This exam is closed-book and mostly closed-notes. You may, however, use an 8 1/2 by 11 sheet of paper with hand-written notes for reference. (Both sides of the sheet are OK.)

Please write your name only on this page. Do not turn the page until instructed, in order that everyone may have the same time. Then, be sure to look at all problems before deciding which one to do first. Some problems are easier than others, so plan your time accordingly. You have 120 minutes to work.

Write the answer to each problem on the page on which that problem appears. You may also request additional paper, which should be labeled with your test number and the problem number.

If you are stuck, ask for help. At worst, I’ll offer to sell you a hint for some points.

Name: ______________________________

<table>
<thead>
<tr>
<th>Problem</th>
<th>Page</th>
<th>Possible</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>80</strong></td>
<td></td>
</tr>
</tbody>
</table>
1. [16 Points] Consider a direct-mapped cache with a total capacity of 32 words and a block size of four words. Assume the cache is initially empty and that we are using word addresses (as in the homework), so that words have addresses 0, 1, 2, etc. For the following sequence of addresses, indicate which are misses and which are hits. Also show which addresses (if any) are contained in each location of the cache at the end. (To receive partial credit for an incorrect answer, you should also show your work, including in particular how the cache contents evolves, e.g., by crossing out addresses that are replaced.)

0 11 50 0 35 27 0 10
2. [16 Points] Three friends each tries improving a system design; the three friends names are Cathy, Alice, and Betty. The existing system has a 64KB direct-mapped cache with 4-byte blocks. Cathy tries increasing the cache size to 128KB. Alice instead tries increasing the associativity to 4-way. Betty instead tries increasing the block size to 64 bytes.

(a) When the three friends measure the performance of their respective systems,
   i. one finds that the number of compulsory misses has gone down,
   ii. one finds that the number of capacity misses has gone down, and
   iii. one finds that the number of conflict misses has gone down.

Which one saw each effect? Write the initial (C, A, or B) next to each of the above.

(b) Regarding space overhead,
   i. One friend reports that her change requires more tags be stored than in the original, but each one is now slightly smaller.
   ii. One friend reports that her change requires the same number of tags as the original, but each is now slightly larger.
   iii. One friend reports that her design requires fewer tags than the original.

Who made each of these reports? Again, write the initial (C, A, or B) next to each of the above.
3. [ **16 Points** ] Suppose we want to design a system to handle 10,000 disk read requests per second. Each request is for a randomly located 8 KB block on disk drives with the following specs:

- 12,000 RPM rotational speed
- 7.3 ms average seek time
- 40 MB/s transfer rate.
- interface is via a Ultra/160 bus that can handle 160 MB/s and up to 31 drives

Assume there is no controller overhead or time spent running software in the processor. Assume also that the workload is spread evenly over the disk drives. Assume that there are 1000 K in an M. How many drives and buses are needed? Justify your answers. In calculating the necessary number of buses, be sure to show that you have considered both limitations of the buses.
4. [ **16 Points** ] Explain the differences between an Ethernet hub, an Ethernet switch, and a router.
5. [16 Points] What is cache coherency? Explain briefly the technique used to maintain cache coherency in typical bus-based multiprocessors. Explain briefly the technique used to maintain cache coherency in typical network-based multiprocessors.