

Content

I.	Motivation	2
	Motivation and the problem analysis.....	2
	What's your plan to deal with the problem	2
II.	Implementation	3
	How do you implement to solve the problem in Nachos.....	3
	You can including some important code segments and comments	3
III.	Result.....	6
	Experiment result and some discussion	6

Reference

1. MrOrz/nachos. [Online]. Available: <https://github.com/MrOrz/nachos>
<https://github.com/taldehyde/os-project-1/blob/master/report.md>
2. Caching and Virtual Memory. [Online]. Available:
http://neuron.csie.ntust.edu.tw/homework/93/OS/homework_3/A9315010-%E7%AC%AC%E4%B9%9D%E7%B5%84/phase%203%20operation.htm

I. Motivation

Motivation and the problem analysis

✓ 若在還沒有改 nachos-4.0 的原始碼以前，同時執行 兩個 program

■ `./nachos -e ../test/matmult -e ../test/sort`

這時會因為超過預先設定的 physical page number (預設為 32)，而產生 core dump。

如下圖所示：

```
r08921005@r08921005-VirtualBox:~/nachos/3/nachos-4.0/code/userprog$ ./nachos -e ../test/matmult -e ../test/sort
Total threads number is 2
Thread ../test/matmult is executing.
Thread ../test/sort is executing.
Assertion failed: line 134 file ../userprog/addrspace.cc
Aborted (core dumped)
```

Fig. 1 修改前所遇到的問題

由於 matmult 與 sort 兩者 thread 所佔用到的 physical page 會超過預設的 32 導致 Aborted

原先 nachos 一次只保留一個 user program 在主記憶體中。

此專題目標在於使用 paging mechanism 來實作 multiprogramming，

也就是在主記憶體中保留多個 user programs，而 pages 只有在這個 program 被執行的時候才會被 loaded，也就是 pure demand paging。

What's your plan to deal with the problem

以 Program matmult 與 Program sort 為例。

輸入 `nachos -e ../test/matmult -e ../test/sort` 之後

1. 先幫這兩個 Program 各創一個 thread 與 page table，也就是生成 thread A 與 thread B。
2. 利用 AddrSpace::Load() 函式把 Program matmult 與 Program sort 之 page table 匯入進硬碟中，同時也把這兩個 thread 丟進 ready queue 裡面，這個時候 OS 會決定要讀取的程式碼在的起始地址與總共要載入多少分頁到硬碟裡。當作業系統開始執行後(執行 `kernel->machine->Run()`後)，選擇其中一個 thread 來執行。當在 page Table 中找不到要跑的 thread 時，發生 page fault，並去磁碟機者到對應的 page (demanded page)與執行 context switching(timer 讓 interrupt 取引發 context switching)，開始執行另外一個 thread。

在此專題中沒有去製造新的 Memory Manager，而是把 virtual address 與 physical address 的對應關係寫在 AddrSpace::Load()裡，利用 Page Size 跟 virtual address 去找對應 physical address。

$$\text{Physical Address} = \text{pageTable} \left[\left(\frac{\text{virtual address}}{\text{PageSize}} \right) \right] \cdot \text{physicalPage} * \text{PageSize} + (\text{virtual address} \% \text{PageSize})$$

排程的 page replacement 使用 LRU(Least Recently Used)。

同時，這邊使用 3 種 table 來維護 physical page 資訊與 swap 進入或換出主記憶體的資訊。

PageTable	One page table per process. Decide your virtual page number. <code>Kernel->machine->pageTable = currentThread ->pageTable</code>
FrameTable	整個 system 一個 紀錄每個 physical page 的資訊
SwapTable	整個 system 一個 紀錄 swap 內每個 sector 的資訊

II. Implementation

How do you implement to solve the problem in Nachos

相較於第二個 project，改動程式碼檔案如下表：

/userprog/userkernel.h /userprog/userkernel.cc	新增 swapDisk
/userprog/exception.cc	新增 PageFaultException
/userprog/addrspace.h /userprog/addrspace.cc	Address Mapping 檢查有沒有閒置的 physical page
/machine/translate.h /machine/translate.cc	實作 LRU
/machine/machine.h	紀錄 sector number, frame paged 和被 main memory, virtual memory 占用的 frame

You can including some important code segments and comments

新增 swapDisk

```
34 FileSystem *fileSystem;
```

Original Text

[/code/userprog/userkernel.h](#)

```
34 FileSystem *fileSystem;  
35 SynchDisk *swapDisk;
```

Changed Text

```
59 fileSystem = new FileSystem();
```

Original Text

[/code/userprog/userkernel.cc](#)

```
59 fileSystem = new FileSystem();  
60 swapDisk = new SynchDisk("New VM disk");
```

Changed Text

```
68 case SC_Sleep:  
69     val=kernel->machine->ReadRegister(4);  
70     cout << "Sleep Time " << val << "(ms) " << endl;  
71     kernel->alarm->WaitUntil(val);  
72     return;
```

Changed Text

[/code/userprog/exception.cc](#)

將程式載入虛擬記憶體，也就是磁碟機中。載入的同時會紀錄 virtual page 的起始位置還有要載入多少分頁到磁碟機中。然而如果可以用的 physical page 已經沒有了，則會把這段程式碼 Load 進 virtual page 裡面，並把它寫到 swap 區域中。

```
32 private:  
33     TranslationEntry *pageTable; // Assume linear page table translation  
34     // for now!  
35     unsigned int numPages; // Number of pages in the virtual  
36     // address space  
37  
38     bool Load(char *fileName); // Load the program into memory  
39     // return false if not found  
40  
41     void InitRegisters(); // Initialize user-level CPU registers,  
42     // before jumping to user code  
43
```

Original Text

[/code/userprog/addrspace.h](#)

```
33 void pageFaultHandle(int);  
34 static bool usedPhysPage[NumPhysPages];  
35 static bool usedVirPage[NumPhysPages];  
36 static bool pageType[NumPhysPages]; //code:1 data:0;  
37 static TranslationEntry *ptrPageTable[NumPhysPages];  
38 // recording which virtual pages are used.  
39 // It can be as large as we want it to be :P  
40  
41 private:  
42     TranslationEntry *pageTable; // Assume linear page table translation  
43     // for now!  
44     unsigned int numPages; // Number of pages in the virtual  
45     // address space  
46  
47     bool Load(char *fileName); // Load the program into memory  
48     // return false if not found  
49  
50     void InitRegisters(); // Initialize user-level CPU registers,  
51     // before jumping to user code  
52  
53     bool pageTableLoaded;  
54
```

Changed Text

<pre> 56: pageTable = new TranslationEntry[NumPhysPages]; 57: for (unsigned int i = 0; i < NumPhysPages; i++) { 58: pageTable[i].virtualPage = i; // for now, virt page # = phys page # 59: pageTable[i].physicalPage = i; 60: // pageTable[i].physicalPage = 0; 61: pageTable[i].valid = TRUE; 62: // pageTable[i].valid = FALSE; 63: pageTable[i].use = FALSE; 64: pageTable[i].dirty = FALSE; 65: pageTable[i].readOnly = FALSE; 66: } 67: 68: // zero out the entire address space 69: bzero(kernel->machine->mainMemory, MemorySize); 70: 71: 72: 73: 74: 75: 76: 77: AddrSpace::~AddrSpace() 78: { 79: for (int i = 0; i < numPages; i++) 80: AddrSpace::usedPhyPage[pageTable[i].physicalPage] = false; 81: 82: 83: 84: 85: 86: 87: 88: 89: 90: 91: 92: 93: 94: 95: 96: 97: 98: unsigned int size; 99: 100: 101: 102: 103: 104: 105: 106: 107: 108: 109: 110: 111: 112: 113: 114: 115: 116: 117: for (unsigned int i = 0, j = 0; i < numPages; i++) { 118: pageTable[i].virtualPage = i; 119: while (j < NumPhysPages && AddrSpace::usedPhyPage[j] == true) 120: j++; 121: AddrSpace::usedPhyPage[j] = true; 122: pageTable[i].physicalPage = j; 123: pageTable[i].valid = true; 124: pageTable[i].use = false; 125: pageTable[i].dirty = false; 126: pageTable[i].readOnly = false; 127: } 128: // end morris add 129: size = numPages * PageSize; 130: 131: ASSERT(numPages <= NumPhysPages); // check we're not trying 132: // to run anything too big -- 133: // at least until we have 134: // virtual memory 135: 136: DEBUG(dbgAddr, "Initializing address space: " << numPages << ", " << size); 137: DEBUG(dbgAddr, "Initializing code segment."); 138: DEBUG(dbgAddr, noffH.code.virtualAddr << ", " << noffH.code.size); 139: executable->ReadAt(140: &(kernel->machine->mainMemory[pageTable[noffH.code.virtualAddr / PageSize].physicalPage * PageSize + (noffH.code.virtualAddr % PageSize)]), 141: noffH.code.size, noffH.code.inFileAddr); 142: 143: 144: 145: 146: 147: 148: 149: 150: 151: 152: 153: 154: 155: 156: 157: 158: 159: 160: 161: 162: 163: 164: 165: 166: 167: 168: 169: 170: 171: 172: 173: 174: 175: 176: 177: 178: 179: 180: 181: 182: 183: 184: 185: 186: 187: 188: 189: 190: 191: 192: 193: 194: 195: 196: 197: 198: 199: 200: 201: 202: 203: 204: 205: 206: 207: 208: 209: 210: 211: 212: 213: 214: 215: 216: 217: 218: 219: 220: 221: 222: 223: 224: 225: 226: 227: 228: 229: 230: 231: 232: 233: 234: 235: 236: 237: 238: 239: 240: 241: 242: 243: 244: 245: 246: 247: 248: 249: 250: 251: 252: 253: 254: 255: 256: 257: 258: 259: 260: 261: 262: 263: 264: 265: 266: 267: 268: 269: 270: 271: 272: 273: 274: 275: 276: 277: 278: 279: 280: 281: 282: 283: 284: 285: 286: 287: 288: 289: 290: 291: 292: 293: 294: 295: 296: 297: 298: 299: 300: 301: 302: 303: 304: 305: 306: 307: 308: 309: 310: 311: 312: 313: 314: 315: 316: 317: 318: 319: 320: 321: 322: 323: 324: 325: 326: 327: 328: 329: 330: 331: 332: 333: 334: 335: 336: 337: 338: 339: 340: 341: 342: 343: 344: 345: 346: 347: 348: 349: 350: 351: 352: 353: 354: 355: 356: 357: 358: 359: 360: 361: 362: 363: 364: 365: 366: 367: 368: 369: 370: 371: 372: 373: 374: 375: 376: 377: 378: 379: 380: 381: 382: 383: 384: 385: 386: 387: 388: 389: 390: 391: 392: 393: 394: 395: 396: 397: 398: 399: 400: 401: 402: 403: 404: 405: 406: 407: 408: 409: 410: 411: 412: 413: 414: 415: 416: 417: 418: 419: 420: 421: 422: 423: 424: 425: 426: 427: 428: 429: 430: 431: 432: 433: 434: 435: 436: 437: 438: 439: 440: 441: 442: 443: 444: 445: 446: 447: 448: 449: 450: 451: 452: 453: 454: 455: 456: 457: 458: 459: 460: 461: 462: 463: 464: 465: 466: 467: 468: 469: 470: 471: 472: 473: 474: 475: 476: 477: 478: 479: 480: 481: 482: 483: 484: 485: 486: 487: 488: 489: 490: 491: 492: 493: 494: 495: 496: 497: 498: 499: 500: </pre>	<pre> 55: ID = (kernel->machine->Identity)++; 56: kernel->machine->Identity = (kernel->machine->Identity)++; 57: 58: 59: 60: 61: 62: 63: 64: 65: 66: 67: 68: 69: 70: 71: 72: 73: 74: 75: 76: 77: 78: 79: 80: 81: 82: 83: 84: 85: 86: 87: 88: 89: 90: 91: 92: 93: 94: 95: 96: 97: 98: 99: 100: 101: 102: 103: 104: 105: 106: 107: 108: 109: 110: 111: 112: 113: 114: 115: 116: 117: 118: 119: 120: 121: 122: 123: 124: 125: 126: 127: 128: 129: 130: 131: 132: 133: 134: 135: 136: 137: 138: 139: 140: 141: 142: 143: 144: 145: 146: 147: 148: 149: 150: 151: 152: 153: 154: 155: 156: 157: 158: 159: 160: 161: 162: 163: 164: 165: 166: 167: 168: 169: 170: 171: 172: 173: 174: 175: 176: 177: 178: 179: 180: 181: 182: 183: 184: 185: 186: 187: 188: 189: 190: 191: 192: 193: 194: 195: 196: 197: 198: 199: 200: 201: 202: 203: 204: 205: 206: 207: 208: 209: 210: 211: 212: 213: 214: 215: 216: 217: 218: 219: 220: 221: 222: 223: 224: 225: 226: 227: 228: 229: 230: 231: 232: 233: 234: 235: 236: 237: 238: 239: 240: 241: 242: 243: 244: 245: 246: 247: 248: 249: 250: 251: 252: 253: 254: 255: 256: 257: 258: 259: 260: 261: 262: 263: 264: 265: 266: 267: 268: 269: 270: 271: 272: 273: 274: 275: 276: 277: 278: 279: 280: 281: 282: 283: 284: 285: 286: 287: 288: 289: 290: 291: 292: 293: 294: 295: 296: 297: 298: 299: 300: 301: 302: 303: 304: 305: 306: 307: 308: 309: 310: 311: 312: 313: 314: 315: 316: 317: 318: 319: 320: 321: 322: 323: 324: 325: 326: 327: 328: 329: 330: 331: 332: 333: 334: 335: 336: 337: 338: 339: 340: 341: 342: 343: 344: 345: 346: 347: 348: 349: 350: 351: 352: 353: 354: 355: 356: 357: 358: 359: 360: 361: 362: 363: 364: 365: 366: 367: 368: 369: 370: 371: 372: 373: 374: 375: 376: 377: 378: 379: 380: 381: 382: 383: 384: 385: 386: 387: 388: 389: 390: 391: 392: 393: 394: 395: 396: 397: 398: 399: 400: 401: 402: 403: 404: 405: 406: 407: 408: 409: 410: 411: 412: 413: 414: 415: 416: 417: 418: 419: 420: 421: 422: 423: 424: 425: 426: 427: 428: 429: 430: 431: 432: 433: 434: 435: 436: 437: 438: 439: 440: 441: 442: 443: 444: 445: 446: 447: 448: 449: 450: 451: 452: 453: 454: 455: 456: 457: 458: 459: 460: 461: 462: 463: 464: 465: 466: 467: 468: 469: 470: 471: 472: 473: 474: 475: 476: 477: 478: 479: 480: 481: 482: 483: 484: 485: 486: 487: 488: 489: 490: 491: 492: 493: 494: 495: 496: 497: 498: 499: 500: </pre>
---	--

Original Text

Changed Text

</code/userprog/addrspace.cc>

主要更改在 Load function.

1. 主要執行程式碼會 allocated to main memory，相關 page table 紀錄。

2. 如果 frame number in main memory 足夠，the pages 會存在 main memory；

如果不夠，最後一個 pages 會寫入 virtual memory by Write Sector.

相關 page table 會被記錄並設定為 invalid.

```
40 // page is referenced or modified.
41 bool dirty; // This bit is set by the hardware every time the
42 // page is modified.
```

Original Text

```
40 // page is referenced or modified.
41 bool dirty; // This bit is set by the hardware every time the
42 // page is modified.
43 int i0;
44
45 int LRU_times; //for Least Recently used algorithm
46
```

Changed Text

`/code/machine/translate.h`
Add ID and LRU_times in class TranslationEntry

```
187 int i;
188 unsigned int vpn, offset;
189 TranslationEntry *entry;
190 unsigned int pageFrame;
```

Original Text

```
216 DEBUG(dbgAddr, "Invalid virtual page # " << virtAddr);
217 return PageFaultException;
```

```
188 int i, j;
189 unsigned int vpn, offset;
190 TranslationEntry *entry;
191 unsigned int pageFrame;
192
193 int Swap_out_page; //find the page Swap_out_page
194
220 printf("Page fault Happen!\n");
221 kernel->stats->numPageFaults += 1; // nachos pagefault counter +1
222 j = 0;
223 while (kernel->machine->Occupied_frame[j] != FALSE && j < NumPhysPages)
224     j += 1; // find valid frame space
225
226 if (j < NumPhysPages) { // if find valid frame space, save the page in virtua
l memory into physical memory
227     char* buffer; //save page temporary
228     buffer = new char[PageSize];
229     pageTable[vpn].physicalPage = j; // save physical memory position
230     pageTable[vpn].valid = TRUE; // page has already in physical memory
231     kernel->machine->Occupied_frame[j] = TRUE;
232     kernel->machine->FrameName[j] = pageTable[vpn].ID;
233     kernel->machine->main_tab[j] = &pageTable[vpn]; // save the page pointer
234
235     pageTable[vpn].LRU_times++; //for LRU
236
237     kernel->Swap_Area->ReadSector(pageTable[vpn].virtualPage, buffer); // rea
d data from swap area
238     bcopy(buffer, &mainMemory[j * PageSize, PageSize]; // save data into phy
sical memory
239 }
240 else {
241     char* buffer1;
242     char* buffer2;
243     buffer1 = new char[PageSize];
244     buffer2 = new char[PageSize];
245
246     //Random
247     //Swap_out_page = (rand()%32);
248
249     //LRU
250
251     int min = pageTable[0].LRU_times;
252     Swap_out_page = 0;
253     for (int cc = 0; cc < 32; cc++) {
254         if (min > pageTable[cc].LRU_times) {
255             min = pageTable[cc].LRU_times;
256             Swap_out_page = cc;
257         }
258     }
259     pageTable[Swap_out_page].LRU_times++;
260
261     printf("PageId swap out!\n", Swap_out_page);
262     bcopy(&mainMemory[Swap_out_page * PageSize, buffer1, PageSize);
263     kernel->Swap_Area->ReadSector(pageTable[vpn].virtualPage, buffer2);
264     bcopy(buffer2, &mainMemory[Swap_out_page * PageSize, PageSize);
265     kernel->Swap_Area->WriteSector(pageTable[vpn].virtualPage, buffer1);
266
267     main_tab[Swap_out_page]->virtualPage = pageTable[vpn].virtualPage;
268     main_tab[Swap_out_page]->valid = FALSE;
269
270     pageTable[vpn].valid = TRUE;
271     pageTable[vpn].physicalPage = Swap_out_page;
272     kernel->machine->FrameName[Swap_out_page] = pageTable[vpn].ID;
273     main_tab[Swap_out_page] = &pageTable[vpn];
274     printf("Finish the page replacement!\n");
275 }
276
```

Changed Text

`/code/machine/translate.cc`

LRU implement.

將 page table 的 valid bit 以確保 page 的位置，i.e. in main memory or in virtual memory.

如果 page 在 virtual memory，檢查 main memory 中現在是否有可用空間。

如果 main memory 沒有空間，將發生 page 替換。

2 buffers 拿來宣告 swap in/out。LRU，使用 for loop 來查找 least used page

靠 implied by bcopy, ReadSector, and WriteSector.完成

讀取 virtual memory，並複製到 main memory，選擇 the swap out page 到 virtual memory

```
63 #define PCReg 34 // Current program counter
64 #define NextPCReg 35 // Next program counter (for branch delay)
65 #define PrevPCReg 36 // Previous program counter (for debugging)
```

Original Text

```
139 int Identity;
140 int SectorNum; //record sector number
141 int FrameName[NumPhysPages];
142 bool Occupied_frame[NumPhysPages]; //record which frame in the main memory is occupie
d.
143 bool Occupied_virpage[NumPhysPages];
144
145 // start for page replacement //
146 int LRU_times[NumPhysPages]; //for LRU
147 bool reference_bit[NumPhysPages]; //for second chance algorithm.
148 // end //
149
150 TranslationEntry *main_tab[NumPhysPages];
151
```

```
67 #define PCReg 34 // Current program LRU_timesen
68 #define NextPCReg 35 // Next program LRU_timesen (for branch delay)
69 #define PrevPCReg 36 // Previous program LRU_timesen (for debugging)
```

Changed Text

`/code/machine/machine.h`

紀錄 sector number, frame paged, 和被 main memory, virtual memory 占用的 frame.

III. Result

Experiment result and some discussion

1. 以 Random replacement 方式同時執行 `/test/matmult.c` 與 `/test/sort.c`

`./nachos -e ../test/matmult -e ../test/sort`

執行時發現 sort 有誤 必須做一個更正

```
r08921005@r08921005-VirtualBox:~/nachos/3/nachos-4.0/code/userprog$ ./nachos -e ../test/matmult -e ../test/sort
Total threads number is 2
Thread ../test/matmult is executing.
Thread ../test/sort is executing.

...

Finish the page replacement!
Page fault Happen!
Page12 swap out!
Finish the page replacement!
Page fault Happen!
Page13 swap out!
Finish the page replacement!
Page fault Happen!
Page14 swap out!
Finish the page replacement!
Page fault Happen!
Page15 swap out!
Finish the page replacement!
return value:7220
No threads ready or runnable, and no pending interrupts.
Assuming the program completed.
Machine halting!

Ticks: total 7691030, idle 1365646, system 6325380, user 4
Disk I/O: reads 80, writes 102
Console I/O: reads 0, writes 0
Paging: faults 80
Network I/O: packets received 0, sent 0
r08921005@r08921005-VirtualBox:~/nachos/3/nachos-4.0/code/userprog$
```

Fig. 2 修改後，由於 virtual address 與 physical address 的維護加上 random replacement 的方式，可以成功執行，並輸出 7220 (matmult) 與 1023(sort)

Conclusion

Page Fault 發生的次數和排程的演算法關係非常的大，這次先使用隨機的方式找到 victim，之後可以再進一步的測試在課堂上有交到過的各種演算法，來測試不同演算法發生 page fault 的次數，來決定最佳的排成方式，找到一個最好的結果。