**ZX ZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZXZZXZXZXZXZXZXZXZXZXZXZXZXZXDepartment of Electrical Engineering**

|  |  |
| --- | --- |
| **Faculty Member:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** | **Dated: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** |
|  |  |
| **Course/Section:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** | **Semester: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** |
|  |  |

**Computer Organization and**

**Achitecture (EE321)**

**Lab # RAID**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Name** | **Reg. no.** | **Report Marks / 10** | **Viva Marks / 5** | **Total/15** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**Lab # 11 RAID**

**Objectives:** The aim of this lab is to develop understanding of following component of Computer Architecture.

* Reviews the basic reliability provided by a [block-interleaved parity disk array](http://www.ecs.umass.edu/ece/koren/architecture/Raid/basicRAID.html#RAID Level 5)
* Lists and discusses three factors that can determine the potential reliability of disk arrays.

**Lab Instructions**

* This lab activity comprises two parts, namely Lab tasks, and Post-Lab Viva session.
* Each group to upload completed lab report on LMS for grading.
* The students will start lab task and demonstrate each lab task separately for step-wise evaluation
* There are related questions in this activity give complete answers. Also provide complete code and results.

1. **RAID Reliability**

Reliability of any I/O system has become as important as its performance and cost. This part of the tutorial: 

* Reviews the basic reliability provided by a [block-interleaved parity disk array](http://www.ecs.umass.edu/ece/koren/architecture/Raid/basicRAID.html#RAID Level 5)
* Lists and discusses three factors that can determine the potential reliability of disk arrays.

Redundancy in disk arrays is motivated by the need to fight disk failures. Two key factors MTTF (Mean-Time-to-Failure) and MTTR (Mean-Time-to-Repair) are of primary concern in estimating the reliability of any disk. Following are some formulae for the mean time between failures:

**RAID level 5**

MTTF(disk) 2

------------------

N\*(G-1)\*MTTR(disk)

**Disk array with two redundant disk per parity group (eg: P+Q redundancy)**

MTTF(disk) 3

-------------------------

N\*(G-1)\*(G-2)\* (MTTR(disk) 2 )

N - total number of disks in the system   
G - number of disks in the parity group

**Factors affecting Reliability**

Three factors that can dramatically affect the reliability of disk arrays are:

* System crashes
* Uncorrectable bit-errors
* Correlated disk failures

**System Crashes**

System crash refers to any event such as a power failure, operator error, hardware breakdown, or software crash that can interrupt an I/O operation to a disk array.

Such crashes can interrupt write operations, resulting in states where the data is updated and the parity is not updated or vice versa. In either case, parity is inconsistent and cannot be used in the event of a disk failure. Techniques such as redundant hardware and power supplies can be applied to make such crashes less frequent.

System crashes can cause parity inconsistencies in both bit-interleaved and block-interleaved disk arrays, but the problem is of practical concern only in block-interleaved disk arrays.For, reliability purposes, system crashes in block-interleaved disk arrays are similar to disk failures in that they may result in the loss of the correct parity for stripes that were modified during the crash.

**Uncorrectable bit-errors**

Most uncorrectable bit-errors are generated because data is incorrectly written or gradually damaged as the magnetic media ages. These errors are detected only when we attempt to read the data.

Our interpretation of uncorrectable bit error rates is that they represent the rate at which errors are detected during reads from the disk during the normal operation of the disk drive.

One approach that can be used with or without redundancy is to try to protect against bit errors by predicting when a disk is about to fail. VAXsimPLUS, a product from DEC, monitors the warnings issued by disks and notifies an operator when it feels the disk is about to fail.

**Correlated disk failures**

**Causes:** Common environmental and manufacturing factors.

For example, an accident might sharply increase the failure rate for all disks in a disk array for a short period of time. In general, power surges, power failures and simply switching the disks on and off can place stress on the electrical components of all affected disks. Disks also share common support hardware; when this hardware fails, it can lead to multiple, simultaneous disk failures.

Disks are generally more likely to fail either very early or very late in their lifetimes.

*Early failures* are frequently caused by transient defects which may not have been detected during the manufacturer's burn-in process.*Late failures* occur when a disk **M** wears out.  Correlated disk failures greatly reduce the reliability of disk arrays by making it much more likely that an initial disk failure will be closely followed by additional disk failures before the failed disk can be reconstructed.

**Mean-Time-To-Data-Loss(MTTDL)**

Following are some formulae to calculate the mean-time-to-data-loss ([MTTDL](http://www.ecs.umass.edu/ece/koren/architecture/Raid/glossary.html#MTTDL)). In a block-interleaved parity-protected disk array, data loss is possible through the following three common ways:

* double disk failures
* system crash followed by a disk failure
* disk failure followed by an uncorrectable bit error during reconstruction

The above three failure modes are the hardest failure combinations, in that we, currently, don't have any techniques to protect against them without sacrificing performance.

RAID Level 5

|  |  |
| --- | --- |
| Double Disk Failure | MTTF(disk1) \* MTTF(disk2)  -----------------------  N \* (G-1) \* MTTR(disk) |
| System Crash + Disk Failure | MTTF(system) \* MTTF(disk)  -----------------------  N \* MTTR(system) |
| Disk Failure + Bit Error | MTTF(disk)  -----------------------  N \* (1 - ( p(disk)) (G-1) ) |
| Software RAID | harmonic sum of the above |
| Hardware RAID | harmonic sum of above excluding system crash + disk failure |
|  | |

### P+Q disk Array

|  |  |
| --- | --- |
| Triple Disk Failures | MTTF(disk) \* (MTTF(disk2) \* MTTF(disk3)  ----------------------------------  N \* (G-1) \* (G-2) \* MTTR(disk) 2 |
| System Crash + Disk Failure | MTTF(system) \* MTTF(disk)  --------------------------  N \* MTTR(system) |
| Double disk failure + Bit error | MTTF(disk) \* MTTF(disk2)  ----------------------------------  N\*(G-1)\*(1-(p(disk)) (G-2) )\* MTTR(disk) |
| Software RAID | harmonic sum of the above |
| Hardware RAID | harmonic sum excluding system crash +disk failure |
|  | |

p(disk) = The probability of reading all sectors on a disk (derived from disk size, sector size, and BER)

**Lab Task 1**

**Online Simulator**: http://www.ecs.umass.edu/ece/koren/architecture/Raid/reliability\_tool.html

Consider the following system with disks 1,2 and 3.

-- 10 Disk 1s with 1,000,000-hour MTTF

-- 19 Disk 2s with 450,000-hour MTTF

-- 1 Disk 3s with 700,000-hour MTTF

-- MTTR for single disk is 8000 hours and 13,000 hours for the system

-- p(disk) is 1

(a) Find the MTTF of the system.

18,639

(b) Use the MTTF found above, to find the Reliability and MTTDL. Which RAID level is best for each failure characteristic

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **MTTDL** | **RAID5** | | **P+Q** | |
| Reliability | MTTDL | Reliability | MTTDL |
| Double/Triple Disk Failure | 143678 | 64655 | 641420 | 202047 |
| System Crash + Disk Failure | 143678 | 47792 | 641420 | 47792 |
| Single/Double Failure + Bit Error | 143678 | Infinity | 641420 | Infinity |

**Lab Task 2**

1. **RAID Redundant Array of Inexpensive Disks**

The need for RAID can be summarized in two points given below. The two keywords are Redundant and Array.

* An array of multiple disks accessed in parallel will give greater throughput than a single disk.
* Redundant data on multiple disks provides fault tolerance.

Provided that the RAID hardware and software perform true parallel accesses on multiple drives, there will be a performance improvement over a single disk.

With a single hard disk, you cannot protect yourself against the costs of a disk failure, the time required to obtain and install a replacement disk, reinstall the operating system, restore files from backup tapes, and repeat all the data entry performed since the last backup was made.

With multiple disks and a suitable redundancy scheme, your system can stay up and running when a disk fails, and even while the replacement disk is being installed and its data restored. To create an optimal cost-effective RAID configuration, we need to simultaneously achieve the following goals: 

* Maximize the number of disks being accessed in parallel.
* Minimize the amount of disk space being used for redundant data.
* Minimize the overhead required to achieve the above goals.

**Online Simulator**: <http://www.ecs.umass.edu/ece/koren/architecture/Raid/raidiator.html>

1. With 250 disks in the disk array, what is the highest and lowest storage efficiency? Which RAID architecture offers the best and the worst efficiency?
2. 50
3. 96.8
4. 99.6
5. 99.6
6. 99.6

(b) Consider a system with 5 blocks of data and one parity disk. Let a block read be 1 cycle and block write be 2 cycles. Let logical XOR take 1 cycle. When a block is being written into the system, list the following for RAID3 and RAID4 systems.

(i) What are the operations that will be performed?

(ii) What is the total cycle time for the operation?

1. What is the main difference between RAID4 and RAID5? What problem does the RAID5 solve?

**Lab Task 3**

**Online Simulator**: <http://www.coastalworks.com/raid/raid.html>

 1.  First begin simulations by choosing RAID 0 and running the program without changing the sectors per actual disc, sequence length, or failure probabilities of each disc.

2.      Leave show trace checked and start simulation.  Below is a screenshot of the initial simulation without changes made.

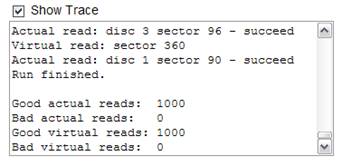


Figure 2 Initial RAID 0 Simulation

3.      Next simulate a failure probability per disk of 0.1, keeping all other factors the same for RAID 0.

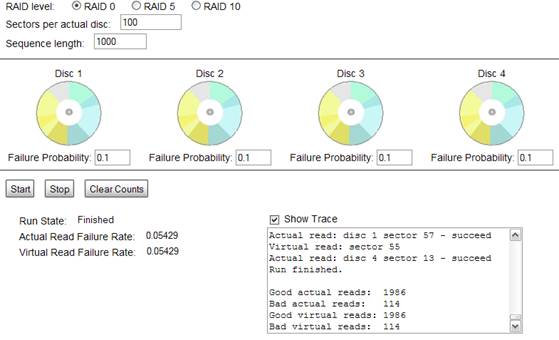
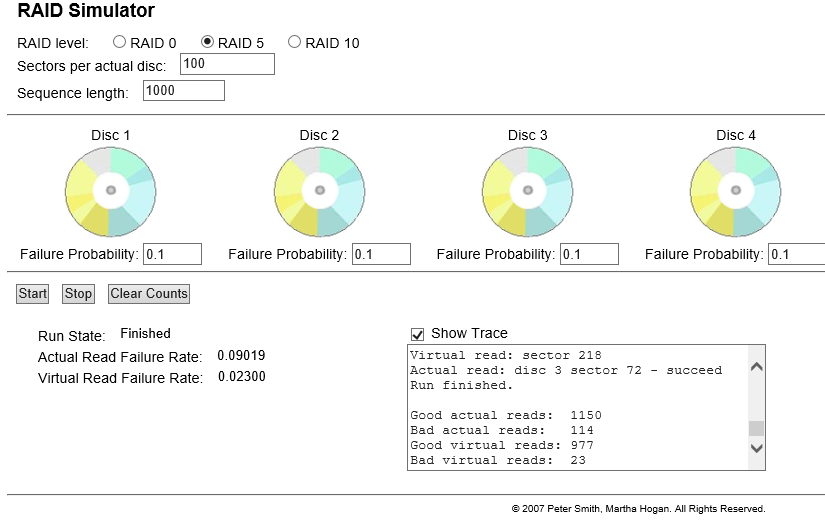
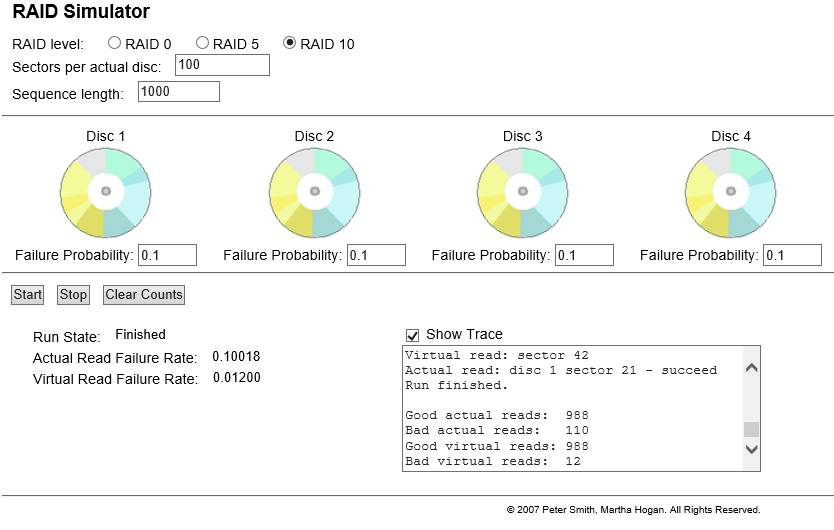


Figure 3 RAID 0 0.1 Failure Probability

4.      Explain the results for the bad virtual/actual reads using the failure rates?

5.      Repeat this simulation for both RAID 5 and RAID 10 and take screen shots of the trace results.





6.      Explain the results using the data provided for RAID 5 and 10.

7.      Continue simulating the different RAID levels changing the Failure Probabilities of each of the 4 Discs.

8.      What do the different RAID levels mean?

9.  If you continue to simulate the same RAID level with all the same options, what happens? Why?