Specification and verification in the field: Applying formal methods to BPF just-in-time compilers in the Linux kernel

Luke Nelson, Jacob Van Geffen, Emina Torlak, and Xi Wang,
University of Washington

https://www.usenix.org/conference/osdi20/presentation/nelson

This paper is included in the Proceedings of the 14th USENIX Symposium on Operating Systems Design and Implementation
November 4–6, 2020
978-1-939133-19-9

Open access to the Proceedings of the 14th USENIX Symposium on Operating Systems Design and Implementation is sponsored by USENIX
Specification and verification in the field: Applying formal methods to BPF just-in-time compilers in the Linux kernel

Luke Nelson, Jacob Van Geffen, Emina Torlak, and Xi Wang
University of Washington

Abstract

This paper describes our experience applying formal methods to a critical component in the Linux kernel, the just-in-time compilers (“JITs”) for the Berkeley Packet Filter (BPF) virtual machine. We verify these JITs using Jitterbug, the first framework to provide a precise specification of JIT correctness that is capable of ruling out real-world bugs, and an automated proof strategy that scales to practical implementations. Using Jitterbug, we have designed, implemented, and verified a new BPF JIT for 32-bit RISC-V, found and fixed 16 previously unknown bugs in five other deployed JITs, and developed new JIT optimizations; all of these changes have been upstreamed to the Linux kernel. The results show that it is possible to build a verified component within a large, unverified system with careful design of specification and proof strategy.

1 Introduction

Downloading application code into the OS kernel is a general approach to extensibility [26]. To extend the kernel, the application submits a program written in a dedicated language, and the kernel executes this program using an interpreter, or translates it into machine code for native execution via a just-in-time (JIT) compiler [3]. Berkeley Packet Filter (BPF) [31] is one such language, and it is used to implement a wide variety of extensions for the Linux kernel, including networking [38], security [79], and tracing [35], among many other services [18, 57].

Given the prevalence of BPF code and its execution in the OS kernel, the correctness of BPF JIT compilers (or simply “JITs”) is critical for the system. Compared to the BPF interpreter, using the JITs is both more efficient and more resistant to speculative attacks [84], leading major Linux distributions to remove the BPF interpreter from the kernel in favor of the JITs [9]. But the JITs are more susceptible to subtle correctness bugs due to their complexity (§3).

This paper presents a formal approach to building JITs in the kernel with high assurance of correctness. We develop Jitterbug, a framework for writing JITs and proving them correct. Using Jitterbug, we design, implement, and verify a BPF JIT for RV32, the 32-bit RISC-V architecture [96]. We also port the existing JITs for Arm32, Arm64, RV64, x86-32, and x86-64 to Jitterbug, uncovering 16 previously unknown bugs. We write patches that fix these bugs and introduce new optimizations, all of which are verified to be correct. The BPF JIT for RV32, bug fixes, and optimizations have been upstreamed to the Linux kernel.

Jitterbug is designed to meet three competing requirements: deployability of verified JITs with minimal changes to the Linux kernel; proof automation to support rapid verification of JITs; and separability of verified JITs from any verification artifacts, making the resulting code auditable by kernel developers with no background in formal methods. Each of these requirements comes with its own challenges and trade-offs.

First, BPF JITs and their generated code interact with a monolithic kernel via an existing interface, which was not designed for verification. As Jitterbug emphasizes deployability, it cannot adopt the clean-slate design favored by previous verification efforts [33, 65, 81, 94] or change this interface to simplify verification. Therefore, it needs a correctness specification that is both capable of ruling out real-world bugs and amenable to verification. Developing such a specification is challenging even for clean-slate designs with strong simplifying assumptions, and it is the core technical challenge addressed by Jitterbug.

Second, verification needs to catch up with increasing functionality and optimization of BPF JITs. Jitterbug thus prioritizes proof automation to free developers from the burden of writing manual proofs and to enable rapid verification in the code review process. Prior work has shown success in scaling automated verification to systems whose code does not change in response to input [68, 70]. But verifying a JIT is particularly challenging, because it requires reasoning about not only the behavior of the JIT itself, but also that of the machine code generated by the JIT for input BPF programs.

Third, kernel development emphasizes the efficiency and clarity of source code, whereas formal development emphasizes managing code complexity to make verification tractable.
Jitterbug must resolve the tension and make the two development processes cleanly separable. While formal development can use specific tools and artifacts such as specifications, the final implementation of a JIT needs to be C code that can be reviewed assuming no knowledge of formal methods, and can be compiled using a standard toolchain.

To address these challenges, Jitterbug makes the following contributions:

- A precise stepwise specification for JIT correctness (§4). The specification models both BPF and target architectures as abstract machines, and it formulates JIT correctness as the behavioral equivalence of running the machines with a source BPF instruction and the target instructions produced by the JIT, respectively. The specification assumes that a JIT translates a single source instruction at a time. This assumption matches real-world BPF JIT implementations and obviates the need to reason about translating entire programs.
- An automated proof strategy that scales to practical BPF JITs (§5). Building on Serval [68], Jitterbug uses symbolic evaluation [10, 89] to produce a satisfiability query that encodes the semantics of a JIT implementation, the semantics of source BPF code, and the semantics of target machine code produced by the JIT. It then discharges the query using an SMT solver [21]. Since Serval was designed to reason about systems whose code is statically known, it cannot be used to verify symbolic instructions (e.g., with symbolic fields, at symbolic addresses) generated by the symbolic evaluation of a JIT. Jitterbug addresses this challenge with a symbolic evaluation strategy that can reason about such symbolic code.
- An approach to writing JITS in a domain-specific language (DSL) based on C (§6). The Jitterbug DSL is a shallow embedding of a structured subset of C in Rosette [88, 89], which extends Racket [29] for symbolic reasoning. That is, the Jitterbug DSL implements a subset of C as a Rosette library. We write new JITS in the DSL, which simplifies verification and enables synthesis of JIT optimizations [59, 82]. Jitterbug automates the step of translating JITS written in the DSL to C through an (unverified) extraction mechanism. We verify existing JITS by manually translating their C code to Rosette.
- Experience with using Jitterbug to build a BPF JIT for RV32, find and fix bugs in five existing BPF JITS, perform code review, develop optimizations, and port a JIT for a stack machine [65], all with low verification overhead (§7). One of the bugs has led to a clarification in the RISC-V instruction-set manual. We report on the iterative process of improving Jitterbug and upstreaming JIT code to the Linux kernel.

To our knowledge, Jitterbug is the first to provide a specification that rules out bugs in practical JIT implementations, and a proof strategy that scales automated verification to a class of compilers. It demonstrates the feasibility of building a verified component (i.e., the BPF JIT) within a large, unverified system under active development (i.e., the Linux kernel), through careful design of specification and proof strategy. This paper describes our design decisions and the rationale behind them (§8).

2 Related work

Code downloading for extensible systems. The Xerox Alto allows applications to customize and optimize the system through microcode [51, 85]. It pioneered the use of packet filters for demultiplexing, debugging, and monitoring.

The CMU/Stanford Packet Filter [62] introduced a stack-based virtual machine into the 4.3BSD kernel to interpret packet filters. To enable more efficient implementations, the Berkeley Packet Filter (BPF) [61] adopts a register-based virtual machine instead, which consists of two 32-bit registers and a scratch memory. BPF has gained a wide adoption in BSD and Linux kernels. Besides BPF, DTrace [12] and Lua on NetBSD [90] are two other in-kernel virtual machines.

A redesign of BPF in the Linux kernel started in 2014, first as an optimization of the internal representation of BPF instructions for 64-bit architectures [83]. It has since grown into a full RISC-like virtual machine, with 64-bit general-purpose registers, flexible control flow (e.g., bounded loops and BPF-to-BPF calls), and safe access to kernel memory. The generality and expressiveness have led to an explosion of tools and systems based on BPF, ranging from networking [38], security [79], tracing [35], to storage [7], virtualization [1, 71], and hardware offloading [43]. The new design is also called “extended BPF” or simply “BPF” in the Linux kernel, while the original design is referred to as classic BPF to avoid ambiguity. Unless otherwise noted, we follow this terminology and use BPF to refer to the new design. This paper focuses on building verified JITS for BPF.

More generally, the exokernels [26] demonstrate a diverse set of mechanisms for code downloading, such as accelerating packet filtering using JIT compilation [25], sandboxing machine code [92] using software-based fault isolation [77, 91], and analyzing file-system metadata using an in-kernel virtual machine [40]. Other extensibility mechanisms include using safe languages [6, 28, 55] and proof-carrying code [67].

Correctness of JIT compilation. Just-in-time compilation (JIT) is a well-studied dynamic code generation technique dating back to Lisp [3, 42] and regular expressions in the QED text editor [76, 87]. It has also been used for dynamically typed languages [14], emulators [5], and specialization [60, 73].

This paper considers JITS that are realized as static compilers, using static register allocation and performing no garbage collection for memory management. In contrast to sophisticated dynamic code generation systems such as those for Java or JavaScript, this simplicity makes static JITS applicable to a restricted environment such as the kernel [24].

There is a rich literature on compiler correctness. Readers may refer to Young [98] and Leroy [54] for overviews. Compilers, especially optimizing compilers, have multiple intermediate representations and translation passes, whereas the JITS considered in this paper are much simpler and resemble a one-pass compiler. On the other hand, compilers usually
output assembly code, relying on a separate assembler and linker (e.g., GNU as and ld) to produce final machine code. The JITs run in the kernel and directly produce machine code, effectively combining a compiler, assembler, and linker.

The closest efforts in this area are the verified JITs by Myreen [65] and Jitk [94]. The former translates code in a simple stack-based instruction set to x86-32 (see §7), and is verified using the HOL4 theorem prover [80]. The JIT is implemented in HOL4 and translated to x86-32 machine code by a separate compiler [66]. Jitk builds on the CompCert verified compiler [53] to translate classic BPF to assembly, and is verified using the Coq theorem prover [86]. The JIT is implemented in Coq and extracted to OCaml code; it runs in user space rather than in the kernel due to the dependency on the OCaml runtime, an assembler, and a linker. Both efforts employ clean-slate designs, require manual proofs, and do not have a C implementation. Jitterbug is inspired by these efforts and shares the goal of building verified JITs, but prioritizes applicability to existing systems, proof automation, and implementation that can be reviewed independent of verification.

Compiler testing and fuzzing tools employ effective strategies to randomly generate input programs and check for miscompilation [58]. Csmith [97] and EMI fuzzers [52] have been used to find hundreds of bugs in GCC and LLVM. Kernel fuzzers such as syzkaller and trinity support generation of random BPF programs [23]. Serval [68] implements a bug finder for the compilation of BPF arithmetic and bitwise instructions. These tools generally do not exhaust all execution paths, thus providing no correctness guarantees for JITs.

### Designing verified systems for deployment

Deployability is a desirable goal for formally verified systems, but it requires navigating an extra set of design trade-offs. As the first verified general-purpose microkernel, seL4 [46] pioneered many aspects of the design and deployment processes. For instance, it introduced a Haskell prototype as the bridge between formal methods and kernel developers, separating verification artifacts from the C implementation [45]. It has been deployed as a hypervisor to retrofit unverified, legacy software to power safety-critical systems [37, 47]. Another example is CompCert, the first verified C compiler. It has been integrated into the development process of control software for safety-critical systems [41, 53], replacing unverified compilers that were configured to disable optimizations due to risk concerns.

Cryptographic libraries are an attractive target for verification due to their essential role in security. For example, verified code from EverCrypt/HACL* [75, 99] and Fiat-Crypto [27] is used by Mozilla and Google, respectively. Amazon’s s2n TLS implementation [16] is verified via a combination of manual and automated proofs.

Jitterbug presents a case study in applying formal methods to the BPF JITs in the Linux kernel. It shares these design challenges and addresses them with a precise specification and a proof strategy that scales to practical JIT implementations.

### 3 Case study

This section presents a brief overview of BPF and a case study of the BPF JIT bugs in the Linux kernel, which helped motivate the design of Jitterbug.

#### 3.1 An overview of BPF

The BPF virtual machine consists of 12 explicit 64-bit registers: general-purpose registers R0–R9, a frame pointer R10 that points to a stack memory region, and an internal register AX used by the kernel for rewrites (e.g., constant blinding against JIT spraying attacks [8]). It maintains a program counter PC and a tail-call counter TCC; the latter bounds the number of tail calls (to another BPF program without returning).

Currently, there are a total of 115 instruction opcodes, which can be categorized into the following:
- ALU (arithmetic and bitwise) instructions,
- JMP (unconditional and conditional jump) instructions,
- MEM (1-, 2-, 4-, and 8-byte memory access, and 4- and 8-byte atomic exchange-and-add) instructions,
- CALL to a kernel function or another BPF program; and
- TAIL_CALL and EXIT, which transfer control to another BPF program and the kernel, respectively.

Figure 1 depicts the execution of a BPF program. The input to a BPF program is provided by the kernel. Prologue and epilogue refer to initialization and cleanup code, respectively, for bridging the kernel. The BPF calling convention specifies that R0 holds the return value, R1–R5 pass arguments, and R6–R9 are preserved across the call.

User processes may share data with BPF programs by creating BPF maps in the kernel, which are key/value stores of different data types. Maps may be accessed concurrently by BPF programs and user processes. Though there have been discussions, BPF has so far chosen not to specify a memory consistency model to avoid performance penalties [19].

Each BPF program consists of a sequence of instructions in bytecode (GCC/LLVM can compile C code to BPF). Upon receiving a BPF program from user space, the kernel invokes a checker to analyze whether the program is safe (e.g., free of division by zero, unbounded loops, and uninitialized register accesses) [34]; we refer to it as the BPF checker (rather than “BPF verifier” as by the Linux kernel to avoid ambiguity). If the BPF checker deems the program safe, the kernel invokes the JIT for compilation and attaches the resulting machine
/* rd[0]: upper 32 bits of the destination register */
rd[1]: lower 32 bits of the destination register

if (val < 32) {
    /* rd[1] = rd[0] >> val */
    emit(ARM_MOV_SI(rd[1], rd[0], SRTYPE_LSR, val), ctx);
} else if (val == 32) {
    /* rd[0] = rd[0] >> val */
    emit(ARM_MOV_SI(rd[0], rd[0], SRTYPE_LSR, val), ctx);
} else {
    /* rd[1] = rd[0] >> (val - 32) */
    emit(ARM_MOV_SI(rd[1], rd[0], SRTYPE_LSR, val - 32), ctx);
}

Figure 2: Incorrect result with zero val for RSH64_IMM (Arm32).

instructions to various hook points in the kernel for execution; otherwise, the kernel rejects the program. The JIT therefore considers safe programs only.

3.2 Bugs in BPF JITs

We manually inspected every commit to the BPF JITs in the Linux kernel from May 2014 (when the new BPF design was introduced) to April 2020, and categorized those that fixed JIT correctness bugs for Arm32, Arm64, RV64, x86-64, and x86-64; those for RV32 will be discussed in §7. We consider “correctness bugs” as JITs producing erroneous machine instructions, and exclude non-correctness bugs (e.g., memory leaks during JIT compilation) from the study. In total, there are 41 commits that fixed 82 JIT correctness bugs during this period. See §A.2 for a complete list.

Below we describe some representative bugs we have found using Jitterbug. These bugs are difficult to find even for veteran developers, and were not caught by the existing test suite. They can lead to security vulnerabilities, since the resulting machine instructions run in the kernel and may process input from untrusted sources. For clarity, BPF instructions and registers are in uppercase, while target machine ones are in lowercase.

Subtle architectural semantics. Figure 2 shows an excerpt of the Arm32 JIT for RSH64_IMM, the BPF logical right shift instruction of a 64-bit register by an immediate. Since the target architecture is 32-bit, the JIT uses two machine registers, represented by rd[0] and rd[1], to hold the upper and lower 32 bits of a 64-bit BPF register, respectively. The BPF checker ensures that the shift amount val is within the range [0, 63]. The emitted instructions work as follows:

- when the shift amount val is less than 32, the result of the upper half is simply rd[0] >> val, and the result of the lower half is rd[1] >> val combined with the bits shifted from the upper half, rd[0] << (32 - val);

- the result of the upper half is simply zero, as all the bits are shifted out, and the result of the lower half holds the bits shifted from the upper half.

One subtlety in Arm32 is that a zero immediate in the lsr (logical shift right) instruction means right-shift by 32 bits (i.e., shifting all bits out) [2]: 8F5.1.103. Therefore, when the shift amount val is zero, the instructions produced by the JIT incorrectly set the destination register to zero, instead of behaving as a no-op. This is further complicated by inconsistent semantics in Arm32: a zero immediate in the shift left instruction means a no-op. We fixed the bug by changing the JIT to emit no instructions when val is zero.

Figure 3 shows another subtle bug in the RV64 JIT. Using a pair of auipc+jalr instructions is a standard way to support pc-relative call with a 32-bit offset on RISC-V [96]:

• auipc t1, imm20 appends 12 low-order zero bits to a 20-bit immediate, sign-extends the 32-bit value to 64 bits, adds the sign-extended value to the address of the instruction, and writes the result in register t1;
• jalr ra, imm12(t1) jumps to a target address obtained by adding a sign-extended 12-bit immediate to the register t1 and clearing the least-significant bit of the result for alignment; the address of the instruction following jalr is written to register ra.

One misconception is that auipc+jalr can reach any 32-bit offset in the range [−231, 231 − 1] on 64-bit RISC-V (RV64), by using certain imm20 and imm12 values. Part of the confusion stems from the “RV32I base integer instruction set” chapter in the RISC-V instruction-set manual indicating that auipc+jalr “can jump anywhere in a 32-bit pc-relative address range.” But the same does not hold on RV64: both auipc and jalr sign-extend their results to 64 bits, causing the reachable offset range to shift by −211. Therefore, the range check on rvoff in the JIT is incorrect, which can lead to an off-target jump.

Our report prompted the RISC-V instruction-set manual to add the following clarification: “Note that the set of address offsets that can be formed by pairing LUI with LD, AUIPC with JALR, etc. in RV64 is [−231 − 211, 231 − 211 − 1].” We fixed the bug in the JIT by using the clarified range for checking rvoff.

Subtle machine state. Figure 4 shows an excerpt of the x86-64 JIT for compiling BPF’s JSET64_REG and JSET32_REG (in the form BPF_JMP[32]/BPF_JSET/BPF_X in C). The semantics of “JSET64_REG DST, SRC, OFF” is to perform a conditional
jump when DST&SRC (“bitwise” and) of two 64-bit BPF registers is non-zero and fall through otherwise; the semantics of “JSET32_REG DST, SRC, OFF” is similar, using only the lower 32 bits of both DST and SRC.

Due to the limited number of registers on x86-32, the JIT spills some BPF registers on the stack. For simplicity, suppose that both DST and SRC are on the stack (i.e., both dstk and sstk are true). In this case, the JIT emits instructions to load the lower 32 bits of DST and SRC to eax and ecx respectively. It also emits instructions to load the upper 32 bits to edx and ebx for JSET64_REG; the two registers are uninitialized for JSET32_REG.

One way to implement JSET32_REG is to emit a bitwise and of eax and ecx, followed by a conditional jump if the result is non-zero (i.e., the zf bit in the eflags register is clear). But the JIT emits extra and or instructions that also use edx and ebx, which are uninitialized for JSET32_REG, incorrectly modifying eflags. The bug was not caught by the BPF selftests suite because none of the tests “polluted” edx and ebx with values that would cause the behavior to change. We fixed the bug by moving the last two EMT2 statements under a condition that is*_jmp64 is true.

There are other bugs in the excerpt: when DST is mapped to x86 registers and not spilled on the stack (i.e., dstk is false), the emitted instructions incorrectly clobber the registers, while the semantics of the BPF instructions requires DST not to change. We fixed the bugs by loading DST to eax and ecx, regardless of whether DST is on the stack.

Subtle instruction encoding. Below is an encoding bug in the x86-32 JIT for the BPF LDXB instruction, which loads a byte from memory. As its semantics requires the result to be zero-

```c
if (is_jmp64) {
    EMT3(0xC7, add_1reg(0xC0, IA32_EBX, IA32_EAX),
        ADD_VAR(dst_hi)); /* eax <- dst_hi */
} else {
    EMT2(0x23, add_2reg(0xC0, dst_hi, sreg_hi));
    EMT2(0x23, add_2reg(0xC0, sreg_lo, dreg_lo));
    goto emit_cond_jmp; /* emit conditional jump */
}
```

Figure 4: Incorrect eflags value for JSET32_REG (x86-32).

extended to 64 bits, the JIT attempts to emit “mov dst_hi, 0” to clear the upper 32 bits, using the following C code:

```c
EMT3(0xC7, add_1reg(0xC0, dst_hi), 0);
```

Notice that EMT3 emits 3 bytes, but a correct “mov dst_hi, 0” expects 6 bytes: the opcode 0xC7, the ModR/M byte formed by add_1reg(0xC0, dst_hi), followed by 4 bytes of zeros as the immediate. The consequence is not merely an incorrect mov: it also “swallows” 3 bytes from the next instruction, breaking the instruction stream and altering the meaning of the subsequent instructions. We fixed the bug by emitting “xor dst_hi, dst_hi” instead, which is also shorter (2 bytes).

### 3.3 Summary

Compared to the bugs in classic BPF JITs [15, 94], those in today’s BPF JITs are more sophisticated due to the increased power of the BPF virtual machine. On the other hand, architecture-independent checks for BPF programs such as division by zero are now performed by the BPF checker, eliminating the need for the JITs to consider such cases.

While the Arm and RISC-V JITs emit instructions using well-defined macros (e.g., Figure 2) or functions (e.g., Figure 3), the x86 JITs directly emit raw bytes (e.g., Figure 4), partly due to the lack of a uniform instruction format on x86. Jitterbug therefore needs to model the semantics of their target architectures precisely; for x86, this means reasoning at the level of raw instruction bytes.

### 4 Specification

Jitterbug aims to rule out subtle bugs in BPF JITs through a formal specification, which is the focus of this section.

We begin with an intuitive description of what it means for a JIT to be correct. At a high level, running the machine code emitted by a JIT for a given source program should be equivalent to running a BPF interpreter with that source program. For example, both should compute the same return value and invoke the same kernel functions with the same arguments; any deviation indicates a bug. Jitterbug captures this intuition as a JIT correctness specification (§4.1).

Specifications like this are usually proved by induction, and the key to carrying out the proof is finding the right inductive invariant—a property preserved by the JIT translation of each individual source instruction. Inspired by the structure of the existing BPF JITs in the Linux kernel, Jitterbug introduces a stepwise specification that serves as our inductive invariant. As shown in Figure 5, this specification consists of a set of properties satisfied by individual translation steps, such as the generation of machine code for a single BPF instruction. Using the Lean theorem prover [22], we prove that any JIT that satisfies the stepwise specification implies our intuitive notion of correctness. This proof serves as the metatheory for...
Jitterbug (§4.2). The stepwise specification itself is proved automatically for each JIT.

To illustrate how to apply the stepwise specification to prevent bugs, we use the BPF JIT for RV32 as an example. We also analyze alternative JIT implementations to demonstrate the generality of the specification (§4.3).

We end this section with a discussion of the limitations of Jitterbug’s specification and how it relates to prior compiler correctness specifications (§4.4).

4.1 JIT correctness

Formalizing JIT correctness requires formalizing the behavior of the JIT, source BPF programs, and target machine programs, as follows.

First, we model a JIT as a function JITCompile. It takes a source program code_s and JIT context ctx as input, and returns either a target program code_t on success, denoted as JITCompile(code_s,ctx) = code_t; or ⊥, indicating compilation error. Both source and target programs are represented as partial maps from addresses to instructions; some addresses may be unmapped. We define code ⊆ code_t to mean that any address that maps to some instruction in code maps to the same instruction in code_t.

The JIT context ctx is an implementation-defined data structure. It usually contains compiler configurations (e.g., the base address of the target program allocated by the kernel, denoted by ctx[base]) and analysis results of the source program, which are used by the JIT for code generation. We assume that the JIT context is well-formed with respect to the source program; this assumption is captured using a predicate wellFormed(code_s,ctx) specified by JIT developers. For example, one may specify that ctx[base] is properly aligned.

Next, we model the execution of both source and target programs as abstract machines, described by a set of states Σ and a state transition function step. Given a state σ ∈ Σ, we write σ[·] to refer to a specific component of the state. For example, σ[pc] is the value of the program counter.

The step function takes as input a state σ, a program code, and an oracle denoted by nd. The oracle nd is an infinite sequence of nondeterministically chosen bytes, which are used for modeling external interactions with the kernel (e.g., values loaded from BPF maps or returned by calls to kernel functions). Given these inputs, the step function produces the next state and a trace of externally visible events generated by executing the instruction at the program counter, code[σ[pc]]. The execution gets stuck if it triggers undefined behavior (e.g., the address σ[pc] is unmapped in code_t). As shorthands, we write (σ, code, nd) → (σ’, tr) to mean step(σ, code, nd) = (σ’, tr), and (σ, code, nd) →* (σ’, tr) to mean that state σ’ is reachable from zero or more applications of step starting from state σ, with concatenated trace tr.

The exact content of events is defined by each machine. For example, consider the BPF machine in Jitterbug. It defines the following events: load(addr, val), store(addr, val), call(addr, args, val), atomic_begin, and atomic_end. It models each memory load as returning a fresh value provided by the BPF maps and producing a load event in the trace, since BPF maps may be modified outside the execution of a BPF program (§3.1). Each step may produce zero or more events. For example, the execution of XADD32 (32-bit atomic exchange-and-add) produces atomic_begin, load, store, and atomic_end.

This model assumes read-only code, which prohibits jits that produce self-modifying code [65]. It also assumes that the execution of a program is deterministic [53: §2.1], since the next state is uniquely determined by the current state, code, and oracle. Both assumptions match the BPF JIts in Linux.

In order to reason about the start and end of execution, each machine defines two predicates:

• initial(x,ctx,σ), where σ is an initial state for input x and JIT context ctx; and
• final(σ’,v), where σ’ is a final state with return value v.

Recall that the JIT considers only safe source programs. For example, the Linux kernel rejects BPF programs that the BPF checker deems unsafe (§3.1). We capture this guarantee with a predicate safe(code), which specifies that executing code always reaches a final state (i.e., the execution terminates without triggering any undefined behavior):

∀ x, σ, nd. initial(x,ctx,σ) →
∃ σ’, tr, v. (σ, code, nd) →* (σ’, tr) ∧ final(σ’, v).

In addition, since a target program generated by the JIT runs within the kernel, it must behave like a regular function and preserve the corresponding calling convention: for example, stack pointer and callee-saved registers must hold the same values before and after the execution. We capture these requirements in the architectural safety predicate A(σ_f, σ_f’), which constrains the initial and final values of all preserved target registers r to be the same, i.e., σ_f[r] = σ_f’[r].

Using our model, we define JIT correctness as follows.

**Definition 1 (JIT correctness).** A JIT is correct if for any safe source program code_s, well-formed JIT context ctx, and target program code_t generated by the JIT such that safe(code_s) ∧ wellFormed(code_s,ctx) ∧ JITCompile(code_s,ctx) = code_t, the following two conditions hold:
1. The execution of source program \( code_S \) and that of target program \( code_T \) produce the same trace and return value.

\[
\forall x, \sigma_S, \sigma_T, nd, tr, v.
\]

\[
\text{initial}_S(x, ctx, \sigma_S) \land \text{initial}_T(x, ctx, \sigma_T) \rightarrow
\]

\[
\left( \exists \sigma'_S. \langle \sigma'_S, code_S, nd \rangle \Rightarrow^* \langle \sigma'_S, tr \rangle \land \text{final}_T(\langle \sigma'_T, v \rangle) \right) \leftrightarrow
\]

\[
\left( \exists \sigma'_T. \langle \sigma'_T, code_T, nd \rangle \Rightarrow^* \langle \sigma'_T, tr \rangle \land \text{final}_T(\langle \sigma'_T, v \rangle) \right).
\]

2. Any final state reachable by executing target program \( code_T \) satisfies architectural safety.

\[
\forall x, \sigma_T, \sigma'_T, nd, tr, v. \text{initial}_T(x, ctx, \sigma_T) \land \text{final}_T(\langle \sigma'_T, v \rangle) \land
\]

\[
\left( \sigma'_T, code_T, nd \right) \Rightarrow^* \langle \sigma'_T, tr \rangle \Rightarrow A(\sigma_T, \sigma'_T).
\]

The first property can be viewed as a bisimulation between source and target machines [54; §2]: the JIT produces a target program that preserves the behavior of the source program, and any behavior of the target program is permitted by the source program. Additionally, given that the source program is safe, this property implies that the target program produced by the JIT is safe (i.e., terminates without undefined behavior). The second property further requires the target program to correctly save and restore the corresponding architectural state. Both guarantees are critical for in-kernel execution.

### 4.2 Stepwise specification

Given Definition 1, our goal is to devise a stepwise specification (i.e., an inductive invariant) that both implies JIT correctness and is amenable to automated verification. We achieve this goal by imposing structure on the JIT compilation process so that we can reason about the correctness of individual compilation steps, as follows.

Inspired by the existing BPF JITS in the Linux kernel, we suppose that the JIT generates a target program in a per-instruction fashion. Specifically, the target program consists of machine instructions for the prologue, each source instruction, and the epilogue (Figure 1). We do not assume any particular code layout. For example, one may produce the target program sequentially:

\[
\text{code}_T = \text{EmitPrologue}(ctx)
\]

\[
\text{for } i \text{ in } [0, \text{codes}_S[-1]]:
\]

\[
\text{code}_T \leftarrow \text{EmitInstruction}(ctx, i, \text{codes}_S[i])
\]

\[
\text{code}_T \leftarrow \text{EmitEpilogue}(ctx)
\]

We formalize our assumptions about the JIT below.

**Definition 2 (JIT assumptions).** We assume that for any safe source program \( code_S \), well-formed JIT context \( ctx \), and target program \( code_T \) produced by a JIT such that \( \text{safe}(code_S) \wedge \text{wf}(code_S, ctx) \wedge \text{JITcompile}(code_S, ctx) = code_T \), the target program \( code_T \) contains the machine instructions produced by each translation step:

- \( \exists p. \text{EmitPrologue}(ctx) = p \wedge p \subseteq code_T \).
- \( \forall i, \text{insn}. \text{codes}_S[i] = \text{insn} \rightarrow \exists p. \text{EmitInstruction}(ctx, i, \text{insn}) = p \wedge p \subseteq code_T \).
- \( \exists p. \text{EmitEpilogue}(ctx) = p \wedge p \subseteq code_T \).

With these assumptions, the stepwise specification boils down to the correctness of each translation step: EmitPrologue, EmitInstruction, and EmitEpilogue. Jitterbug allows developers to provide two relations as invariants maintained by their JIT implementations:

- \( \sigma_S \xrightarrow{ctx} \sigma_T \) relates source state \( \sigma_S \) and target state \( \sigma_T \) with respect to JIT context \( ctx \). For example, it may specify that the value of a BPF register in \( \sigma_S \) is equal to that of the machine register the JIT uses to realize the BPF register in \( \sigma_T \).
- \( I_{\text{ctx}}(\sigma_{T_0}, \sigma_T) \) relates initial target state \( \sigma_{T_0} \) and non-final target state \( \sigma_T \) with respect to JIT context \( ctx \). For example, the prologue usually saves callee-saved registers to a designated memory region; \( I_{\text{ctx}} \) may specify that the values of callee-saved registers in \( \sigma_{T_0} \) are equal to those in that region in \( \sigma_T \).

Below we describe the correctness definition for each translation step. We denote the empty trace as \( \epsilon \).

**Definition 3 (Prologue correctness).** A JIT emits a correct prologue if executing the prologue results in a target state that establishes the invariants, and produces an empty trace:

\[
\forall code_S, ctx, p, x, \sigma_S, \sigma_T, nd, tr, v. \text{wf}(code_S, ctx) \wedge
\]

\[
\text{EmitPrologue}(ctx) = p \wedge
\]

\[
\text{initial}_S(x, ctx, \sigma_S) \land \text{initial}_T(x, ctx, \sigma_T) \rightarrow
\]

\[
\exists \sigma'_T. \langle \sigma'_T, p, nd \rangle \Rightarrow^* \sigma'_T, \epsilon \land \langle \sigma_S \xrightarrow{ctx} \sigma_T \rangle \land I_{\text{ctx}}(\sigma_{T_0}, \sigma_T).
\]

**Definition 4 (Per-instruction correctness).** A JIT emits correct target instructions for a given source instruction if executing the emitted instructions results in a target state that preserves the invariants, and produces the same trace as executing the source instruction:

\[
\forall code_S, ctx, i, insn, p, \sigma_S, \sigma_T, \sigma_{T_0}, nd, tr, v. \text{wf}(code_S, ctx) \land
\]

\[
\text{code}_S[i] = \text{insn} \land \sigma_S[\text{pc}] = i \land
\]

\[
\text{EmitInstruction}(ctx, i, \text{insn}) = p \wedge
\]

\[
\langle \sigma_S, code_S, nd \rangle \Rightarrow^* \langle \sigma'_S, tr \rangle \land \langle \sigma_S \xrightarrow{ctx} \sigma_T \rangle \land I_{\text{ctx}}(\sigma_{T_0}, \sigma_T) \rightarrow
\]

\[
\exists \sigma'_T. \langle \sigma'_T, p, nd \rangle \Rightarrow^* \sigma'_T, \epsilon \land \langle \sigma_S \xrightarrow{ctx} \sigma_T \rangle \land I_{\text{ctx}}(\sigma_{T_0}, \sigma_T).
\]

**Definition 5 (Epilogue correctness).** A JIT emits a correct epilogue if executing the epilogue results in a final target state that satisfies architectural safety, and produces the same return value as in the source final state and an empty trace:

\[
\forall code_S, ctx, p, \sigma_S, v, \sigma_T, \sigma_{T_0}, nd, v. \text{wf}(code_S, ctx) \land
\]

\[
\text{EmitEpilogue}(ctx) = p \wedge
\]

\[
\text{final}_S(\sigma_S, v) \land \langle \sigma_S \xrightarrow{ctx} \sigma_T \rangle \land I_{\text{ctx}}(\sigma_{T_0}, \sigma_T) \rightarrow
\]

\[
\exists \sigma'_T. \langle \sigma'_T, p, nd \rangle \Rightarrow^* \sigma'_T, \epsilon \land \text{final}_T(\sigma'_T, v) \land A(\sigma_{T_0}, \sigma'_T).
\]

Together, these three properties imply JIT correctness given the JIT assumptions. We prove the following theorem in Lean:

**Theorem 1 (Stepwise soundness).** JIT assumptions \land prologue correctness \land per-instruction correctness \land epilogue correctness \rightarrow JIT correctness.

With Theorem 1 as a meta-theorem, Jitterbug proves the correctness of a JIT implementation by proving the properties in
Definitions 3, 4, and 5 via automated verification (see §5). The JIT context well-formedness \( \text{wf} \) and assumptions are assumed to be correct and trusted. The invariants \((\sim_{\text{ctx}} \text{ and } L_{\text{ctx}})\) are untrusted: if incorrect invariants are provided, verification fails.

### 4.3 Applying the stepwise specification

The stepwise specification is parameterized by assumptions (well-formedness of JIT context \( \text{wf} \)) and invariants \((\sim_{\text{ctx}} \text{ and } L_{\text{ctx}})\), which reflect how JIT developers intend to establish correctness. We illustrate how to apply the stepwise specification to the BPF JIT for RV32 by specifying the assumptions and invariants regarding registers, program counters, and memory. We also describe how one may specify them for the alternative JIT implementations we have considered.

Figure 6 shows the design of the BPF JIT for RV32. The upper half denotes the BPF state, including registers \((R0-R10, AX)\), counters (tail-call counter \( \text{TCC} \) and program counter \( PC \)), a stack memory region, and maps of shared data (§3.1). The lower half denotes the RV32 state, including registers \((fp, \text{sp}, a0-a7, s1-s6, t6; \text{those not mapped to BPF registers are omitted})\), a machine program counter \( PC \), and memory.

#### Registers.

Since BPF registers are 64-bit and RV32 is a 32-bit architecture, the JIT realizes each BPF register using either a pair of RV32 registers (e.g., \( R1 \) using \( a0 \) and \( a1 \)) or 64 bits in the “spilled registers” memory region (e.g., \( R6 \)). This register mapping is static and pre-determined, eliminating the need for register allocation at compilation time. Other BPF JITs in the Linux kernel use similar register mappings.

The register mapping is handcrafted to achieve good performance. For instance, recall that BPF designates \( R1-R5 \) to pass function-call arguments, while the RISC-V calling convention uses \( a0-a7 \), plus the stack if needed [20]. To minimize register save and restore, the JIT realizes \( R1-R4 \) using \( a0-a7 \). For \( R5 \), the JIT emits instructions to push the corresponding \( s3, s4 \) to the “argument” memory region before the call.

To specify the relation \( \sigma_S \sim_{\text{ctx}} \sigma_T \) between source and target states, let \( \phi_{\text{reg}}(ctx, \sigma_T, r) \) denote the value stored at the target location(s) to which a BPF register \( r \) is mapped (e.g., \( R1 \) mapped to \( a0, a1 \) with respect to JIT context \( ctx \)). A strawman approach is to require a strict equivalence: \( \sigma_S[r] = \phi_{\text{reg}}(ctx, \sigma_T, r) \) for every BPF register \( r \). With this relation, the stepwise specification would require that if every BPF register and the mapped locations contain equivalent values initially, their values remain equivalent after executing a BPF instruction and the emitted machine instructions, respectively. One such example is the partial specification used by the BPF bug finder in Serval [68: §7]; the specification is partial because it does not support reasoning about control flow (e.g., program counters) or memory and cannot be used to prove JIT correctness.

While it is useful for finding bugs, the strawman relation is too restrictive for verification. First, if a BPF program does not use a certain register, it should be safe for the JIT to skip emitting code for initializing the corresponding target locations, but doing so violates the strict equivalence. Second, the relation is difficult to establish in the presence of calls. To see why, consider the BPF register \( R1 \), which is not preserved across a BPF CALL instruction (§3.1). \( R1 \) is thus considered uninitialized after the call as per the BPF semantics (the BPF checker ensures that \( R1 \) will be written to before any further use). On the other hand, \( R1 \) is mapped to \( a0, a1 \), both of which hold the return value after the call as per the RISC-V calling convention (the JIT emits instructions to further copy their values to \( s1, s2 \) to match the BPF calling convention for \( R0 \)). Therefore, \( R1 \) and the corresponding \( a0, a1 \) do not hold equivalent values after the call, which violates the strict equivalence.

To relax the strict equivalence and give the JIT more freedom regarding uninitialized BPF registers, we augment the state of the BPF machine with an initialized set, which represents the set of registers that are initialized at this point: the set is updated based on the semantics of each BPF instruction. For example, \( \text{DST} \) is added to the set after “MOV64_\text{IMM} \text{DST, IMM},” as it is written to by the instruction. Similarly, \( R1-R5 \) are removed from the set after \( \text{CALL} \), as they are not preserved across the call and become uninitialized. In doing so, it suffices to require equivalence \( \sigma_S[r] = \phi_{\text{reg}}(ctx, \sigma_T, r) \) for every BPF register \( r \in \sigma_S[\text{initialized}] \), effectively excluding uninitialized ones.

#### Program counters.

Let \( \phi_{\text{pc}}(ctx, i) \) denote the target address to which the \( i \)-th BPF instruction is mapped in JIT context \( ctx \). This is useful for a JIT to implement the compilation of jump instructions. It also allows us to relate program counters in BPF and machine states as an invariant \( \phi_{\text{pc}}(ctx, \sigma_S[\text{pc}]) = \sigma_T[\text{pc}] \).

To define \( \phi_{\text{pc}} \), one simple approach is to require the JIT to emit a fixed number of machine instructions for each BPF instruction (e.g., by padding with NOPs) [33]. In this case, we have \( \phi_{\text{pc}}(ctx, i) = ctx[\text{base}] + i \times N \), where \( ctx[\text{base}] \)
is the starting address of the emitted machine instructions determined by JIT context and \( N \) is a pre-determined number of machine instructions large enough to compile any BPF instruction. This is simple to specify and implement, but the emitted code wastes space and CPU cycles.

A more efficient approach is to emit a variable number of machine instructions for each BPF instruction. For example, the BPF JITs in the Linux kernel maintain an offset table in the JIT context to map each BPF instruction index to an offset into the emitted code; in this case \( \phi_{pc}(ctx, i) \) is defined by simply consulting the offset table. The JITs construct the offset table by repeating the compilation process until the table converges, or fail if an upper bound on the number of iterations is reached (e.g., 16 in the BPF JIT for RV32).

For flexibility, we choose not to specify how to construct the offset table in the JIT context. Instead, we specify the property a valid JIT context should satisfy. A key observation is that such JITs emit consecutive blocks of machine instructions, one block for each BPF instruction. As a result, the difference between the target addresses for a BPF instruction \( code_S[i] \) and its successor \( code_S[i + 1] \) must be equal to the number of bytes emitted for \( code_S[i] \). We capture the observation using the well-formedness predicate \( \Psi \) over source program \( code_S \) and JIT context \( ctx \) for any \( i \)-th BPF instruction:

\[
\text{EmitInstruction}(ctx, i, code_S[i]) = p \\
|p| = \phi_{pc}(ctx, i + 1) - \phi_{pc}(ctx, i).
\]

Here \(|p|\) denotes the length of machine instructions \( p \) (in bytes). This allows for both NOP-padding and the more sophisticated JIT implementations such as those in the Linux kernel. Note that this is an assumption on the validity of the JIT context, which does not rule out bugs in the construction of the offset table (see §4.4). A JIT may validate the offset table by checking that this predicate holds at compilation time.

**Memory.** One approach to relating the memory state of source and target machines, denoted by \( \sigma_S[\text{mem}] \) and \( \sigma_T[\text{mem}] \), respectively, is to require \( \sigma_S[\text{mem}](a) = \sigma_T[\text{mem}](a) \) for every address \( a \) [65], where memory is modeled as a map from addresses to values. But this approach assumes a closed system (see §7 for such a JIT) and does not fit BPF. For example, both user processes and other BPF programs may concurrently modify memory to which a BPF program has access; therefore, consecutive loads from the same address in the BPF program may return different values. A further complication is that BPF does not specify a memory consistency model (§3.1), effectively assuming that of the underlying architecture.

We observe that a BPF JIT does not need to reason about the behavior of concurrent memory accesses [13, 56]. Instead, the goal is to faithfully translate BPF memory accesses to ones in the target machine, which is a simpler task. Based on this observation, we employ a hybrid approach to specify the invariants for BPF JITs using traces and maps, as follows.

Each target machine models memory as consisting of two disjoint parts, one corresponding to shared memory and the other for internal use (Figure 6). The memory layout used by a JIT determines which target addresses are shared and which are internal. The internal memory is simply a map from addresses to values, since it is private to each execution and the effects are not externally visible. The shared memory captures memory-related effects using events in a trace (§4.1). Since the BPF machine adopts the memory model of the underlying architecture, Jitterbug relates the traces of the BPF and target machines by using the same memory model for both; i.e., the BPF and target traces are drawn from the same set of all possible memory events. Given this correspondence, it suffices to require the traces produced by the BPF and target machines to be identical.

The requirement of having identical traces suffices for the BPF JITs. One exception is that older versions of the BPF JIT for Arm64 use Arm’s exclusive access instructions in a busy loop [2: §B2], which violates the requirement. Newer versions of the JIT have switched to using atomic instructions, which satisfies the requirement. We decide not to relax the requirement of having identical traces to keep the specification simple.

### 4.4 Discussion and limitations

Jitterbug’s JIT correctness (Definition 1), especially the use of traces, is inspired by the specification of CompCert [54]. Jitterbug’s specification differs in the following ways. First, in-kernel execution imposes stricter requirements on the source program (e.g., determinism, termination, and absence of undefined behavior), allowing us to prove stronger properties. Second, Jitterbug uses fine-grained models of target architectures to precisely reason about low-level state (e.g., program counter and stack pointer), whereas CompCert uses a more abstract semantics for assembly [64: §5] and relies on a separate assembler and linker. Third, the per-instruction compilation process of such JITs enables us to develop a stepwise specification amenable to automated verification.

Jitterbug trusts the correctness of the assumptions (§4.2). Therefore, it cannot catch bugs in the JIT context (e.g., offset-table construction) or layout of the target program. We manually examine the existing BPF JIT correctness bugs in the Linux kernel (§3.2), and determine that out of the 82 bugs, the specification can catch all but two bugs, both in offset-table construction. This shows the effectiveness of the specification.

Jitterbug’s specification permits “null” JIT implementations that fail on all source programs; we use existing test suites (e.g., the BPF selftests) to assess the feature completeness of JITs. It focuses on the JIT and cannot rule out bugs in the BPF checker, memory management for code images, or how the kernel uses the JIT. It does not model the instruction cache or memory permissions, relying on the kernel to correctly flush the cache and set up permissions. It does not provide any guarantees against microarchitectural timing channels [32, 48].
In a restricted setting where the JIT emits straight-line code without any branches, \( f(\bar{x}) \) is simply the number of emitted instructions. The BPF bug finder in Serval and synthesis-based superoptimizers [72] all assume this setting and use the corresponding basic realization of \( f \). But Jitterbug considers JItTs that can emit code with branches, and when executing such code, the target machine can take a different number of steps depending on the input state. This rules out the straightforward realization of \( f \) that counts the number of emitted instructions.

To illustrate the challenge of computing \( f \) in our setting, consider the instructions emitted by the RV32 JIT for the BPF instruction \"JNE64_REG DST, SRC, OFF\" (jump to offset \( OFF \) if the values in \( DST \) and \( SRC \) differ). The JIT may emit different blocks of RV32 instructions, conditioned on whether it spills the registers (requiring \textit{lw} to load from stack) or the offset requires a far jump (\textit{jal} or \textit{auipc}+\textit{jalr}). Figure 8 shows three examples of these blocks and the \( f(\bar{x}) \) values for executing them, which vary depending on the register state and instructions. In general, constructing \( f \) requires human insight [63, 98], so Jitterbug allows JIT developers to provide a manually constructed \( f \). In practice, however, Jitterbug can automate the construction of \( f \) for BPF JITs as follows.

To compute \( f \), Jitterbug requires the target interpreter to maintain the symbolic program counter in the form \texttt{base+offset}, where \texttt{base} is the (symbolic) starting address of instructions. Maintaining this form is straightforward for most instructions. For instructions with subtler semantics, the interpreter achieves this by rewriting the program counter via \textit{symbolic optimization} [68, 74]. For example, RISC-V’s \textit{jalr} sets the least-significant bit of the program counter to zero (§3.2), causing it to take the form \texttt{(base+offset) & mask}. The interpreter rewrites this expression by dropping the mask and checking that the resulting expression is equivalent (i.e., that the program counter is properly aligned).

Given a program counter of the form \texttt{base+offset}, Jitterbug provides a reusable procedure for constructing \( f \) through symbolic evaluation. It extracts the \texttt{offset} from the program counter expression and applies a simple rule: stop symbolic evaluation either if the \texttt{offset} is concrete but leaves the block of emitted instructions, or if the \texttt{offset} becomes symbolic. The intuition is that branching with a symbolic offset likely leaves the block, because the JIT generally produces such branching instructions by consulting the \texttt{offset} table in the JIT context (§4.3), which is considered symbolic for verification. Internal branching tends to have a concrete \texttt{offset}, for which symbolic evaluation continues.

This procedure guesses an \( f \) for verifying that the target machine reaches a desired state after taking \( f(\bar{x}) \) steps. It does not guarantee to find the right \( f \) if one exists, though it is sufficient for all the JITs we have studied and works well in practice. The procedure is untrusted: choosing a wrong \( f \) causes the target machine to either get stuck or enter a state that violates the inductive invariant, but it does not cause an erroneous JIT implementation to pass verification.

---

**Figure 7:** Jitterbug’s verification pipeline. Shaded boxes denote inputs provided by JIT developers.

**5 Proving JIT correctness**

Jitterbug extends automated verification to JIT correctness, a form of compiler correctness tailored for in-kernel execution. This section describes how Jitterbug achieves the automation.

As shown in Figure 7, Jitterbug provides the stepwise specification and the executable semantics (i.e., interpreters) for BPF and various architectures. It asks JIT developers for a JIT implementation, and assumptions and invariants regarding the implementation. All the inputs are written in the Rosette language (the JIT is written in the DSL described in §6).

Jitterbug builds on Serval for automated verification [68]. It invokes Rosette to reduce all the inputs to symbolic constraints via symbolic evaluation, and an SMT solver to check the satisfiability of these constraints. For symbolic evaluation to terminate, the JIT implementation and the execution of both interpreters must be free of unbounded loops [88]. The BPF JITs satisfy this requirement.

Below we highlight three key challenges in automated verification of JITs and how Jitterbug addresses these challenges.

**Instantiation of existential quantifiers.** To prove the stepwise specification, Jitterbug has to construct \textit{some} execution of target instructions emitted by the JIT and show that the execution exhibits the same behavior as that of a source instruction. Automating the construction is challenging.

To see why, consider the specification for per-instruction correctness (§4.2). Letting \( \bar{x} \) stand for the universally quantified variables in Definition 4, the specification says that the target machine executes some finite number of steps, \( k \), to produce a state \( \sigma_T' \) that satisfies the inductive invariant \( P \). Making the number of steps \( k \) explicit, we can write the correctness formula as

\[
\forall \bar{x}. \exists k. P(\bar{x}, k), \text{or equivalently,} \exists f. \forall \bar{x}. P(\bar{x}, f(\bar{x})),
\]

where \( f \) is a Skolem function that computes the right \( k \) for each combination of the variables \( \bar{x} \). The verification problem that Jitterbug solves therefore involves constructing \( f \). In other words, Jitterbug must determine the number of steps to run symbolic evaluation with emitted instructions, and this value \( f(\bar{x}) \) may depend on the source program, JIT context, source and target states, etc.

---

In a restricted setting where the JIT emits straight-line code without any branches, \( f(\bar{x}) \) is simply the number of emitted instructions. The BPF bug finder in Serval and synthesis-based superoptimizers [72] all assume this setting and use the corresponding basic realization of \( f \). But Jitterbug considers JItTs that can emit code with branches, and when executing such code, the target machine can take a different number of steps depending on the input state. This rules out the straightforward realization of \( f \) that counts the number of emitted instructions.

To illustrate the challenge of computing \( f \) in our setting, consider the instructions emitted by the RV32 JIT for the BPF instruction \"JNE64_REG DST, SRC, OFF\" (jump to offset \( OFF \) if the values in \( DST \) and \( SRC \) differ). The JIT may emit different blocks of RV32 instructions, conditioned on whether it spills the registers (requiring \textit{lw} to load from stack) or the offset requires a far jump (\textit{jal} or \textit{auipc}+\textit{jalr}). Figure 8 shows three examples of these blocks and the \( f(\bar{x}) \) values for executing them, which vary depending on the register state and instructions. In general, constructing \( f \) requires human insight [63, 98], so Jitterbug allows JIT developers to provide a manually constructed \( f \). In practice, however, Jitterbug can automate the construction of \( f \) for BPF JITs as follows.

To compute \( f \), Jitterbug requires the target interpreter to maintain the symbolic program counter in the form \texttt{base+offset}, where \texttt{base} is the (symbolic) starting address of instructions. Maintaining this form is straightforward for most instructions. For instructions with subtler semantics, the interpreter achieves this by rewriting the program counter via \textit{symbolic optimization} [68, 74]. For example, RISC-V’s \textit{jalr} sets the least-significant bit of the program counter to zero (§3.2), causing it to take the form \texttt{(base+offset) & mask}. The interpreter rewrites this expression by dropping the mask and checking that the resulting expression is equivalent (i.e., that the program counter is properly aligned).

Given a program counter of the form \texttt{base+offset}, Jitterbug provides a reusable procedure for constructing \( f \) through symbolic evaluation. It extracts the \texttt{offset} from the program counter expression and applies a simple rule: stop symbolic evaluation either if the \texttt{offset} is concrete but leaves the block of emitted instructions, or if the \texttt{offset} becomes symbolic. The intuition is that branching with a symbolic offset likely leaves the block, because the JIT generally produces such branching instructions by consulting the \texttt{offset} table in the JIT context (§4.3), which is considered symbolic for verification. Internal branching tends to have a concrete \texttt{offset}, for which symbolic evaluation continues.

This procedure guesses an \( f \) for verifying that the target machine reaches a desired state after taking \( f(\bar{x}) \) steps. It does not guarantee to find the right \( f \) if one exists, though it is sufficient for all the JITs we have studied and works well in practice. The procedure is untrusted: choosing a wrong \( f \) causes the target machine to either get stuck or enter a state that violates the inductive invariant, but it does not cause an erroneous JIT implementation to pass verification.
Symbolic evaluation of symbolic instructions. As shown in Figure 8, the JIT may emit different blocks of target instructions for a given source instruction opcode. When Jitterbug symbolically evaluates a JIT implementation, it produces a symbolic representation of all of these instruction blocks. This representation takes the form of symbolic instructions that may contain symbolic values in register and immediate fields. To verify the JIT, Jitterbug must then evaluate the target interpreter on both a symbolic input state and a symbolic program. This is in contrast to prior work on verifying systems code such as Serval, where the input state is symbolic but the program itself is concrete (e.g., all register and immediate fields are concrete bytes). Reasoning about symbolic programs both magnifies existing challenges to scaling verification and creates new ones. We discuss one example of each.

The first challenge is path explosion. While common to all tools based on symbolic evaluation, this problem becomes exponentially worse in the presence of symbolic code. For example, the BPF JIT for RV64 compiles LD64_IMM to a variable number of instructions to load a 64-bit immediate in chunks, selecting each instruction based on the chunk value and destination register. This amounts to reasoning about a total of 2,181 types of blocks of RV64 instructions for downstream stages, out of which 307 are feasible, applied to all possible input instructions (roughly, $2^{64}$. Symbolic execution [17, 44], which explores individual paths separately, is thus not a good fit for this verification pipeline.

Jitterbug instead adopts Rosette’s strategy for symbolic evaluation [89: §4] to merge the program state at each control-flow join, but it forces a split on every possible (concrete) opcode of symbolic instructions. The intuition is that both the JIT and interpreters tend to handle each opcode separately; splitting on the opcode enables opportunities for concrete evaluation. This strategy works well in practice: it avoids path explosion and leads to easier-to-solve constraints.

The second challenge is that Jitterbug interpreters, unlike those in Serval, must be designed to work on both symbolic state and symbolic instructions. Failing to do so both causes state explosion and produces constraints difficult for SMT solving. For example, the interpreters in Serval represent the CPU state using a vector of bitvectors (one bitvector per register), and encode accessing register $r_i$ as indexing into the vector using integer $i$. This is suitable for concrete instructions, where $i$ is concrete and the generated constraints are restricted to the theory of bitvectors. But with symbolic instructions, a symbolic register index $i$ causes symbolic evaluation to produce constraints that also use the theory of (mathematical) integers. Mixing integers and bitvectors is expensive for solving and can lead to verification bottlenecks [39: §3].

We develop interpreters that account for symbolic instructions and thus can work with Jitterbug. For example, we carefully avoid integers in instruction semantics to restrict resulting constraints to the theories of equality with uninterpreted functions and bitvectors, a decidable fragment of first-order logic. Additionally, recall that the BPF JItS for x86 emit raw bytes (§3.2) and thus require a decoder for verification. We implement an x86 decoder that works on symbolic bytes. The development process is guided by using symbolic profiling to identify verification performance bottlenecks [10] and applying symbolic optimization to fine-tune symbolic evaluation [68: §4].

Axiomatization of expensive SMT operations. Both BPF and machine interpreters provide arithmetic instructions for multiplication, division, and remainder. Reasoning about these operations is expensive for SMT solvers [4, 49], and has been a source of timeouts in verification practice [30, 36].

To avoid such expensive reasoning, Jitterbug takes a standard axiomatization approach [50: §3.2] by replacing these bi-vector operations with uninterpreted functions $\text{mul}_n, \text{div}_n,$ and $\text{rem}_n$ ($n = 32, 64$) in instruction semantics. For example, the BPF JIT for RV32 translates BPF’s $\text{DIV32_REG}$ into using RISC-V’s $\text{divu}$; both instructions are encoded using uninterpreted function $\text{div}_{32}$ (with variations for handling division by zero). Proving the correctness of this translation does not require the semantics of division, thereby scaling verification.

This approach is less general than using SMT’s built-in bi-vector operations, because it ignores the semantics of these operations and might reject valid JIT implementations. For example, the JIT may reorder the operands of a multiplication in emitted instructions; for target architectures lacking native instructions for remainder or 64-bit multiplication, the JIT may emit instructions that emulate the behavior. Proving such a JIT correct requires additional properties about the operations.
Jitterbug captures these properties using the following axioms, which are sufficient to verify all the JIts we have studied:

- commutativity of \( \text{mul} \): \( \text{mul}_n(x, y) = \text{mul}_n(y, x) \);
- remainder: \( \text{rem}_n(x, y) = x - \text{mul}_n(\text{div}_n(x, y), y) \);
- commutativity of \( \text{mul} \): \( \text{mul}_n(x, y) = \text{mul}_n(y, x) \); and decomposition of \( \text{mul} \): \( \text{mul}_{64}(x, y) = (\text{mul}_{128}(x_0, y_0) + \text{mul}_{128}(x_1, y_1)) \) \( \text{mul}_{128}(x_0, y_0) \) and \( \text{mul}_{128}(x_1, y_1) \).

Here \( x \) and \( y \) are bitvectors; subscripts “lo” and “hi” denote the lower and upper half bits, respectively; \( \oplus \) denotes bitvector concatenation; and \( \text{mul}_n \) is an auxiliary uninterpreted function for modeling the upper bits of a product. For example, x86’s 32-bit unsigned multiplication instruction stores a 64-bit product in registers edx and eax; the x86 interpreter encodes their values using \( \text{mul}_1 \) and \( \text{mult} \).

These axioms are shared by the verification of the BPF JIts across architectures. As a sanity check, we formalize and manually prove them using Lean [22].

### 6 Implementing a JIT

**DSL.** Figure 9 shows an excerpt of the BPF JIT for RV32, written in the Jitterbug DSL. The DSL is implemented as a Rosette library and reflects a structured subset of C: booleans, (machine) integers, array accesses (“\( @ \)”), as well as conditional and switch statements. This subset is minimal and sufficient to support the development of the BPF JIT for RV32. It does not support address-of, dereference, or unstructured constructs (e.g., goto or fallback in switch).

Jitterbug extracts the final C code from JIT fragments written in the DSL, a code template (including glue code not covered by the DSL), and a type mapping (not shown here; both array accesses to bpf2rv32 and calls to bpf_get_reg64 return a value of type “\( \text{const s8 * } \)”). Jitterbug does not perform any type checking for the C code.

Using the DSL simplifies verification by avoiding the need to model the C semantics. One can also “escape” from the DSL to use the full Rosette language, though in that case Jitterbug is unable to perform C code extraction; we leverage this to simplify porting the existing BPF JIts from C to Jitterbug.

**Synthesis.** As an application of Jitterbug’s specification and verification, we use Rosette’s support for program synthesis to optimize the BPF JIT for RV32 [59]. We do so by synthesizing JIT fragments written in (a subset of) the DSL, where each fragment takes as input a BPF instruction with a given opcode (e.g., \( \text{ADD64\_REG} \)) and emits a short sequence of RV32 instructions with equivalent behavior. We use the standard approach of writing program sketches [11, 82] to compactly define a space of JIT fragments for compiling ALU instructions. The synthesizer searches this space for the shortest fragment that satisfies per-instruction correctness (Definition 4), according to the Jitterbug verifier.

/ (func (emit_alu_r64 dst src ctx op) (var [tmp @ bpf2rv32 TMP_REG_1] [tmp2 @ bpf2rv32 TMP_REG_2] [rd @ bpf_get_reg64 dst tmp1 ctx] [rs @ bpf_get_reg64 src tmp2 ctx]) (switch op | {BPF_ADD} (cond [(equal rd rs) (emit (rv_add(hi rd) (hi rd) (hi rs)) ctx) (emit (rv_sltu(RV_REG_T0, lo rd, lo rs)) ctx) (emit (rv_add(lo rd) (hi rd) (hi rs)) ctx) (emit (rv_add(hi rd) (hi rd) (hi rd)) ctx)]) (else (emit (rv_add(lo rd) (hi rd) (lo rs)) ctx) (emit (rv_sltu(RV_REG_T0, lo rd, lo rs)) ctx) (emit (rv_add(lo rd) (hi rd) (hi rd)) ctx) (emit (rv_add(hi rd) (hi rd) (hi rd)) ctx) (emit (rv_add(lo rd) (hi rd) (hi rd)) ctx) (emit (rv_add(hi rd) (hi rd) (hi rd)) ctx) (emit (rv_add(lo rd) (hi rd) (hi rd) (RV_REG_T0)) ctx)])])

(a) JIT implementation written in the DSL.

```c
void emit_alu_r64(const s8 *dst, const s8 *src, struct rv_jit_context *ctx, const u8 op) {
  // clang-format off
  if (op) {
    switch (op) {
      case BPF_ADD:
        if (rd == rs) {
          emit(rv_slli(RV_REG_T0, lo rd, 31) ctx);
          emit(rv_slli(hi rd, hi rd, 1) ctx);
          emit(rv_or(hi rd, RV_REG_T0, hi rd) ctx);
          emit(rv_srli(RV_REG_T0, lo rd, 31) ctx);
        } else {
          emit(rv_add(lo rd, lo rd, lo rs) ctx);
          emit(rv_sltu(RV_REG_T0, lo rd, lo rs) ctx);
          emit(rv_add(hi rd, hi rd, hi rs) ctx);
          emit(rv_add(hi rd, hi rd, RV_REG_T0) ctx);
        }
        break;
    }
  ...
}
// clang-format on
}@emit_alu_r64|
```

(b) C code template, where \( \{ \ldots \} \) expands to generated code.

```c
void emit_alu_r64(const s8 *dst, const s8 *src, struct rv_jit_context *ctx, const u8 op) {
  const s8 *tmp1 = bpf2rv32(TMP_REG_1);
  const s8 *tmp2 = bpf2rv32(TMP_REG_2);
  const s8 *rd = bpf_get_reg64(dst, tmp1, ctx);
  const s8 *rs = bpf_get_reg64(src, tmp2, ctx);
  switch (op) {
    case BPF_ADD:
      if (rd == rs) {
        emit(rv_slli(RV_REG_T0, lo rd, 31) ctx);
        emit(rv_slli(hi rd, hi rd, 1) ctx);
        emit(rv_or(hi rd, RV_REG_T0, hi rd) ctx);
        emit(rv_srli(RV_REG_T0, lo rd, 31) ctx);
      }
      ...}
```

(c) Final (extracted) JIT implementation in C.

Figure 9: Excerpt of the BPF JIT for RV32 for compiling the “ADD64\_REG DST, SRC” instruction.

Using this approach, we found two JIT fragments better than our manual implementation for compiling ADD64\_REG (Figure 9) and SUB64\_REG. In each case, the synthesized fragment emitted four instructions, whereas our manual implementation emitted five. We adopted the synthesized fragments in the JIT.

### 7 Experience

Figure 10 shows the code size of the Jitterbug framework and the interpreters for verifying JIts, all written in Rosette. We wrote the interpreters in an idiomatic way [36: §3.3], adding instructions as needed. We borrowed part of the BPF and RV64 semantics from Serval [68], but rewrote the interpreters to support symbolic instructions (§5); we wrote the others from scratch. We developed the metatheory for JIT correctness and the bitvector axiomatization using 1,492 lines of Lean code.
The primary application of Jitterbug is a new BPF JIT for RV32, which we wrote in the DSL, proved against the stepwise specification, and extracted to a C implementation. To validate the generality of Jitterbug, we ported the existing BPF JITS for RV64, Arm32, Arm64, x86-32, and x86-64 in the Linux kernel to Jitterbug for verification. Each port was line-by-line transcription from C to the DSL (and Rosette), emitting the same instructions as the original JIT. These ports did not cover the support for legacy instruction sets (e.g., those lacking atomic instructions mentioned in §4.3), compiling TAIL_CALL, or optimizing register saving.

**Figure 10** lists the line counts for Jitterbug’s components.

**Figure 11** lists the line counts and per-opcode verification time (in seconds) of the BPF JITs for six architectures.

The BPF JIT for RV32. We chose to implement a BPF JIT for RV32 because there was not one in the Linux kernel. The development took five iterations of code reviews.

The first two iterations occurred in June 2019. We sent an initial implementation to kernel developers to gather feedback and gauge interest. The implementation was written in Rosette and manually translated to C, and was unverified. The feedback was positive, with suggestions to add support for eliminating zero extensions [93], an optimization BPF had just introduced.

We submitted the third implementation in February 2020. It was switched to using the DSL (§6), which was less prone to errors in manual translation. It passed the BPF selftests suite, and was verified against an early version of the specification. One of the suggestions from kernel developers was to factor out the common code to be shared among the JITS, such as the per-instruction structure (§4.2). We addressed the suggestions in the next two iterations, after which the JIT was accepted into the Linux kernel (see §A.1.1).

Throughout this process, we refined both the specification and the implementation. The early version of the specification missed two bugs that were also missed by testing. The first bug was an off-by-one error for TAIL_CALL: the emitted instructions limited the TCC (tail-call counter) to 32, rather than the correct value 33. The second bug was that the JIT did not maintain 16-byte alignment of the stack as per the calling convention. We found the two bugs once we completed the specification.

Automated verification supported this development process in two ways. First, it minimized the proof burden for developing the JIT, which must be feature-complete for deployability. Second, it enabled us to catch up with new features being introduced (e.g., support for eliminating zero extensions) and address code reviews by kernel developers in a timely manner.

**Code review.** As listed in §A.1.2, we found 16 new bugs in the existing BPF JITS, wrote patches that fix these bugs, and verified the fixed code. In addition, we found two new bugs in the Arm64 instruction encoding library, a core component shared by BPF and other kernel subsystems (e.g., KVM). We wrote new test cases to be included in the BPF selftests suite. This is useful for catching similar bugs across the JITS, as various “bots” are running selftests continuously for the Linux kernel. Finding subtle bugs in well-tested code shows the effectiveness of the specification and verification.

The main effort for porting and verifying these JITS was in writing the target interpreters for Jitterbug. Verifying the JITS for Arm32 and Arm64 took one week each. Verifying the JITS for x86-32 and x86-64 took three weeks in total, due to the complexity of the x86 interpreter (e.g., instruction decoding). Translating C code to Rosette was mechanical and straightforward, though mistakes in manual translation might hide bugs; extending Jitterbug to work on C code is future work. For specification, we adopted the assumptions and invariants for the JIT for RV32 and adjusted them accordingly.

In our experience, automated verification is key to rapid code review using formal methods. As an example, in December 2019, the developer of the BPF JIT for RV64 submitted patches to add support for far jumps. We ported the patches to Jitterbug and verified their correctness within days of the patch submission. We reported the verification results to kernel developers; the patches were accepted with our review.

<table>
<thead>
<tr>
<th>Component (in Rosette)</th>
<th>Lines of code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jitterbug framework</td>
<td>1,825</td>
</tr>
<tr>
<td>BPF interpreter</td>
<td>471</td>
</tr>
<tr>
<td>Arm32 interpreter</td>
<td>1,265</td>
</tr>
<tr>
<td>Arm64 interpreter</td>
<td>1,166</td>
</tr>
<tr>
<td>RISC-V interpreter (32- and 64-bit)</td>
<td>1,571</td>
</tr>
<tr>
<td>x86 interpreter (32- and 64-bit)</td>
<td>2,299</td>
</tr>
</tbody>
</table>

**Figure 10:** Line counts of Jitterbug’s components.

<table>
<thead>
<tr>
<th>JIT impl.</th>
<th>Spec.</th>
<th>Per-opcode verification time</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>DSL</td>
<td>Min</td>
</tr>
<tr>
<td>RV32</td>
<td>1,964</td>
<td>1,420</td>
</tr>
<tr>
<td>RV64</td>
<td>1,862</td>
<td>1,225</td>
</tr>
<tr>
<td>Arm32</td>
<td>1,620</td>
<td>839</td>
</tr>
<tr>
<td>Arm64</td>
<td>1,025</td>
<td>653</td>
</tr>
<tr>
<td>x86-32</td>
<td>1,683</td>
<td>1,074</td>
</tr>
<tr>
<td>x86-64</td>
<td>1,382</td>
<td>644</td>
</tr>
</tbody>
</table>

**Figure 11:** Line counts and per-opcode verification time (in seconds) of the BPF JITs for six architectures.
Optimization. Another advantage of verification is that it enables developing complex optimizations by providing a high degree of confidence in their correctness. As listed in §A.1.3, we developed 12 patches optimizing the existing BPF JITs. Like code review, we verified the correctness of these optimizations by manually translating the C code to Rosette.

One of the optimizations adds support for RISC-V compressed instruction-set extension (RVC) to the BPF JIT for RV64. RVC improves code density and reduces instruction cache misses by adding short 2-byte instructions for common operations [95: §5], but it poses a challenge to verification: the JIT may choose either base (4-byte) or RVC (2-byte) for emitting each instruction, depending on the immediate value or registers. This leads to an exponential increase in the number of paths in the JIT, emitted instructions, and machine state (e.g., variable code lengths causing the program counter to take different values). Developing and verifying this optimization took approximately 3 weeks, following the proof strategy described in §5 to scale verification.

Beyond BPF JITs. While Jitterbug focuses on the BPF JITs, we also applied it to a JIT for a stack machine to x86-32. We ported the “version 1” JIT described by Myreen [65] to the Jitterbug DSL and extracted it to C code; the port emitted the same x86-32 instructions and was able to run the example as in the paper (Jitterbug does not support the “version 2” JIT that emits self-modifying code). For specification, we excluded registers from the invariants, since the stack machine had no registers; and modeled memory as a map from addresses to values without using traces, since the stack machine had no shared memory (§4.3). For verification, we wrote an interpreter for the stack machine and reused the x86 interpreter provided by Jitterbug. This process took one day.

Jitterbug reported two bugs in the JIT implementation: the offsets for two conditional jump instructions are given as 5 in the original paper, but we concluded that the correct value should be 8. We fixed the offsets and verification succeeded. We believe that both are typos in the paper, as our fixed JIT is consistent with the paper’s HOL4 code and proof.

8 Reflection and conclusion

Our work on Jitterbug was inspired by an earlier effort, started in 2015, to use the Coq theorem prover to develop a verified BPF JIT for x86-64. We chose to implement the JIT itself in x86-64 so as to minimize the trusted computing base. In hindsight, this was a mistake: doing so required reasoning about low-level machine state for both the compiler and emitted code, which hindered the completion of the proof; and the JIT implementation was impractical to audit and deploy due to the lack of C code and the optimizations seen in the Linux kernel. We suspended this effort two years later, in 2017.

Our interest in BPF JITs was revived with the development of symbolic profiling [10] and optimization [68, 74], which together demonstrated a systematic approach for scaling automated verification of low-level code. As an experiment, we wrote a bug finder for BPF JITs in Serval, which checked for strict equivalence of registers (§4.3) over straight-line instructions (§5). It enabled us to find 15 bugs regarding ALU instructions in two BPF JITs, although it was insufficient for verification or finding the bugs described in §3.2 due to the lack of a correctness specification and proof strategy (e.g., support for symbolic instructions).

For Jitterbug, we spent most of our effort devising a specification of JIT correctness that is general enough to cover a broad range of in-kernel JITs (e.g., without requiring padding emitted instructions), expressive enough to catch real bugs, and amenable to automated verification. We found the use of the Lean theorem prover valuable for navigating this trade-off, developing several iterations of the stepwise specification and a proof that implies the correctness of compiling entire programs. It also improved our confidence in the axiomatization of bitvector operations.

A key lesson from Jitterbug is that deciding what not to verify is as important as deciding what to verify. For instance, while ideally we would write and verify the BPF JIT for RV32 in C directly, the use of the DSL enabled us to fine-tune symbolic evaluation, which was critical for scaling verification. If we could not scale verification to JITs written in the DSL, verifying JITs written in C would surely be out of reach. Inspired by seL4 [78] and Ironclad [36], we bridged the resulting gap through validation, separately verifying that the instruction encoding functions in C emitted the same bytes as their original DSL code.

Through this paper, we presented our experience with specifying and verifying BPF JITs, a critical and rapidly evolving component in the Linux kernel. Our experience demonstrates, for the first time, the feasibility of extending automated verification to a restricted but practically important class of compilers. The source code of Jitterbug and the JITs is publicly available at https://github.com/uw-unsat/jitterbug.

Acknowledgments

We thank Anish Athalye, Tej Chajed, Gernot Heiser, Zhihao Jia, Noah Moroze, Jared Roesch, Zach Tatlock, Nickolai Zeldovich, the anonymous reviewers, and our shepherd, Bryan Parno, for feedback that improved this paper. We also thank Andrew Waterman and Claire Wolf for discussion on RISC-V, and H. Peter Anvin, Daniel Borkmann, Will Deacon, Brian Gerst, Song Liu, Andy Shevchenko, Alexei Starovoitov, Björn Töpel, Jiong Wang, Yanqing Wang, and Marc Zynigier for reviewing our patches to the Linux kernel. This work was supported by NSF awards CCF-1651225, CCF-1836724, and CNS-1844807, and by a gift from the VMware University Research Fund.
References


43. Jakub Kicinski and Nicolaas Viljoen. eBPF hardware offload to SmartNICs: cls_bpf and XDP. In the 3rd Technical Conference on Linux Networking, Tokyo, Japan, October 2016.


Conference on Theorem Proving in Higher Order Logics (TPHOLS), pages 28–32, Montreal, Canada, August 2008.


## Artifact appendix

### Patches to the Linux kernel developed using Jitterbug

The following tables list the upstreamed patches to the Linux kernel that we have developed using Jitterbug.

#### Development of the BPF JIT for RV32

<table>
<thead>
<tr>
<th>Commit</th>
<th>Architecture</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5f316b5e699f</td>
<td>RV32</td>
<td>Add RV32G eBPF JIT</td>
</tr>
<tr>
<td>ca6cb5447ec</td>
<td>RV32</td>
<td>Factor common RISC-V JIT code</td>
</tr>
<tr>
<td>745abfaa9af</td>
<td>RV32</td>
<td>Fix tail call count off by one in RV32 BPF JIT</td>
</tr>
<tr>
<td>91f658587a96</td>
<td>RV32</td>
<td>Fix stack layout of JITed code on RV32</td>
</tr>
</tbody>
</table>

#### Bug fixes and new test cases

<table>
<thead>
<tr>
<th>Commit</th>
<th>Architecture</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5f316b5e699f</td>
<td>RV32</td>
<td>Add RV32G eBPF JIT</td>
</tr>
<tr>
<td>ca6cb5447ec</td>
<td>RV32</td>
<td>Factor common RISC-V JIT code</td>
</tr>
<tr>
<td>745abfaa9af</td>
<td>RV32</td>
<td>Fix tail call count off by one in RV32 BPF JIT</td>
</tr>
<tr>
<td>91f658587a96</td>
<td>RV32</td>
<td>Fix stack layout of JITed code on RV32</td>
</tr>
</tbody>
</table>

#### Optimizations for existing BPF JITs

<table>
<thead>
<tr>
<th>Commit</th>
<th>Architecture</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>bb9562cf5c67</td>
<td>Arm32</td>
<td>Fix bugs with ALU64 RSH, ARSH BPF_K shift by 0</td>
</tr>
<tr>
<td>4178417cc535</td>
<td>Arm32</td>
<td>Fix offset overflow for BPF_MEM BPF_DW</td>
</tr>
<tr>
<td>745abfaa9af</td>
<td>Arm64</td>
<td>Fix two bugs in encoding 32-bit logical immediates</td>
</tr>
<tr>
<td>6fa632e719ee</td>
<td>x86-32</td>
<td>Fix bug with ALU64 LSH, RSH, ARSH BPF_K shift by 0</td>
</tr>
<tr>
<td>68a8357ec15b</td>
<td>x86-32</td>
<td>Fix bug with ALU64 LSH, RSH, ARSH BPF_X shift by 0</td>
</tr>
<tr>
<td>80f1f853635</td>
<td>x86-32</td>
<td>Fix bug with JMP32 JSET BPF_X checking upper bits</td>
</tr>
<tr>
<td>5fa9a98fb103</td>
<td>x86-32</td>
<td>Fix incorrect encoding in BPF_LDX zero-extension</td>
</tr>
<tr>
<td>504e7eb6475</td>
<td>x86-32</td>
<td>Fix clobbering of dst for BPF_JSET</td>
</tr>
<tr>
<td>aee194b14d2</td>
<td>x86-32</td>
<td>Add test for lower 8-bit registers in BPF_STX BPF_B</td>
</tr>
<tr>
<td>d2b6c3ab70db</td>
<td></td>
<td>Add test for JMP32 JSET BPF_X with upper bits set</td>
</tr>
<tr>
<td>93e5fb18eccc</td>
<td></td>
<td>Add tests for shifts by zero</td>
</tr>
<tr>
<td>ac8786c72eba</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

60 14th USENIX Symposium on Operating Systems Design and Implementation USENIX Association
A.2 Bug-fixing commits in BPF JITs in the Linux kernel (May 2014–April 2020)

The following table lists bug-fixing commits in the BPF JITs in the Linux kernel for Arm32, Arm64, RV64, x86-32, and x86-64. The superscripts \(J\) and \(S\) mark those for fixing bugs found using Jitterbug and the BPF bug finder in Serval, respectively.

<table>
<thead>
<tr>
<th>Commit</th>
<th>Architecture</th>
<th>Year</th>
<th>Description</th>
</tr>
</thead>
</table>
| ALU:
  bb9562cf5c67\(^J\) | Arm32 | 2020 | Fix bugs with alu64 rsh, arsh bpf_k shift by 0 |
  14e589ff4aa3 | Arm64 | 2015 | Fix mod-by-zero case |
  251599e1d690 | Arm64 | 2015 | Fix div-by-zero case |
  d6393bcb30c | Arm64 | 2015 | Fix endianness conversion bugs |
  1e4df6bf7281 | Arm64 | 2015 | Fix signedness bug in loading VTMbit immediate |
  1e692f09e591\(^S\) | RV64 | 2019 | Clear high 32 bits for alu32 add/sub/neg/lsh/rsh/arsh |
  fe121ee531d1 | RV64 | 2019 | Clear target register high 32-bits for and/or/xor on alu32 |
  1e692f09e591\(^S\) | RV64 | 2019 | Fix bug with alu64 lsh, rsh, arsh bpf_k shift by 0 |
  68a8357ec15b | RV64 | 2019 | Fix bug with alu64 lsh, rsh, arsh bpf_x shift by 0 |
  b9aa0b35d878 | RV64 | 2019 | Fix bug for bpf_alu64 | bpf_neg |
  343f8453b59 | RV64 | 2019 | Fix from_be16 and from_le16/32 instructions |
  2b589a7e2bd3 | Arm32 | 2018 | Correct check_imm24 |
  ddc65a44bb4b | Arm64 | 2017 | Fix jit branch offset related to ldimm64 |
  8ee539ddea0 | Arm64 | 2015 | Fix out-of-bounds read in bpf2a64_offset() |
  59e7eb6475\(^J\) | x86-32 | 2020 | Fix clobbering of dst for bpf_jset |
  80f1f850365\(^S\) | x86-32 | 2020 | Fix bug with jmp32 jset bpf_x checking upper bits |
  711ae1fbbf88 | x86-32 | 2019 | Fix bug for bpf_jmp | bpf_jsgt, bpf_jsle, bpf_jslt, bpf_jsge |
  7c2e988f400e | x86-64 | 2019 | Fix x64 jit code generation for jmp to 1st insn |
  4178417cc535\(^J\) | Arm32 | 2020 | Fix offset overflow for bpf_mem bpf_dw |
  ec19e02b343d | Arm32 | 2018 | Fix ldx instructions |
  8233a67b2a2a | Arm64 | 2019 | Remove prefetch insn in vadd mapping |
  7605cde1b9d | Arm64 | 2017 | Use separate register for state in stxr |
  5aca6f1f86 | x86-32 | 2020 | Fix logic error in bpf_ldx zero-extension |
  5f3a963b10f3\(^J\) | x86-32 | 2020 | Fix incorrect encoding in bpf_ldx zero-extension |
  aee194b14dd2\(^J\) | x86-64 | 2020 | Fix encoding for lower 8-bit registers in bpf_stx bpf_b |
  8c11ea5ce13d | Arm64 | 2018 | Fix getting subprog addr from aux for calls |
  489553dd13a8\(^J\) | RV64 | 2020 | Fix offset range checking for auipc+jalr on rv64 |
  02088d9b392f | Arm32 | 2018 | Fix register saving |
  f448f32cc1fd | Arm32 | 2018 | Fix tail call jumps |
  51c92fbb1b146 | Arm64 | 2014 | Lift restriction on last instruction |
  16338a9b3ac3 | Arm64 | 2018 | Fix out of bounds access in tail call |
  a2284d912bf3 | Arm64 | 2018 | Fix stack_depth tracking in combination with tail calls |
  d8b54110ee94 | Arm64 | 2017 | Fix faulty emission of map access in tail calls |
  96bc4432f5ad | RV64 | 2019 | Limit to 33 tail calls |
  7630e6d475e | x86-64 | 2014 | Fix epilogue generation for ebpf programs |
  90ac4cdd8ec8 | x86-64 | 2017 | Fix bpf_tail_call() x64 jit |
  24823abb93ebf | x86-64 | 2015 | Fix general protection fault when tail call is invoked |
  Prologue and epilogue:
  d122be5d2348 | Arm32 | 2018 | Fix stack alignment |
  f1003b787c0 | RV64 | 2019 | Fix broken bpf tail calls |
  9e4e5b5c8666 | x86-32 | 2018 | Fix regression caused by commit 24dea04767e6 |
  fe8d9571dc50 | x86-64 | 2019 | Fix stack layout of jited bpf code |