CSE P 501 – Compilers

Java Implementation – JVMs, JITs &c

Hal Perkins

Winter 2008

Agenda

- Java virtual machine architecture
- .class files
- Class loading
- Execution engines
  - Interpreters & JITs – various strategies
- Exception Handling

Java Implementation Overview

- Java compiler (javac et al) produces machine-independent .class files
  - Target architecture is Java Virtual Machine (JVM) – simple stack machine
- Java execution engine (java)
  - Loads .class files (often from libraries)
  - Executes code
    - Either interprets stack machine code or compiles to native code (JIT)

JVM Architecture

- Abstract stack machine
- Implementation not required to use JVM specification literally
  - Only requirement is that execution of .class files has specified effect
  - Multiple implementation strategies depending on goals
    - Compilers vs interpreters
    - Optimizing for servers vs workstations

JVM Data Types

- Primitive types
  - byte, short, int, long, char, float, double, boolean
- Reference types
  - Non-generic only (more on this later)

JVM Runtime Data Areas (1)

- Semantics defined by the JVM Specification
  - Implementer may do anything that preserves these semantics
- Per-thread data
  - pc register
  - Stack
    - Holds frames (details below)
    - May be a real stack or may be heap allocated

JVM Runtime Data Areas (2)

- Per-VM data – shared by all threads
  - Heap – objects allocated here
  - Method area – per-class data
    - Runtime constant pool
    - Field and method data
    - Code for methods and constructors
- Native method stacks
  - Regular C-like stacks or equivalent

Frames

- Created when method invoked; destroyed when method completes
- Allocated on stack of creating thread
- Contents
  - Local variables
  - Operand stack for JVM instructions
  - Reference to runtime constant pool
    - Symbolic data that supports dynamic linking
  - Anything else the implementer wants

Representation of Objects

- Implementer's choice
  - JVM spec 3.7: “The Java virtual machine does not mandate any particular internal structure for objects”
  - Likely possibilities
    - Data + pointer to Class object
    - Pair of pointers: one to heap-allocated data, one to Class object

JVM Instruction Set

- Stack machine
- Byte stream
- Instruction format
  - 1 byte opcode
  - 0 or more bytes of operands
- Instructions encode type information
  - Verified when class loaded

Instruction Sampler (1)

- Load/store
  - Transfer values between local variables and operand stack
  - Different opcodes for int, float, double, addresses
  - Load, store, load immediate
    - Special encodings for load0, load1, load2, load3 to get compact code for first few local vars

Instruction Sampler (2)

- Arithmetic
  - Again, different opcodes for different types
    - byte, short, char & boolean use int instructions
  - Pop operands from operand stack, push result onto operand stack
  - Add, subtract, multiply, divide, remainder, negate, shift, and, or, increment, compare
- Stack management
  - Pop, dup, swap

Instruction Sampler (3)

- Type conversion
  - Widening – int to long, float, double; long to float, double, float to double
  - Narrowing – int to byte, short, char; double to int, long, float, etc.

Instruction Sampler (4)

- Object creation & manipulation
  - New class instance
  - New array
  - Static field access
  - Array element access
  - Array length
  - Instanceof, checkcast

Instruction Sampler (5)

- Control transfer
  - Unconditional branch – goto, jsr (originally used to implement finally blocks)
  - Conditional branch – ifeq, iflt, ifnull, etc.
  - Compound conditional branches - switch

Instruction Sampler (6)

- Method invocation
  - invokevirtual
  - invokeinterface
  - invokespecial (constructors, superclass, private)
  - invokestatic
- Method return
  - Typed value-returning instructions
  - Return for void methods

Instruction Sampler (7)

- Exceptions: athrow
- Synchronication
  - Model is *monitors* (cf any standard operating system textbook)
  - monitorenter, monitorexit
  - Memory model greatly cleaned up in Java 5

JVM and Generics

- Surprisingly, JVM has no knowledge of generic types
  - Not checked at runtime, not available for reflection, etc.
- Compiler *erases* all generic type info
  - Resulting code is pre-generics Java
  - Objects are class Object in resulting code & appropriate casts are added
- Only one instance of each type-erased class – no code expansion/duplication (as in C++ templates)

Generics and Type Erasure

- Why did they do that?
  - Compatibility: need to interop with existing code that doesn't use generics
    - Existing non-generic code and new generic libraries, or
    - Newly written code and older non-generic classes
- Tradeoffs: only reasonable way to add generics given existing world, but
  - Generic type information unavailable at runtime (casts, instanceof, reflection)
  - Can't create new instance or array of generic type
- C#/CLR is different – generics reflected in CLR

Class File Format

- Basic requirements are tightly specified
- Implementations can extend
  - Examples: data to support debugging or profiling
  - JVMs must ignore extensions they don't recognize
- Very high-level, symbolic, lots of metadata – much of the symbol table/type/other attribute data produced by a compiler front end
  - Supports dynamic class loading
  - Allows runtime compilation (JITs), etc.

Contents of Class Files (1)

- Starts with magic number (0xCAFEBABE)
- Constant pool - symbolic information
  - String constants
  - Class and interface names
  - Field names
- All other operands and references in the class file are referenced via a constant pool offset
- Constant pool is essentially a “symbol table” for the class

Contents of Class Files (2)

- Class and superclass info
  - Index into constant pool
- Interface information
  - Index into constant pool for every interface this class implements
- Fields declared in this class proper, but not inherited ones (includes type info)
- Methods (includes type info)
  - Includes byte code instructions for methods that are not native or abstract

Constraints on Class Files (1)

- Long list; verified at class load time
  - ∴ execution engine can assume valid, safe code
- Some examples of static constraints
  - Target of each jump must be an opcode
  - No jumps to the middle of an instruction or out of bounds
  - Operands of load/store instructions must be valid index into constant pool
  - new is only used to create objects; anewarray is only used to create arrays
  - Only invokespecial can call a constructor
  - Index value in load/store must be in bounds
  - Etc. etc. etc.

Constraints on Class Files (2)

- Some examples of structural constraints
  - Instructions must have appropriate type and number of arguments
  - If instruction can be executed along several paths, operand stack must have same depth at that point along all paths
  - No local variable access before being assigned a value
  - Operand stack never exceeds limit on size
  - No pop from empty operand stack
  - Execution cannot fall off the end of a method
  - Method invocation arguments must be compatible with method descriptor
  - Etc. etc. etc. etc.

Class Loaders

- One or more class loader (instances of ClassLoader or its derived classes) is associated with each JVM
- Responsible for loading the bits and preparing them
- Different class loaders may have different policies
  - Eager vs lazy class loading, cache binary representations, etc.
- May be user-defined, or the initial built-in bootstrap class loader

Readying .class Files for Execution

- Several distinct steps
  - Loading
  - Linking
    - Verification
    - Preparation
    - Resolution of symbolic references
  - Initialization

Loading

- Class loader locates binary representation of the class and reads it (normally a .class file, either in the local file system, or in a .jar file, or on the net)
- Once loaded, a class is identified in the JVM by its fully qualified name + class loader id
  - A good class loader should always return the same class object given the same name
  - Different class loaders generally create different class objects even given the same class name

Linking

- Combines binary form of a class or interface type with the runtime state of the JVM
- Always occurs after loading
- Implementation has flexibility on timing
  - Example: can resolve references to other classes during verification (static) or only when actually used (lazy)
  - Requirement is that verification must precede initialization, and semantics of language must be respected
    - No exceptions thrown at unexpected places, for example

Linking: Verification

- Checks that binary representation is structurally correct
  - Verifies static and structural constraints (see above for examples)
  - Goal is to prevent any subversion of the Java type system
- May causes additional classes and interfaces to be loaded, but not necessarily prepared or verified

Linking: Preparation

- Creation of static fields & initialization to default values
- Implementations can optionally precompute additional information
  - Method tables, for example

Linking: Resolution

- Check symbolic references and, usually, replace with direct references that can be executed more efficiently

Initialization

- Execute static initializers and initializers for static fields
- Direct superclass must be initialized first
- Constructor(s) not executed here
  - Done by a separate instruction as part of new, etc.

Virtual Machine Startup

- Initial class specified in implementation-defined manner
  - Command line, IDE option panel, etc.
- JVM uses bootstrap class loader to load, link, and initialize that class
- `public static void main(String[])` method of initial class is executed to drive all further execution

Execution Engines

- Basic Choices
  - Interpret JVM bytecodes directly
  - Compile bytecodes to native code, which then executes on the native processor
    - Just-In-Time compiler (JIT)

Hybrid Implementations

- Interpret or use very simple compiler most of the time
- Identify “hot spots” by dynamic profiling
  - Often per-method counter incremented on each call
  - Timer-based sampling, etc.
- Run optimizing JIT on hot code
  - Data-flow analysis, standard compiler middle-end optimizations, back-end instruction selection/ scheduling & register allocation
  - Need to balance compilation cost against responsiveness, expected benefits
    - Different tradeoffs for desktop vs server JVMs

Memory Management

- JVM includes instructions for creating objects and arrays, but not deleting
- Garbage collection used to reclaim no-longer needed storage (objects, arrays, classes, …)
- Strong type system means GC can have exact information
  - .class file includes type information
  - GC can have exact knowledge of layouts since these are internal to the JVM
- More details next hour

Escape Analysis

- Another optimization based on observation that many methods allocate local objects as temporaries
- Idea: Compiler tries to prove that no reference to a locally allocated object can "escape"
  - Not stored in a global variable or object
  - Not passed as a parameter

Using Escape Analysis

- If all references to an object are local, it doesn't need to be allocated on the heap in the usual manner
  - Can allocate storage for it in local stack frame
    - Essentially zero cost
  - Still need to preserve the semantics of new, constructor, etc.

Exception Handling

- Goal: should have zero cost if no exceptions are thrown
  - Otherwise programmers will subvert exception handling with the excuse of "performance"
- Corollary: cannot execute any exception handling code on entry/exit from individual methods or try blocks

Implementing Exception Handling

- Idea: Original compiler generates table of exception handler information in the .class file
  - Entries include start and end of section of code array protected by this handler; argument type
  - Order of entries is significant
- When exception is thrown, JVM searches exception table for first matching argument type that has a pc range that includes the current execution location

Summary

- That's the overview – many more details, obviously, if you want to implement a JVM
- Primary reference: Java Virtual Machine Specification, 2nd ed, A-W, 1999. Available online: http://java.sun.com/docs/books/jvms/
- Many additional research papers & studies all over the web and in conference proceedings