Production Breakpoints

Dale Hamel

08/16/19 11:03:12 PM UTC

Contents

This document is also available in epub and pdf format if you prefer.

Status

- Still very much under heavy development and experimentation
- Shifted direction to be tied to Ruby TracePoints
- Original implementation tied to eval to re-interpret method code is still more functional, but less practical for production use.

Injecting breakpoints

There are multiple ways to inject breakpoints into running code, each with different tradeoffs.

Unmixer Approach

I set out to prove if I could build a line-targeting debugger with USDT tracepoints, and did so via modifying the class hierarchy, generating an override method on-the-fly.

I was able to do this with "eval", and be passing the current binding to a block that calls eval, I was able to wrap the original code in my own block handler, but still be able to execute it in the current context!

This was a proof of concept at least. However, I quickly learned from my colleague Alan Wu, that this type of monkeypatching had dangerous implications to the ruby method cache which I had not considered.

Ruby "production breakpoints" would originally use its AST parsing to rewrite the source code of a method with the targeted lines to include a wrapper around those lines. This is still the most powerful way to do this from outside of the RubyVM I have been able to determine.

The method is redefined by prepending a module with the new definition to the parent of the original method, overriding it. To undo this, the module can be 'unprepended' restoring the original behavior.

Using the Ruby TracePoint API

Ruby now supports tracing only particular methods [1] instead of all methods, and looks to be aiming to add a similar sort of debugging support natively.

The docs are not very thorough in the official rubydoc for this [2], but it is further documented in the [3] prelude.rb file:

```
# call-seq:
#    trace.enable(target: nil, target_line: nil, target_thread:
↪  nil)    -> true or false
#    trace.enable(target: nil, target_line: nil, target_thread:
↪  nil) { block }  -> obj
#
# Activates the trace.
#
# Returns +true+ if trace was enabled.
# Returns +false+ if trace was disabled.
#
#   trace.enabled?  #=> false
#   trace.enable    #=> false (previous state)
#                   #   trace is enabled
#   trace.enabled?  #=> true
#   trace.enable    #=> true (previous state)
#                   #   trace is still enabled
#
# If a block is given, the trace will only be enabled within
↪  the scope of the
# block.
#
#    trace.enabled?
#    #=> false
#
#    trace.enable do
#      trace.enabled?
#      # only enabled for this block
#    end
#
#    trace.enabled?
#    #=> false
#
# +target+, +target_line+ and +target_thread+ parameters are
↪  used to
# limit tracing only to specified code objects. +target+ should
↪  be a
# code object for which RubyVM::InstructionSequence.of will
↪  return
# an instruction sequence.
#
#    t = TracePoint.new(:line) { |tp| p tp }
#
#    def m1
#      p 1
```

```
  #    end
  #
  #    def m2
  #      p 2
  #    end
  #
  #    t.enable(target: method(:m1))
  #
  #    m1
  #    # prints #<TracePoint:line@test.rb:5 in `m1'>
  #    m2
  #    # prints nothing
  #
  # Note: You cannot access event hooks within the +enable+
  ↪   block.
  #
  #    trace.enable { p tp.lineno }
  #    #=> RuntimeError: access from outside
  #
```

Alan Wu [4] tipped me off to this, and showed by an initial prototype gist [5], which succinctly shows a basic usage of this:

```
class Foo
  def hello(arg = nil)
    puts "Hello #{arg}"
  end
end

one = Foo.new
two = Foo.new

one.hello
two.hello

trace = TracePoint.new(:call) do |tp|
  puts "intercepted! arg=#{tp.binding.local_variable_get(:arg)}"
end
trace.enable(target: Foo.instance_method(:hello)) do
  one.hello(:first)
  two.hello(:second)
end
```

From this, we can build handlers that run using the RubyVM TracePoint object. This object has full access to the execution context of the caller it would seem.

Specifying breakpoints

A global config value:

```
ProductionBreakpoints.config_file
```

Can be set to specify the path to a JSON config, indicating the breakpoints that are to be installed:

```
{
  "breakpoints": [
    {
      "type": "inspect",
      "source_file": "test/ruby_sources/config_target.rb",
      "start_line": 7,
      "end_line": 9,
      "trace_id": "config_file_test"
    }
  ]
}
```

These values indicate:

- `type`: the built-in breakpoint handler to run when the specified breakpoint is hit in production.
- `source_file`: the source repository-root relative path to the source file to install a breakpoint within. (note, the path of this source folder relative to the host / mount namespace is to be handled elsewhere by the caller that initiates tracing via this gem)
- `start_line`: The first line which should be evaluated from the context of the breakpoint.
- `end_line`: The last line which should be evaluated in the context of the breakpoint
- `trace_id`: A key to group the output of executing the breakpoint, and filter results associated with a particular breakpoint invocation

Many breakpoints can be specified. Breakpoints that apply to the same file

are added and removed simultaneously. Breakpoints that are applied but not specified in the config file will be removed if the config file is reloaded.

Built-in breakpoints

Latency

Output: Integer, nanoseconds elapsed

The latency breakpoint provides the elapsed time from a monotonic source, for the duration of the breakpoint handler.

This shows the time required to execute the code between the start and end lines, within a given method.

Handler definition:

```
def handle(caller_binding, &block)
  return super(caller_binding, &block) unless
  ↪ @tracepoint.enabled?
  start_time = StaticTracing.nsec
  val = super(caller_binding, &block)
  duration = StaticTracing.nsec - start_time
  @tracepoint.fire(duration)
  return val
end
```

Locals

Output: String, key,value via ruby inspect

The 'locals' breakpoint shows the value of all locals.

NOTE: due to limitations in eBPF, there is a maximum serializable string size. Very complex objects cannot be efficiently serialized and inspected.

```
def handle(caller_binding, &block)
  return super(caller_binding, &block) unless
  ↪ @tracepoint.enabled?
```

```
        val = super(caller_binding, &block)
        locals = caller_binding.local_variables
        locals.delete(:local_bind) # we define this, so we'll get
↪   rid of it
        vals = locals.map { |v| [v,
↪   caller_binding.local_variable_get(v) ]}.to_h
        @tracepoint.fire(vals.inspect)
        return val
      end
```

Inspect

Output: String, value via ruby inspect

The `inspect` command shows the inspected value of whatever the last expression evaluated to within the breakpoint handler block.

```
      def handle(caller_binding, &block)
        return super(caller_binding, &block) unless
        ↪  @tracepoint.enabled?
        val = super(caller_binding, &block)
        @tracepoint.fire(val.inspect)
        return val
      end
```

Ustack

Output: String, simplified stack caller output

Internals

Installing breakpoints

This gem leverages the ruby-static-tracing [6] gem [7] which provides the **secret sauce** that allows for plucking this data out of a ruby process, using the kernel's handling of the Intel x86 "Breakpoint" int3 instruction, you can learn more about that in the USDT Report [8]

For each source file, a 'shadow' ELF stub is associated with it, and can be easily found by inspecting the processes open file handles.

After all breakpoints have been specified for a file, the ELF stub can be generated and loaded. To update or remove breakpoints, this ELF stub needs to be reloaded, which requires the breakpoints to be disabled first. To avoid this, the scope could be changed to be something other than file, but file is believed to be nice and easily discoverable for now.

The tracing code will noop, until a tracer is actually attached to it, and should have minimal performance implications.

Ruby Override technique via Unmixer

The Unmixer gem hooks into the ruby internal header API, and provides a back-door into the RubyVM source code to `unprepend` classes or modules from the global hierarchy.

An anonymous module is created, with the modified source code containing our breakpoint handler.

To enable the breakpoint code, this module is prepended to the original method's parent. To undo this, the module is simply 'unprepended', a feature unmixer uses to tap into Ruby's ancestry hierarchy via a native extension.

```
def install
  @injector_module =
  ↪ build_redefined_definition_module(@node)
```

```
      @ns.prepend(@injector_module)
    end

    # FIXME saftey if already uninstalled
    def uninstall
      @ns.instance_eval{ unprepend(@injector_module) }
      @injector_module = nil
    end
```

Dynamically redefined methods

We define a 'handler' and a 'finisher' block for each breakpoint we attach.

Presently, we don't support attaching multiple breakpoints within the same function, but we could do so if we applied this as a chain of handers followed by a finalizer, but that will be a feature for later. Some locking should exist to ensure the same method is not overriden multiple times until this is done.

These hooks into our breakpoint API are injected into the method source from the locations we used the ruby AST libraries to parse:

```
    def build_redefined_definition_module(node)

      # This is the metaprogramming to inject our breakpoint
      ↪ handler around the original source code
      handler = "local_bind=binding;
↪ ProductionBreakpoints.installed_breakpoints[:#{@trace_id}].handle(local_bind)"

      # This is needed to keep the execution of the remaining
      ↪ lines of the method within the same binding
      finisher =
↪ "ProductionBreakpoints.installed_breakpoints[:#{@trace_id}].finish(local_bind)"

      # This injects our handler and finisher blocks into the
      ↪ original source code, treating the code
      # in between as string literals to be evaluated
      injected =
↪ @parser.inject_metaprogramming_handlers(handler, finisher,
                                          node.first_lineno,
↪ node.last_lineno, @start_line, @end_line)
      #ProductionBreakpoints.logger.debug(injected)
      Module.new { module_eval{ eval(injected); eval('def
      ↪ production_breakpoint_enabled?; true; end;') } }
    end
```

And the outcome looks something like this:

```
# def some_method
# local_bind=binding;
↪   ProductionBreakpoints.installed_breakpoints[:test_breakpoint_install].handle(local_bind)
↪   do
# <<-EOS
#       a = 1
#       sleep 0.5
#       b = a + 1
# EOS
# end
#
↪   ProductionBreakpoints.installed_breakpoints[:test_breakpoint_install].finish(local_bind)
↪   do
# <<-EOS
# EOS
# end
#     end
```

This is called when the breakpoint is handled, in order to evaluate the whole method within the original, intended context:

```
      # Allows for specific handling of the selected lines
      def handle(caller_binding)
        eval(yield, caller_binding)
      end

      # Execute remaining lines of method in the same binding
      def finish(caller_binding)
        eval(yield, caller_binding)
      end
```

This caller binding is taken at the point our handler starts, so it's propagated from the untraced code within the method. We use it to evaluate the original source within the handler and finalizer, to ensure that the whole method is evaluated within the original context / binding that it was intended to. This should make the fact that there is a breakpoint installed transparent to the application.

The breakpoint handler code (see above) is only executed when the handler is attached, as they all contain an early return if the shadow "ELF" source doesn't have a breakpoint installed, via the `enabled?` check.

Experimentation with caching iseq to reduce overhead of eval

- Make iseq eval consistent with Kernel eval [9] is an attempt to revive an older patch to allow for evaluate an instruction sequence against an arbitrary binding. More details of this patch are below. The patch is based on draft work from Nobu [10] a few years ago on Ruby Feature 12093.

Ruby allows passing a binding to Kernel.eval, to arbitratily execute strings as ruby code in the context of a particular binding. Ruby also has the ability to precompile strings of Ruby code, but it does not have the ability to execute this within the context of a particular binding, it will be evaluated against the binding that it was compiled for.

This patch changes the call signature of eval, adding an optional single argument, where none are currently accepted. This doesn't change the contract with existing callers, so all tests pass. However, there is a major flaw in the design.

Why the patch doesn't work

Ultimately, this patch has been rejected for good reason, but I figured I'd explain what the obstacle is in case someone can figure out a clever solution!

Take for example this test case:

```
def bind
  a = 1
  b = 2
  binding
end
```

```
iseq = RubyVM::InstructionSequence.compile("a + b")
val = iseq.eval(bind)
```

We will get an error!

```
NameError: undefined local variable or method `a'
```

This is because compiled source may contain undefined references which may be assumed to be method calls or local variables. In the ruby instruction sequence, it is assumed to be a method call:

```
== disasm: #<ISeq:<compiled>@<compiled>:1 (1,0)-(1,5)> (catch:
↪  FALSE)
0000 putself
↪  (   1)[Li]
0001 opt_send_without_block       <callinfo!mid:a, argc:0,
↪  FCALL|VCALL|ARGS_SIMPLE>, <callcache>
0004 putself
0005 opt_send_without_block       <callinfo!mid:b, argc:0,
↪  FCALL|VCALL|ARGS_SIMPLE>, <callcache>
0008 opt_plus                     <callinfo!mid:+, argc:1,
↪  ARGS_SIMPLE>, <callcache>
0011 leave
```

In the other test case, where the variables are wrapped in a struct:

```
obj = Struct.new(:a, :b).new(1, 2)
bind = obj.instance_eval {binding}
iseq = RubyVM::InstructionSequence.compile("a + b")
val = iseq.eval(bind)
```

They are accessible because the struct is able receive the call, where the local variables in the binding object above wouldn't.

If a way can be devised for local variables to be added to the binding as above but in a more elegant / transparent way, this approach could regain its efficacy.

Patch Description

The rejected patch updated the signature and docs to:

```
/*
 *  call-seq:
 *     iseq.eval([binding]) -> obj
 *
 *  Evaluates the instruction sequence and returns the result.
 *
```

```
 *      RubyVM::InstructionSequence.compile("1 + 2").eval #=> 3
 *
 *  If <em>binding</em> is given, which must be a Binding object,
↪  the
 *  evaluation is performed in its context.
 *
 *      obj = Struct.new(:a, :b).new(1, 2)
 *      bind = obj.instance_eval {binding}
 *      RubyVM::InstructionSequence.compile("a + b").eval(bind)
↪  #=> 3
 */
```

Note that this is based on the signature of the Kernel.eval method:

```
/*
 *  call-seq:
 *     eval(string [, binding [, filename [,lineno]]])  -> obj
 *
 *  Evaluates the Ruby expression(s) in <em>string</em>. If
 *  <em>binding</em> is given, which must be a Binding object,
↪  the
 *  evaluation is performed in its context. If the optional
 *  <em>filename</em> and <em>lineno</em> parameters are present,
↪  they
 *  will be used when reporting syntax errors.
 *
 *     def get_binding(str)
 *       return binding
 *     end
 *     str = "hello"
 *     eval "str + ' Fred'"                      #=> "hello Fred"
 *     eval "str + ' Fred'", get_binding("bye")  #=> "bye Fred"
 */
```

Where the:

- First argument `string` to `Kernel.eval` is not necessary in the iseq version, as it is implied as part of `self` that was used to compile the iseq.
- Second argument, binding, is optional. This becomes the first argument of iseq.eval, as reasoned above
- Third optional argument, filename, is specified when an iseq is created so is not needed
- Fourth optional argument, lineno, is also specified when an iseq is created so is not needed

To implement this new call signature, the definition of `iseqw_eval` is updated to check the number of arguments.

```
static VALUE
iseqw_eval(int argc, const VALUE *argv, VALUE self)
{
    VALUE scope;

    if (argc == 0) {
        rb_secure(1);
        return rb_iseq_eval(iseqw_check(self));
    }
    else {
        rb_scan_args(argc, argv, "01", &scope);
        rb_secure(1);
        return rb_iseq_eval_in_scope(iseqw_check(self), scope);
    }
}
```

If no arguments are specified, it does what it always did. If an argument is specified, it scans for the argument and uses it as the binding, for a new VM method `rb_iseq_eval_in_scope` :

```
VALUE
rb_iseq_eval_in_scope(const rb_iseq_t *iseq, VALUE scope)
{
    rb_execution_context_t *ec = GET_EC();
    rb_binding_t *bind = Check_TypedStruct(scope,
↪   &ruby_binding_data_type);

    vm_set_eval_stack(ec, iseq, NULL, &bind->block);

    /* save new env */
    if (iseq->body->local_table_size > 0) {
      vm_bind_update_env(scope, bind, vm_make_env_object(ec,
↪   ec->cfp));
    }

    return vm_exec(ec, TRUE);
}
```

This definition is based on the approach used under the hood for Kernel.eval, when it calls `eval_string_with_scope`:

```
static VALUE
eval_string_with_scope(VALUE scope, VALUE src, VALUE file, int
↪   line)
{
    rb_execution_context_t *ec = GET_EC();
```

```
    rb_binding_t *bind = Check_TypedStruct(scope,
↪   &ruby_binding_data_type);
    const rb_iseq_t *iseq = eval_make_iseq(src, file, line, bind,
    ↪   &bind->block);
    if (!iseq) {
        rb_exc_raise(ec->errinfo);
    }

    vm_set_eval_stack(ec, iseq, NULL, &bind->block);

    /* save new env */
    if (iseq->body->local_table_size > 0) {
        vm_bind_update_env(scope, bind, vm_make_env_object(ec,
↪   ec->cfp));
    }

    /* kick */
    return vm_exec(ec, TRUE);
}
```

[1] “Ruby 2.6 adds targetted tracing.” [Online]. Available: https://bugs.ruby-lang.org/issues/15289

[2] “Ruby 2.6 Method Tracing Docs.” [Online]. Available: https://ruby-doc.org/core-2.6/TracePoint.html#method-i-enable

[3] S. Hiroshi, N. Nakada, and K. Sasada, “Ruby 2.6 targetted tracing prelude docs.” [Online]. Available: https://github.com/ruby/ruby/blame/master/prelude.rb#L140-L173

[4] A. Wu, “Alan wu’s Github.” [Online]. Available: https://github.com/XrXr

[5] “Demo of targetted tracing by Alan Wu.” [Online]. Available: https://gist.github.com/XrXr/f6cc2f1a4b6d3341e3d3fc749bcd5255

[6] “Ruby Static Tracing Github.” [Online]. Available: http://github.com/dalehamel/ruby-static-tracing

[7] “Ruby Static Tracing Gem.” [Online]. Available: https://rubygems.org/gems/ruby-static-tracing/

[8] D. Hamel, “USDT Report Doc.” [Online]. Available: https://bpf.sh/usdt-report-doc/index.html

[9] “Make iseq eval consistent with Kernel eval.” [Online]. Available: https://github.com/ruby/ruby/pull/2298

[10] “Nobu’s original patch for iseq.eval_with.” [Online]. Available: https://github.com/ruby/ruby/pull/2298/commits/5d0c8c83d8fdea16c403abbc80c160ddddb92ef8