Working with Native Libraries in Java

Vladimir Ivanov HotSpot JVM Compiler Oracle Corp.

@iwan0www OpenJDK: vlivanov 14.10.2016



Why?

- LAPACK
 - Linear Algebra PACKage

```
L A P A C K
L -A P -A C -K
L A P A -C -K
L -A P -A -C K
L A -P -A C K
L -A -P A C -K
```



Why?

- LAPACK
 - Linear Algebra PACKage
 - written in Fortran 90
 - highly optimized
 - "The original goal of the LAPACK was to ... run efficiently on sharedmemory vector and parallel processors."

```
L A P A C K
L -A P -A C -K
L A P A -C -K
L -A P -A -C K
L A -P -A C K
L -A -P A C -K
```



How?

LAPACK

- 1. invoke library code
- 2. pass data into library
- 3. access data from Java

```
L A P A C K
L -A P -A C -K
L A P A -C -K
L -A P -A -C K
L A -P -A C K
L -A -P -A C K
```



Overview

- Existing
 - Java Native Interface (JNI) & JNR library
 - java.nio.DirectByteBuffer
 - sun.misc.Unsafe (get*/set*)
- JDK9
 - j.l.i.VarHandle views over ByteBuffers
- Future
 - Project Panama



Native Code



Native Code

- LAPACK
 - 1. invoke library code
 - 2. pass data into library
 - 3. access data from Java

$$B = A^2$$



@since 1.1



Usage scenario

```
class LibC {
            User Code
                                   static native long getpid();
             JNI call
lava
C/native
             JNI impl
                                 jlong JNICALL Java_LibC_getpid(
                                   JNIEnv* env, jclass c) {
           Target Library
                                    return getpid();
```



Upcall



Data access

```
jlong JNICALL Java ...(JNIEnv* env,
                        jclass cls,
                        jobject obj) {
jfieldID fid = env->GetFieldID(cls, "f", "J");
ilong result = env->GetLongField(obj, fid);
jlong result = env->SetLongField(obj, fid, 10);
```

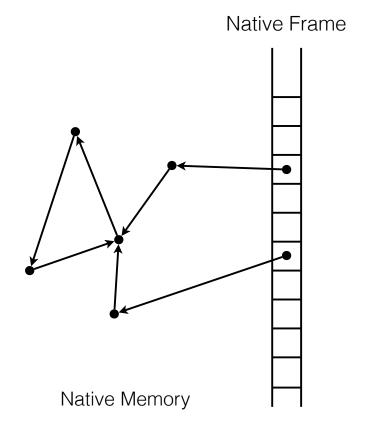


Native API: JNIEnv

- Operations on
 - Classes
 - Strings
 - Arrays
 - Monitors

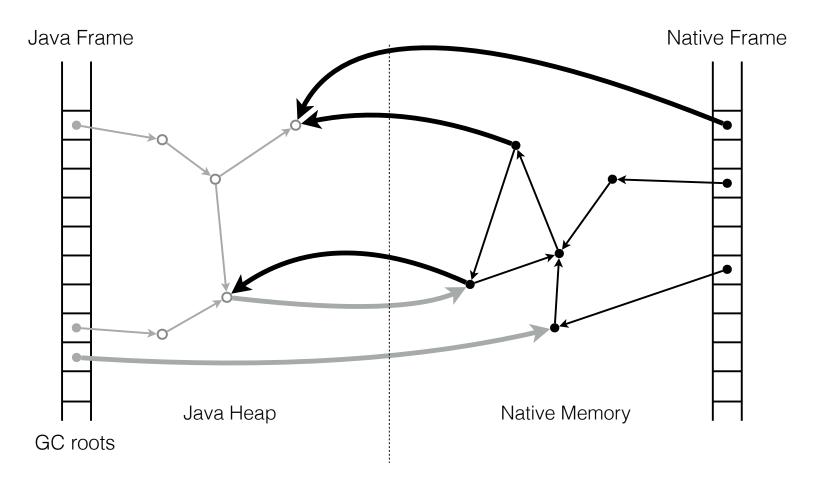


Java Frame Java Heap

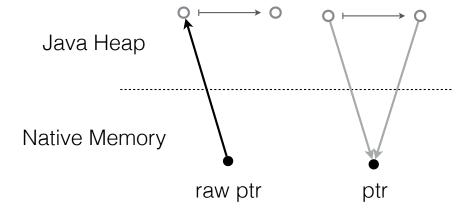




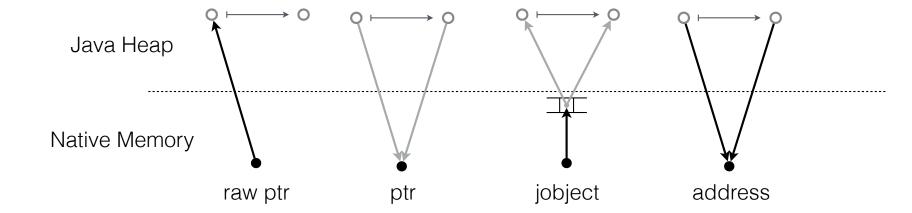
GC roots

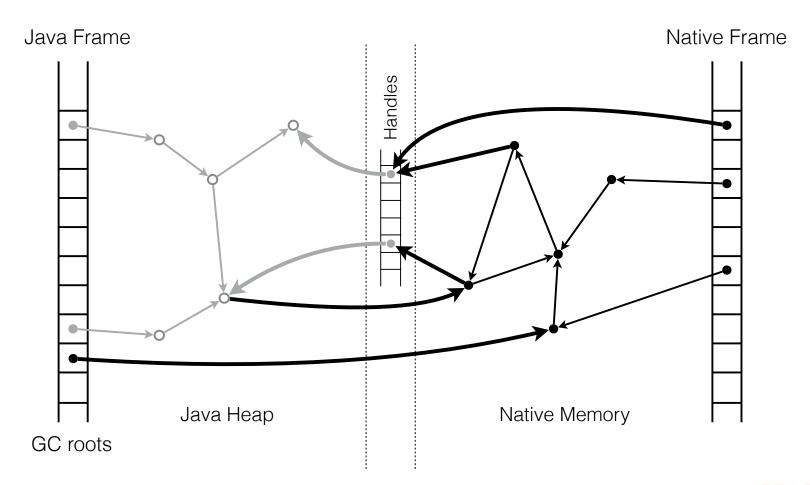






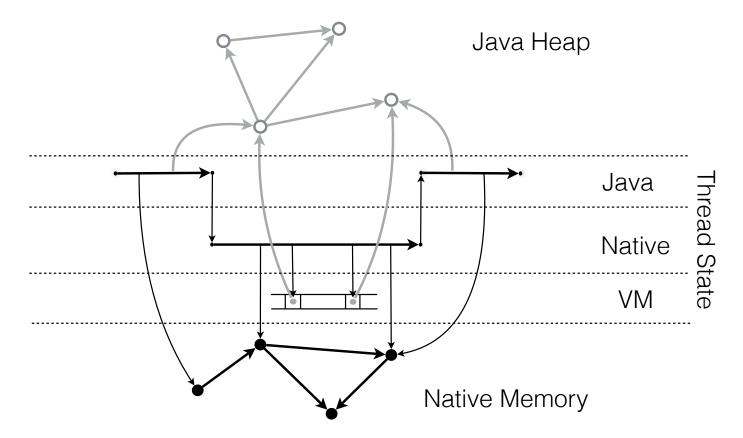








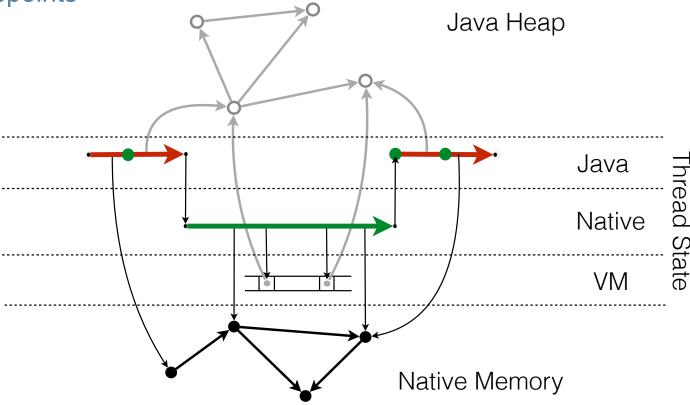
Anatomy of JNI call





Anatomy of JNI call

Safepoints





Pros

- seamless integration
 - looks like a Java method
- rich native API to interact with Java

Cons

- manual binding
- invocation overhead



Victim of its own success?





Sum array elements

```
jint JNICALL Java_...(JNIEnv *env, jclass c, jobject arr) {
   jint len = (*env)->GetArrayLength(env, arr);
   jbyte* a = (*env)->GetPrimitiveArrayCritical(env, arr, 0);
   ...
   return sum;
}
```

	empty	sum 1	sum 10³	sum 10 ⁶
JNI	11.4±0.3 ns	178.0±7.1 ns	798±32 ns	641±51 µs



/* @since 7 */



Sum array elements

```
jint JNICALL JavaCritical_...(jint length, jbyte* first) {
    ...
    return sum;
}
```

	empty	sum 1	sum 10 ³	sum 10 ⁶
JNI	11.4±0.3 ns	178.0±7.1 ns	798±32 ns	641±51 μs
CriticalJNI	11.4±0.3 ns	17.2±0.8 ns	680±22 ns	636±12 μs

Limitations

only static, non-synchronized methods supported



- only static, non-synchronized methods supported
- no JNIEnv*
 - hence, no upcalls or access to Java heap



- only static, non-synchronized methods supported
- no JNIEnv*
- arguments: primitives or primitive arrays
 - [I => (length, I*)
 - null => (0, NULL)



- only static, non-synchronized methods supported
- no JNIEnv*
- arguments: primitives or primitive arrays
 - [I => (length, I*)
 - null => (0, NULL)
- no object arguments



- only static, non-synchronized methods supported
- no JNIEnv*
- arguments: primitives or primitive arrays
 - [I => (length, I*)
 - null => (0, NULL)
- no object arguments
- used only in optimized code
 - 2 versions are needed: ordinary JNI & critical JNI versions



Hard cases

int printf(const char *format, ...)



Hard cases

```
void qsort(
  void* base,
  size_t nel,
  size_t width,
  int (*cmp)(const void*, const void*));
```



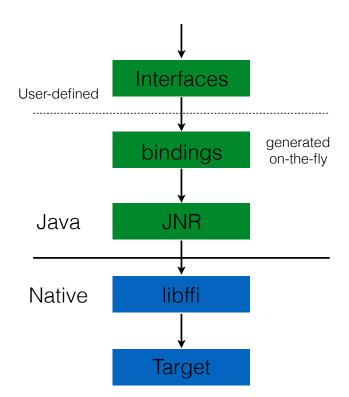
JNR

Java Native Runtime



JNR

Usage scenario



```
public interface LibC {
  @pid_t long getpid();
LibC lib = LibraryLoader
    .create(LibC.class)
    .load("c");
libc.getpid()
```



DEMO

- native call
 - getpid
- structs
 - gettimeofday
- upcalls
 - qsort



JNR

- Pros
 - automatic binding of native methods
- Cons
 - manual interface extraction
 - doesn't scale
 - still uses JNI to perform native calls



Better JNI

Easier, safer, faster!



"If non-Java programmers find some library useful and easy to access, it should be similarly accessible to Java programmers."

John Rose, JVM Architect,

Oracle Corporation



Project Panama

"Bridging the gap"







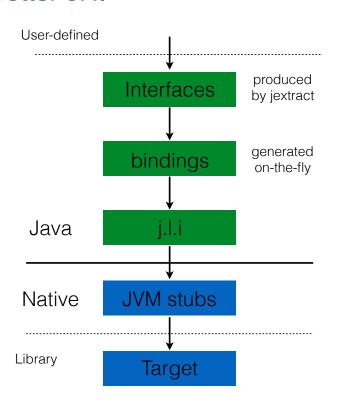


Better JNI

pid_t get_pid();



Easier



```
public interface LibC {
    long getpid();
LibC libc = Library
     .load(LibC.class, "c");
libc.getpid();
```

Easier

```
public interface LibC {
    long getpid();
}
LibC libc = Library.load(LibC.class, "c" /* lib_name */ );
libc.getpid();
```



Faster

Better JNI

callq 0x1057b2eb0; getpid entry



Faster

```
MethodType mt = MethodType.methodType(int.class); // pid_t
MethodHandle mh =
    MethodHandles.lookup().findNative("getpid", mt);
int pid = (int)mh.invokeExact();
```

	getpid
JNI	$13.7 \pm 0.5 \text{ns}$
Direct call	$3.4 \pm 0.2 \text{ ns}$



Safer

- no crashes
- no leaks
- no hangs
- no privilege escalation
- no unguarded casts





















Trust Levels

Better JNI



Untrusted



Trust Levels

Better JNI



Trusted



Trust Levels

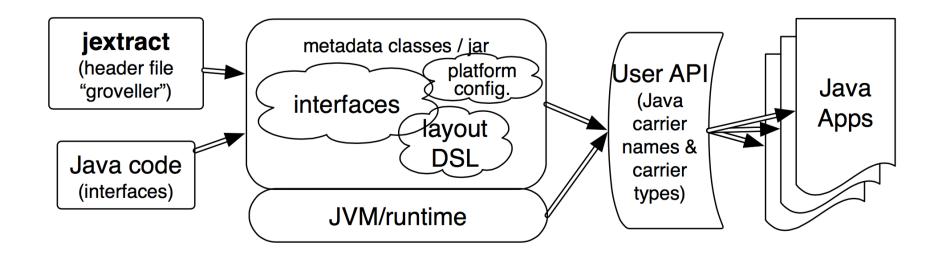
Better JNI



Privileged



Usage

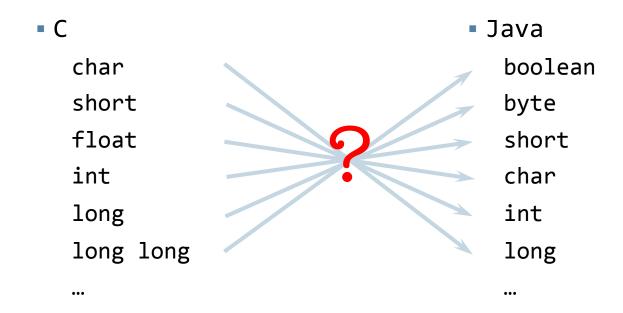




gettimeofday

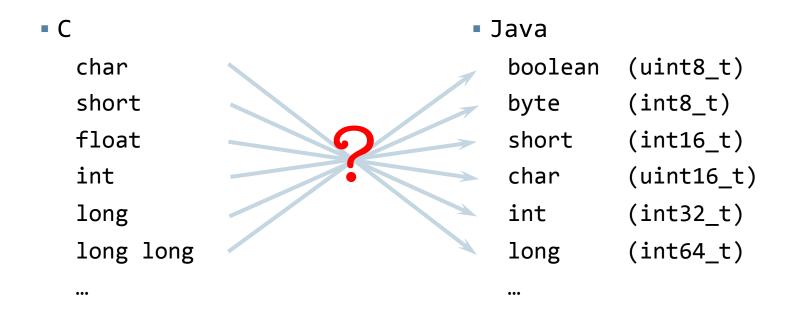


Carrier Types





Carrier Types



\$ jextract time.h

```
interface Time {
interface Timeval {
                              interface Timezone {
    long tv sec$get();
                                   int tz ...$get();
    void tv sec$set(long);
                                   void tz ...$set(int);
    long tv usec$get();
                                   int tz ...$get();
    void tv usec$set(long);
                                   void tz ...$set(int);
int gettimeofday(Timeval, Timezone);
```

Foreign Layouts

- Native data requires special address arithmetic
 - Native layouts should not be built into the JVM
 - Native types are unsafe, so trusted code must manage the bits
- Solution: A metadata-driven Layout API

- As a bonus, layouts other than C and Java are naturally supported
 - Network protocols, specialized in-memory data stores, mapped files, etc.



Better JNI

Data Layout

```
interface Timeval {
...
    @Offset(offset=0L)
    long tv_sec$get();
...
    @Offset(offset=64L)
    long tv_usec$get();
```

- work on Layout Definition Language (LDL) is in progress
 - https://github.com/J9Java/panama-docs/blob/master/StateOfTheLDL.html
 - http://cr.openjdk.java.net/~jrose/panama/minimal-ldl.html



```
Library lib = Library.create("c");
Time time = lib.create(Time.class);
Timeval tval = lib.create(Timeval.class);
int res = time.gettimeofday(tval, null);
if (res == 0) {
 long tv sec = tval.tv sec$get();
  long tv usec = tval.tv usec$get();
} else { /* error handling */ }
```



```
Library lib = Library.create("c");
Time time = lib.create(Time.class);
Timeval tval = lib.create(Timeval.class);
int res = time.gettimeofday(tval, null);
if (res == 0) {
 long tv sec = tval.tv sec$get();
  long tv usec = tval.tv usec$get();
} else { /* error handling */ }
```



```
Library lib = Library.create("c");
Time time = lib.create(Time.class);
Timeval tval = lib.create(Timeval.class);
int res = time.gettimeofday(tval, null);
if (res == 0) {
 long tv sec = tval.tv sec$get();
  long tv usec = tval.tv usec$get();
} else { /* error handling */ }
```



```
Library lib = Library.create("c");
Time time = lib.create(Time.class);
Timeval tval = lib.create(Timeval.class);
int res = time.gettimeofday(tval, null);
if (res == 0) {
  long tv sec = tval.tv sec$get();
  long tv usec = tval.tv usec$get();
} else { /* error handling */ }
```



Explicit management

```
Timeval tval = null;
try {
  tval = lib.create(Timeval.class);
  int res = time.gettimeofday(tval, null);
  if (res == 0) {
    long tv sec = tval.tv sec$get();
    long tv usec = tval.tv usec$get();
 } else { /* error handling */ }
} finally {
  if (tval != null) {
    lib.free(tval);
    tval = null;
```

Try-with-resources

```
interface Timeval extends AutoCloseable { ... }
try (Timeval tval = lib.create(Timeval.class)) {
  int res = time.gettimeofday(tval, null);
  if (res == 0) {
    long tv sec = tval.tv sec$get();
    long tv usec = tval.tv usec$get();
 } else { /* error handling */ }
```



Scoped memory

```
try (Scope scope = lib.createScope()) {
   TimeVal tval = scope.create(TimeVal.class);
   int res = time.gettimeofday(tval, null);
   if (res == 0) {
      long tv_sec = tval.tv_sec$get();
      long tv_usec = tval.tv_usec$get();
   } else { /* error handling */ }
}
```



Scoped memory

```
TimeVal tval = null;

try (Scope scope = lib.createScope()) {
   tval = scope.create(TimeVal.class);
   int res = time.gettimeofday(tval, null);
} // end of scope
```



Scoped memory

```
TimeVal tval = null;
try (Scope scope = lib.createScope()) {
  tval = scope.create(TimeVal.class);
  int res = time.gettimeofday(tval, null);
} // end of scope
// Access attemps out of scope
long tv sec = tval.tv sec$get(); // liveness checks!
long tv usec = tval.tv usec$get(); // liveness checks!
```

"Civilizer"

```
interface Timeval {
  void gettimeofday(Timeval, Timezone) throws ErrNo;
}
```



"Civilizer"

```
interface Timeval {
 void gettimeofday(Timeval, Timezone) throws ErrNo;
try (Timeval tval = lib.create(Timeval.class)) {
  time.gettimeofday(tval, null); // throws exception
  long tv sec = tval.tv sec$get();
  long tv usec = tval.tv usec$get();
```

Variadic Function

Better JNI

int printf(const char *format, ...)



jextract + Civilizer

```
// int printf(const char *format, ...)
interface Stdio {
...
    // "Raw"
    int printf(Pointer<Byte> format, byte[] args);
```



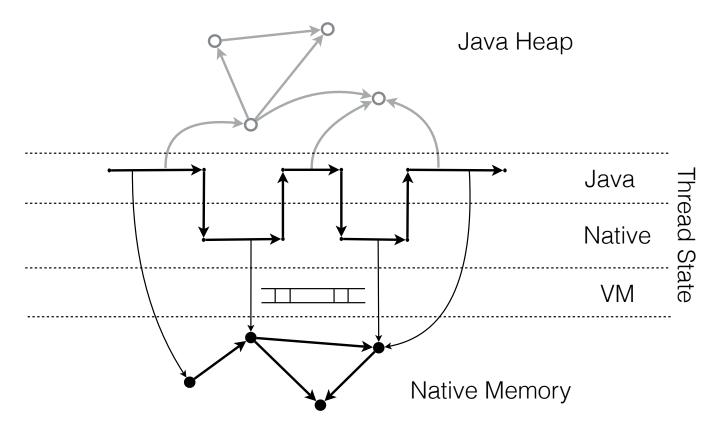
jextract + Civilizer

Better JNI

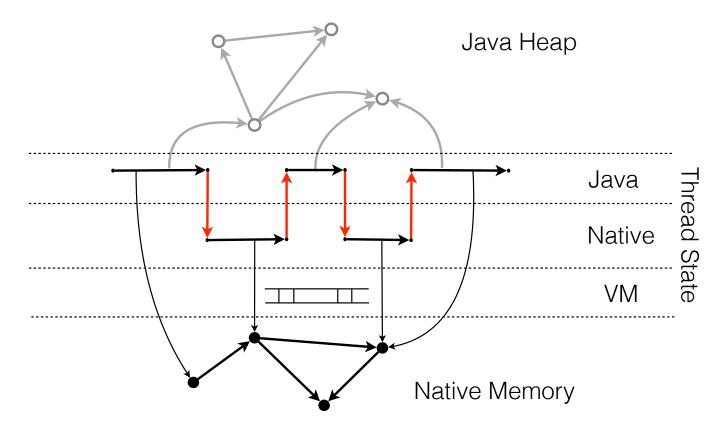
```
// int printf(const char *format, ...)
interface Stdio {
...
    // "Raw"
    int printf(Pointer<Byte> format, byte[] args);
    // "Civilized"
    void printf(String format, Object... args);
```



```
void run(MyClass obj) {
  obj.nativeFunc1(); // checks & state trans.
  obj.nativeFunc2(); // checks & state trans.
  obj.nativeFunc3(); // checks & state trans.
}
```









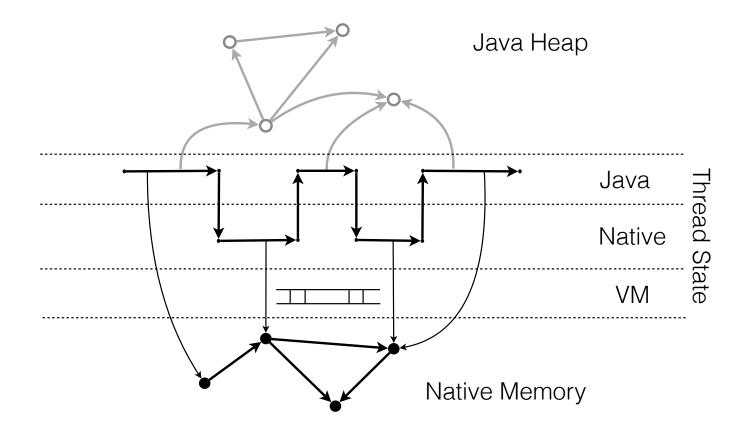
```
void run(MyClass obj) {
  obj.f1(); // NPE
  obj.f2(); // NPE
  obj.f3(); // NPE
}
```



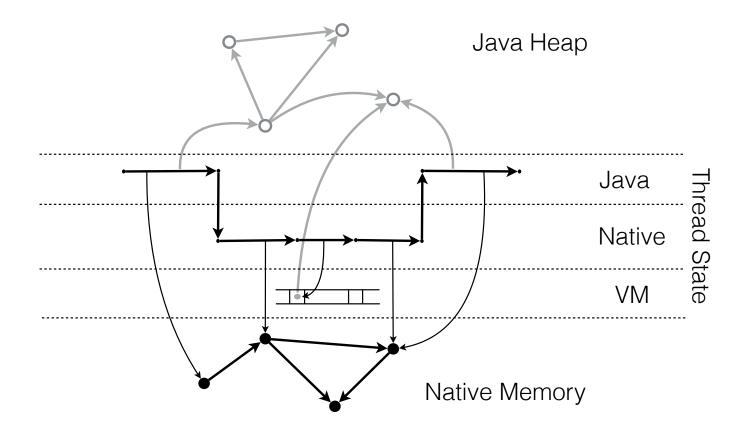
```
void run(MyClass obj) {
  if (obj == null) jump throwNPE_stub;
  call MyClass::f(obj);
  call MyClass::f1(obj);
  call MyClass::f3(obj);
}
```

```
void run(MyClass obj) {
  obj.nativeFunc1(); // checks & state trans.
  obj.nativeFunc2(); // checks & state trans.
  obj.nativeFunc3(); // checks & state trans.
}
```

```
void run(MyClass obj) {
  if (!performChecks()) jump failed stub;
  call transJavaToNative();
  MyClass::nativeFunc1(env, obj);
  MyClass::nativeFunc2(env, obj);
  MyClass::nativeFunc3(env, obj);
  call transNativeToJava();
```









Better JNI

Easier, Safer, Faster!

- Native access between the JVM and native APIs
 - Native code via FFIs
 - Native data via safely-wrapped access functions
 - Tooling for header file API extraction and API metadata storage
- Wrapper interposition mechanisms, based on JVM interfaces
 - add (or delete) wrappers for specialized safety invariants
- Basic bindings for selected native APIs



Native Data



Native Data

- LAPACK
 - 1. invoke library code
 - 2. pass data into library
 - 3. access data from Java

$$B = A^2$$



JNI

@since 1.1



NIO

@since 1.4



NIO

"New I/O"

- Provides access to the low-level I/O operations
 - Buffers for bulk memory operations
 - on-heap and off-heap
 - Character set encoders and decoders
 - Channels, a new primitive I/O abstraction
 - File interface
 - supports locks and memory mapping of files
 - Multiplexed, non-blocking I/O



- java.nio.ByteBuffer / CharBuffer / ...
 - MappedByteBuffer extends ByteBuffer
 - memory-mapped region of a file
 - DirectByteBuffer extends MappedByteBuffer
 - malloc'ed native memory
 - HeapByteBuffer
 - backed by byte[]



```
ByteBuffer dbb = ByteBuffer.allocateDirect(size);
while (dbb.hasRemaining()) {
  dbb.putInt(...); // init
LAPACK.square(dbb.address(), size); // invoke
dbb.rewind(); // reset position
while (dbb.hasRemaining()) {
  int i = dbb.getInt(); // read result
```



```
ByteBuffer dbb = ByteBuffer.allocateDirect(size);
while (dbb.hasRemaining()) {
  dbb.putInt(...); // init
LAPACK.square(dbb.address(), size); // invoke
dbb.rewind(); // reset position
while (dbb.hasRemaining()) {
  int i = dbb.getInt(); // read result
```



```
ByteBuffer dbb = ByteBuffer.allocateDirect(size);
while (dbb.hasRemaining()) {
  dbb.putInt(...); // init
LAPACK.square(dbb.address(), size); // invoke
dbb.rewind(); // reset position
while (dbb.hasRemaining()) {
  int i = dbb.getInt(); // read result
```



```
ByteBuffer dbb = ByteBuffer.allocateDirect(size);
while (dbb.hasRemaining()) {
  dbb.putInt(...); // init
LAPACK.square(dbb.address(), size); // invoke
dbb.rewind(); // reset position
while (dbb.hasRemaining()) {
  int i = dbb.getInt(); // read result
```



```
ByteBuffer dbb = ByteBuffer.allocateDirect(size);
while (dbb.hasRemaining()) {
  dbb.putInt(...); // init
LAPACK.square(dbb.address(), size); // invoke
dbb.rewind(); // reset position
while (dbb.hasRemaining()) {
  int i = dbb.getInt(); // read result
```

```
ByteBuffer dbb = ByteBuffer.allocateDirect(size);
while (dbb.hasRemaining()) {
  dbb.putInt(...); // init
LAPACK.square(dbb); // invoke
while (dbb.hasRemaining()) {
  int i = dbb.getInt(); // read result
```



- < 2GiB</p>
 - ByteBuffer.allocateDirect(int size)
 - ByteBuffer.allocate(int size)



- < 2GiB</p>
 - ByteBuffer.allocateDirect(int size)
- Stateful
 - Buffer.position
 - not thread-safe



- < 2GiB</p>
 - ByteBuffer.allocateDirect(int size)
- Stateful
 - Buffer.position
 - not thread-safe
- Resource deallocation
 - GC-based (Cleaner) memory management



- < 2GiB
 - ByteBuffer.allocateDirect(int size)
- Stateful
 - Buffer.position
 - not thread-safe
- Resource deallocation
 - GC-based (Cleaner) memory management
- Zeroing
 - on initialization



- < 2GiB</p>
 - ByteBuffer.allocateDirect(int size)
- Stateful
 - Buffer.position
 - not thread-safe
- Resource deallocation
 - GC-based (Cleaner) memory management
- Zeroing
 - on initialization
- Bounds checking



Anti-JNI



Use case	Example methods
Concurrency primitives	compareAndSwap*
Serialization	allocateInstance
Efficient memory management, layout, and access	allocateMemory/freeMemory get*/put*
Interoperate across the JVM boundary	get*/put*
•••	••••



- Unsafe.get*/put*
 - getInt(Object base, long offset)
 - putInt(Object base, long offset, int value);



- Unsafe.get*/put*
 - getInt(Object base, long offset)
 - putInt(Object base, long offset, int value);

- double-register addressing mode
 - getInt(o, offset) == o + offset
 - getInt(null, address) == address



- Unsafe.get*/put*
 - getInt(Object base, long offset)
 - putInt(Object base, long offset, int value);
- double-register addressing mode
 - getInt(o, offset) == o + offset
 - getInt(null, address) == address
- long allocateMemory(long size)

void freeMemory(long address)



```
long buf = UNSAFE.allocateMemory(size);

LAPACK.square(buf, size);

for (long l = 0; l < size; l =+ 4) {
  int i = UNSAFE.getInt(null, buf + 1);
}</pre>
```

UNSAFE.putInt(new Object(), 0L, 0)



UNSAFE.putInt(null, 0L, 0)



Object UNSAFE.getObject(**long** address)



long UNSAFE.getAddress(long address)



Unsafe =?= Fast





public native Object allocateInstance(Class<?> cls) throws ...;



Array index vs Raw offset

```
long[] base = new long[...];
int idx = ...;
```



Array index vs Raw offset

```
long[] base = new long[...];
int idx = ...;

// "Naïve" version
long value = base[idx];
```



Array index vs Raw offset

```
long[] base = new long[...];
int idx = ...;
// "Naïve" version
long value = base[idx];
// Highly optimized
long offset = (((long) idx) << SCALE + OFFSET)</pre>
long value = Unsafe.getLong(base, offset);
```

Array index vs Raw offset: 32-bit platform

```
long[] base = new long[...];
int idx = ...;
// "Naïve" version
long value = base[idx];
// Highly optimized
long offset = (((long) idx) << SCALE + OFFSET)</pre>
long value = Unsafe.getLong(base, offset);
```

- Missing optimizations
 - <u>JDK-8078629</u>: "VM should constant fold Unsafe.get*() loads from final fields"



- How many of you have used the Unsafe API?

John Rose, JVM Architect, Oracle JVM Language Summit 2014

ORACLE

Copyright © 2016, Oracle and/or its affiliates. All rights reserved

- How many of you have used the Unsafe API?

- A lot of you. Gosh. I'm sorry.

John Rose, JVM Architect, Oracle JVM Language Summit 2014



OpenJDK

OpenJDK FAQ Installing Contributing Sponsoring Developers' Guide

Mailing lists IRC · Wiki

Bylaws · Census Legal

JEP Process

search

Source code

Mercurial Bundles (6)

Groups

(overview)
2D Graphics
Adoption
AWT
Build
Compiler
Conformance
Core Libraries
Governing Board
HotSpot
Internationalization
IMX

NetBeans Projects

Members Networking

JEP 260: Encapsulate Most Internal APIs

Author Mark Reinhold
Owner Chris Hegarty
Created 2015/08/03 18:29

Updated 2015/10/02 17:20

Type Feature Status Candidate Scope JDK

Discussion jigsaw dash dev at openjdk dot java dot net

Effort M Duration L Priority 1

Reviewed by Alan Bateman, Alex Buckley, Brian Goetz, John Rose, Paul

Sandoz

Release 9

Issue 8132928

Summary

Make most of the JDK's internal APIs inaccessible by default but leave a few critical, widely-used internal APIs accessible, until supported replacements exist for all or most of their functionality.



sun.misc.Unsafe

Use case	Example methods
Concurrency primitives	compareAndSwap*
Serialization	allocateInstance (ReflectionFactory.newConstructorForSerialization)
Efficient memory management, layout, and access	allocateMemory/freeMemory get*/put*
Interoperate across the JVM boundary	get*/put*

sun.misc.Unsafe

Use case	Replacement
Concurrency primitives	JEP 193 Variable Handles
Serialization	Reboot JEP 187 Serialization Improvements
Efficient memory management, layout, and access	Project Panama, Project Valhalla, Arrays 2.0, Better GC
Interoperate across the JVM boundary	Project Panama, JEP 191 Foreign Function Interface



java.lang.invoke.

VarHandle

@since 9

JEP 193: Variable Handles



VarHandle

ByteBuffer View

MethodHandles.Lookup:

"Produces a VarHandle giving access to elements of a ByteBuffer viewed as if it were an array of elements of a different primitive component type to that of byte, such as int[] or long[]."



VarHandle

ByteBuffer View

```
VarHandle VH =
    MethodHandles.byteBufferViewVarHandle(
         int[].class,
         ByteOrder.nativeOrder() == ByteOrder.BIG_ENDIAN);
ByteBuffer dbb = ByteBuffer.allocateDirect(size);
int v = (int)VH.get(dbb, idx);
```



java.nio.ByteBuffer vs VarHandle View

	DirectByteBuffer	VarHandle
Size	< 2 GiB	< 2 GiB
State	Yes	No
Resource management	GC-based	No (delegates to DBB)
Zeroing	Yes	No (delegates to DBB)
Atomics/Fences/	No	Yes
Bound checks	Yes (optimized)	Yes (optimized)



Optimized Bounds Checks

int[]

```
// null check + (index u< array.length)
return array[index];</pre>
```



Optimized Bounds Checks

int[]: Unsafe access

```
// bounds and null check
if (index < 0 || index >= array.length)
  throw new ...();

long offset = BASE + (((long) index) << 2);
return UNSAFE.getInt(array, offset);</pre>
```



Optimized Bounds Checks

int[]: Unsafe access

```
// bounds (u<) and null check
index = Objects.checkIndex(index, array.length);
long offset = BASE + (((long) index) << 2);</pre>
return UNSAFE.getInt(array, offset);
@HotSpotIntrinsicCandidate
public static int checkIndex(int index, int length, ...);
```



Summary

- Existing
 - Java Native Interface (JNI) & JNR library
 - java.nio.DirectByteBuffer
 - sun.misc.Unsafe (get*/set*)
- JDK9
 - j.l.i.VarHandle views over ByteBuffers
- Future
 - Project Panama



Project Panama

OpenJDK

Foreign Function Interface Data Layout Control Vector API Arrays 2.0

http://openjdk.java.net



Project Panama

panama-dev@openjdk.java.net

http://hg.openjdk.java.net/panama/panama

OpenJDK

http://openjdk.java.net



Safe Harbor Statement

The preceding is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.



MAKE THE FUTURE JAVA



ORACLE*