

```

(gdb)
main () at test.c:11
11      ptr=(int*)malloc(sizeof(int));
(gdb)
malloc (nbytes=4) at umalloc.c:69
69      nunits = (nbytes + sizeof(Header) - 1)/sizeof(Header) + 1;
(gdb)
70      if((prevp = freep) == 0){
(gdb)
69      nunits = (nbytes + sizeof(Header) - 1)/sizeof(Header) + 1;
(gdb)
70      if((prevp = freep) == 0){
(gdb)
71          base.s.ptr = freep = prevp = &base;
(gdb)
72          base.s.size = 0;
(gdb)
86          if(p == freep)
(gdb)
87              if((p = morecore(nunits)) == 0)
(gdb)
morecore (nu=4096) at umalloc.c:54
54      p = sbrk(nu * sizeof(Header));
(gdb)
sbrk () at usys.S:29
29      SYSCALL(sbrk)

```

Figure 1: GDB steps

From the GDB tracing we find that malloc returns to sbrk. For the purpose of the exercise, the parsing portion will be ignored to focus on the concepts of malloc.

```

63 void*
64 malloc(uint nbytes)
65 {
66     Header *p, *prevp;
67     uint nunits;
68
69     nunits = (nbytes + sizeof(Header) - 1)/sizeof(Header) + 1;
70     if((prevp = freep) == 0){
71         base.s.ptr = freep = prevp = &base;
72         base.s.size = 0;
73     }
74     for(p = prevp->s.ptr; ; prevp = p, p = p->s.ptr){
75         if(p->s.size >= nunits){
76             if(p->s.size == nunits)
77                 prevp->s.ptr = p->s.ptr;
78             else {
79                 p->s.size -= nunits;
80                 p += p->s.size;
81                 p->s.size = nunits;
82             }
83             freep = prevp;
84             return (void*)(p + 1);
85         }
86         if(p == freep)
87             if((p = morecore(nunits)) == 0)
88                 return 0;
89     }
90 }

```

Figure 2: umalloc.c malloc

To begin, the user calls on the function malloc from umalloc.c. It is evident that given valid conditions it will adjust the pointers and size of the header. Once adjusted and the new header is valid, it will call on morecore.

```

46 static Header*
47 morecore(uint nu)
48 {
49     char *p;
50     Header *hp;
51
52     if(nu < 4096)
53         nu = 4096;
54     p = sbrk(nu * sizeof(Header));
55     if(p == (char*)-1)
56         return 0;
57     hp = (Header*)p;
58     hp->s.size = nu;
59     free((void*)(hp + 1));
60     return freep;
61 }

```

Figure 3: umalloc.c morecore

morecore then gives the new unit size (minimum 4K for page size) to sbrk.

```

4  #define SYSCALL(name) \
5      .globl name; \
6      name: \
7          movl $SYS_ ## name, %eax; \
8          int $T_SYSCALL; \
9          ret
.....
29 SYSCALL(sbrk)

```

Figure 4: usys.S user sbrk

```

13 #define SYS_sbrk 12

```

Figure 5: syscall.h SYS\_sbrk define

The function sbrk can be found from the usys.S defined library of functions. This utilizes syscall.h to return the syscall number, resulting in the appropriate arguments and registers before performing the trap.

```

25 // These are arbitrarily chosen, but with care not to overlap
26 // processor defined exceptions or interrupt vectors.
27 #define T_SYSCALL 64 // system call

```

Figure 6: traps.h T\_SYSCALL define

The pushed arguments are then compared to traps.h definitions to determine which type it is, in this case a T\_SYSCALL. At this point privilege/level is changed as the trap function changes from user to kernel.

```

14 // Bootstrap processor starts running C code here.
15 // Allocate a real stack and switch to it, first
16 // doing some setup required for memory allocator to work.
17 int
18 main(void)
19 {
20     kinit1(end, P2V(4*1024*1024)); // phys page allocator
21     kvmalloc(); // kernel page table
30     tvinit(); // trap vectors

```

Figure 7: main.c main

```

17 void
18 tvinit(void)
19 {
20     int i;
21
22     for(i = 0; i < 256; i++)
23         SETGATE(idt[i], 0, SEG_KCODE<<3, vectors[i], 0);
24     SETGATE(idt[T_SYSCALL], 1, SEG_KCODE<<3, vectors[T_SYSCALL], DPL_USER);
25
26     initlock(&tickslock, "time");
27 }

```

Figure 8: trap.c tvinit

```

207 // Set up a normal interrupt/trap gate descriptor.
208 // - istrap: 1 for a trap (= exception) gate, 0 for an interrupt gate.
209 // - interrupt gate clears FL_IF, trap gate leaves FL_IF alone
210 // - sel: Code segment selector for interrupt/trap handler
211 // - off: Offset in code segment for interrupt/trap handler
212 // - dpl: Descriptor Privilege Level -
213 //       the privilege level required for software to invoke
214 //       this interrupt/trap gate explicitly using an int instruction.
215 #define SETGATE(gate, istrap, sel, off, d) \
216 { \
217     (gate).off_15_0 = (uint)(off) & 0xffff; \
218     (gate).cs = (sel); \
219     (gate).args = 0; \
220     (gate).rsv1 = 0; \
221     (gate).type = (istrap) ? STS_TG32 : STS_IG32; \
222     (gate).s = 0; \
223     (gate).dpl = (d); \
224     (gate).p = 1; \
225     (gate).off_31_16 = (uint)(off) >> 16; \
226 }

```

Figure 9: mmu.h SETGATE

It is important to note that the trap table has been initialized during the boot process of the system, which allows it so that any point after boot to execute the defined trap functions. This is done with defining the offsets of the trap vectors.

```

50 // Common CPU setup code.
51 static void
52 mpmain(void)
53 {
54     cprintf("cpu%d: starting %d\n", cpuid(), cpuid());
55     idtinit(); // load idt register
56     xchg(&(mycpu()->started), 1); // tell startothers() we're up
57     scheduler(); // start running processes
58 }

```

Figure 10: main.c mpmain

```

29 void
30 idtinit(void)
31 {
32     lidt(idt, sizeof(idt));
33 }

```

Figure 11: trap.c idtinit

```

76 static inline void
77 lidt(struct gatedesc *p, int size)
78 {
79     volatile ushort pd[3];
80
81     pd[0] = size-1;
82     pd[1] = (uint)p;
83     pd[2] = (uint)p >> 16;
84
85     asm volatile("lidt (%0)" : : "r" (pd));
86 }

```

Figure 12: x86.h lidt

The previous set of figures lists how the processor itself (hardware) accesses the table from memory, which in turn allows for some work to be done at the hardware level. While this sequence of trap table tracing was not necessary, it explains priori the access of the trap from memory (and sequentially this occurs before user call as it is set during boot). At this point the trap table is set up, and the tracing of sbrk (and malloc) may continue.

```

317 .globl vector64
318 vector64:
319     pushl $0
320     pushl $64
321     jmp alltraps

```

Figure 13: vectors.S global vector

```

3   # vectors.S sends all traps here.
4   .globl alltraps
5   ~ alltraps:
6       # Build trap frame.
7       pushl %ds
8       pushl %es
9       pushl %fs
10      pushl %gs
11      pushal
12
13      # Set up data segments.
14      movw $(SEG KDATA<<3), %ax
15      movw %ax, %ds
16      movw %ax, %es
17
18      # Call trap(tf), where tf=%esp
19      pushl %esp
20      call trap
21      addl $4, %esp

```

Figure 14: trapasm.S alltraps

Continuing from sbrk when the trap is called, the first thing performed is a push of the trap number followed by a saving of registers to the stack. Afterwards, these registers are adjusted to access kernel memory for trap handling.

```

36 void
37 trap(struct trapframe *tf)
38 {
39     if(tf->trapno == T_SYSCALL){
40         if(myproc()->killed)
41             exit();
42         myproc()->tf = tf;
43         syscall();
44         if(myproc()->killed)
45             exit();
46         return;
47     }

```

Figure 15: trap.c trap

The trap number read 64, or T\_SYSCALL, thus focusing on that segment of the trap function. It is at this point where the table of system calls is referred to for the statements to perform.

```

131 void
132 syscall(void)
133 {
134     int num;
135     struct proc *curproc = myproc();
136
137     num = curproc->tf->eax;
138     if(num > 0 && num < NELEM(syscalls) && syscalls[num]) {
139         curproc->tf->eax = syscalls[num]();
140     } else {
141         cprintf("%d %s: unknown sys call %d\n",
142                 curproc->pid, curproc->name, num);
143         curproc->tf->eax = -1;
144     }
145 }

```

Figure 16: syscall.c syscall

```

107 static int (*syscalls[])(void) = {
108     [SYS_fork]    sys_fork,
109     [SYS_exit]    sys_exit,
119     [SYS_sbrk]    sys_sbrk,

```

Figure 17: syscall.c syscalls[]

In the function it is found that `eax` is referred to, better known as the syscall number (12). In this case, `sys_sbrk` is called and its return value is set back to `eax`.

```

45 int
46 sys_sbrk(void)
47 {
48     int addr;
49     int n;
50
51     if(argint(0, &n) < 0)
52         return -1;
53     addr = myproc()->sz;
54     if(growproc(n) < 0)
55         return -1;
56     return addr;
57 }

```

Figure 18: sysproc.c sys\_sbrk



Now beginning with the syscall for sbrk, the sub-function argomt is done to validate the argument for malloc.

```
48 // Fetch the nth 32-bit system call argument.
49 int
50 argint(int n, int *ip)
51 {
52     return fetchint((myproc()->tf->esp) + 4 + 4*n, ip);
53 }
```

Figure 19: syscall.c argint

```
16 // Fetch the int at addr from the current process.
17 int
18 fetchint(uint addr, int *ip)
19 {
20     struct proc *curproc = myproc();
21
22     if(addr >= curproc->sz || addr+4 > curproc->sz)
23         return -1;
24     *ip = *(int*)(addr);
25     return 0;
26 }
```

Figure 20: syscall.c fetchint

As provided with in-code documentation, it pulls the process state (myproc is the structure) and fetches the nth 32-bit syscall argument. In fetchint, it validates the address for bounds and size before returning (rewriting the pointer). Now exploring the myproc:

```
55 // Disable interrupts so that we are not rescheduled
56 // while reading proc from the cpu structure
57 struct proc*
58 myproc(void) {
59     struct cpu *c;
60     struct proc *p;
61     pushcli();
62     c = mycpu();
63     p = c->proc;
64     popcli();
65     return p;
66 }
```

Figure 21: proc.c myproc

```

96 // Pushcli/popcli are like cli/sti except that they are matched:
97 // it takes two popcli to undo two pushcli. Also, if interrupts
98 // are off, then pushcli, popcli leaves them off.
99
100 void
101 pushcli(void)
102 {
103     int eflags;
104
105     eflags = readeflags();
106     cli();
107     if(mycpu()->ncli == 0)
108         mycpu()->intena = eflags & FL_IF;
109     mycpu()->ncli += 1;
110 }

```

Figure 22: spinlock.c pushcli

```

94 static inline uint
95 readeflags(void)
96 {
97     uint eflags;
98     asm volatile("pushfl; popl %0" : "=r" (eflags));
99     return eflags;
100 }

```

Figure 23: x86.h readeflags

```

108 static inline void
109 cli(void)
110 {
111     asm volatile("cli");
112 }

```

Figure 24: x86.h cli

```

35 // Must be called with interrupts disabled to avoid the caller being
36 // rescheduled between reading lapicid and running through the loop.
37 struct cpu*
38 mycpu(void)
39 {
40     int apicid, i;
41
42     if(readeflags() & FL_IF)
43         panic("mycpu called with interrupts enabled\n");
44
45     apicid = lapicid();
46     // APIC IDs are not guaranteed to be contiguous. Maybe we should have
47     // a reverse map, or reserve a register to store &cpus[i].
48     for (i = 0; i < ncpu; ++i) {
49         if (cpus[i].apicid == apicid)
50             return &cpus[i];
51     }
52     panic("unknown apicid\n");
53 }

```

Figure 25: proc.h cpu

```

100 int
101 lapicid(void)
102 {
103     if (!lapic)
104         return 0;
105     return lapic[ID] >> 24;
106 }

```

Figure 26: lapic.c lapicid

```

112 void
113 popcli(void)
114 {
115     if(readeflags() & FL_IF)
116         panic("popcli - interruptible");
117     if(--mycpu()->ncli < 0)
118         panic("popcli");
119     if(mycpu()->ncli == 0 && mycpu()->intena)
120         sti();
121 }

```

Figure 27: spinlock.c popcli

```

114 static inline void
115 sti(void)
116 {
117     asm volatile("sti");
118 }

```

Figure 28: x86.h sti

When calling myproc, pushcli disables interrupts to allow mycpu to retrieve the local APIC, which in turn retrieves the desired process. popcli ends this set of actions by removing the paired pushcli.

Returning to sys\_sbrk (Figure 18), the next performed function is growproc.

```

156 // Grow current process's memory by n bytes.
157 // Return 0 on success, -1 on failure.
158 int
159 growproc(int n)
160 {
161     uint sz;
162     struct proc *curproc = myproc();
163
164     sz = curproc->sz;
165     if(n > 0){
166         if((sz = allocuvm(curproc->pgdir, sz, sz + n)) == 0)
167             return -1;
168     } else if(n < 0){
169         if((sz = deallocuvm(curproc->pgdir, sz, sz + n)) == 0)
170             return -1;
171     }
172     curproc->sz = sz;
173     switchuvm(curproc);
174     return 0;
175 }

```

Figure 29: proc.c growproc

As visible, it will either allocate or deallocate n bytes from the process through the two sub-functions.

```

219 // Allocate page tables and physical memory to grow process from oldsz to
220 // newsz, which need not be page aligned. Returns new size or 0 on error.
221 int
222 allocuvm(pde_t *pgdir, uint oldsz, uint newsz)
223 {
224     char *mem;
225     uint a;
226
227     if(newsz >= KERNBASE)
228         return 0;
229     if(newsz < oldsz)
230         return oldsz;
231
232     a = PGROUNDUP(oldsz);
233     for(; a < newsz; a += PGSIZE){
234         mem = kalloc();
235         if(mem == 0){
236             cprintf("allocuvm out of memory\n");
237             deallocuvm(pgdir, newsz, oldsz);
238             return 0;
239         }
240         memset(mem, 0, PGSIZE);
241         if(mappages(pgdir, (char*)a, PGSIZE, V2P(mem), PTE_W|PTE_U) < 0){
242             cprintf("allocuvm out of memory (2)\n");
243             deallocuvm(pgdir, newsz, oldsz);
244             kfree(mem);
245             return 0;
246         }
247     }
248     return newsz;
249 }

```

Figure 30: vm.c allocuvm

```

131 #define PGROUNDUP(sz)  (((sz)+PGSIZE-1) & ~(PGSIZE-1))

```

Figure 31: mmu.h PGROUNDUP

Simple rounds up the page table given an argument.

```
79 // Allocate one 4096-byte page of physical memory.
80 // Returns a pointer that the kernel can use.
81 // Returns 0 if the memory cannot be allocated.
82 char*
83 kalloc(void)
84 {
85     struct run *r;
86
87     if(kmem.use_lock)
88         acquire(&kmem.lock);
89     r = kmem.freelist;
90     if(r)
91         kmem.freelist = r->next;
92     if(kmem.use_lock)
93         release(&kmem.lock);
94     return (char*)r;
95 }
```

Figure 32: kalloc.c kalloc

```

20 // Acquire the lock.
21 // Loops (spins) until the lock is acquired.
22 // Holding a lock for a long time may cause
23 // other CPUs to waste time spinning to acquire it.
24 void
25 acquire(struct spinlock *lk)
26 {
27     pushcli(); // disable interrupts to avoid deadlock.
28     if(holding(lk))
29         panic("acquire");
30
31     // The xchg is atomic.
32     while(xchg(&lk->locked, 1) != 0)
33         ;
34
35     // Tell the C compiler and the processor to not move loads or stores
36     // past this point, to ensure that the critical section's memory
37     // references happen after the lock is acquired.
38     __sync_synchronize();
39
40     // Record info about lock acquisition for debugging.
41     lk->cpu = mycpu();
42     getcallerpcs(&lk, lk->pcs);
43 }

```

Figure 33: spinlock.c acquire

```

88 // Check whether this cpu is holding the lock.
89 int
90 holding(struct spinlock *lock)
91 {
92     return lock->locked && lock->cpu == mycpu();
93 }

```

Figure 34: spinlock.c holding

```

120 static inline uint
121 xchg(volatile uint *addr, uint newval)
122 {
123     uint result;
124
125     // The + in "+m" denotes a read-modify-write operand.
126     asm volatile("lock; xchgl %0, %1" :
127                 "+m" (*addr), "=a" (result) :
128                 "1" (newval) :
129                 "cc");
130     return result;
131 }

```

Figure 35: x86.h xchg

```

70 // Record the current call stack in pcs[] by following the %ebp chain.
71 void
72 getcallerpcs(void *v, uint pcs[])
73 {
74     uint *ebp;
75     int i;
76
77     ebp = (uint*)v - 2;
78     for(i = 0; i < 10; i++){
79         if(ebp == 0 || ebp < (uint*)KERNBASE || ebp == (uint*)0xffffffff)
80             break;
81         pcs[i] = ebp[1]; // saved %eip
82         ebp = (uint*)ebp[0]; // saved %ebp
83     }
84     for(; i < 10; i++)
85         pcs[i] = 0;
86 }

```

Figure 36: spinlock.c getcallerpcs



```

45 // Release the lock.
46 void
47 release(struct spinlock *lk)
48 {
49     if(!holding(lk))
50         panic("release");
51
52     lk->pcs[0] = 0;
53     lk->cpu = 0;
54
55     // Tell the C compiler and the processor to not move loads or stores
56     // past this point, to ensure that all the stores in the critical
57     // section are visible to other cores before the lock is released.
58     // Both the C compiler and the hardware may re-order loads and
59     // stores; __sync_synchronize() tells them both not to.
60     __sync_synchronize();
61
62     // Release the lock, equivalent to lk->locked = 0.
63     // This code can't use a C assignment, since it might
64     // not be atomic. A real OS would use C atomics here.
65     asm volatile("movl $0, %0" : "+m" (lk->locked) : );
66
67     popcli();
68 }

```

Figure 37: spinlock.c release

The process of kalloc is extensive, in which the process itself will allocate a page of memory and return the pointer. The sub-function acquire will wait for a lock (checks through holding) before continuing to manipulate the memory through moving the free pointer to next. The lock is then released and pointer is returned. In short, kalloc retrieves the free-list pointer, moves the head of said pointer to next and returns the pointer for usage.

Continuing in allocuvm (Figure 30), assuming valid memory it will allocate memory through memset:

```

4 void*
5 memset(void *dst, int c, uint n)
6 {
7     if ((int)dst%4 == 0 && n%4 == 0){
8         c &= 0xFF;
9         stosl(dst, (c<<24)|(c<<16)|(c<<8)|c, n/4);
10    } else
11        stosb(dst, c, n);
12    return dst;
13 }

```

Figure 38: string.c memset

```

51 static inline void
52 stosl(void *addr, int data, int cnt)
53 {
54     asm volatile("cld; rep stosl" :
55                 "=D" (addr), "=c" (cnt) :
56                 "0" (addr), "1" (cnt), "a" (data) :
57                 "memory", "cc");
58 }

```

Figure 39: x86.h stosl

```

42 static inline void
43 stosb(void *addr, int data, int cnt)
44 {
45     asm volatile("cld; rep stosb" :
46                 "=D" (addr), "=c" (cnt) :
47                 "0" (addr), "1" (cnt), "a" (data) :
48                 "memory", "cc");
49 }

```

Figure 40: x86.h stosb

memset validates the destination against size, in which depending on condition will store the appropriate string (long vs byte).

Returning to allocuvn (Figure 30) are conditions to check allocuvn through sub-functions mappages and V2P.

```

57 // Create PTEs for virtual addresses starting at va that refer to
58 // physical addresses starting at pa. va and size might not
59 // be page-aligned.
60 static int
61 mappages(pde_t *pgdir, void *va, uint size, uint pa, int perm)
62 {
63     char *a, *last;
64     pte_t *pte;
65
66     a = (char*)PGROUNDDOWN((uint)va);
67     last = (char*)PGROUNDDOWN(((uint)va) + size - 1);
68     for(;;){
69         if((pte = walkpgdir(pgdir, a, 1)) == 0)
70             return -1;
71         if(*pte & PTE_P)
72             panic("remap");
73         *pte = pa | perm | PTE_P;
74         if(a == last)
75             break;
76         a += PGSIZE;
77         pa += PGSIZE;
78     }
79     return 0;
80 }

```

Figure 41: vm.c mappages

The function creates page table entries by walking the table by increasing the page size every iteration.

```

32 // Return the address of the PTE in page table pgdir
33 // that corresponds to virtual address va. If alloc!=0,
34 // create any required page table pages.
35 static pte_t *
36 walkpgdir(pde_t *pgdir, const void *va, int alloc)
37 {
38     pde_t *pde;
39     pte_t *pgtab;
40
41     pde = &pgdir[PDX(va)];
42     if(*pde & PTE_P){
43         pgtab = (pte_t*)P2V(PTE_ADDR(*pde));
44     } else {
45         if(!alloc || (pgtab = (pte_t*)kalloc()) == 0)
46             return 0;
47         // Make sure all those PTE_P bits are zero.
48         memset(pgtab, 0, PGSIZE);
49         // The permissions here are overly generous, but they can
50         // be further restricted by the permissions in the page table
51         // entries, if necessary.
52         *pde = V2P(pgtab) | PTE_P | PTE_W | PTE_U;
53     }
54     return &pgtab[PTX(va)];
55 }

```

Figure 42: vm.c walkpgdir

Returns valid addresses of the page table corresponding to virtual addresses through validation of the page indices.

```

113 // page directory index
114 #define PDX(va) (((uint)(va) >> PDXSHIFT) & 0x3FF)

```

Figure 43: mmu.h PDX

```

11 #define V2P(a) (((uint) (a)) - KERNBASE)
12 #define P2V(a) (((void *) (a)) + KERNBASE)

```

Figure 44: memlayout.h V2P and P2V

```

145 // Address in page table or page directory entry
146 #define PTE_ADDR(pte) ((uint)(pte) & ~0xFFF)

```

Figure 45: mmu.h PTE\_ADDR

The three functions simply return the page index (PDX given virtual address), or the desired address given the opposite. The address itself is returned through PTE\_ADDR.

Returning to allocuvm (Figure 30), in the event that the condition is true or growproc was given a negative argument, the function proceeds to deallocuvm.

```
251 // Deallocate user pages to bring the process size from oldsz to
252 // newsz. oldsz and newsz need not be page-aligned, nor does newsz
253 // need to be less than oldsz. oldsz can be larger than the actual
254 // process size. Returns the new process size.
255 int
256 deallocuvm(pde_t *pgdir, uint oldsz, uint newsz)
257 {
258     pte_t *pte;
259     uint a, pa;
260
261     if(newsz >= oldsz)
262         return oldsz;
263
264     a = PGROUNDUP(newsz);
265     for(; a < oldsz; a += PGSIZE){
266         pte = walkpgdir(pgdir, (char*)a, 0);
267         if(!pte)
268             a = PGADDR(PDX(a) + 1, 0, 0) - PGSIZE;
269         else if((*pte & PTE_P) != 0){
270             pa = PTE_ADDR(*pte);
271             if(pa == 0)
272                 panic("kfree");
273             char *v = P2V(pa);
274             kfree(v);
275             *pte = 0;
276         }
277     }
278     return newsz;
279 }
```

Figure 46: vm.c deallocuvm

Similar to allocuvm with walking, with the exception of kfree instead of increasing size.

```

55 // Free the page of physical memory pointed at by v,
56 // which normally should have been returned by a
57 // call to kalloc(). (The exception is when
58 // initializing the allocator; see kinit above.)
59 void
60 kfree(char *v)
61 {
62     struct run *r;
63
64     if((uint)v % PGSIZE || v < end || V2P(v) >= PHYSTOP)
65         panic("kfree");
66
67     // Fill with junk to catch dangling refs.
68     memset(v, 1, PGSIZE);
69
70     if(kmem.use_lock)
71         acquire(&kmem.lock);
72     r = (struct run*)v;
73     r->next = kmem.freelist;
74     kmem.freelist = r;
75     if(kmem.use_lock)
76         release(&kmem.lock);
77 }

```

Figure 47: kalloc.c kfree

The opposite of kalloc, freeing the memory and changing the freelist accordingly.

Assuming that de/allocuvn was successful, it will return the new size to growproc (Figure 29), which will set the new size and execute switchuvn.

```

155 // Switch TSS and h/w page table to correspond to process p.
156 void
157 switchvm(struct proc *p)
158 {
159     if(p == 0)
160         panic("switchvm: no process");
161     if(p->kstack == 0)
162         panic("switchvm: no kstack");
163     if(p->pgdir == 0)
164         panic("switchvm: no pgdir");
165
166     pushcli();
167     mycpu()->gdt[SEG_TSS] = SEG16(STS_T32A, &mycpu()->ts,
168                                 sizeof(mycpu()->ts)-1, 0);
169     mycpu()->gdt[SEG_TSS].s = 0;
170     mycpu()->ts.ss0 = SEG_KDATA << 3;
171     mycpu()->ts.esp0 = (uint)p->kstack + KSTACKSIZE;
172     // setting IOPL=0 in eflags *and* iomb beyond the tss segment limit
173     // forbids I/O instructions (e.g., inb and outb) from user space
174     mycpu()->ts.iomb = (ushort) 0xFFFF;
175     ltr(SEG_TSS << 3);
176     lcr3(V2P(p->pgdir)); // switch to process's address space
177     popcli();
178 }

```

Figure 48: vm.c switchvm

```

88 static inline void
89 ltr(ushort sel)
90 {
91     asm volatile("ltr %0" : : "r" (sel));
92 }

```

Figure 49: x86.h ltr

```

141 static inline void
142 lcr3(uint val)
143 {
144     asm volatile("movl %0,%%cr3" : : "r" (val));
145 }

```

Figure 50: x86.h lcr3

Per documentation, switches the tables so to correspond to correspond to the process, done by manipulating mycpu structure.

Afterwards, growproc returns and begins recursively returning back to user. The returned integer (0 success, 1 fail) is passed as the eax variable and returned to trap.

```
23  # Return falls through to trapret...
24  .globl trapret
25  trapret:
26      popal
27      popl %gs
28      popl %fs
29      popl %es
30      popl %ds
31      addl $0x8, %esp # trapno and errcode
32      iret
```

Figure 51: trapasm.S trapret

Everything from the stack prior to the trap call is popped, the privilege/level is set back to user, and the instruction returns back to main. The process results in either a changed process size or a returned error which can be handled accordingly. This returns up to morecore (Figure 3) and validates this pointer and returns 0 or the header freep accordingly.