A NICE Way to Test OpenFlow Applications
EPFL Technical Report EPFL-REPORT-169211

Marco Canini*, Daniele Venzano*, Peter Perešíni*, Dejan Kostić*, and Jennifer Rexford†

*EPFL †Princeton University

Abstract

The emergence of OpenFlow-capable switches enables exciting new network functionality, at the risk of programming errors that make communication less reliable. The centralized programming model, where a single controller program manages the network, seems to reduce the likelihood of bugs. However, the system is inherently distributed and asynchronous, with events happening at different switches and end hosts, and inevitable delays affecting communication with the controller. In this paper, we present efficient, systematic techniques for testing unmodified controller programs. Our NICE tool applies model checking to explore the state space of the entire system—the controller, the switches, and the hosts. Scalability is the main challenge, given the diversity of data packets, the large system state, and the many possible event orderings. To address this, we propose a novel way to augment model checking with symbolic execution of event handlers (to identify representative packets that exercise code paths on the controller). We also present a simplified OpenFlow switch model (to reduce the state space), and effective strategies for generating event interleavings likely to uncover bugs. Our prototype tests Python applications on the popular NOX platform. In testing three real applications—a MAC-learning switch, in-network server load balancing, and energy-efficient traffic engineering—we uncover eleven bugs.

1 Introduction

While lowering the barrier for introducing new functionality into the network, Software Defined Networking (SDN) also raises the risks of software faults (or *bugs*). Even today's networking software—written and extensively tested by equipment vendors, and constrained (at least somewhat) by the protocol standardization process—can have bugs that trigger Internet-wide outages [1, 2]. In contrast, programmable networks will offer a much wider range of functionality, through software created by a diverse collection of network operators and third-party developers. The ultimate success of SDN, and enabling technologies like OpenFlow [3], depends on having effective ways to test applications in pursuit of achieving high reliability. In this paper, we present NICE, a tool that efficiently uncovers bugs in OpenFlow programs, through a combination of model checking and symbolic execution. Building on our position paper [4] that argues for automating the testing of OpenFlow applications, we introduce several new contributions summarized in Section 1.3.

1.1 Bugs in OpenFlow Applications

An OpenFlow network consists of a distributed collection of switches managed by a program running on a logically-centralized controller, as illustrated in Figure 1. Each switch has a flow table that stores a list of rules for processing packets. Each rule consists of a pattern (matching on packet header fields) and actions (such as forwarding, dropping, flooding, or modifying the packets, or sending them to the controller). A pattern can require an "exact match" on all relevant header fields (*i.e.*, a *microflow* rule), or have "don't care" bits in some fields (*i.e.*, a *wildcard* rule). For each rule, the switch maintains traffic counters that measure the bytes and packets processed so far. When a packet arrives, a switch selects the highest-priority matching rule, updates the counters, and performs the specified action(s). If no rule matches, the switch sends the packet header to the controller and awaits a response on what actions to take. Switches also send event messages, such as a "join" upon joining the network, or "port change" when links go up or down.

The OpenFlow controller (un)installs rules in the switches, reads traffic statistics, and responds to events. For each event, the controller program defines a handler, which may install rules or issue requests for traffic statistics. Many OpenFlow applications[1] are writ-

[1] In this paper, we use the terms "OpenFlow application" and "con-

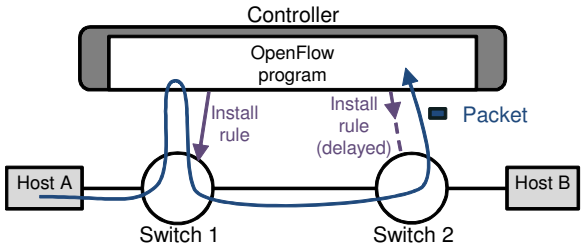


Figure 1: An example of OpenFlow network traversed by a packet. In a plausible scenario, due to delays between controller and switches, the packet does not encounter an installed rule in the second switch.

ten on the NOX controller platform [5], which offers an OpenFlow API for Python and C++ applications. These general-purpose programs can perform arbitrary computation and maintain arbitrary state. A growing collection of controller applications support new network functionality [6–11], over OpenFlow switches available from several different vendors. Our goal is to create an efficient tool for systematically testing these applications.

On the surface, the centralized programming model should reduce the likelihood of bugs. Yet, the system is inherently distributed and asynchronous, with events happening at multiple switches and inevitable delays affecting communication with the controller. To reduce overhead and delay, applications push as much packet-handling functionality to the switches as possible. A common programming idiom is to respond to a packet arrival by installing a rule for handling subsequent packets in the data plane. Yet, a race condition can arise if additional packets arrive while installing the rule. A program that implicitly expects to see just one packet may behave incorrectly when multiple arrive [4]. In addition, many applications install rules at multiple switches along a path. Since rules are not installed atomically, some switches may apply new rules before others install theirs. Figure 1 shows an example where a packet reaches an intermediate switch before the relevant rule is installed. This can lead to unexpected behavior, where an intermediate switch directs a packet to the controller. As a result, an OpenFlow application that works correctly most of the time can misbehave under certain event orderings.

1.2 Challenges of Testing OpenFlow Apps

Testing OpenFlow applications is challenging because the behavior of a program depends on the larger environment. The end-host applications sending and receiving traffic—and the switches handling packets, installing rules, and generating events—all affect the program running on the controller. The need to consider the larger environment leads to an extremely large state space, which "explodes" along three dimensions:

troller program" interchangeably.

Large space of switch state: Switches run their own programs that maintain state, including the many packet-processing rules and associated counters and timers. Further, the set of packets that match a rule depends on the presence or absence of other rules, due to the "match the highest-priority rule" semantics. As such, testing OpenFlow applications requires an effective way to capture the large state space of the switch.

Large space of input packets: Applications are *data-plane* driven, *i.e.*, programs must react to a huge space of possible packets. The OpenFlow specification allows switches to match on source and destination MAC addresses, IP addresses, and TCP/UDP port numbers, as well as the switch input port; future generations of switches will match on even more fields. The controller can perform arbitrary processing based on other fields, such as TCP flags or sequence numbers. As such, testing OpenFlow applications requires effective techniques to deal with large space of inputs.

Large space of event orderings: Network events, such as packet arrivals and topology changes, can happen at any switch at any time. Due to communication delays, the controller may not receive events in order, and rules may not be installed in order across multiple switches. Serializing rule installation, while possible, would significantly reduce application performance. As such, testing OpenFlow applications requires efficient strategies to explore a large space of event orderings.

To simplify the problem, we could require programmers to use domain-specific languages that prevent certain classes of bugs. However, the adoption of new languages is difficult in practice. Not surprisingly, most OpenFlow applications are written in general-purpose languages, like Python, Java, and C++. Alternatively, programmers could create abstract models of their applications, and use formal-methods techniques to prove properties about the system. However, these models are time-consuming to create and easily become out-of-sync with the real implementation. In addition, existing model-checking tools like SPIN [12] and Java PathFinder (JPF) [13] cannot be directly applied because they require explicit developer inputs to resolve the data-dependency issues and sophisticated modeling techniques to leverage domain-specific information. They also suffer state-space explosion, as we show in Section 7. Instead, we argue that testing tools should operate directly on *unmodified* OpenFlow applications written in *general-purpose languages*, and leverage *domain-specific knowledge* to improve scalability.

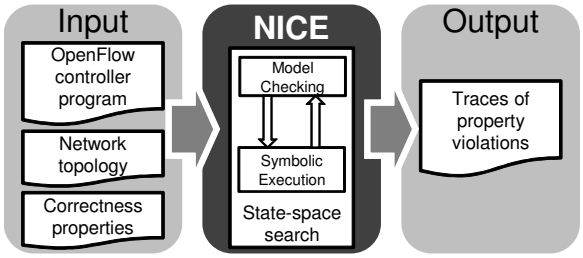


Figure 2: Given an OpenFlow program, a network topology, and correctness properties, NICE performs a state-space search and outputs traces of property violations.

1.3 NICE Research Contributions

To address these scalability challenges, we present NICE (*No bugs In Controller Execution*)—a tool that tests unmodified controller programs by automatically generating carefully-crafted streams of packets under many possible event interleavings. To use NICE, the programmer supplies the controller program, and the specification of a topology with switches and hosts. The programmer can instruct NICE to check for generic correctness properties such as no forwarding loops or no black holes, and optionally write additional, application-specific correctness properties (*i.e.*, Python code snippets that make assertions about the global system state). By default, NICE systematically explores the space of possible system behaviors, and checks them against the desired correctness properties. The programmer can also configure the desired search strategy. In the end, NICE outputs property violations along with the traces to deterministically reproduce them. The programmer can also use NICE as a simulator to perform manually-driven, step-by-step system executions or random walks on system states.

Our design uses explicit state, software model checking [13–16] to explore the state space of the entire system—the controller program, the OpenFlow switches, and the end hosts—as discussed in Section 2. However, applying model checking "out of the box" does not scale. While simplified models of the switches and hosts help, the main challenge is the event handlers in the controller program. These handlers are data dependent, forcing model checking to explore all possible inputs (which doesn't scale) or a set of "important" inputs provided by the developer (which is undesirable). Instead, we extend model checking to *symbolically execute* [17, 18] the handlers, as discussed in Section 3. By symbolically executing the packet-arrival handler, NICE identifies equivalence classes of packets—ranges of header fields that determine unique paths through the code. NICE feeds the network a representative packet from each class by adding a state transition that inject the packet. To reduce the space of event orderings, we propose several domain-specific search strategies that generate event interleavings that are likely to uncover bugs in the controller program, as discussed in Section 4.

Bringing these ideas together, NICE combines model checking (to explore system execution paths), symbolic execution (to reduce the space of inputs), and search strategies (to reduce the space of event orderings). The programmer can specify correctness properties as snippets of Python code that operate on system state, or select from a library of common properties, as discussed in Section 5. Our NICE prototype tests unmodified applications written in Python for the popular NOX platform, as discussed in Section 6. Our performance evaluation in Section 7 shows that: (i) even on small examples, NICE is five times faster than approaches that apply state-of-the-art tools (ii) our OpenFlow-specific search strategies reduce the state space by up to 20 times, and (iii) the simplified switch model brings a 7-fold reduction on its own. In Section 8, we apply NICE to three real OpenFlow applications and uncover *11* bugs. Most of the bugs we found are design flaws, which are inherently less numerous than simple implementation bugs. In addition, at least one of these applications was tested using unit tests. Section 9 discusses the trade-off between testing coverage and the overhead of symbolic execution. Section 10 discusses related work, and Section 11 concludes the paper with a discussion of future research directions.

2 Model Checking OpenFlow Applications

The execution of a controller program depends on the underlying switches and end hosts; the controller, in turn, affects the behavior of these components. As such, testing is not just a simple matter of exercising every path through the controller program—we must consider the state of the larger system. The need to systematically explore the space of system states, and check correctness in each state, naturally leads us to consider *model checking* techniques. To apply model checking, we need to identify the system states and the transitions from one state to another. After a brief review of model checking, we present a strawman approach for applying model checking to OpenFlow applications, and proceed by describing changes that make it more tractable.

2.1 Background on Model Checking

Modeling the state space. A distributed system consists of multiple *components* that communicate asynchronously over message *channels*, *i.e.*, first-in, first-out buffers (*e.g.*, see Chapter 2 of [19]). Each component has a set of variables, and the *component state* is an assignment of values to these variables. The *system state* is the composition of the component states. To capture in-flight

```
1 state_stack = []; explored_states = []; errors = []
2 initial_state = create_initial_state()
3 for t in initial_state.enabled_transitions:
4   state_stack.push([initial_state, t])
5 while len(state_stack) > 0:
6   state, transition = choose(state_stack)
7   try:
8     next_state = run(state, transition)
9     check_properties(next_state)
10    if next_state not in explored_states:
11      explored_states.add(next_state)
12      for t in state.enabled_transitions:
13        state_stack.push([next_state, t])
14  except PropertyViolation as e:
15    errors.append([e, trace])
```

Figure 3: Pseudo-code of the basic model-checking loop.

messages, the system state also includes the contents of the channels. A *transition* represents a change from one state to another (*e.g.*, due to sending a message). At any given state, each component maintains a set of enabled transitions, *i.e.*, the state's possible transitions. For each state, the enabled system transitions are the union of enabled transitions at all components. A *system execution* corresponds to a sequence of these transitions, and thus specifies a possible behavior of the system.

Model-checking process. Given a model of the state space, performing a search is conceptually straightforward. Figure 3 shows the pseudo-code of the model-checking loop. First, the model checker initializes a stack of states with the initial state of the system. At each step, the checker chooses one state from the stack and one of its enabled transitions. After executing that transition, the checker tests the correctness properties on the newly reached state. If the new state violates a correctness property, the checker saves the error and the execution trace. Otherwise, the checker adds the new state to the set of explored states (unless the state was added earlier) and schedules the execution of all transitions enabled in this state (if any). The model checker can run until the stack of states is empty, or until detecting the first error.

2.2 Transition Model for OpenFlow Apps

Model checking relies on having a model of the system, *i.e.*, a description of the state space. This requires us to identify the states and transitions for each component—the controller program, the OpenFlow switches, and the end hosts. However, we argue that applying existing model-checking techniques imposes too much work on the developer and leads to an explosion in the state space.

```
1  ctrl_state = {} # State of the controller is a global variable (a hashtable)
2  def packet_in(sw_id, inport, pkt, bufid): # Handles packet arrivals
3    mactable = ctrl_state[sw_id]
4    is_bcast_src = pkt.src[0] & 1
5    is_bcast_dst = pkt.dst[0] & 1
6    if not is_bcast_src:
7      mactable[pkt.src] = inport
8    if (not is_bcast_dst) and (mactable.has_key(pkt.dst)):
9      outport = mactable[pkt.dst]
10     if outport != inport:
11       match = {DL_SRC: pkt.src, DL_DST: pkt.dst,
             DL_TYPE: pkt.type, IN_PORT: inport}
12       actions = [OUTPUT, outport]
13       install_rule(sw_id, match, actions, soft_timer=5,
             hard_timer=PERMANENT) # 2 lines optionally
14       send_packet_out(sw_id, pkt, bufid) # combined in 1 API
15       return
16   flood_packet(sw_id, pkt, bufid)
17 def switch_join(sw_id, stats): # Handles when a switch joins
18   if not ctrl_state.has_key(sw_id):
19     ctrl_state[sw_id] = {}
20 def switch_leave(sw_id): # Handles when a switch leaves
21   if ctrl_state.has_key(sw_id):
22     del ctrl_state[sw_id]
```

Figure 4: Pseudo-code of a MAC-learning switch, based on the `pyswitch` application. The `packet_in` handler learns the input port associated with each non-broadcast source MAC address; if the destination MAC address is known, the handler installs a forwarding rule and instructs the switch to send the packet according to that rule; and otherwise floods the packet. The switch join/leave events initialize/delete a table mapping addresses to switch ports. Note that another API (not shown) wraps an OpenFlow protocol optimization that combines into a single one the two operations: modifying the flow table and processing the packet that caused the modification.

2.2.1 Controller Program

Modeling the controller as a transition system seems relatively straightforward. A controller program is structured as a set of event handlers (*e.g.*, packet arrival and switch join/leave for the MAC-learning application in Figure 4), that interact with the switches using a standard interface, and these handlers execute atomically. As such, we can model the state of the program as the values of its global variables (*e.g.*, `ctrl_state` in Figure 4), and treat each event handler as a transition. To execute a transition, the model checker can simply invoke the associated event handler. For example, receiving a packet-in message from a switch enables the `packet_in` transition, and the model checker can execute the transition by invoking the corresponding event handler.

However, the behavior of event handlers is often data-dependent. In line 7 of Figure 4, for instance, the `packet_in` handler assigns `mactable` only for unicast source MAC addresses, and either installs a forward-

ing rule or floods a packet depending on whether or not the destination MAC address is known. This leads to different system executions. Unfortunately, model checking does not cope well with data-dependent applications (*e.g.*, see Chapter 1 of [19]). Since enumerating all possible inputs is intractable, a brute-force solution would require developers to specify a set of "relevant" inputs based on their knowledge of the application. Hence, a controller transition would be modeled as a pair consisting of an event handler and a concrete input. This is clearly undesirable. NICE overcomes this limitation by using *symbolic execution* to automatically identify the relevant inputs, as discussed in Section 3.

2.2.2 OpenFlow Switches

To test the controller program, the system model must include the underlying switches. Yet, switches run complex software, and this is not the code we intend to test. A strawman approach for modeling the switch is to start with an existing reference OpenFlow switch implementation (*e.g.*, [20]), define the switch state as the values of all variables, and identify transitions as the portions of the code that process packets or exchange messages with the controller. However, the reference switch software has a large amount of state (*e.g.*, several hundred KB), not including the buffers containing packets and OpenFlow messages awaiting service; this aggravates the state-space explosion problem. Importantly, such a large program has many sources of nondeterminism and it is difficult to identify them automatically [16].

Instead, we create a switch model that omits inessential details. Indeed, creating models of some parts of the system is common to many standard approaches for applying model checking. Further, in our case, this is a one-time effort that does not add burden on the user. Following the OpenFlow specification [21], we view a switch as a set of communication channels, transitions that handle data packets and OpenFlow messages, and a flow table.

Simple communication channels: Each channel is a first-in, first-out buffer. Packet channels have an optionally-enabled fault model that can drop, duplicate, or reorder packets, or fail the link. The channel with the controller offers reliable, in-order delivery of OpenFlow messages, except for optional switch failures. We do not run the OpenFlow protocol over SSL on top of TCP/IP, allowing us to avoid intermediate protocol encoding/decoding and the substantial state in the network stack.

Two simple transitions: The switch model supports `process_pkt` and `process_of` transitions—for processing data packets and OpenFlow messages, respectively. We enable these transitions if at least one packet channel or the OpenFlow channel is non empty, respectively. To match the controller program's expectations about the environment, our switch model includes buffers that temporarily store packets awaiting further instruction from the controller. However, to improve scalability, we do not include these buffers in our definition of the state space. A final simplification we make is in the `process_pkt` transition. Here, the switch dequeues the first packet from each packet channel, and processes *all* these packets according to the flow table. So, multiple packets at different channels are processed as a single transition. This optimization is safe because the model checker already systematically explores the possible orderings of packet arrivals at the switch.

Merging equivalent flow tables: A flow table can easily have two states that appear different but are semantically equivalent, leading to a larger search space than necessary. For example, consider a switch with two microflow rules. These rules do not overlap—no packet would ever match both rules. As such, the order of these two rules is not important. Yet, simply storing the rules as a list would cause the model checker to treat two different orderings of the rules as two distinct states. Instead, as often done in model checking (*e.g.*, [22]), we construct a canonical representation of the flow table that derives a unique order of rules with overlapping patterns.

2.2.3 End Hosts

Modeling the end hosts is tricky, because hosts run arbitrary applications and protocols, have large state, and have behavior that depends on incoming packets. We could require the developer to provide the host programs, with a clear indication of the transitions between states. Instead, NICE provides simple programs that act as clients or servers for a variety of protocols including Ethernet, ARP, IP, and TCP. These models have explicit transitions and relatively little state. For instance, the default client has two basic transitions—`send` (initially enabled; can execute C times, where C is configurable) and `receive`—and a counter of sent packets. The default server has the `receive` and the `send_reply` transitions; the latter is enabled by the former. A more realistic refinement of this model is the mobile host that includes the `move` transition that moves the host to a new <switch, port> location. The programmer can also customize the models we provide, or create new models.

3 Symbolic Execution of Event Handlers

To systematically test the controller program, we must explore all of its possible transitions. Yet, the behavior of an event handler depends on the inputs (*e.g.*, the MAC addresses of packets in Figure 4). Rather than explore all

possible inputs, NICE identifies which inputs would exercise different execution paths through an event handler. Systematically exploring all code paths naturally leads us to consider *symbolic execution* (SE) techniques. After a brief review of symbolic execution, we describe how we apply symbolic execution to controller programs. Then, we explain how NICE combines model checking and symbolic execution to explore the state space effectively.

3.1 Background on Symbolic Execution

Symbolic execution runs a program with symbolic variables as inputs (*i.e.*, any values). The symbolic execution engine tracks the use of symbolic variables and records the constraints on their possible values. For example, the engine does not learn the value of `is_bcast_src` in line 4 of Figure 4, but instead learns that `is_bcast_src` is "`pkt.src[0] & 1`". At any branch, the engine queries a solver for two assignments of symbolic inputs—one that satisfies the branch predicate and one that satisfies its negation (*i.e.*, takes the "else" branch)— and logically forks the execution to follow the feasible paths. For example, the engine determines that to reach line 7 of Figure 4, the source MAC address must have its eighth bit set to zero. The engine then updates the path constraint, *i.e.*, the conjunction of all constraints on symbolic variables that led execution down that path. The values of symbolic variables that are sufficient for execution to take a path derive from that path constraint.

Unfortunately, symbolic execution does not scale well because the number of code paths can grow exponentially with the number of branches and the size of the inputs. Also, symbolic execution does not explicitly model the state space, which can cause repeated exploration of the same system state. In addition, despite exploring all *code paths*, symbolic execution does not explore all *system execution paths*, such as different event interleavings. Techniques exist that can add artificial branching points to a program to inject faults or explore different event orderings [18, 23], but at the expense of extra complexity. As such, symbolic execution is not a sufficient solution for testing OpenFlow applications. Instead, NICE uses model checking to explore system execution paths (and detect repeated visits to the same state [24]), and symbolic execution to determine which inputs would exercise a particular state transition.

3.2 Symbolic Execution of OpenFlow Apps

Applying symbolic execution to the controller event handlers is relatively straightforward, with two exceptions. First, to handle the diverse inputs to the `packet_in` handler, we construct *symbolic packets*. Second, to minimize the size of the state space, we choose a *concrete* (rather than symbolic) representation of controller state.

Symbolic packets. The main input to the `packet_in` handler is the incoming packet. To perform symbolic execution, NICE must identify which (ranges of) packet header fields determine the path through the handler. Rather than view a packet as a generic array of symbolic bytes, we introduce *symbolic packets* as our symbolic data type. A symbolic packet is a group of symbolic integer variables that each represents a header field. To reduce the overhead for the constraint solver, we maintain each header field as an individual symbolic variable (*e.g.*, a MAC address is a 6-byte variable), which reduces the number of variables. Yet, we still allow byte- and bit-level accesses to the fields. We also apply domain knowledge to further constrain the possible values of header fields (*e.g.*, the MAC and IP addresses used by the hosts and switches in the system model, as specified by the input topology). Finally, the fields are lazily-initialized so that we reduce the overhead for the constraint solver by omitting the unused fields. This also tells us whether the program is agnostic to particular protocols (*e.g.*, ignoring transport header fields), allowing us to select a simpler host model for generating the input packets.

Concrete controller state. The execution of the event handlers also depends on the controller state. For example, the code in Figure 4 reaches line 9 only for unicast destination MAC addresses stored in `mactable`. Starting with an empty `mactable`, symbolic execution cannot find an input packet that forces the execution of line 9; yet, with a non-empty table, certain packets could trigger line 9 to run, while others would not. As such, we must incorporate the global variables into the symbolic execution. We choose to represent the global variables in a concrete form. We apply symbolic execution by using these concrete variables as the initial state and by marking as symbolic the packets and statistics arguments to the handlers. The alternative of treating the controller state as symbolic would require a sophisticated type-sensitive analysis of complex data structures (*e.g.*, [24]), which is computationally expensive and difficult for an untyped language like Python. In addition, having purely symbolic controller state could cause NICE to test spurious states that are not reachable in practice due to the constraints imposed by the larger environment.

3.3 Combining SE With Model Checking

With all of NICE's parts in place, we now describe how we combine model checking (to explore system execution paths) and symbolic execution (to reduce the space of inputs). Figure 5 shows the unfolding of controller's state-space graph. At any given controller state, we want to identify the packets that each client should

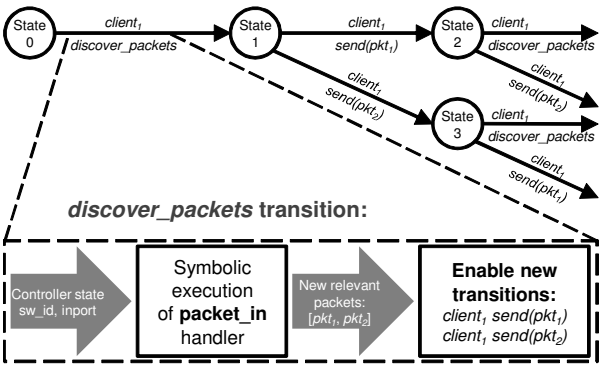


Figure 5: Example of how NICE identifies relevant packets and uses them as new enabled send packet transitions of $client_1$. For clarity, the circled states refer to the controller state only.

send—specifically, the set of packets that exercise all feasible code paths on the controller in that state. To do so, we create a special end-host transition called `discover_packets` that is initially enabled. When executed, this transition invokes the symbolic-execution engine to symbolically execute the `packet_in` handler.

NICE executes the handler symbolically starting from the initial state defined by (i) the concrete controller state (State 0 in Figure 5) and (ii) a concrete "context" (*i.e.*, the switch and input port that identify the host's location). For every feasible code path in the handler, the symbolic-execution engine finds an equivalence class of packets that exercise it. For each equivalence class, we instantiate one *concrete packet* (referred to as the relevant packet) and enable a corresponding `send` transition for the client. While this example focuses on the `packet_in` handler, we apply similar techniques to deal with traffic statistics, by introducing a special `discover_stats` transition that symbolically executes the statistics handler with symbolic integers as arguments. Other handlers, related to topology changes, operate on concrete inputs (*e.g.*, the switch and port ids).

Figure 6 shows the pseudo-code of our search-space algorithm, which extends extends the basic model-checking loop of Figure 3 in two main ways.

Initialization (lines 3-5): For each host (or "client"), the algorithm (i) creates an empty map for storing the relevant packets for a given controller state and (ii) enables the `discover_packets` transition.

Checking process (lines 12-18): Upon reaching a new state, the algorithm checks for each client (line 15) whether a set of relevant packets already exists. If not, it enables the `discover_packets` transition. In addition, it checks (line 17) if the controller has a `process_stat` transition enabled in the newly-reached state, meaning that the controller is awaiting a

```
1  state_stack = []; explored_states = []; errors = []
2  initial_state = create_initial_state()
3  for client in initial_state.clients
4    client.packets = {}
5    client.enable_transition(discover_packets)
6  for t in initial_state.enabled_transitions:
7    state_stack.push([initial_state, t])
8  while len(state_stack) > 0:
9    state, transition = choose(state_stack)
10   try:
11     next_state = run(state, transition)
12     ctrl = next_state.ctrl # Reference to controller in next_state
13     ctrl_state = state(ctrl) # Stringified controller state in next_state
14     for client in state.clients:
15       if not client.packets.has_key(ctrl_state):
16         client.enable_transition(discover_packets, ctrl)
17     if process_stats in ctrl.enabled_transitions:
18       ctrl.enable_transition(discover_stats, state, sw_id)
19     check_properties(next_state)
20     if next_state not in explored_states:
21       explored_states.add(next_state)
22       for t in state.enabled_transitions:
23         state_stack.push([next_state, t])
24   except PropertyViolation as e:
25     errors.append([e, trace])
26 def discover_packets_transition(client, ctrl):
27   sw_id, inport = switch_location_of(client)
28   new_packets = SymbolicExecution(ctrl, packet_in,
         context=[sw_id, inport])
29   client.packets[state(ctrl)] = new_packets
30   for packet in client.packets[state(ctrl)]:
31     client.enable_transition(send, packet)
32 def discover_stats_transition(ctrl, state, sw_id):
33   new_stats = SymbolicExecution(ctrl, process_stats,
         context=[sw_id])
34   for stats in new_stats:
35     ctrl.enable_transition(process_stats, stats)
```

Figure 6: Pseudo-code of the state-space search algorithm used in NICE for finding errors. The highlighted parts, including the special "*discover*" transitions, are our additions to the basic model-checking loop of Figure 3.

response to a previous query for statistics. If so, the algorithm enables the `discover_stats` transition.

Invoking the `discover_packets` (lines 26-31) and `discover_stats` (lines 32-35) transitions allows the system to evolve to a state where new transitions become possible—one for each path in the packet-arrival or statistics handler. This allows the model checker to reach new controller states, allowing symbolic execution to again uncover new classes of inputs that enable additional transitions, and so on.

By symbolically executing the controller event handlers, NICE can automatically infer the test inputs for enabling model checking without developer input, at the expense of some limitations in coverage of the system

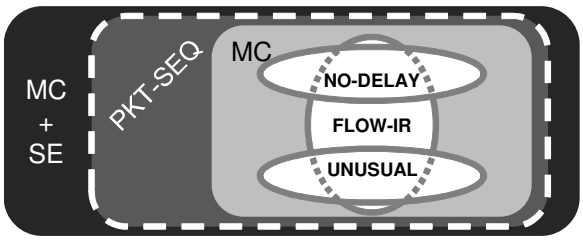


Figure 7: Illustration of the state space explored by our search strategies in relation to the entire state space.

state space which we discuss later in Section 9.

4 OpenFlow-Specific Search Strategies

Even with our optimizations from the last two sections, the model checker cannot typically explore the entire state space, since events can occur in so many different orders. Thus, we propose domain-specific heuristics that substantially reduce the space of event orderings while focusing on scenarios that are likely to uncover bugs. Figure 7 graphically summarizes the state space explored by the proposed search strategies in relation to the entire state space. Most of the strategies operate on the event interleavings produced by model checking, except for PKT-SEQ which reduces the state-space explosion due to the transitions uncovered by symbolic execution.

PKT-SEQ: Relevant packet sequences. The effect of discovering new relevant packets and using them as new enabled `send` transitions is that each end-host generates a potentially-unbounded tree of packet sequences. To make the state space finite and smaller, this heuristic reduces the search space by bounding the possible end host transitions (indirectly, bounding the tree) along two dimensions, each of which can be fine tuned by the user. The first is merely *the maximum length of the sequence*, or in other words, the depth of the tree. Effectively, this places a hard limit to the issue of infinite execution trees due to symbolic execution. The second is the *maximum number of outstanding packets*, or in other words, the length of a packet burst. For example, if $client_1$ in Figure 5 is allowed only a 1-packet burst, this heuristic would disallow both `send(`pkt_2`)` in State 2 and `send(`pkt_1`)` in State 3. Effectively, this limits the level of "packet concurrency" within the state space. To introduce this limit, we conceive each end host has a counter c, and when $c = 0$, the end host cannot send any more packet until the counter is replenished. As we are dealing with multiple communicating end hosts, we find it natural to use as the default behavior increasing c by one unit for every received packet. However, this behavior can be modified in more complex end host models, *e.g.*, to mimic the TCP flow and congestion controls.

NO-DELAY: Instantaneous rule updates. When using this simple heuristic, NICE treats each communication between a switch and the controller as a single atomic action (*i.e.*, not interleaved with any other transitions). In other words, the global system runs in "lock step." This heuristic is useful during the early stages of development to find basic design errors, rather than race conditions or other concurrency-related problems. For instance, this heuristic would allow the developer to realize that installing a rule prevents the controller from seeing other packets that are important for program correctness. For example, a MAC-learning application that installs forwarding rules based only on the destination MAC address would prevent the controller from seeing some packets with new source MAC addresses.

UNUSUAL: Uncommon delays and reorderings. With this heuristic, NICE only explores event orderings with unusual and unexpected delays, with the goal of uncovering race conditions. For example, if an event handler in the controller installs rules in switches 1, 2, and 3, the heuristic explores transitions that reverse the order by allowing switch 3 to install its rule first, followed by switch 2 and then switch 1. This heuristic uncovers bugs like the example in Figure 1.

FLOW-IR: Flow independence reduction. Many OpenFlow applications treat different groups of packets independently; that is, the handling of one group is not affected by the presence or absence of another. In this case, NICE can reduce the search space by exploring only one relative ordering between the events affecting each group. To use this heuristic, the programmer provides `isSameFlow`, a Python function that takes two packets (and the switch and input port) as arguments and returns whether the packets belong to the same group. For example, in some scenarios different microflows are independent, whereas other programs may treat packets with different destination MAC addresses independently.

Summary. PKT-SEQ is complementary to other strategies in that it only reduces the number of `send` transitions rather than the possible kind of event orderings. PKT-SEQ is enabled by default and used in our experiments (unless otherwise noted). The other heuristics can be selectively enabled and arbitrarily combined.

5 Specifying Application Correctness

Correctness is not an intrinsic property of a system—a specification of correctness states what the system should (or should not) do, whereas the implementation determines what it actually does. NICE allows programmers to specify correctness properties as Python code snippets, and provides a library of common properties (*e.g.*, no forwarding loops or blackholes).

5.1 Customizable Correctness Properties

Testing correctness involves asserting safety properties (*"something bad never happens"*) and liveness properties (*"eventually something good happens"*), defined more formally in Chapter 3 of [19]. Checking for safety properties is relatively easy, though sometimes writing an appropriate predicate over all state variables is tedious. As a simple example, a predicate could check that the collection of flow rules does not form a forwarding loop or a black hole. Checking for liveness properties is typically harder because of the need to consider a possibly infinite system execution. In NICE, we make the inputs finite (*e.g.*, a finite number of packets, each with a finite set of possible header values), allowing us to check some liveness properties. For example, NICE could check that, once two hosts exchange at least one packet in each direction, no further packets go to the controller (a property we call "StrictDirectPaths"). Checking this liveness property requires knowledge not only of the system state, but also which transitions have executed.

To check both safety and liveness properties, NICE allows correctness properties to (i) access the system state, (ii) register callbacks invoked by NICE to observe important transitions in system execution, and (iii) maintain local state. In our experience, these features offer enough expressiveness for specifying correctness properties. For ease of implementation, these properties are represented as snippets of Python code that make assertions about global system state. NICE invokes these snippets after each transition. For example, to check the StrictDirectPaths property, the code snippet would have local state variables that keep track of whether a pair of hosts has exchanged at least one packet in each direction, and would flag a violation if a subsequent packet triggers a `packet_in` event at the controller. When a correctness check signals a violation, the tool records the execution trace that recreates the problem.

5.2 Library of Correctness Properties

NICE provides a library of correctness properties applicable to a wide range of OpenFlow applications. A programmer can select properties from a list, as appropriate for the application. Writing these correctness modules can be challenging because the definitions must be robust to communication delays between the switches and the controller. Many of the definitions must intentionally wait until a "safe" time to test the property to prevent natural delays from erroneously triggering a violation of the property. Providing these modules as part NICE can relieve the developers from the challenges of specifying correctness properties precisely, though creating any custom modules would require similar care.

• *NoForwardingLoops*: This property checks that each packet goes through any given <switch, input port> pair at most once.

• *NoBlackHoles*: This property states that no packets should be dropped in the network, and is implemented by checking that every packet that enters the network ultimately leaves the network or is consumed by the controller itself (for simplicity, we disable optional packet drops and duplication on the channels). To account for flooding, the property enforces a zero balance between the packet copies and packets consumed.

• *DirectPaths*: This property checks that, once a packet has successfully reached its destination, future packets of the same flow do not go to the controller. Effectively, this checks that the controller successfully establishes a direct path to the destination as part of handling the first packet of a flow. This property is useful for many OpenFlow applications, though it does not apply to the MAC-learning switch, which requires the controller to learn how to reach both hosts before it can construct unicast forwarding paths in either direction.

• *StrictDirectPaths*: This property checks that, after two hosts have successfully delivered at least one packet of a flow in each direction, no successive packets reach the controller. This checks that the controller has established a direct path in *both* directions between the two hosts.

• *NoForgottenPackets*: This property checks that all switch buffers are empty at the end of system execution. A program can easily violate this property by forgetting to tell the switch how to handle a packet. This can eventually consume all the available buffer space for packets awaiting controller instruction; after a timeout, the switch may discard these buffered packets. [2]. A short-running program may not run long enough for the queue of awaiting-controller-response packets to fill, but the *NoForgottenPackets* property easily detects these bugs.

6 Implementation Highlights

We have built a prototype implementation of NICE written in Python so as to seamlessly support OpenFlow controller programs for the popular NOX controller platform (which provides an API for Python).

As a result of using Python, we face the challenge of doing symbolic execution for a dynamic, untyped language. This task turned out to be quite challenging from an implementation perspective. To avoid modifying the Python interpreter, we implement a derivative technique

[2] In our tests of the ProCurve 5406zl OpenFlow switch, we see that, once the buffer becomes full, the switch starts sending the *entire contents* of new incoming packets to the controller, rather than buffering them. After a ten-second timeout, the switch deletes the packets that are buffered awaiting instructions from the controller.

of symbolic execution called *concolic execution* [25][3], which executes the code with concrete instead of symbolic inputs. Alike symbolic execution, it collects constraints along code paths and tries to explore all feasible paths. Another consequence of using Python is that we incur a significant performance overhead, which is the price for favoring usability. We plan to improve performance in a future release of the tool.

NICE consists of three parts: (i) a model checker, (ii) a concolic-execution engine, and (iii) a collection of models including the simplified switch and several end hosts. We now briefly highlight some of the implementation details of the first two parts: the model checker and concolic engine, which run as different processes.

Model checker details. To checkpoint and restore system state, NICE takes the approach of remembering the sequence of transitions that created the state and restores it by replaying such sequence, while leveraging the fact that the system components execute deterministically. State-matching is doing by comparing and storing hashes of the explored states. The main benefit of this approach is that it reduces memory consumption and, secondarily, it is simpler to implement. Trading computation for memory is a common approach for other model-checking tools (*e.g.*, [15, 16]). To create state hashes, NICE serializes the state via the cPickle module and applies the built-in hash function to the resulting string. We surmise it would be possible to reduce NICE running time by storing the serialized state itself, at the cost of higher memory usage.

Concolic execution details. A key step in concolic execution is tracking the constraints on symbolic variables during code execution. To achieve this, we first implement a new "symbolic integer" data type that tracks assignments, changes and comparisons to its value while behaving like a normal integer from the program point of view. We also implement arrays (tuples in Python terminology) of these symbolic integers. Second, we reuse the Python modules that naturally serve for debugging and disassembling the byte-code to trace the program execution through the Python interpreter.

Further, before running the code symbolically, we normalize and instrument it since, in Python, the execution can be traced at best with single code-line granularity. Specifically, we convert the source code into its abstract syntax tree (AST) representation and then manipulate this tree through several recursive passes that perform the following transformations: (i) we split composite branch predicates into nested if statements to work around shortcut evaluation, (ii) we move function calls before conditional expressions to ease the job for the STP constraint solver [26], (iii) we instrument branches to inform the concolic engine on which branch is taken, (iv) we substitute the built-in dictionary with a special stub that exposes the constraints, and (v) we intercept and remove sources of nondeterminism (*e.g.*, seeding the pseudo-random number generator). The AST tree is then converted back to source code for execution.

[3]Concolic stands for concrete + symbolic.

7 Performance Evaluation

Here we present an evaluation of how effectively NICE copes with the large state space in OpenFlow.

Experimental setup. We run the experiments on the simple topology of Figure 1, where the end hosts behave as follows: host A sends a "*layer-2 ping*" packet to host B which replies with a packet to A. The controller runs the MAC-learning switch program of Figure 4. We report the numbers of transitions and unique states, and the execution time as we increase the number of concurrent pings (a pair of packets). We run all our experiments on a machine set up with Linux 2.6.32 x86_64 that has 64 GB of RAM and a clock speed of 2.6 GHz. Our prototype implementation does not yet make use of multiple cores.

Benefits of simplified switch model. We first perform a full search of the state space using NICE as a depth-first search model checker (NICE-MC, without symbolic execution) and compare to NO-SWITCH-REDUCTION: doing model-checking without a canonical representation of the switch state. Effectively, this prevents the model checker from recognizing that it is exploring semantically equivalent states. These results, shown in Table 1, are obtained without using any of our search strategies. We compute ρ, a metric of state-space reduction due to using the simplified switch model, as $\frac{Unique(\text{NO-SWITCH-REDUCTION}) - Unique(\text{NICE-MC})}{Unique(\text{NO-SWITCH-REDUCTION})}$.

We observe the following:

• In both samples, the number of transitions and of unique states grow roughly exponentially (as expected). However, using the simplified switch model, the unique states explored in NICE-MC only grow with a rate that is about half the one observed for NO-SWITCH-REDUCTION.

• The efficiency in state-space reduction ρ scales with the problem size (number of pings), and is substantial (factor of seven for three pings).

Heuristic-based search strategies. Figure 8 illustrates the contribution of NO-DELAY and FLOW-IR in reducing the search space relative to the metrics reported for the full search (NICE-MC). We omit the results for UNUSUAL as they are similar. The state space reduction is again significant; about factor of four for three pings. In summary, our switch model and these heuristics result in a 28-fold state space reduction for three pings.

Comparison to other model checkers. Next, we con-

Pings	NICE-MC Transitions	NICE-MC Unique states	NICE-MC CPU time	NO-SWITCH-REDUCTION Transitions	NO-SWITCH-REDUCTION Unique states	NO-SWITCH-REDUCTION CPU time	ρ
2	470	268	0.94 [s]	760	474	1.93 [s]	0.38
3	12,801	5,257	47.27 [s]	43,992	20,469	208.63 [s]	0.71
4	391,091	131,515	36 [m]	2,589,478	979,105	318 [m]	0.84
5	14,052,853	4,161,335	30 [h]	-	-	-	-

Table 1: Dimensions of exhaustive search in NICE-MC vs. model-checking without a canonical representation of the switch state, which prevents recognizing equivalent states. Symbolic execution is turned off in both cases. NO-SWITCH-REDUCTION did not finish with five pings in four days.

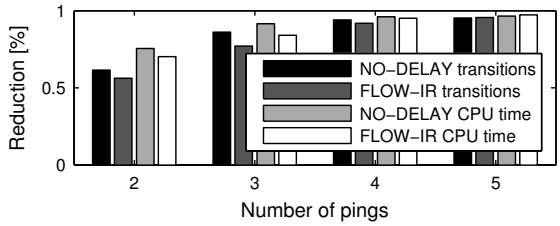


Figure 8: Relative state-space search reduction of our heuristic-based search strategies vs. NICE-MC.

trast NICE-MC with two state-of-the-art model checkers, SPIN [12] and JPF [13]. We create system models in PROMELA and Java that replicate as closely as possible the system tested in NICE. For clarity, we present the details of these modeling efforts in Appendix A and we summarize the results here:

• As expected, by using an abstract model of the system, SPIN performs a full search more efficiently than NICE. Of course, state-space explosion still occurs: *e.g.*, with 7 pings, SPIN runs of out memory. This validates our decision to maintain hashes of system states instead of keeping entire system states.

• SPIN's partial-order reduction (POR)[4], decreases the grow rate of explored transitions by only 18%. This is because POR is applied to the granularity level that cannot be refined to distinguish between independent network flows.

• Taken "as is", JPF is already slower than NICE by a factor of 290 with 3 pings. The reason is that JPF uses Java threads to represent system concurrency. However, JPF leads to too many possible thread interleavings to explore even in our small example.

• Even with our extra effort in rewriting the Java model to explicitly expose possible transitions, JPF is 5.5 times slower than NICE using 4 pings.

These results suggest that NICE, in comparison to the other model-checkers, strikes a good balance between (i) capturing system concurrency at the right level of granularity, (ii) simplifying the state space and (iii) allowing testing of unmodified controller programs.

[4]POR is a well-known technique for avoiding exploring unnecessary orderings of transitions (*e.g.*, [27]).

8 Experiences With Real Applications

In this section, we report on our experiences applying NICE to three real applications—a MAC-learning switch, a server load-balancer, and energy-aware traffic engineering—and uncovering eleven bugs.

8.1 MAC-learning Switch (PySwitch)

Our first application is the `pyswitch` software included in the NOX distribution. The application implements MAC learning, coupled with flooding to unknown destinations, common in Ethernet switches. Realizing this functionality seems straightforward (*e.g.*, the pseudocode in Figure 4), yet NICE automatically detects three violations of correctness conditions.

BUG-I: Host unreachable after moving. This fairly subtle bug is triggered when a host B moves from one location to another. Before B moves, host A starts streaming to B, which causes the controller to install a forwarding rule. When B moves, the rule stays in the switch as long as A keeps sending traffic, because the soft timeout does not expire. As such, the packets do not reach B's new location. This serious correctness bug violates the *NoBlackHoles* condition. If the rule had a *hard* timeout, the application would eventually flood packets and reach B at its new location; then, B would send return traffic that would trigger MAC learning, allowing future packets to follow a direct path to B. While this "bug fix" prevents persistent packet loss, the network still experiences *transient* loss until the hard timeout expires. Designing a new *NoBlackHoles* condition that is robust to transient loss is part of our ongoing work.

BUG-II: Delayed direct path. The `pyswitch` also violates the *StrictDirectPaths* condition, leading to suboptimal performance. The violation arises after a host A sends a packet to host B, and B sends a response packet to A. This is because `pyswitch` installs a forwarding rule in one direction—from the sender (B) to the destination (A), in line 13 of Figure 4. The controller does *not* install a forwarding rule for the other direction until seeing a subsequent packet from A to B. For a three-way packet exchange (*e.g.*, a TCP handshake), this performance bug directs 50% more traffic than necessary to the controller. Anecdotally, fixing this bug can easily

introduce another one. The naïve fix is to add another `install_rule` call, with the addresses and ports reversed, after line 14, for forwarding packets from A to B. However, since the two rules are not installed atomically, installing the rules in this order can allow the packet from B to reach A before the switch installs the second rule. This can cause a subsequent packet from A to reach the controller. A correct fix would install the rule for traffic from A first, before allowing the packet from B to A to traverse the switch. With this "fix", the resulting program satisfies the *StrictDirectPaths* property.

BUG-III: Excess flooding. When we test `pyswitch` on a topology that contains a cycle, the program violates the *NoForwardingLoops* property. This is not surprising, since `pyswitch` does not construct a spanning tree.

8.2 Web Server Load Balancer

Data centers rely on load balancers to spread incoming requests over service replicas. Previous work created a load-balancer application that uses wildcard rules to divide traffic based on the client IP addresses to achieve a target load distribution [9]. The application can dynamically adjust the load distribution by installing new wildcard rules; during the transition, old transfers complete at their existing servers while new requests are handled according to the new distribution. We test this application with one client and two servers connected to a single switch. The client opens a TCP connection to a virtual IP address corresponding to the two replicas. In addition to the default correctness properties, we create an application-specific condition *FlowAffinity* that verifies that all packets of a single TCP connection go to the same server replica. Here we report on the bugs NICE found in the original code.

BUG-IV: Next TCP packet always dropped after reconfiguration. Having observed a violation of the *NoForgottenPackets* property, we identified a bug where the application neglects to handle the "next" packet of each flow—for both ongoing transfers and new requests—after a change in the load-balancing policy. Despite correctly installing the forwarding rule for each flow, the application does *not* instruct the switch to forward the packet that triggered the `packet_in` handler. Since the TCP sender ultimately retransmits the lost packet, the program does successfully handle each Web request, making it hard to notice the bug. The bug degrades performance and, for a long execution trace, would ultimately exhaust the switch's space for buffering packets awaiting controller action.

BUG-V: Some TCP packets dropped after reconfiguration. After fixing **BUG-IV**, NICE detected another *NoForgottenPackets* violation, due to a race condition. In switching from one load-balancing policy to another, the application sends multiple updates to the switch for each existing rule: (i) a command to remove the existing forwarding rule followed by (ii) commands to install one or more rules (one for each group of affected client IP addresses) that direct packets to the controller. Since these commands are not executed atomically, packets arriving between the first and second step do not match either rule. The OpenFlow specification prescribes that packets that do not match any rule should go to the controller. Although the packets go to the controller either way, these packets arrive with a different "reason code" (*i.e.*, `NO_MATCH`). As written, the `packet_in` handler ignores such (unexpected) packets, causing the switch to hold them until the buffer fills. This appears as a packet loss to the end hosts[5]. To fix this bug, the program should reverse the two steps, installing the new rules (perhaps at a lower priority) before deleting the existing ones.

BUG-VI: ARP packets forgotten during address resolution. Another *NoForgottenPackets* violation uncovered two bugs that are similar in spirit to the previous one. The controller program handles client ARP requests. Despite sending the correct reply, the program neglects to discard the ARP request packet. A similar problem occurs for server-generated ARP messages.

BUG-VII: Duplicate SYN packets during transitions. A *FlowAffinity* violation detected a subtle bug that arises only when a connection experiences a duplicate (*e.g.*, retransmitted) SYN packet while the controller changes from one load-balancing policy to another. During the transition, the controller inspects the "next" packet of each flow, and assumes a SYN packet implies the flow is new and should follow the new load-balancing policy. Under duplicate SYN packets, some packets of a connection (arriving before the duplicate SYN) may go to one server, and the remaining packets to another, leading to a broken connection. The authors of [9] acknowledge this possibility (see footnote #2 in their paper), but only realized this problem after careful consideration.

8.3 Energy-Efficient Traffic Engineering

OpenFlow enables a network to reduce energy consumption [10,28] by selectively powering down links and redirecting traffic to alternate paths during periods of lighter load. REsPoNse [28] precomputes several routing tables (the default is two), and makes an online selection for each flow. The NOX implementation has an *always-on* routing table (that can carry all traffic under low demand) and an *on-demand* table (that serves additional traffic under higher demand). Under high load, the flows

[5]To understand the impact, consider a switch with 1 Gb/s links, 850-byte frames, and a flow-table update rate of 257 rules/s (as widely reported for the HP 5406zl). That would lead to 150 dropped packets per switch port.

should probabilistically split evenly over the two classes of paths. The application learns the link utilizations by querying the switches for port statistics. Upon receiving a packet of a new flow, the `packet_in` handler chooses the routing table, looks up the list of switches in the path, and installs a forwarding rule at each hop.

For testing with NICE, we install a network topology with three switches in a triangle, one sender host at one switch and two receivers at another switch. The third switch lies on the on-demand path. We define the following application-specific correctness property:

• *UseCorrectRoutingTable*: This property checks that the program, upon receiving a packet from an ingress switch, issues the installation of rules to all and just the switches on the appropriate path for that packet, as determined by the network load, It uses the source and destination addresses to determine the path (as does the application code). Enforcing this is important, because if it is violated, the network might be configured to carry more traffic than it physically can, degrading the performance of end-host applications running on top of the network.

NICE found several bugs in this application:

BUG-VIII: The first packet of a new flow is dropped. A violation of *NoForgottenPackets* revealed a bug that is almost identical to **BUG-IV**. The `packet_in` handler installed a rule but neglected to instruct the switch to forward the packet that triggered the event.

BUG-IX: The first few packets of a new flow can be dropped. After fixing **BUG-VIII**, NICE detected another violation of the *NoForgottenPackets* property at the second switch in the path. Since the `packet_in` handler installs an end-to-end path when the first packet of a flow enters the network, the program implicitly assumes that intermediate switches would never direct packets to the controller. However, with communication delays in installing the rules, the packet could reach the second switch before the rule is installed. Although these packets trigger `packet_in` events, the handler implicitly ignores them, causing the packets to buffer at the intermediate switch. This bug is hard to detect because the problem only arises under certain event orderings. Simply installing the rules in the reverse order, from the last switch to the first, is not sufficient—differences in the delays for installing the rules could still cause a packet to encounter a switch that has not (yet) installed the rule. A correct "fix" should either handle packets arriving at intermediate switches, or use "barriers" to ensure that rule installation completes at all intermediate hops before allowing the packet to depart the ingress switch.

BUG-X: Only on-demand routes used under high load. NICE detects a *CorrectRoutingTableUsed* violation that prevents on-demand routes from being used properly. The program updates an extra routing table in the port-statistic handler (when the network's perceived energy state changes) to either always-on or on-demand, in an effort to let the remainder of the code simply reference this extra table when deciding where to route a flow. Unfortunately, this made it impossible to split flows equally between always-on and on-demand routes, and the code directed all new flows over on-demand routes under high load. A "fix" was to abandon the extra table and choose the routing table on per-flow basis.

BUG-XI: Packets can be dropped when the load reduces. After fixing **BUG-IX**, NICE detected another violation of the *NoForgottenPackets*. When the load reduces, the program recomputes the list of switches in each always-on path. Under delays in installing rules, a switch not on these paths may send a packet to the controller, which ignores the packet because it fails to find this switch in any of those lists.

BUG	PKT-SEQ only	NO-DELAY	FLOW-IR	UNUSUAL
I	23 / 0.02	23 / 0.02	23 / 0.02	23 / 0.02
II	18 / 0.01	18 / 0.01	18 / 0.01	18 / 0.01
III	11 / 0.01	16 / 0.01	11 / 0.01	11 / 0.01
IV	386 / 3.41	1661 / 9.66	321 / 1.1	64 / 0.19
V	22 / 0.05	Missed	21 / 0.02	60 / 0.18
VI	48 / 0.05	48 / 0.06	31 / 0.04	49 / 0.07
VII	297k / 1h	191k / 39m	Missed	26.5k / 5m
VIII	23 / 0.03	22 / 0.02	23 / 0.03	23 / 0.02
IX	21 / 0.03	17 / 0.02	21 / 0.03	21 / 0.02
X	2893 / 35.2	Missed	2893 / 35.2	2367 / 25.6
XI	98 / 0.67	Missed	98 / 0.67	25 / 0.03

Table 2: Comparison of the number of transitions / running time to the first violation that uncovered each bug. Time is in seconds unless otherwise noted.

8.4 Overhead of Running NICE

In Table 2, we summarize how many seconds NICE took (and how many state transitions were explored) to discover the *first property violation* that uncovered each bug, under four different search strategies. Note the numbers are generally small because NICE quickly produces simple test cases that trigger the bugs. One exception, **BUG-VII**, is found in 1 hour by doing a PKT-SEQ-only search but UNUSUAL can detect it in just 5 minutes. Our search strategies are also generally faster than PKT-SEQ-only to trigger property violations, except in one case (**BUG-IV**). NO-DELAY takes longer for **BUG-IV** because the latter is faster to explore a sequence of transitions where the network reconfiguration event happens at the right time for experiencing a *NoForgottenPackets* violation. FLOW-IR does not produce benefits for the last four bugs because these are uncovered by test cases that do not involve using multiple flows. Also, note that only in few cases (**BUG-IV**, **BUG-X** and **BUG-XI**) the heuristic-based strategies experience false negatives. Expectedly, these race condition bugs are missed by NO-

DELAY, which does not consider rule installation delays.

Finally, the reader may find that some of the bugs we found—like persistently leaving some packets in the switch buffer—are relatively simple and their manifestations could be detected with run-time checks performed by the controller platform. However, the programmer would not know what caused it. For example, a run-time check that flags a "no forgotten packets" error due to **BUG-IV** or **BUG-V** would not tell the programmer what was special about this particular system execution that triggered the error. Subtle race conditions are very hard to diagnose, so having a (preferably small) example trace—like NICE produces—is crucial.

9 Coverage vs. Overhead Trade-Offs

Testing is inherently incomplete, walking a fine line between good coverage and low overhead. As part of our ongoing work, we want to explore further how to best leverage symbolic execution in NICE. We here discuss some limitations of our current approach.

Concrete execution on the switch: In identifying the equivalence classes of packets, the algorithm in Figure 6 implicitly assumes the packets reach the controller. However, depending on the rules already installed in the switch, some packets in a class may reach the controller while others do not. This leads to two limitations. First, if *no* packets in an equivalence class would go to the controller, generating a representative packet from this class was unnecessary. This leads to some loss in efficiency. Second, if *some* (but not all) packets go to the controller, we may miss an opportunity to test a code path through the handler by inadvertently generating a packet that stays in the "fast path" through the switches. This leads to some loss in both efficiency and coverage. We could overcome these limitations by extending symbolic execution to include our simplified switch model and performing "symbolic packet forwarding" across multiple switches. We chose not to pursue this approach because (i) symbolic execution of the flow-table code would lead to a path-explosion problem, (ii) including these variables would increase the overhead of the constraint solver, and (iii) rules that modify packet headers would further complicate the symbolic analysis. Still, we are exploring "symbolic forwarding" as future work, by leveraging reachability-analysis techniques [29].

Concrete global controller variables: In symbolically executing each event handler, NICE could miss complex dependencies *between* handler invocations. This is a byproduct of our decision to represent global controller variables in a concrete form. In some cases, one call to a handler could update the variables in a way that affects the symbolic execution of a second call (to the same handler, or a different one). Symbolic execution of the second handler would start from the *concrete* global variables, and may miss an opportunity to recognize additional constraints on packet header fields. We could overcome this limitation by running symbolic execution across multiple handler invocations, at the expense of a significant explosion in the number of code paths. Or, we could revisit our decision to represent global variables in a concrete form. As future work, we are considering ways to efficiently represent global variables symbolically.

Infinite execution trees in symbolic execution: Symbolically unrolling a "for loop" in a program can lead to an arbitrarily large state space. In our context, such an infinite execution tree [24] arises if each state has at least one input that modifies the controller state. This is an inherent limitation of symbolic execution, whether applied independently or in conjunction with model checking. To address this limitation, we explicitly bound the state space by limiting the size of the input (*e.g.*, a limit on the number of packets) and devise OpenFlow-specific search strategies that explore the system state space efficiently. These heuristics offer a tremendous improvement in efficiency, at the expense of some loss in coverage.

10 Related Work

Bug finding. While model checking [12–16] and symbolic execution [17, 18, 25] are automatic techniques, a drawback is that they typically require a closed system, *i.e.*, a system (model) together with its environment. Typically, the creation of such environment is a manual process (*e.g.*, [23]). NICE re-uses the idea of model checking—systematic state-space exploration—and combines it with the idea of symbolic execution—exhaustive path coverage—to avoid pushing the burden of modeling the environment on the user. Also, NICE is the first to demonstrate the applicability of these techniques for testing the dynamic behavior of OpenFlow networks. Finally, NICE makes a contribution in managing state-space explosion for this specific domain.

Khurshid *et al.* [24] enable a model checker to perform symbolic execution. Both our and their work share the spirit of using symbolic variables to represent data from very large domains. Our approach differs in that it uses state matching and symbolic execution in a selective way for uncovering possible transitions given a certain controller state. As a result, we (i) reduce state-space explosion due to feasible code paths because not all code is symbolically executed, and (ii) preserve system state as a first-class notion that is used to further reduce the search of the state-space.

OpenFlow and network testing. Frenetic [30] is a domain-specific language for OpenFlow that aims to eradicate a large class of programming faults. Using Fre-

netic requires the network programmer to learn extensions to Python to support the higher-layer abstractions.

OFRewind [31] enables recording and replay of events for troubleshooting problems in production networks due to closed-source network devices. However, it does not automate the testing of OpenFlow controller programs.

Mai *et al.* [32] use static analysis of network devices forwarding information bases to uncover problems in the data plane. FlowChecker [33] applies symbolic model checking techniques on a manually-constructed network model based on binary decision diagrams to detect misconfigurations in OpenFlow forwarding tables. We view these works as orthogonal to ours since they both aim to analyze a snapshot of the data plane.

Bishop *et al.* [34] examine the problem of testing the specification of end host protocols. NICE tests the network itself, in a new domain of software defined networks. Kothari *et al.* [35] use symbolic execution and developer input to identify protocol manipulation attacks for network protocols. In contrast, NICE combines model checking with symbolic execution to identify relevant test inputs for injection into the model checker.

11 Conclusion

We built NICE, a tool for automating the testing of OpenFlow applications that combines model checking and concolic execution to quickly explore the state space of unmodified controller programs written for the popular NOX platform. Further, we devised a number of new, domain-specific techniques for mitigating the state-space explosion that plagues approaches such as ours. We contrast NICE with an approach that applies off-the-shelf model checkers to the OpenFlow domain, and demonstrate that NICE is five times faster even on small examples. We applied NICE to implementations of three important applications, and found 11 bugs. A release of NICE will be made publicly available.

We plan to use the simplified switch model as the basis for testing the OpenFlow controller program with real switch implementations: run n-versions of OpenFlow switches side-by-side with the switch model and automatically detect deviant behaviors.

Acknowledgments.

We are grateful to Stefan Bucur, Olivier Crameri, Johannes Kinder, Viktor Kuncak, Darko Marinov and Sharad Malik for useful discussions and comments on earlier drafts of this work. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement 259110.

A Model-checking with Existing Tools

SPIN [12] is one of the most popular tools for verifying the correctness of distributed software models. In this case, these are written in a high-level modeling language called PROMELA. This language exposes nondeterminism as a first-class concept, making it easier to model the concurrency in OpenFlow. However, using this language proficiently is non-trivial and it took several person-days to implement the model of the simple OpenFlow system (Figure 1). To capture the system concurrency at the right level of granularity, we use the `atomic` language feature to model each transition as a single atomic computation that cannot be interleaved to any other transition. In practice, this behavior cannot be faithfully modeled due to the blocking nature of `channels` in PROMELA. To enable SPIN's POR be most effective, we assign exclusive rights to the processes involved in each communication channel.

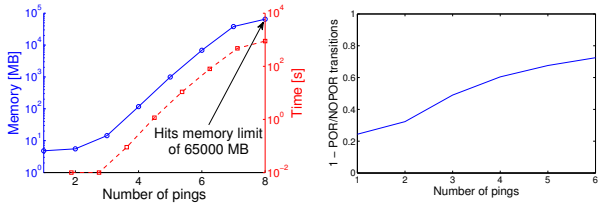
Figure 9a shows the memory usage and elapsed time for the exhaustive search with POR as we increase the number of packets sent by host 1. As expected, we observe an exponential increase in computational resources until SPIN reaches the memory limit when checking the model with 8 pings (*i.e.*, 16 packets).

To see how effective POR is, we compare in Figure 9b the number of transitions explored with POR vs. without POR (NOPOR) while we vary the number of pings. In relative terms, POR's efficiency increases, although with diminishing returns, from 24% to 73% as we inject more packets that are identical to each other. The benefits due to POR on elapsed time follow a similar trend and POR can finish 6 pings in 28% of time used by NOPOR. However, NOPOR hits the memory limit at 7 pings, so POR only adds one extra ping.

Finally, we test if POR can reduce the search space by taking advantage of one simple rule of independence for the networking domain: *i.e.*, packets involving disjoint pairs of source and destination addresses are completely independent. Unfortunately, we observe that there is no reduction when we inject two packets with distinct address pairs compared to the case with identical packets. This is because SPIN uses the accesses to communication channels to derive the independence of events.

Java PathFinder (JPF) [13] is one among the first modern model checkers which use the implementation in place of the model. We follow two approaches to model the system by porting our Python code to Java.

In the first approach, we naively use threads to capture nondeterminism, hoping that JPF's automatic state-space reduction techniques would cope with different thread creation orders of independent transitions. However, in our case, the built-in POR is not very efficient in removing unnecessary network event interleav-



(a) Memory usage and elapsed time (log y-scales). **(b) Efficiency of POR.**

Figure 9: SPIN: Exponential increase in computational resources partially mitigated by POR.

Ping	Time [s]	Unique states	End states	Mem. [MB]
1	0	55	2	17
2	9	20638	134	140
3	13689	25470986	2094	1021

Table 3: JPF: Exhaustive search on thread-based model.

Ping	Time [s]	Unique states	End states	Mem. [MB]
1	0	1	1	17
2	1	691	194	33
3	16	29930	6066	108
4	11867	16392965	295756	576

Table 4: JPF: Exhaustive search on choice-based model.

ings because thread interleaving happens at finer granularity than event interleavings. To solve this problem, we tune this model by using the `beginAtomic()` and `endAtomic()` functions provided by JPF. As this still produces too many possible interleavings, we further introduced a global lock.

In a second approach to further refine the model, we capture nondeterminism via JPF's choice generator: `Verify.getInt()`. This gives a significant improvement over threads, mainly because we are able to specify precisely the granularity of interleavings. However, this second modeling effort is non trivial since we are manually enumerating the state space and there are several caveats in this case too. For example, explicit choice values should not be saved on the stack as the choice value may became a part of the global state, thus preventing reduction. The vector of possible transitions must also be sorted[6].

Table 3 illustrates the state space explosion when using the thread-based model. Unfortunately, as show in Table 4, the choice-based model improves only by 1 ping the size of the model that we can explore within a comparable time period ($\approx$ 4 hours).

[6]We order events by their states' hash values.

References

[1] AfNOG Takes Byte Out of Internet. `http://goo.gl/HVJw5`.

[2] Reckless Driving on the Internet. `http://goo.gl/otilX`.

[3] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker, and J. Turner. OpenFlow: Enabling Innovation in Campus Networks. *SIGCOMM Comput. Commun. Rev.*, 38:69–74, March 2008.

[4] M. Canini, D. Kostić, J. Rexford, and D. Venzano. Automating the Testing of OpenFlow Applications. In *WRiPE*, 2011.

[5] N. Gude, T. Koponen, J. Pettit, B. Pfaff, M. Casado, N. McKeown, and S. Shenker. NOX: Towards an Operating System for Networks. *SIGCOMM Comput. Commun. Rev.*, 38:105–110, July 2008.

[6] M. Casado, M. J. Freedman, J. Pettit, J. Luo, N. Gude, N. McKeown, and S. Shenker. Rethinking Enterprise Network Control. *IEEE/ACM Transactions on Networking*, 17(4), August 2009.

[7] A. Nayak, A. Reimers, N. Feamster, and R. Clark. Resonance: Dynamic Access Control for Enterprise Networks. In *WREN*, 2009.

[8] N. Handigol et al. Plug-n-Serve: Load-Balancing Web Traffic using OpenFlow, August 2009. SIGCOMM Demo.

[9] R. Wang, D. Butnariu, and J. Rexford. OpenFlow-Based Server Load Balancing Gone Wild. In *HotICE*, 2011.

[10] B. Heller, S. Seetharaman, P. Mahadevan, Y. Yiakoumis, P. Sharma, S. Banerjee, and N. McKeown. ElasticTree: Saving Energy in Data Center Networks. In *NSDI*, 2010.

[11] D. Erickson et al. A Demonstration of Virtual Machine Mobility in an OpenFlow Network, August 2008. SIGCOMM Demo.

[12] G. Holzmann. *The Spin Model Checker - Primer and Reference Manual*. Addison-Wesley, Reading Massachusetts, 2004.

[13] W. Visser, K. Havelund, G. Brat, S. Park, and F. Lerda. Model Checking Programs. *Automated Software Engineering*, 10(2):203–232, 2003.

[14] M. Musuvathi and D. R. Engler. Model Checking Large Network Protocol Implementations. In *NSDI*, 2004.

[15] C. E. Killian, J. W. Anderson, R. Jhala, and A. Vahdat. Life, Death, and the Critical Transition: Finding Liveness Bugs in Systems Code. In *NSDI*, 2007.

[16] J. Yang, T. Chen, M. Wu, Z. Xu, X. Liu, H. Lin, M. Yang, F. Long, L. Zhang, and L. Zhou. MODIST: Transparent Model Checking of Unmodified Distributed Systems. In *NSDI*, 2009.

[17] C. Cadar, D. Dunbar, and D. R. Engler. KLEE: Unassisted and Automatic Generation of High-Coverage Tests for Complex Systems Programs. In *OSDI*, 2008.

[18] S. Bucur, V. Ureche, C. Zamfir, and G. Candea. Parallel Symbolic Execution for Automated Real-World Software Testing. In *EuroSys*, 2011.

[19] C. Baier and J.-P. Katoen. *Principles of Model Checking*. The MIT Press, 2008.

[20] Open vSwitch: An Open Virtual Switch. `http://openvswitch.org`.

[21] OpenFlow Switch Specification. `http://www.openflow.org/documents/openflow-spec-v1.1.0.pdf`.

[22] A. Sobeih, M. D'Amorim, M. Viswanathan, D. Marinov, and J. C. Hou. Assertion Checking in J-Sim Simulation Models of Network Protocols. *Simulation*, 86, 2010.

[23] R. Sasnauskas, O. Landsiedel, M. H. Alizai, C. Weise, S. Kowalewski, and K. Wehrle. KleeNet: Discovering Insidious Interaction Bugs in Wireless Sensor Networks Before Deployment. In *IPSN*, 2010.

[24] S. Khurshid, C. S. Păsăreanu, and W. Visser. Generalized Symbolic Execution for Model Checking and Testing. In *TACAS*, 2003.

[25] P. Godefroid, N. Klarlund, and K. Sen. DART: Directed Automated Random Testing. In *PLDI*, 2005.

[26] V. Ganesh and D. L. Dill. A Decision Procedure for Bit-Vectors and Arrays. In *CAV*, 2007.

[27] C. Flanagan and P. Godefroid. Dynamic Partial-Order Reduction for Model Checking Software. In *POPL*, 2005.

[28] N. Vasić, D. Novaković, S. Shekhar, P. Bhurat, M. Canini, and D. Kostić. Identifying and using energy-critical paths. In *CoNEXT*, 2011.

[29] G. Xie, J. Zhang, D. Maltz, H. Zhang, A. Greenberg, G. Hjalmtysson, and J. Rexford. On Static Reachability Analysis of IP Networks. In *IEEE INFOCOM*, 2005.

[30] N. Foster, R. Harrison, M. J. Freedman, C. Monsanto, J. Rexford, A. Story, and D. Walker. Frenetic: A Network Programming Language. In *ICFP*, 2011.

[31] A. Wundsam, D. Levin, S. Seetharaman, and A. Feldmann. OFRewind: Enabling Record and Replay Troubleshooting for Networks. In *USENIX ATC*, 2011.

[32] H. Mai, A. Khurshid, R. Agarwal, M. Caesar, P. B. Godfrey, and S. T. King. Debugging the Data Plane with Anteater. In *SIGCOMM*, 2011.

[33] E. Al-Shaer and S. Al-Haj. FlowChecker: Configuration Analysis and Verification of Federated OpenFlow Infrastructures. In *SafeConfig*, 2010.

[34] S. Bishop, M. Fairbairn, M. Norrish, P. Sewell, M. Smith, and K. Wansbrough. Rigorous Specification and Conformance Testing Techniques for Network Protocols, as applied to TCP, UDP, and Sockets. In *SIGCOMM*, 2005.

[35] N. Kothari, R. Mahajan, T. Millstein, R. Govindan, and M. Musuvathi. Finding Protocol Manipulation Attacks. In *SIGCOMM*, 2011.