

Virtual Memory

Ming <rf37535@gmail.com>

About me

- 戴宗明 (Ming)
- 師大資工四年級
- 研究領域：
 - Software-defined Network, Named-data Network
 - Neural Network
 - High Performance Computing
- Github: <https://github.com/CorcovadoMing>
- Email: rf37535@gmail.com
- Facebook: <https://www.facebook.com/nicktaming>

為什麼要了解記憶體？

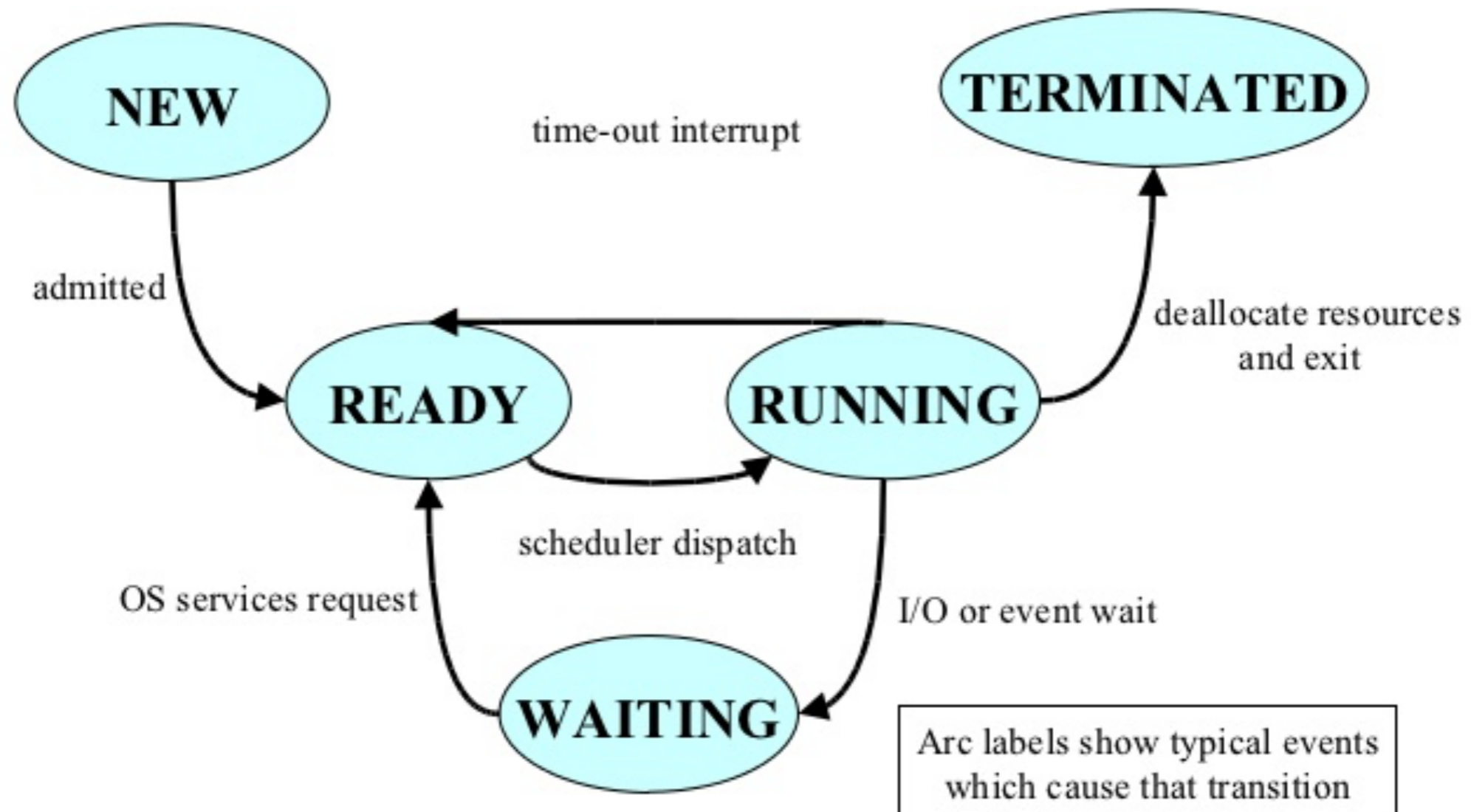
如果我撰寫高階語言，仍需要
關心記憶體配置嗎？

幾乎大部分的效率瓶頸都在記
憶體上！
(IO-Bounded)

什麼是 Virtual Memory ?

Virtual Memory 想要解決的問題是什麼？

The Five State Model



Questions

- 為什麼我們能夠開啟兩個相同的程式？
- 記憶體不夠怎麼辦？
- Virtual 的意思是什麼？

“All problems in computer science can be solved by another level of indirection”

記憶體不夠怎麼辦？

- 分享實體記憶體
- 假設電腦有 4G 記憶體，A 程式執行需要 0.5G，B 程式要 3G，C 程式要 1G
- 直接的做法是將 0.5G 分配給 A，3G 分配給 B，AB 可以同時執行
- 這個做法的問題在哪？

可能的問題

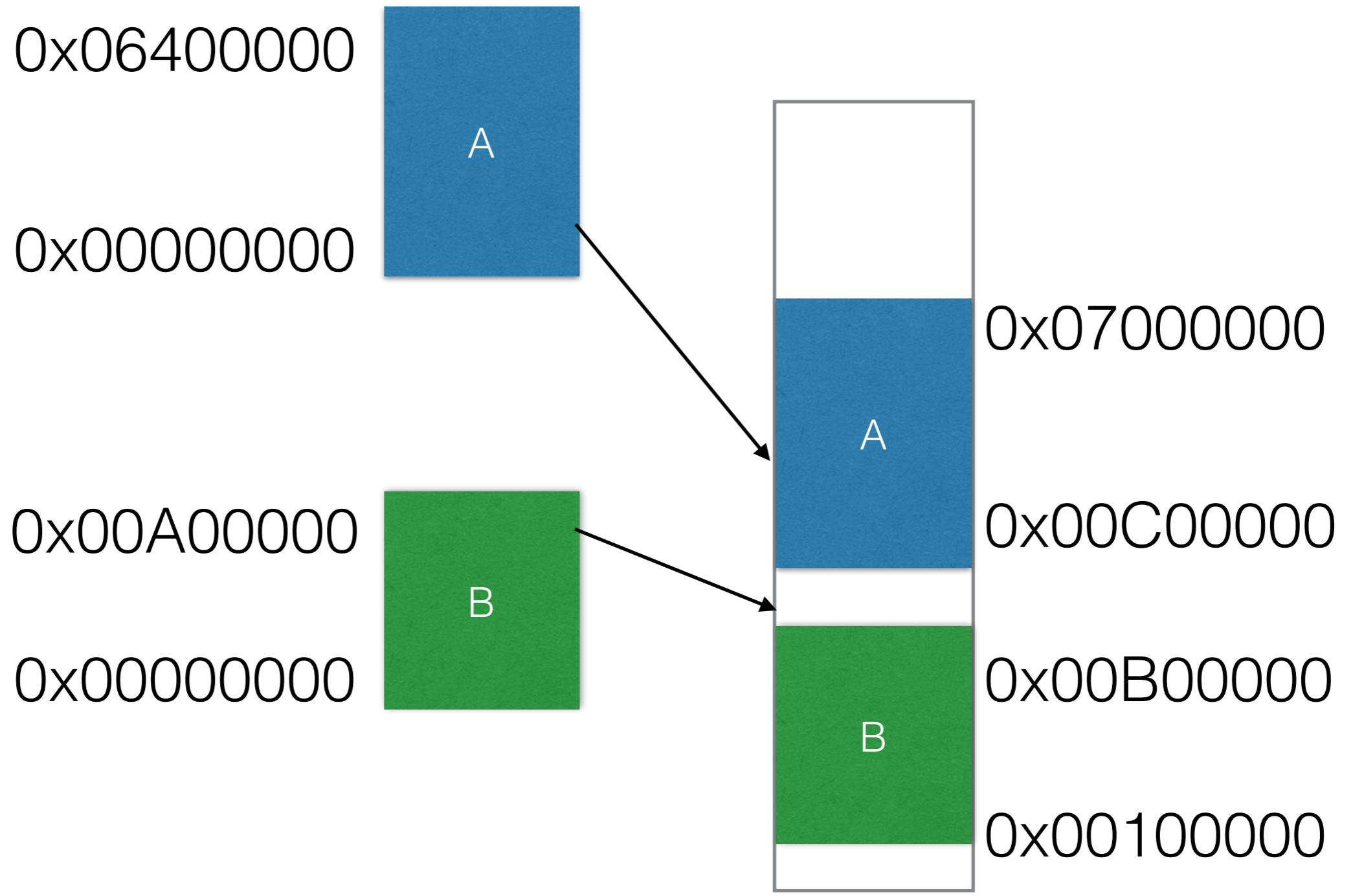
- 位址空間不隔離
- 記憶體使用效率低
- 程式執行的位址不穩定（程式指令 jump 時，若目標地址是絕對位址時？）

關於隔離

- 每個程式有自己的位址空間
- 每個程式都不需要關心其他程式
- 虛擬位址

Segmentation

- 每個人都有自己的虛擬空間
- 載入時再分配



若程式要求不屬於它的空間時會發生什麼事？

```
Ming@MBA ~$ ./a.out  
[1] 22887 segmentation fault ./a.out  
x Ming@MBA ~$
```

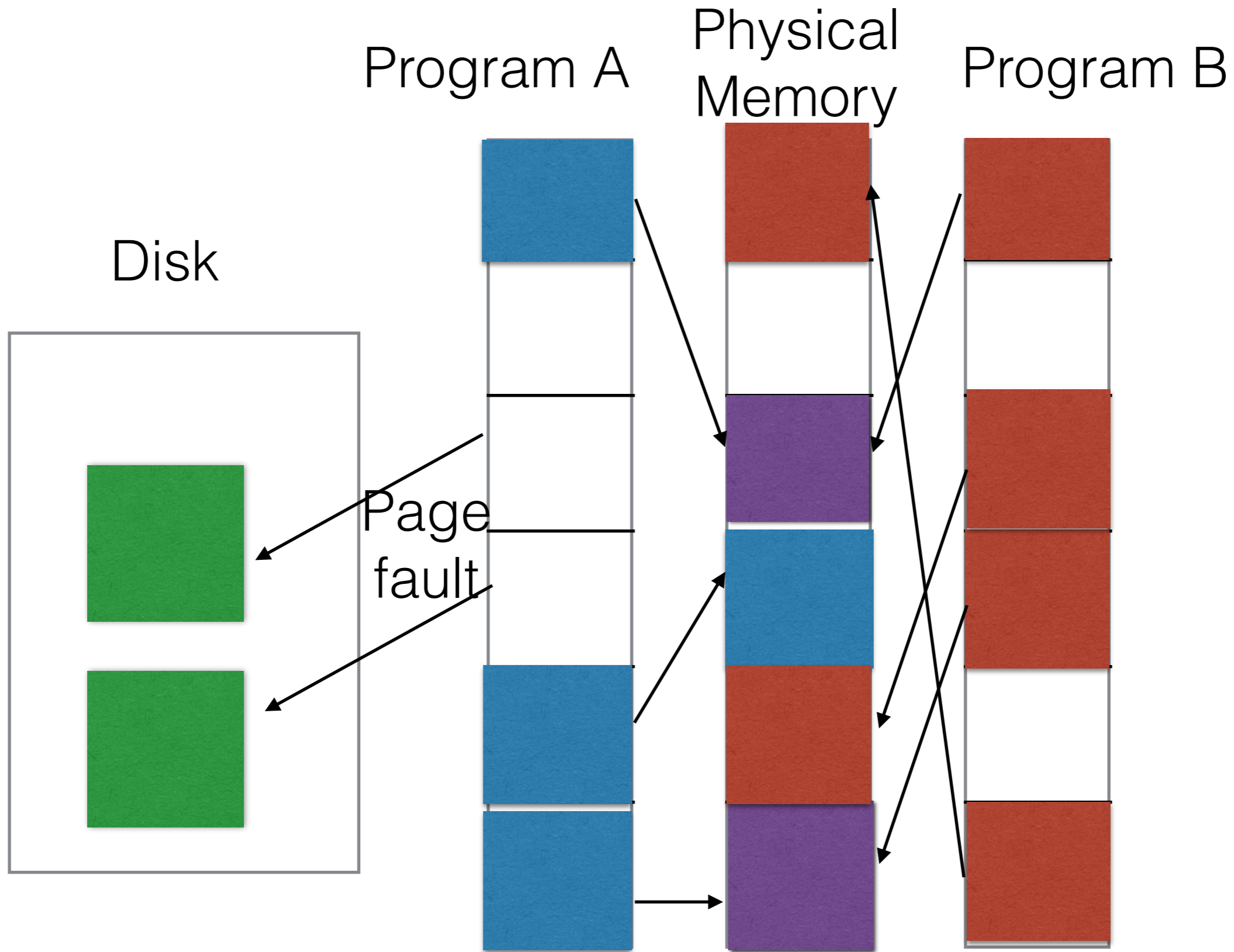
這種方法解決的哪些問題？

可能的問題

- ~~位址空間不隔離~~
- 記憶體使用效率低
- ~~程式執行的位址不穩定 (程式指令 jump 時，若目標地址是絕對位址時?)~~

Paging

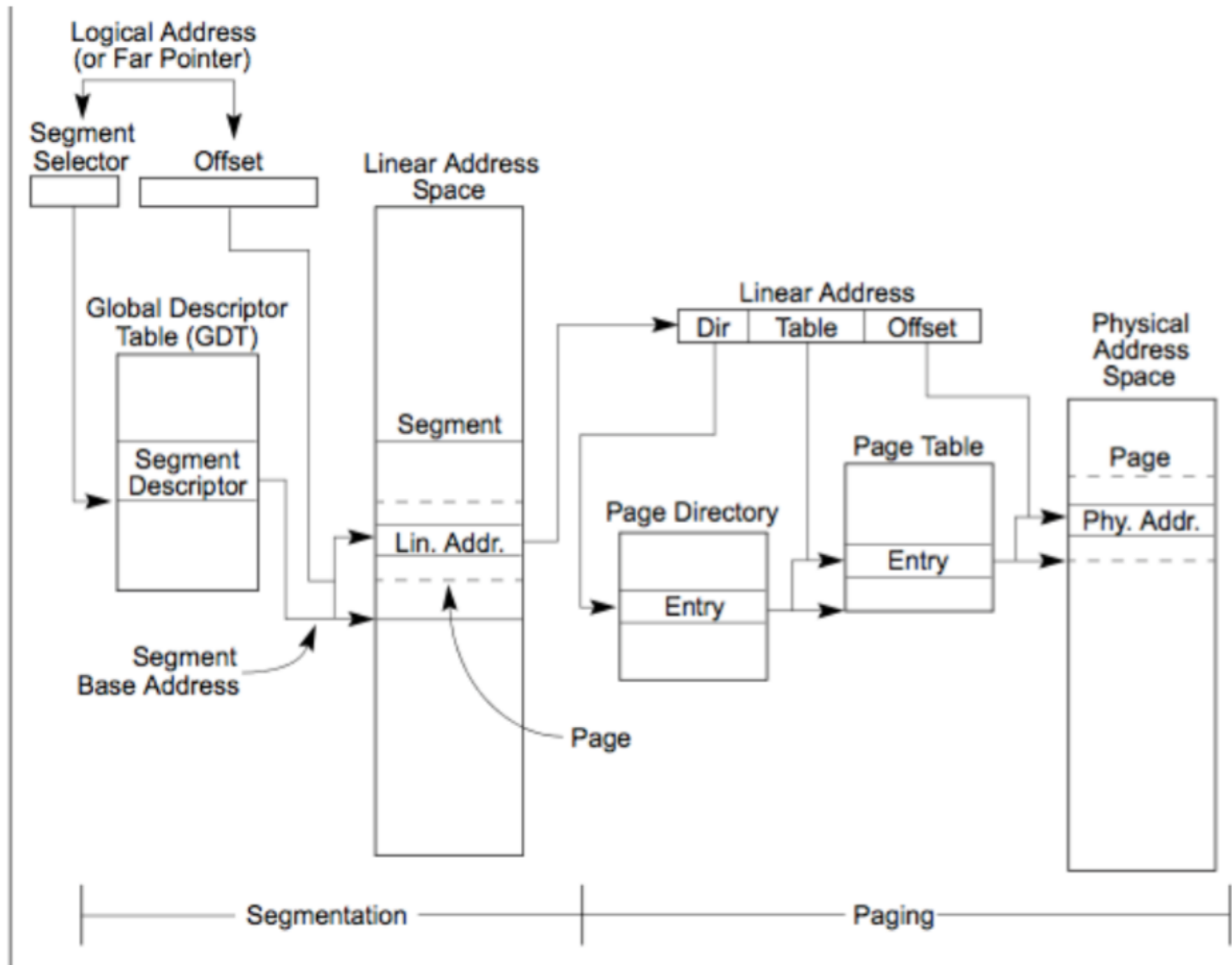
- 將位址切割成固定大小的頁面



如何決定虛擬到實體的
對應？

MMU

(Memory Management Unit)



問題

- 每次 Paging 都要經過 MMU 太複雜了
- 如何改善？

TLB

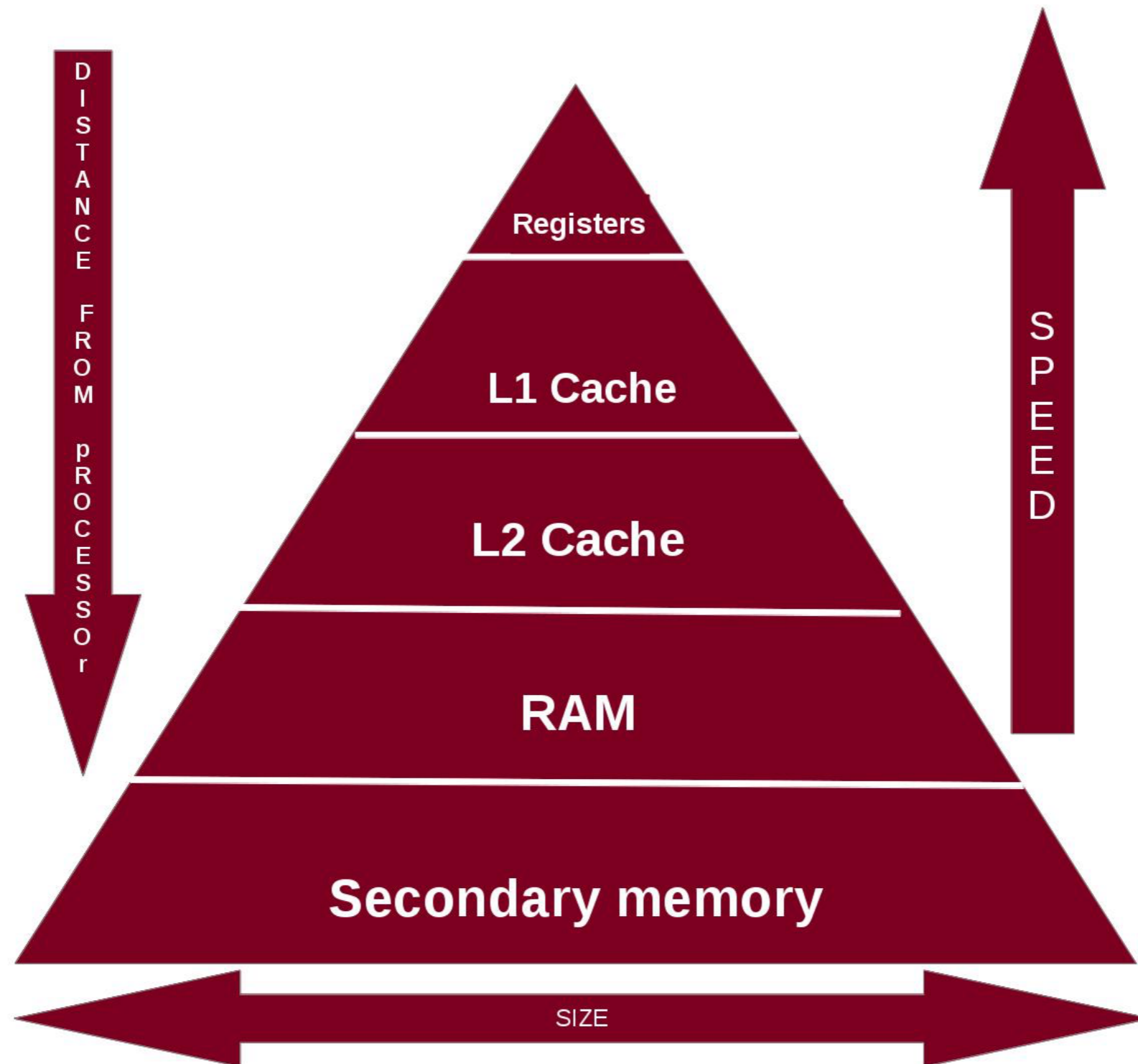
(Translation Lookaside Buffer)

- 查表法
- Cache

Worst case ?

Cache Miss
(TLB Miss)

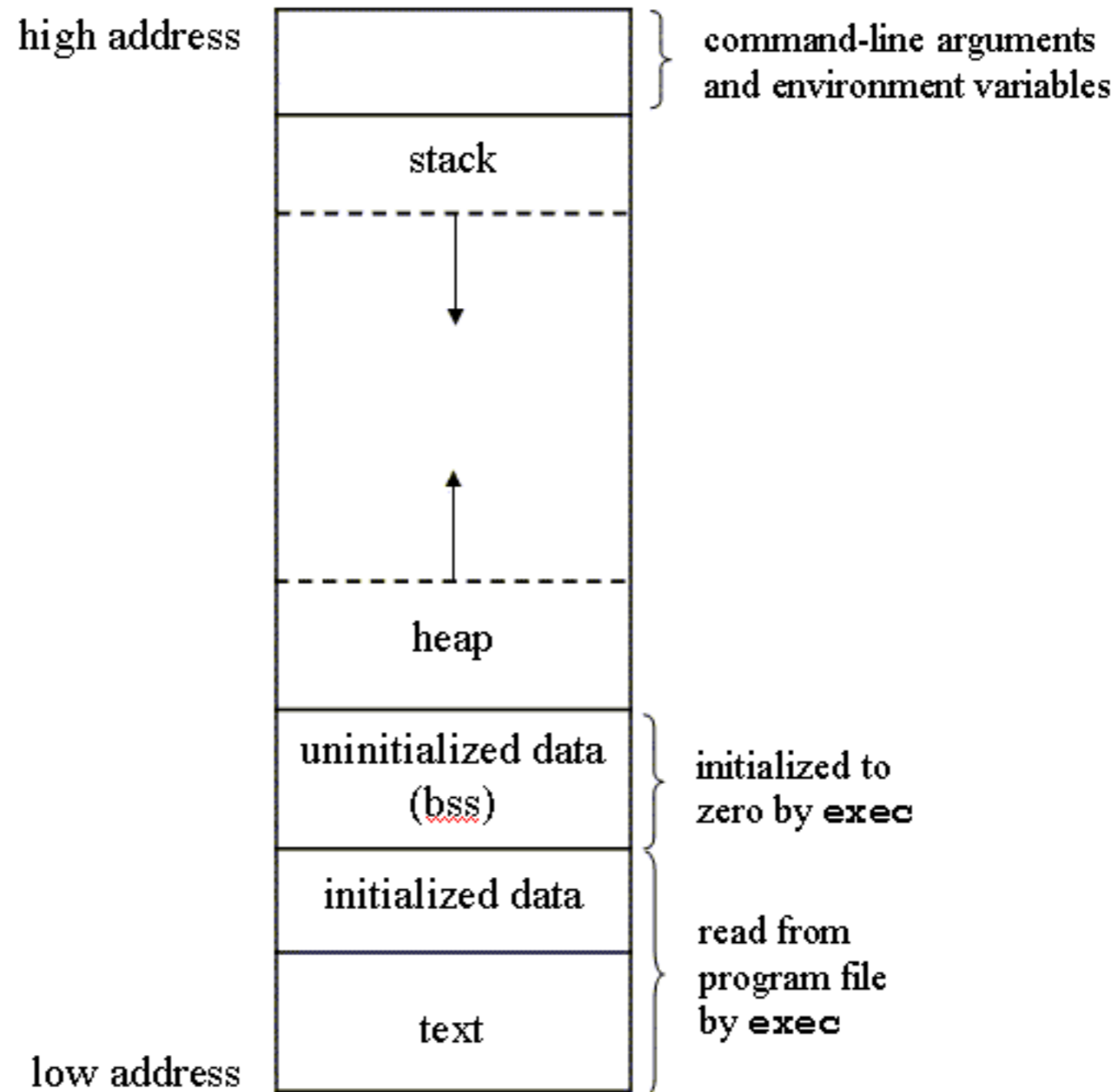
Background: 記憶體階層



剖析目的檔

- File Header
- .text section
- .data section
- .bss section

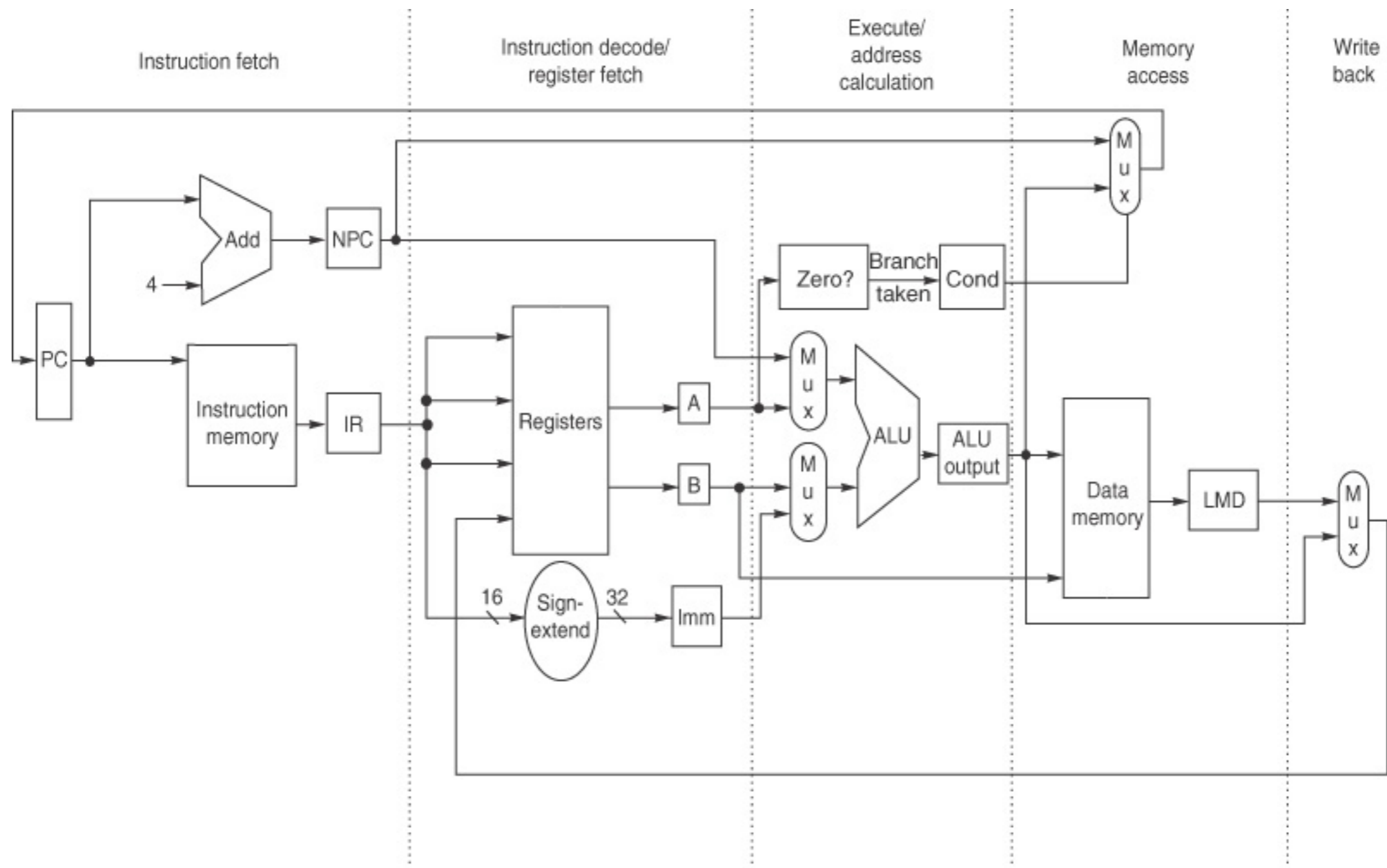
Virtual Memory Address Space of Programs



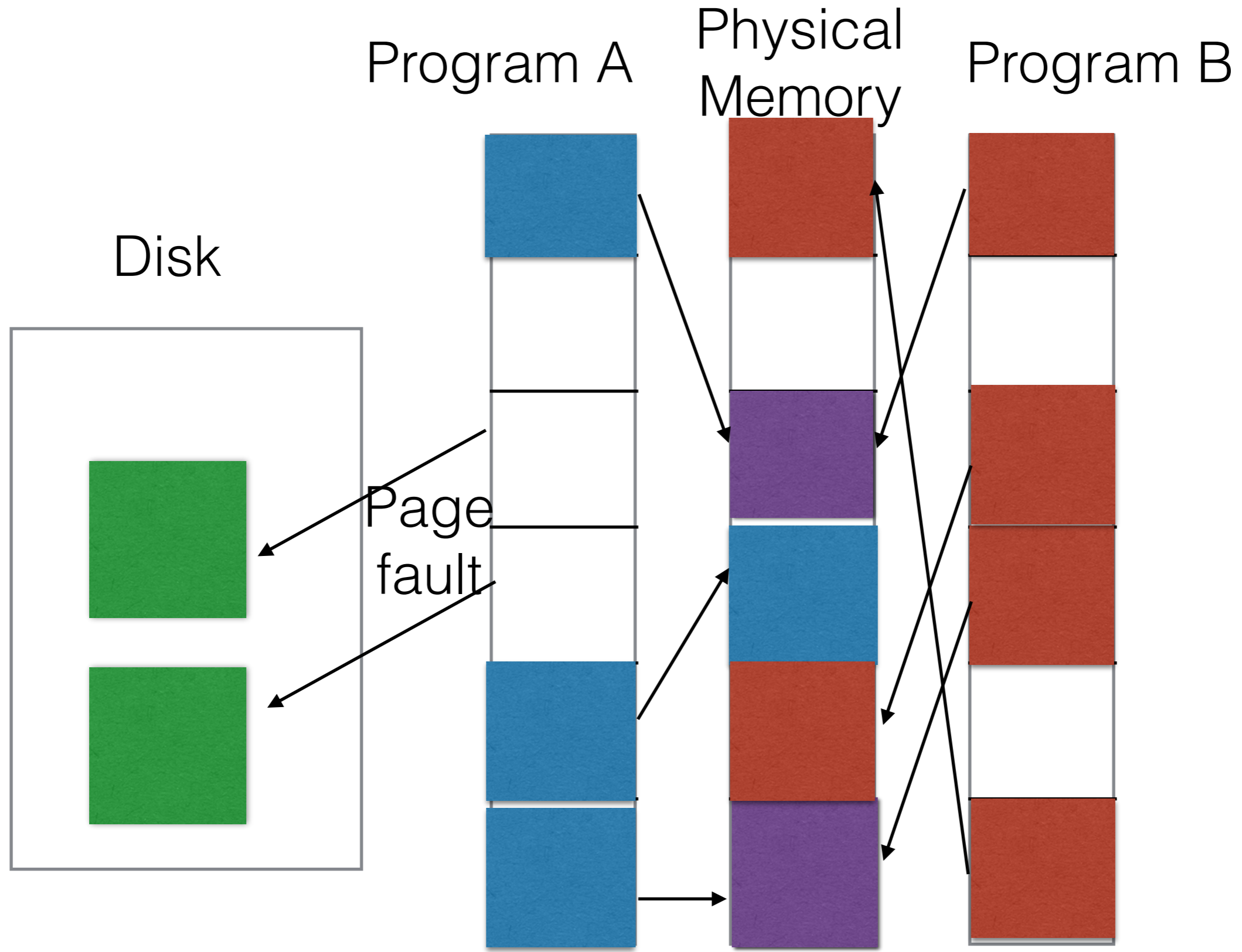
與 Cache Miss 有什麼
關係？

資料快取 & 指令快取

- .data : 可讀可寫
- .text : 唯讀



Recall: Page fault



Conclusion

- 為何需要 TLB，MMU
- 如何撰寫高效率的程式

You may want to know more about ...

- Linker and Loader
- CPU scheduling
- Memory safety
- Compiler

Questions ?