iOS 6 Kernel Security: A Hacker's Guide

by Mark Dowd and Tarjei Mandt mdowd@azimuthsecurity.com tm@azimuthsecurity.com



Introduction

- iOS 6 recently released
- Large focus on security improvements particularly kernel hardening
- Primarily targets strategies employed in "jailbreaks"
- This talk provides an overview of the new kernel-based mitigations
- Explores new techniques for attacking iOS 6



Topics Covered

- Part 1 Defense
 - Heap Hardening Strategies
 - Stack Cookies
 - Information Leaking Mitigations
 - Address Space Layout Randomization (ASLR)
 - User/Kernel address space hardening
- Part 2 Offense
 - Information Leaking
 - Heap Strategies



Randomization Algorithm

- First, a word on randomness...
- Used to derive random numbers for stack cookie, heap cookies, kernel map ASLR, and pointer obfuscation
- Random seed generated (or retrieved) during boot loading (iBoot)
- Combined with current time to get random value



Randomization Algorithm

```
unsigned long long
GetRandomValue(unsigned long long time, unsigned long long seed)
   unsigned int time low
                               = time & 0xFFFFFFF;
   unsigned int time high = (time >> 32) & 0xFFFFFFFF;
   unsigned int result low;
   unsigned int result high;
   unsigned int tmp;
    // calculate low DWORD of output
    tmp = (time low & 0xFF) << 8;
    result low = (time low ^ tmp) ^ (time low << 16);
    // calculate high DWORD of output
    tmp = (seed \& 0xFF) << 16;
    result high = tmp ^ (result low ^ time high);
    tmp = (result_low >> 8) ^ 0xFF;
    result high = ROTATE RIGHT(result high, tmp);
    // done
    return (result high << 32) | result low;
```



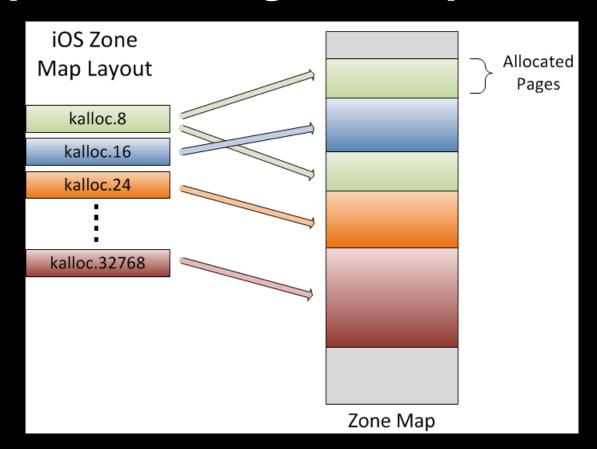
Heap Hardening

- Heap has been hardened to prevent wellknown attack strategies
- Three mitigations put in place
 - Pointer validation
 - Block poisoning
 - Freelist integrity verification
- Specific to the zone allocator (zalloc(), used by kalloc(), MALLOC(), MALLOC_ZONE())



- Quick recap of old exploitation techniques required
 - Covered in the past extensively by Stefan Esser,
 Nemo, probably others
- Zone allocations divided in to fixed-size zones (kalloc.8, kalloc.16, ... kalloc.32768)
 - Specialized zones also utilized for specific tasks (eg. Pmap_zone, vm_map_copy_zone, etc)
- Zone allocates more pages on demand

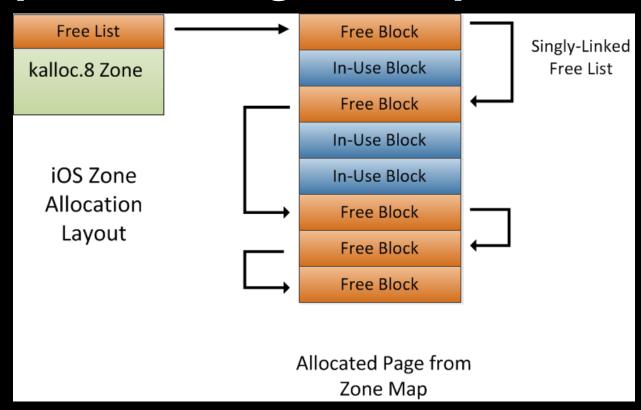






- Zone allocates blocks of pages on demand
 - Divides memory in to element-size blocks
 - All blocks initially added to zone's free list
- Zone free list maintained as singly linked list
 - First DWORD of free block overwritten with "next" pointer when it is freed
- Allocations simply remove elements from the free list







- Previous exploitation techniques rely on overwriting free list pointers in free blocks
 - Future allocation can return arbitrary memory block
- Typical strategy: Add a pointer to sysent
 - Add new system call
 - Invoke new system call
 - Profit



Heap Hardening – Pointer Validation

- Goal: Prevent invalid pointers being entered in to kalloc() zone's freelist
- Additional checks performed on pointers passed to zfree()
 - Also performed as part of validation on pointers in freelist during allocation (zalloc())



Heap Hardening – Pointer Validation

- Pointer verified to be in kernel memory (0x80000000 < ptr < 0xFFFEFFF)
- If allows_foreign is set in zone, no more validation performed
 - Currently event_zone,vm_map_entry_reserved_zone, vm_page_zone
- If pointer is within kernel image, allow (??)
- Otherwise, ensure pointer is within zone_map



Heap Hardening – Block Poisoning

- Goal: Prevent UAF-style attacks
- Strategy involves filling blocks with sentinel value (0xdeadbeef) when being freed
 - Performed by add_to_zone() called from zfree()
- Only performed on selected blocks
 - Block sizes smaller than cache line size of processor (e.g. 32 bytes on A5/A5X devices)
 - Can override with "-zp", "-no-zp", "zp-factor" boot parameters

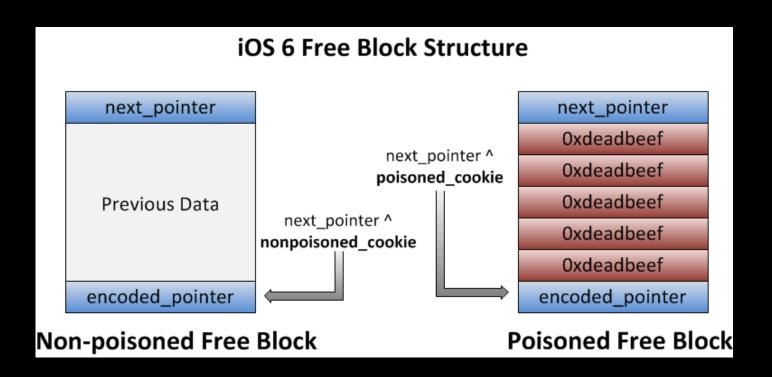


- Goal: Prevent heap overwrites from being exploitable
- Two random values generated at boot time (zone_bootstrap())
 - 32-bit cookie for "poisoned blocks"
 - 31-bit cookie for "non-poisoned blocks"
 - Low bit is clear
- Values serve as validation cookies



- Freelist pointers at the top of a free block are now validated by zalloc()
 - Work performed by alloc_from_zone()
- Encoded next pointer placed at end of block
 - XOR'd with poisoned_cookie or nonpoisoned_cookie







- zalloc() ensures next_pointer matches encoded pointer at end of block
 - Tries both cookies
 - If poisoned cookie matches, check whole block for modification of sentinel (0xdeadbeef) values
 - Cause kernel panic if either check fails
- Next pointer and cookie replaced by 0xdeadbeef when allocated
 - Possible information leak protection



Heap Hardening – Primitives

- OSUnserializeXML() could previously be used to perform kernel heap feng shui
 - Technique presented by Stefan Esser in «iOS Kernel Heap Armageddon» at SyScan 2012
- Allowed precise allocation and freeing of kalloc zone data
- Also possible to force persistent allocations by wrapping the reference count



Heap Hardening - Primitives



Heap Hardening - Primitives

- Duplicate dictionary keys no longer result in freeing of the original key/value
- Dictionary entries can no longer be pinned to memory using multiple references
- In both cases, the plist dictionary is considered invalid

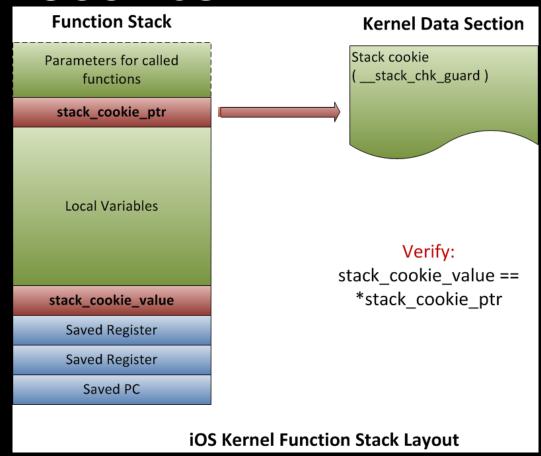


- Goal: Prevent stack overflow exploitation
- Only applied to functions with structures/buffers
- Random value generated during early kernel initialization (arm_init())
- 24-bit random value
 - 32-bit value really, but 2nd byte zeroed out
 - Presumably string copy prevention



- Generated stack cookie placed directly after saved registers at bottom of stack frame
- Pointer to cookie saved at top of stack frame
 - Or in a register if convenient
 - Space above stack cookie pointer used for called functions if necessary







- Function epilog verifies saved stack cookie
 - Generated value found by following saved pointer
- Verification failure results in kernel panic

```
__TEXT:__text:800051FC __epilog
                                                                   CODE XREF: sub_80004F98+2B41j
__TEXT:__text:800051FC
                                                                   sub_80004F98+486⊥ j
__TEXT:___text:800051FC
                                        LDR
                                                 RO, [SP,#0x2CC+stack_cookie_ptr]
__TEXT:__text:800051FE
                                        LDR
                                                 RO, [RO]
__TEXT:__text:80005200
                                                 R1, [SP,#0x2CC+stack_cookie]
                                        LDR
__TEXT:__text:80005202
                                        CMP
                                                 R0. R1
                                                                  ; check stack cookie validity
__TEXT:__text:80005204
                                        ITTTT EQ
__TEXT:__text:80005206
                                        MOVEQ
                                                 RO. R4
__TEXT:__text:80005208
                                        ADDEQ.W SP, SP, #0x2B4
__TEXT:__text:8000520C
                                        POPEQ.W {R8,R10,R11}
__TEXT:__text:80005210
                                        POPEQ
                                                 {R4-R7, PC}
__TEXT:__text:80005212
                                        BL
                                                 ___stack_chk_fail
```



Goals:

- Prevent disclosure of kernel base
- Prevent disclosure of kernel heap addresses

Strategies:

- Disables some APIs
- Obfuscate kernel pointers for some APIs
- Zero out pointers for others



- Previous attacks relied on zone allocator status disclosure
 - host_zone_info() / mach_zone_info()
 - Stefan Esser described using this for heap "feng shui" (https://media.blackhat.com/bh-us-11/Esser/BH_US_11_Esser_Exploiting_The_iOS_Kernel_Slides.pdf)
- APIs now require PE_i_can_has_debugger() access



- Several APIs disclose kernel object pointers
 - mach_port_kobject()
 - mach_port_space_info()
 - vm_region_recurse() / vm_map_region_recurse()
 - vm_map_page_info()
 - proc_info (PROC_PIDREGIONINFO, PROC_PIDREGIONPATHINFO, PROC_PIDFDPIPEINFO, PROC_PIDFDSOCKETINFO, PROC_PIDFILEPORTSOCKETINFO)
 - fstat() (when querying pipes)
 - sysctl(net.inet.*.pcblist)



- Need these APIs for lots of reasons
 - Often, underlying APIs rather than exposed ones listed previously
- Strategy: Obfuscate pointers
 - Generate 31 bit random value at boot time
 - lowest bit always 1
 - Add random value to real pointer



```
int
fill_pipeinfo(struct pipe * cpipe, struct pipe_info * pinfo)
{
    ... code ...
    pinfo->pipe_handle = (uint64_t)((uintptr_t)cpipe);
    pinfo->pipe_peerhandle = (uint64_t)((uintptr_t)(cpipe->pipe_peer));
    pinfo->pipe_status = cpipe->pipe_state;

PIPE_UNLOCK(cpipe);
    return (0);
}
```



```
_TEXT:__text:801EC942
                                                        R1, (heap_random_value_ptr - 0x801EC94E); heap_random_value_ptr
                                       MOU
__TEXT:__text:801EC94A
                                       ADD
                                                        R1, PC; heap_random_value_ptr
TEXT: __text:801EC94C
                                       LDR
                                                        R1, [R1]
TEXT:__text:801EC94E
                                       STR
                                                        R3, [R5,#0×20]
                                                        R6, R3, #0x1F
_TEXT:__text:801EC950
                                       ASRS
_TEXT:__text:801EC952
                                       STR
                                                        R6, [R5,#0×24]
_TEXT:__text:801EC954
                                       STMIA.W
                                                        R4, {R0,R2,R3,R6}
_TEXT:__text:801EC958
                                       ADD.W
                                                        R4, R5, #0x38
_TEXT:__text:801EC95C
                                                        R4, {R0,R2,R3,R6}
                                       STMIA.W
_TEXT:__text:801EC960
                                       MOUS
                                                        RO, #0
_TEXT:__text:801EC962
                                       LDR
                                                        R2, [R1]; R2 = heap_random_value
                                                        R6, R2, R8; R6 = (unsigned long) cpipe + heap_random_value;
_TEXT:__text:801EC964
                                       ADD.W
                                                        R6, [R5,#0x88]; set pipe_handle to cpipe + heap_random_value
_TEXT:__text:801EC968
                                       STR.W
_TEXT:__text:801EC96C
                                       STR.W
                                                        RO, [R5, #0x8C]; set pipe_handle_peer to NULL
TEXT:__text:801EC970
                                                        R2, [R8, #0x30]
                                       LDR.W
```



- Other APIs disclose pointers unnecessarily
 - Zero them out
- Used to mitigate some leaks via sysctl
 - Notably, known proc structure infoleak



Kernel ASLR

- Goal: Prevent attacker's from modifying/utilizing data at known (fixed) addresses
- Strategy is two-fold
 - Randomize kernel image base
 - Randomize base of kernel_map (sort of)



Kernel ASLR – Kernel Image

- Kernel base randomized by boot loader (iBoot)
 - Random data generated
 - SHA-1 hash of data taken
 - Byte from SHA-1 hash used to calculate kernel "slide"
- Kernel is rebased using the formula: 0x01000000 + (slide_byte * 0x00200000)
 - If slide is 0, static offset of 0x21000000 is used



Kernel ASLR – Kernel Image

```
ROM:5FF19CF8
ROM:5FF19CF8 loc_5FF19CF8
                                                       ; CODE XREF: Image3_RelocateImage+3A<sup>†</sup>j
                                       RO, SP, #0x3C+slide; address of output buffer
ROM:5FF19CF8
                              ADD
                              MOUS
ROM:5FF19CFA
                                       R4, #0
ROM:5FF19CFC
                              MOUS
                                                       ; number of random bytes required
                                       R1. #1
                              STRB.W
ROM:5FF19CFE
                                      R4, [SP,#0x3C+slide]
ROM:5FF19D02
                              MOU
                                       R5, R9
ROM:5FF19D04
                                       iBoot_GetRandomBytes
ROM:5FF19D08
                              CBZ
                                       R0, loc_5FF19D10
ROM:5FF19D0A
                              LDR
                                       RO: =0x5FF44EB8
                              STR
ROM:5FF19D0C
                                       R4, [R0]
                                                       ; failed to generate random bytes, just make slide 0
ROM:5FF19D0E
                                       loc_5FF19D28
ROM:5FF19D10
ROM:5FF19D10
                                                       ; CODE XREF: Image3_RelocateImage+54fj
ROM:5FF19D10 loc_5FF19D10
ROM:5FF19D10
                              LDR
                                       R2, =0x5FF44EB8
ROM:5FF19D12
                              MOV.W
                                       RO. #0×21000000
ROM:5FF19D16
                                      R1, [SP,#0x3C+slide]
ROM:5FF19D1A
                              CMP
                                       R1, #0
                              ITT NE
ROM:5FF19D1C
ROM:5FF19D1E
                              MOUNE.W RO.
                                          #0×1000000
ROM:5FF19D22
                              ADDNE.W R0, R0, R1,LSL#21 ; slide = (slide_byte << 21) + 0x00100000
ROM:5FF19D26
                              STR
                                                       ; store slide value for later use
```



Kernel ASLR – Kernel Image

- Calculated value added to kernel preferred base later on
- Result:
 - Kernel can be rebased at 1 of 256 possible locations
 - Base addresses are 2MB apart
 - Example: 0x81200000, 0x81400000, ... 0xA1000000
- Adjusted base passed to kernel in boot args structure (offset 0x04)



- Used for kernel allocations of all types
 - kalloc(), kernel_memory_allocate(), etc
- Spans all of kernel space (0x80000000 -> 0xFFFEFFF)
- Kernel-based maps are submaps of kernel_map
 - zone_map, ipc_kernel_map, etc

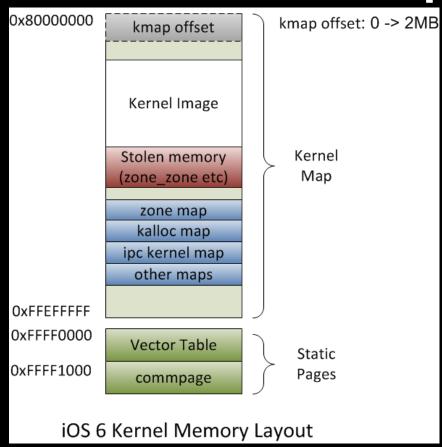


- Strategy involves randomizing the base of kernel_map
 - Random 9-bit value generated right after kmem_init() (which establishes kernel_map)
 - Multiplied by page size
 - Resulting value used as size for initial kernel_map allocation
 - 9 bits = 512 different allocation size possibilities



- Future kernel_map (including submap) allocations pushed forward by random amount
 - Allocation silently removed after first garbage collection (and reused)
- Behavior can be overridden with "kmapoff" boot parameter





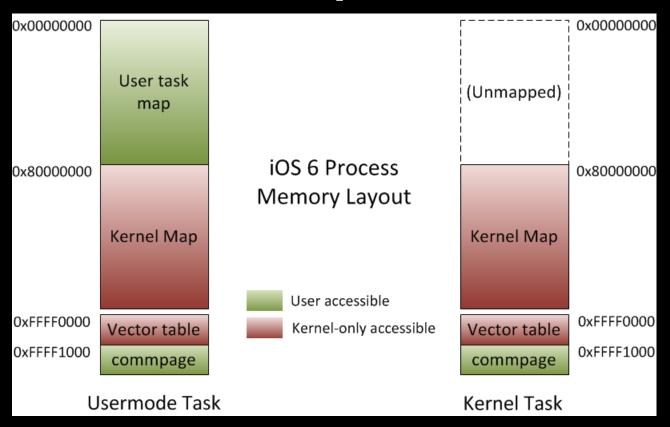


- Goal: Prevent NULL/offset-to-NULL dereference vulnerabilities
- Previously, kernel mapped in to user-mode address space
- NULL-dereferences were prevented by forcing binaries to have __PAGE_ZERO section
 - Does not prevent offset-to-NULL problems



- kernel_task now has its own address space while executing
 - Transitioned to with interrupt handlers
 - Switched between during copyin() / copyout()
- User-mode pages therefore not accessible while executing in kernel mode
- Impossible to accidentally access them







- BUG iOS 5 and earlier had very poor user/kernel validation in copyin() / copyout()
 - Only validation: usermode pointer < 0x80000000
 - Length not validated
- Pointer + length can be > 0x80000000 (!)
 - Can potentially read/write to kernel memory
- Limitation: Device must have > 512M to map 0x7FFF000
 - iPad 3 / iPhone 5



```
EXPORT _copyout
                                           ; CODE XREF: sub_8000D64C+Cfp
_copyout
                                           ; __TEXT:__text:800132F6<sup>†</sup>p ...
                CMP
                                  R2. #0
                                  RO. #0
                 MOUEQ
                 BXEQ
                                  LR
                 CMP
                                  R1. #0x80000000
                 BCS
                                  _error
                 STMFD
                                  SP!, {R4}
                 ADR
                                  R3, fault_handler_routine
                                  p15, 0, R12,c13,c0, 4
                 MRC
                                  R4, [R12,#0x220]
                 LDR
                 STR
                                  R3, [R12,#0x220]
                CMP
                                  R2. #0×10
                 BLT
                                  butewise_copu
                                  R3, R0, R1
                 ORR
                                  R3. #3
                 TST
                                  butewise_copy
                 BNE
                                  R2, R2, #8
                 SUB
```



- iOS 6 added security checks
 - Integer overflow/signedness checks
 - Conservative maximum length
 - Pointer + length < 0x80000000</p>
- iOS 6 still vulnerable!
 - If copy length <= 0x1000, pointer + length check not performed
 - Can read/write to first page of kernel memory



```
; CODE XREF: sub_8000E490+CTp
_copyout
                                           ; __mach_trap_vm_allocate+4ATp ...
var_8
                 = -8
                 CMP
                                  R2, #0
                                  RO, #0
                 MOVEQ
                 BXEQ
                                  LR
                 CMP
                                 R1, #0x80000000
                 BCS
                                  loc_80088278
                                  R2, #0x1000
                 CMP
                 BLS
                                  do copy
                 STMFD
                                  SP!, {R4-R7,LR}
                 MOU
                                  R4, R0
                                  R5, R1
                 MOU
                                  R6, R2
                 MOU
                                  R7, SP, #0x14+var_8
                 ADD
                 BLX
                                  copy_validate
                 CMP
                                  RO, #0
                 LDMNEFD
                                  SP!, {R4-R7,PC}
                 MOV
                                  R0, R4
                 MOU
                                  R1, R5
                 MOV
                                  R2, R6
                                  SP!, {R4-R7,LR}
                 LDMFD
```



- Is anything in the first page of memory?
 - Initially contains kmap offset allocation, but that is removed after first garbage collection
 - Some things allocate to kernel map directly
 - HFS
 - kalloc() blocks of >= 256k
- Create a pipe, specify buffers > 0x7FFFF000
- Bonus: If memory is not mapped, kernel will not panic (safely return EFAULT)



- Memory is no longer RWX
 - Kernel code cannot be directly patched
 - Heap is non-executable
 - Stack is non-executable



Kernel Attacks: Overview

- Protections kill most of the known attack strategies
 - Syscall table overwrites
 - Patching kernel code
 - Attacking key data structures (randomized locations)
- Need something new!



Kernel Attacks: Overview

- Generally, exploit will require information leaking followed by corruption
- Corruption primitives dictate strategy
 - Write in to adjacent buffer (overflow)
 - Write to relative location from buffer
 - Write to arbitrary location
- Different types of primitives will be considered separately



- Leaking the kernel base is really useful
- Kext_request() allows applications to request information about kernel modules
 - Divided into active and passive operations
- Active operations (load, unload, start, stop, etc.) require privileged (root) access
 - Secure kernels (i.e. iOS) remove ability to load kernel extensions



- Passive operations were originally unrestricted in < iOS 6
 - Allowed unprivileged users to query kernel and module base addresses



- iOS 6 inadvertently removed some limitations
 - Only load address requests disallowed



- We can use kKextRequestPredicateGetLoaded
 - Returns load addresses and mach-o header dumps (base64 encoded)
 - Load address / Mach-O segment headers are obscured to hide ASLR slide
 - Mach-O section headers are not!
 - Reveals virtual addresses of loaded kernel sections



Request

<dict><key>Kext Request Predicate</key><string>Get Loaded Kext Info</string></dict>



Kernel Attacks: Heap Corruption

- Standard heap overflow tricks no longer work
 - Overwriting freelist pointer results in validation step failing
- Exploitation requires new strategies
 - Information leak to find heap address/cookies
 - Control structure manipulation
- Depends on corruption primitives



- Overflowing metadata is useful
 - Various control structures can be targeted instead
 - Requires some heap grooming (may or may not be difficult depending on block size)
- Various heap attacking primitives can be combined to gain code execution



Introducing vm_map_copy_t

```
struct vm map copy {
              type;
#define VM MAP COPY ENTRY LIST
#define VM MAP COPY OBJECT
#define VM MAP COPY KERNEL BUFFER
   vm object offset t offset;
   vm map size t
                   size;
   union {
       struct vm map header hdr; /* ENTRY LIST */
                  object; /* OBJECT */
       vm object t
       struct {
       void
                    *kdata; /* KERNEL BUFFER */
      vm size t kalloc size; /* size of this copy t */
       } c k;
   } c u;
```

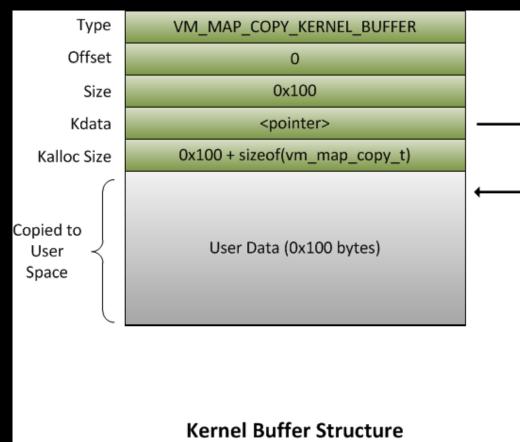


- Kernel buffers allocated by vm_map_copyin() if size < 4096
- Creating them is easy
 - Send messages to a mach port with ool_descriptors in them
 - They are persistent until the message is received
- Corrupting these structures are useful for information leaking and exploitation

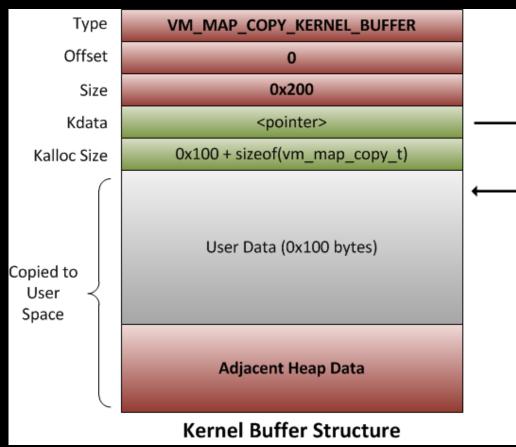


- Primitive 1: Adjacent Disclosure
 - Overwrite size parameter of vm_map_copy_t
 - Receive the message corresponding to the map
 - Returns memory past the end of your allocated buffer
- Bonus: Overwritten size is not used in kfree()
 - No side effects









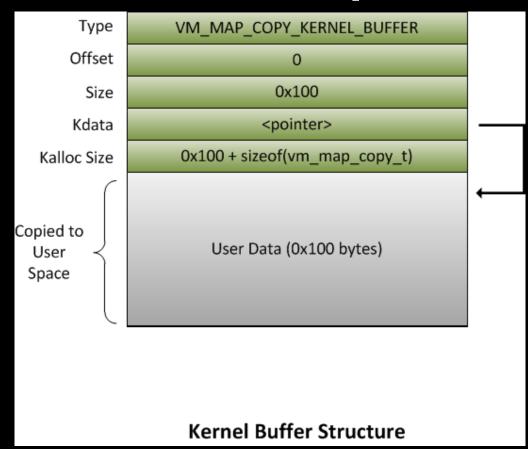


- Primitive 2: Arbitrary Memory Disclosure
 - Overwrite size and pointer of adjacent vm_map_copy_t
 - Receive message, read arbitrary memory from kernel
- No side effects
 - Data pointer (cpy_kdata) is never passed to kfree() (the vm_map_copy_t is)
 - Leave kalloc_size alone!

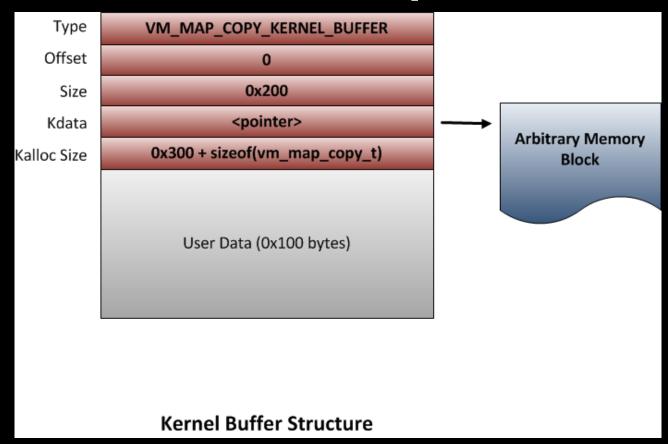


- Primitive 3: Extended Overflow
 - Overwrite kalloc_size with larger value
 - Passed to kfree() block entered in to wrong zone (eg. kalloc.256 instead of kalloc.128)
 - Allocate block from poisoned zone
 - Profit





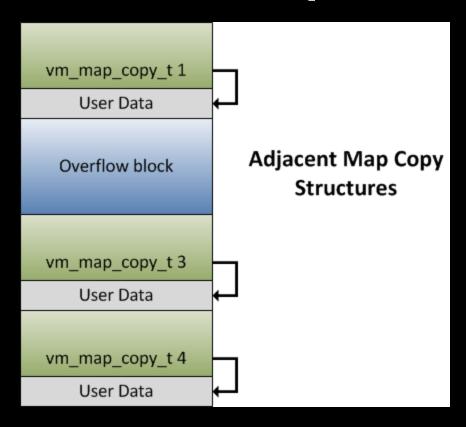




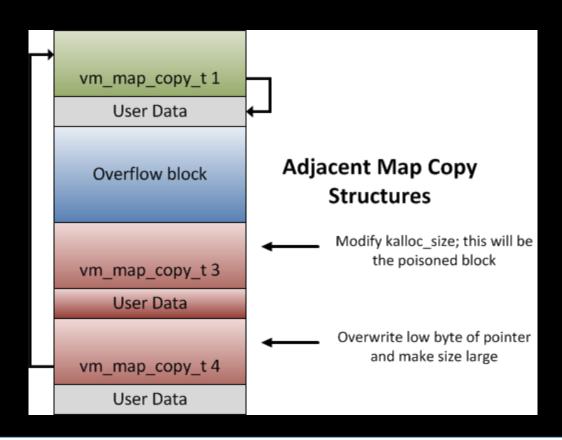


- Primitive 4: Find our own address + Overflow
 - Mix and match primitive 1 and 3
 - Overwrite one whole vm_map_copy_t, changing kalloc_size to be suitably large
 - Overflow in to adjacent vm_map_copy_t, partially overwriting pointer / length
 - Free second copy (revealing pointers to itself)
 - Free first copy, creating poisoned kalloc block at known location











Conclusion

- iOS 6 mitigations significantly raise the bar
 - Many of the old tricks don't work
 - A variety of bugs likely to be (reliably) unexploitable now
- Presented strategies provide useful mechanisms for exploiting iOS 6
- Thanks!

