Department of Computer Science
School of Engineering and Applied Science
The George Washington University

Doctoral Preliminary Examination
Spring 2002

Duration of Exam: 5 hours
The exam consists of four parts. You must answer one question for each part.

Use a separate examination book for each question.
If you have some doubts regarding the contents of a given question, do not ask the proctor for clarification; clearly specify your assumptions and proceed with your answer.
You are not allowed to use any notes or other materials for this exam.

PART A

You must answer one question (Question 1 or Question 2) from this part of the exam.

Question 1:
A pipelined RISC computer has the following characteristics:

1) A 32-bit fixed-point data word that uses two's complement notation.
2) 64 general-purpose integer registers called R registers.
3) A 32-bit register HI that holds the most significant product bits of a fixed-point multiply instruction.
4) A 32-bit register LO that holds the least significant product bits of a fixed-point multiply instruction.
5) For a fixed-point division instruction, the register HI receives the 32-bit remainder and the register LO receives the 32-bit quotient.
6) The 32-bit instruction format is:

8-bits	6-bits	6-bits	6-bits	6-bits
Op-code	R_d	R_s	R_t	Function

7) The following table defines the instructions:

Operation	Definition	Registers	
ADD	Add integer	$R_d = R_s + R_t$	
SUB	Subtract integer	$R_d = R_s - R_t$	
MUL	Multiply integer	$R_s \times R_t$	Most significant bits of product in register HI: least significant bits of product in register LO
DIV	Divide integer	R_s / R_t	Quotient in register LO; remainder in register HI
MFHI	Move from HI	R_d	Move contents of register HI into register R_d.
MFLO	Move from LO	R_d	Move contents of register LO into register R_d.
BE	Branch on equal	R_d , $(R_s) = (R_t)$.	If the contents of register R_s equals the contents of register R_t, the address of the next instruction is in register R_d.

8) The computer has a single branch-delay slot after a branch instruction is executed.
9) The computer uses a five stage pipeline that is defined as:

Instruction Fetch	Instruction Decode and Register Fetch	Instruction Execution	Memory Store or Fetch	Write Back to Registers.

10) The computer does not have any hardware data forwarding (shelving), but relies on an optimizing assembler to reorder instructions so as to obtain maximum throughput.

11) The computer does have data hazard and branch hazard detection hardware and will insert bubbles (stalls) into the pipeline when necessary.
12) All instructions require only one cycle to perform the operation.

The following assembly language program is to be run on the computer specified above:

1) ADD R_3 , R_1, R_2
2) SUB R_4, R_5, R_3
3) MUL R_7, R_8
4) MFLO R_{11}
5) ADD R_9, R_6, R_3
6) ADD R_{10}, R_4, R_1
7) SUB R_6, R_1, R_2
8) ADD R_{16}, R_6, R_1
9) BE R_{14}, R_{11}, R_3
10) ADD R_{15}, R_3, R_4

Instruction 1) cannot be moved and is the first instruction executed.

a) If you are the optimizing assembler, show the order to execute the instructions so as to minimize the total time to execute the instructions.
b) How many machine cycles will it take to execute the instructions as shown in the program above? Show how you obtain this figure.
c) How many machine cycles will it take to execute the program you have shown for part a) above? Show how you obtain this figure.

Question 2:

The key objective of a computer architect is to design an architecture that improves performance – where performance is typically measured in terms of the execution time of the application program. The past two decades have seen the development of processor architectures that have performance that beats the performance of early supercomputers. A key enabling concept has been the execution of multiple instructions in each clock cycle – a concept that is seen implemented in processors ranging from pipelined processors, to superscalar processors (such as the PowerPC) to the latest class of architectures 'EPIC-Explicitly Parallel Instruction Computing' (such as Intel's Itanium IA-64 class of processors). These processors provide sophisticated techniques for executing instructions – such as speculation, and predication. Their actual performance however, depends on how the instruction scheduling is performed and can be limited due to factors such as branches, and dependencies.
This questions requires you to discuss some of these problems and solutions that have been provided.

- Why do branches limit the performance and discuss ONE method for handling branches in instruction level parallel processors. Give an example to illustrate your answer.
- What kinds of dependencies can exist in a program and how do they limit performance ? Give an example to illustrate your answer.
- How can speculative and predicated execution improve performance, and what new challenges and problems do they create ? Give a simple example (in high level code if you prefer) to illustrate how (a) speculation and (b) predication can improve performance.

PART B

You must answer one question (Question 3 or Question 4 or Question 5) from this part of the exam.

Question 3:

We would like to design a domain-specific architecture for a web-based course registration system (**CRS**). This type of systems is essential for university management information systems and online distance learning systems. The system includes three types of objects: administrators, instructors, and students.

- Functionalities of administrators include:
    - Get the list of all instructors.
    - Add/Update courses
    - Add/update instructor details
    - Etc.
- Functionalities of instructors include:
    - Create and update quizzes
    - Enter Student grades
    - Etc.
- Functionalities of students include:
    - Check grades
    - Submit evaluation forms
    - Etc.

The above lists of functionalities of all categories are not complete. You need to add additional functionalities to this problem domain.

1. State the major elements of a domain-specific software architecture (DSSA).

2. Provide a complete DSSA for **CRS**.

3. What kind of metrics would you use to evaluate two DSSAs for a given domain name?

Question 4:

Having recently graduated from GW's Department of Computer Science, you have joined the Bigger Is Better (BIB) Corporation (formed from HP, Compaq, IBM, Unisys, Sun, and Apple) to help design and write a new operating system for a scalable multiprocessor system. In CS210, you acquired a set of tools for constructing large software systems and are eager to use these tools in this project. Individual operating system services will be assigned to different processors within this new computer. Each service is implemented wholly on one processor, but must communicate with other services (and processors).

As the system architect, you have to define the overall structure and describe the design down to the second architectural level. You have concluded that five (5) major services are required: Processor Management, Storage Management, Communications Management, I/O Management, and User Management are required. (a) Describe the architectural framework which enables these services to integrate and interoperate. (b) For each service, describe the software architecture that best supports that service and give three (3) detailed reasons for each service why you have selected a software architecture for that service. (c) Interfaces between the services will be based on design patterns. Describe five (5) design patterns that you would use in implementing the interfaces and provide a detailed rationale for your choice. (d) Finally, you need to specify a programming language (or languages) to implement the new OS. Which one would you choose and why? Provide three reasons for choosing the language(s) you recommend.

Question 5:

Iteration statements (FOR, WHILE, and their variants) appear in all programming languages; but there are often significant differences between them, in various languages.
Pick three languages that you know, and, for each language, describe in detail how the iteration statements work in that language.
Do not choose, among your three languages, two languages (such as C and C++) in which the iteration statements are substantially the same. Also give as many advantages and disadvantages as you can, of the various iteration statements when compared to one another.

PART C

You must answer one question (Question 6 or Question 7) from this part of the exam.

Question 6:

Many generalizations of binary search trees, such as 2-3 trees, 2-3-4 trees, red-black trees, and AVL trees, have been devised. Every one of these generalizations has been devised to address, in one way or another, one specific problem with ordinary binary search trees.

(A). Describe in detail what this problem is.

(B). Choose one of the generalizations, and describe in detail how it solves this problem.

Question 7:

a- What is the time complexity of the greedy-based minimum spanning tree algorithm for arbitrary weighted undirected graphs?

b- Can a faster algorithm be found if all the edges have the same weight? If so, indicate briefly the general approach and give the corresponding time complexity.

c- Assume now that each edge has a weight equal to *a* or *b*, where *a* and *b* are two given positive constants. Give an $O(n+E)$ algorithm to find a minimum spanning tree, where n is the number of nodes and E the number of edges in the input graph**.**

PART D

You must answer one question out of Questions 1 through 7 for this part of the exam. The question you pick MUST be different from the three questions you answered in Parts A,B, and C.