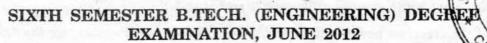
Name....

Reg. No



IT 04 604—COMPUTER ARCHITECTURE

Time: Three Hours

Maximum: 100 Marks

- Explain the three basic variations in instruction encoding.
  - 2 State Amdahl's Law and describe how speedup can be improved.
  - 3 What is meant by instruction level parallelism? Describe how it can be increased.
  - 4 Explain how vector performance can be enhanced.
  - Assume that the hit time of a two-way set-associative first-level data cache is 1.1 times faster than a four-way set-associative cache of the same size. The miss rate falls from 0.049 to 0.044 for an 8 KB data cache, Assume a hit is 1 clock cycle and that the cache is the critical path for the clock. Assume the miss penalty is 10 clock cycles to the L2 cache for the two-way set-associative cache, and that the L2 cache does not miss. Which has the faster average memory access time? Indicate assumptions made, if any.
  - 6 What are the different levels in a typical memory hierarchy? Compare their features.
  - 7 Describe how multiple computers can be connected together.
  - 8 When is a memory system said to be coherent?

 $(8 \times 5 = 40 \text{ marks})$ 

- 2. (a) Explain the five stage pipeline with suitable diagram. What are the basic performance issues in pipelining?
  - (b) Explain the various instruction set architectures with suitable diagrams. For a sample instruction, show the code sequence in each of the instruction set architecture.

 $(1 \times 15 = 15 \text{ marks})$ 

- 3. (a) Explain Dynamic Scheduling Using Tomasulo's Approach. Give suitable diagram and example.
  - (b) What are the goals of the multiple-issue processors? What are the different forms of Multiple-issue processors? Explain the characteristic features of each form.

 $(1 \times 15 = 15 \text{ marks})$ 

- 4. (a) Explain the various ways of reducing cache miss rates.
  - (b) Imagine that the storage system is configured to contain two 40 GB disks in a RAID 0 array; that is, the data is striped in blocks of 8 KB equally across the two disks with no redundancy.
    - (i) How will the 40 GB of data be allocated across the disks? Given a random request workload over a total of 40 GB, what is the expected service time of each request?

Turn over

C 28766

- (ii) How can queuing theory be used to model this storage system?
- (iii) What is the average utilization of each disk?
- (iv) On average, how much time does each request spend waiting for the disk?
- (v) What is the mean number of requests in each queue?
- (vi) What is the average response time for the disk requests?

 $(1 \times 15 = 15 \text{ marks})$ 

- 5. (a) What is Multiprocessor Cache Coherence? Explain the basic Schemes for enforcing Coherence.
  - (b) What is the need for synchronization? What are the hardware primitives used to achieve synchronization? Explain how synchronization is achieved using these primitives.

 $(1 \times 15 = 15 \text{ marks})$