



ΑΛΕΞΑΝΔΡΕΙΟ Τ.Ε.Ι. ΘΕΣΣΑΛΟΝΙΚΗΣ
ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΦΑΡΜΟΓΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ



Πτυχιακή εργασία

«Προσομοίωση της Τεχνολογίας RAID με την χρήση του
Macromedia Flash MX»



της φοιτήτριας
Καρακίζου Βασιλικής
Αρ. Μητρώου: 04/2488

Επιβλέπων καθηγητής
Βαφειάδης Αντώνιος

Θεσσαλονίκη 2009

ΠΡΟΛΟΓΟΣ

Η διαρκής ανάπτυξη της τεχνολογίας καθώς και η ανάγκη για καλύτερη εξυπηρέτηση των υπολογιστικών σκοπών μας συντέλεσαν στην έντονη επιθυμία για βελτίωση των υπολογιστικών συστημάτων.

Σ' ένα σύστημα, κατά κύριο λόγο, βελτιώνεται η απόδοση του επεξεργαστή και της κύριας μνήμης, που φανερώνεται αν παρατηρήσει κανείς τους μεγάλους ρυθμούς βελτίωσης τους. Παρόλα αυτά κάτι που έχει παραμένει σταθερό και με μικρό ρυθμό βελτίωσης της απόδοσής του είναι η δευτερεύουσα αποθήκευση, δηλαδή ο σκληρός δίσκος. Ωστόσο, οι σχεδιαστές δίσκων γνωρίζουν ότι αν ένα στοιχείο δεν μπορεί να βελτιωθεί, τότε μπορεί να καλυτερεύσει η απόδοσή του αποθηκευτικού συστήματος αν προστεθούν κι άλλα ίδια στοιχεία. Άρα, προσθέτοντας πολλούς δίσκους που σχηματίζουν συστοιχίες, οι οποίες λειτουργούν ανεξάρτητα και παράλληλα, πετυχαίνεται αύξηση της απόδοσης της δευτερεύουσας αποθήκευσης και κατ' επέκταση της απόδοσης του συστήματος.

Έτσι προέκυψε το RAID (Redundant Array of Independent Disks), που είναι ένα τυποποιημένο σχήμα σχεδίασης δεδομένων πολλαπλών δίσκων, το οποίο δημιουργήθηκε για να μπορούν να οργανωθούν τα δεδομένα των δίσκων κατά συγκεκριμένο τρόπο. Μ' αυτό τον τρόπο αποφεύγεται η δυσκολία στην ανάπτυξη σχημάτων δεδομένων που θα μπορούσαν να χρησιμοποιηθούν σε πλήθος πλατφορμών και λειτουργικών συστημάτων.

Τα επίπεδα RAID είναι επτά και δεν συνδέονται με ιεραρχική σχέση αλλά καθορίζουν διαφορετικές αρχιτεκτονικές σχεδίασης.

Η μεγάλη συνεισφορά της αρχιτεκτονικής RAID είναι ότι μπορεί και αντιμετωπίζει αποτελεσματικά την ανάγκη για πλεονασμό. Όσο όμως υπάρχουν περισσότερες συσκευές, τόσο αυξάνεται η πιθανότητα βλάβης. Το RAID με αποθηκευμένες πληροφορίες ισοτιμίας μπορεί να επαναφέρει δεδομένα που έχουν χαθεί.

Τελικά, με το RAID πετυχαίνεται βελτίωση απόδοσης καθώς και μεγάλη αξιοπιστία.

Στη συνέχεια θα παρουσιαστούν όλα τα επίπεδα RAID με λεπτομέρειες το καθένα ως προς τη διαμόρφωσή του, τον τρόπο λειτουργίας, τα πλεονεκτήματα και τα μειονεκτήματα του καθενός.

ΠΕΡΙΛΗΨΗ

Η παρούσα πτυχιακή εργασία εκπονήθηκε κατά το διδακτικό έτος 2008-2009 και αφορά την τεχνολογία RAID. Είναι χωρισμένη σε συνολικά 9 κεφάλαια.

Στο πρώτο κεφάλαιο γίνεται μία εισαγωγή σ' όλα τα αποθηκευτικά μέσα. Υπάρχει εκτενής αναφορά στον τρόπο λειτουργίας των κυριότερων συσκευών αποθήκευσης παλιότερων όπως και πιο εξελιγμένων τεχνολογικά.

Επόμενο είναι το δεύτερο κεφάλαιο το οποίο παρουσιάζει πρώτη φορά την τεχνολογία RAID αναφέροντας τα διάφορα επίπεδα του και δίνει μια πρώτη εικόνα στον αναγνώστη για τη λειτουργία του και τους τρόπους σχηματισμού του. Επίσης γίνεται μια μικρή αναφορά στο ιστορικό του RAID.

Στο τρίτο κεφάλαιο αναλύονται όλα τα βασικά επίπεδα RAID επισημαίνοντας τα κύρια πλεονεκτήματα και μειονεκτήματα του κάθε επιπέδου.

Στο επόμενο, τέταρτο κεφάλαιο αναλύονται τα κυριότερα εμφωλιασμένα επίπεδα RAID, όλα όσα δηλαδή αναπτύχθηκαν από την ένωση δύο βασικών επιπέδων RAID. Δίνονται σχηματικά και αναφέρεται η χρήση τους.

Έπειτα γίνεται μια αναφορά σε διάφορα μη πρότυπα επίπεδα RAID. Είναι διάφορες διαμορφώσεις RAID από κάποιες εταιρίες, οι οποίες τις δημιούργησαν με βάση τα πρότυπα επίπεδα για δική τους αποκλειστική χρήση σύμφωνα με τις ανάγκες τους.

Στο έκτο κεφάλαιο περιγράφονται όλοι οι πιθανοί τρόποι υλοποίησης του RAID σύμφωνα με τα σημερινά τεχνολογικά δεδομένα.

Ακολούθως δίνονται επιγραμματικά τα μέτρα μέτρησης της αξιοπιστίας, μέτρα που χρησιμοποιήθηκαν κατά την ανάλυση κάθε επιπέδου RAID.

Στο όγδοο κεφάλαιο αναφέρονται τα διάφορα προβλήματα που αντιμετωπίζουν οι συστοιχίες RAID σε ειδικές περιπτώσεις, καθώς και κάποιες πιθανές λύσεις τους.

Τέλος στο ένατο κεφάλαιο γίνεται μια μικρή περίληψη του πρακτικού κομματιού της εργασίας που εκπονήθηκε χρησιμοποιώντας το εργαλείο Macromedia Flash MX.

ΠΕΡΙΛΗΨΗ ΣΤΗΝ ΑΓΓΛΙΚΗ

This subject has been written in 2008-2009 and introduces the RAID technology. It is divided in 9 total chapters.

The first chapter contains a general reference to all storage devices. There is an extensive reference to operation of the most proper storage devices.

Next there is the second chapter that is presenting for first time the RAID technology giving all its levels and it's giving a first clue to the reader for its operation and the various configurations. Also it is given a small history reference.

In the third chapter there are assayed all the standard RAID levels pointing out the advantages and the disadvantages of each level.

In the next, fourth level, all nested levels are described with figures. Nested levels are all RAID levels that came out of the combination of two basic levels.

Then there are described some non standard RAID levels. There are some RAID configurations build by companies, in order to use them for their own special needs.

In the sixth chapter there are described all the possible ways to setup a RAID configuration according to the nowadays technology.

Afterwards are giving all the reliability measures.

In the eighth chapter are described all the problems that an array would face in special occasions, and also it is given every possible solutions.

At the end, in the ninth chapter there is a small summary of the practical piece.

ΕΥΡΕΤΗΡΙΟ ΠΕΡΙΕΧΟΜΕΝΩΝ

Πρόλογος	σελ. 2
Περίληψη	σελ. 3
Περίληψη στα Αγγλικά	σελ. 4
Ευρετήριο Περιεχομένων	σελ. 5
Ευρετήριο Σχημάτων και Πινάκων	σελ. 7
Εισαγωγή	σελ. 9
Κεφάλαιο 1 : Συσκευές αποθήκευσης	σελ. 12
1.1. Τύποι συσκευών αποθήκευσης δεδομένων	σελ. 12
1.1.1. Κασέτα μαγνητικής εγγραφής	σελ. 12
1.1.2. Σκληρός δίσκος	σελ. 12
1.1.2.1. Χωρητικότητα και ταχύτητα προσπέλασης	σελ. 13
1.1.2.2. Δομή	σελ. 14
1.1.2.3. Τρόπος αποθήκευσης	σελ. 15
1.1.2.4. Είδη σκληρών δίσκων	σελ. 15
1.1.3. CD-ROM/DVD-ROM	σελ. 15
1.1.3.1. Ιστορία	σελ. 16
1.1.3.2. Φυσικές ιδιότητες και χαρακτηριστικά	σελ. 16
1.1.3.3. Είδη δίσκων DVD	σελ. 18
1.1.4. Αυτοματοποιημένες βιβλιοθήκες κασετών	σελ. 20
1.2. Δίαυλοι - Συνδέοντας συσκευές I/O με την CPU-Μνήμη	σελ. 23
1.2.1. Σχεδιασμός Διαύλων	σελ. 25
1.2.2. Πρότυπα Διαύλων	σελ. 28
1.2.3. Παραδείγματα Διαύλων	σελ. 29
1.3. Μέτρα μέτρησης I/O	σελ. 30
Κεφάλαιο 2: Γενικά για το RAID	σελ. 32
2.1. Βασικές Αρχές	σελ. 35
2.2. Οργάνωση Δίσκων	σελ. 38
2.3. Ιστορία	σελ. 41
Κεφάλαιο 3: Standard επίπεδα RAID	σελ. 42
3.1. RAID 0	σελ. 42
3.1.1. Ρυθμός σφαλμάτων RAID 0	σελ. 43
3.1.2. Επίδοση RAID 0	σελ. 44
3.2. RAID 1	σελ. 46
3.2.1. Ρυθμός σφαλμάτων RAID 0	σελ. 48
3.2.2. Επίδοση RAID 0	σελ. 48
3.3. RAID 2	σελ. 50
3.4. RAID 3	σελ. 51
3.4.1. Επίδοση RAID 3	σελ. 52
3.5. RAID 4	σελ. 53
3.5.1. Επίδοση RAID 4	σελ. 54
3.5.2. Υλοποίηση	σελ. 55
3.6. RAID 5	σελ. 55
3.6.1. Χειρισμός ισοτιμίας RAID 5	σελ. 57
3.6.2. Επίδοση RAID 5	σελ. 58
3.6.3. Ρυθμός Σφαλμάτων RAID 5	σελ. 59
3.6.4. Υλοποίηση	σελ. 61

3.7. RAID 6	σελ. 61
3.7.1. Επίδοση(Ταχύτητα)	σελ. 62
3.7.2. Αποδοτικότητα(Πιθανή Απώλεια Χωρητικότητας)	σελ. 63
3.7.3. Υλοποίηση	σελ. 63
Κεφάλαιο 4: Εμφωλευμένα Επίπεδα RAID	σελ. 65
4.1 Εμφωλιασμός	σελ. 65
4.2. RAID 0+1	σελ. 65
4.2.1. RAID 0+1, έξι δίσκων	σελ. 65
4.3. RAID 10(RAID 1+0)	σελ. 67
4.3.1. Πλεονασμός και ικανότητα επανάκτησης απώλειας	σελ. 68
4.3.2. Επίδοση	σελ. 68
4.3.3. Απόδοση(Πιθανή απώλεια χωρητικότητας)	σελ. 68
4.4. RAID 0+3 και 3+0	σελ. 69
4.4.1. RAID 0+3	σελ. 69
4.4.2. RAID 3+0	σελ. 69
4.5. RAID 50(RAID 5+0)	σελ. 70
4.6. RAID 60(RAID 6+0)	σελ. 71
Κεφάλαιο 5 : Μη πρότυπα επίπεδα RAID	σελ. 73
Κεφαλαίο 6 : Υλοποιήσεις	σελ. 76
6.1. Βασισμένο σε λειτουργικό σύστημα("Λογισμικό RAID")	σελ. 76
6.2. Βασισμένο σε Hardware	σελ. 79
6.3. Βασισμένο σε Firmware - driver RAID	σελ. 80
6.4. Network-attached storage	σελ. 81
6.5. Hot spares	σελ. 81
Κεφάλαιο 7 : Όροι Αξιοπιστίας	σελ. 82
7.1. Ρυθμός Αποτυχίας	σελ. 82
7.2. Μέσος όρος απώλειας δεδομένων(Mean time to data loss-MTTDL)	σελ. 82
7.3. Μέσος χρόνος επανάκτησης(Mean time to recovery-MTTR)	σελ. 82
7.4. Ρυθμός μη ανακτήσιμου σφάλματος bit(Unrecoverable bit error rate-UBE)	σελ. 83
7.5. Αξιοπιστία εγγραφής cache(write cache reliability)	σελ. 83
7.6. Ατομική βλάβη εγγραφής(Atomic write failure)	σελ. 83
Κεφάλαιο 8 : Προβλήματα με το RAID	σελ. 84
8.1. Συσχετιζόμενα Σφάλματα	σελ. 84
8.2. Ατομικότητα	σελ. 84
8.3. Αξιοπιστία cache εγγραφής	σελ. 86
8.4. Συμβατότητα συσκευών	σελ. 87
8.5. Επανάκτηση δεδομένων σε περίπτωση εσφαλμένης συστοιχίας	σελ. 88
8.6. Αλγόριθμοι επανάκτησης σφαλμάτων	σελ. 88
Κεφάλαιο 9 : Σχηματική αναπαράσταση της τεχνολογίας RAID με την χρήση του Macromedia Flash MX 2004	σελ. 90
9.1. Γενικά για το Macromedia Flash MX 2004	σελ. 90
9.2. Υλοποίηση της τεχνολογίας RAID με τη χρήση του εργαλείου Macromedia Flash MX 2004	σελ. 91
Βιβλιογραφία	σελ. 93

ΕΥΡΕΤΗΡΙΟ ΣΧΗΜΑΤΩΝ ΚΑΙ ΠΙΝΑΚΩΝ

ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ

Πίνακας	Περιγραφή	Σελίδα
1.1	Είδη DVD	σελ. 18
1.2	Οι κυριότερες επιλογές ενός διαύλου	σελ. 25
1.3	Σύνοψη των διαύλων I/O	σελ. 29
1.4	Σύνοψη διαύλων CPU-Μνήμη που βρίσκονται σε servers	σελ. 30
2.1	Επίπεδα RAID, η διαθεσιμότητά τους, και η ανοχή σε σφάλματα στους επιπλέον δίσκους.	σελ. 33

ΕΥΡΕΤΗΡΙΟ ΣΧΗΜΑΤΩΝ

Σχήμα	Περιγραφή	Σελίδα
1.1	Τυπική συναλλαγή ανάγνωσης διαύλων	σελ. 24
1.2	Ένας δίαυλος χωριστής συναλλαγής	σελ. 26
1.3	Ένας κάτοχος εκτελεί μια εγγραφή	σελ. 27
1.4	Διάγραμμα λοξότητας ρολογιού και ποικιλίας ταχύτητας συσκευών I/O	σελ. 28
1.5	Μοντέλο παραγωγού-server χρόνου απόκρισης και παραγωγικότητας	σελ. 31
3.1	Διάγραμμα εγκατάστασης RAID 0	σελ. 43
3.2	Διάγραμμα εγκατάστασης RAID 1	σελ. 47
3.3	Σύστημα RAID 2 με 39 δίσκους	σελ. 50
3.4	Διάγραμμα εγκατάστασης RAID 3 με blocks των 6 bytes και 2 bytes ισοτιμίας	σελ. 52
3.5	Διάγραμμα εγκατάστασης RAID 4 με αφιερωμένη ισοτιμία	σελ. 54
3.6	Διάγραμμα εγκατάστασης RAID 5 με διαμοιραζόμενη ισοτιμία.	σελ. 57
3.7	Διάγραμμα εγκατάστασης RAID 6	σελ. 62
4.1	Διάγραμμα εγκατάστασης RAID 0+1	σελ. 66
4.2	Διάγραμμα εγκατάστασης RAID 10	σελ. 67
4.3	Διάγραμμα RAID 0+3	σελ. 69
4.4	Εγκατάσταση RAID 50	σελ. 70
4.5	Εγκατάσταση RAID 60	σελ. 72

ΕΙΣΑΓΩΓΗ

Η Είσοδος/Εξοδος (Input/Output - I/O) υπήρξε πάντα το “ορφανό” παιδί της αρχιτεκτονικής υπολογιστών. Ιστορικά, τα I/O είναι παραμελημένα από τους υποστηρικτές της CPU, ωστόσο αυτή η επιφύλαξη απέναντι τους βασίζεται στο πιο διαδεδομένο μέτρο μέτρησης της απόδοσης, τον χρόνο της CPU. Η ποιότητα ενός συστήματος I/O - είτε είναι το καλύτερο ή το χειρότερο στον κόσμο – δεν μπορεί να μετρηθεί με τον χρόνο CPU, ο οποίος εξ ορισμού αγνοεί το I/O. Είναι εμφανές ότι κατατάσσεται ως δευτερεύουσας σημασίας ακόμη και από την ετικέτα *περιφερειακό* που αποδίδεται σε συσκευές I/O.

Αυτή η στάση αντικρούεται από την κοινή λογική. Ένας υπολογιστής χωρίς συσκευές I/O είναι σαν ένα αυτοκίνητο χωρίς ρόδες – δεν μπορεί κανείς να πάει πολύ μακριά χωρίς αυτές. Και όσο ενδιαφέρον κι αν είναι ο χρόνος της CPU, ο χρόνος απόκρισης – ο χρόνος μεταξύ του χρονικού σημείου που ο χρήστης πληκτρολογεί μία εντολή και του χρονικού σημείου που τελικά εμφανίζεται – είναι σίγουρα ένα καλύτερο μέτρο μέτρησης της απόδοσης. Ένας πελάτης που αγοράζει έναν Η/Υ ενδιαφέρεται για τον χρόνο απόκρισης, ακόμη και αν δεν ενδιαφέρεται ο σχεδιαστής CPU.

Η εκδίκηση του I/O είναι εδώ! Υποθέστε ότι έχουμε μια διαφορά ανάμεσα στον χρόνο της CPU και στον χρόνο απόκρισης της τάξεως του 10%, και επιταχύνουμε την CPU με ένα παράγοντα του 10, ενώ αγνοούμε το I/O. Ο Νόμος του Amdahl μας ενημερώνει ότι θα πάρουμε μία πενταπλάσια επιτάχυνση, ενώ έχει σπαταληθεί το μισό του δυναμικού της CPU. Όμοια, κάνοντας την CPU 100 φορές πιο γρήγορη χωρίς να βελτιώσει κάποιος το I/O θα επιταχυνθεί το υπολογιστικό σύστημα μόλις 10 φορές, σπαταλώντας το 90% του δυναμικού. Αν η απόδοση της CPU βελτιώνεται κατά 55% κάθε χρόνο και το I/O δεν βελτιωθεί, κάθε εργασία θα έρθει σε μία κατάσταση κατά την οποία ο ολικός χρόνος ενός υπολογισμού θα καθορίζεται απ’ τον χρόνο αναμονής έως ότου ολοκληρωθούν οι I/O λειτουργίες (I/O bounding). Έτσι πιο πολύς χρόνος θα καταναλώνεται στις αιτήσεις δεδομένων παρά στην επεξεργασία τους. Δεν θα υπάρχει κανένας λόγος να αγοραστούν γρηγορότερες CPUs – και άρα καμία δουλειά για τους σχεδιαστές CPU.

Κάποιοι υπαινίσσονται ότι η προκατάληψη απέναντι στα I/O είναι βάσιμη. Υποστηρίζουν ότι η ταχύτητα I/O δεν έχει άμεσο ενδιαφέρον, αφού υπάρχει πάντα κάποια διεργασία να τρέχει όσο μία άλλη διεργασία περιμένει για ένα περιφερειακό.

Υπάρχουν πολλά σημεία για απάντηση. Πρώτον, αυτό είναι ένα επιχείρημα ότι η επίδοση μετριέται ως παραγωγικότητα(throughput) - πολλές εργασίες ανά ώρα - παρά ως χρόνος απόκρισης. Απλά, αν οι χρήστες δεν ενδιαφέρονταν για τον χρόνο απόκρισης, τα αλληλεπιδραστικά λογισμικά δεν θα είχαν επινοηθεί ποτέ, και δεν θα υπήρχαν ούτε σταθμοί εργασιών ή προσωπικοί υπολογιστές σήμερα. Μπορεί επίσης να αποδειχθεί ακριβό από άποψη κόστους να βασίζεται κάποιος στην εκτέλεση άλλων διεργασιών όσο περιμένει για μια ενέργεια I/O, αφού η κύρια μνήμη πρέπει να είναι μεγάλη αλλιώς η κίνηση σελιδοποίησης από την εναλλαγή διεργασιών θα αυξήσει πραγματικά το I/O. Επιπλέον, σε κάθε προσωπικό υπολογιστή υπάρχει μόνο ένα άτομο για κάθε CPU, και άρα τρέχουν λιγότερες διεργασίες απ' ότι σ' ένα σύστημα στο οποίο μοιράζονται πολλοί χρήστες έναν υπολογιστικό πόρο, εδώ την CPU(time-sharing); πολλές φορές η μόνη διεργασία σε αναμονή είναι ο άνθρωπος! Και κάποιες εφαρμογές, όπως οι επεξεργασίες συναλλαγών, δηλαδή η επεξεργασία πληροφοριών χωρισμένες σε ανεξάρτητες λειτουργίες, βάζουν αυστηρά όρια στον χρόνο απόκρισης ως μέρος της ανάλυσης επίδοσης.

Άρα, η επίδοση I/O μπορεί να περιορίσει την επίδοση του συστήματος και την αποτελεσματικότητα[1].

Σ' αυτό το σημείο παρουσιάζεται μια νέα τεχνολογία η οποία αναπτύσσεται τα τελευταία χρόνια, και αφορά τη χρήση μιας "περιοχής" πολλαπλών δίσκων σε σύνδεση, για να βελτιωθεί η απόδοση και να ενισχυθεί η αξιοπιστία τους. Αυτό το σύστημα ονομάζεται Redundant Arrays of Inexpensive Disks (RAID) και βελτιώνει την απόδοση και την αξιοπιστία αποθηκεύοντας δεδομένα σε πολλούς φυσικούς οδηγούς. Η αξιοπιστία βελτιώνεται μέσω του πλεονασμού της πληροφορίας και η απόδοση βελτιώνεται από το γεγονός, ότι πολλοί οδηγοί είναι διαθέσιμοι για χρήση. Έτσι με την παρούσα πτυχιακή εργασία αναπτύσσεται αυτή ακριβώς η τεχνολογία. Γίνεται μια σχηματική αναπαράστασή της και αναλύεται ως προς την λειτουργία της, την επίδοση της ενώ επισημαίνονται όλα τα πλεονεκτήματα και μειονεκτήματα του κάθε επιπέδου της.

Πτυχιακή εργασία της φοιτήτριας Καρακίζου Βασιλικής

Στόχος της εργασίας είναι μετά την ανάγνωση της και σε συνδυασμό με το πρακτικό μέρος της, ο αναγνώστης να έχει κατανοήσει την λειτουργία κάθε επιπέδου RAID, και να μπορεί να ξεχωρίζει ποιο επίπεδο ταιριάζει σε ποια περίπτωση.

ΚΕΦΑΛΑΙΟ 1: ΣΥΣΚΕΥΕΣ ΑΠΟΘΗΚΕΥΣΗΣ

1.1. Τύποι συσκευών αποθήκευσης δεδομένων

Σ' αυτό το κεφάλαιο θα επικεντρωθούμε στις συσκευές με την υψηλότερη χωρητικότητα όπως οι μαγνητικοί δίσκοι, οι μαγνητικές κασέτες, τα CD-ROMs, και οι αυτοματοποιημένες βιβλιοθήκες κασετών.

1.1.1. Κασέτα μαγνητικής εγγραφής

Στις αρχές του 1950, οι μαγνητικές ταινίες πρωτοχρησιμοποιήθηκαν από την IBM για αποθήκευση δεδομένων. Ο αποθηκευτικός χώρος των μαγνητικών ταινιών ήταν 10.000 φορές μεγαλύτερος των καρτών, γεγονός το οποίο συντέλεσε στην άμεση επιτυχία του μέσου και το καθιέρωσε ως το δημοφιλέστερο αποθηκευτικό μέσο της δεκαετίας του 1980.

Η κασέτα μαγνητικής εγγραφής (compact cassette) ανήκει στην κατηγορία των μαγνητικών ταινιών. Ωστόσο αξίζει ειδικής μνείας μιας και χρησιμοποιήθηκε πάρα πολύ. Η κασέτα κυκλοφόρησε το 1963 από την Philips. Χρειάστηκαν όμως επτά χρόνια για να γίνει το γνωστό και δημοφιλές μέσο αποθήκευσης. Η κυκλοφορία των υπολογιστών ZX Spectrum, Commodore 64 και του Amstrad CPC συντέλεσε κατά πολύ στην διάδοση της. Μία κασέτα 90 λεπτών μπορούσε να αποθηκεύσει περίπου 700kB έως και 1MB δεδομένων στην κάθε της πλευρά. Για την αποθήκευση δεδομένων ενός DVD θα χρειαζόνταν δηλαδή 4500 κασέτες και 281 μέρες[5].

1.1.2. Σκληρός δίσκος

Ο σκληρός δίσκος είναι ένα μαγνητικό αποθηκευτικό μέσο - συσκευή που χρησιμοποιείται στους ηλεκτρονικούς υπολογιστές, στις ψηφιακές βιντεοκάμερες, στα φορητά mp3 players, επιτραπέζια ψηφιακά βίντεο, κονσόλες παιχνιδομηχανών, ψηφιακούς επίγειους και δορυφορικούς τηλεοπτικούς δέκτες κ.τ.λ. και αποτελείται από μια ή περισσότερες μαγνητικές πλακέτες. Η αποθήκευση στους μαγνητικούς δίσκους βασίζεται σε μια τεχνολογία 40 και πλέον ετών, η οποία έχει βελτιωθεί και ακόμη βελτιώνεται ριζικά. Οι σκληροί δίσκοι συνεχίζουν να μικραίνουν σε μέγεθος, να μεγαλώνουν σε χωρητικότητα και να πετυχαίνουν όλο

και πιο σπουδαίες ταχύτητες μεταφοράς. Τα τελευταία δέκα χρόνια, έχει γίνει τρομακτική ανάπτυξη στην τεχνολογία των σκληρών δίσκων. Ένας σκληρός δίσκος αποθηκεύει μεγάλες ποσότητες δεδομένων και η συνήθης χωρητικότητα των σκληρών δίσκων που κυκλοφορούν στο εμπόριο είναι 80 GB έως 1 TB. Για μεγαλύτερες χωρητικότητες που αγγίζουν τα 4 TB (terabyte) χρησιμοποιούνται κυκλώματα πολλαπλών σκληρών δίσκων, με τη μορφή συρταρωτής διάταξης. Η ταχύτητα προσπέλασης των δεδομένων είναι ταχύτερη από το DVD/R/RW αλλά πολύ πιο αργή από την μνήμη του υπολογιστή.

Οι σκληροί δίσκοι χρησιμοποιούνται στους υπολογιστές για την αποθήκευση δεδομένων, κυρίως προγραμμάτων και αρχείων που είναι απαραίτητο να διατηρηθούν, σε αντίθεση με την μνήμη RAM όπου τα δεδομένα διαγράφονται με την διακοπή τροφοδοσίας ηλεκτρικού ρεύματος. Επίσης όλοι οι σκληροί δίσκοι πλέον, έχουν ενσωματωμένη μνήμη(Cache RAM) για προσωρινή αποθήκευση που η χωρητικότητα της ξεκινά από 8MB και φτάνει τα 32MB.

1.1.2.1 Χωρητικότητα και ταχύτητα προσπέλασης

Η χρήση άκαμπτων υλικών και η σφράγιση της μονάδας προσδίδει πολύ μεγαλύτερη αντοχή στους δίσκους σε σχέση με τη δισκέτα. Κατά συνέπεια, οι σκληροί δίσκοι μπορούν να αποθηκεύσουν, να προσπελάσουν και να μεταφέρουν πολύ περισσότερα δεδομένα και σε λιγότερο χρόνο από ότι οι δισκέτες.

Τον Ιανουάριο του 2008, ένας τυπικός σκληρός δίσκος για υπολογιστή desktop, μπορούσε να αποθηκεύσει από 120 έως 1000 GB δεδομένων, να περιστραφεί με ταχύτητες 5.400 έως 10.000 rpm και να μεταφέρει δεδομένα με ρυθμό 1 Gbit/s ή ταχύτερο. (1 GB = 10^9 B; 1 Gbit/s = 10^9 bit/s)

Τον Ιούλιο του 2008, η μέγιστη χωρητικότητα έφτασε το 1,5 TB.

Ο ταχύτερος σκληρός δίσκος επαγγελματικών προδιαγραφών περιστρέφεται με 10.000 έως 15.000 rpm, και μπορεί να επιτύχει μεταφορά δεδομένων με ρυθμό πάνω από 1,6 Gbit/s και σταθερή ταχύτητα μεταφοράς μέχρι 125 Mbytes/sec. Οι δίσκοι που περιστρέφονται με αυτήν την ταχύτητα έχουν μικρότερα platters λόγω της αντίστασης του αέρα και κατά συνέπεια έχουν και μικρότερη χωρητικότητα από την αντίστοιχη των δίσκων για desktop.

Οι σκληροί δίσκοι για φορητούς υπολογιστές, που είναι συγκριτικά μικρότεροι σε μέγεθος από τους desktop ή τους επαγγελματικούς, τείνουν να είναι και πιο αργοί αλλά και με λιγότερη χωρητικότητα. Ένας τυπικός δίσκος για φορητό υπολογιστή περιστρέφεται με 5.400 rpm, ενώ τα μοντέλα που περιστρέφονται με 7.200 είναι ελαφρώς πιο ακριβά. Εξαιτίας του μικρού μεγέθους τους, αυτοί οι δίσκοι έχουν αρκετά μικρότερη χωρητικότητα από τους μεγάλους δίσκους για desktop.

Ο σκληρός δίσκος παίζει πολύ σημαντικό ρόλο στην απόδοση ενός υπολογιστή και πιθανώς μεγαλύτερο από αυτόν, που οι περισσότεροι πιστεύουν. Η ταχύτητα με την οποία ο υπολογιστής αρχίζει και η ταχύτητα με την οποία φορτώνει τα διάφορα προγράμματα για να τα εκτελέσει είναι άμεσα σχετιζόμενες με την ταχύτητα του σκληρού δίσκου. Η απόδοση του σκληρού δίσκου είναι επίσης αποφασιστική όταν χρησιμοποιείται προεργασία (multitasking) ή όταν υπάρχει ανάγκη επεξεργασία μεγάλων ποσοτήτων δεδομένων, για παράδειγμα, γραφικές εργασίες, επεξεργασία video ή ήχου κλπ. Επίσης, σε αρκετά λειτουργικά συστήματα ο δίσκος λειτουργεί σαν ένα μέρος της κύριας μνήμης (RAM) με την τεχνική της εικονικής μνήμης (virtual memory), έτσι ένας αργός δίσκος θα επιβραδύνει πολύ την ταχύτητα του συστήματος. Ένα πρόβλημα στο σκληρό δίσκο είναι ίσως το πιο σημαντικό σφάλμα που μπορεί να συμβεί σε ένα υπολογιστικό σύστημα, και αυτό γιατί συνοδεύεται από απώλεια μεγάλης ποσότητας δεδομένων.

1.1.2.2. Δομή

Ένας σκληρός δίσκος αποτελείται από:

- **μαγνητικά επίπεδα(platter)** κατασκευασμένα από μέταλλο ή πλαστικό και επικαλυμμένα από ένα λεπτό στρώμα οξειδίου του σιδήρου ή άλλο μαγνητικό υλικό.
- τον **άξονα κίνησης** γύρω από τον οποίο περιστρέφονται οι μαγνητικοί δίσκοι με την ίδια ταχύτητα.
- **κεφαλές ανάγνωσης/εγγραφής** επάνω σε βραχίονες πάνω και κάτω από κάθε επιφάνεια δίσκου, που μετακινούνται εμπρός-πίσω. Ο συνδυασμός

της κίνησης των βραχιόνων με την κίνηση των δίσκων, επιτρέπουν στις κεφαλές να έχουν πρόσβαση σε όλα τα σημεία των δίσκων.

- **ηλεκτρονικά εξαρτήματα** που εξυπηρετούν την λειτουργία του σκληρού δίσκου, επικοινωνώντας με τον υπολογιστή και αναλαμβάνοντας την κίνηση των κεφαλών και τη μεταφορά των δεδομένων.

1.1.2.3. Τρόπος αποθήκευσης

Τα δεδομένα αποθηκεύονται στον σκληρό δίσκο ως ακολουθίες bit (αφού οι υπολογιστές λειτουργούν με το δυαδικό σύστημα). Οι κεφαλές γράφουν κάθε bit αλλάζοντας το μαγνητικό πεδίο στην επιφάνεια των μαγνητικών δίσκων και το διαβάζουν απλώς αναγνωρίζοντας το μαγνητικό πεδίο. Κάθε bit δεδομένων καταλαμβάνει τον δικό του χώρο στην επιφάνεια του δίσκου, ωστόσο οι ακολουθίες bit που αποτελούν τα δεδομένα, δεν είναι απαραίτητο να εγγράφονται σειριακά στον δίσκο, αλλά είναι δυνατό να κατακερματιστούν και να εγγραφούν σε διάφορες θέσεις.

1.1.2.4. Είδη σκληρών δίσκων

Με βάση το πρωτόκολλο επικοινωνίας και το interface οι δίσκοι διακρίνονται σε: IDE, SATA, SATA II, SCSI[6].

1.1.3. CD - ROM/DVD - ROM

Το CD (Compact Disk) είναι ένας οπτικός δίσκος που χρησιμοποιείται για αποθήκευση ψηφιακών δεδομένων ή για αποθήκευση ψηφιακού ήχου. Το CD, είναι διαθέσιμο στην αγορά από τα τέλη του 1982 και παραμένει ως τυποποιημένο φυσικό μέσο για τις εμπορικές μουσικές καταγραφές έως σήμερα. Το μουσικό CD αποτελείται από μια ή περισσότερες στερεοφωνικές διαδρομές που αποθηκεύονται χρησιμοποιώντας τη δεκαεξάμπιτη κωδικοποίηση PCM σε ένα ρυθμό δειγματοληψίας 44,1 kHz ανά κανάλι. Με βάση την αρχική τους τυποποίηση τα CDs έχουν μια διάμετρο 120 χιλ. ενώ η δυνατότητα αποθήκευσης τους σε ασυμπιεστο ψηφιακό ήχο είναι 74 λεπτά(ενώ σήμερα έχει επεκταθεί στα 80 λεπτά). Υπάρχουν επίσης δίσκοι 80 χιλ., οι οποίοι αποθηκεύουν περίπου 20 λεπτά

ψηφιακού ασυμπίεστου ήχου. Η τεχνολογία compact disc προσαρμόστηκε αργότερα για χρήση σε ηλεκτρονικούς υπολογιστές ως συσκευή ανάγνωσης αρχείων, γνωστή ως CD-ROM όπου υπήρχε η δυνατότητα μιας μόνο ανάγνωσης και CD-R με δυνατότητα μιας μόνο εγγραφής/πολλών αναγνώσεων και CD-RW για πολλαπλές εγγραφές και αναγνώσεις. Τα CD-ROM/CDR/CDRW ως συσκευές δεν παράγονται πλέον αφού έχουν αντικατασταθεί πλήρως από τα DVD ROM/DVDR+/- /DVDRW. Το CD και οι επεκτάσεις του είναι εξαιρετικά επιτυχείς: το 2004, οι ετήσιες παγκόσμιες πωλήσεις του CD-Audio, του CD-ROM και του CD-R έφθασαν περίπου στους 30 δισεκατομμύρια δίσκους. Η χωρητικότητα ενός CD φθάνει τα 800 Mbyte.

Το DVD (Digital Versatile Disc), σε μετάφραση Ψηφιακός Ευέλικτος Δίσκος ή κατά ελεύθερη μετάφραση Ψηφιακός Πολυμορφικός Δίσκος, είναι ένα οπτικό μέσο αποθήκευσης μεγάλης χωρητικότητας. Συνήθως χρησιμοποιείται για την αποθήκευση εικόνας, ήχου, δεδομένων και κυριότερα βίντεο. Τα DVD μοιάζουν πολύ με τα CD καθώς έχουν το ίδιο σχήμα (με διάμετρο 120 ή σπανιότερα 80 χιλιοστά) αλλά έχουν μορφοποιηθεί με τέτοιο τρόπο ώστε να χωρούν πολύ περισσότερα δεδομένα από ένα CD.

1.1.3.1. Ιστορία

Το DVD είναι μια συγχώνευση του Super Disc (SD) που αναπτύχθηκε από τις εταιρείες Matsushita Electric, Toshiba και του Multimedia CD (MMCD) από τις εταιρείες Sony, Philips. Το 1995 αποφασίστηκε κατόπιν πιέσεων των Microsoft, Intel, Apple και IBM να καταλήξουν σε μία μορφή, αυτή του γνωστού μας DVD. Το 1996 εμφανίζονται τα πρώτα DVD-Video με κύριο χαρακτηριστικό τους την αντιγραφική τεχνολογία Content Scrambling System (CSS), η οποία αποδείχθηκε λίγο αργότερα αρκετά αδύναμη.

1.1.3.2. Φυσικές ιδιότητες και χαρακτηριστικά

Το DVD δεν διαφέρει εμφανισιακά από το CD, έχει και αυτό πάχος 1,2 χιλιοστά και διάμετρο 12 εκατοστά. Έχει υιοθετηθεί από τις κονσόλες παιχνιδιών Playstation2, XBOX με αρκετές ιδιαιτερότητες όσο αφορά την προστασία κατά τις αντιγραφές. Η απόσταση μεταξύ της σπειροειδούς διαδρομής που είναι γραμμένα τα δεδομένα

είναι 0.74 μικρόμετρα στο DVD, 1.6 μικρόμετρα στο CD και 0,32 στο Blu-ray. (1000 μικρόμετρα είναι ένα χιλιοστό του μέτρου). Η απόσταση μεταξύ των bits στο DVD είναι 0,4 μικρόμετρα, 0.834 στο CD και 0,14 στο Blu-ray. Το μήκος κύματος τις κόκκινης ακτίνας λέιζερ που διαβάζει το DVD είναι 640 nm. Το αντίστοιχο στο Blu-ray το μήκος της μπλε ακτίνας λέιζερ είναι 405 nm και στο κοινό CD 780 nm. Το μήκος που έχει το σπειροειδές αυλάκι που γράφονται τα δεδομένα έχει μήκος 17,5 χιλιόμετρα DVD-5 (4,7 Gbyte) και 25 στο DVD-9 (8,5 Gbyte). Αντίστοιχα στο CD είναι από 6 έως 7,5 χιλιόμετρα και στο blu-ray 27 στα 25 Gbyte και 54 χιλιόμετρα στα 50 Gbyte.

Στην πρώτη στρώση του DVD τα δεδομένα γράφονται από το κέντρο προς την περιφέρεια του, ενώ στην δεύτερη στρώση -όταν αυτή υπάρχει- με τον αντίστροφο τρόπο για να αποφεύγεται η καθυστέρηση στην ταχύτητα μεταφοράς δεδομένων. Ο κώδικας ελέγχου αποφυγής σφαλμάτων error correction code (ECC) είναι ανώτερος από αυτών των CD, έτσι τα DVD θεωρούνται αρκετά πιο ανθεκτικά στην κακομεταχείριση.

Όλα τα DVD έχουν τρεις περιοχές την Lead-in την κύρια ενδιάμεση περιοχή εγγραφής των δεδομένων και την Lead-out. Οι περιοχές Lead-in και Lead-out καθορίζουν την αρχή και το τέλος των δεδομένων. Είναι πολύ σημαντικές στις περιπτώσεις που γράφουμε επαναλαμβανόμενα ένα DVD (multisession recording) (όπως στις ψηφιακές βιντεοκάμερες με DVD) ή όταν θέλουμε να προσθαφαιρέσουμε δεδομένα. Επειδή ένα multisession DVD δεν είναι απόλυτα συμβατό κάθε φορά πρέπει να καθορίζονται (finalised) τα Lead-in, Lead-out.

Η παγκόσμια ζήτηση για DVD-Recordable σύμφωνα με την Japan Recording-Media Industries Association αναμένεται να είναι για το 2007 5,7 δισεκατομμύρια μονάδες, για DVD re-writable 500 εκατομμύρια μονάδες, για DVD 8cm 110 εκατομμύρια μονάδες ενώ για τα CD-Recordable θα πέσει στις 6,5 δισεκατομμύρια μονάδες.

1.1.3.3. Είδη δίσκων DVD

Πίνακας 1.1 Είδη DVD

DVD 12εκατοστών	Μονής στρώσης	Διπλής στρώσης
Μονής όψης	DVD-5	DVD-9
Διπλής όψης	DVD-10	DVD-18

DVD 12cm

- Single Layer DVD-5 4,7GB μονής επίστρωσης (4.700.000.000 bytes = 4,38 Gbyte πραγματικά). Είναι τα πλέον διαδεδομένα. Η αποδοχή τους από τους καταναλωτές και η μαζική παραγωγή τους έχει ρίξει τις τιμές στα επίπεδα των CD Recordable.
- DVD-R SL
- DVD-R/RW επανεγγράψιμο
- DVD+R SL
- DVD+R/RW επανεγγράψιμο
- DVD-RAM SL (Random Access Memory) version 2.1. Η δομή του πλησιάζει αυτή του σκληρού δίσκου. Δεν χρειάζεται η χρήση εξειδικευμένων προγραμμάτων. Έχουν πολύ μεγαλύτερη αντοχή από τα κοινά DVD Recordable. Είναι ιδανικά σε βιντεοκάμερες και επιτραπέζια DVD. Δυστυχώς είναι σχετικά ακριβό και δεν αναγνωρίζεται από όλες τις συσκευές DVD. Τα τεχνικά χαρακτηριστικά του δεν επιτρέπουν χαμηλό κόστος κατασκευής. Η χρήση σκληρών δίσκων στις βιντεοκάμερες και επιτραπέζια DVD κάνει το μέλλον τους αβέβαιο.
- Dual Layer
- DVD-9 8,5GB διπλής επίστρωσης (8.500.000.000 bytes = 7.92 Gbyte πραγματικά)
- DVD-R DL
- DVD+R DL
- DVD-RAM DL version 2.1

- DVD-10 9,4GB μονής επίστρωσης δύο όψεων (9.400.000.000 bytes = 8,74 Gbyte πραγματικά)
- DVD-18 17 GB διπλής επίστρωσης δύο όψεων (17.000.000.000 bytes = 15.93 Gbyte πραγματικά)
- DVD 8cm ή Mini-DVD Η χρήση τους γίνεται από ψηφιακές βιντεοκάμερες.
- Μονής όψης 1,4Gbyte (1.400.000.000 bytes = 1,3 Gbyte πραγματικά)
- Διπλής όψης 2,8Gbyte (2.800.000.000 bytes = 2.61 Gbyte πραγματικά)

DVD-Βίντεο

Οι δίσκοι DVD-Video είναι δίσκοι DVD που αποθηκεύουν βίντεο. Για να λειτουργήσουν χρειάζονται έναν αναγνώστη DVD (γνωστό ως DVD-player, για τηλεόραση ή DVD-drive, για υπολογιστή) με αποκωδικοποιητή MPEG-2. Οι ταινίες DVD του εμπορίου χρησιμοποιούν μια σύνθεση της κωδικοποίησης MPEG-2 συμπίεσης βίντεο, ώστε οι ταινίες αυτές να χωράνε στο DVD έτσι ώστε να μην χάνεται η απαραίμιλλη ποιότητα της εικόνας, χάρη στην οποία το DVD έχει γίνει ευρέως διαδεδομένο . Παρά την μεγάλη του χωρητικότητα (8,5 GB) οι περισσότερες ταινίες πρέπει να συμπιεστούν για να χωρέσουν σε ένα DVD. Ο συνολικός ρυθμός μεταφοράς δεδομένων και όχι μόνο του βίντεο, είναι συνήθως μεταβαλλόμενος και κυμαίνεται από 2 έως 10 Mbit/s. Η μεγαλύτερη ποιότητα εικόνας που μπορεί να αναπαράγει ένα DVD αγγίζει τα 10 MBit/s ανά δευτερόλεπτο, το οποίο απλά σημαίνει εξαιρετική ποιότητα εικόνας χωρίς κανένα ψεγάδι. Τις περισσότερες φορές όμως η συμπίεση κάνει τις ταινίες να αναπαράγονται μέχρι και στα 2 MBits ανά δευτερόλεπτο, ωστόσο και πάλι χάρη στην κωδικοποίηση σπανίως η εικόνα ενός DVD φαίνεται να έχει κακή ποιότητα στον θεατή.

Υποστηρίζονται:

- Κώδικες βίντεο MPEG-1 έως 1.8 Mbit/s και MPEG-2 έως 9.8 Mbit/s
- Ανάλυση βίντεο για το ευρωπαϊκό PAL σύστημα τηλεόρασης 720 × 576 ή 352 × 288 γραμμών στα 25 καρέ το δευτερόλεπτο.

- Κώδικες ήχου : PCM έως 6 Mbit/s , DTS έως 1,5 Mbit/s , MPEG-1 Audio Layer II έως 912 kbit/s και ο πολυκάναλος AC-3 έως 448 kbit/s. Ο Dolby AC-3 παρόλο το χαμηλό bitrate του ,είναι ο συνηθέστερος με εξαιρετική ποιότητα ήχου.
- Μέχρι 32 διαφορετικοί υπότιτλοι.

DVD Audio

Το DVD Audio προσφέρει πολυκάναλο ήχο με ρυθμό δειγματοληψίας που φτάνει τα 192 kHz 24bit ανά κανάλι και ρυθμό μεταφοράς δεδομένων που αγγίζει τα 9.6 Mbits. Ο ήχος δεν είναι πάντα συμπιεσμένος επειδή το DVD έχει έως 12 φορές μεγαλύτερη χωρητικότητα από το CD. Τα DVD Audio players εξακολουθούν να είναι αρκετά ακριβά. Μία πιο οικονομική λύση είναι η χρήση του υπολογιστή εφόσον η κάρτα ήχου ανταποκρίνεται στις δυνατότητες του DVD Audio.[7][8]

1.1.4. Αυτοματοποιημένες βιβλιοθήκες κασετών

Οι Αυτοματοποιημένες Βιβλιοθήκες Κασετών(Automated Tape Libraries - ATLs) είναι ολοκληρωμένα συστήματα υλικού και λογισμικού που έχουν σχεδιαστεί για να επιτρέπουν την αυτοματοποιημένη πρόσβαση σε terabytes-to-petabytes δεδομένων που βρίσκονται σε φυσίγγια μαγνητικών ταινιών, χωρίς να απαιτείται ανθρώπινη εργασία ("βιβλιοθηκάριος κασετών") για να φορτώσει φυσικά μια κασέτα σε ένα δίσκο όταν κάποιος χρήστης το απαιτήσει. Τα ATLs χρησιμοποιούνται ευρέως ως υλικό που επιτρέπει το backup και την ανάκτηση, καθώς και συστήματα αποθήκευσης βασισμένα σε κασέτες όπως το Hierarchical Storage Management (HSM). Σημειώστε ότι οι αυτοματοποιημένες βιβλιοθήκες υπάρχουν και μέσα μη κασετών (π.χ. οπτικά μέσα βιβλιοθήκες), μοιράζοντας πολλά από τα ίδια χαρακτηριστικά των ATLs όπως περιγράφονται παρακάτω.

Τα περισσότερα ATLs μοιράζονται ένα κοινό σύνολο χαρακτηριστικών γνωρισμάτων από την παρακάτω λίστα, με διαφορετικά σχήματα και μεγέθη ανάλογα με τον πωλητή / μοντέλο / ρύθμιση:

- Οι φυσικές μονάδες μαγνητοταινίας χρησιμοποιούνται για την ανάγνωση και την εγγραφή δεδομένων σε κασέτες. Οι βιβλιοθήκες κασετών ποικίλουν σε

μέγεθος σε σχέση με πόσες μονάδες μαγνητοταινίας υποστηρίζονται - από ένα μόνο μηχανισμό δίσκου, σε δεκάδες/εκατοντάδες οδηγεί σε μεγαλύτερες βιβλιοθήκες.

Οι περισσότερες τεχνολογίες κασετών έχουν προσαρμοστεί για χρήση μέσα σε ένα ή περισσότερα ATL, συμπεριλαμβανομένου του Digital Linear Tape, LTO Ultrium, Sony και άλλες 4mm / 8mm μορφές.

- Οι διεπαφές καναλιών συνδέουν τις μονάδες μαγνητοταινίας για να φιλοξενήσουν σταθμούς (π.χ. ένας παράλληλος δίαυλος SCSI) ή σε κοινό χώρο αποθήκευσης δικτύου (όπως Fibre Channel SAN ύφασμα ή ένα ύφασμα FICON).

Σε ορισμένες εφαρμογές, οι διεπαφές καναλιού είναι εξαιρετικά εικονικές από τη βιβλιοθήκη ή από μια εξωτερική εικονική μηχανή, επιτρέποντας εκτεταμένη προσαρμογή του τρόπου με τον οποίο οι οδηγοί κασετών φαίνονται από τους διακομιστές μέσα από διάφορες διεπαφές, και προσφέρουν ένα βαθμό καναλιού "γεφυρώνοντας" εντός της βιβλιοθήκης. Σε άλλες εφαρμογές, η μονάδα διασύνδεσης καναλιού δεν είναι τίποτα περισσότερο από ένα εξελιγμένο κανάλι ινών(Fibre Channel), FICON ή δίαυλο SCSI που κρέμεται από το πίσω μέρος ενός φυσικού δίσκου.

- Οι χρονοθυρίδες ταινιών("slots") που κρατάνε τις ταινίες όταν βρίσκονται εντός της βιβλιοθήκης και όχι σε μια ταινία. Τα ATLs έρχονται σε ένα ευρύ φάσμα ικανοτήτων, από δεκάδες έως δεκάδες χιλιάδες των slots.
- Φυσίγγια κασετών("ταινίες"). Αυτά είναι φυσίγγια φυσικών ταινιών που αποθηκεύουν τα δεδομένα. Σε ATLs, τα μέσα ενημέρωσης συχνά παρέχουν προ-εγκατεστημένες ή χειροκίνητες ετικέτες barcode για να βοηθήσουν τη βιβλιοθήκη να καθορίζει τη μοναδική ταυτότητα κάθε μέσου.
- Ένας ρομποτικός μηχανισμός, ή "συλλέκτης", που κινεί τις ταινίες από τις θυρίδες στους δίσκους, από θυρίδα σε θυρίδα, και από θυρίδα σε CAP. Σε γενικές γραμμές, αυτή είναι η "αυτοματοποίηση" του ATL.
- Τα φυσίγγια έχουν πρόσβαση σε πόρτες, ή CAPs, τα οποία ενεργούν ως μηχανισμοί εισαγωγής/εξαγωγής για να προσθέσουν ή να αφαιρέσουν τις ταινίες από τη βιβλιοθήκη.

- Μηχανισμοί pass-through. Στις βιβλιοθήκες που κλιμακώνονται ή εξαπλώνονται προσθέτοντας πολλαπλά μέσα αποθήκευσης βιβλιοθήκης μαζί, χρειάζεται συχνά ένας μηχανισμός που να επιτρέπει τη διέλευση των ταινιών μεταξύ των ενοτήτων; αυτές είναι συνήθως αναφέρονται ως "pass-through ports".
- Σύστημα ελέγχου βιβλιοθήκης. Αυτή είναι η "κόλλα" του υλικού και του λογισμικού που βοηθά τις θυρίδες, τους οδηγούς, τα pickers, τα CAPs και οτιδήποτε άλλο στη βιβλιοθήκη να λειτουργεί ως ένα σύνολο. Το σύστημα ελέγχου της βιβλιοθήκης παρέχει γενικά κάποια μορφή διασύνδεσης με το "έξω κόσμος" για να καταστεί δυνατός ο έλεγχος της βιβλιοθήκης από λογισμικό τρίτων, είτε μέσω του προτύπου εντολών autoloader SCSI, ή μέσω ενός συγκεκριμένου API. Εκτός από τον έλεγχο της βιβλιοθήκης, μια σημαντική λειτουργία του συστήματος ελέγχου της βιβλιοθήκης είναι ο καθορισμός των περιεχομένων των μέσων της ATL, με βάση το ID του barcode / μέσου.

Στις περισσότερες περιπτώσεις, το σύνολο των λειτουργιών / εργαλείων που παρέχει μια ολοκληρωμένη βιβλιοθήκη "ελέγχου" μοιράζεται μεταξύ υλικού που παρέχεται από τρίτους και προϊόντα λογισμικού που χρησιμοποιούν τη βιβλιοθήκη, για ένα σκοπό (π.χ. μια εφεδρικό/ανάκτησης εργαλείο όπως Veritas NetBackup ή EMC/ Legato NetWorker). Το σημείο αυτό είναι σημαντικό γιατί μία βιβλιοθήκη αυτοματοποιημένων ταινιών, από μόνη της, χωρίς καμία πρόσθετη υποστήριξη του συστήματος ή του λογισμικού, γενικά στερείται αντικειμένου.

- Βιβλιοθήκη πίνακα ελέγχου. Μερικές βιβλιοθήκες διαθέτουν οθόνες που ενημερώνουν τι συμβαίνει στο εσωτερικό της βιβλιοθήκης.
- Ορισμένες βιβλιοθήκες προσφέρουν τη δυνατότητα, στο "υλικό", να κατατμηθούν σε πολλαπλές ανεξάρτητες φυσικές βιβλιοθήκες, που περιέχουν μια ξεχωριστή σειρά από μονάδες μαγνητοταινίας και φυσίγγια χρονοθυρίδων.

Χαρακτηριστικοί πωλητές ATL είναι οι IBM, StorageTek (νυν Sun), η HP, SpectraLogic, Quantum, και ADIC.[9]

1.2. Δίαυλοι – Συνδέοντας συσκευές I/O με την CPU/Μνήμη

Σ' ένα υπολογιστικό σύστημα, τα ποικίλα υποσυστήματα πρέπει να έχουν διεπαφές το ένα με το άλλο; για παράδειγμα, η μνήμη και η CPU πρέπει να επικοινωνούν, και το ίδιο και οι συσκευές CPU και I/O. Αυτό γίνεται με τον *δίαυλο*. Ο δίαυλος εξυπηρετεί ως ένας διαμοιραζόμενος σύνδεσμος επικοινωνίας ανάμεσα στα υποσυστήματα. Τα δύο κύρια πλεονεκτήματα των διαύλων είναι το χαμηλό κόστος και η προσαρμοστικότητα. Οι νέες συσκευές μπορούν εύκολα να προστεθούν και τα περιφερειακά μπορούν να μετακινηθούν μεταξύ συστημάτων υπολογιστών που χρησιμοποιούν ένα δίαυλο με τα ίδια τεχνικά χαρακτηριστικά. Το κόστος είναι χαμηλό, αφού ένα μόνο σετ καλωδίων διαμοιράζεται με πολλαπλούς τρόπους.

Το κύριο μειονέκτημα ενός διαύλου είναι ότι δημιουργεί το φαινόμενο "λαιμό μπουκαλιού", πιθανόν μειώνοντας τη μέγιστη παραγωγικότητα του συστήματος. Όταν μια ενέργεια I/O πρέπει να περάσει από ένα κεντρικό δίαυλο, ο περιορισμός του εύρους ζώνης είναι τόσο πραγματικός όσο - και κάποιες φορές πιο σοβαρός από - το εύρος ζώνης της μνήμης. Σε εμπορικά συστήματα, που γίνεται εκτεταμένη χρήση I/O, και σε υπερυπολογιστές, όπου είναι απαραίτητο οι τιμές I/O να είναι υψηλές επειδή η επίδοση της CPU είναι υψηλή, ο σχεδιασμός ενός συστήματος διαύλων είναι μεγάλη πρόκληση καθώς πρέπει να ικανοποιεί τις απαιτήσεις του επεξεργαστή.

Ένας λόγος που ο σχεδιασμός των διαύλων είναι τόσο δύσκολος είναι ότι η μέγιστη ταχύτητα διαύλων μειώνεται κατά πολύ από φυσικούς παράγοντες:

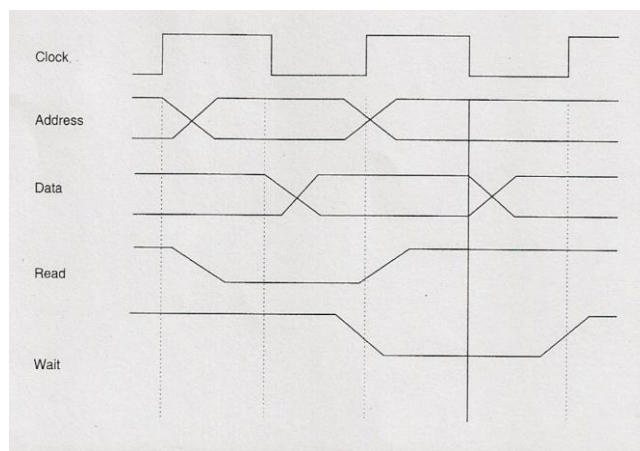
- a) το μήκος του διαύλου και
- b) ο αριθμός των συσκευών(και, άρα η φόρτωση διαύλων).

Αυτά τα φυσικά όρια εμποδίζουν την αυθαίρετη επιτάχυνση των διαύλων. Η επιθυμία για υψηλές τιμές παραγωγικότητας I/O μπορεί επίσης να οδηγήσει σε συγκρουόμενες απαιτήσεις σχεδιασμού.

Οι δίαυλοι παραδοσιακά κατατάσσονται ως *δίαυλοι CPU-Μνήμης* και *δίαυλοι I/O*. Οι δίαυλοι I/O μπορούν να είναι παρατεταμένοι σε μήκος, μπορούν να έχουν πολλούς τύπους συσκευών συνδεδεμένων σ' αυτούς, να έχουν μεγάλη έκταση σε εύρος ζώνης δεδομένων των συσκευών συνδεδεμένων σ' αυτούς, και φυσιολογικά

ακολουθούν ένα πρότυπο διαύλου. Οι δίαυλοι CPU-Μνήμης, απ' την άλλη πλευρά, είναι μικροί σε μήκος, λειτουργούν γενικά με μεγάλη ταχύτητα, και ταιριάζουν στο σύστημα μνήμης για να μεγιστοποιούν το εύρος ζώνης CPU-Μνήμης. Κατά τη φάση του σχεδιασμού, ο σχεδιαστής ενός διαύλου CPU-Μνήμης ξέρει όλους τους τύπους συσκευών που πρέπει να ενωθούν μεταξύ τους, ενώ ο σχεδιαστής I/O πρέπει να αποδεχτεί τις συσκευές, οι οποίες ποικίλουν σε λανθάνουσα και ικανότητες εύρους ζώνης. Λανθάνουσα ονομάζεται η καθυστέρηση μεταξύ μιας εκτελούμενης εντολής συναλλαγής και του υλικού που πραγματικά εκτελεί την εναλλαγή της τάσης από υψηλή σε χαμηλή και το αντίστροφο. Για να μειωθεί το κόστος, κάποιοι υπολογιστές έχουν ένα ενιαίο δίαυλο και για τη μνήμη και για τις συσκευές I/O.

Μία τυπική συναλλαγή διαύλου περιλαμβάνει δύο μέρη: την αποστολή της διεύθυνσης και τη λήψη ή αποστολή δεδομένων. Οι συναλλαγές διαύλων ορίζονται συνήθως από το τι κάνουν στην μνήμη: μία συναλλαγή ανάγνωσης μεταφέρει δεδομένα από την μνήμη(είτε στην CPU είτε στην συσκευή I/O), και μία συναλλαγή εγγραφής γράφει δεδομένα στη μνήμη. Σε μία συναλλαγή ανάγνωσης, η διεύθυνση στέλνεται πρώτα μέσω του διαύλου στη μνήμη, μαζί με τα κατάλληλα σήματα ελέγχου υποδεικνύοντας μία ανάγνωση. Στο σχήμα 1.1, π.χ. μια μονάδα ζητά από τη μνήμη να εκτελέσει μια διαδικασία ανάγνωση στέλνοντας στον δίαυλο το σήμα read και τη διεύθυνση. Η μνήμη απαντάει επιστρέφοντας τα δεδομένα στο δίαυλο με τα ακατάλληλα σήματα ελέγχου, σ' αυτή την περίπτωση υποδεικνύοντας το σήμα αναμονής.



Σχήμα 1.1 Τυπική συναλλαγή ανάγνωσης σε σύγχρονο δίαυλο.

Μία συναλλαγή εγγραφής απαιτεί από την CPU ή τη συσκευή I/O να στείλουν και τη διεύθυνση και τα δεδομένα και δεν απαιτεί καμιάν επιστροφή δεδομένων. Συνήθως η πρέπει να περιμένει τη στιγμή ανάμεσα στην αποστολή της διεύθυνσης και στην λήψη των δεδομένων στην ανάγνωση, αλλά η CPU συχνά δεν περιμένει τις εγγραφές.

1.2.1. Σχεδιασμός Διαύλων

Ο σχεδιασμός ενός διαύλου παρουσιάζει διάφορες επιλογές, όπως δείχνει ο πίνακας 1.2. Όπως τα υπόλοιπα μέρη ενός υπολογιστικού συστήματος, οι αποφάσεις θα εξαρτώνται από το κόστος και τους στόχους απόδοσης. Οι τρεις πρώτες επιλογές στον πίνακα είναι ξεκάθαρες επιλογές - ξεχωριστές διευθύνσεις και γραμμές δεδομένων, μεγαλύτερες γραμμές δεδομένων, και οι μεταφορές πολλαπλών λέξεων όλες δίνουν μεγαλύτερη επίδοση σε μεγαλύτερο κόστος.

Πίνακας 1.2 Οι κυριότερες επιλογές ενός διαύλου. Το πλεονέκτημα των ξεχωριστών διευθύνσεων και διαύλων δεδομένων είναι κυρίως στις εγγραφές.

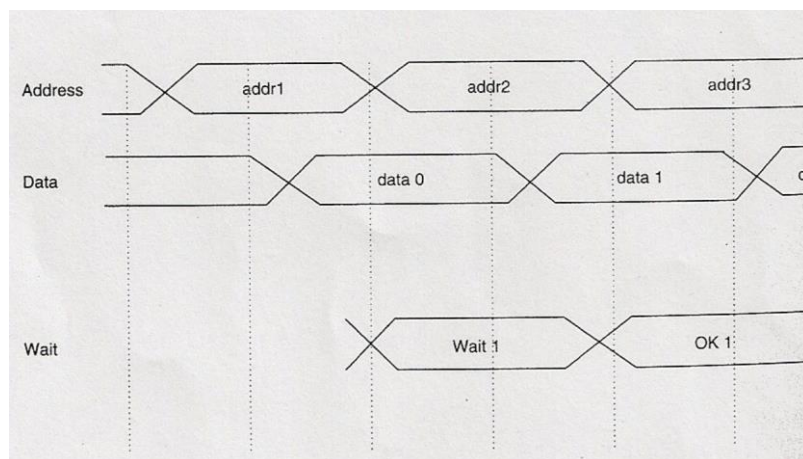
Επιλογή	Υψηλή απόδοση	Χαμηλό κόστος
Πλάτος διαύλου	Ξεχωριστές διευθύνσεις και γραμμές δεδομένων	Πολυπλεξία διευθύνσεων και δεδομένων
Πλάτος δεδομένων	Πιο φαρδύς πιο γρήγορος	Πιο στενός πιο φθηνός
Μέγεθος μεταφοράς	Πολλές λέξεις έχουν λιγότερη υπερφόρτωση διαύλου	Η μεταφορά μίας λέξης είναι πιο απλή
Κάτοχοι διαύλου	Πολλαπλοί(απαιτεί διαιτησία)	Ένας κάτοχος(όχι διαιτησία)
Χωριστές συναλλαγές(;))	Ναι-ξεχωριστά πακέτα αίτησης και απάντησης χρησιμοποιούν μεγαλύτερο εύρος ζώνης(χρειάζεται περισσότερους κυρίους	Όχι-η συνεχής σύνδεση είναι φθηνότερη και έχει λιγότερη λανθάνουσα
Χρονισμός	Σύγχρονος	Ασύγχρονος

Το επόμενο αντικείμενο στον πίνακα 1.2 αφορά τον αριθμό των κατόχων διαύλων. Αυτές είναι οι συσκευές που μπορούν να εισάγουν μια συναλλαγή ανάγνωσης ή εγγραφής; η CPU, για παράδειγμα, είναι πάντα ένας βασικός κάτοχος διαύλου. Ένας δίαυλος έχει πολλαπλούς κατόχους όταν υπάρχουν πολλαπλές CPUs ή

όταν συσκευές I/O μπορούν να εισάγουν μία συναλλαγή διαύλου. Αν υπάρχουν πολλαπλοί κάτοχοι, ένα σύστημα διαιτησίας απαιτείται ανάμεσα στους κατόχους για να αποφασιστεί ποιος θα πάρει τον δίαυλο επόμενος. Η διαιτησία είναι μία απαραίτητη απαίτηση, στην περίπτωση που έχουμε συσκευές συνδεδεμένες η μία μετά την άλλη (daisy-chained), π.χ. SCSI.

Με πολλαπλούς κατόχους, ένας δίαυλος μπορεί να προσφέρει υψηλότερο εύρος ζώνης μεταφοράς πακέτων, σε αντίθεση με την κατάσταση κατά την οποία ο δίαυλος χρησιμοποιείται για μια μοναδική συναλλαγή. Αυτή η τεχνική λέγεται διαίρεση συναλλαγών (split transactions).

Το σχήμα 1.2 δείχνει ένα δίαυλο διαίρεσης-συναλλαγής. Η συναλλαγή ανάγνωσης χωρίζεται σε μια συναλλαγή αίτησης-ανάγνωσης που περιέχει την διεύθυνση και σε μία συναλλαγή απάντησης-μνήμης που περιέχει τα δεδομένα. Κάθε συναλλαγή πρέπει τώρα να επισημανθεί έτσι ώστε η CPU και η μνήμη να μπορούν να πουν τι είναι τι. Οι χωριστές συναλλαγές καθιστούν τον δίαυλο διαθέσιμο για τους άλλους κατόχους όσο η μνήμη διαβάζει τις λέξεις από την αιτούμενη διεύθυνση. Επίσης η CPU πρέπει να παίζει τον ρόλο του διαιτητή στον δίαυλο για να στείλει τα δεδομένα και η μνήμη πρέπει να παίζει επίσης το ρόλο του διαιτητή στον δίαυλο ώστε να επιστρέψει τα δεδομένα. Άρα, μία διαιρεμένη συναλλαγή διαύλου έχει υψηλότερο εύρος ζώνης, αλλά συνήθως έχει υψηλότερο χρόνο απόκρισης απ' ότι ένας δίαυλος που χρησιμοποιείται κατά την διάρκεια μιας μοναδικής συναλλαγής.



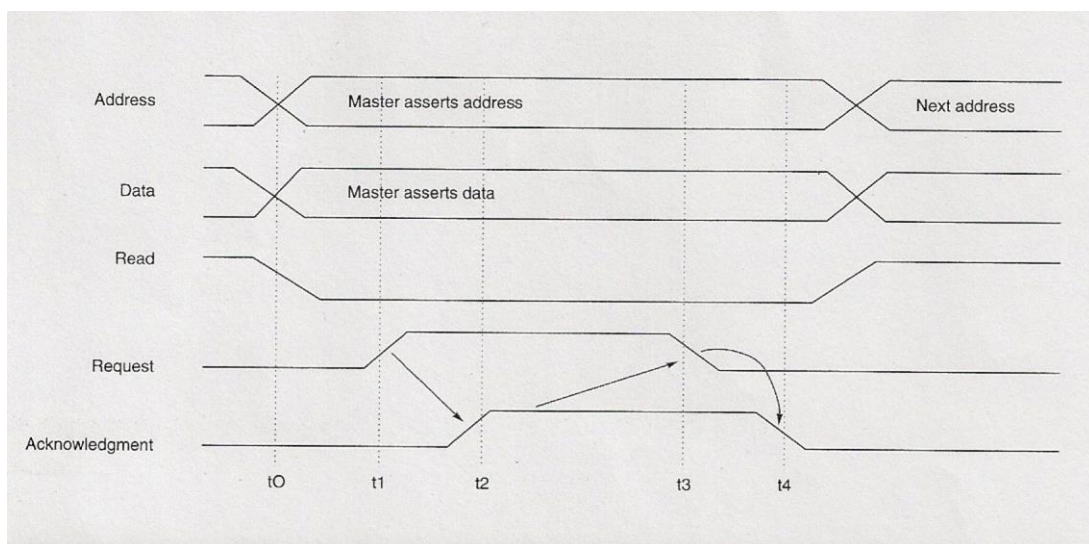
Σχήμα 1.2 Ένας δίαυλος χωριστής-συναλλαγής. Εδώ η διεύθυνση στον δίαυλο αντιστοιχεί επόμενη πρόσβαση στη μνήμη.

Το τελευταίο αντικείμενο στο πίνακα 1.2, είναι ο χρονισμός, ο οποίος αναφέρεται αν ο δίαυλος είναι σύγχρονος ή ασύγχρονος. Αν ένας δίαυλος είναι σύγχρονος, περιλαμβάνει ένα ρολόι στις γραμμές ελέγχου και ένα σταθερό πρωτόκολλο για διευθύνσεις και δεδομένα σχετικά με το ρολόι.

Αυτοί οι δίαυλοι μπορούν να είναι γρήγοροι και φθηνοί. Έχουν δύο κύρια μειονεκτήματα, παρόλα αυτά. Τα πάντα πάνω στον δίαυλο πρέπει να τρέχουν με τον ίδιο ρυθμό ρολογιού, και εξαιτίας προβλημάτων λοξότητας ρολογιού, οι σύγχρονοι δίαυλοι δεν μπορούν να είναι μακριοί. Η λοξότητα ρολογιού(clock skew) είναι ένα φαινόμενο που συμβαίνει στα σύγχρονα κυκλώματα, κατά το οποίο το σήμα του ρολογιού, που στέλνεται απ' το ρολόι, καταφθάνει σε διάφορα μέρη του κυκλώματος σε διαφορετικές χρονικές στιγμές.

Τέλος οι δίαυλοι CPU-Μνήμης είναι πάντοτε σύγχρονοι.

Ένας ασύγχρονος δίαυλος, απ' την άλλη πλευρά, δεν είναι χρονισμένος. Αντ' αυτού, χρονίζεται μόνος του, χρησιμοποιεί πρωτόκολλα χειραφιάς(handshaking) μεταξύ του αποστολέα διαύλου και του παραλήπτη. Το σχήμα 1.3 δείχνει τα βήματα ενός κατόχου που εκτελεί μία εγγραφή σε ασύγχρονο δίαυλο.

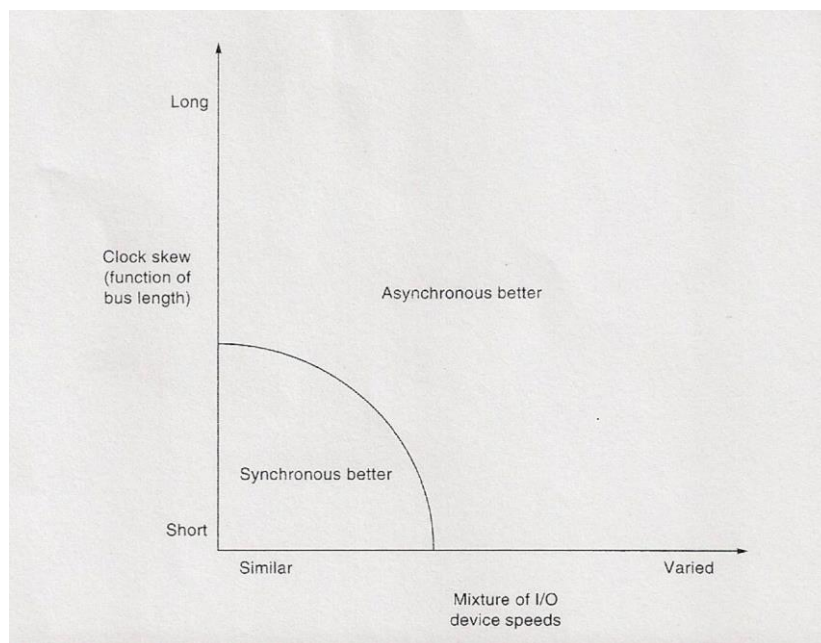


Σχήμα 1.3 Ένας κάτοχος εκτελεί μία εγγραφή. Η κατάσταση της συναλλαγής σε κάθε χρονική στιγμή είναι η ακόλουθη. Ο κάτοχος πήρε τον έλεγχο και διεκδικεί διεύθυνση, κατεύθυνση, και δεδομένα, μετά περιμένει ένα συγκεκριμένο ποσοστό χρόνου για τους slaves να αποκωδικοποιήσουν τον στόχο;

t1: Ο κάτοχος διεκδικεί την γραμμή αίτησης; t2: ο slave διεκδικεί την αναγνώριση, δείχνοντας ότι παραλείφθηκαν δεδομένα; t3: ο κάτοχος απελευθερώνει μία αίτησης; t4 ο slave απελευθερώνει μία αναγνώριση

Ο ασυγχρονισμός κάνει πιο εύκολη την εγκατάσταση πολλών συσκευών και επίσης μπορεί να επιμηκυνθεί ο δίαυλος χωρίς την ανησυχία της λοξότητας ρολογιού ή άλλων προβλημάτων συγχρονισμού. Χρησιμοποιώντας έναν σύγχρονο δίαυλο, ο οποίος συνήθως είναι γρηγορότερος από έναν ασύγχρονο καθώς αποφεύγεται η υπερφόρτωση του συγχρονισμού του διαύλου για κάθε συναλλαγή. Η επιλογή του σύγχρονου από τον ασύγχρονο δίαυλο έχει επιπτώσεις όχι μόνο στο εύρος ζώνης των δεδομένων αλλά επίσης και στην χωρητικότητα του συστήματος I/O όπως στην φυσική απόσταση και στον αριθμό των συσκευών που μπορούν να συνδεθούν στον δίαυλο; οι ασύγχρονοι δίαυλοι μπορούν να ανταπεξέλθουν καλύτερα στις τεχνολογικές αλλαγές. Οι δίαυλοι I/O είναι περισσότερο πιο πρόσφοροι για ασυγχρονισμό παρά οι δίαυλοι CPU – μνήμης. Το σχήμα 1.4 προτείνει πότε να χρησιμοποιηθεί πότε η μια τεχνική και πότε η άλλη.

Οι σύγχρονοι δίαυλοι είναι καλύτεροι όταν η απόσταση είναι μικρή και οι συσκευές I/O στον δίαυλο μεταφέρουν δεδομένα με τις ίδιες ταχύτητες.



Σχήμα 1.4 Διάγραμμα λοξότητας ρολογιού και ποικιλίας ταχύτητας συσκευών I/O. Οι σύγχρονοι δίαυλοι είναι καλύτεροι όταν η απόσταση είναι μικρή και οι συσκευές I/O στον δίαυλο μεταφέρουν δεδομένα με τις ίδιες ταχύτητες.

1.2.2. Πρότυπα Διαύλων

Ο αριθμός και η ποικιλία των συσκευών I/O δεν είναι σταθεροί στα περισσότερα συστήματα υπολογιστών, επιτρέποντας έτσι στους πελάτες να προσαρμόσουν

τους υπολογιστές στις ανάγκες τους. Όπως η διεπαφή στην οποία είναι συνδεδεμένες οι συσκευές, ο δίαυλος I/O μπορεί επίσης να θεωρηθεί σαν μία επέκταση διαύλου για να προστίθενται συσκευές I/O με τον καιρό. Τα πρότυπα που επιτρέπουν στον κατασκευαστή υπολογιστών και στον σχεδιαστή συσκευών I/O να δουλεύουν ανεξάρτητα, συνεπώς, παίζουν μεγάλο ρόλο στον καθορισμό της επιλογής των διαύλων. Όσο ο σχεδιαστής του υπολογιστή - συστήματος και ο σχεδιαστής συσκευής I/O συμφωνούν στις απαιτήσεις, κάθε συσκευή I/O μπορεί να συνδεθεί σε κάθε υπολογιστή. Μάλιστα, κάθε πρότυπο διαύλου I/O είναι η τεκμηρίωση που ορίζει πως να συνδεθούν μεταξύ τους.

Οι μηχανές πολλές φορές γίνονται τόσο δημοφιλής που οι δίαυλοι I/O τους γίνονται de facto πρότυπα; παραδείγματα είναι το PDP-11 Unibus και το IBM PC-AT Bus. Αν έχουν κατασκευαστεί πολλές συσκευές I/O για ένα δημοφιλές μηχάνημα, πολλοί κατασκευαστές υπολογιστών θα δημιουργήσουν την διεπαφή I/O τους έτσι ώστε αυτές οι συσκευές να μπορούν να κουμπώσουν πάνω στα μηχανήματά τους. Η Ευφυής Περιφερειακή Διεπαφή(Intelligent Peripheral Interface-IPI) και το Ethernet είναι παραδείγματα των προτύπων που προκύπτουν από την συνεργασία των κατασκευαστών. Αν τα πρότυπα είναι επιτυχή, τελικά επικυρώνονται από οργανισμούς όπως ο ANSI ή ο IEEE. Περιστασιακά, ένα πρότυπο διαύλου έρχεται στην κορυφή κατευθείαν από την επιτροπή προτύπων - ένα παράδειγμα είναι το PCI.

1.2.3. Παραδείγματα διαύλων

Ο πίνακας 1.3 συνοψίζει τα χαρακτηριστικά πέντε διαύλων I/O, και ο πίνακας 1.4 συνοψίζει τρεις διαύλους CPU-μνήμης που βρίσκονται σε servers.

Πίνακας 1.3 Σύνοψη των διαύλων I/O. Οι δύο πρώτοι ξεκίνησαν ως δίαυλοι CPU-μνήμης και εξελίχθηκαν σε δίαυλους I/O.

	S bus	MicroChannel	PCI	IPI	SCSI 2
Πλάτος δεδομένων	32 bits	32 bits	32 έως 64 bits	16 bits	8 έως 16 bits
Ρυθμός ρολογιού	16 έως 25 MHz	Ασύγχρονο	33 MHz	Ασύγχρονο	10 MHz ή ασύγχρονο

	S bus	MicroChannel	PCI	IPI	SCSI 2
Αριθμός master διαύλων	Πολλοί	Πολλοί	Πολλοί	Ένας	Πολλοί
Εύρος ζώνης, αναγνώσεις 32-bits	33 MB/sec	20 MB/sec	33 MB/sec	25 MB/sec	20 MB/sec ή 6 MB/sec
Εύρος ζώνης, κορυφή	89 MB/sec	75 MB/sec	111 MB/sec	25 MB/sec	20 MB/sec ή 6 MB/sec
Πρότυπο	Κανένα	-	-	ANSI X3.129	ANSI X3.131

Πίνακας 1.4 Σύνοψη διαύλων CPU-μνήμης που βρίσκονται σε servers. Όλοι χρησιμοποιούν χωριστές συναλλαγές για να ενισχύσουν το εύρος ζώνης. Ο αριθμός των θυρίδων στον δίαυλο για τους masters ή τους slaves είναι 16, 9, και 10, αντίστοιχα.

	HP Summit	SGI Challenge	Sun XDBus
Πλάτος δεδομένων	128 bits	256 bits	144 bits
Ρυθμός ρολογιού	60MHz	48 MHz	66 MHz
Αριθμός master διαύλων	Πολλοί	Πολλοί	Πολλοί
Εύρος ζώνης, κορυφή	960 MB/sec	1200 MB/sec	1056 MB/sec
Πρότυπο	Κανένα	Κανένα	Κανένα

1.3. Δείκτες μέτρησης I/O

Η επίδοση I/O έχει κάποιους δείκτες μέτρησης που δεν υπάρχουν αντίστοιχα στο σχεδιασμό CPU.

Ένα απ' αυτά είναι η *πολυμορφία*:

- Ποιες συσκευές I/O μπορούν να συνδεθούν σ' ένα υπολογιστικό σύστημα;

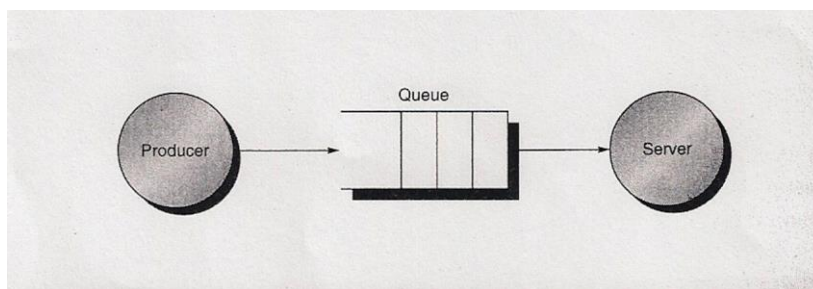
Ένα άλλο είναι η *χωρητικότητα*:

- Πόσες συσκευές I/O μπορούν να συνδεθούν σ' ένα υπολογιστικό σύστημα;

Επιπρόσθετα σ' αυτούς τους μοναδικούς δείκτες μέτρησης, τα παραδοσιακά μέτρα μέτρησης της επίδοσης είναι, ο χρόνος απόκρισης και η παραγωγικότητα, τα οποία εφαρμόζονται και στα I/O. (Η παραγωγικότητα I/O ονομάζεται πολλές φορές

εύρος ζώνης I/O, και ο χρόνος απόκρισης ονομάζεται πολλές φορές *λανθάνουσα*.) Το επόμενο σχήμα 1.5 δείχνει ένα απλό μοντέλο παραγωγού - διακομιστή. Ο παραγωγός δημιουργεί εργασίες για να εκτελεστούν και τις τοποθετεί στο buffer; ο διακομιστής παίρνει τις εργασίες από τον buffer(χρησιμοποιεί τεχνική πρώτος-μέσα-πρώτος-έξω, FIFO) και τις εκτελεί.

Ο χρόνος απόκρισης ξεκινά όταν μία εργασία τοποθετείται στο buffer και τερματίζεται όταν ολοκληρώνεται απ' τον server. Η παραγωγικότητα είναι ο αριθμός των εργασιών που ολοκληρώθηκαν απ' τον sever σε μια μονάδα του χρόνου.



Σχήμα 1.5 Μοντέλο παραγωγού-server χρόνου απόκρισης και παραγωγικότητας.

Ο χρόνος απόκρισης ορίζεται ως ο χρόνος μίας εργασίας από τη στιγμή που θα τοποθετηθεί στο buffer μέχρι ο server να τελειώσει την εργασία. Η παραγωγικότητα είναι απλώς ο μέσος αριθμός των εργασιών που ολοκληρώνονται από τον server σε μία περίοδο. Για να πάρουμε την μεγαλύτερη δυνατή παραγωγικότητα, ο server δεν θα γίνει ποτέ αδρανής, και άρα το buffer δεν θα είναι ποτέ άδειο. Ο χρόνος απόκρισης, απ' την άλλη πλευρά, μετράει τον χρόνο που ξοδεύεται στο buffer και άρα μειώνεται όταν ο buffer αδειάζει.[1]

ΚΕΦΑΛΑΙΟ 2 : ΓΕΝΙΚΑ ΓΙΑ ΤΟ RAID

Κατά τον σχεδιασμό επεξεργαστών λαμβάνονται υπόψη και η παραγωγικότητα και ο χρόνος απόκρισης. Όμως, σε ότι έχει σχέση με την αποθήκευση λαμβάνονται υπόψη η αξιοπιστία και η διαθεσιμότητα. Ο όρος αξιοπιστία χρησιμοποιείται συχνά εναλλακτικά με την διαθεσιμότητα κάτι που δεν είναι καθόλου σωστό.

Αν το σύστημα λειτουργεί συνεχώς λέμε ότι είναι αξιόπιστο.

Αν κάτι δεν λειτουργεί σωστά, αλλά ο χρήστης μπορεί να χρησιμοποιήσει ακόμη το σύστημα, φαίνεται δηλαδή σαν να λειτουργεί το σύστημα λέμε ότι το σύστημα είναι διαθέσιμο και άρα φαίνεται ακόμη αξιόπιστο. Εδώ είναι μια πιο ξεκάθαρη διάκριση:

Αξιοπιστία – Υπάρχει κάτι που δεν λειτουργεί;

Διαθεσιμότητα – Είναι το σύστημα διαθέσιμο για τον χρήστη;

Προσθέτοντας υλικό μπορεί κάποιος συνεπώς να βελτιώσει την διαθεσιμότητα(για παράδειγμα, το ECC στην μνήμη), αλλά δεν μπορεί να βελτιώσει την αξιοπιστία(το DRAM ακόμη δεν λειτουργεί). Η αξιοπιστία μπορεί να βελτιωθεί μόνο καλυτερεύοντας τις συνθήκες περιβάλλοντος, χρησιμοποιώντας περισσότερο αξιόπιστα υλικά, ή χρησιμοποιώντας λιγότερα υλικά. Μία άλλη έννοια, η ακεραιότητα δεδομένων, αναφέρεται στη σταθερή αποκατάσταση των δεδομένων όταν χάνεται μια πληροφορία εξαιτίας κάποιου σφάλματος. Αυτό είναι πολύ σημαντικό για κάποιες εφαρμογές.

Μία καινοτομία που βελτιώνει και την διαθεσιμότητα και την απόδοση των συστημάτων αποθήκευσης είναι οι συστοιχίες δίσκων. Το επιχείρημα για τις συστοιχίες είναι ότι αφού το κόστος ανά megabyte είναι ανεξάρτητο από το μέγεθος του δίσκου, η προσδοκόμενη παραγωγικότητα μπορεί να αυξηθεί προσθέτοντας πολλούς σκληρούς. Για παράδειγμα, η παραγωγικότητα του NFS(Network File system) αυξάνεται σ' ένα σύστημα όσο εξαπλώνεται από 9 δίσκους σε 109 δίσκους. Το μειονέκτημα των συστοιχιών είναι ότι με πολλές συσκευές, μειώνεται η αξιοπιστία: N συσκευές γενικά έχουν το $1/N$ την αξιοπιστία μίας μόνο συσκευής.

Οπότε, αφού μία συστοιχία δίσκων δεν μπορεί ποτέ να είναι πιο αξιόπιστη από τον μικρότερο αριθμό των μεγαλύτερων δίσκων, όταν κάθε δίσκος έχει τον ίδιο ρυθμό σφαλμάτων, η διαθεσιμότητα μπορεί να βελτιωθεί προσθέτοντας επιπλέον δίσκους. Αν ένας μόνο δίσκος αποτύχει, η χαμένη πληροφορία μπορεί να ανακατασκευαστεί από την επιπλέον πληροφορία. Ο μόνος κίνδυνος είναι να συμβεί κι άλλο σφάλμα δίσκου μεταξύ της χρονικής στιγμής που ο δίσκος αποτύχει και της στιγμής που αντικαθίσταται (ορίζεται ως μέση ώρα επιδιόρθωσης, ή MTTR). Αφού ο μέσος χρόνος σφάλματος (MTTF) των δίσκων είναι πέντε ή περισσότερα χρόνια, και ο MTTR μετριέται σε ώρες, ο πλεονασμός μπορεί να κάνει την διαθεσιμότητα 100 δίσκων πολύ μεγαλύτερη από αυτή ενός μόνο δίσκου. Αυτά τα συστήματα έχουν γίνει γνωστά με το ακρωνύμιο RAID, που σημαίνει Redundant Array of Inexpensive Disks.

Πίνακας 2.1 Επίπεδα RAID, η διαθεσιμότητα τους, και η ανοχή σε σφάλματα στους επιπλέον δίσκους.

Επίπεδο RAID	Ανοχή σφαλμάτων	Δίσκοι δεδομένων (παράδειγμα)	Δίσκοι ελέγχου
0 Μη πλεονασμός	0	8	0
1 Κατοπτρισμός	1	8	8
2 ECC μνήμη	1	8	4
3 Ισοτιμία παρεμβολής σε επίπεδο bit	1	8	1
4 Ισοτιμία παρεμβολής σε επίπεδο block	1	8	1
5 Ισοτιμία παρεμβολής με αφιερωμένη ισοτιμία	1	8	1
6 P+Q Πλεονασμός	2	8	2

Το RAID είναι ένα ακρωνύμιο που πρωτοορίστηκε από τους David A. Patterson, Garth A. Gibson και Randy Katz στο Πανεπιστήμιο της Καλιφόρνια University, στο Berkeley το 1987 για να περιγράψει την Πλεονασματική Διάταξη Φθηνών Δίσκων (Redundant Array of Inexpensive Disks), μία τεχνολογία που επιτρέπει στους

χρήστες υπολογιστών να πετύχουν υψηλά επίπεδα αποθηκευτικής αξιοπιστίας με συσκευές αποθήκευσης χαμηλού κόστους και λιγότερο αξιόπιστους σκληρούς δίσκους, μέσω της τεχνικής διευθέτησης των συσκευών σε συστοιχίες για πλεονασμό.

Πιο πρόσφατα, οι εμπορικοί αντιπρόσωποι των κατασκευαστών RAID αναθώρησαν τον όρο σε Πλεονασματική Διάταξη Ανεξάρτητων Δίσκων (Redundant Array of Independent Disks), ένα πιο βολικό όρο για να αποφεύγεται η προσδοκία του χαμηλού κόστους εξαιτίας του όρου "φθηνό".

Η φράση "RAID" χρησιμοποιείται τώρα ως όρος για συστήματα υπολογιστών αποθήκευσης δεδομένων που μπορούν να διαχωρίσουν και να αναπαράγουν δεδομένα μεταξύ πολλαπλών σκληρών δίσκων. Οι διάφοροι συνδυασμοί RAID περιλαμβάνουν δύο βασικούς σχεδιαστικούς στόχους:

- a) αυξημένη αξιοπιστία των δεδομένων ή
- b) αυξημένη επίδοση εισόδου / εξόδου.

Όταν διάφοροι φυσικοί δίσκοι έχουν συσταθεί για να χρησιμοποιηθεί η τεχνολογία RAID, λέγονται ότι είναι σε συστοιχία RAID. Αυτή η συστοιχία διανέμει τα δεδομένα σε πολλούς δίσκους, αλλά η συστοιχία η ίδια θεωρείται από τον χρήστη του υπολογιστή και το λειτουργικό σύστημα ως ένας μόνο λογικός δίσκος. Το RAID μπορεί να δημιουργηθεί για να εξυπηρετήσει πολλές διαφορετικές χρήσεις.

Η ανάγκη για την τεχνολογία RAID προέκυψε από την ανάγκη:

- για μεγαλύτερη παραγωγικότητα (μια συστοιχία μπορεί να προσπελαστεί παράλληλα)
- για μεγαλύτερη διαθεσιμότητα

Με ένα δίσκο δεδομένων, σε περίπτωση βλάβης και σε περίπτωση που διαθέτουμε backup, χάνουμε σημαντικό χρόνο για να αποκαταστήσουμε τα δεδομένα στην αρχική τους μορφή, μειώνοντας έτσι τη διαθεσιμότητα του συστήματος.

Με πολλαπλούς δίσκους και διαλέγοντας το κατάλληλο σχήμα RAID ένα σύστημα εξακολουθεί και λειτουργεί, ακόμα και όταν ένας δίσκος χαλάσει.

Η επιλογή ενός σχήματος RAID γίνεται με τέτοιο τρόπο ώστε να πληρούνται οι παρακάτω συνθήκες:

- ελαχιστοποίηση του χρόνου που απαιτείται για τη διαχείριση του
- μεγιστοποίηση του διαθέσιμου χώρου που απομένει μετά την εγκατάσταση
- μεγιστοποίηση του πλήθους των δίσκων που μπορούν να χρησιμοποιηθούν παράλληλα

2.1. Βασικές αρχές

Το RAID συνδυάζει δύο ή περισσότερους φυσικούς σκληρούς δίσκους σε μία μόνο λογική μονάδα χρησιμοποιώντας είτε ειδικό υλικό ή λογισμικό. Το ειδικό υλικό σχεδιάζεται έτσι ώστε να παρουσιάζεται στο σύστημα ως ένας σκληρός δίσκος, ενώ το λειτουργικό σύστημα δεν γνωρίζει τις τεχνικές ρυθμίσεις. Για παράδειγμα, μπορεί κάποιος να ρυθμίσει μία συστοιχία RAID 5 του ενός TB χρησιμοποιώντας τρεις δίσκους των 500 GB με υλικό RAID. Στο λειτουργικό σύστημα θα παρουσιαστεί απλώς ένας δίσκος του ενός TB. Το ειδικό λογισμικό υλοποιείται μέσα στο λειτουργικό σύστημα και παρουσιάζει τον δίσκο RAID ως έναν μόνο δίσκο σε εφαρμογές που τρέχουν στο λειτουργικό σύστημα.

Υπάρχουν τρεις βασικές έννοιες στο RAID:

- a) ο *κατοπτρισμός*, η αντιγραφή των δεδομένων σε περισσότερους από έναν δίσκους;
- b) ο *παραλληλισμός*, ο διαμερισμός των δεδομένων σε περισσότερους από έναν δίσκους;
- c) και η *διόρθωση λαθών*, όπου επιπλέον δεδομένα αποθηκεύονται για να επιτρέπουν σε λάθη να ανιχνεύονται και πιθανόν να διορθώνονται. (γνωστό ως ανοχή σε σφάλματα).

Τα διάφορα επίπεδα RAID χρησιμοποιούν μία ή περισσότερες από αυτές τις τεχνικές, ανάλογα με τις απαιτήσεις του συστήματος.

Στην περίπτωση που χρησιμοποιούμε επιπλέον δίσκους, λέμε ότι έχουμε πλεονασμό.

Ο κατοπτρισμός είναι μία από τις δύο τεχνικές πλεονασμού που χρησιμοποιούνται στο RAID(η άλλη είναι η τεχνική της ισοτιμίας). Σ' ένα σύστημα RAID που χρησιμοποιεί κατοπτρισμό, όλα τα δεδομένα στο σύστημα γράφονται ταυτόχρονα σε δύο σκληρούς δίσκους αντί για ένα. Η αρχή πίσω απ' τον κατοπτρισμό είναι ότι υπάρχει 100% πλεονασμός δεδομένων και παρέχεται πλήρη προστασία από σφάλματα που μπορεί να συμβούν σ' οποιοδήποτε δίσκο. Ο κατοπτρισμός πάντα απαιτεί ζυγό αριθμό δίσκων για εμφανείς λόγους.

Η τεχνική της ισοτιμίας είναι ο πιο γνωστός και πολυχρησιμοποιημένος κώδικας ανίχνευσης λάθους. Αυτός ο κώδικας αποτελείται από την πληροφορία και ένα bit ελέγχου. Η λειτουργία του είναι απλή. Δίνεται στο bit ισοτιμίας μία τιμή(1 ή 0) ώστε το άθροισμα όλων των μονάδων του κώδικα να είναι περιττός αριθμός(odd parity). Η πληροφορία έχει μεταφερθεί σωστά αν ο δέκτης μετρήσει όλα τα bit με τιμή 1 και βρεθεί περιττός αριθμός. Αν βρεθεί ζυγός αριθμός τότε κάποιο bit δεν μεταφέρθηκε σωστά αλλά δεν μπορεί να εντοπιστεί η θέση του. Επίσης αν κατά τη μεταφορά αλλοιωθούν δύο bits τότε η τεχνική της ισοτιμίας δεν μπορεί να τα ανιχνεύσει. Αντίστοιχα με την περιπτή ισοτιμία υπάρχει και η άρτια ισοτιμία(even parity).

Κατά την εμφάνιση μίας βλάβης ενός δίσκου, γίνεται προσπέλαση του δίσκου ισοτιμίας και τα δεδομένα αναδομούνται μέσω των συσκευών που απομένουν. Μόλις αντικατασταθεί ο δίσκος που υπέστη τη βλάβη, τα χαμένα δεδομένα μπορούν να επαναφερθούν στο νέο δίσκο και η λειτουργία να συνεχιστεί. Η αναδόμηση δεδομένων είναι απλή. Η πληροφορία ισοτιμίας σ' ένα περιβάλλον RAID υπολογίζεται χρησιμοποιώντας την συνάρτηση Boolean XOR.

Ας εξετάσουμε συστοιχία πέντε δίσκων, στην οποία οι δίσκοι X0 έως X3 περιέχουν δεδομένα, ενώ ο X4 είναι ο δίσκος ισοτιμίας. Η ισοτιμία του ιστού bit υπολογίζεται ως εξής:

$$X4(i) = X3(i) \oplus X2(i) \oplus X1(i) \oplus X0(i) \quad (2.1)$$

Όπου \oplus είναι η συνάρτηση αποκλειστικού 'Η.

Ας υποθέσουμε ότι ο οδηγός X_1 έχει υποστεί βλάβη. Αν προσθέσουμε το $X_4 \oplus X_1$ και στα δύο μέρη της προηγούμενης εξίσωσης, έχουμε

$$X_1(i) = X_4(i) \oplus X_3(i) \oplus X_2(i) \oplus X_0(i) \quad (2.2)$$

Συνεπώς, τα περιεχόμενα κάθε ζώνης δεδομένων του X_1 μπορούν να αναπαραχθούν από τα περιεχόμενα των αντιστοίχων ζωνών των υπολοίπων δίσκων της συστοιχίας. Ο κανόνας αυτός ισχύει για τα επίπεδα RAID 3 ως 6.

Τα κυριότερα προβλήματα μειωμένης απόδοσης στους αποθηκευτικούς δίσκους σχετίζονται με τον αργό ρυθμό των μηχανικών συστατικών που χρησιμοποιούνται για να τοποθετούν και να μεταφέρουν δεδομένα. Αφού μια συστοιχία RAID περιέχει πολλούς δίσκους, η απόδοση μπορεί να βελτιωθεί χρησιμοποιώντας το υλικό(hardware) παράλληλα σ' όλους αυτούς τους δίσκους. Για παράδειγμα, αν χρειάζεται κάποιος να διαβάσει ένα μεγάλο αρχείο αντί να το "τραβήξει" όλο από έναν μόνο σκληρό δίσκο, είναι πιο γρήγορο να το "κόψει" σε κομμάτια, να αποθηκεύσει κάποια απ' αυτά σε κάθε δίσκο στη συστοιχία, και μετά να χρησιμοποιήσει όλους τους δίσκους για να διαβάσει πίσω το αρχείο όταν χρειαστεί. Αυτή η τεχνική ονομάζεται παραλληλισμός(striping).

Ο παραλληλισμός μπορεί να γίνει σε επίπεδο byte, ή σε blocks. Ο παραλληλισμός σε επίπεδο byte σημαίνει ότι το αρχείο σπάει σε "κομμάτια μεγέθους byte". Το πρώτο byte του αρχείου στέλνεται στον πρώτο δίσκο, το δεύτερο byte στον δεύτερο δίσκο, και ου το καθεξής. Κάποιες φορές ο παραλληλισμός σε επίπεδο byte γίνεται ως ένας τομέας(sector) των 512 bytes. Ο παραλληλισμός σε επίπεδο block σημαίνει ότι το κάθε αρχείο σπάει σε blocks ενός συγκεκριμένου μεγέθους και αυτά τα blocks διανέμονται σε διάφορους δίσκους. Το μέγεθος των blocks που χρησιμοποιείται ονομάζεται επίσης μέγεθος ζώνης(stripe size).

Ο βασικός στόχος του RAID μπορεί να είναι:

- a) η βελτίωση της αξιοπιστίας και η διαθεσιμότητα των δεδομένων, διασφαλίζοντας ότι τα σημαντικά δεδομένα θα είναι διαθέσιμα πιο συχνά από ότι καθόλου διαθέσιμα, π.χ. μία βάση δεδομένων με εντολές πελατών,

- b) ή απλώς η βελτίωση της ταχύτητας πρόσβασης στα αρχεία, π.χ. για ένα σύστημα που παρέχει προγράμματα τηλεόρασης με video on demand σε πολλούς θεατές.

Η διαμόρφωση επηρεάζει την αξιοπιστία και την απόδοση με διαφορετικούς τρόπους. Το πρόβλημα με την χρήση πολλών δίσκων είναι ότι είναι πιο πιθανό ένας δίσκος να αποτύχει, αλλά χρησιμοποιώντας τον έλεγχο σφαλμάτων ολόκληρο το σύστημα μπορεί να γίνει πιο αξιόπιστο αφού μπορεί να επιβιώσει και να διορθώσει ένα σφάλμα. Ο βασικός κατοπτρισμός μπορεί να επιταχύνει την ανάγνωση δεδομένων, αφού το σύστημα μπορεί να διαβάσει διαφορετικά δεδομένα και από τους δύο δίσκους, αλλά μπορεί να είναι αργό στην εγγραφή αν η διαμόρφωση απαιτεί ότι και οι δύο δίσκοι πρέπει να επιβεβαιώνουν ότι τα δεδομένα γράφτηκαν σωστά. Ο παραλληλισμός συχνά χρησιμοποιείται για απόδοση, όπου επιτρέπει σε ακολουθίες δεδομένων να διαβάζονται από διαφορετικούς δίσκους στον ίδιο χρόνο. Ο έλεγχος σφαλμάτων τυπικά θα επιβραδύνει το σύστημα αφού τα δεδομένα πρέπει να διαβάζονται από διάφορα μέρη και να συγκρίνονται.

2.2. Οργάνωση δίσκων

Ο πλεονασμός επιτυγχάνεται είτε γράφοντας τα ίδια δεδομένα σε πολλούς δίσκους(γνωστό ως κατοπτρισμός), ή γράφοντας επιπλέον δεδομένα(γνωστό ως δεδομένα ισοτιμίας) στην συστοιχία, υπολογισμένα έτσι ώστε το σφάλμα ενός(ή πιθανότατα και περισσότερων ανάλογα με τον τύπο του RAID) δίσκου στη συστοιχία δεν θα προκαλέσει απώλεια δεδομένων. Ένας κατεστραμμένος δίσκος μπορεί να αντικατασταθεί από κάποιον άλλο καινούργιο, και τα χαμένα δεδομένα ανακατασκευάζονται από τα δεδομένα που απομένουν και τα δεδομένα ισοτιμίας. Οργανώνοντας τους δίσκους σε πλεονάζουσα συστοιχία μειώνεται η χρησιμοποιούμενη χωρητικότητα αποθήκευσης. Για παράδειγμα, μία συστοιχία RAID 1 δύο δίσκων χάνει τη μισή απ' την ολική της χωρητικότητα η οποία σε διαφορετική περίπτωση θα ήταν διαθέσιμη χρησιμοποιώντας και τους δύο δίσκους ξεχωριστά, και μία συστοιχία RAID 5 με αρκετούς δίσκους χάνει την χωρητικότητα του ενός δίσκου. Άλλες RAID συστοιχίες είναι διατεταγμένες έτσι ώστε να είναι πιο γρήγορες στην εγγραφή και στην ανάγνωση, από ό, τι ένας μόνο δίσκος.

Υπάρχουν διάφοροι συνδυασμοί αυτών των προσεγγίσεων που δίνουν διαφορετικά σημεία προστασίας έναντι στην απώλεια δεδομένων, την χωρητικότητα και την ταχύτητα. Τα επίπεδα RAID 0, 1, και 5 είναι τα συνηθέστερα που βρίσκονται, και καλύπτουν τις περισσότερες απαιτήσεις.

- Το RAID 0(Disk Striping) είναι το πιο απλό επίπεδο RAID, το οποίο δεν παρέχει πλεονασμό. Τα δεδομένα διαχωρίζονται σε ζώνες και διανέμονται σ' όλους τους δίσκους της συστοιχίας. Το μέγεθος κάθε ζώνης ορίζεται απ' τον χρήστη. Απαιτούνται τουλάχιστον 2 δίσκοι, ίδιου τύπου και μεγέθους. Δεν παρέχεται πλεονασμός, αλλά αυτό το επίπεδο χρησιμοποιείται για να αυξήσει την απόδοση.
- Το RAID 1 (Mirroring) διπλασιάζει τα δεδομένα κάθε δίσκου σ' όλη την συστοιχία, παρέχοντας πλήρη πλεονασμό. Δύο δίσκοι καθένας μπορεί να αποθηκεύσει ακριβώς τα ίδια δεδομένα, στον ίδιο χρόνο, και όλες τις φορές. Τα δεδομένα δεν χάνονται, εφόσον ένας δίσκος λειτουργεί. Σε κάθε περίπτωση, τα περιεχόμενα κάθε δίσκου της συστοιχίας είναι πανομοιότυπα με αυτά οποιουδήποτε άλλου δίσκου της συστοιχίας.
- Το RAID 2(Bit-level striping with Hamming code ECC) είναι το μόνο επίπεδο RAID που δεν χρησιμοποιεί καμία από τις πρότυπες τεχνικές δηλαδή κατοπτρισμό, παραλληλισμό ή/και ισοτιμία. Το RAID 2 χρησιμοποιεί μία παρόμοια τεχνική με τον παραλληλισμό με ισοτιμία. Υλοποιείται διαχωρίζοντας τα δεδομένα σε επίπεδο bit και τα διανέμει σ' όλους τους δίσκους. Απαιτούνται N δίσκοι που θα περιέχουν δεδομένα και m δίσκοι που θα περιέχουν τις πληροφορίες πλεονασμού. Η συστοιχία θα αποτελείται από N+m δίσκους.
- Τα RAID 3 και 4 (Striped disks with dedicated parity) συνδυάζουν τρεις ή περισσότερους δίσκους μ' ένα τρόπο ώστε να προστατεύουν τα δεδομένα από απώλεια οποιουδήποτε δίσκου. Η ανοχή στα σφάλματα πετυχαίνεται προσθέτοντας έναν ακόμη δίσκο στη συστοιχία, ο οποίος χρησιμοποιείται μόνο για να αποθηκεύει τις πληροφορίες ισοτιμίας. Η χωρητικότητα της συστοιχίας αυξάνεται κατά έναν δίσκο. Μια συστοιχία RAID 3 ή 4 απαιτεί τουλάχιστον τρεις δίσκους: δύο για να αποθηκεύουν τα δεδομένα και έναν για την πληροφορία ισοτιμίας.

- Το RAID 5 (Striped disks with distributed parity) συνδυάζει τρεις ή περισσότερους δίσκους με τρόπο που να προστατεύει τα δεδομένα από απώλεια οποιουδήποτε δίσκου; είναι παρόμοιο με το RAID 3 αλλά η πληροφορία ισοτιμίας δεν αποθηκεύεται σ' ένα αφιερωμένο δίσκο, αλλά διανέμεται σ' όλους τους δίσκους της συστοιχίας. Μόνο ένας δίσκος μπορεί να μην λειτουργεί σε οποιαδήποτε υλοποίηση RAID 5 χωρίς να χαθούν δεδομένα.
- Το RAID 6 (Striped disks with dual parity) συνδυάζει τέσσερις ή παραπάνω δίσκους μ' ένα τρόπο ώστε να μπορεί να ανακάμψει από τις απώλειες δύο δίσκων.
- Το RAID 10 (ή 1+0) χρησιμοποιεί και διαγραμμισμένους δίσκους και κατοπτρισμό. Το "01" ή "0+1" πολλές φορές διακρίνεται από το "10" ή "1+0" : ένα διαγραμμισμένο σετ από δίσκους κάτοπτρα και ένα κατοπτρισμένο σετ από διαγραμμισμένους δίσκους.

Το RAID μπορεί να περιλαμβάνει σημαντικό υπολογισμό κατά την ανάγνωση και τη γραφή πληροφοριών. Με πραγματικό υλικό RAID, ένα ξεχωριστός ελεγκτής κάνει όλες τις εργασίες αυτού του υπολογισμού. Σε άλλες περιπτώσεις, το λειτουργικό σύστημα ή απλούστεροι και λιγότερο δαπανηροί ελεγκτές απαιτούν από τον επεξεργαστή του κεντρικού υπολογιστή να κάνει τον υπολογισμό, κάτι που μειώνει την επίδοση του υπολογιστή σε εργασίες επεξεργαστή. Απλούστερους ελεγκτές RAID μπορούν να παρέχουν μόνο τα επίπεδα 0 και 1, τα οποία απαιτούν λιγότερη επεξεργασία.

Τα συστήματα RAID με πλεονασμό συνεχίζουν να εργάζονται χωρίς διακοπή όταν ένας(ή μερικές φορές και περισσότεροι, ανάλογα με τον τύπο του RAID) δίσκος της συστοιχίας αποτύχει, αν και μετά είναι ευάλωτοι σε περαιτέρω σφάλματα. Όταν ο χαλασμένος δίσκος αντικαθίσταται από ένα νέο η συστοιχία ξαναχτίζεται ενώ το σύστημα συνεχίζει να λειτουργεί κανονικά. Ορισμένα συστήματα θα πρέπει να σταματήσουν την λειτουργία τους κατά την αφαίρεση ή προσθήκη ενός δίσκου; άλλοι υποστηρίζουν hot swapping, επιτρέποντας στους δίσκους να αντικατασταθούν χωρίς να απενεργοποιηθούν. Το RAID με hot swapping χρησιμοποιείται συχνά σε συστήματα με υψηλή διαθεσιμότητα, όπου είναι

σημαντικό το σύστημα να λειτουργεί όσο το δυνατόν για περισσότερο χρονικό διάστημα.

2.3. Ιστορία

Στον Norman Ken Ouchi της IBM απονεμήθηκε το 1978 η αμερικάνικη πατέντα 4,092,732 με τον τίτλο "Σύστημα για επανάκτηση δεδομένων αποθηκευμένα σε εσφαλμένη μονάδα μνήμης". Οι αξιώσεις γι' αυτήν την πατέντα περιγράφουν ότι αργότερα θα οριστεί ως RAID 5 με πλήρεις εγγραφές γραμμών. Αυτή η πατέντα του 1978 επισημαίνει επίσης ότι ο κατοππρισμός δίσκων ή διπλασιασμός(αργότερα θα ορισθεί ως RAID 1)και η προστασία με αφιερωμένη ισοτιμία(που αργότερα θα ονομαστεί RAID 4) δεν ήταν έγκυρα εκείνο τον καιρό.

Ο όρος RAID πρωτοορίστηκε από τους David A. Patterson, Garth A. Gibson και Randy Katz στο Πανεπιστήμιο της Καλιφόρνια University, στο Berkeley το 1987. Μελέτησαν την πιθανότητα χρήσης δύο ή περισσότερων δίσκων ώστε να εμφανίζονται ως μία ενιαία συσκευή στο σύστημα και εκδώσαν μία μελέτη: "Μια μελέτη για Πλεονασματικές Συστοιχίες Φθηνών Δίσκων" τον Ιούνιο του 1988 στο συνέδριο SIGMOD.

Αυτή η προδιαγραφή πρότεινε έναν αριθμό πρωτότυπων επιπέδων RAID, ή συνδυασμών δίσκων. Κάθε ένα είχε θεωρητικά πλεονεκτήματα και μειονεκτήματα. Με τα χρόνια, εμφανίστηκαν διαφορετικές υλοποιήσεις του RAID. Πολλά διέφεραν ουσιαστικά από τα πρωτότυπα επίπεδα RAID, αλλά τα ονόματα παρέμειναν. Αυτό μπορεί να μπερδέψει κάποιον, αφού μία υλοποίηση του RAID 5, για παράδειγμα, μπορεί να διαφέρει ουσιαστικά από άλλη. Τα RAID 3 και RAID 4 συχνά συγχέονται και χρησιμοποιούνται εναλλακτικά.

ΚΕΦΑΛΑΙΟ 3 : STANDARD ΕΠΙΠΕΔΑ RAID

Ένας αριθμός από πρότυπα σχέδια έχει εξελιχθεί και αναφέρεται ως *επίπεδα RAID*. Αρχικά υπήρχαν πέντε επίπεδα RAID, αλλά υλοποιήθηκαν πολλές διαφορετικές εκδοχές, δημιουργήθηκαν πολλά εμφωλευμένα επίπεδα και πολλά μη-πρότυπα επίπεδα.

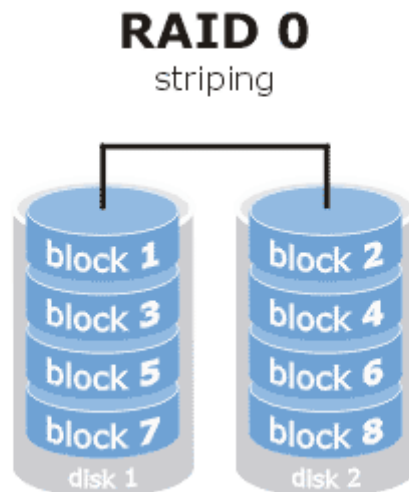
3.1. RAID 0

Μία συστοιχία RAID 0 (Striping without parity) διαχωρίζει τα δεδομένα ομοιόμορφα σε δύο ή περισσότερους δίσκους(διαγραμμισμένους, striped) χωρίς να υπάρχει πληροφορία ισοτιμίας για πλεονασμό δεδομένων. Είναι σημαντικό να σημειώσει κάποιος ότι το RAID 0 δεν ήταν ένα από τα πρωτότυπα επίπεδα RAID και δεν παρέχει πλεονασμό δεδομένων. Το RAID 0 χρησιμοποιείται για να αυξήσει την απόδοση, παρόλο που μπορεί να χρησιμοποιηθεί και ως ένας τρόπος για να δημιουργηθεί ένας μικρός αριθμός μεγάλων εικονικών δίσκων από έναν μεγάλο αριθμό μικρών φυσικών δίσκων.

Σύμφωνα με το RAID 0, τα δεδομένα του χρήστη και του συστήματος κατανέμονται σε όλους τους δίσκους της συστοιχίας. Αυτό έχει ένα αξιοσημείωτο πλεονέκτημα σε σχέση με τη χρήση ενός μεγάλου δίσκου: Αν εκκρεμούν δύο διαφορετικές αιτήσεις I/O για διαφορετικά blocks δεδομένων, τότε υπάρχει η πιθανότητα τα blocks που ζητούνται να βρίσκονται σε διαφορετικούς δίσκους. Συνεπώς, οι δύο αιτήσεις μπορούν να εκδοθούν παράλληλα, μειώνοντας έτσι το χρόνο τοποθέτησης σε ουρά της I/O.

Ένα RAID 0 μπορεί να δημιουργηθεί με δίσκους διαφορετικών μεγεθών, αλλά ο αποθηκευτικός χώρος που προστίθεται στην συστοιχία από κάθε δίσκο μειώνεται στο μέγεθος του μικρότερου δίσκου. Για παράδειγμα, αν ένας δίσκος 120 GB παραλληλιστεί με έναν δίσκο των 100 GB, το μέγεθος της συστοιχίας θα είναι 200 GB.

$$\begin{aligned} \text{Size} &= 2 \cdot \min(120 \text{ GB}, 100 \text{ GB}) && (3.1) \\ &= 2 \cdot 100 \text{ GB} = 200 \text{ GB} \end{aligned}$$



Σχήμα 3.1 Διάγραμμα εγκατάστασης RAID 0

3.1.1. Ρυθμός σφαλμάτων RAID 0

Παρόλο που το RAID 0 δεν προσδιορίστηκε στην αρχική μελέτη RAID, μία ιδεατή υλοποίηση του θα χώριζε τις λειτουργίες I/O σε ίσου μεγέθους blocks και θα τις διένειμε ομοιόμορφα σ' όλους τους δίσκους. Οι υλοποιήσεις του RAID 0 με περισσότερους από δύο δίσκους, είναι πάλι πιθανές, παρόλο που η αξιοπιστία της συστοιχίας μειώνεται ανάλογα με το μέγεθος κάθε μέλους.

Η αξιοπιστία μιας συστοιχίας RAID 0 είναι ίση με την μέση αξιοπιστία του κάθε δίσκου διαιρεμένη με τον αριθμό των δίσκων της συστοιχίας:

$$MTTF_{\text{group}} \approx MTTF_{\text{disk}} / \text{number} \quad (3.2)$$

Η αξιοπιστία, η οποία μετρείται από τον Μέσο Χρόνο Σφάλματος ΜΧΣ (Mean Time to Failure, MTTF) ή από τον Μέσο Χρόνο Μεταξύ Σφαλμάτων ΜΧΜΣ (Mean Time Between Failures, MTBF) είναι περίπου αντιστρόφως ανάλογη με τον αριθμό των μελών - οπότε σε μια συστοιχία δύο δίσκων η αξιοπιστία είναι ίση περίπου με το μισό της αξιοπιστίας ενός δίσκου. Αν υπήρχε μία πιθανότητα 0.05 ότι ο δίσκος θα έσφαλε σε τρία χρόνια, σε μία συστοιχία δύο δίσκων, αυτή η πιθανότητα θα ανέβαινε στο

$$\begin{aligned} P(\text{τουλάχιστον-ένας-να-σφάλει}) &= 1 - P(\text{κανένας-δεν-σφάλει}) && (3.3) \\ &= 1 - (1-0.05)^2 \\ &= 0.0975 \\ &\approx 0.1 \end{aligned}$$

Ο λόγος γι' αυτό είναι ότι το σύστημα αρχείων είναι διανεμημένο σ' όλους τους δίσκους. Όταν ένας δίσκος σταματήσει να λειτουργεί, το σύστημα αρχείων δεν μπορεί να ανταπεξέλθει μια τόσο μεγάλη απώλεια δεδομένων και συνοχής αφού τα δεδομένα είναι "διαγραμμισμένα" σ' όλους τους δίσκους.(τα δεδομένα δεν μπορούν επανέλθουν χωρίς τον εσφαλμένο δίσκο) Τα δεδομένα μπορούν να επανέλθουν χρησιμοποιώντας διάφορα ειδικά εργαλεία. Παρόλα αυτά, αυτά τα δεδομένα θα είναι ανολοκλήρωτα και πολύ πιθανόν κατεστραμμένα, ενώ η επανάκτηση τους είναι πολύ δαπανηρή και καθόλου εγγυημένη.

3.1.2. Επίδοση RAID 0

Παρόλο που το μέγεθος του block μπορεί τεχνικά να είναι τόσο μικρό όσο ένα byte, είναι σχεδόν πάντα το πολλαπλάσιο ενός τμήματος ενός σκληρού δίσκου π.χ. 512 bytes. Αυτό επιτρέπει σε κάθε δίσκο να αναζητεί ανεξάρτητα όταν γίνεται ανάγνωση και εγγραφή δεδομένων με τυχαίο τρόπο στον δίσκο. Το πόσο ανεξάρτητα ενεργούν οι δίσκοι εξαρτάται από το τρόπο πρόσβασης του κάθε επιπέδου συστήματος αρχείων. Για αναγνώσεις και εγγραφές που είναι μεγαλύτερες από το μέγεθος γραμμής(stripe size), όπως η αντιγραφή αρχείων ή το playback ενός βίντεο, οι δίσκοι θα ψάχνουν στο ίδιο σημείο κάθε δίσκου, ώστε ο χρόνος αναζήτησης της συστοιχίας να είναι ο ίδιος όπως και αυτός ενός μόνο δίσκου. Για αναγνώσεις και εγγραφές που είναι μικρότερες από το μέγεθος γραμμής, όπως προσβάσεις σε βάσεις δεδομένων, οι δίσκοι θα είναι ικανοί να αναζητούν ανεξάρτητα. Αν οι τομείς πρόσβασης είναι διασκορπισμένοι ομοιόμορφα σε δύο δίσκους, ο εμφανής χρόνος αναζήτησης της συστοιχίας θα είναι ο μισός από αυτόν ενός μόνο δίσκου(υποθέτοντας ότι οι δίσκοι στην συστοιχία έχουν πανομοιότυπα χαρακτηριστικά χρόνου πρόσβασης). Η ταχύτητα

μεταφοράς της συστοιχίας θα είναι η συνολική ταχύτητα μεταφοράς όλων των δίσκων μαζί, μειωμένη μόνο από την ταχύτητα του ελεγκτή RAID.

Η χρήση του RAID 0 για επίτευξη υψηλού ρυθμού μεταφοράς δεδομένων, πρέπει να ικανοποιεί δύο απαιτήσεις. Πρώτον, πρέπει να υπάρχει υψηλή ικανότητα μεταφοράς κατά μήκος ολόκληρης της διαδρομής ανάμεσα στην κεντρική μνήμη και τους μεμονωμένους οδηγούς δίσκου. Αυτό περιλαμβάνει τους εσωτερικούς διαύλους ελεγκτή, τους διαύλους I/O του υπολογιστικού συστήματος, τους προσαρμογείς I/O και τους διαύλους κεντρικής μνήμης.

Η δεύτερη απαίτηση είναι το γεγονός ότι η εφαρμογή πρέπει να υποβάλλει αιτήσεις I/O, που να αξιοποιούν τις δυνατότητες της συστοιχίας δίσκων. Η απαίτηση αυτή ικανοποιείται, αν η συνήθης αίτηση αφορά μεγάλους όγκους λογικά παρακείμενων δεδομένων, σε σύγκριση με το μέγεθος μίας ζώνης. Στην περίπτωση αυτή, μία αίτηση I/O περιλαμβάνει την παράλληλη μεταφορά δεδομένων από πολλαπλούς δίσκους, αυξάνοντας τον αποτελεσματικό ρυθμό μεταφοράς, σε σύγκριση με τη μεταφορά ενός δίσκου.

Σε περιβάλλοντα που είναι προσανατολισμένα στις συναλλαγές, ο χρήστης συνήθως ενδιαφέρεται περισσότερο για το χρόνο απόκρισης παρά για το ρυθμό μεταφοράς. Ο χρόνος I/O μίας μεμονωμένης αίτησης I/O για μικρό όγκο δεδομένων καθορίζεται κυρίως από τη μετακίνηση των κεφαλών του δίσκου(χρόνος αναζήτησης) και από την κίνηση του δίσκου(λανθάνων χρόνος περιστροφής).

Επίσης, σε περιβάλλοντα συναλλαγών, μπορεί να υπάρχουν εκατοντάδες αιτήσεις I/O ανά δευτερόλεπτο. Η συστοιχία δίσκων μπορεί να παράσχει υψηλούς ρυθμούς εκτέλεσης I/O, εξισορροπώντας το φορτίο I/O μεταξύ πολλαπλών δίσκων. Αποτελεσματική εξισορρόπηση φορτίου επιτυγχάνεται, μόνο αν υπάρχουν συνήθως πολλαπλές εκκρεμείς αιτήσεις I/O. Αυτό, με τη σειρά του, συνεπάγεται ότι υπάρχουν πολλαπλές ανεξάρτητες εφαρμογές ή ότι υπάρχει εφαρμογή που είναι προσανατολισμένη στις συναλλαγές, η οποία είναι ικανή να υποβάλλει πολλαπλές ασύγχρονες αιτήσεις I/O. Η απόδοση, επίσης, επηρεάζεται από το μέγεθος ζώνης. Αν το μέγεθος ζώνης είναι σχετικά μεγάλο, έτσι ώστε μία αίτηση I/O να περιλαμβάνει μια πρόσβαση στο δίσκο, τότε μπορεί να γίνει η παράλληλη

διαχείριση των αιτήσεων I/O που βρίσκονται σε αναμονή, μειώνοντας το χρόνο τοποθέτησης κάθε αίτησης στην ουρά.

Το RAID 0 είναι χρήσιμο για εγκαταστάσεις όπως μεγάλοι NFS(Network File System) Διακομιστές μόνο ανάγνωσης (read-only NFS Servers) όπου η φόρτωση(mounting) πολλών δίσκων είναι χρονοβόρα ή αδύνατη και ο πλεονασμός είναι άσχετος.

Το RAID 0 χρησιμοποιείται επίσης και σε κάποια συστήματα παιχνιδιών όπου επιθυμείται απόδοση ενώ η ακεραιότητα των δεδομένων δεν τόσο σημαντική. Παρόλα αυτά, πραγματικές δοκιμές(real-world tests) με παιχνίδια έδειξαν ότι τα κέρδη απόδοσης του RAID 0 είναι μικρά, παρόλο που κάποιες εφαρμογές desktop θα ωφεληθούν.

Μελέτες συμπέραναν ότι: "Ο παραλληλισμός δεν αυξάνει πάντοτε την απόδοση (σε ορισμένες περιπτώσεις είναι πραγματικά πιο αργό από ότι μία εγκατάσταση μη-RAID), αλλά στις περισσότερες περιπτώσεις θα παράγει μία σημαντική αύξηση απόδοσης".

3.2. RAID 1

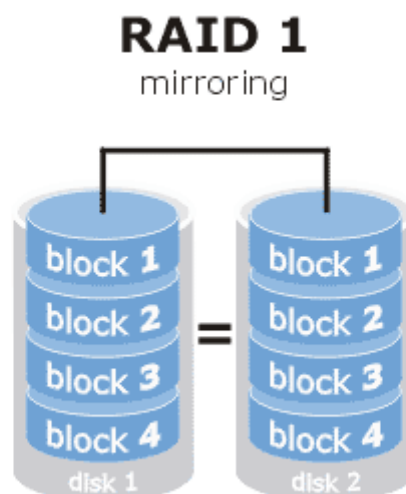
Το RAID 1 δημιουργεί ένα ακριβές αντίγραφο(ή καθρέπτη) των δεδομένων δύο δίσκων. Αυτό είναι χρήσιμο όταν η απόδοση ανάγνωσης ή η αξιοπιστία είναι πιο σημαντικά από την χωρητικότητα δεδομένων. Τέτοια συστοιχία μπορεί να είναι τόσο μεγάλη όσο το μικρότερο μέλος του δίσκου. Ένα κλασικό ζευγάρι καθρεπτών RAID 1(βλέπε Σχήμα 3.2) περιέχει δύο δίσκους, το οποίο αυξάνει την αξιοπιστία από έναν δίσκο. Αφού κάθε μέλος περιέχει ένα ολοκληρωμένο αντίγραφο των δεδομένων, και μπορεί να διευθυνσιοποιηθεί ανεξάρτητα, η αξιοπιστία αυξάνεται ως δύναμη του αριθμού των περιεχόμενων αντιγράφων.

Η οργάνωση RAID 1 έχει αρκετά πλεονεκτήματα:

1. Οποιοσδήποτε από τους δύο δίσκους που περιέχουν τα δεδομένα μπορεί να εξυπηρετήσει μία αίτηση ανάγνωσης. Επιλέγεται εκείνος που έχει το μικρότερο χρόνο αναζήτησης και το μικρότερο λανθάνοντα χρόνο περιστροφής.

2. Μία αίτηση εγγραφής απαιτεί την ενημέρωση και των δύο αντίστοιχων ζωνών, κάτι που όμως μπορεί να πραγματοποιηθεί παράλληλα. Άρα, η απόδοση εγγραφής εξαρτάται από την πιο αργή από τις δύο εγγραφές(π.χ. αυτή που έχει το μεγαλύτερο χρόνο αναζήτησης και το μεγαλύτερο λανθάνοντα χρόνο περιστροφής). Ωστόσο, δεν υπάρχει "ποινή εγγραφής" στο RAID 1. Τα επίπεδα RAID 2 ως 6 περιλαμβάνουν τη χρήση bits ισοτιμίας. Συνεπώς, όταν ενημερώνεται μία ζώνη, το λογισμικό διαχείρισης της ισοτιμίας πρέπει πρώτα να υπολογίσει και να ενημερώσει τα bits ισοτιμίας και ακολούθως την πραγματική ζώνη.
3. Η επαναφορά από μία βλάβη είναι απλή. όταν κάποιος δίσκος υφίσταται βλάβη, μπορεί κανείς να προσπελάσει τα δεδομένα από το δεύτερο δίσκο.

Το βασικό μειονέκτημα του RAID 1 είναι το κόστος. Απαιτεί το διπλάσιο χώρο δίσκου του λογικού δίσκου που υποστηρίζει. Εξαιτίας αυτού, ο σχηματισμός RAID 1 είναι πιθανό να περιορίζεται σε οδηγούς που αποθηκεύουν λογισμικό και δεδομένα συστήματος, καθώς και άλλα ιδιαίτερως σημαντικά αρχεία. Στις περιπτώσεις αυτές, το RAID 1 παρέχει υποστήριξη εφεδρικών αντιγράφων όλων των δεδομένων σε πραγματικό χρόνο, έτσι ώστε σε περιπτώσεις βλάβης δίσκου, όλα τα σημαντικά δεδομένα να είναι άμεσα διαθέσιμα.



Σχήμα 3.2 Διάγραμμα μιας εγκατάστασης RAID 1

3.2.1. Ρυθμός σφαλμάτων RAID 1

Ως ένα απλό παράδειγμα, θεωρήστε ένα RAID 1 με δύο πανομοιότυπα μοντέλα ενός σκληρού δίσκου με 0.05 πιθανότητα ότι ο δίσκος θα σφάλει σε τρία χρόνια. Υποθέτοντας ότι τα σφάλματα είναι στατιστικά ανεξάρτητα, τότε η πιθανότητα να σφάλουν και οι δύο δίσκοι σε διάρκεια τριών χρόνων είναι

$$P(\text{και-οι-δύο-σφάλουν}) = (0.05)^2 = 0.0025 \quad (3.4)$$

Άρα, η πιθανότητα να χαθούν όλα τα δεδομένα είναι 0.0025 αν ο πρώτος δίσκος δεν αντικατασταθεί ποτέ. Αν ένας από τους δύο δίσκους αποτύχει, κανένα δεδομένο δεν θα χαθεί, προϋποθέτοντας ότι ο εσφαλμένος δίσκος θα αντικατασταθεί πριν αποτύχει και ο δεύτερος.

Παρόλα αυτά, αφού δύο πανομοιότυποι δίσκοι χρησιμοποιούνται και αφού τα μοτίβα χρήσης τους είναι επίσης πανομοιότυπα, οι αποτυχίες τους δεν μπορούν να εννοηθούν ανεξάρτητες. Άρα η πιθανότητα να χαθούν όλα τα δεδομένα είναι σημαντικά υψηλότερη από 0.0025 αλλά μικρότερη από 0.05.

3.2.2. Επίδοση RAID 1

Επειδή όλα τα δεδομένα υπάρχουν σε δύο ή περισσότερα αντίγραφα, κάθε ένα με το δικό του υλικό, η απόδοση ανάγνωσης μπορεί να ανεβεί περίπου σαν γραμμικό πολλαπλάσιο του αριθμού των αντιγράφων. Άρα, μία συστοιχία RAID 1 με δύο δίσκους μπορεί να διαβάζει δύο διαφορετικά μέρη την ίδια χρονική στιγμή, παρόλο που δεν μπορούν όλες οι υλοποιήσεις του RAID 1 να το κάνουν αυτό. Για να μεγιστοποιηθούν τα οφέλη απόδοσης του RAID 1, χρειάζονται ανεξάρτητοι ελεγκτές δίσκων, ένας για κάθε δίσκο. Κάποιοι αναφέρονται σ' αυτή τη τεχνική ως διαχωρισμός ή διπλασιασμός. Κατά την ανάγνωση, και οι δύο δίσκοι μπορούν να προσπελαστούν ξεχωριστά και οι αιτούμενοι τομείς μπορούν να χωριστούν όμοια στους δίσκους. Οπότε για τον κοινό κατοπτρισμό δύο δίσκων, θεωρητικά, αυτό θα διπλασίαζε τον ρυθμό μεταφοράς κατά την ανάγνωση. Ο εμφανής χρόνος πρόσβασης της συστοιχίας θα ήταν ο μισός από αυτόν ενός μόνο δίσκου. Σε αντίθεση με το RAID 0, αυτό θα ισχύσει για όλες τις πατέντες πρόσβασης, αφού

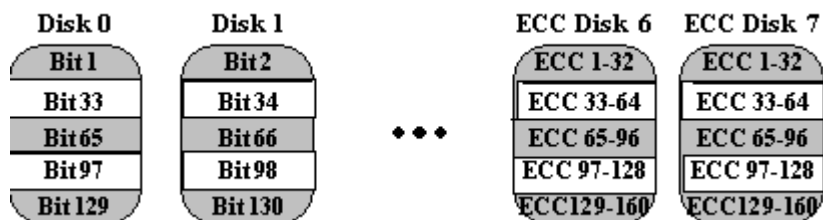
όλα τα δεδομένα υπάρχουν στους δίσκους. Στην πραγματικότητα, η ανάγκη να μετακινηθούν οι κεφαλές των δίσκων στο επόμενο block(να παραλείψουν τα block που έχουν διαβαστεί από άλλους δίσκους) μπορεί να μετριάσει αποδοτικά τα πλεονεκτήματα ταχύτητας για πρόσβαση ακολουθίας. Η απόδοση ανάγνωσης μπορεί να βελτιωθεί επιπλέον προσθέτοντας δίσκους στους καθρέπτες δίσκους. Πολλοί παλιοί ελεγκτές IDE RAID1 διαβάζουν μόνο απ' τον ένα δίσκο του ζευγαριού, οπότε η απόδοση ανάγνωσής τους είναι πάντα αυτή του ενός δίσκου. Κάποιες παλιότερες υλοποιήσεις RAID 1 μπορούν επίσης να διαβάσουν και τους δύο δίσκους ταυτόχρονα και να συγκρίνουν τα δεδομένα ώστε να εντοπίζουν σφάλματα. Ο εντοπισμός λαθών και η διόρθωσή τους στους μοντέρνους δίσκους κάνει αυτό λιγότερο χρήσιμο σε περιβάλλοντα που απαιτούν μέτρια διαθεσιμότητα. Κατά την εγγραφή, η συστοιχία συμπεριφέρεται ως ένας δίσκος, αφού όλοι οι δίσκοι καθρέπτες πρέπει να εγγραφούν με τα δεδομένα

Το RAID 1 έχει πολλά διαχειριστικά πλεονεκτήματα. Για παράδειγμα, σε κάποια περιβάλλοντα, είναι πιθανό να «χωρίζει κάποιος τον δίσκο καθρέπτη»: διακηρύσσοντας έναν δίσκο ως ανενεργό, παίρνοντας ένα αντίγραφο ασφαλείας αυτού του δίσκου, και ύστερα «ξαναχτίζοντας» τον δίσκο καθρέπτη. Αυτό είναι χρήσιμο σε περιπτώσεις που το σύστημα αρχείων πρέπει να είναι συνεχώς διαθέσιμο. Αυτό απαιτεί ότι η εφαρμογή υποστηρίζει ανάκτηση από την εικόνα των δεδομένων στον δίσκο στο σημείο που χωρίστηκε ο δίσκος καθρέπτης. Αυτή η διαδικασία είναι λιγότερο κρίσιμη στην παρουσία του χαρακτηριστικού "στιγμιότυπο"(snapshot) σε κάποια συστήματα αρχείων, στα οποία κάποιος χώρος εξυπηρετεί για αλλαγές, παρουσιάζοντας ένα στατικό σημείο του συστήματος αρχείων. Εναλλακτικά, ένας νέος δίσκος μπορεί να αντικατασταθεί ώστε ο ανενεργός δίσκος μπορεί να κρατηθεί με τον ίδιο τρόπο όπως ένα παραδοσιακό αντίγραφο ασφαλείας. Για να κρατηθεί ο πλεονασμός κατά τη δημιουργία αντιγράφου ασφαλείας, κάποιοι ελεγκτές υποστηρίζουν την πρόσθεση και τρίτου δίσκου στο ενεργό ζευγάρι. Αφού ολοκληρωθεί η ανάκτηση του τρίτου δίσκου, γίνεται ανενεργό και λαμβάνεται αντίγραφο ασφαλείας όπως περιγράφεται παραπάνω.

3.3. RAID 2

Το RAID 2 είναι το "μαύρο πρόβατο" της οικογένειας RAID, επειδή είναι το μόνο επίπεδο RAID που δεν χρησιμοποιεί καμία από τις πρότυπες τεχνικές δηλαδή κατοπτρισμό, παραλληλισμό ή/και ισοτιμία. Το RAID 2 χρησιμοποιεί μία παρόμοια τεχνική με τον παραλληλισμό με ισοτιμία. Υλοποιείται διαχωρίζοντας τα δεδομένα σε επίπεδο bit και τα διανέμει σ' όλους τους δίσκους. Το επίπεδο 2 κάνει χρήση του κώδικα Hamming(7,4) ο οποίος κωδικοποιεί τέσσερα bits δεδομένων σε 7 bits προσθέτοντας τρία bits ισοτιμίας. Οπότε μία τυπική υλοποίηση RAID 2 επιτρέπει να χρησιμοποιηθούν 7 δίσκοι, με τους 4 να χρησιμοποιούνται για αποθήκευση δεδομένων και τους 3 για την διόρθωση σφαλμάτων. Γενικά απαιτούνται N δίσκοι που θα περιέχουν δεδομένα και m δίσκοι που θα περιέχουν τις πληροφορίες πλεονασμού. Η συστοιχία θα αποτελείται από N+m δίσκους.

Δηλαδή αν έχουμε ένα block δίσκου 4 bit χρειαζόμαστε 4 δίσκους , ένα για κάθε bit και 3 δίσκους για τα 3 bit ισοτιμίας του Hamming. Το 1988 η εταιρεία UNISYS κατασκεύασε ένα πειραματικό RAID 2 σύστημα για να αποθηκεύει block δίσκου με μήκος όσο και το μήκος μια θέσης μνήμης (32 bits). Χρειάστηκε 32 δίσκους για τα δεδομένα και 7 για τα bit ισοτιμίας του κώδικα Hamming. Εμπορικό προϊόν όμως δεν κατασκευάστηκε ποτέ.



Σχήμα 3.3 Σύστημα RAID 2 με 39 δίσκους

Όπως αναφέρθηκε παραπάνω τα bits πλεονασμού υπολογίζονται χρησιμοποιώντας κώδικες Hamming, μία μορφή κώδικα διόρθωσης σφαλμάτων(Error Correction Code – ECC). Κάθε φορά που κάτι γράφεται στη συστοιχία αυτοί οι κώδικες υπολογίζονται και γράφουν τα δεδομένα σε αφιερωμένους δίσκους ECC; όταν τα δεδομένα διαβάζονται πίσω, αυτοί οι ECC κώδικες διαβάζονται επίσης για να επιβεβαιωθεί ότι δεν συνέβη κανένα σφάλμα κατά τη διάρκεια της εγγραφής των δεδομένων. Αν ένα σφάλμα ενός μόνο bit

συμβαίνει, τότε μπορεί να διορθωθεί εκείνη την στιγμή. Το RAID 2 είναι το μόνο επίπεδο RAID, πέρα από κάποιες υλοποιήσεις του RAID 6, που μπορεί αυτόματα να επανακτήσει ακριβή δεδομένα από φθορά σε δεδομένα του ενός bit. Τα άλλα επίπεδα RAID μπορούν να εντοπίσουν την φθορά σε δεδομένα του ενός bit, ή μπορούν κάποιες φορές να επανακτήσουν τα χαμένα δεδομένα, αλλά δεν μπορούν να επιλύσουν αντιφάσεις αξιόπιστα ανάμεσα σε bits ισοτιμίας και bits δεδομένων χωρίς την ανθρώπινη συμβολή.

Κατά την ανάγνωση, υπάρχει ταυτόχρονη πρόσβαση σε όλους τους δίσκους. Τα δεδομένα που έχουν ζητηθεί και ο αντίστοιχος κώδικας διόρθωσης σφαλμάτων παραδίδονται στον ελεγκτή συστοιχίας. Αν υπάρχει σφάλμα του ενός bit, ο ελεγκτής το αναγνωρίζει και το διορθώνει άμεσα, έτσι ώστε να μην επιβραδύνεται ο χρόνος πρόσβασης. Κατά την εγγραφή, πρέπει να υπάρχει πρόσβαση σε όλους τους δίσκους δεδομένων και τους δίσκους ισοτιμίας, ώστε να πραγματοποιηθεί η λειτουργία εγγραφής.

Η φθορά πολλαπλών-bit είναι πιθανή αλλά πολύ σπάνια. Το RAID 2 μπορεί να εντοπίσει την φθορά σε διπλό bit αλλά όχι να την διορθώσει.

Το Επίπεδο 2 είναι το μόνο επίπεδο RAID που έχει οριστεί στην πρωτότυπη μελέτη του Berkley και δεν χρησιμοποιείται σήμερα, για πολλούς λόγους. Έχει πολύ ακριβό κόστος καθώς απαιτεί πολλούς δίσκους. Επίσης απαιτείται ένας σύνθετος ελεγκτής, ειδικής κατασκευής και φυσικά με υψηλό κόστος.

Η επίδοση του RAID 2 είναι κατώτερη σε περιβάλλοντα συναλλαγών εξαιτίας του παραλληλισμού σε επίπεδο bit. Μπορεί να είναι αποτελεσματικό μόνο σε περιβάλλοντα που προκύπτουν πολλά σφάλματα. Έχοντας ως δεδομένη την υψηλή αξιοπιστία των μεμονωμένων δίσκων και των οδηγών δίσκων, το RAID 2 θεωρείται υπερβολικό και δεν υλοποιείται.

Σήμερα δεν υπάρχουν εμπορικές υλοποιήσεις του RAID 2.

3.4. RAID 3

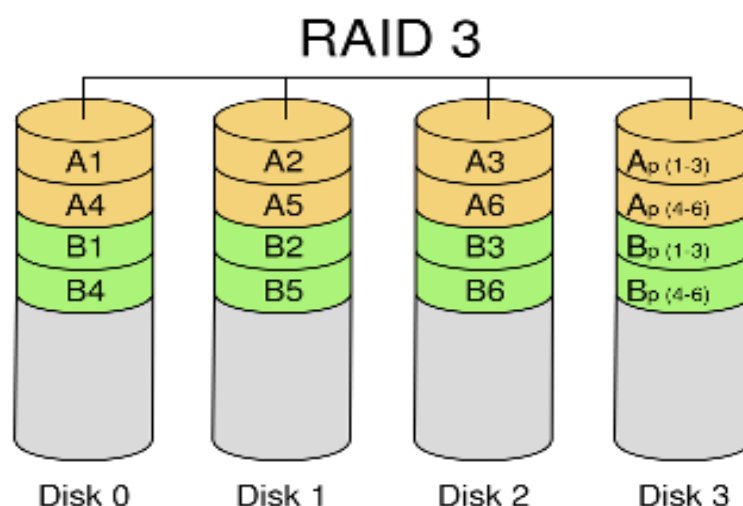
Στο RAID 3(Byte-level striping with dedicated parity), τα δεδομένα διανέμονται στους δίσκους σε επίπεδο byte, δηλαδή τα δεδομένα κατανέμονται σε μικρές ζώνες και διανέμονται σε N δίσκους δεδομένων. Επίσης χρειάζεται ένας

αφιερωμένος δίσκος στον οποίο στέλνεται η πληροφορία ισοτιμίας, οπότε η συστοιχία θα αποτελείται από $N+1$ δίσκους. Συνήθως το ελάχιστο που χρειάζονται είναι 3 δίσκοι. Ο ακριβής αριθμός των byte που στέλνονται σε κάθε ζώνη ποικίλει αλλά συνήθως είναι κάτω από 1024. Η πληροφορία ισοτιμίας στέλνεται στον αφιερωμένο δίσκο ισοτιμίας, αλλά αν συμβεί κάποιο σφάλμα σε οποιοδήποτε δίσκο στη συστοιχία μπορεί να αντιμετωπιστεί.

3.4.1. Επίδοση RAID 3

Η χαρακτηριστική απόδοση του RAID 3 είναι πολύ σταθερή, σε αντίθεση με τα υψηλότερα επίπεδα RAID. Το μέγεθος μίας ζώνης, είναι μικρότερο από το μέγεθος ενός τομέα ή από το block του ΛΣ, και για ανάγνωση και για εγγραφή, ενώ ολόκληρη η ζώνη έχει προσβασιμότητα κάθε φορά. Η απόδοση της συστοιχίας, άρα θα είναι όμοια με την απόδοση ενός μόνο δίσκου στην συστοιχία εκτός από τις τιμές μεταφοράς, που πολλαπλασιάζονται με τον αριθμό των δίσκων δεδομένων.

Στο παράδειγμα(Σχήμα 3.3), μία αίτηση για το block A που αποτελείται από τα bytes A1-A6 θα χρειαστεί και τους τρεις δίσκους για να ψάξει στην αρχή(A1) και να απαντήσει με τα περιεχόμενα τους. Μία ταυτόχρονη αίτηση για το block B θα έπρεπε να περιμένει.



Σχήμα 3.4 Διάγραμμα εγκατάστασης RAID 3 με blocks των 6-bytes και δύο bytes ισοτιμίας

Η απόδοση ανάγνωσης με τυχαίο τρόπο είναι καλή αλλά όχι τέλεια εξαιτίας του παραλληλισμού σε επίπεδο byte. Η απόδοση σε συνεχόμενες αναγνώσεις είναι πολύ καλή. Στις εγγραφές υπάρχει μειωμένη απόδοση, ειδικά στις εγγραφές με τυχαίο τρόπο εξαιτίας του παραλληλισμού σε επίπεδο byte, του υπολογισμού της πληροφορίας ισοτιμίας καθώς και του φαινομένου “Λαιμός Μπουκαλιού” που συμβαίνει στο δίσκο ισοτιμίας.

Ένα από τα μειονεκτήματα του RAID 3 είναι ότι γενικά δεν μπορεί να εξυπηρετήσει πολλές αιτήσεις ταυτόχρονα. Αυτό γίνεται γιατί κάθε block δεδομένων θα διαμοιραστεί, εξ ορισμού, σ' όλα τα μέλη της συστοιχίας και θα διανεμηθεί στην ίδια τοποθεσία. Οπότε, κάθε λειτουργία I/O απαιτεί δραστηριότητα σε κάθε δίσκο.

Το κόστος του RAID 4 είναι μέτριο αφού χρειάζεται συνήθως ένας ελεγκτής υλικού, καθώς και τρεις δίσκοι. Εφαρμογές που απαιτούν τις υψηλότερες τιμές μεταφοράς, ειδικά εφαρμογές που εξυπηρετούν και επεξεργάζονται μεγάλα αρχεία, για παράδειγμα: multimedia, μη συμπιεσμένο μοντάζ βίντεο, κ.τ.λ. Το RAID 3 χρησιμοποιείται συχνά για εφαρμογές που θα χρησιμοποιούνταν το RAID 0, αλλά η έλλειψη της ανοχής σε σφάλματα του RAID 0 το έκανε μη αποδεκτό.

3.5. RAID 4

Το RAID 4(Block-level striping with dedicated parity) χρησιμοποιεί παραλληλισμό επιπέδου block με ένα αφιερωμένο δίσκο ισοτιμίας. Τα δεδομένα διανέμονται σ' όλους τους δίσκους σε blocks, δηλαδή οι ζώνες έχουν όλες ένα συγκεκριμένο μέγεθος.

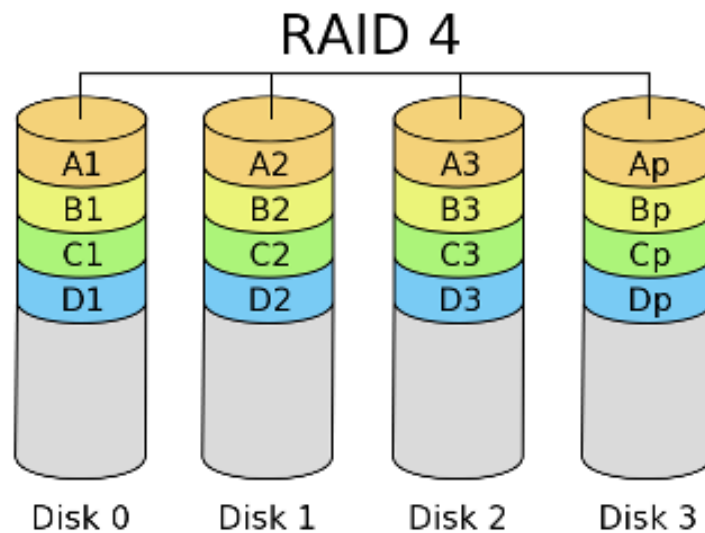
Αυτό επιτρέπει σε κάθε δίσκο της συστοιχίας να ενεργεί ανεξάρτητα όταν αιτείται ένα μόνο block. Αν το επιτρέψει ο ελεγκτής δίσκου, μία συστοιχία RAID 4 μπορεί να εξυπηρετήσει πολλαπλές αιτήσεις ταυτόχρονα. Εξαιτίας αυτού, οι συστοιχίες ανεξάρτητης προσπέλασης είναι πιο κατάλληλες για εφαρμογές που απαιτούν υψηλούς ρυθμούς αιτήσεων I/O και σχετικά λιγότερο κατάλληλες για εφαρμογές που απαιτούν υψηλούς ρυθμούς μεταφοράς δεδομένων. Το RAID 4 μοιάζει με το RAID 3 εκτός απ' το γεγονός ότι διαμοιράζει σε επίπεδο block και όχι σε επίπεδο byte. Γενικά, το RAID 4 υλοποιείται με υποστήριξη υλικού για υπολογισμούς ισοτιμίας, και απαιτούνται γενικά

$$S_{\min} * (N-1) \quad (3.5)$$

δίσκοι, με S_{\min} το μέγεθος του μικρότερου δίσκου και N τον αριθμό των δίσκων στη συστοιχία.

Στις υλοποιήσεις συνήθως χρησιμοποιούνται τουλάχιστον τρεις δίσκοι για μία ολοκληρωμένη διαμόρφωση RAID 4.

Στο παράδειγμα(Σχήμα 3.4), μία αίτηση ανάγνωσης για το block A1 θα εξυπηρετούνταν από τον δίσκο 0. Μία ταυτόχρονη αίτηση ανάγνωσης για το block B1 θα πρέπει να περιμένει, αλλά μία αίτηση για το block B2 θα μπορούσε να εξυπηρετηθεί ταυτόχρονα από τον δίσκο 1.



Σχήμα 3.5 Διάγραμμα μίας εγκατάστασης RAID 4 με αφιερωμένη ισοτιμία. Κάθε χρώμα αναπαριστά την ομάδα των block στο αντίστοιχο block ισοτιμίας(γραμμή – stripe).

3.5.1. Επίδοση RAID 4

Η επίδοση των λειτουργιών ανάγνωσης, είτε είναι τυχαίες είτε είναι συνεχόμενες είναι πολύ καλή.

Δυστυχώς για την λειτουργία της εγγραφής ο δίσκος ισοτιμίας γίνεται "λαιμός μπουκαλιού", περιλαμβάνεται δηλαδή μια ποινή εγγραφής, αφού ταυτόχρονες εγγραφές στα A1 και B2 εκτός απ' τις εγγραφές στους αντίστοιχους δίσκους τους, χρειάζεται επίσης να εγγραφούν και οι δύο στον δίσκο ισοτιμίας. Κάθε φορά που προκύπτει μία εγγραφή, το λογισμικό διαχείρισης συστοιχίας πρέπει να ενημερώνει όχι μόνο τα δεδομένα χρήστη αλλά και τα αντίστοιχα bits ισοτιμίας.

Έτσι σε κάθε εγγραφή το RAID 4 διοχετεύει μεγάλο φόρτο εργασίας στον δίσκο ισοτιμίας.

Στη περίπτωση εγγραφής I/O μεγαλύτερου μεγέθους, που περιλαμβάνει ζώνες σε όλους τους οδηγούς δίσκου, η ισοτιμία υπολογίζεται εύκολα, χρησιμοποιώντας μόνο τα νέα bits δεδομένων. Άρα, ο δίσκος ισοτιμίας μπορεί να ενημερωθεί εύκολα παράλληλα με τους οδηγούς δεδομένων, χωρίς να υπάρχουν επιπλέον αναγνώσεις και εγγραφές.

Σε κάθε περίπτωση, κάθε λειτουργία εγγραφής πρέπει να περιλαμβάνει το δίσκο ισοτιμίας, ο οποίος επομένως μπορεί να αποτελέσει ανασχετικό παράγοντα.

3.5.2. Υλοποίηση

Το κόστος του RAID 4 είναι μέτριο αφού χρειάζεται συνήθως ένας ελεγκτής υλικού, καθώς και τρεις δίσκοι. Το RAID 4 δεν χρησιμοποιείται συχνά όσο το RAID 3 ή το RAID 5 επειδή είναι κατά κάποιον τρόπο ο “συμβιβασμός” μεταξύ των δύο επιπέδων. Δεν απευθύνεται σε συγκεκριμένες ομάδες όπως τα άλλα δύο επίπεδα. Χρησιμοποιείται συχνά σε εφαρμογές που μπορούν να χρησιμοποιηθούν τα RAID 3 ή RAID 5.

3.6. RAID 5

Ένα από τα πιο δημοφιλή επίπεδα RAID, το επίπεδο RAID 5 (Block-level striping with distributed parity) διανέμει σε ζώνες και τα δεδομένα και την πληροφορία ισοτιμίας σε τρεις ή παραπάνω δίσκους. Είναι όμοιο με το RAID 4 εκτός απ’ το ότι ανταλλάσει τον αφιερωμένο δίσκο ισοτιμίας με ένα αλγόριθμο διανομής της ισοτιμίας, γράφοντας τα δεδομένα και τα blocks ισοτιμίας σ’ όλους τους δίσκους στη συστοιχία. Απαιτούνται λοιπόν N+1 δίσκοι, οι οποίοι χρησιμοποιούνται όλοι για δεδομένα και πληροφορίες ισοτιμίας. Συνήθως για να επιτευχθεί το RAID 5 χρειάζονται το ελάχιστο 3 δίσκοι. Σε κάποιες υλοποιήσεις, μπορεί να γίνει μία υποβαθμισμένη συστοιχία δίσκων RAID 5 (τρεις δίσκοι, απ’ τους οποίους μόνο οι δύο είναι ενεργοί), ενώ το mdadm υποστηρίζει μία πλήρως μεθοδική εγκατάσταση RAID 5 (μη υποβαθμισμένη) με δύο δίσκους - που λειτουργούν σαν ένα αργό RAID 1, αλλά μπορεί να επεκταθεί με περισσότερους τόμους.

Η λύση του RAID 5 αφαιρεί το φαινόμενο "λαιμός μπουκαλιού" που συμβαίνει στον αφιερωμένο δίσκο ισοτιμίας(στο επίπεδο 4), βελτιώνοντας ελαφρώς την απόδοση εγγραφής και επιτρέποντας καλύτερα τον παραλληλισμό σ' ένα περιβάλλον πολλαπλών συναλλαγών, παρόλο που η υπερφόρτωση που συμβαίνει με την ισοτιμία συνεχίζει να καθυστερεί τις εγγραφές.

Το RAID 5 πέτυχε την δημοτικότητά του εξαιτίας του χαμηλού κόστους πλεονασμού. Αυτό μπορεί να το δει κανείς συγκρίνοντας τον αριθμό των δίσκων που χρειάζονται για να επιτευχθεί μία δεδομένη χωρητικότητα.

Αν ποικίλουν οι δίσκοι σε χωρητικότητα, ο μικρότερος απ' αυτούς θέτει το όριο οπότε, η χωρητικότητα σε χρήση μιας συστοιχίας RAID 5 είναι

$$(N-1)*S_{min} \quad (3.6)$$

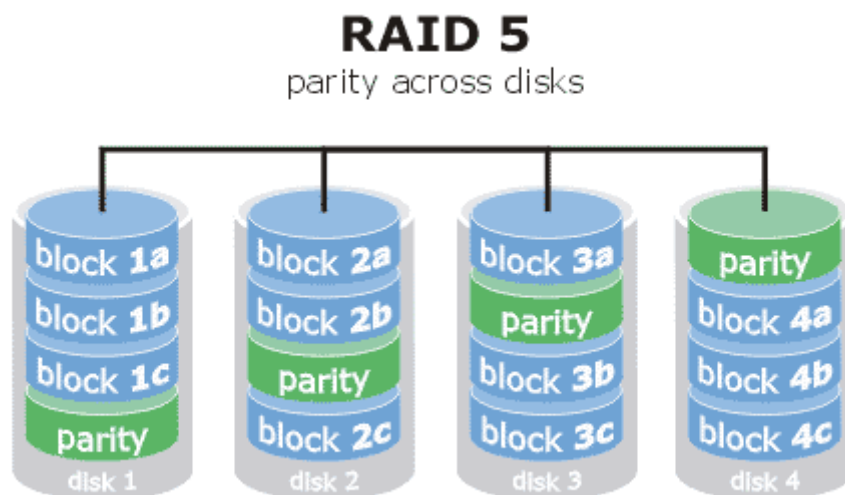
όπου N είναι ο συνολικός αριθμός των δίσκων στη συστοιχία και S_{min} η χωρητικότητα του μικρότερου δίσκου στη συστοιχία.

Ο αριθμός των σκληρών δίσκων που μπορούν να ανήκουν σε μια συστοιχία είναι θεωρητικά απεριόριστος.

Για παράδειγμα, τέσσερις δίσκοι του 1 TB μπορούν γίνουν μία συστοιχία των 2 TB είτε με RAID 1 ή με RAID 1+0, αλλά οι ίδιοι τέσσερις δίσκοι μπορούν να χρησιμοποιηθούν για να χτιστεί μία συστοιχία των 3 TB με RAID 5.

Παρόλο που το RAID 5 υλοποιείται συχνά μ' ένα ελεγκτή δίσκου, μπορεί επίσης να υλοποιηθεί και σε επίπεδο λειτουργικού συστήματος, π.χ., χρησιμοποιώντας τους Δυναμικούς Δίσκους των Windows(Windows Dynamic Disks) ή με το mdadm σε Linux. Κάποιες υλοποιήσεις χρησιμοποιούν υποστήριξη υλικού για υπολογισμούς ισοτιμίας(κάρτες υλικού RAID) και κάποιες χρησιμοποιούν τον κεντρικό επεξεργαστή του συστήματος(ελεγκτές βασισμένοι στην μητρική κάρτα).

Στο σχήμα 3.5, μία αίτηση ανάγνωσης για το block 1a θα εξυπηρετούνταν από τον δίσκο 1. Μία ταυτόχρονη αίτηση ανάγνωσης για το block 1b θα έπρεπε να περιμένει, αλλά μία αίτηση ανάγνωσης για το 2b θα μπορούσε να εξυπηρετηθεί ταυτόχρονα από τον δίσκο 2.



Σχήμα 3.6 Διάγραμμα της εγκατάστασης RAID 5 με διαμοιραζόμενη ισοτιμία.

3.6.1 Χειρισμός ισοτιμίας RAID 5

Οι παράλληλες σειρές από block (ένα block σε κάθε ένα από τους δίσκους στην συστοιχία) ονομάζονται όπως είπαμε και νωρίτερα **ζώνες**(stripes). Αν ένα άλλο block, ή κάποιο μέρος του, γράφεται στην ίδια ζώνη, το block ισοτιμίας, ή κάποιο μέρος του, υπολογίζεται ξανά και ξαναγράφεται. Για μικρές εγγραφές, αυτό απαιτεί.

- Να διαβαστούν τα παλιά δεδομένα του block.
- Να διαβαστεί η παλιά ισοτιμία του block.
- Σύγκριση του παλιού block δεδομένων με μία αίτηση εγγραφής. Για κάθε bit που αλλάζει(από 0 σε 1, ή από 1 σε 0) στο block δεδομένων, αλλάζει και το αντίστοιχο bit στο block ισοτιμίας.
- Να εγγραφούν τα νέα block δεδομένων.
- Να εγγραφεί το νέο block ισοτιμίας.

Ο δίσκος που χρησιμοποιούνται στο RAID 4 για τα block ισοτιμίας πλέον χρησιμοποιείται και για αποθήκευση δεδομένων ενώ blocks ισοτιμίας διανέμονται σ' όλους τους δίσκους, από εκεί και ο όρος **διαγραμματισμένα block ισοτιμίας**.

Τα block ισοτιμίας δεν διαβάζονται κατά την ανάγνωση δεδομένων, αφού αυτό θα ήταν περιττή υπερφόρτωση και θα μείωνε την απόδοση. Τα block ισοτιμίας διαβάζονται, παρόλα αυτά, όταν χρησιμοποιούνται, μια λειτουργία ανάγνωσης των blocks και το αντίστοιχο block ισοτιμίας στη ίδια ζώνη για να ανακατασκευάσουν έναν τομέα που δεν λειτουργεί. Οπότε, αν ένας δίσκος στη συστοιχία αποτύχει, τα block ισοτιμίας από τους δίσκους που απομένουν συνδυάζονται μαθηματικά με τα block δεδομένων για να ανακατασκευαστούν τα δεδομένα στον εσφαλμένο δίσκο εκείνη την στιγμή.

Αυτό ονομάζεται κάποιες φορές Προσωρινή Κατάσταση Ανάκτησης Δεδομένων (Interim Data Recovery Mode). Ο υπολογιστής γνωρίζει ότι ένας δίσκος έχει σταματήσει να λειτουργεί, αλλά αυτό συμβαίνει μόνο επειδή το λειτουργικό σύστημα μπορεί να πληροφορήσει τον διαχειριστή ότι ο δίσκος χρειάζεται αντικατάσταση; οι εφαρμογές που τρέχουν στον υπολογιστή δεν γνωρίζουν για το σφάλμα. Η ανάγνωση και η εγγραφή στη συστοιχία των δίσκων συνεχίζει, με κάποια μείωση της απόδοσης βέβαια.

3.6.2. Επίδοση RAID 5

Οι υλοποιήσεις RAID 5 υποφέρουν από μειωμένη απόδοση όταν αντιμετωπίζουν φόρτο εργασίας το οποίο περιλαμβάνει πολλές εγγραφές οι οποίες είναι μικρότερες από την χωρητικότητα μιας μόνο ζώνης. Αυτό συμβαίνει επειδή η ισοτιμία πρέπει να αναβαθμιστεί σε κάθε εγγραφή, απαιτώντας ακολουθίες ανάγνωσης-τροποποίησης-έγγραφής και για τα block δεδομένων και για το block ισοτιμίας. Πιο περίπλοκες υλοποιήσεις μπορεί να περιλαμβάνουν μια μνήμη non-volatile cache – δηλαδή που διατηρεί τις πληροφορίες ακόμη και όταν δεν έχει ρεύμα – για εγγραφές που μειώνει τις επιπτώσεις απόδοσης των ενημερώσεων ισοτιμίας.

Η τυχαία απόδοση εγγραφής είναι μειωμένη, ειδικά σε υψηλά επίπεδα συγχρονισμού συχνά σε μεγάλες βάσεις δεδομένων με πολλούς χρήστες. Ο κύκλος ανάγνωσης-τροποποίησης-έγγραφής του RAID 5 επιβάλλει κυρώσεις σε τυχαίες εγγραφές.

Τα προβλήματα απόδοσης μπορεί να είναι τόσο σοβαρά που κάποιοι ειδικοί βάσεων δεδομένων ίδρυσαν μια ομάδα με όνομα BAART-The Battle Against Any Raid Five.

Η απόδοση ανάγνωσης του RAID 5 είναι σχεδόν τόσο καλή όσο αυτή του RAID 0 για ίδιο αριθμό δίσκων. Ειδικότερα είναι εξαιρετική για τυχαίες ενέργειες ανάγνωσης σε blocks με μεγάλο μέγεθος ζώνης. Σε συνεχόμενες αναγνώσεις είναι πολύ καλή, ειδικά σε blocks με μικρό μέγεθος ζώνης. Εκτός από τα block ισοτιμίας, η διανομή των δεδομένων στους δίσκους ακολουθεί το ίδιο μοτίβο όπως το RAID 0. Ο λόγος που το RAID 5 είναι ελαφρώς αργότερο είναι ότι οι δίσκοι πρέπει να παραλείψουν τα block ισοτιμίας.

Σε περίπτωση σφάλματος συστήματος ενώ υπάρχουν ενεργές εγγραφές, η ισοτιμία μιας ζώνης μπορεί να γίνει ασυνεπής με τα δεδομένα. Αν αυτό δεν εντοπιστεί και επισκευαστεί πριν ο δίσκος ή το block καταστραφεί, τότε θα προκληθεί απώλεια δεδομένων ενώ η λανθασμένη ισοτιμία θα χρησιμοποιηθεί για να ξαναφτιαχτεί το χαμένο block στη ζώνη. Αυτή η πιθανή ευπάθεια είναι γνωστή κάποιες φορές ως **λευκή τρύπα**. Για να μειωθεί η πιθανότητα να συμβεί αυτό το γεγονός χρησιμοποιούνται συχνά μνήμες cache battery-backed και παρόμοιες τεχνικές.

3.6.3. Ρυθμός Σφαλμάτων RAID 5

Ο μέγιστος αριθμός των δίσκων σε μια συστοιχία RAID 5 είναι θεωρητικά απεριόριστος, αλλά είναι κοινή πρακτική να υπάρχει όριο στον αριθμό των δίσκων. Μεγαλύτερες συστοιχίες έχουν μεγαλύτερη πιθανότητα να συμβεί σ' αυτές ένα ταυτόχρονο διπλό σφάλμα, οπότε χρειάζεται περισσότερος χρόνος για να ξαναχτιστεί η συστοιχία, και υπάρχει μεγαλύτερη πιθανότητα να αντιμετωπιστεί ένας μη ανακτήσιμος τομέας δίσκου κατά την επανάκτηση του RAID 5. Όσο μεγαλώνει ο αριθμός των δίσκων του RAID 5, ο ΜΧΜΣ(Μέσος Χρόνος Μεταξύ Σφαλμάτων, Mean Time Between Failures, MTBF) μπορεί να γίνει μικρότερος από αυτόν ενός μόνο δίσκου. Αυτό συμβαίνει όταν η πιθανότητα ενός δεύτερου δίσκου να σφάλει από N-1 δίσκους, μέσα στο χρόνο που χρειάζεται για να εντοπιστεί, να αντικατασταθεί και να ξαναδημιουργηθεί ο πρώτος εσφαλμένος δίσκος, μεγαλώνει σε σχέση με την πιθανότητα ενός μόνο δίσκου να αποτύχει.

Μία επιδείνωση σ' αυτό το ζήτημα είναι ένας σχετικά στάσιμος ρυθμός ανεπανόρθωτων σφαλμάτων ανάγνωσης τα τελευταία χρόνια, ο οποίος είναι κατά κανόνα της τάξεως του ενός σφάλματος σε 10^4 bits για δίσκους SATA. Ενώ το μέγεθος των δίσκων έχει αυξηθεί δραστικά (> 1TB) τα πρόσφατα χρόνια, στην πραγματικότητα είναι πιθανό με μία συστοιχία ~10 TB να συμβεί ένα μη αναστρέψιμο σφάλμα ανάγνωσης κατά την ανακατασκευή ενός RAID 5. Κάποια τέτοια πιθανά σφάλματα μπορούν να αποφευχθούν σε συστήματα RAID τα οποία ελέγχουν τους δίσκους τους αυτόματα και περιοδικά σε στιγμές μη μεγάλης ζήτησης. Η χρήση δαπανηρών δίσκων με μικρότερα μεγέθη και καλύτερους ρυθμούς σφαλμάτων (περίπου 10^{-15}) βοήθησε, αλλά ακόμη και αυτές οι εγκαταστάσεις RAID 5 θα υποφέρουν από το ίδιο πρόβλημα τα επόμενα χρόνια εκτός αν βελτιωθούν οι ρυθμοί σφαλμάτων ανάγνωσης με ρυθμό που να ισοδυναμεί με το μέγεθος των δίσκων.

Τα SSD's (Solid-state drives) παρουσιάζουν έναν επαναστατικό παρά εξελικτικό τρόπο για να αντιμετωπίσουν αυτό το πρόβλημα. Με την ενθάρρυνση από πολλούς κατασκευαστές Flash-SSD (Flash – solid state drive), το JEDEC προετοιμάζεται να θέσει πρότυπα μέσα στο 2009 για να μετρήσει το UBER (Uncorrectable Bit Error Rates, Ρυθμός Μη Διορθώσιμων Bit) και τους ρυθμούς σφαλμάτων bit των “γραμμών” (ρυθμοί σφαλμάτων πριν το ECC).

Αλλά ακόμη και οι οικονομικής κλάσης Intel X25-E SDD παρουσιάζουν στους σημερινούς δίσκους 15K-RPM HDD ρυθμό μη διορθώσιμων bit ίσο με 1 τομέα σε 10^{15} bits και MTBF ίσο με 2 εκατομμύρια ώρες. Η πολύ πιο γρήγορη παραγωγικότητα των SSD's, που δείχνει ρυθμό σφαλμάτων της τάξεως των 10^{15} bits, θα είχε ως αποτέλεσμα την μείωση του MTBF. Η STEC υποστηρίζει ότι οι δίσκοι Zeus SSD's πετυχαίνουν ρυθμούς έως 200 φορές καλύτερους από τους σημερινούς της κλάσης HDD's. Τα SSD's DRAM έχουν ακόμη λιγότερα σφάλματα και παρόμοια ανωριμότητα σε τεχνολογία.

Το RAID 6 είναι μια εναλλακτική λύση που παρέχει προστασία διπλής ισοτιμίας ενώ επιπλέον ενεργοποιεί μεγάλους αριθμούς δίσκων ανά ομάδα. Κάποιες υλοποιήσεις έχουν έναν ανταλλακτικό δίσκο που μπορεί να χρησιμοποιηθεί για να ξαναχτιστεί αμέσως ένας χαλασμένος δίσκος.

3.6.4. Υλοποίηση

Το κόστος του RAID 5 είναι μέτριο, και συχνά πιο μειωμένο από αυτό του RAID 3 ή του RAID 4 εξαιτίας της μεγάλης δημοτικότητάς του, και ειδικά αν χρειάζεται ειδικό λογισμικό RAID. Επειδή απαιτείται μεγάλος υπολογισμός ισοτιμίας, το λογισμικό RAID 5 μπορεί να επιβαρύνει το σύστημα σημαντικά.

Το RAID 5 είναι για πολλούς ο ιδανικός συνδυασμός καλής απόδοσης, καλής ανοχής σφαλμάτων και υψηλής χωρητικότητας και επάρκειας αποθηκευτικού χώρου. Χρησιμοποιείται καλύτερα σε συστήματα επεξεργασίας συναλλαγών και για λειτουργίες “γενικού σκοπού”, όπως επίσης και για εφαρμογές βάσεων δεδομένων, και άλλα επιχειρηματικά συστήματα. Για εφαρμογές εγγραφής, είναι καλύτερες επιλογές τα RAID 1 ή RAID 1+0, ενώ η επίδοση του RAID 5 θα αρχίζει να μειώνεται σε περιβάλλοντα με πολλές εγγραφές.

3.7. RAID 6

Στο επίπεδο RAID 6(Block-level striping with *dual* distributed parity) τα blocks δεδομένων και ισοτιμίας διανέμονται σε μια συστοιχία παρόμοια με το RAID 5, με την διαφορά ότι εδώ υπολογίζονται 2 blocks ισοτιμίας για κάθε ζώνη δεδομένων. Χρησιμοποιούνται N+2 δίσκοι, οι οποίοι όλοι περιέχουν δεδομένα και πληροφορίες ισοτιμίας. Συνήθως χρειάζονται το ελάχιστο 4 δίσκοι. Ο στόχος αυτού του διπλασιασμού της πληροφορίας της ισοτιμίας είναι να βελτιωθεί η ανοχή σε σφάλματα. Το RAID 6 μπορεί να χειριστεί σφάλματα οποιονδήποτε δύο δίσκων στη συστοιχία, ενώ όλα τα άλλα επίπεδα RAID μπορούν να χειριστούν το πολύ ένα σφάλμα. Το πλεονέκτημα του RAID 6 είναι το γεγονός ότι παρέχει εκπληκτικά υψηλή διαθεσιμότητα δεδομένων. Προκειμένου να χαθούν δεδομένα πρέπει να παρουσιάσουν βλάβη τρεις δίσκοι εντός του ΜΧΕ(Μέσου Χρόνου Επιδιόρθωσης, Mean Time To Repair, MTTR).

Το RAID 6 επεκτείνει το RAID 5 προσθέτοντας ένα επιπρόσθετο block ισοτιμίας. Επιπλέον χρησιμοποιεί παραλληλισμό σε επίπεδο block με 2 block ισοτιμίας να διαμοιράζονται σ' όλους τους δίσκους. Δεν ήταν από τα πρωτότυπα επίπεδα RAID. Το RAID 6 μπορεί να εννοηθεί ως μια ειδική περίπτωση του κώδικα Reed-Solomon.

Δύο διαφορετικά *σύνδρομα* χρειάζεται να υπολογιστούν για να αποτραπεί η απώλεια δύο οποιονδήποτε δίσκων δεδομένων. Ένα απ' αυτά το **P** μπορεί να είναι ένα απλό XOR των δεδομένων στις ζώνες, όπως στο RAID 5. Ένα δεύτερο, ανεξάρτητο σύνδρομο, το **Q** είναι πιο περίπλοκο.

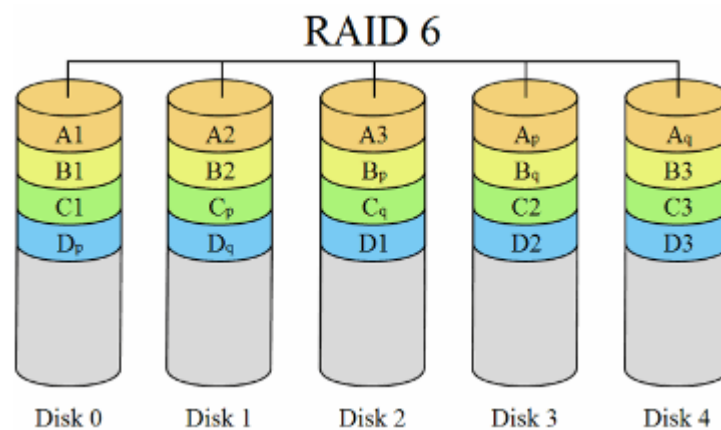
Για n δίσκους δεδομένων D_0, D_1, \dots, D_{n-1}

$$\mathbf{P} = D_0 + D_1 + D_2 + \dots + D_{n-1} \quad (3.7)$$

$$\mathbf{Q} = g^0 \cdot D_0 + g^1 \cdot D_1 + g^2 \cdot D_2 + \dots + g^{n-1} \cdot D_{n-1} \quad (3.8)$$

Στην παραπάνω παράσταση, ο υπολογισμός του **P** είναι απλά ένα XOR για κάθε ζώνη. Στο πεδίο Galois, το + είναι το XOR.

Το **Q** αναφέρεται ως ένας κώδικας Reed-Solomon.



Σχήμα 3.7 Διάγραμμα της εγκατάστασης RAID 6.

3.7.1. Επίδοση(Ταχύτητα)

Το RAID 6 δεν διαθέτει κυρώσεις απόδοσης για λειτουργίες ανάγνωσης, αφού παρέχει εξαιρετική απόδοση σε τυχαίες αναγνώσεις, ειδικά όταν υπάρχει μεγάλο μέγεθος ζώνης. Επίσης παρέχει πολύ καλή απόδοση σε συνεχόμενες αναγνώσεις γενικά καλύτερη όταν υπάρχει μικρό μέγεθος ζώνης. Στις λειτουργίες εγγραφής η απόδοση μειώνεται, ειδικά στις τυχαίες εγγραφές, εξαιτίας της υπερφόρτωσης που συμβαίνει με τους υπολογισμούς ισοτιμίας. Το RAID 6 υφίσταται σημαντική ποινή εγγραφής, διότι κάθε εγγραφή επηρεάζει δύο blocks ισοτιμίας. Έλεγχοι σύγκρισης απόδοσης δείχνουν ότι ο ελεγκτής RAID 6 μπορεί να υποστεί πτώση μεγαλύτερη

του 30% αναφορικά με τη γενική απόδοση εγγραφής σε σύγκριση με την υλοποίηση RAID 5.

Η απόδοση ποικίλει ανάλογα με το πώς υλοποιείται το RAID 6 στην αρχιτεκτονική των κατασκευαστών σε υλικό, σε firmware ή χρησιμοποιώντας firmware και εξειδικευμένους ASICs για υπολογισμούς ισοτιμίας. Μπορεί να είναι τόσο γρήγορο όσο το RAID 5 με ένα λιγότερο δίσκο(ίδιο αριθμό δεδομένων).

3.7.2. Αποδοτικότητα(Πιθανή Απώλεια Χωρητικότητας)

Η απόδοση του RAID 6 σε χωρητικότητα δεν θεωρείται καλύτερη απ' ότι το RAID 5 με ένα δίσκο hot spare ειδικά όταν χρησιμοποιείται σε ένα μικρό αριθμό δίσκων. Όσο όμως οι συστοιχίες γίνονται μεγαλύτερες και έχουν πιο πολλούς δίσκους, η απώλεια στη χωρητικότητα γίνεται λιγότερο σημαντική και η πιθανότητα να χαθούν δεδομένα γίνεται μεγαλύτερη.

Το RAID 6 παρέχει προστασία από απώλεια δεδομένων κατά την διαδικασία της επανάκτησης δεδομένων; όταν δηλαδή χάνεται ένας δεύτερος δίσκος, ή αντιμετωπίζεται κάποιο ελαττωματικό block ανάγνωσης, ή όταν μια τυχαία ανθρώπινη λειτουργία αντικαθιστά τον λάθος δίσκο κατά την προσπάθεια να αντικαταστήσει ένα χαλασμένο δίσκο.

Η χρησιμοποιούμενη χωρητικότητα του RAID 6 είναι

$$(N-2) * S_{min} \quad (3.9)$$

όπου N είναι ο συνολικός αριθμός των δίσκων στη συστοιχία και S_{min} είναι η χωρητικότητα του μικρότερου δίσκου στη συστοιχία.

3.7.3. Υλοποίηση

Σύμφωνα με την Ένωση Βιομηχανιών Δικτυακής Αποθήκευσης (Storage Networking Industry Association - SNIA), ο ορισμός του RAID 6 είναι: "Κάθε μορφή του RAID που μπορεί να συνεχίσει να εκτελεί αιτήσεις αναγνώσεων και εγγραφών σε όλους τους εικονικούς δίσκους της συστοιχίας RAID κατά την παρουσία δύο ταυτόχρονων σφαλμάτων δίσκων. Διάφορες μέθοδοι, συμπεριλαμβανομένου των υπολογισμών διπλών δεδομένων ελέγχου(ισοτιμία και

Reed-Solomon), ορθογώνιος διπλή ισοτιμία δεδομένων και διαγώνια ισοτιμία έχουν χρησιμοποιηθεί για την υλοποίηση RAID Επίπεδου 6".

Θεωρητικά, το RAID 6 είναι ιδανικό για τις ίδιες εφαρμογές που χρησιμοποιεί το RAID 5, αλλά σε περιπτώσεις που απαιτείται επιπρόσθετη ανοχή σε σφάλματα. Στη πράξη όμως, το RAID 6 δεν χρησιμοποιείται συχνά, αφού λίγες εταιρίες είναι πρόθυμες να δαπανήσουν επιπλέον κόστη για να διασφαλιστούν από ένα σχετικά σπάνιο γεγονός - δεν είναι τόσο συχνό δύο δίσκοι να μην λειτουργούν ταυτόχρονα(εκτός αν συμβεί κάτι που θα σταματήσει την λειτουργία ολόκληρης της συστοιχίας οπότε το RAID 6 δεν θα βοηθήσει έτσι κι αλλιώς). Επίσης, η αύξηση του hot swapping και των χαρακτηριστικών αυτόματης επανάκτησης για το RAID 5, παραγκώνισαν το RAID 6, αφού αυτά τα εξελιγμένα χαρακτηριστικά έκαναν μία συστοιχία RAID 5 να μπορεί να επανακτηθεί από ένα σφάλμα σε πολύ μικρό χρονικό διάστημα(όταν χωρίς αυτά το RAID 5 απαιτεί πολύ χρόνο εκτός λειτουργίας, δίνοντας το πλεονέκτημα στο RAID 6). Σε πιο δαπανηρές λύσεις RAID, το RAID 6 χάνει έδαφος από τις εμφωλευμένες λύσεις όπως το RAID 10 που παρέχει κάποιο βαθμό ανοχής πολλαπλών σφαλμάτων ενώ βελτιώνει και την απόδοση.

ΚΕΦΑΛΑΙΟ 4 : ΕΜΦΩΛΕΥΜΕΝΑ ΕΠΙΠΕΔΑ RAID

Τα επίπεδα του εμφωλευμένου RAID, γνωστό και ως υβριδικό RAID, συνδυάζουν δύο ή περισσότερα πρότυπα επίπεδα RAID για να κερδίσουν επίδοση ή/και επιπρόσθετο λογισμικό.

4.1. Εμφωλιασμός

Όταν συνδυάζουμε επίπεδα RAID, ένας τύπος RAID που παρέχει πλεονασμό συνδυάζεται με το RAID 0 για να προάγει την επίδοση. Με αυτές τις διαμορφώσεις είναι προτιμότερο να έχουμε στην κορυφή το RAID 0 και τη συστοιχία πλεονασμού από κάτω, επειδή χρειάζεται να ανακτηθούν λιγότεροι δίσκοι αν σταματήσει να λειτουργεί ένας δίσκος.

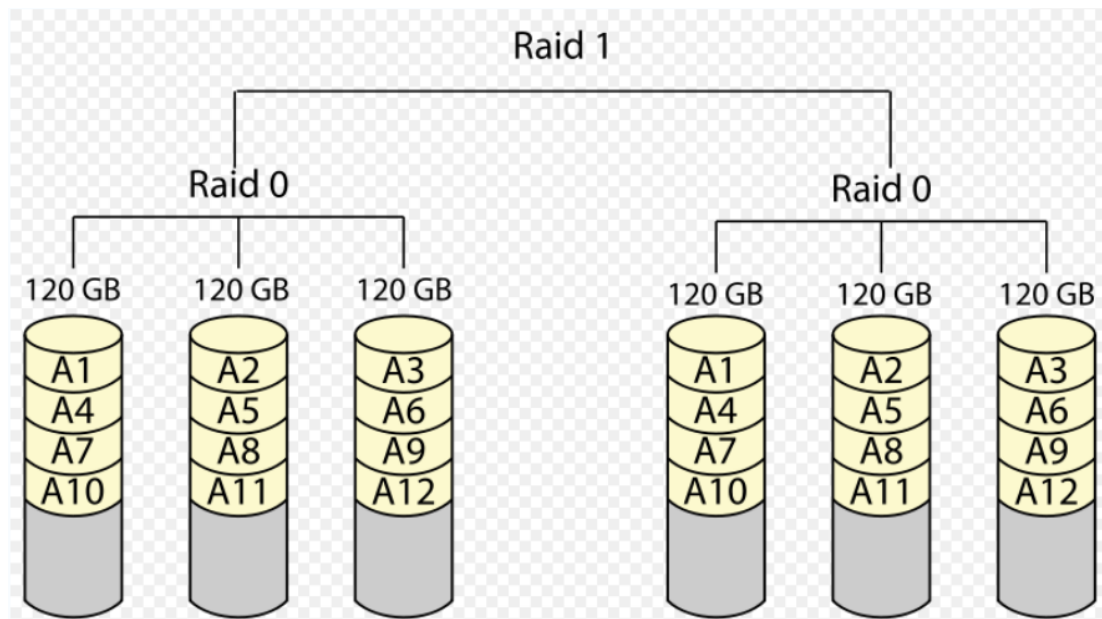
4.2. RAID 0+1

Το RAID 0+1(επίσης ονομάζεται RAID 01, για να μην συγκρίνεται με το RAID 10), είναι ένα επίπεδο RAID που χρησιμοποιείται και για διπλασιασμό και για διαμοιρασμό δεδομένων ανάμεσα στους δίσκους. Η διαφορά μεταξύ του RAID 0+1 και του RAID 1+0 είναι η τοποθεσία του κάθε συστατικού RAID – το RAID 0+1 είναι ένας καθρέπτης ζωνών. Το μέγεθος της συστοιχίας RAID 0+1 μπορεί να υπολογιστεί όπως ακολούθως όπου n είναι ο αριθμός των δίσκων(πρέπει να είναι ζυγός αριθμός) και c είναι η χωρητικότητα του μικρότερου δίσκου στη συστοιχία:

$$\text{Size} = (n * c) \div 2 \quad (4.1)$$

4.2.1. RAID 0+1, έξι δίσκων

Φανταστείτε ένα παράδειγμα RAID 0+1: 6 δίσκοι των 120 GB πρέπει να διαμορφωθούν σε μια συστοιχία RAID 0+1. Παρακάτω είναι ένα παράδειγμα 2 συστοιχιών 360 GB επιπέδου 0 που αντικατοπτρίζονται, δημιουργώντας συνολικό αποθηκευτικό χώρο 360 GB.



Σχήμα 4.1 Διάγραμμα εγκατάστασης RAID 0+1. Τα A1,A2,κτλ κάθε ένα αναπαριστά ένα block δεδομένων; κάθε στήλη αναπαριστά έναν δίσκο.

Ο μέγιστος αποθηκευτικός χώρος είναι 360 GB, που κατανέμεται στις δύο συστοιχίες RAID 0. Το πλεονέκτημα είναι ότι όταν ένας σκληρός δίσκος αποτύχει σ' ένα από τα επίπεδα 0 των συστοιχιών, τα απολεσθέντα δεδομένα μπορούν να μεταφερθούν από την άλλη συστοιχία. Παρόλα αυτά, προσθέτοντας έναν επιπλέον δίσκο σε μια ζώνη απαιτεί να προσθέσει κάποιος ένα δίσκο στις άλλες ζώνες για να ισορροπηστεί η χωρητικότητα στις συστοιχίες.

Δεν είναι τόσο αυτόνομο όσο το RAID 10 και δεν μπορεί να ανεχτεί δύο ταυτόχρονα σφάλματα δίσκων, εκτός αν ο δεύτερος δίσκος είναι απ' την ίδια ζώνη όπως ο πρώτος. Αν ένας μόνο δίσκος αποτύχει, κάθε ένας απ' τους μηχανισμούς στην άλλη ζώνη είναι ένα μοναδικό σημείο σφάλματος. Επίσης, μόλις ο μηχανισμός σφαλμάτων αντικατασταθεί, για να ξαναχτιστούν τα δεδομένα του, πρέπει να λαμβάνουν μέρος στην επανάκτηση όλοι οι δίσκοι στη συστοιχία. Η εξαίρεση σ' αυτό είναι, αν όλοι οι δίσκοι είναι συνδεδεμένοι στον ελεγκτή RAID, τότε ο ελεγκτής μπορεί να επιτύχει την ίδια επανάκτηση σφαλμάτων όπως το RAID 10 αφού μπορεί να έχει πρόσβαση ακόμη και σε "φανταστικούς" δίσκους σε κάθε συστοιχία RAID 0. Αν συγκρίνει κάποιος τα διαγράμματα μεταξύ RAID 0+1 και RAID 1+0 και αγνοήσει τις γραμμές από πάνω απ' τους δίσκους μπορεί να δει ότι το μόνο διαφορετικό είναι ότι οι δίσκοι διανέμονται. Αν ο ελεγκτής έχει μια άμεση

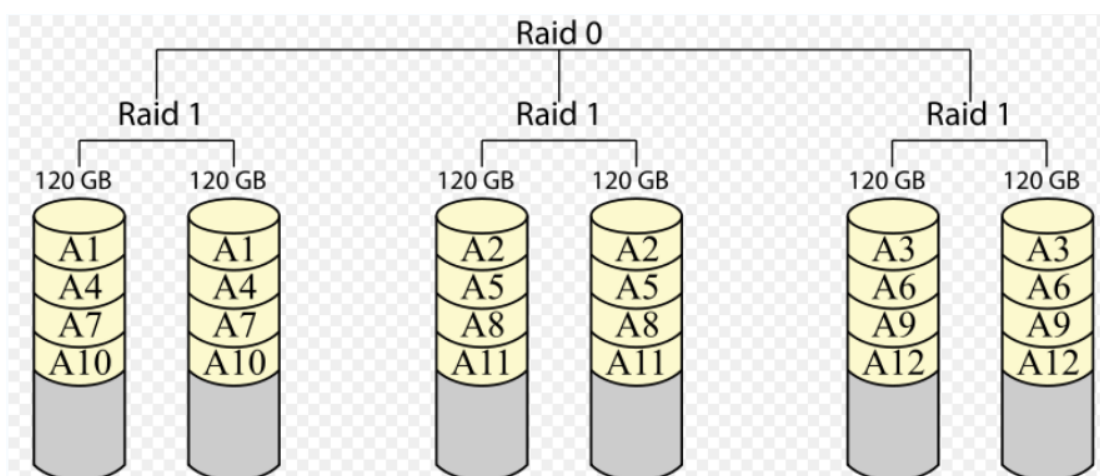
σύνδεση με κάθε δίσκο μπορεί να κάνει το ίδιο. Σ' αυτή τη μοναδική περίπτωση δεν υπάρχει διαφορά ανάμεσα στο RAID 0+1 και RAID 10.

Επιπρόσθετα, οι τεχνολογίες διόρθωσης σφαλμάτων σε επίπεδο bit δεν μπορούν να διατηρηθούν με τις ταχείες ανάπτυξης χωρητικότητας δίσκων, έχοντας ως αποτέλεσμα να υπάρχουν μεγαλύτερα ρίσκα στην αντιμετώπιση σφαλμάτων των μέσων. Στην περίπτωση που ένας εσφαλμένος δίσκος δεν αντικαθίσταται σε μια διαμόρφωση RAID 0+1, ένα μόνο αδιόρθωτο σφάλμα μέσου που συμβαίνει σ' ένα σκληρό δίσκο κάτοπτρο θα έχει ως αποτέλεσμα την απώλεια δεδομένων.

Δεδομένων αυτών των αυξανόμενων ρίσκων με το RAID 0+1 πολλές επιχειρήσεις και περιβάλλοντα επιχειρήσεων έχουν ξεκινήσει να αξιολογούν εγκαταστάσεις RAID πιο ανεκτικές σε σφάλματα και να προσθέτουν βασικούς δίσκους ισοτιμίας. Ανάμεσα στις πιο υποσχόμενες είναι οι υβριδικές προσεγγίσεις όπως το RAID 51 ή το RAID 61.

4.3. RAID 10(RAID 1+0)

Το RAID 1+0, κάποιες φορές ονομάζεται RAID 1&0, ή RAID 10, είναι παρόμοιο με το RAID 0+1 με τη διαφορά ότι τα επίπεδα RAID που χρησιμοποιούνται είναι ανεστραμμένα – το RAID 10 είναι μια ζώνη από κάτοπτρα. Παρακάτω είναι ένα παράδειγμα από 3 συλλογές των 120GB επιπέδου 1 συστοιχιών που ενώνονται μαζί για να σχηματίζουν 360 GB συνολικό χώρο αποθήκευσης:



Σχήμα 4.2 Διάγραμμα εγκατάστασης RAID 10. Τα A1, A2, κ.τ.λ. κάθε ένα αναπαριστά ένα block δεδομένων; κάθε στήλη αναπαριστά έναν δίσκο.

4.3.1. Πλεονασμός και ικανότητα επανάκτησης απώλειας δεδομένων

Όλοι εκτός από έναν δίσκο σε κάθε συστοιχία RAID 1 μπορούν να αποτύχουν χωρίς να βλάψουν τα δεδομένα. Παρόλα αυτά, αν ο εσφαλμένος δίσκος δεν αντικατασταθεί, ο μοναδικός σκληρός δίσκος που λειτουργεί στη συστοιχία θα γίνει ένα μοναδικό σημείο σφάλματος για ολόκληρη τη συστοιχία. Αν αυτός ο σκληρός δίσκος αποτύχει, όλα τα αποθηκευμένα δεδομένα στη συστοιχία θα χαθούν. Όπως και η περίπτωση με το RAID 0+1, αν ένας εσφαλμένος δίσκος δεν αντικατασταθεί στη διαμόρφωση με το RAID 10 τότε ένα μοναδικό μη διορθώσιμο σφάλμα που συμβαίνει στον σκληρό δίσκο-κάτοπτρο θα έχει ως αποτέλεσμα απώλεια δεδομένων. Κάποια διανύσματα RAID 10 επιλύουν αυτό το πρόβλημα χρησιμοποιώντας ένα δίσκο "hot spare", ο οποίος αυτόματα αντικαταστεί και ξαναχτίζει τον εσφαλμένο δίσκο στη συστοιχία. Δεδομένων αυτών των αυξανόμενων ρίσκων με το RAID 0+1 πολλές επιχειρήσεις και περιβάλλοντα επιχειρήσεων έχουν ξεκινήσει να αξιολογούν εγκαταστάσεις RAID πιο ανεκτικές σε σφάλματα και να προσθέτουν βασικούς δίσκους ισοτιμίας. Ανάμεσα στις πιο υποσχόμενες είναι οι υβριδικές προσεγγίσεις όπως το RAID 50 ή το RAID 60.

4.3.2. Επίδοση(Ταχύτητα)

Το RAID 10 είναι συχνά η πρώτη επιλογή σε βάσεις δεδομένων με υψηλό φόρτο, επειδή η έλλειψη του υπολογισμού της ισοτιμίας του δίνει υψηλότερες ταχύτητες εγγραφής.

4.3.3. Απόδοση(Πιθανή απώλεια χωρητικότητας)

Η χρησιμοποιούμενη χωρητικότητα μιας συστοιχίας RAID 10 είναι:

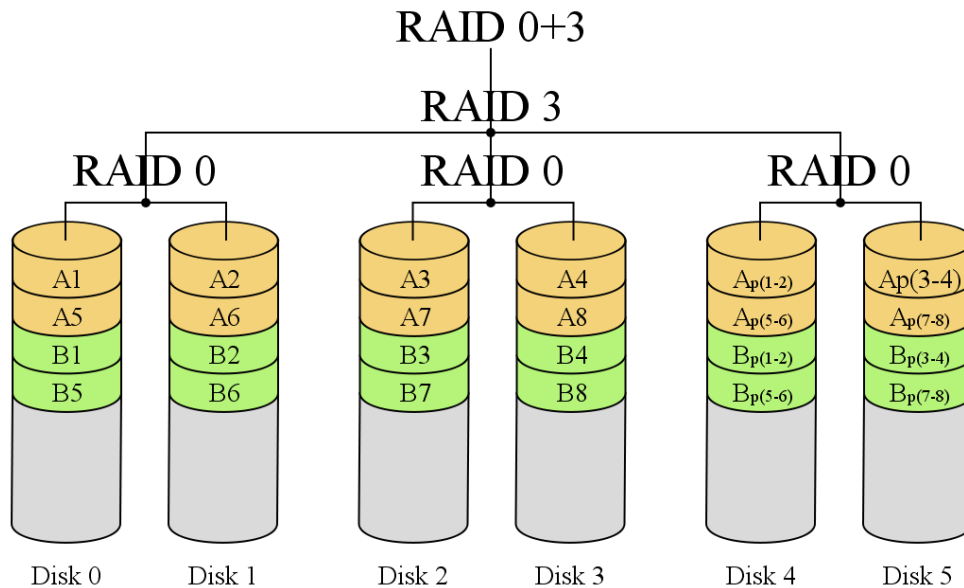
$$(N/2) * S_{min}, \quad (4.2)$$

όπου N είναι ο ολικός αριθμός των δίσκων της συστοιχίας και S_{min} είναι η χωρητικότητα του μικρότερου δίσκου στη συστοιχία.

4.4. RAID 0+3 και 3+0

4.4.1. RAID 0+3

Το επίπεδο RAID 0+3 ή επίπεδο RAID 03 είναι μια συστοιχία με αφιερωμένη ισοτιμία διανεμημένη σε δίσκους. Κάθε block δεδομένων στο επίπεδο RAID 3 σπάει σε συστοιχίες RAID 0 όπου τα μικρότερα κομμάτια διανέμονται στους δίσκους.



Σχήμα 4.3 Διάγραμμα RAID 0+3

4.4.2. RAID 3+0

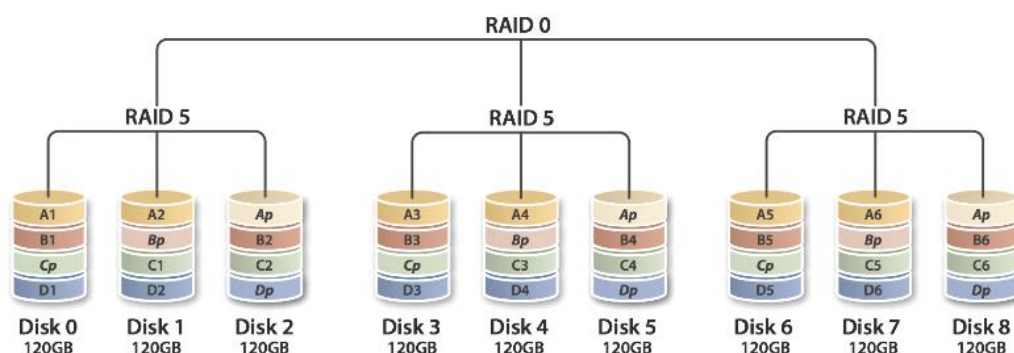
Το επίπεδο 30 είναι επίσης γνωστό ως παραλληλισμός αφιερωμένων ισοτιμιών. Είναι ένας συνδυασμός του RAID επιπέδου 3 και του RAID επιπέδου 0. Το RAID 30 παρέχει υψηλές τιμές μεταφοράς δεδομένων, συνδυασμένες με υψηλή αξιοπιστία δεδομένων. Το RAID 30 υλοποιείται καλύτερα σε 2 συστοιχίες δίσκων RAID 3 με τα δεδομένα να διανέμονται και στις δύο συστοιχίες δίσκων. Το RAID 30 σπάει τα δεδομένα σε μικρότερα blocks, και μετά διανέμει τα blocks των δεδομένων σε κάθε συστοιχία RAID 3. Το RAID 3 σπάει τα δεδομένα σε μικρότερα blocks, υπολογίζοντας την ισοτιμία εκτελώντας μια αποκλειστική OR στα blocks, και μετά γράφει τα blocks σ' όλους τους δίσκους, εκτός από έναν, στη συστοιχία. Το bit ισοτιμίας δημιουργείται χρησιμοποιώντας αποκλειστική OR και μετά γράφεται στον τελευταίο δίσκο κάθε συστοιχίας RAID3. Το μέγεθος κάθε block καθορίζεται από την παράμετρο του μεγέθους ζώνης, το οποίο ορίζεται όταν δημιουργείται η συστοιχία RAID.

Ένας δίσκος από κάθε συστοιχία RAID 3 μπορεί να αποτύχει. Μέχρι να αντικατασταθούν οι εσφαλμένοι δίσκοι, οι άλλοι δίσκοι στις συστοιχίες που εντοπίστηκε το σφάλμα είναι το σημείο του σφάλματος για ολόκληρη τη συστοιχία RAID 30. Με άλλα λόγια, αν ένας απ' αυτούς τους δίσκους αποτύχει, όλα τα δεδομένα στην συστοιχία θα χαθούν. Ο χρόνος που δαπανάται για την ανάκτηση(ανίχνευση και απάντηση στο σφάλμα δίσκου, και η διαδικασία ξαναχτισίματος στον νέο δίσκο) αναπαριστά μια περίοδο ευπάθειας της συστοιχίας RAID.

4.5. RAID 50(RAID 5+0)

Το RAID 50 συνδυάζει τον ευθύ παραλληλισμό επιπέδου block του RAID 0 με την διανεμημένη ισοτιμία του RAID 5. Αυτή είναι μία ισοτιμία RAID 0 που διανέμεται σε στοιχεία RAID 5. Χρειάζεται τουλάχιστον 6 δίσκους. Παρακάτω είναι ένα παράδειγμα όπου 3 συλλογές των 240 GB RAID 5 συνδυάζονται για να φτιάξουν συνολικό χώρο 720 GB.

Τα A1, B1, κ.τ.λ. το καθένα εκπροσωπεί ένα block δεδομένων; κάθε στήλη αντιπροσωπεύει ένα δίσκο; τα Ap, Bp, κ.τ.λ. αντιπροσωπεύουν κάθε ισοτιμία για κάθε χωριστό RAID 5 και μπορούν να αντιπροσωπεύουν διαφορετικές τιμές σε όλο το RAID 5 (δηλαδή, το Ap για το A1 και A2 μπορεί να διαφέρει από Ap για A3 και A4).



Σχήμα 4.4 Εγκατάσταση RAID 50

Ένας δίσκος από κάθε συστοιχία RAID 5 μπορεί να σταματήσει να λειτουργεί χωρίς να υπάρχει απώλεια δεδομένων. Παρόλα αυτά, αν ο ελαττωματικός δίσκος δεν αντικατασταθεί, οι υπόλοιποι δίσκοι στη συστοιχία τότε θα γίνουν σημείο σφάλματος για ολόκληρη τη συστοιχία. Αν ένας απ' αυτούς τους δίσκους αποτύχει, όλα τα δεδομένα σ' ολόκληρη την συστοιχία θα χαθούν. Ο χρόνος που ξοδεύεται για την επανάκτηση(ανίχνευση και εντοπισμός του δίσκου με το σφάλμα, και η διαδικασία επανάκτησης στον καινούργιο δίσκο) αναπαριστά μια περίοδο ευπάθειας για το σετ του RAID.

Η διαμόρφωση των συστοιχιών RAID θα έχει αντίκτυπο σ' ολόκληρη την ανοχή σφαλμάτων. Μια κατασκευή τριών ομάδων RAID 5 των 7 δίσκων έχει υψηλότερη χωρητικότητα και απόδοση χωρητικότητας, αλλά μπορεί να ανεχθεί το πολύ τρία πιθανά σφάλματα δίσκων. Επειδή η αξιοπιστία του συστήματος εξαρτάται από τη γρήγορη αντικατάσταση του χαλασμένου δίσκου ώστε η συστοιχία να ξαναχτιστεί, είναι συχνό να κατασκευάζονται τρία σετ RAID 5 των 6 δίσκων το καθένα με hot spare που μπορεί αμέσως να ξεκινήσει τη διαδικασία επανάκτησης της συστοιχίας σε περίπτωση σφάλματος. Η συστοιχία καταβάλλει μέγιστη προσπάθεια να διαβάσει κάθε bit για να ξαναχτίσει τη συστοιχία ακριβώς τη στιγμή που είναι πιο ευπαθής. Μια κατασκευή των 7 σετ RAID 5 τριών δίσκων μπορεί να χειριστεί έως επτά σφάλματα αλλά έχει λιγότερη χωρητικότητα και απόδοση χωρητικότητας.

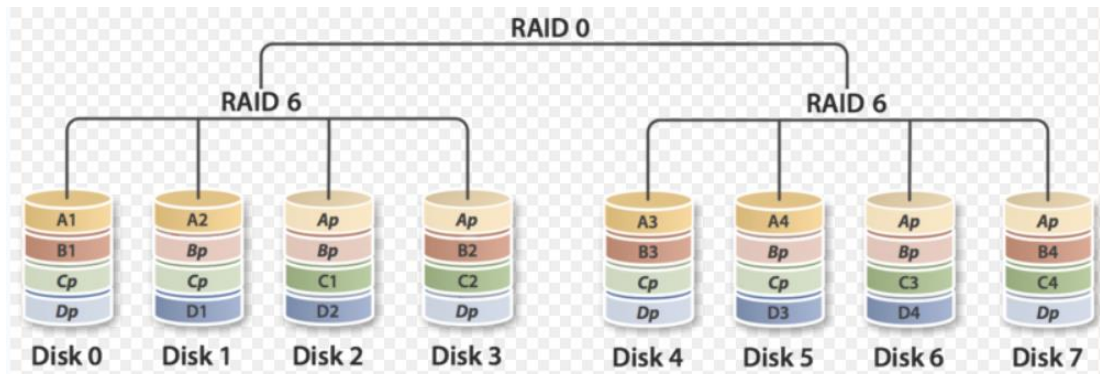
Το RAID 50 βελτιώνει την απόδοση του RAID 5 ειδικά κατά την διάρκεια εγγραφών, και παρέχει καλύτερη ανοχή σφαλμάτων απ' ότι ένα μόνο επίπεδο RAID. Αυτό το επίπεδο συνίσταται για εφαρμογές που απαιτούν υψηλή ανοχή σφαλμάτων, χωρητικότητα και απόδοση τυχαίας τοποθέτησης.

Όσο ο αριθμός των δίσκων σε μια συστοιχία RAID αυξάνεται, και η χωρητικότητα των δίσκων αυξάνεται, αυτό έχει αντίκτυπο στο χρόνο επανάκτησης σφαλμάτων αντίστοιχα όσο αυξάνεται το διάστημα για να ξαναχτιστεί ένα RAID.

4.6. RAID 60(RAID 6+0)

Μία συστοιχία RAID 60 συνδυάζει το διαμοιρασμό σε επίπεδο block του RAID 0 με την διπλή αφιερωμένη ισοτιμία του RAID 6. Οπότε μια συστοιχία RAID 0 διανέμεται σε στοιχεία RAID 6. Απαιτούνται το ελάχιστο 8 δίσκοι.

Παρακάτω είναι ένα παράδειγμα όπου 2 συλλογές από RAID 6 των 240 GB διανέμονται μαζί για να φτιάξουν 480 GB συνολικού αποθηκευτικού χώρου:



Σχήμα 4.5 Εγκατάσταση RAID 60(RAID 6+0). Αποτελείται από 2 σετ των 4 δίσκων το καθένα.

Όπως και στο RAID 6, δύο δίσκοι από κάθε συστοιχία RAID 6 μπορούν να σταματήσουν να λειτουργούν χωρίς απώλεια δεδομένων. Επίσης αν συμβούν σφάλματα όσο ένας δίσκος ξαναχίτζεται σε μια συστοιχία RAID 6, δεν θα έχει ως αποτέλεσμα την απώλεια δεδομένων. Το RAID 60 βελτίωσε την ανοχή σε σφάλματα, οπότε οποιοδήποτε δύο δίσκοι μπορούν να σταματήσουν να λειτουργούν χωρίς απώλεια δεδομένων και περισσότεροι από τέσσερις συνολικά με την προϋπόθεση ότι μόνο 2 από κάθε συστοιχία RAID 6 αποτύχουν.

Ο παραλληλισμός βοηθάει να αυξηθεί η χωρητικότητα και η απόδοση χωρίς να προσθέτουν δίσκοι σε κάθε συστοιχία RAID 6(το οποίο θα μείωνε την διαθεσιμότητα των δεδομένων και θα είχε αντίκτυπο στην απόδοση).Το RAID 60 βελτιώνει την απόδοση του RAID 6. Παρόλο που το RAID 60 είναι ελαφρώς αργότερο από το RAID 50 σε περιπτώσεις εγγράφων εξαιτίας του επιπρόσθετου φόρτου των περισσότερων υπολογισμών της ισοτιμίας και όταν η ασφάλεια των δεδομένων ενδιαφέρει, η πτώση της επίδοσης μπορεί να γίνει αμελητέα.

ΚΕΦΑΛΑΙΟ 5 : ΜΗ - ΠΡΟΤΥΠΑ ΕΠΙΠΕΔΑ RAID

Πολλές διαμορφώσεις εκτός από τα βασικά αριθμητικά επίπεδα RAID είναι πιθανές, και πολλές εταιρείες, οργανισμοί και ομάδες έχουν δημιουργήσει τις δικές τους μη - πρότυπες διαμορφώσεις, σε πολλές περιπτώσεις σχεδιασμένες να καλύπτουν τις ειδικές ανάγκες ενός μικρού εξειδικευμένου γκρουπ. Τα περισσότερα από αυτά τα μη πρότυπα επίπεδα RAID είναι ιδιοσκευάσματα.

Μερικές απ' τις πιο εξέχουσες τροποποιήσεις είναι:

- Η Storage Computer Corporation χρησιμοποιεί το RAID 7, το οποίο προσθέτει κρυφή μνήμη στο RAID 3 και στο RAID 4 για να βελτιώσει την επίδοση I/O.
- Η EMC Corporation προσέφερε το RAID S ως μια εναλλακτική για το RAID 5 στα συστήματα Symmetrix (τα οποία δεν υποστηρίζονται πλέον από τις τελευταίες εκδόσεις του Enguity, το λειτουργικό σύστημα της Symmetrix).
- Το σύστημα αρχείων ZFS, διαθέσιμο στα Solaris, OpenSolaris, FreeBSD και Mac OS X, προσφέρει το RAID-Z, το οποίο επιλύει το πρόβλημα εγγραφών(τρυπών) του RAID 5.
- Το Advanced Data Guarding (ADG) της Hewlett-Packard είναι μια μορφή του RAID 6.
- Το Data ONTAP της NetApp χρησιμοποιεί το RAID-DP(το οποίο αναφέρεται και ως "διπλή", "διπλή", ή "διαγώνια" ισοτιμία), είναι μια μορφή του RAID 6, αλλά σε αντίθεση με πολλές υλοποιήσεις του RAID 6, δεν χρησιμοποιεί αφιερωμένη ισοτιμία όπως στο RAID 5. Αντ' αυτού, χρησιμοποιούνται 2 μοναδικοί δίσκοι ισοτιμίας, με ξεχωριστούς υπολογισμούς ισοτιμίας. Αυτή είναι μια τροποποίηση του RAID 4 με έναν επιπλέον δίσκο ισοτιμίας.
- Η Τριπλή Ισοτιμία της Accusys(RAID TP) υλοποιεί τρεις ανεξάρτητες ισοτιμίες επεκτείνοντας τους αλγορίθμους RAID 6 στους FC-SATA και στους SCSI-SATA ελεγκτές RAID για να ανεχτούν τα σφάλματα τριών δίσκων.

- Το RAID MD RAID 10(RAID 10)υλοποιεί έναν γενικό δίσκο RAID που προεπιλέγει ένα πρότυπο RAID 1+0 με τέσσερις δίσκους, αλλά μπορεί να έχει οποιοδήποτε αριθμό δίσκων. Το MD RAID 10 μπορεί να τρέξει παραλληλισμένο και κατοπτρισμένο μόνο με 2 δίσκους με διάταξη F2(κατοπτρισμός με διαμερισμένες αναγνώσεις, το κανονικό λογισμικό Linux RAID 1 δεν διανέμει τις αναγνώσεις αλλά μπορεί να διαβάσει παράλληλα)
- Το X-RAID της Infrant(τώρα μέλος της Netgear) προσφέρει δυναμική επέκταση ενός τόμου RAID 5 χωρίς να πρέπει να κρατήσει back up ή να αποκαταστήσει το υπάρχων περιεχόμενο. Απλώς προσθέτει μεγάλους δίσκους έναν κάθε φορά, τον αφήνει να συγχρονίσει, και μετά προσθέτει τον επόμενο δίσκο μέχρι να εγκατασταθούν όλοι οι δίσκοι. Η συνολική χωρητικότητα των δίσκων αυξάνεται χωρίς την περίοδο που το σύστημα είναι εκτός λειτουργίας.(Πρέπει να σημειωθεί ότι αυτό είναι δυνατό σε Linux, όταν χρησιμοποιείται το εργαλείο Mdmadm. Επίσης ήταν δυνατό στο EMC Clariion για πολλά χρόνια.)
- Το BeyondRAID, που δημιουργήθηκε από την Data Robotics και χρησιμοποιείται στις σειρές προϊόντων Drobo, υλοποιεί και κατοπτρισμό και διαμερισμό ταυτόχρονα ή ανεξάρτητα ανάλογα με τον δίσκο και το περιεχόμενο δεδομένων. Προσφέρει επεκτασιμότητα χωρίς να χρειαστεί να ξαναπαραμετροποιηθεί. Επίσης προσφέρει την δυνατότητα να ανακατεύει και να ταιριάζει(mix and match) τα μεγέθη των δίσκων και έχει τη δυνατότητα να ξαναταξινομεί τους δίσκους. Υποστηρίζει τα συστήματα αρχείων NTFS, HFS+, FAT 32 και EXT3. Χρησιμοποιεί επίσης λεπτή τοποθέτηση(thin positioning) για να επιτρέπει σε δίσκους άνω των 16 TB την υποστήριξη ανάλογα με το λειτουργικό σύστημα του host.
- Οι σειρές συστοιχιών EVA της Hewlett-Packard υλοποιούν τα vRAID – vRAID-0, vRAID-1, vRAID-5 και vRAID-6. Οι EVA επιτρέπουν στους δίσκους να τοποθετούνται σε ομάδες(που αποκαλούνται Ομάδες Δίσκων) που σχηματίζουν μια πίσινα από block δεδομένων στην κορυφή των οποίων υλοποιείται το επίπεδο RAID. Κάθε ομάδα δίσκου μπορεί να έχει “εικονικούς δίσκους” ή LUNs οποιοδήποτε τύπου vRAID, συμπεριλαμβανομένου των ανακατεμένων τύπων vRAID στη ίδια ομάδα

δίσκων- ένα μοναδικό χαρακτηριστικό. Τα επίπεδα n RAID είναι πολύ όμοια με τα εμφωλευμένα επίπεδα RAID – το n RAID 1 είναι στην πραγματικότητα το RAID 1+0, το n RAID 5 είναι στην πραγματικότητα το RAID 5+0, κ.τ.λ. Επίσης, κάποιοι δίσκοι μπορούν να προστατευθούν εκείνη τη στιγμή, ως εκ τούτου επιτρέπουν την δυναμική απόδοση και την χωρητικότητα να μεγαλώσουν.

- Η IBM έχει υλοποιήσει το RAID 1 E(Επιπέδου 1 Αυξημένο). Με ένα ζυγό αριθμό δίσκων είναι όμοιο με τη συστοιχία RAID 10, αλλά αντίθετα με τη συστοιχία RAID 10, μπορεί επίσης να υλοποιηθεί με μονό αριθμό δίσκων. Σε κάθε περίπτωση, ο τελικός διαθέσιμος αποθηκευτικό χώρος είναι $n/2$. Απαιτεί το ελάχιστο τρεις δίσκους.

ΚΕΦΑΛΑΙΟ 6 : ΥΛΟΠΟΙΗΣΕΙΣ

Η διανομή των δεδομένων σε πολλαπλούς δίσκους μπορεί να επιτευχθεί είτε με ειδικό υλικό ή με λογισμικό. Όταν η διαδικασία γίνεται με λογισμικό, αυτό το λογισμικό μπορεί να είναι μέρος του λειτουργικού συστήματος ή μπορεί να είναι μέρος του firmware και των drivers που παρέχονται με την κάρτα.

6.1. Βασισμένο σε λειτουργικό σύστημα(“Λογισμικό RAID”)

Οι υλοποιήσεις με λογισμικό παρέχονται τώρα από πολλά λειτουργικά συστήματα. Ένα επίπεδο λογισμικού υπάρχει πάνω από τους drivers των συσκευών δίσκων(γενικά βασισμένο σε block) και παρέχει ένα αφαιρετικό επίπεδο μεταξύ των λογικών δίσκων(RAIDs) και των φυσικών δίσκων. Τα πιο συχνά επίπεδα είναι το RAID 0 και το RAID 1 ενώ ακολουθούν τα RAID 1+0, RAID 0+1 και RAID 5.

- Το Mac OS X Server της Apple υποστηρίζει RAID 0,RAID 1,RAID 5 και RAID 1+0.
- Το FreeBSD υποστηρίζει RAID 0, RAID 1, RAID 3 και RAID 5 και όλες τις διαστρωματώσεις των παραπάνω μέσω των ενοτήτων GEOM και ccd. Επίσης υποστηρίζουν το RAID 0, RAID 1, RAID-Z και RAID Z2(όμοια με τα RAID 5 και RAID 6 αντίστοιχα), συν τους εμφωλευμένους συνδυασμούς αυτών μέσω του ZFS.
- Το Linux υποστηρίζει το RAID 0, το RAID 1, το RAID 4, το RAID 5, το RAID 6 και όλες τις διαστρωματώσεις των παραπάνω.
- Τα λειτουργικά συστήματα για servers της Microsoft υποστηρίζουν 3 επίπεδα RAID; το RAID 0, το RAID 1 και το RAID 5. Κάποια από τα λειτουργικά συστήματα της Microsoft για σταθερούς υπολογιστές υποστηρίζουν το RAID όπως το Windows XP Professional, το οποίο υποστηρίζει το RAID επιπέδου 0. Επιπλέον υποστηρίζει τη σύνδεση πολλαπλών δίσκων αλλά μόνο αν χρησιμοποιούνται δυναμικοί δίσκοι και τόμοι. Το Windows XP χρησιμοποιεί RAID 0,1 και 5 με ένα απλό patch αρχείων. Η λειτουργικότητα του RAID στα Windows είναι πιο αργή απ’ ότι το RAID υλικού, αλλά επιτρέπει στη συστοιχία RAID να μετακινείται σε άλλο μηχάνημα χωρίς να υπάρχουν προβλήματα συμβατότητας.

- Το NetBSD υποστηρίζει τα RAID 0, RAID 1, RAID 4 και RAID 5(και οποιοδήποτε εμφωλευμένο συνδυασμό αυτών όπως 1+0) μέσω της υλοποίησης λογισμικού του, που ονομάζεται RAIDFrame.
- Το OpenBSD έχει ως στόχο να υποστηρίζει τα RAID 0, RAID 1, RAID 4 και RAID 5 μέσω της υλοποίησης λογισμικού του, Softraid.
- Το OpenSolaris και το Solaris 10 υποστηρίζουν το RAID 0, το RAID 1 και το RAID 5(ή το όμοιο με το "RAID Z" που βρίσκεται μόνο στο ZFS), και το RAID 6(και κάθε εμφωλευμένο συνδυασμό αυτών όπως το 1+0) μέσω του ZFS και τώρα έχει την ικανότητα να κάνει boot από έναν τόμο ZFS και σε x86 και σε UltraSPARC. Μέσω SVM, το Solaris 10 και οι νεότερες εκδόσεις του υποστηρίζουν το RAID 1 για το σύστημα αρχείων εκκίνησης, και προσθέτει υποστήριξη για RAID 0 και RAID 5(και διάφορους εμφωλευμένους συνδυασμούς) δίσκους δεδομένων.

Το RAID λογισμικού έχει πλεονεκτήματα και μειονεκτήματα σε σύγκριση με το RAID υλικού. Το λογισμικό πρέπει να τρέχει σ' ένα host server ρυθμισμένο να αποθηκεύει, και ο επεξεργαστής του server πρέπει να αφιερώνει το χρόνο επεξεργασίας για να τρέχει στο λογισμικό RAID. Αυτό είναι αμελητέο για το RAID 0 και RAID 1, αλλά μπορεί να γίνει σημαντικό όταν χρησιμοποιούνται συστοιχίες βασισμένες σε ισοτιμία, οπότε είτε υπάρχει πρόσβαση σε πολλές συστοιχίες την ίδια στιγμή ή τρέχουν πολλοί δίσκοι. Επιπλέον όλοι οι δίαυλοι μεταξύ του επεξεργαστή και του ελεγκτή δίσκου πρέπει να μεταφέρουν επιπλέον δεδομένα, που απαιτούνται από το RAID, κάτι που μπορεί να προκαλέσει συμφόρηση.

Μια άλλη ανησυχία με το RAID που βασίζεται σε λειτουργικό σύστημα είναι η διαδικασία εκκίνησης(boot). Μπορεί να είναι δύσκολο ή αδύνατο να εγκαταστήσει κάποιος τη διαδικασία εκκίνησης αφού μπορεί να σφάλει σε άλλο δίσκο αν αποτύχει η συνήθης διαδικασία. Τέτοια συστήματα μπορεί να απαιτούν χειροκίνητη παρέμβαση για να κάνουν μια συσκευή bootable ξανά μετά από ένα σφάλμα. Υπάρχουν εξαιρέσεις σ' αυτό, όπως το LILO bootloader για τα Linux, ο loader για FreeBSD, και κάποιες διαμορφώσεις του GRUB bootloader που μπορούν να καταλάβουν το RAID - 1 και μπορούν να φορτώσουν τον πυρήνα. Αν το BIOS αναγνωρίσει τον πρώτο κατεστραμμένο δίσκο και αναφέρει bootstrapping

στον επόμενο δίσκο, αυτό το σύστημα θα επανέλθει χωρίς παρέμβαση. Ένας ελεγκτής RAID υλικού έχει προγραμματιστεί για να αποφασίσει αν ένας δίσκος είναι χαλασμένος ή αν θα περάσει στον επόμενο δίσκο.

Πολλοί ελεγκτές υλικού RAID είναι εφοδιασμένοι επίσης με μνήμη cache που λειτουργεί με ρεύμα από μπαταρία. Για την ασφάλεια των δεδομένων στα μοντέρνα συστήματα ο χρήστης του λογισμικού RAID μπορεί να χρειάζεται να απενεργοποιήσει τη μνήμη cache των εγγραφών(αλλά κάποιοι δίσκοι έχουν δική τους μπαταρία/πυκνωτές στις μνήμες cache write-back, ένα UPS, και/ή υλοποιούν ατομικότητα με πολλούς τρόπους, κ.τ.λ.). Απενεργοποιώντας τη μνήμη cache εγγραφής έχει ως αντίκτυπο μια ποινή απόδοσης που μπορεί, ανάλογα με το φόρτο εργασίας και το πόσο καλά υποστηρίζεται η ουρά εντολών στο σύστημα δίσκου, να είναι σημαντική. Η μνήμη cache εφοδιασμένη με μπαταρία στον ελεγκτή RAID είναι μια λύση για να υπάρχει ασφαλής write-back cache.

Τέλος το RAID που βασίζεται σε λογισμικό χρησιμοποιεί διαμορφώσεις ειδικά για το λειτουργικό σύστημα ώστε να μην μπορεί γενικά να χρησιμοποιηθεί σε partitions που μοιράζονται σε λειτουργικά συστήματα ως μέρος μιας εγκατάστασης για πολλές διαδικασίες εκκίνησης. Παρόλα αυτά, αυτό επιτρέπει στους δίσκους RAID να μετακινούνται από το ένα PC στο άλλο με λειτουργικό σύστημα ή σύστημα αρχείων ίδιου τύπου, το οποίο μπορεί να είναι πιο δύσκολο όταν χρησιμοποιούμε RAID υλικού,

π.χ. #1: Όταν ένας υπολογιστής χρησιμοποιεί έναν ελεγκτή υλικού RAID από έναν κατασκευαστή και ένας άλλος υπολογιστής χρησιμοποιεί έναν ελεγκτή από διαφορετικό κατασκευαστή, οι δίσκοι δεν μπορούν να αλλαχθούν μεταξύ τους,

π.χ. #2: Αν ο ελεγκτής υλικού σταματήσει να λειτουργεί πριν από τους δίσκους, τα δεδομένα μπορεί να γίνουν μη ανακτήσιμα εκτός αν αποκτηθεί ένας ελεγκτής υλικού του ίδιου τύπου, άσχετα αν το RAID είναι βασισμένο σε firmware ή λογισμικό.

Οι πιο πολλές υλοποιήσεις βασισμένες στο λειτουργικό σύστημα επιτρέπουν στις συστοιχίες RAID να δημιουργηθούν από partitions παρά από φυσικούς δίσκους. Για παράδειγμα ένας διαχειριστής μπορεί να διαιρέσει έναν αριθμό δίσκων σε 2 partitions ανά δίσκο, σε partitions καθρέπτες σ' όλους τους δίσκους και να διανέμει ένα τόμο σ' όλες τις partitions καθρέπτες για να μιμηθεί τη RAID 1E διαμόρφωση

της IBM. Χρησιμοποιώντας partitions μ' αυτό τον τρόπο επιτρέπεται επίσης να αναμιχθούν τα επίπεδα αξιοπιστίας στην ίδια συστοιχία δίσκων. Για παράδειγμα, κάποιος μπορεί να έχει ένα πολύ εύρωστο RAID 1 partition για σημαντικά αρχεία, και ένα λιγότερο εύρωστο RAID 5 ή RAID 0 partition για λιγότερο σημαντικά αρχεία.

Χρησιμοποιώντας 2 partitions στον ίδιο δίσκο στο ίδιο RAID, είναι παρόλα αυτά επικίνδυνο.

π.χ. #1: Έχοντας όλα τα partitions ενός RAID -1 στον ίδιο δίσκο, προφανώς, όλα τα δεδομένα γίνονται μη προσβάσιμα αν ένας μόνο δίσκος αποτύχει.

π.χ. #2: Σε μια συστοιχία RAID 5 που έχει συσταθεί από τέσσερις δίσκους 250+250+250+250 GB, με τον δίσκο των 500 GB να χωρίζεται σε 2 partitions 250 GB, μια αποτυχία αυτού του δίσκου θα καταργήσει τα 2 partitions από τη συστοιχία, προκαλώντας όλα τα δεδομένα που κρατούνται σ' αυτόν να χαθούν.

6.2. Βασισμένο σε Hardware

Οι ελεγκτές υλικού RAID χρησιμοποιούν διαφορετικές διατάξεις δίσκων, οπότε δεν είναι σύνηθες να επεκτείνονται οι ελεγκτές από διαφορετικούς κατασκευαστές. Δεν απαιτούν πόρους επεξεργαστή, το BIOS μπορεί να bootάρει απ' αυτούς, και η ενσωμάτωση σ' ένα δίσκο μπορεί να προσφέρει καλύτερη αντιμετώπιση σφαλμάτων.

Μια υλοποίηση υλικού RAID απαιτεί το λιγότερο ένα ελεγκτή RAID για ειδικές περιπτώσεις. Σ' ένα σύστημα desktop αυτό μπορεί να είναι μια κάρτα επέκτασης PCI, μια κάρτα επέκτασης PCI-e ή να είναι ενσωματωμένο σε μια μητρική κάρτα. Μπορούν να χρησιμοποιηθούν ελεγκτές που υποστηρίζουν τους περισσότερους τύπους δίσκων - IDE/ATA, SATA, SCSI, SSA, Fibre Channel, πολλές φορές ακόμη και ο συνδυασμός τους. Ο ελεγκτής και οι δίσκοι μπορεί να βρίσκονται σ' ένα δικό τους περίβλημα, παρά μέσα στον υπολογιστή. Αυτό το περίβλημα μπορεί να είναι άμεσα συνδεδεμένο με τον υπολογιστή, ή συνδεδεμένο μέσω SAN(Storage Area Network). Ο ελεγκτής υλικού χειρίζεται τη διαχείριση των δίσκων, και εκτελεί οποιουσδήποτε υπολογισμούς ισοτιμίας που απαιτούνται από το επιλεγμένο επίπεδο RAID. Οι περισσότερες υλοποιήσεις υλικού παρέχουν μια

μνήμη cache εγγραφής/ανάγνωσης, η οποία, ανάλογα με το φόρτο εργασίας I/O, θα βελτιώσει την απόδοση. Στα περισσότερα συστήματα η cache εγγραφής είναι non-volatile, δηλαδή μπορεί να διατηρήσει τις πληροφορίες ακόμη και όταν δεν διαρρέεται από ρεύμα (π.χ. προστατευμένη με μπαταρία), οπότε οι εκκρεμείς εγγραφές δεν χάνονται σε περίπτωση πτώσης ρεύματος.

Οι υλοποιήσεις υλικού παρέχουν εγγυημένη απόδοση, δεν προσθέτουν υπερφόρτωση στο τοπικό συγκρότημα CPU και μπορούν να υποστηρίξουν πολλά λειτουργικά συστήματα, αφού ο ελεγκτής απλώς παρουσιάζει ένα λογικό δίσκο στο λειτουργικό σύστημα.

Οι υλοποιήσεις υλικού υποστηρίζουν επίσης hot swapping, επιτρέποντας στους ελαττωματικούς δίσκους να αντικαθίστανται όταν τρέχει το σύστημα.

6.3. Βασισμένοι σε Firmware-driver RAID

Οι συστοιχίες RAID που είναι βασισμένες στο λειτουργικό σύστημα δεν προστατεύουν πάντα την διαδικασία εκκίνησης και γενικά δεν είναι πρακτική αυτή η υλοποίηση σε εκδόσεις desktop των Windows. Οι ελεγκτές υλικού RAID είναι ακριβοί και ιδιόκτητοι. Οι φθηνοί “ελεγκτές RAID” παρουσιάστηκαν χωρίς να περιέχουν ένα chip ελεγκτή RAID, αλλά απλώς ένα πρότυπο chip ελεγκτή δίσκου με ειδικό firmware και drivers. Κατά τη διάρκεια του πρώιμου σταδίου booting, το RAID υλοποιείται από το firmware. Όταν ένας πυρήνας λειτουργικού συστήματος με λειτουργία προστασίας όπως το Linux ή μια μοντέρνα έκδοση των Microsoft Windows φορτώνεται, αναλαμβάνουν οι drivers.

Αυτοί οι ελεγκτές περιγράφονται από τους κατασκευαστές ως ελεγκτές RAID, και γίνεται σπάνια γνωστό στους χρήστες, ότι το φορτίο της διαδικασίας RAID μεταφέρεται στην κεντρική μονάδα επεξεργασίας του host υπολογιστή, όχι στον ίδιο τον ελεγκτή RAID. Οπότε σ’ αυτούς τους ελεγκτές εισάγεται η προαναφερόμενη υπερφόρτωση της CPU, η οποία δεν επηρεάζει τους ελεγκτές υλικού. Οι ελεγκτές firmware συχνά χρησιμοποιούν μόνο συγκεκριμένους τύπους σκληρών δίσκων στις συστοιχίες RAID τους, π.χ. SATA για το Intel Matrix RAID, αφού δεν υπάρχει υποστήριξη ούτε για SCSI ούτε για PATA στις μοντέρνες γέφυρες Intel ICH. Παρόλα αυτά ούτε οι κατασκευαστές μητρικών καρτών υλοποιούν ελεγκτές RAID εκτός της γέφυρας σε κάποιες μητρικές κάρτες. Πριν

από την εισαγωγή των firmware ελεγκτών, ένας "ελεγκτής RAID" υπαινισσόταν ότι ο ελεγκτής έκανε την επεξεργασία, οπότε ο καινούργιος τύπος έγινε γνωστός ως "Ψεύτικο RAID"(fake RAID) παρόλο που το RAID από μόνο του υλοποιείται σωστά. Η Adaptec τους αποκαλεί "Host RAID".

6.4. Network – attached storage

Παρόλο που δεν συνδέεται άμεσα με το RAID, η αποθήκευση σε δίκτυο Network – attached storage(NAS) είναι ένα περίβλημα που περιέχει δίσκους και τον απαραίτητο εξοπλισμό για να τους κάνει διαθέσιμους σ' ένα δίκτυο υπολογιστών, συνήθως Ethernet. Το περίβλημα είναι βασικά ένας αφιερωμένος υπολογιστής, σχεδιασμένος να λειτουργεί στο δίκτυο χωρίς οθόνη ή πληκτρολόγιο. Περιέχει ένα ή παραπάνω δίσκους; πολλαπλοί δίσκοι μπορούν να διαμορφωθούν σε RAID.

6.5. Hot spares

Και το RAID υλικού και το RAID λογισμικού με πλεονασμό μπορούν να υποστηρίξουν τη χρήση των hot spares δίσκων, αν ένας δίσκος φυσικά είναι εγκατεστημένος στη συστοιχία, ο οποίος είναι μη ενεργός μέχρι ένας ενεργός δίσκος να αποτύχει. Τότε το σύστημα αυτόματα αντικαθιστά τον χαλασμένο δίσκο με έναν spare, ξαναχτίζοντας τη συστοιχία συμπεριλαμβανομένου του δίσκου spare. Αυτό μειώνει τον μέσο χρόνο επανάκτησης(MTTR), παρόλο που δεν τον εξαλείφει τελείως. Μεταγενέστερες αποτυχίες στην ίδια συστοιχία RAID με πλεονασμό πριν επανακτηστεί πλήρως η συστοιχία μπορούν να προκαλέσουν απώλεια δεδομένων, αφού η επανάκτηση μπορεί να πάρει πολλές ώρες, ειδικά σε απασχολημένα συστήματα.

Η ταχεία αντικατάσταση των εσφαλμένων δίσκων είναι σημαντική, αφού όλοι οι δίσκοι μιας συστοιχίας θα είχαν το ίδιο μερίδιο χρήσης, και μπορεί να τείνουν να αποτύχουν περίπου στον ίδιο χρόνο παρά με τυχαία σειρά. Το RAID 6 χωρίς το spare δίσκο χρησιμοποιεί τον ίδιο αριθμό δίσκων όπως το RAID 5 με hot spare και προστατεύει τα δεδομένα από ταυτόχρονες αποτυχίες παραπάνω από 2 δίσκων, αλλά απαιτεί έναν πιο αναβαθμισμένο ελεγκτή RAID. Επιπρόσθετα, ένα hot spare μπορεί να διαμοιραστεί σε πολλές συστοιχίες RAID.

ΚΕΦΑΛΑΙΟ 7 : ΟΡΟΙ ΑΞΙΟΠΙΣΤΙΑΣ

7.1.Ρυθμός Αποτυχίας(Failure rate)

Ο ρυθμός αποτυχίας έχει νόημα αν έχει οριστεί η βλάβη. Αν μια βλάβη έχει οριστεί ως απώλεια ενός δίσκου(υπολογιστικός ρυθμός σφάλματος), ο ρυθμός αποτυχίας θα είναι το άθροισμα των ρυθμών αποτυχίας RAID των ξεχωριστών δίσκων. Σ' αυτή τη περίπτωση ο ρυθμός αποτυχίας του RAID θα είναι μεγαλύτερος απ' ότι ο ρυθμός αποτυχίας των επιμέρους δίσκων. Απ' την άλλη πλευρά, αν έχει οριστεί μια αποτυχία ως απώλεια δεδομένων(ρυθμός αποτυχίας συστήματος), τότε ο ρυθμός αποτυχίας του RAID θα είναι λιγότερος απ' ότι αυτός των επιμέρους δίσκων. Πόσο λιγότερος εξαρτάται από τον τύπο του RAID.

7.2. Μέσος χρόνος απώλειας δεδομένων(Mean time to data loss-*MTTDL*)

Μ' αυτό τον όρο εννοούμε τον μέσο χρόνο πριν από την απώλεια δεδομένων σε μια δεδομένη συστοιχία. Ο μέσος χρόνος απώλειας δεδομένων ενός δεδομένου RAID μπορεί να είναι υψηλότερος ή λιγότερος απ' ότι αυτός των περιεχομένων σκληρών δίσκων, ενώ εξαρτάται από τον τύπο του RAID που χρησιμοποιείται. Διάφορες εκθέσεις υποθέτουν ότι οι χρόνοι απώλειας δεδομένων είναι εκθετικά διανεμημένοι. Αυτό σημαίνει ότι το 63,2% όλων των χαμένων δεδομένων θα συμβεί μεταξύ του χρόνου 0 και του MTTDL.

7.3. Μέσος χρόνος επανάκτησης (Mean time to recovery-*MTTR*)

Σε συστοιχίες που περιλαμβάνουν πλεονασμό για αξιοπιστία, αυτός είναι ο χρόνος που ακολουθεί μια αποτυχία για να αποκαταστήσει μια συστοιχία στη φυσιολογική της λειτουργία με ανοχή σφαλμάτων. Αυτό περιλαμβάνει χρόνο για να αντικατασταθεί ο εσφαλμένος δίσκος όπως και χρόνο για να ξαναχτιστεί η συστοιχία(π.χ. να αντιγράψουν τα δεδομένα για πλεονασμό).

7.4. Ρυθμός Μη Ανακτήσιμου σφάλματος bit (Unrecoverable bit error rate-UBE)

Αυτός είναι ο ρυθμός κατά τον οποίο ένας δίσκος θα είναι μη ικανός να επανακτήσει τα δεδομένα μετά την εφαρμογή του κώδικα CRC και πολλαπλών προσπαθειών.

7.5. Αξιοπιστία εγγραφής cache (Write cache reliability)

Κάποια συστήματα RAID χρησιμοποιούν RAM cache εγγραφής για να αυξήσουν την απόδοση. Μια πτώση του ρεύματος μπορεί να προκαλέσει απώλεια δεδομένων εκτός αν αυτό το είδος του buffer του δίσκου συμπληρώνεται με μια μπαταρία για να σιγουρευτεί ότι το buffer έχει αρκετό χρόνο για να γράψει από τη RAM πίσω στο δίσκο.

7.6. Ατομική βλάβη εγγραφής (Atomic write failure)

Γνωστό και ως torn εγγραφές, torn σελίδες, ανολοκλήρωτες εγγραφές, διακοπτόμενες εγγραφές κτλ.

ΚΕΦΑΛΑΙΟ 8 : ΠΡΟΒΛΗΜΑΤΑ ΜΕ ΤΟ RAID

8.1. Συσχετιζόμενα Σφάλματα

Η θεωρία πίσω απ' την διόρθωση λαθών στο RAID υποθέτει ότι τα σφάλματα των δίσκων είναι ανεξάρτητα. Δεδομένων αυτών των υποθέσεων είναι πιθανό να υπολογίσει κάποιος, πόσο συχνά αποτυχαίνουν οι δίσκοι και για να σχηματίσει κάποιος μια συστοιχία η οποία θα κάνει απίθανη την απώλεια δεδομένων.

Στη πράξη, οι δίσκοι συχνά της ίδια ηλικίας, έχουν παρόμοιο προσδόκιμο ζωής. Αφού πολλές αποτυχίες συμβαίνουν εξαιτίας μηχανικών θεμάτων, τα οποία είναι πιο πιθανά σε μεγαλύτερης ηλικίας δίσκους, αυτό παραβιάζει αυτές τις υποθέσεις και οι αποτυχίες συσχετίζονται στατιστικά. Στη πράξη τότε, οι πιθανότητες μιας δεύτερης αποτυχίας πριν να έχει επανακτηθεί η πρώτη δεν είναι τόσο απίθανη όσο μπορεί να υποτεθεί, και η απώλεια δεδομένων μπορεί στη πράξη να συμβεί με συγκεκριμένες τιμές. Οι περισσότεροι σκληροί δίσκοι έχουν μια προσδοκώμενη ζωή των πέντε χρόνων. Παρόλα αυτά, οι χρήστες πρέπει να γνωρίζουν ότι οι δίσκοι χτίζονται με διαφορετικά επίπεδα ευρωστίας, ανάλογα με την επιθυμητή εφαρμογή. Τα κανάλια καλωδίων εμπορικής κλάσης και οι δίσκοι SAS σχεδιάζονται γενικά να αντέχουν την σκληρή χρήση σε μια συστοιχία, αλλά οι δίσκοι κλάσης desktop είναι λιγότερο εύρωστοί και αν χρησιμοποιούνται σε μια συστοιχία μπορούν να μικρύνουν την ζωή τους σημαντικά.

8.2. Ατομικότητα

Αυτή είναι μια βλάβη λιγότερο κατανοητή και αναφέρεται σπάνια για συστήματα με πλεονάζουσα αποθήκευση που δεν χρησιμοποιούν χαρακτηριστικά συναλλαγής. Ο μελετητής βάσεων δεδομένων Jim Gray έγραψε το "Update in Place is a Poison Apple" κατά την διάρκεια των νεότερων ημερών των σχεσιακών βάσεων δεδομένων εμπορευματοποίησης. Παρόλα αυτά αυτή η προειδοποίηση αγνοήθηκε και παραγκωνίστηκε με την έλευση του RAID, το οποίο εξέλαβαν πολλοί μηχανικοί λογισμικού ως επίλυση όλων των προβλημάτων ακεραιότητας και αξιοπιστίας των δεδομένων αποθήκευσης. Πολλά προγράμματα αναβαθμίζουν

ένα αντικείμενο αποθήκευσης εκείνη τη στιγμή; που σημαίνει ότι γράφουν μια νέα έκδοση του αντικειμένου στις διευθύνσεις του ίδιου δίσκου όπως η παλιά έκδοση του αντικειμένου. Όσο το λογισμικό μπορεί να καταγράψει κάποιες πληροφορίες δέλτα κάπου αλλού, η αποθήκευση περιμένει να παρουσιάσει "ατομική σημασιολογία εγγραφών", που σημαίνει ότι η εγγραφή των δεδομένων είτε συμβαίνει ολοκληρωτικά σ' αυτό ή δεν συμβαίνει καθόλου.

Παρόλα αυτά, πολύ λίγα συστήματα αποθήκευσης παρέχουν υποστήριξη για ατομικές εγγραφές, και ακόμα λιγότερα προσδιορίζουν τον ρυθμό των σφαλμάτων παρέχοντας αυτή τη σημασιολογία. Προσέξτε ότι κατά τη διάρκεια της πράξης της εγγραφής ενός αντικειμένου, μια συσκευή αποθήκευσης RAID θα γράφει συνήθως όλα τα επιπλέον αντίγραφα στο αντικείμενο παράλληλα, παρόλο που επικαλυπτόμενες ή εναλλασσόμενες εγγραφές είναι πιο κοινές όταν ένας μόνο επεξεργαστής RAID είναι υπεύθυνος για πολλούς δίσκους. Άρα ένα σφάλμα που συμβαίνει κατά τη διάρκεια μιας διεργασίας εγγραφής μπορεί να αφήσει επιπλέον αντίγραφα σε διαφορετικές καταστάσεις, και επιπλέον μπορεί να αφήσει τα αντίγραφα ούτε στην παλιά κατάσταση ούτε στη καινούργια. Η άγνωστη κατάσταση σφάλματος είναι ότι οι καταγραφές δέλτα βασίζονται στα πρωτότυπα δεδομένα είτε είναι στη παλιά κατάσταση ή στη καινούργια ώστε να ενεργοποιούν την λογική αλλαγή, ενώ κάποια συστήματα αποθήκευσης παρέχουν μια ατομική σημασιολογία εγγραφής στον δίσκο RAID.

Παρόλο που η cache εγγραφής μπορεί να λύσει το πρόβλημα μερικώς είναι εφαρμόσιμη μόνο σε περίπτωση πτώσης του ρεύματος.

Ενώ η υποστήριξη συναλλαγής δεν είναι διεθνώς παρούσα στο υλικό RAID, πολλά λειτουργικά συστήματα περιλαμβάνουν υποστήριξη συναλλαγής για να προστατεύονται ενάντια στην απώλεια δεδομένων κατά τη διάρκεια μιας διακοπής εγγραφής. Η Novell Netware, ξεκινώντας με την έκδοση 3.x., περιλαμβάνει ένα σύστημα εντοπισμού συναλλαγών. Η Microsoft παρουσίασε τον εντοπισμό συναλλαγών μέσω του ημερολογιακού(journaling) χαρακτηριστικού NTFS. Το Ex+4 έχει ημερολογιακό σύστημα αρχείων με έλεγχο προσθέσεων. Μπορεί επίσης να έχει ημερολογιακό σύστημα αρχείων χωρίς τον έλεγχο προσθέσεων, εκτός απ' την επιλογή "προσάρτησης – μόνο" ή ext3cow(copy on write). Το ημερολόγιο στο σύστημα αρχείων WAFL της NetApp δίνει ατομικότητα αλλά δεν ανανεώνει τα δεδομένα εγκαίρως, όπως κάνει το ZFS. Μια εναλλακτική μέθοδος για το

journaling είναι οι soft updates που χρησιμοποιούνται σε μερικές υλοποιήσεις του UFS αντλούμενες από BSD συστήματα. Οι soft updates είναι μια προσέγγιση των υπολογιστικών συστημάτων για να διατηρηθεί η ακεραιότητα ενός δίσκου από σφάλματα ή πτώση του ρεύματος. Λειτουργούν ταξινομώντας τις εγγραφές μεταδεδομένων σε αντίθεση με τον διπλασιασμό τους, που χρησιμοποιούν τα ημερολογιακά συστήματα αρχείων.

Η ατομικότητα μπορεί να παρουσιαστεί ως ένας τομέας σφάλματος ανάγνωσης. Κάποιες υλοποιήσεις RAID προστατεύουν ενάντια σ' αυτή τη κατάσταση βλάβης ξαναφτιάχνοντας τον χάρτη των ελαττωματικών τομέων, χρησιμοποιώντας τα επιπλέον δεδομένα για να αποσπάρουν ένα καλό αντίγραφο δεδομένων, και να ξαναγράψουν τα σωστά δεδομένα στον νέο αντικατεστημένο με νέο χάρτη τομέα. Ο ρυθμός UBE(Μη ανακτήσιμα Bit Σφάλματα) τυπικά προσδιορίζεται στο 1 bit στο 10^{15} για επιχειρησιακής κλάσης δίσκους (SCSI, FC, SAS) και 1 bit στο 10^{14} για δίσκους κλάσης desktop(IDE/ATA/PATA/SATA).

Οι αυξανόμενες χωρητικότητες των δίσκων και τα μεγάλα γκρουπ RAID 5 έχουν οδηγήσει σε αυξανόμενη ανικανότητα να ξαναχτιστεί με επιτυχία ένα γκρουπ RAID μετά από ένα σφάλμα δίσκου εξαιτίας ενός μη ανακτήσιμου τομέα που βρέθηκε στους εναπομείναντες δίσκους. Τα σχήματα διπλής προστασίας όπως το RAID 6 προσπαθούν να διευθυνσιοποιήσουν το θέμα, αλλά υποφέρουν από πολύ υψηλή ποινή εγγραφών.

8.3. Αξιοπιστία Cache Εγγραφής

Το σύστημα δίσκων μπορεί να αναγνωρίσει τη λειτουργία εγγραφής μόλις τα δεδομένα βρίσκονται μέσα στη cache, χωρίς να περιμένει να γραφτούν φυσικά. Αυτό τυπικά συμβαίνει σε παλιά, μη-ημερολογιακά(non-journaling) συστήματα όπως το FAT 32, ή αν επιλεγεί η επιλογή "εγγραφή – πίσω"(writeback) στα Linux/Unix χωρίς καμιά προστασία όπως η επιλογή soft updates(για να προωθηθεί η ταχύτητα I/O όσο ανταλλάσσεται η αξιοπιστία των δεδομένων). Μια διακοπή ρεύματος ή αν "παγώσει" το σύστημα όπως το BSOD(Blue Screen of Death) μπορεί να σημαίνει σημαντική απώλεια οποιουδήποτε δεδομένου που βρίσκεται στην ουρά στη cache.

Συχνά μια μπαταρία προστατεύει την cache εγγραφής, που επιλύει το πρόβλημα. Αν μια εγγραφή αποτύχει εξαιτίας πτώσης του ρεύματος, ο ελεγκτής μπορεί να ολοκληρώσει τις ανολοκλήρωτες εγγραφές μόλις επανακτηθεί. Αυτή η λύση έχει όμως κάποια πιθανά αίτια σφάλματος: η μπαταρία μπορεί να τελειώσει, το ρεύμα μπορεί παραμένει πεσμένο για πολύ ώρα, οι δίσκοι μπορεί να μετακινηθούν σ' άλλο ελεγκτή, ο ελεγκτής ο ίδιος μπορεί να σταματήσει να λειτουργεί. Κάποια συστήματα δίσκων παρέχουν την ικανότητα να τεστάρεται η μπαταρία περιοδικά, παρόλα αυτά αυτή η ενέργεια αφήνει το σύστημα χωρίς μια πλήρως φορτισμένη μπαταρία για πολλές ώρες.

Μια επιπρόσθετη έγνοια υπάρχει για την αξιοπιστία της cache εγγραφής, ειδικά όσων αφορά συσκευές εφοδιασμένες με μια write-back cache – ένα σύστημα cache το οποίο αναφέρει τα δεδομένα ως γραμμένα μόλις εγγραφούν στη cache, αντίθετα με το μη πτητικό(non-volatile) μέσο. Η ασφαλέστερη κατάσταση cache είναι η write-through, η οποία αναφέρει τις συναλλαγές ως εγγεγραμμένες όταν αυτές γραφτούν στο μη-πτητικό μέσο.

8.4. Συμβατότητα Συσκευών

Οι διαμορφώσεις δίσκων σε διαφορετικούς ελεγκτές RAID δεν είναι απαραίτητο να είναι συμβατές, οπότε μπορεί να είναι πιθανό να μην διαβαστεί το RAID σε διαφορετικό υλικό. Συνεπώς ένα σφάλμα υλικού χωρίς δίσκους μπορεί να απαιτεί να χρησιμοποιηθεί όμοιο υλικό, ή back-up δεδομένων, για να επανακτηθούν τα δεδομένα. Το RAID λογισμικού, παρόλα αυτά, όπως υλοποιείται στον πυρήνα Linux, επιλύει αυτό το πρόβλημα, αφού η εγκατάσταση δεν βασίζεται στο υλικό, αλλά τρέχει σε κοινούς ελεγκτές δίσκων. Επιπρόσθετα, το λογισμικό δίσκων RAID 1(και κάποιοι δίσκοι RAID 1, για παράδειγμα ο Silicon Image S744) μπορούν να διαβαστούν σαν κανονικοί δίσκοι, ώστε να μην απαιτείται κανένα σύστημα RAID για να ανακτηθούν τα δεδομένα. Οι εταιρίες επανάκτησης δεδομένων έχουν μεγάλη δυσκολία στην ανάκτηση δεδομένων από δίσκους RAID, με εξαίρεση των δίσκων RAID 1 με συμβατική κατασκευή δεδομένων.

8.5. Επανάκτηση δεδομένων σε περίπτωση εσφαλμένης συστοιχίας

Σε μεγαλύτερες χωρητικότητες δίσκων οι πιθανότητες βλάβης των δίσκων κατά τη διάρκεια της επανάκτησης δεν είναι αμελητέα. Σ' αυτή τη περίπτωση η δυσκολία της εξαγωγής δεδομένων από έναν εσφαλμένο δίσκο πρέπει να ληφθεί υπόψη. Μόνο το RAID 1 αποθηκεύει όλα τα δεδομένα σε κάθε δίσκο. Παρόλο που μπορεί να εξαρτάται απ' τον ελεγκτή, κάποιοι δίσκοι RAID 1 μπορούν να διαβαστούν ως ένας συμβατικός δίσκος. Αυτό σημαίνει ότι ένας "χτυπημένος" δίσκος RAID 1, παρόλο χαλασμένος, μπορεί να επανακτηθεί εύκολα χρησιμοποιώντας ένα πρόγραμμα λογισμικού επανάκτησης ή το CHKDSK. Αν η ζημιά είναι πιο σοβαρή, τα δεδομένα μπορούν να επανακτηθούν από επαγγελματίες ειδικούς σκληρών δίσκων. Το RAID 5 και άλλες διαμεριζόμενες ή παραλληλιζόμενες συστοιχίες παρουσιάζουν σοβαρά στην επανάκτηση δεδομένων σε περίπτωση που η συστοιχία σταματήσει να λειτουργεί.

8.6. Αλγόριθμοι επανάκτησης σφαλμάτων δίσκων

Πολλοί μοντέρνοι δίσκοι έχουν εσωτερικούς αλγορίθμους επανάκτησης σφαλμάτων οι οποίοι μπορεί να χρειαστούν πάνω από ένα λεπτό να επανακτήσουν και να ξαναεντοπίσουν δεδομένα τα οποία ο δίσκος δεν μπορεί να διαβάσει. Πολλοί ελεγκτές RAID θα θέσουν εκτός λειτουργίας έναν μη-ανταποκρινόμενο δίσκο σε 8 δευτερόλεπτα. Αυτό μπορεί να προκαλέσει μια συστοιχία να θέσει εκτός λειτουργίας ένα δίσκο που λειτουργούσε επειδή δεν δόθηκε αρκετός χρόνος να ολοκληρωθεί η εσωτερική διαδικασία επανάκτησης σφαλμάτων, αφήνοντας την υπόλοιπη συστοιχία ευάλωτη. Οι εμπορικοί δίσκοι μειώνουν το χρόνο επανάκτησης λαθών και προλαβαίνουν αυτό το πρόβλημα, αλλά οι desktop δίσκοι μπορεί να είναι αρκετά επικίνδυνοι γι' αυτό το λόγο. Μια λύση είναι γνωστή για τους δίσκους Western Digital. Ένα εργαλείο που λέγεται WDTLER.exe μπορεί να μειώσει τον χρόνο επανάκτησης σφαλμάτων για ένα επιτραπέζιο δίσκο WD ώστε να μην τεθεί εκτός λειτουργίας απ' την συστοιχία γι' αυτό το λόγο. Η λειτουργία ενεργοποιεί το TLER(Time Limited Error Recovery) το οποίο μειώνει τον χρόνο επανάκτησης δεδομένων σε 7 δευτερόλεπτα. Οι εμπορικοί δίσκοι Western Digital από κατασκευής τους έχουν ενεργοποιημένο το

Πτυχιακή εργασία της φοιτήτριας Καρακίζου Βασιλικής

TLER για να αποφευχθεί η πτώση από μια συστοιχία RAID. Παρόμοιες τεχνολογίες χρησιμοποιούνται απ' την Seagate, Samsung, και Hitachi.

ΚΕΦΑΛΑΙΟ 9 : ΣΧΗΜΑΤΙΚΗ ΑΝΑΠΑΡΑΣΤΑΣΗ ΤΗΣ ΤΕΧΝΟΛΟΓΙΑΣ RAID ΜΕ ΤΗΝ ΧΡΗΣΗ ΤΟΥ Macromedia Flash MX 2004

9.1. Γενικά για το Macromedia Flash MX 2004

Το Adobe Flash (παλαιότερα γνωστό και ως Macromedia Flash) είναι μια πλατφόρμα πολυμέσων που αρχικά αποκτήθηκε από την Macromedia και σήμερα αναπτύσσεται και διανέμεται από την Adobe Systems. Από την εισαγωγή του το 1996, το Flash έχει γίνει μια δημοφιλής μέθοδος για να προσθέσει κάποιος κινούμενα σχέδια και γενικά διαδραστικότητα σε ιστοσελίδες. Το Flash χρησιμοποιείται συνήθως για τη δημιουργία animation, διαφημίσεων, και διάφορων συστατικών Flash ιστοσελίδων, για να ενσωματώνει βίντεο σε ιστοσελίδες, και πιο πρόσφατα, για να αναπτύσσει πλούσιες εφαρμογές Διαδικτύου.

Το Flash μπορεί να χειριστεί γραφικά διανύσματος και raster, και υποστηρίζει αμφίδρομη μετάδοση ήχου και βίντεο. Περιέχει μια γλώσσα δέσμης ενεργειών που ονομάζεται ActionScript. Διάφορα προϊόντα λογισμικού, συστήματα και συσκευές μπορούν να δημιουργήσουν ή να εμφανίζουν περιεχόμενο σε Flash, συμπεριλαμβανομένου του Adobe Flash Player, το οποίο είναι διαθέσιμο δωρεάν για τα πιο κοινά προγράμματα περιήγησης στο Web, μερικά κινητά τηλέφωνα και άλλες ηλεκτρονικές συσκευές (με τη χρήση Flash Lite). Το συγγραφικό πρόγραμμα πολυμέσων Adobe Flash Professional χρησιμοποιείται για τη δημιουργία περιεχομένου για την Πλατφόρμα Δέσμευσης Adobe (Adobe Engagement Platform), όπως δικτυακές εφαρμογές, παιχνίδια και ταινίες, και περιεχόμενο για κινητά τηλέφωνα και άλλες φορητές συσκευές. Αρχεία σε μορφή .swf, που παραδοσιακά ονομάζονται "Shockwave Flash" ταινίες, "ταινίες Flash" ή "παιχνίδια Flash", συνήθως έχουν .swf επέκταση αρχείου και μπορούν να αποτελέσουν το αντικείμενο μιας ιστοσελίδας, το οποίο "παίζει" αυστηρά σε ένα αυτόνομο Flash Player, ή να ενσωματώνεται σε ένα Προβολέα (Projector), ο οποίος είναι μια αυτό-εκτελούμενη ταινία Flash (με την επέκταση .exe των Microsoft Windows ή .hqx για Macintosh). Τα αρχεία Flash Video έχουν επέκταση αρχείου .flv και χρησιμοποιούνται είτε από τα εσωτερικά

αρχεία .swf ή παίζουν με έναν player που αναγνωρίζει τα .flv, όπως το VLC, ή το QuickTime και το Windows Media Player αν προστεθούν εξωτερικοί codecs.

9.2. Υλοποίηση της τεχνολογίας RAID με την χρήση του εργαλείου της Macromedia Adobe Flash

Για την δημιουργία της πτυχιακής εργασίας χρησιμοποιήθηκε ως εργαλείο το Adobe Flash MX.

Λόγω του ότι το θέμα της πτυχιακής εργασίας αφορά τα επίπεδα RAID, για να γίνει η σχηματική τους αναπαράσταση χρησιμοποιήθηκε μία σκηνή(scene) για κάθε επίπεδο, καθώς και από μια επιπλέον για χρήσιμες πληροφορίες σχετικά με το κάθε επίπεδο. Οι δύο σκηνές ενώνονται με 2 κουμπιά(buttons). Επίσης υλοποιήθηκαν και κάποια μη βασικά επίπεδα RAID, τα λεγόμενα εμφωλευμένα, που κάθε ένα καταλαμβάνει μια σκηνή. Όλες οι σκηνές ενώνονται με μία βασική, το κύριο μενού.

Το κύριο μενού, είναι μια σκηνή που αποτελείται από 15 κουμπιά(7 τα βασικά επίπεδα RAID και 8 τα εμφωλευμένα) για τη εύκολη πλοήγηση του χρήστη σε κάθε επίπεδο. Επίσης υπάρχει μια επιλογή για αλλαγή της γλώσσας στην Αγγλική καθώς και αντίστροφα στην Ελληνική.

Κάθε επίπεδο RAID αποτελείται από 4 κύρια κουμπιά: 3 κουμπιά επιλογής που αφορούν την ταινία, δηλ. ένα το play που ξεκινάει την αναπαράσταση των δεδομένων, ένα το stop που σταματάει την αναπαράσταση και το reset που γυρίζει την αναπαράσταση στην αρχική κατάσταση. Επιπλέον υπάρχει το back με το οποίο επιστρέφει ο χρήστης στο αρχικό μενού.

Τα σχήματα, δηλαδή, οι αναπαραστάσεις των σκληρών δίσκων, είναι κινούμενα σχέδια(movie clips) στα οποία έχει γραφτεί ο κατάλληλος κώδικας σε Actionscript, τη γλώσσα του Flash.

Ο χρήστης από το κύριο μενού επιλέγει το επίπεδο RAID για το οποίο θέλει να ενημερωθεί. Μόλις φορτωθεί η σελίδα με το αντίστοιχο επίπεδο ο χρήστης βλέπει μία αρχική κατάσταση κατά την οποία υπάρχουν τα τέσσερα κουμπιά δράσης, δύο(ή περισσότεροι) δίσκοι στους οποίους θα καταλήξουν τα δεδομένα και έναν κεντρικό ο οποίος περιέχει τα δεδομένα έτσι όπως καταφθάνουν από την CPU. Ο

χρήστης πατάει το κουμπί play και η αναπαράσταση ξεκινάει. Ανάλογα με το επίπεδο RAID τα δεδομένα διανέμονται με διαφορετικό τρόπο στους δίσκους. Ο χρήστης μπορεί να σταματήσει την ταινία πατώντας το κουμπί stop ή να την ξαναγυρίσει στην αρχική κατάσταση με το κουμπί reset. Μπορεί επίσης να βρει κάποιες βασικές πληροφορίες για το αντίστοιχο επίπεδο πατώντας το κουμπί info. Τέλος μπορεί να επιστρέψει στο βασικό μενού με το κουμπί back.

Όλες οι σκηνές ακολουθούν το ίδιο μοτίβο για την ευκολότερη πλοήγηση του χρήστη.

ΒΙΒΛΙΟΓΡΑΦΙΑ

• Βιβλιογραφική

1. Hennessy, L. John and Patterson, A. David (1996), Computer Architecture: A Quantitative Approach, Morgan Kaufmann Publishers, Inc.
2. Stallings William (2009), Operating Systems Internals and Design Principles, 6th Edition, Prentice Hall
3. Βαφειάδης Αντώνιος (2002), Εισαγωγή στην Επιστήμη των Υπολογιστών

• Ηλεκτρονική

4. <http://en.wikipedia.org/wiki/RAID>
5. http://el.wikipedia.org/wiki/Μέσο_αποθήκευσης_δεδομένων
6. http://el.wikipedia.org/wiki/Σκληρός_δίσκος
7. <http://en.wikipedia.org/wiki/CD-ROM>
8. <http://en.wikipedia.org/wiki/DVD-ROM>
9. http://it.toolbox.com/wiki/index.php/Automated_Tape_Libraries
10. <http://www.lascon.co.uk/d008005.htm>
11. <http://www.fujitsu.com/global/services/computing/storage/system/glossary/raid/>
12. <http://www.pcguides.com/ref/hdd/perf/raid/levels/index.htm>
13. <http://www.raidrecoveryguide.com/levels1.html>

Πτυχιακή εργασία της φοιτήτριας Καρακίζου Βασιλικής

14. http://www.it.uom.gr/project/mycomputer/storage/magnetic/harddisk/p_reliab.html

15. <http://www.ecs.umass.edu/ece/koren/architecture/Raid/basicRAID.html>

16. <http://www.pcguides.com/ref/hdd/perf/raid/index.htm>