CMSC611: Advanced Computer Architecture
Homework 6 Solutions

Question 1: (50 points)

a) The layout of blocks across disks of the RAID 5 system is like:

```
0 1 2 3 4 5 6 P0
8 9 10 11 12 13 P1
16 17 18 19 20 21 P2
24 25 26 27 28 29 P3
32 33 34 35 36 37 P4
40 41 42 43 44 45 P5
48 49 50 51 52 53 P6
56 57 58 59 60 61 P7
64 65 66 67 68 69 P8
```

b) In order to perform reconstruction, the RAID 5 system needs to read the data or parity blocks from 7 working disks and write the data or parity block to the fail disk. If failures are independent for the eight disks, the expected time until a reconstruction can be estimated based on the MTTF for each disk:

\[
\text{Expected Time} = \frac{\text{MTTF}}{\text{Number of disks}} = \frac{3 \times 10^6 \text{ hours}}{8} = 375,000 \text{ hours}
\]

c) With the peak sequential read rate, each disk can read 400MB/sec. Consider all reads are parallel,

\[
\text{Read Time} = \frac{\text{Disk capacity}}{\text{Read rate}} = \frac{100\text{GB}}{400\text{MB/sec}} = 250\text{s}
\]

With the peak sequential write rate, each disk can write 200MB/sec, so

\[
\text{Write Time} = \frac{\text{Disk capacity}}{\text{Write rate}} = \frac{100\text{GB}}{200\text{MB/sec}} = 500\text{s}
\]

Write can only occur after all reads are completed, so the minimum offline reconstruction time is

\[
\text{Reconstruction}_{\text{offline}} = \text{Read time} + \text{Write time} = 250 + 500 = 750\text{s}
\]

d) Considering the 40MB/sec limitation, the rate of each disk for the read or write is

\[
\frac{\text{Read/Write Rate}}{\text{Number of disks}} = \frac{40\text{MB/sec}}{8} = 5\text{ MB/sec}
\]
If all read and write operations can be parallel, the online reconstruction time is

\[
\text{Reconstruction}_{\text{online}} = \frac{\text{Disk capacity}}{\text{Read/Write rate}} = \frac{100\text{GB}}{5\text{MB/sec}} = 20,000\text{s}
\]

**Question 2:** (50 points)

a) The MESIF state transition diagram is shown as the following:

- **PrRd/-, PrWr/-**
- **PrRd/BusRd**
- **PrRd/BusRdX**
- **BusRdX/Flush**
- **BusRd/Flush**
- **BusRdX/-**
- **PrRd/BusRd**
- **BusRdX/Flush**
- **PrRd/-, BusRd/-**
- **PrRd/-, PrWr/-**

**PrRd:** Read request from processor
**PrWr:** Write request from processor
**BusRd:** Read request from the bus without intent to modify
**BusRdX:** Read request from the bus with intent to modify

The transitions are labeled as “action observed/action performed”.

<table>
<thead>
<tr>
<th>Processor 1</th>
<th>Processor 2</th>
<th>MESIF State</th>
<th>Transition</th>
<th>Bus Message</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read address A</td>
<td>E</td>
<td>PrRd</td>
<td>BusRd</td>
<td></td>
</tr>
<tr>
<td>Read address A</td>
<td>F</td>
<td>PrRd</td>
<td>BusRd</td>
<td></td>
</tr>
<tr>
<td>Read address A</td>
<td>S</td>
<td>BusRd</td>
<td>Flush</td>
<td></td>
</tr>
<tr>
<td>Write to address A</td>
<td>M</td>
<td>PrWr</td>
<td>BusRdX</td>
<td></td>
</tr>
<tr>
<td>Read address A</td>
<td>I</td>
<td>BusRdX</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Read address A</td>
<td>F</td>
<td>PrRd</td>
<td>BusRd</td>
<td></td>
</tr>
<tr>
<td>Read address B</td>
<td>I(A)</td>
<td>PrReplace(A)</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Read address A</td>
<td>F(B)/E(B)</td>
<td>PrRd(B)</td>
<td>BusRd(B)</td>
<td></td>
</tr>
<tr>
<td>Write to address A</td>
<td>S</td>
<td>BusRd</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Write to address A</td>
<td>M</td>
<td>PrWr</td>
<td>BusRdX</td>
<td></td>
</tr>
</tbody>
</table>