Digital Systems
EEE4084F

FINAL EXAM

15 June 2018

Out of 120 marks

SOLUTIONS!!!

Section 1: Short Answers [50 marks]

Q1. [12 marks]

(a) …Briefly explain the concept of bisectional bandwidth and why it is useful in judging the network performance design of a high performance computer system (2011 exam 1.2)

Answer: The bisection bandwidth of a network is the bandwidth at which data may be simultaneously communicated between the two halves of the network; the value is calculated as the product of the bisection width (the number of links between the bisection) and the bandwidth of each link. It is useful in establishing whether the interconnections between arbitrary processing nodes, that may need to communicate, have an adequate bandwidth in the system. [5 marks]

(b) Calculate the bisection bandwidth for the network shown in Figure 1, assuming each link is 1Gbps…

Answer: This is a rectangular mesh structure; it is not a square with both size the same size so you can't simply use sqrt(P) with P=18 and getting a bandwidth of min 4Gbps. Rather, you need to consider the two cuts: left/right P3,P9,P15 | P4,P10,P16 or up/down P7,P8… P12 | P13,P14…P18. As you can see in the diagram if you divide left/right, the max bandwidth is 3 between the two 3x3 network sections. If you divide up/down then bandwidth is doubled to 6. Generally the worst case scenario is applied, so it would be the former case of the left/right cut that has bisection bandwidth of 3Gbps. [5 marks]

(c) i. What is the maximum speed P1 can continuously stream data to P18, assuming a 0us (i.e. zero) time cost for each node…

Answer: This would simply be 1GBps, i.e. the links would be: [1 marks]

P1-P2-P3-P4-P5-P6-P12-P18 and P18-P17-P16-P15-P14-P13-P7-P1.

(c) ii.What is the maximum speed P1 can stream to P18 and P6 can simultaneously stream data to P13…

Answer: This is also 1GBps. You might have thinking from (i) in terms of L-shapes going along the perimeter (which would lead to 500MBps), but using some of the internal nodes you can easily see that it is 1GBps due to the zero routing delay. [1 mark]

Q1.2. [12 marks]

Q1.2(a)

i. The main difference between a FPGA and a PLA is the: the architecture (how the system is configured and programmed), the number of logic elements available, and the programming speed. The FPGA has a more complex architecture, supports more complex designs, usually many types of logic elements. [2]

ii. There is usually a particular programing sequence needed for an FPGA. In particular, if a FPGA board needs to start up without being programmed from a host (e.g. attached PC), there needs to be some way to program the FPGA. This is where a configuration architecture, utilizing a statemachine implemented using a PLA or CPLD, is used in order to read the FPGA program from non-volatile memory (e.g. a EEPROM chip) and to program the FPGA. Furthermore, the PLD/CPLD may also include logic to support programming from a host, i.e. to receive a program sent from the host into a then exercise the necessary programming pins on the FPGA in order to program it. [3]

Q1.2 (b)

Difficulties associated with taking an FPGA design forward to an ASCI design include accounting for differences in propagation delays and operational speeds, different layouts of components; possibly different implementations of components or CLBs that are utilized. Changes in the interconnections and electrical properties of the material used for the ASIC. Futhermore, the tool chains may be quite different and require the designer to undergo a lengthy learning curve to learn how to use the tools effectively. There would also need to be more reliance on simulation, due to the expense of running of physical instances of ASICs; whereas for FPGAs it is just a matter of programming the FPGA and testing it on hardware, using a development kit prototyped board. Risks for ASIC include the potential for having a re-do designs and the expense of additional runs to compensate for design faults. Further there may be the risk of hiring consultants to assist with ASIC design and that it is difficult to predict how long it will take to achieve a final operational ASIC due to the complexity of this practice. [4]

Q1.2 (c) Advantages of parallel code are: potential for increased performance (by doing multiple operations in parallel as opposed to being limited to sequential operation), the potential for redundancy and fault tolerance (e.g. running the same operation on multiple different processors which could be used to work around interference or damage that could cause processors to fail temporarily or permanently). Improved responsiveness / decreased latency, the ability to respond to interrupts more quickly, without necessarily relying on one available processor to handle the request. [3]

Q1.3

1.3 (a)

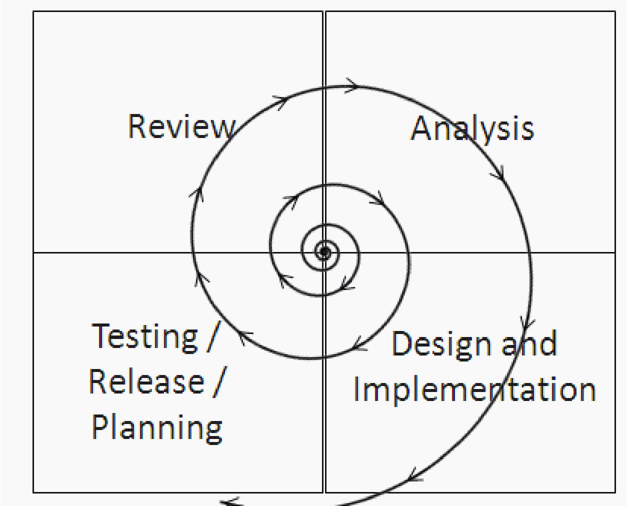
Answer: Reasoning for identifying critical parts / ‘hotspots’: This involves determine where most of the work needs to be done. Most scientific and technical programs accomplish the most substantial portion of the work in only a few small places. So it is more effective to focus on parallelizing these hotspots instead of overall improvement that would take more time – essentially ignoring parts of the program that don’t need much CPU use and can be completed quick enough with just one processor. [3 marks]

1.3 (b) *Description of the spiral model and progression of development:*

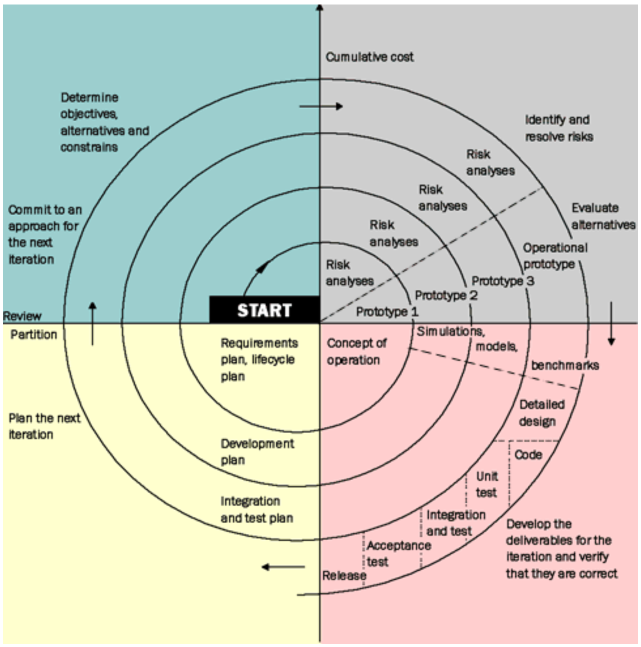
Answer: Major activities repeated:

- Analysis
- Design/implementation/prototyping
- Testing and planning for the next iteration
- Review

The spiral model tends to start small and ‘wind’ into an increasingly more complex and complete product. The diagrams below illustrates suitable spiral model. (note on marking: a level of detail as shown on the left is more expected in a student’s answer, i.e., an indication of where the 4 activities above could be position would be sufficient – the more detailed diagram is just give as a reference for marking). [5 marks]



A spiral* model view of development



1.3 (c) *Description of the spiral model and progression of development:*

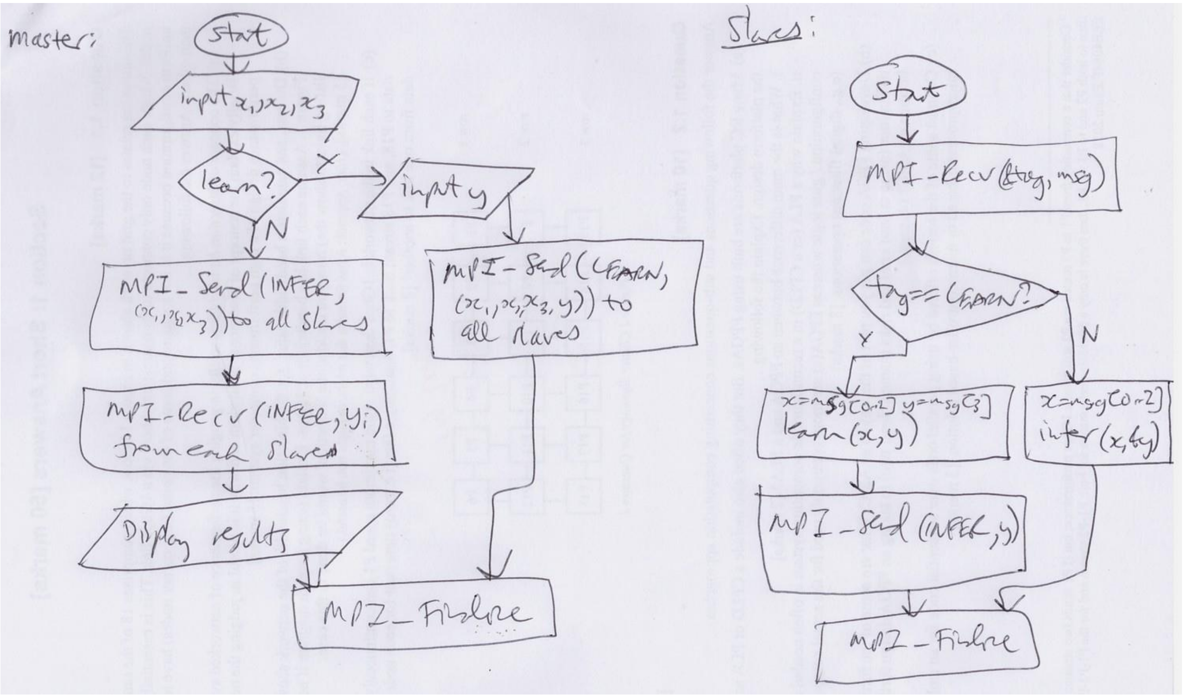
Answer: fine grained [2 marks]

1.3 (d) *Where dynamic work assignment would be applied rather than partitioning work as a preprocess:*

Answer: “Dynamic work assignment” is used for operations where the workload is unknown, or cannot be effectively calculated, before starting the operation. [2 marks]

Q1.4

1.4 (a) A simple diagram is all that is needed to express the student's understanding… e.g.:



1.4 (b) The student would be expected to write a simple main function that first reads in the command line parameters. The solution look something like the following, annotations show marking.

```
/* Program MIPP
    Machine learning Integer Prediction Program (MIPP)
    Given a group of integers this program tries to predict an output
*/

#include <mpi.h>
#include <stdio.h>
#include <string.h>
#include <stdlib.h>

/** Assume these functions are already defined: */
void learn ( int x[3], int y )  {}  /* ML training alg */
void infer ( int x[3], int *y ) {} /* ML inference alg */

/** Define tags for the messages */
#define LEARN 1
#define INFER 2

int mean (int* x, int n)
{
        int sum = 0;
        for (int i=0; i<n; i++) sum+=x[i];
        return sum/n;
}
```

```
/** Main Function */

int main(int argc, char **argv) {
  int my_rank;
  int worldsize;
  int learning = 0;
  MPI_Init(&argc, &argv); /* Start MPI */

  /* Get Rank of this processor */
  MPI_Comm_rank(MPI_COMM_WORLD, &my_rank);

  /* Determined number of processors, needed for voting */
  MPI_Comm_size(MPI_COMM_WORLD, &worldsize);

  printf("Hi! I am processor ID %d of %d processes\n", my_rank, worldsize);

  /* Implement the psedocode as given in the question... */
  if (my_rank == 0) {
     /* This is the Master */
         if (argc<2) {
                 printf("Not enough arguments\n");
                 MPI_Abort(MPI_COMM_WORLD, 1);
                 MPI_Finalize();  /* EXIT MPI */
         }
     if (strcmp(argv[1],"learn")==0) {
               int x[4];
               // y is x[4]!  so that it all fits into one buffer
        learning = 1;
        printf("LEARNING MODE\n");
               if (argc < 5) {
                       printf("No es posible :(  Correct syntax: mipp x1 x2
x3 y)\n");
                       MPI_Abort(MPI_COMM_WORLD, 1);
                       MPI_Finalize();  /* EXIT MPI */
                       }
               x[0] = atoi(argv[1]);
               x[1] = atoi(argv[2]);
               x[2] = atoi(argv[3]);
               x[3] = atoi(argv[4]);
               printf("TRAINING: %d %d %d -> %d\n", x[0], x[1], x[2],
x[3]);
               MPI_Bcast(x, sizeof(x), MPI_BYTE, 0, MPI_COMM_WORLD);
               printf(" master is done because there is no response
needed\n");
        } else {
               int x[3];
               if (strcmp(argv[1],"infer")!=0) {
                       printf("No es posible :(  Correct syntax: mipp x1 x2
x3 y)\n");
                       MPI_Abort(MPI_COMM_WORLD, 1);
                       MPI_Finalize();  /* EXIT MPI */
                       }
        printf("INFER MODE\n");
               x[0] = atoi(argv[1]);
               x[1] = atoi(argv[2]);
               x[2] = atoi(argv[3]);
               printf("INFER: %d %d %d -> ?\n", x[0], x[1], x[2]);
               MPI_Bcast(x, sizeof(x), MPI_BYTE, 0, MPI_COMM_WORLD);
               printf(" master now needs to get back the predictions\n");
               int y[worldsize-1];
               MPI_Status status;
               for (int i=1; i<worldsize; i++) {
                        MPI_Recv(&y[i-1],sizeof(y[i-
1]),MPI_CHAR,i,0,MPI_COMM_WORLD,&status);
                        printf("  -- From %d got %d\n",i,y[i]);
                    }
               /* now it should do voting or avaraging on what is the most
common answer */
               printf("DECISION: %d\n", mean(y,sizeof(y)/sizeof(int)));
        }
  } else {
     /* This is one of the Slaves */
         MPI_Status status;
         if (strcmp(argv[1],"learn")==0) {
                   int x[4];
                   printf("  SLAVE LEARNING!\n");
                       // Note that with broadcast you don't use recv, all
nodes except root receives
                       MPI_Bcast(x, sizeof(x), MPI_BYTE, 0,
MPI_COMM_WORLD);
            printf("  slave %d received: %d %d %d %d\n", my_rank, x[0],
x[1], x[2], x[3]);
                       learn(x,x[3]);
            } else {
                       int x[3];
                       int y = my_rank;
                       printf("  SLAVE INFERRING!\n");
                       // Note that with broadcast you don't use recv, all
nodes except root receives
                       MPI_Bcast(x, sizeof(x), MPI_BYTE, 0,
MPI_COMM_WORLD);
                       printf("  slave %d received: %d %d %d\n", my_rank,
x[0], x[1], x[2]);
                       infer(x,&y);
                       MPI_Send(&y,sizeof(y),MPI_CHAR,0,0,MPI_COMM_WORLD);
                       printf("  -+ From %d sent %d\n",my_rank,y);
               }
         }

  if (my_rank == 0) printf("All done :)\nHasta luego cocodrilo!\n");
  MPI_Finalize();  /* EXIT MPI */
}
```

SECTION 2 [each question worth 5 marks]

Q2.1 (d)

Q2.2 (c)

Q2.3 (b)

Q2.4 (a)

Q2.5 Answer true or false to each question below (each answer is 2 marks).

(a) DeepQA is a natural language processing system. **TRUE**

(b) A GPU is only able to process graphics data, in particular two dimensional matrices, vectors and pixels. **FALSE**

(c) The real-world performance measure is a major telling factor in determining whether or not it was worth the effort to develop a parallel solution. **TRUE**

(d) Amdahl's law states the number of transistors per square inch on ICs doubles every 18 months. **FALSE (that is Moore's Law!)**

(e) A CPU tends to support more threads than a GPU. **FALSE**

SECTION 3: Long Answers [40 marks]

3.1-3.3 These questions will be marked to award solutions that are more optimal and easier to read with high results.

3.4 This is more an essay question. The students should explain the process or state machines that would be utilized to implement modes 2 and 3 of the VCB device. Ideally it is expected that some sample code should be provided to help express how parts of the system will be connected and also how the timing will be managed.

[40 marks]