

Computer architecture

Final review

Jiao jiajia

Email : jaojiajia@shmtu.edu.cn

Shanghai Maritime University

Agenda

- **Exam time & location**
- **Exam basics**
- **Review all of 10 units**
- **Notices about your usual performance**
- **Summary**

Exam time & location

- Time
 - June 23th, Wed morning in the 17th week
 - 8:20 am - 9:55 am
- Location
 - 2B201 (33 students)
 - 2B203 (69 students)

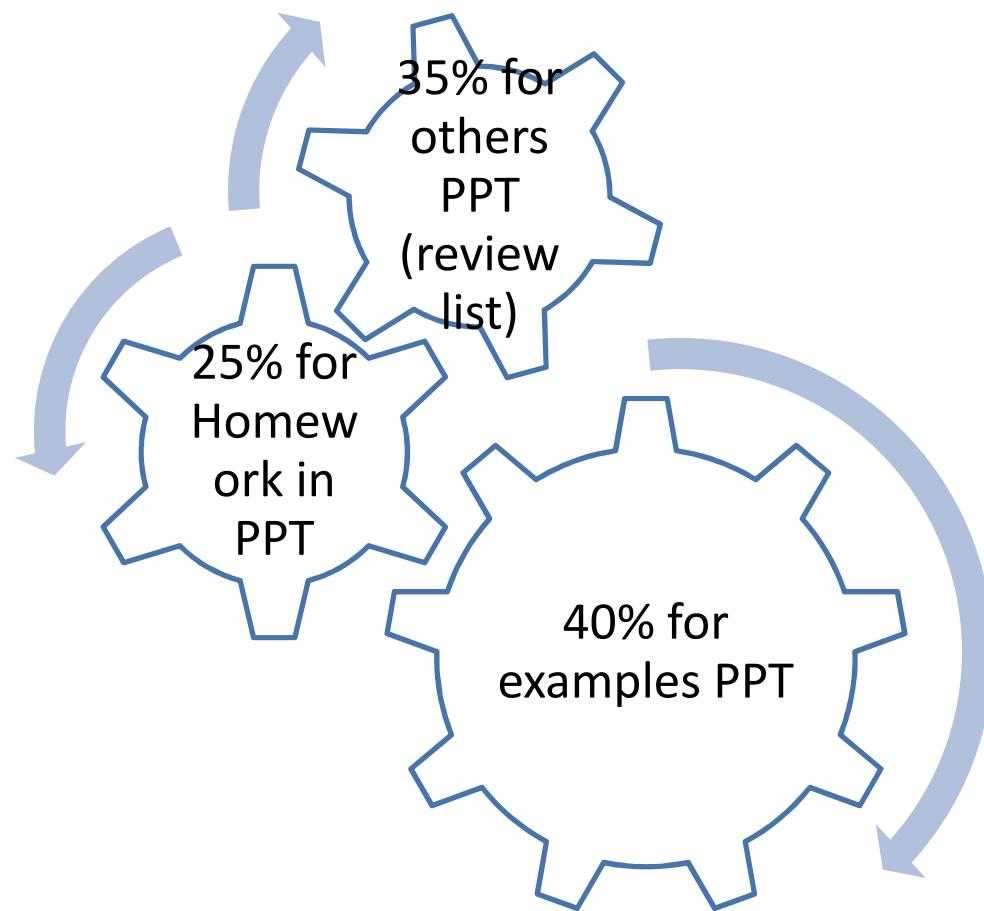
Exam types

- **Closed exam**
- **Three types**
 - 12 choices ($1.5' * 12 = 18'$)
 - 6 analysis ($5' * 4 + 10' + 12' = 42'$)
 - 4 calculation ($8' * 2 + 12' * 2 = 40'$)

Exam requirements

- **On time in the correct location**
- **English is better**
- **List the formulas first**
- **Details as many as possible**
- **No cheating**

Review materials **PPT** & time allocation



Review 10 units

- Unit10: power issue (Charpter1)
- Unit 9: dependability issue (Charpter1)
- Unit 8: GPU (Charpter5)
- Unit 7: interconnection networks(AppendixF)
- Unit 6: multicore architecture(Charpter5)
- Unit 5: memory system(Charpter2+AppendixB)
- Unit4: pipeline technique(Charpter3+AppendixC)
- Unit3: isa(AppendixA)
- Unit2: performance evaluation (Charpter1)
- Unit1: introduction (Charpter1)

Key points

- **Pipeline technique**
 - Basics (idea, classification)
 - Metrics (throughput, speedup, efficiency)
 - Dependence & hazards & Solutions
- **Cache design & optimizations**
 - Cache basics (4 questions)
 - Cache performance evaluation(CPU time, average access time, memory stall cycles)
 - Cache optimizations

Difficult points

- **Multicore**
 - Reasons for to be the processor mainstream
 - Cache coherence
 - Write invalidate vs. write update
 - MSI based state transition for two requests from CPU and bus

Unit10: power issue

1. Why consider power issues?
2. **Power calculation (2 parts)**
3. **Power optimization (4 methods)**

Unit 9: dependability issue

1. Computer dependability
2. **Important terms (fault, error, failure)and their relations**
3. Fault classification
4. Computer two states

Unit 9: dependability issue

5. Dependability measures(two metrics)
6. **More detailed metrics and calculation:
MTTF,FIT,MTBF,MTTR, Availability**
7. Mitigation schemes(redundancy and error coding)

Unit 8: GPU

- 1. GPU architecture design: 3 ideas**
- 2. SIMD processing implementation: 2 options**
3. Interleaving between contexts management:
hardware or software or both
4. GPU memory design
5. GPU programming

Unit7: interconnection networks

- 1. Why interconnection networks?**
- 2. Interconnection networks organizational structure**
- 3. Topology properties and examples**
- 4. Flow control schemes**
- 5. Routing schemes**
- 6. Communication mode**

Unit6: multicore architecture

1. Why multicore?
2. typical multicore architectures: two kinds
3. **communication model for a multicore: two types**
4. performance challenges: two points

Unit6: multicore architecture

5. cache coherence

- issue
- 2 solutions
- 2 protocols
- 2 state transition diagrams based on MSI

6. Memory Coherence & Consistency

7. Multicore programming

Unit 5: memory system

1. Memory hierarchy features
2. Why memory hierarchy?
3. Memory performance metrics(cost, size, average access time, hit rate, hit time, miss rate, miss penalty)
4. Cache-memory vs. memory-secondary storage
5. Cache definition and design principles

Unit 5: memory system

6. Typical cache terms: cache hit, cache miss, cache block....
7. Block placement or originations: 3 types
8. Cache access based on address format
9. Lookup algorithm: 3 options
10. Replacement policy: typical 3 choices

Unit 5: memory system

11. Write hit policy: 2 types
12. Write miss policy: 2 types
13. Cache performance evaluation(average memory access time and CPU time based on memory stall cycles definition and calculation)
14. Cache misses classification: 3 types
15. Cache optimizations(6 basic+10 advanced)

Unit 5: memory system

16. Multilevel cache performance evaluation
(metrics definition and calculation: local miss rate, global miss rate, average memory access time, stall cycles per instruction)
17. Multilevel cache inclusion vs. exclusion
18. Memory technology: SRAM,
DRAM(optimizations), Flash memory(features)
19. Flash cell
20. Flash memory vs. disk

Unit 5: memory system

21.FTL functions

22.Virtual memory definition and jobs

23.Virtual memory vs. cache

24.Virtual memory classification: 2 + hybrid

25.4 Memory Hierarchy Questions Revisited in
virtual memory

26.TLB

Unit4: pipeline technique

1. Pipeline main idea
2. Pipeline terms (pipeline depth, pipeline stage)
3. Pipeline classifications (single function vs. multiple function, static vs. dynamic, linear vs. non-linear)
4. 5-stage instruction pipeline(MIPS): 5 stages
5. Pipeline diagram

Unit4: pipeline technique

6. Pipeline metrics and calculation: 3 measures(speedup efficiency) and 2 cases(uniform or non-uniform)
7. Pipeline bottleneck issue and solutions
8. Dependence definition, classification and analysis
9. Dependence solutons
- 10.Hazard definition, classification and analysis

Unit4: pipeline technique

11.Data hazards

12.Stall or bubble

13.Forwarding

14.Split design

15.Compiler optimization for pipeline scheduling

Unit4: pipeline technique

16.Branch speculation (static vs. dynamic)

17.Branch delay

18.Delayed branch

19.Delay slots

20.Instruction scheduling: 3 types

Unit3: isa

- 1. ISA definition**
- 2. ISA vs. microarchitecture**
3. ISA execution model
4. Instruction format and types
5. Operations, Operand model, Data types and transfers
- 6. Good ISAs: 3 factors**
7. Benchmarks definition
- 8. RISC vs. CISC**

Unit2: Performance evaluation

- 1. Metrics and calculation (throughput, latency, CPU time, Frequency, CPI, MIPS....)**
2. Benchmarks definitions and typical examples
- 3. Amdal's law**
- 4. principle of locality**
5. analytical modeling definition and classification
6. simulations and speedup techniques

Unit1: Introduction

1. Why computer architecture?
2. Computer architecture topics
- 3. Computer architecture definition**

Summary

- **Final exam(40%)**
 - Time
 - Location
 - Details
- Usual performance(60%)
- **Study hard!!!**