TRƯỜNG ĐH BÁCH KHOA – ĐHQG-HCM KHOA ĐIỆN-ĐIỆN TỬ	**THI CUỐI KỲ**	**Học kỳ/năm học**	**2**	**2024-2025**
		Ngày thi		**18/05/2025**
	Môn học	**Thiết kế hệ thống trên chip**		
	Mã môn học	**EE5157**		
	Thời lượng	**90 phút**	**Mã đề**	**1**

Ghi chú: *- Học viên được sử dụng tài liệu giấy (viết tay, photocopy, in). Không được sử dụng: ĐTDĐ, máy tính xách tay, máy tính bảng,....*
- Đề thi gồm 02 trang.

Câu hỏi 1) (L.O.1) (1.5đ):

Cho CPU với thời gian xử lý các tác vụ cho 1 lệnh (instruction) Load như sau:

Instruction Fetch	Instruction Decode	Execute	Memory Access	Write back
180ps	140ps	160ps	220ps	120ps

Tính các thông số: Cycle time, Latency, Throughput, Speed-up (trường hợp sử dụng pipeline so với không dùng pipeline) trong 2 trường hợp sau:
1. Không sử dụng pipeline.
2. Sử dụng pipeline 5 tầng. Giả sử ở mỗi tầng pipeline phát sinh thêm thời gian trễ 15ps (do thanh ghi pipeline).

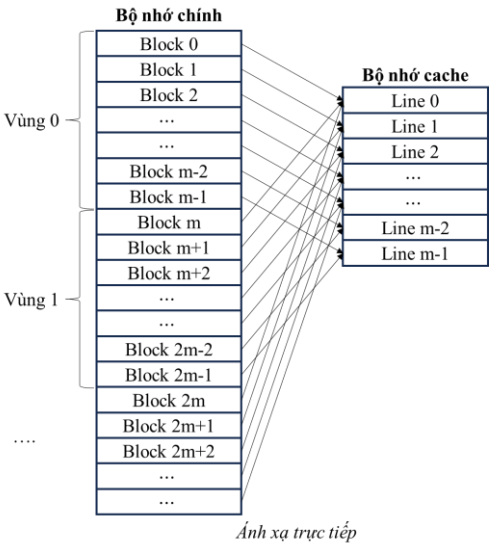
Câu hỏi 2) (L.O.1) (1.0đ):

Máy tính A hoạt động ở tần số 2GHz thực thi một chương trình trong 10s. Các kỹ sư muốn thiết kế một máy tính B có khả năng thực thi chương trình đó chỉ trong thời gian 7s bằng cách tăng tần số so với máy tính A. Tuy nhiên, khi tăng tần số thì số chu kỳ cần thiết để thực thi chương trình trên máy tính B cũng tăng lên 1.4 lần so với số chu kỳ cần thiết để thực thi chương trình trên máy tính A. Hỏi tần số máy tính B phải là bao nhiêu để đạt được thời gian xử lý như mong muốn.

Câu hỏi 3) (L.O.3) (2.0đ):

Cho sơ đồ ánh xạ bộ nhớ giữa bộ nhớ chính (main memory) và bộ nhớ cache (cache memory) theo phương pháp "ánh xạ trực tiếp" (direct mapping) như hình bên. Cấu trúc địa chỉ bộ nhớ chính (main memory address) như sau:

Tag	Index	Block offset
t	*i*	*b*



Trong đó: trường Tag (*t*-bit) dùng để xác định vùng nào của bộ nhớ chính được lưu trữ trong bộ nhớ cache (mỗi vùng của bộ nhớ chính có dung lượng bằng dung lượng bộ nhớ cache); trường Index (*i*-bit) dùng để xác định line trong bộ nhớ cache mà block trong bộ nhớ chính được ánh xạ; trường Block offset (*b*-bit) dùng để phân biệt các byte trong cùng một block/line.

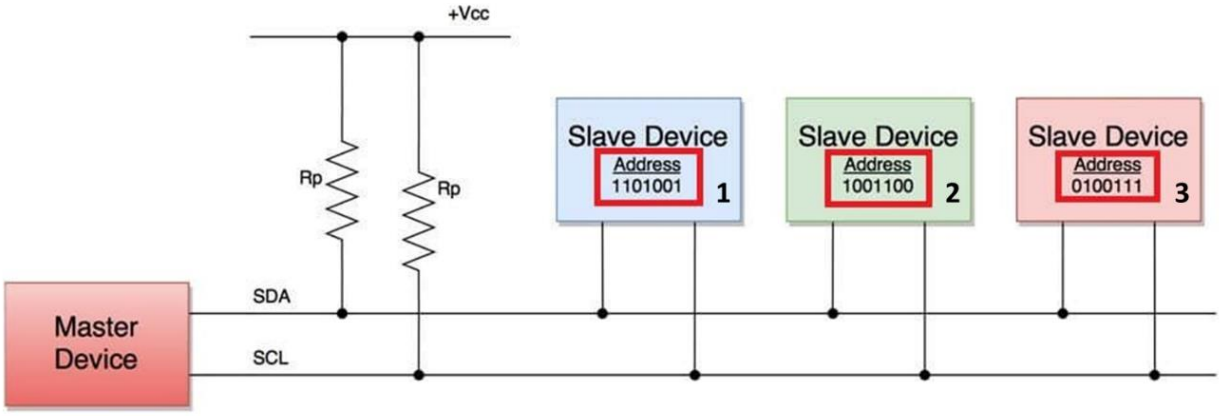
Giả sử bộ nhớ chính (main memory) có dung lượng là 4GB, định địa chỉ theo byte (mỗi byte trong bộ nhớ có một địa chỉ xác định), được tổ chức thành các block, mỗi block có dung lượng là 64 byte. Bộ nhớ cache (cache memory) có dung lượng là 256KB; kích thước mỗi line trong bộ nhớ cache bằng với kích thước mỗi block trong bộ nhớ chính.
1. (1.0đ) Xác định số bit tương ứng với các trường Tag (*t*), Index (*i*) và Block offset (*b*).
2. (1.0đ) Giả sử vi xử lý cần truy xuất từ nhớ có địa chỉ là 180525 (hệ thập phân) trong bộ nhớ chính. Xác định xem từ nhớ này thuộc vùng nào của bộ nhớ chính (thứ tự vùng nhớ được đếm bắt đầu từ 0). Nếu từ nhớ ở địa chỉ trên có tồn tại trong bộ nhớ cache thì sẽ thuộc line nào trong bộ nhớ cache?

Câu hỏi 4) (L.O.3) (1.5đ):

Dynamic power là gì (hình vẽ minh họa, giải thích, công thức tính)? Trình bày chi tiết (vẽ hình, công thức, giải thích) 2 kỹ thuật dùng để giảm Dynamic power? Nhận xét ưu-khuyết điểm của 2 kỹ thuật đó.

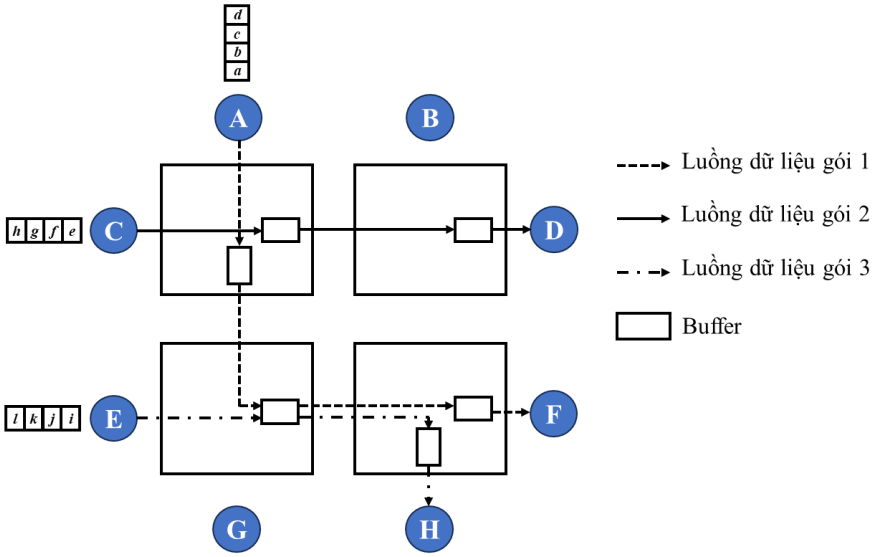
Câu hỏi 5) (L.O.3) (1.5đ):

Cho kết nối giữa Master và các Slave theo giao thức I2C như hình bên dưới. Mô tả quá trình Master muốn gửi liên tục 2 byte data 12H và 34H xuống Slave 2 từ lúc bắt đầu cho đến kết thúc khung truyền.



Câu hỏi 6) (L.O.6) (1.5đ):

Xem xét mạng chuyển mạch gói (packet switching) 2x2 router sử dụng phương pháp Wormhole switching như hình bên. Có 3 gói dữ liệu cần được gửi gồm: gói 1 gồm 4 flit '***abcd***', từ A đến F; gói 2 gồm 4 flit '***efgh***', từ C đến D; và gói 3 gồm 4 flit '***ijkl***', từ E đến H. Giả định rằng định tuyến đã được tính toán như trong hình vẽ. Các router sử dụng bộ đệm (buffer) để lưu các flit. Giả sử mỗi bộ đệm lưu tối đa 2 flit khi đường truyền bị tắc nghẽn (buffer của router kế chưa sẵn sàng nhận). Trường hợp kênh truyền không tắc nghẽn thì buffer chỉ lưu 1 flit. Gọi T là thời gian cần xử lý 1 flit. Ban đầu cả 3 gói 1, 2 và 3 đều gửi flit đầu tiên (header flit), lần lượt là '*a*', '*e*' và '*i*' đến bộ đệm tương ứng. Sau khoảng thời gian T, flit kế tiếp được xử lý. Hãy mô tả hoạt động của mạng ở các thời điểm tiếp theo. Sau bao nhiêu T thì cả 3 gói dữ liệu đến được đích?



Câu hỏi 7) (L.O.6) (1.0đ):

Phân tích hệ thống SoC được thiết kế với công nghệ 65nm, bao gồm các thành phần trên 1 die như sau:

Thành phần	**Diện tích (Đơn vị: *A*)**
CPU core	750
GPU core	300
DMA controller	100
Peripheral I/O (UART, SPI, I2C)	350
Bus và control unit	500
Internal SRAM (1 MB)	4096
L2 cache (256 KB)	1024
1 A = 1481 rbe = $10^6 \times f^2$ (f: feature size in µm)	

Biết rằng 15% diện tích chip dành cho IO pads và diện tích khung cần thiết. Tính diện tích chip khi chưa đóng gói (gross chip area) theo đơn vị mm^2.

--- HẾT ---