there are examples for RAID 0+1 and RAID 10.

0	3	6	...	N-2
1	4	7	...	N-1
2	5	8	...	N
0	3	6	...	N-2
1	4	7	...	N-1
2	5	8	...	N

Figure 2. Data organization on mirror of stripes (RAID 0+1; 6 units)

0	2	4	...	N-1
0	2	4	...	N-1
0	2	4	...	N-1
1	3	5	...	N
1	3	5	...	N
1	3	5	...	N

Figure 3. Data organization on stripe of mirrors (RAID 10; 6 units, 2x3 mirrors)

Advantages of nested RAIDs:

- Increased speed or fault-tolerance;
- RAID may operate in a degraded mode;
- RAID 10 and RAID 0+1 are the most available solutions (some on-board controllers support these RAID types).

Disadvantages

- An expensive solution as most of disk space is used for mirrors;
- Hard to manage/maintain.

Recovery perspectives

- Controller failure/disassembled array: easy to recover all data;
- Damaged unit: recovery chances are close to 100%, if it's possible to virtually assemble at least one stripe set (RAID10, RAID50) or at least one mirror instance (RAID 0+1, RAID 51).

instance (RAID 0+1, RAID 51).

Getting started

Taking into account data recovery chances described above and using efficient software specialized on recovery and reconstruction of RAID systems users can easily recover the data from your RAID storage by themselves. The following information is needed to start RAID scan:

- RAID level;
- Scheme of RAID layout across the units;
- RAID parameters – units order, stripe size, parity distribution (if applicable).

After recovery operation you can save image file of your RAID storage to a safe location and load it into data recovery software for further analysis.
We recommend **UFS Explorer RAID Recovery** as efficient software specialized on RAID systems. This application supports virtual reconstruction and data recovery from:

- RAID arrays levels 0, 1, 0+1, 10, 4, 5, 6, 50, 51, RAID 7 (hardware variation of RAID 4);
- Degraded RAID arrays level 1, 0+1, 10, 5 and 6;
- File systems used for NAS storages and servers with RAID arrays.

Related articles:

- RAID recovery

Last update: 05.07.2018