# VLSI Architecture Design Course (048853)
# Final Exam
# July 4th, 2002

**Electrical engineering Department**

**Student name:______________________ Student number:________________**

**This exam contains TWO questions.**
**The exam duration is 2:30 hours.**
**Please fill the answers ON THE EXAM forms.**

**Please explain or provide a formula for each computation!**

**TAKE YOUR TIME, READ THE QUESTIONS THOUROUGHLY, UNDERSTAND THE CONTENT AND ONLY THAN START TO ANSWER**

**Good luck!**

| | |
|---|---|
| Q1 | |
| Q2 | |
| Total | |

## Question 1 (50%)

**A microprocessor system is given (see Figure 1). The system contains a CPU with a given $CPI_1$ (for 100% cache hit rate), frequency of operation $f_1$, a 256KB cache with 0 access time (CPU-to-cache) and an Hit Rate (HR) =99% for SPECint_base2000 (assume 1 cache access per instruction), and a memory system. The cache is a blocking cache, so none of the misses overlap. The access frequency to memory is $F_{BUS}$=300Mhz and the data access time (request-to-data) is $L_{BUS}$=10 bus cycles.**
**Performance vs. frequency measurements were performed and the results are shown in Figure 2.**

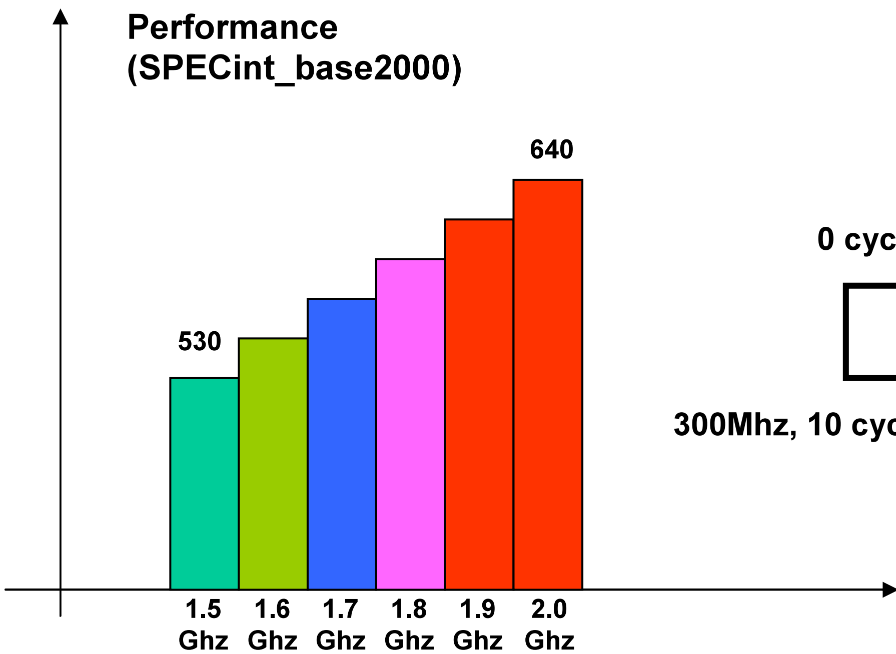



**Figure 2**

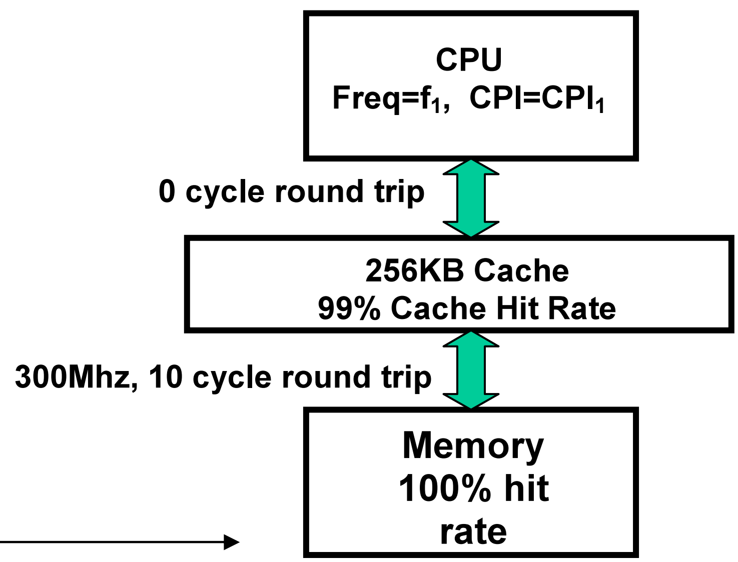



**Figure 1**

**QUESTIONS:**

A. Specify in a formula the contribution of the compute time and the memory time to the total execution time.

B. Calculate the processor's $CPI_1$, (for 100% cache hit rate), using the system characteristics and the performance graph (Figure 2).

**C. The Cache size was increased to 512KB.**
**Figure 3 shows the performance vs. frequency of the system with 512KB cache. What is the Cache Hit Rate ($HR_2$) of the new cache?**

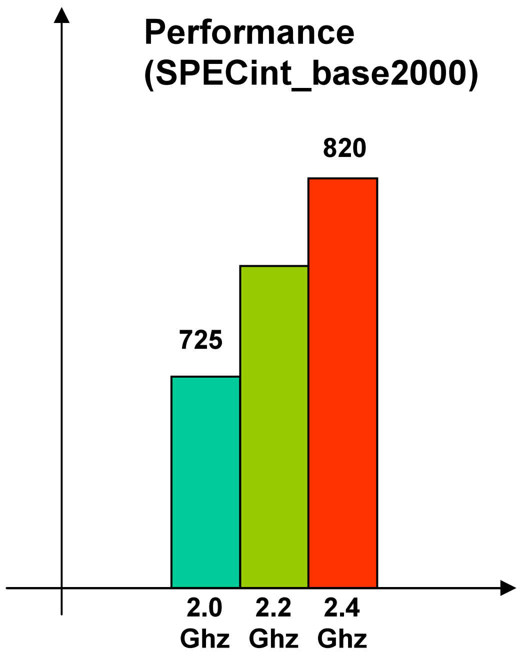
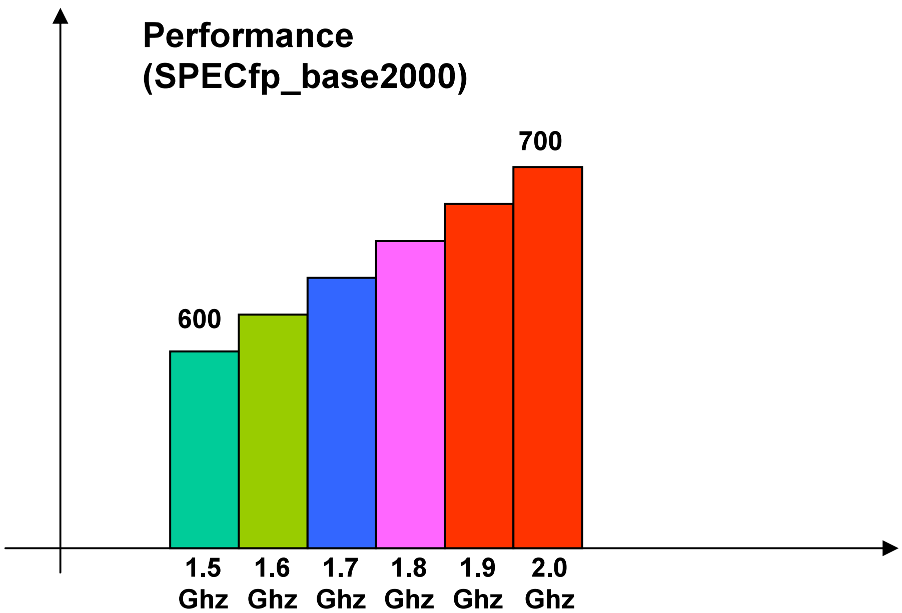



**Figure 3**

**D.** **Figure 4 Shows the Floating Point (FP) performance of the original configuration (256KB Cache). From the figure identify the difference behavior of SPECfp_base2000 benchmarks vs. Integer SPECint_base2000.**




**D1.** **What is the difference between the SPECfp_base2000 and SPECint_base2000 processor behavior?**

**D2.** **Provide reasoning why this difference exist.**

**D3. Knowing integer CPI can you find out the parameters of FP (FP CPI, Cache hit rate…)?**
**If yes, express it. If no, which parameter is missing?**

## BONUS QUESTION

**E. What should the architect do to enable performance growth when CPU frequency continues to increase?**

**F. What is the "break point"? Meaning: when will the direction you provide in E, will "run out of steam"? When do you get into a diminishing return area?**

## Question 2 (50%)

**This question deals with a branch prediction mechanism for a microprocessor. The system consists of the following (see also figure 1):**

1. **Branch predictor (BP) - A two level predictor – chooser :**
   **A local and a global predictors multiplexed by a chooser.**
2. **A standard BTB (branch target address buffer).**
3. **Unconditional branches bypass the BP (i.e., always taken).**
4. **A subroutine return stack, for predicting procedure return address. Bypasses both the BP and the BTB.**
5. **Software prediction (static) tag – overrides the dynamic BP.**
   **Every conditional branch opcode holds 2 bits that translates as:**

| opcode | Action |
|---|---|
| 00 | Null – Uses the BP |
| 01 | Predict taken |
| 10 | Predict not taken |
| 11 | Unpredictable – wait for branch resolution before fetching target |

6. **Fixed target address cache: the compiler can tag (with a special dedicated bit in the opcode) those indirect branches that always dynamically compute the same target address. These target addresses are kept in a special cache, that bypasses the BTB. Otherwise, normal BTB behavior is assumed.**

**Note: The software prediction tags and the fixed target address tags in the opcode are available and decoded at fetch time**

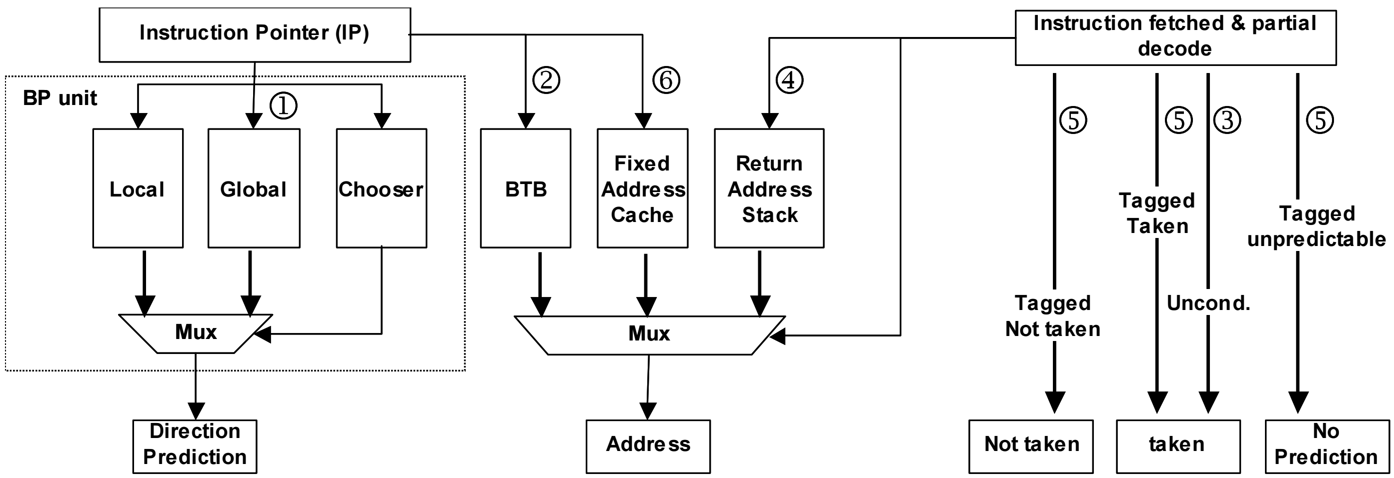



**Figure 1**

**The numbers on the arrows refer to the above numbered items.**

## QUESTIONS:

**A.** **For the above numbered items 3,4,5,6 – explain briefly what is the motivation and the possible gain from using such feature - why it is better be used instead of BP and BTB only, which SW construct can use that? For item 5 explain the use of all the 4 tag options.**

**B. The branches divided into categories:**

| Type | % of dynamic branches | success rate | Access BP? | Access BTB? |
|---|---|---|---|---|
| Unconditional | 25% | | | |
| Direct | 15% | 100% | | |
| Return inst. | 5% | 95% | | |
| Indirect | 5% | 80% | | |
| Conditionals | 75% | | | |
| None | 60% | | | |
| Tagged Taken | 7% | 98% | | |
| Tagged Not taken | 7% | 98% | | |
| Tagged Unpredict. | 1% | | | |
| Total | 100% | -- | | |

- **The overall BP prediction rate is 90%.**
- **The overall BTB success rate is 95%.**
- **The fixed target address cache is used for 50% of the unconditional indirect branches and it is 100% correct.**

**B.1 Why the BTB is needed? Why is it not sufficient to get the target address from the decoded instruction?**

**B.2 What is the percentage of branches reaching the BP? Fill the Access BP column in the table above.**

**B.3 What is the percentage of branches reaching the BTB? Fill the Access BTB column in the above table.**

**B.4** **What is the percentage of the branches with wrong direction prediction?**

**B.5** **What is the percentage of the branches with successful target address prediction? (ignore the tagged not taken branches)**

**B.6.** **Can you deduce the percentage of the branches that is correctly predicted? Explain. Specify a range (Maximum & Minimum).**

**<u>Note: the following sections do not depend on the previous one!</u>**

**C. Assume that every $5^{th}$ instruction is a branch instruction and the branch mis-prediction rate is 6%.**
**The average perfect IPC is 1.5.**
**The pipe length is 8 cycles.**
**The mis-predictions are detected at the last stage of the pipe and cause a full flush of the rest of the pipe. Assume (ideally) that in the next cycle the correct instructions are fetched (perfect L1 instruction cache).**

**C.1 What is the MPI (misses per instructions - defined as the number of branch mis-predictions divided by the total number of instructions executed)?**

**C.2 What is the performance loss due to branch mis-prediction? (execution time with branch mis-prediction penalty divided by execution time with perfect prediction ).**

D. A novel BP (NBP) has a correct prediction rate of 99% and 4 stall cycles (the existing BP has a 94% correct prediction rate and no stall cycles). The NBP is integrated with the existing system, as a 2nd level BP.

Assume that the BP and the NBP prediction decision and success are uncorrelated.

D.1 Why is it not recommended to use the NBP only? Explain.

D.2 How should the two BPs be integrated?

D.3. Compute the MPI with of the new integrated BP.

D.4. Compute the number of stall cycles for the new integrated BP. Consider the following 4 cases:

BP and NBP are both correct:

BP and NBP are both incorrect:

BP incorrect, NBP correct:

BP correct, NBP incorrect:

D.5. What will be the performance loss with the new integrated BP?