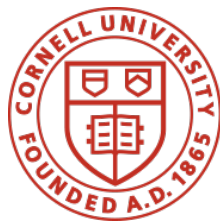


The Next 700 Network Programming Languages

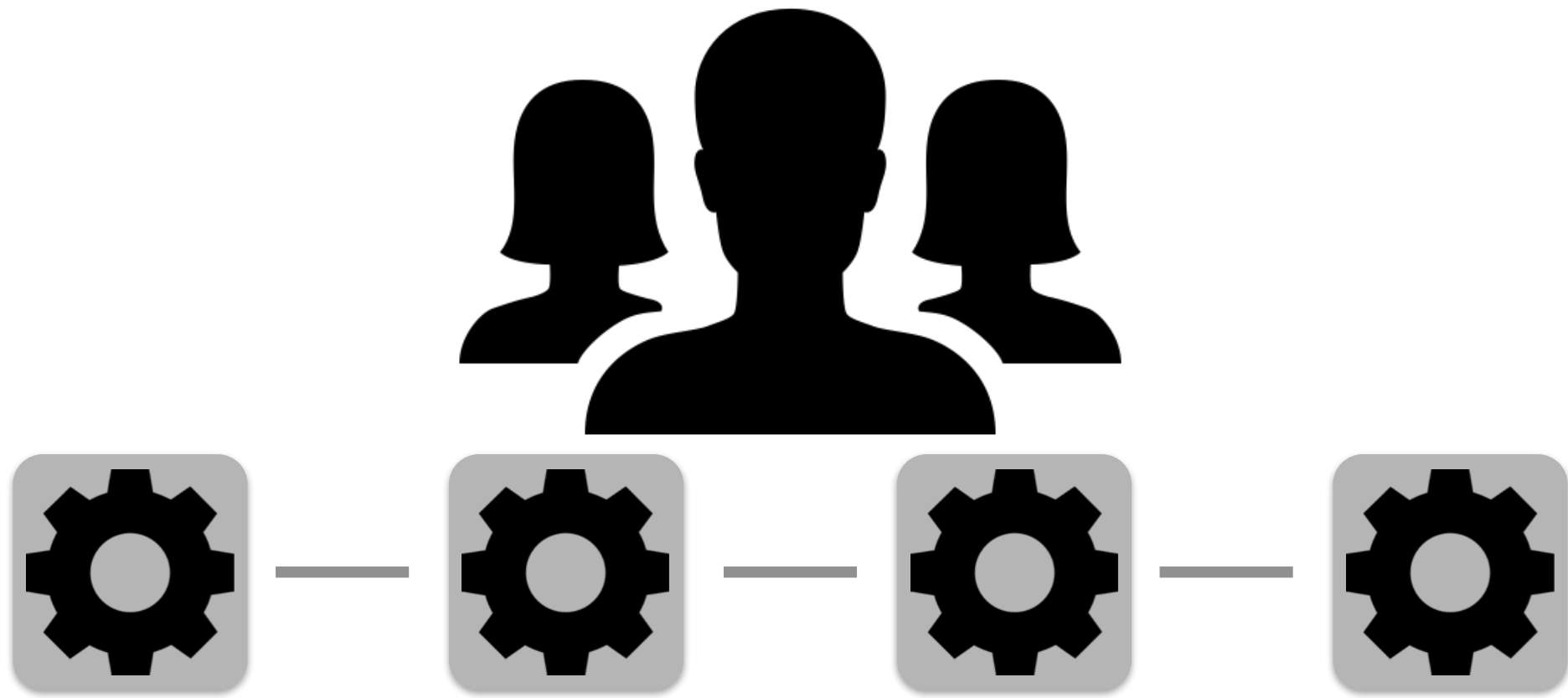
Nate Foster
Cornell University



ACM SIGCOMM NetPL '16

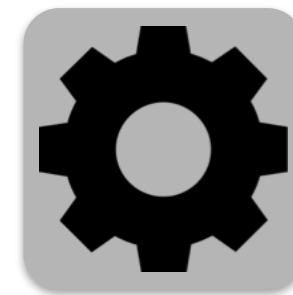
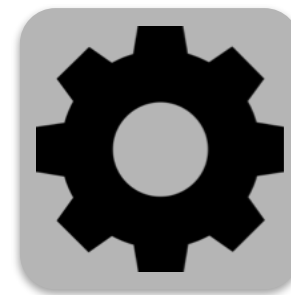


workshop(Networking & Programming Languages)
{ Brazil, 2016 }





Operating Systems

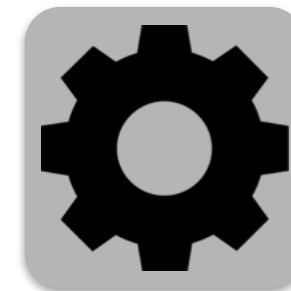




Programming Languages



Operating Systems





The Next 700 Programming Languages

P. J. Landin

Univac Division of Sperry Rand Corp., New York, New York

"... today ... 1,700 special programming languages used to 'communicate' in over 700 application areas."—*Computer Software Issues*, an American Mathematical Association Prospectus, July 1965.

A family of unimplemented computing languages is described that is intended to span differences of application area by a unified framework. This framework dictates the rules about the uses of user-coined names, and the conventions about characterizing functional relationships. Within this framework the design of a specific language splits into two independent parts. One is the choice of written appearances of programs (or more generally, their physical representation). The other is the choice of the abstract entities (such as numbers, character-strings, lists of them, functional relations among them) that can be referred to in the language.

The system is biased towards "expressions" rather than "statements." It includes a nonprocedural (purely functional) subsystem that aims to expand the class of users' needs that can be met by a single print-instruction, without sacrificing the important properties that make conventional right-hand-side expressions easy to construct and understand.

1. Introduction

Most programming languages are partly a way of expressing things in terms of other things and partly a basic set of given things. The Iswim (If you See What I Mean) system is a byproduct of an attempt to disentangle these two aspects in some current languages.

This attempt has led the author to think that many linguistic idiosyncracies are concerned with the former rather than the latter, whereas aptitude for a particular class of tasks is essentially determined by the latter rather than the former. The conclusion follows that many language characteristics are irrelevant to the alleged problem orientation.

Iswim is an attempt at a general purpose system for describing things in terms of other things, that can be problem-oriented by appropriate choice of "primitives." So it is not a language so much as a family of languages, of which each member is the result of choosing a set of primitives. The possibilities concerning this set and what is needed to specify such a set are discussed below.

Iswim is not alone in being a family, even after mere syntactic variations have been discounted (see Section 4). In practice, this is true of most languages that achieve more than one implementation, and if the dialects are well disciplined, they might with luck be characterized as

Presented at an ACM Programming Languages and Pragmatics Conference, San Dimas, California, August 1965.

¹ There is no more use or mention of λ in this paper—cognoscenti will nevertheless sense an undercurrent. A not inappropriate title would have been "Church without lambda."

differences in the set of things provided by the library or operating system. Perhaps had ALGOL 60 been launched as a family instead of proclaimed as a language, it would have fielded some of the less relevant criticisms of its deficiencies.

At first sight the facilities provided in Iswim will appear comparatively meager. This appearance will be especially misleading to someone who has not appreciated how much of current manuals are devoted to the explanation of common (i.e., problem-orientation independent) logical structure rather than problem-oriented specialties. For example, in almost every language a user can coin names, obeying certain rules about the contexts in which the name is used and their relation to the textual segments that introduce, define, declare, or otherwise constrain its use. These rules vary considerably from one language to another, and frequently even within a single language there may be different conventions for different classes of names, with near-analogies that come irritatingly close to being exact. (Note that restrictions on what names can be coined also vary, but these are trivial differences. When they have any logical significance it is likely to be pernicious, by leading to puns such as ALGOL's integer labels.)

So rules about user-coined names is an area in which we might expect to see the history of computer applications give ground to their logic. Another such area is in specifying functional relations. In fact these two areas are closely related since any use of a user-coined name implicitly involves a functional relation; e.g., compare

$x(x+a)$ $f(b+2c)$
where $x = b + 2c$ **where** $f(x) = x(x+a)$

Iswim is thus part programming language and part program for research. A possible first step in the research program is 1700 doctoral theses called "A Correspondence between x and Church's λ -notation."¹

2. The where-Notation

In ordinary mathematical communication, these uses of '**where**' require no explanation. Nor do the following:

$f(b+2c) + f(2b-c)$
where $f(x) = x(x+a)$
 $f(b+2c) + f(2b-c)$
where $f(x) = x(x+a)$
and $b = u/(u+1)$
and $c = v/(v+1)$
 $g(f \text{ where } f(x) = ax^2 + bx + c,$
 $u/(u+1),$
 $v/(v+1))$
where $g(f, p, q) = f(p+2q, 2p-q)$

“Most programming languages are partly a way of expressing things in terms of other things and partly a basic set of given things.”

The Next 700 Programming Languages

P. J. Landin

Univac Division of Sperry Rand Corp., New York, New York

“... today ... 1,700 special programming languages used to ‘communicate’ in over 700 application areas.”—*Computer Software Issues*, an American Mathematical Association Prospectus, July 1965.

A family of unimplemented computing languages is described, each intended to provide a different approach to the problem of expressing things in terms of other things. This framework, rather than the rules about the uses of user-coined names, and the conventions about characterizing functional relationships within this framework the design of a new language splits into two independent parts. One is the choice of written appearances of programs (or more generally, their physical representation). The other is the choice of the abstract entities (such as numbers, characters, strings, lists, etc., functions, relations, and so on) that can be referred to in the language. The latter is picked out towards ‘expressing’ rather than ‘stating’ and includes a nonprocedural (usually functional) subsystem that aims to expand the class of users’ needs that can be met by a single print-instruction, without sacrificing the important properties that make conventional right-hand-side expressions easy to construct and understand.

Introduction

Most programming languages are partly a way of expressing things in terms of other things and partly a basic set of given things. The latter (If you like, you may call it a byproduct of an attempt to disentangle these two aspects in some current languages.

This attempt has led the author to think that many linguistic idiosyncracies are concerned with the former rather than the latter, whereas aptitude for a particular class of task is essentially determined by the latter rather than the former. The conclusion is drawn that many linguistic idiosyncracies are irrelevant to the algebraic problem orientation.

ISWIM is an attempt at a general purpose system for describing things in terms of other things, that can be problem-oriented by appropriate choice of “primitives.” So it is not a language so much as a family of languages, each with members that result from choosing different primitives. The possibilities concerning the choice of what is not to specify such a system are discussed below.

ISWIM is not alone in being a family, but unlike many syntactic variations have been discounted (see Section 4). In practice, this is true of most languages that achieve more than one implementation, and if the dialects are well disciplined, they might with luck be characterized as

Presented at an ACM Programming Languages and Pragmatics Conference, San Dimas, California, August 1965.

¹ There is no more use or mention of λ in this paper—cognoscenti will nevertheless sense an undercurrent. A not inappropriate title would have been “Church without lambda.”

differences in the set of things provided by the library or in the content. Perhaps because 1960 has been launched as a family-oriented programming language, it would have added some of the less relevant criticisms of its deficiencies.

At first sight the facilities provided in ISWIM will appear comparatively meager. This appearance will be especially misleading to someone who has not appreciated how much of current manuals are devoted to the explanation of features (i.e., problem-orientation independent logical structure rather than problem-oriented specializations). For example, almost every language has a convention for checking a main rule about the context in which the name is used and their relation to the textual segments that introduce, define, declare, or otherwise constrain its use. These rules vary considerably from one language to another, and frequently even within a single language there may be different conventions for different classes of names, with no analogies to the corresponding conventions exact. (Note that restrictions on user-coined names can be named also, but these are trivial differences. Who has any logical significance in the way to be discerned, by leading to puns such as ALGOL’s integer labels.)

So rules about user-coined names is an area in which we might expect to see the history of computer applications give ground to their logic. Another such area is in specifying functional relations. In fact, these two areas are closely related since any use of a user-coined name implicitly involves a functional relation, e.g., $x(x+2)$ where $x = b+2c$ where $f(x) = x(x+a)$.

ISWIM is thus part programming language and part program for research. A possible first step in the research program is 1700 doctoral theses called “A Correspondence between λ and Church’s λ -notation.”¹

2. The where notation. In ordinary mathematical communication, these uses of “where” require no explanation, but in the following:

$f(b+2c) + f(2b-c)$
where $f(x) = x(x+a)$
 $f(b+2c) + f(2b-c)$
where $f(x) = x(x+a)$
and $b = u/(u+1)$
and $c = v/(v+1)$
 $g(f \text{ where } f(x) = ax^2 + bx + c,$
 $u/(u+1),$
 $v/(v+1))$
where $g(f, p, q) = f(p+2q, 2p-q)$

NetKAT Language



A domain-specific language for network programming that offers...

- Boolean predicates
- Regular expressions
- Modular composition
- Network-wide visibility and control

... embedded within a language with standard programming constructs (assignment, conditionals, loops, etc.)

NetKAT Language



A domain-specific language for network programming that offers...

- Boolean predicates *← Packet classification!*
- Regular expressions
- Modular composition
- Network-wide visibility and control

... embedded within a language with standard programming constructs (assignment, conditionals, loops, etc.)

NetKAT Language



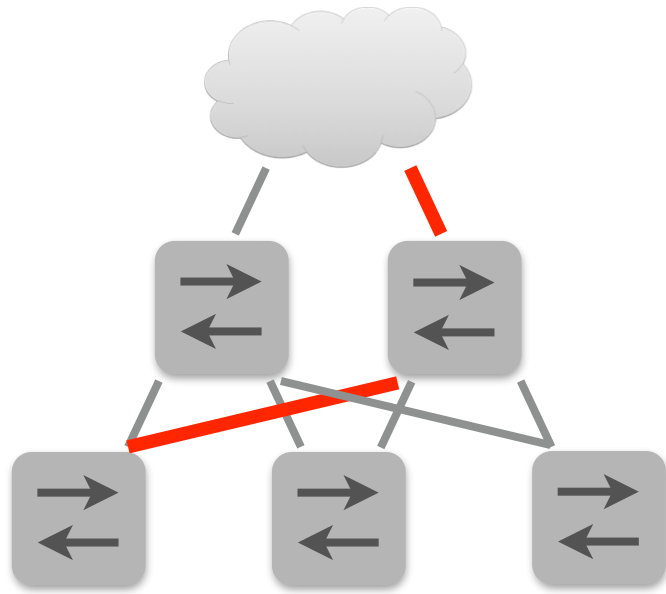
A domain-specific language for network programming that offers...

- Boolean predicates *← Packet classification!*
- Regular expressions *← Forwarding paths!*
- Modular composition
- Network-wide visibility and control

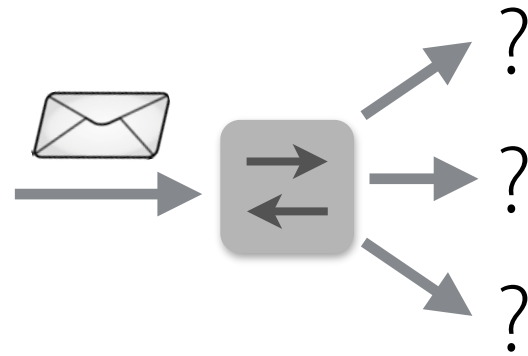
... embedded within a language with standard programming constructs (assignment, conditionals, loops, etc.)

NetKAT Design

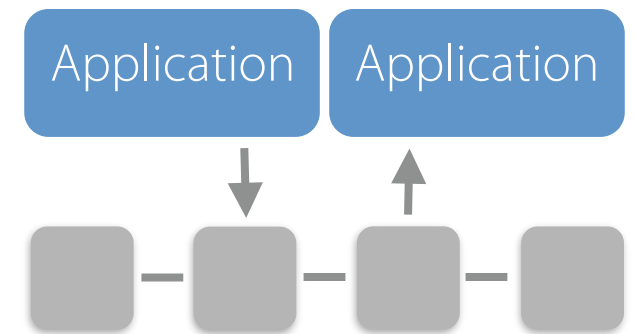
NetKAT Features



Network-wide
Abstractions



Rich Packet
Classification



Modular
Composition

NetKAT Syntax

```
pol ::= false
      | true
      |  $f = n$ 
      |  $f := n$ 
      |  $pol_1 + pol_2$ 
      |  $pol_1 \bullet pol_2$ 
      |  $!pol$ 
      |  $pol^*$ 
      | dup
```

NetKAT Syntax

```
pol ::= false
      | true
      |  $f = n$ 
      |  $f := n$ 
      |  $pol_1 + pol_2$ 
      |  $pol_1 \bullet pol_2$ 
      |  $!pol$ 
      |  $pol^*$ 
      | dup
```

Boolean
Predicates
+
Regular
Expressions
+
Packet
Primitives

Negation may only be applied to *Boolean predicates*:
true, **false**, $f = n$, closed under $+$, \bullet , and $!$

NetKAT Syntax

```
pol ::= false
      | true
      |  $f = n$ 
      |  $f := n$ 
      |  $pol_1 + pol_2$ 
      |  $pol_1 \bullet pol_2$ 
      |  $!pol$ 
      |  $pol^*$ 
      | dup
```

Boolean
Predicates
+
Regular
Expressions
+
Packet
Primitives

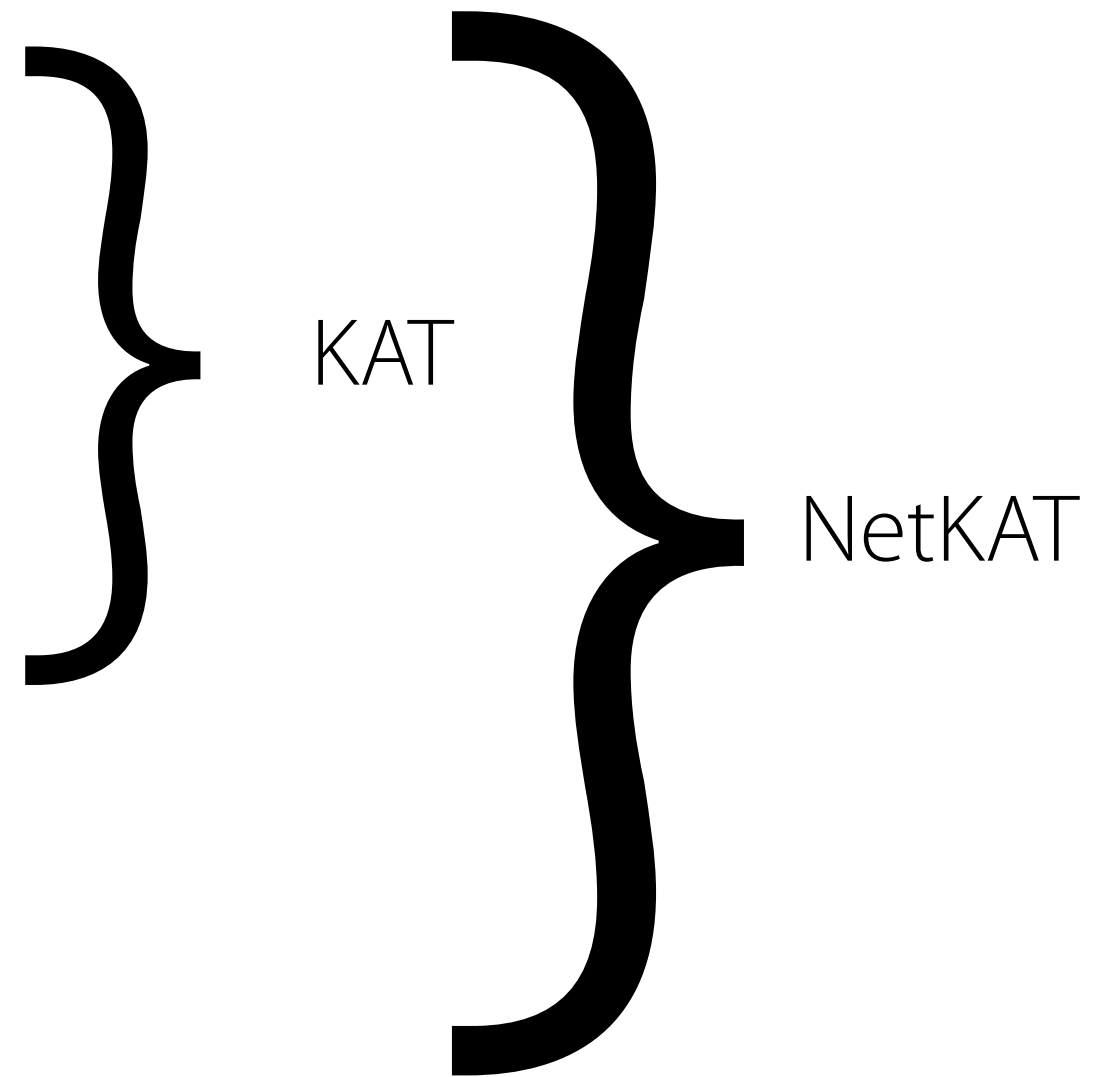
} KAT

Negation may only be applied to *Boolean predicates*:
true, **false**, $f = n$, closed under $+$, \bullet , and $!$

NetKAT Syntax

```
pol ::= false
      | true
      |  $f = n$ 
      |  $f := n$ 
      |  $pol_1 + pol_2$ 
      |  $pol_1 \bullet pol_2$ 
      |  $!pol$ 
      |  $pol^*$ 
      | dup
```

Boolean
Predicates
+
Regular
Expressions
+
Packet
Primitives



Negation may only be applied to *Boolean predicates*:
true, **false**, $f = n$, closed under $+$, \bullet , and $!$

NetKAT Syntax

$\text{pol} ::= \text{false}$

Boolean

if b **then** p_1 **else** $p_2 \triangleq (b \bullet p_1) + (!b \bullet p_2)$

while b **do** $p \triangleq (b \bullet p)^* \bullet !b$

$p^+ \triangleq p \bullet p^*$

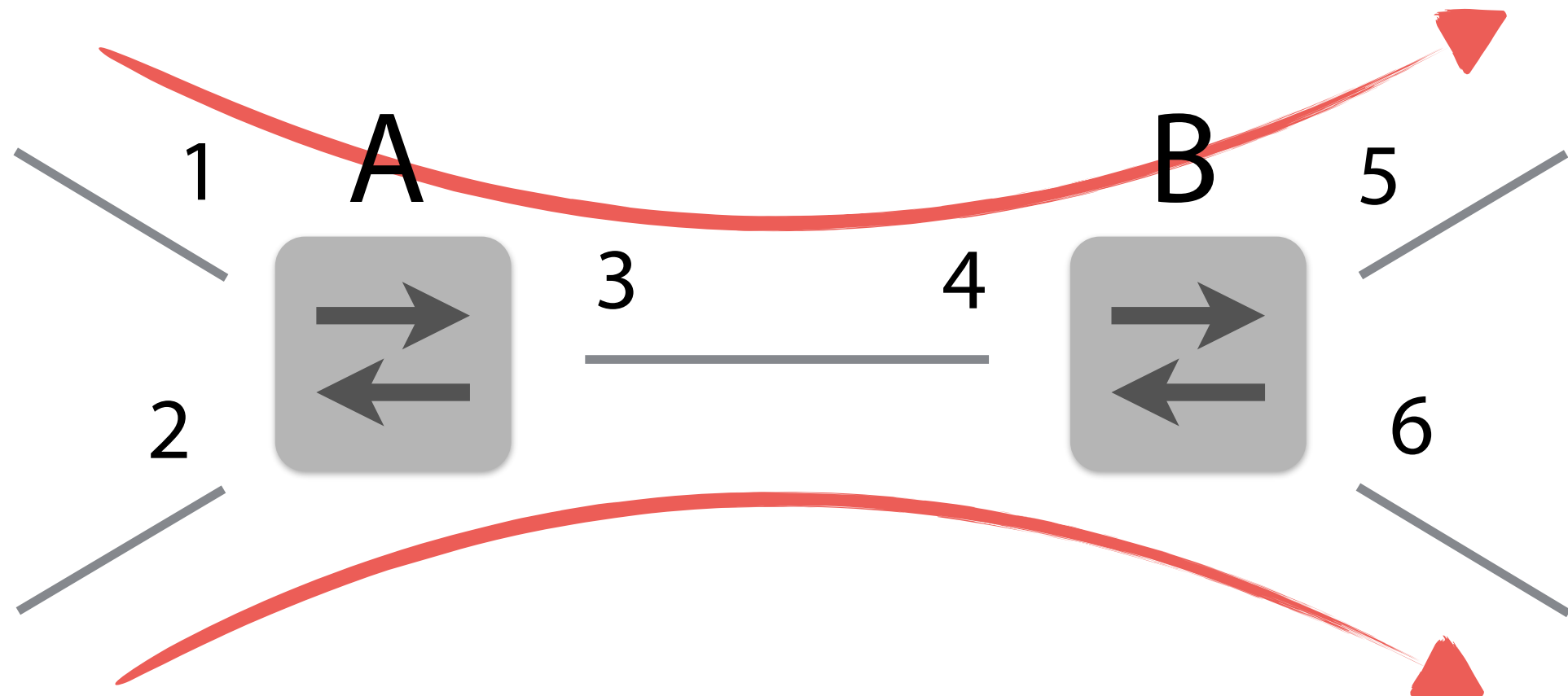
$S \rightarrow S' \triangleq \text{sw} = S \bullet \text{dup} \bullet \text{sw} := S' \bullet \text{dup}$

| **dup**

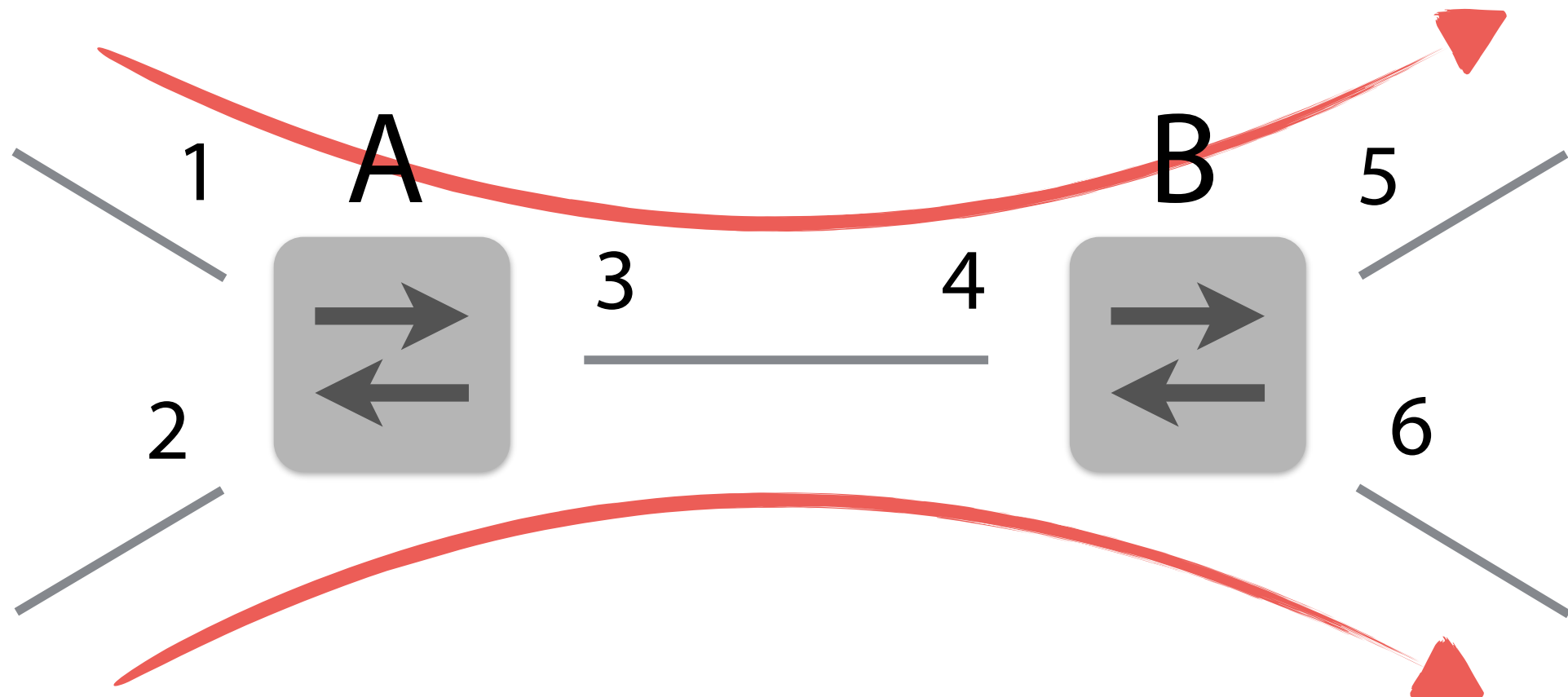
Negation may only be applied to *Boolean predicates*:

true, **false**, $f = n$, closed under $+$, \bullet , and $!$

Example



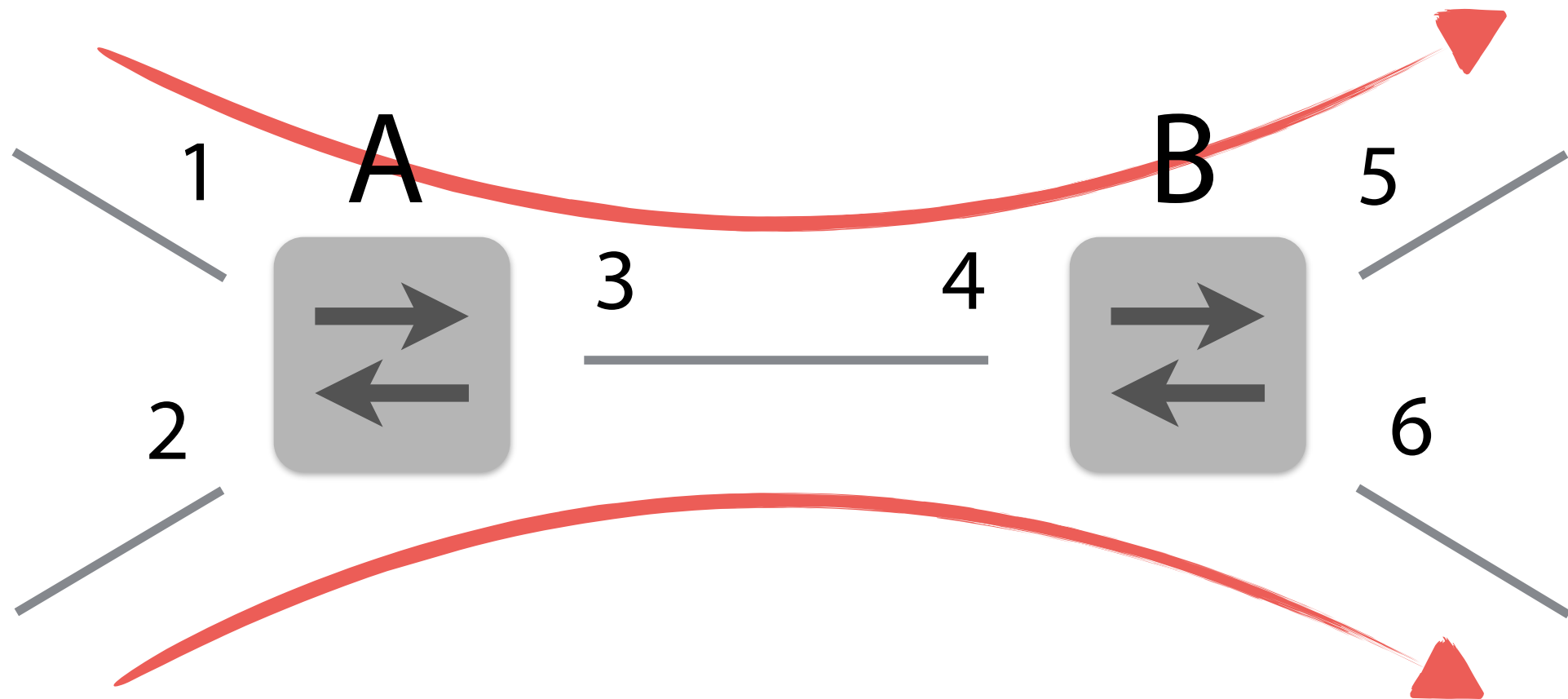
Local Program



port=1 • tag:=1 • port:=3
+
port=2 • tag:=2 • port:=3

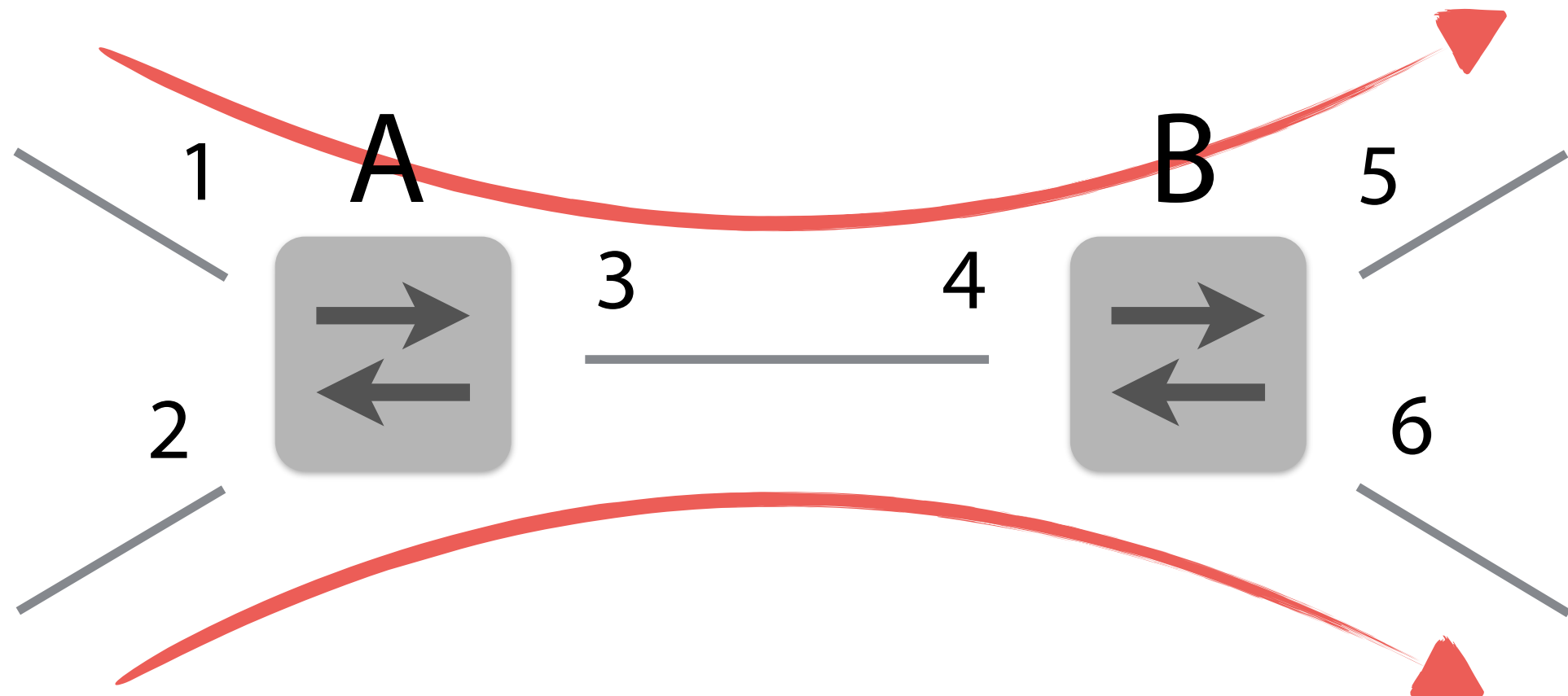
tag=1 • port:=5
+
tag=2 • port:=6

Global Program

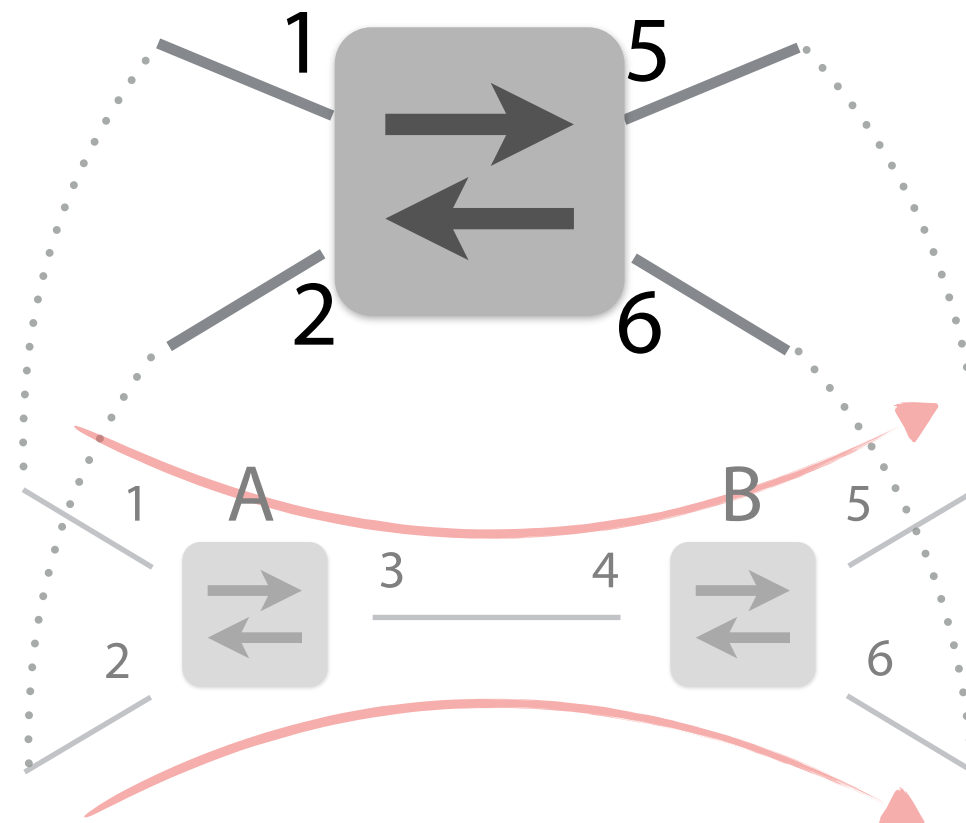


port=1 • **A** → **B** • port:=5
+
port=2 • **A** → **B** • port:=6

Virtual Program

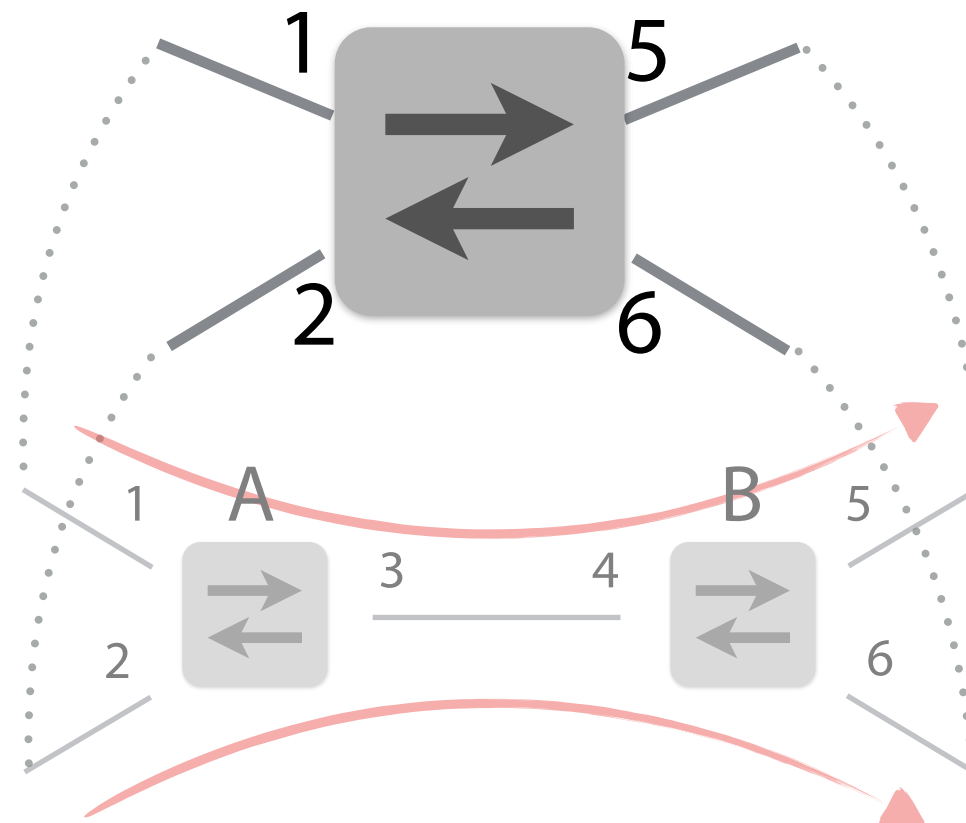


Virtual Program



virtual "big switch"

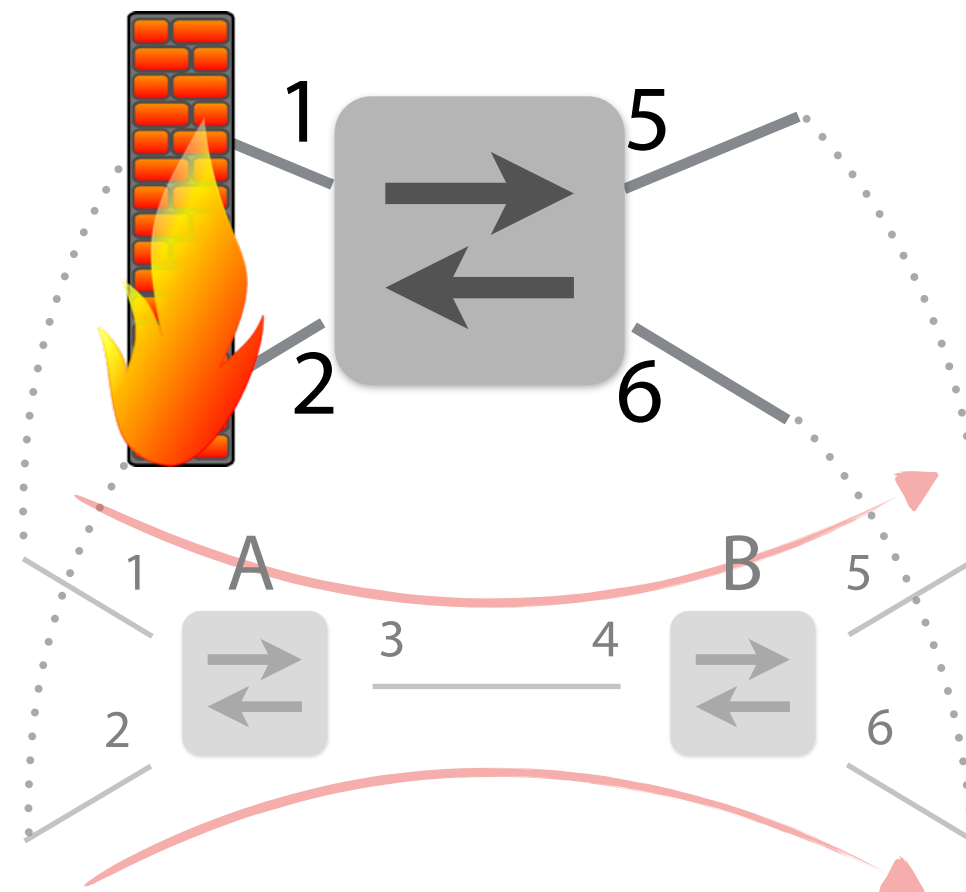
Virtual Program



virtual "big switch"

port=1 • port:=5
+
port=2 • port:=6

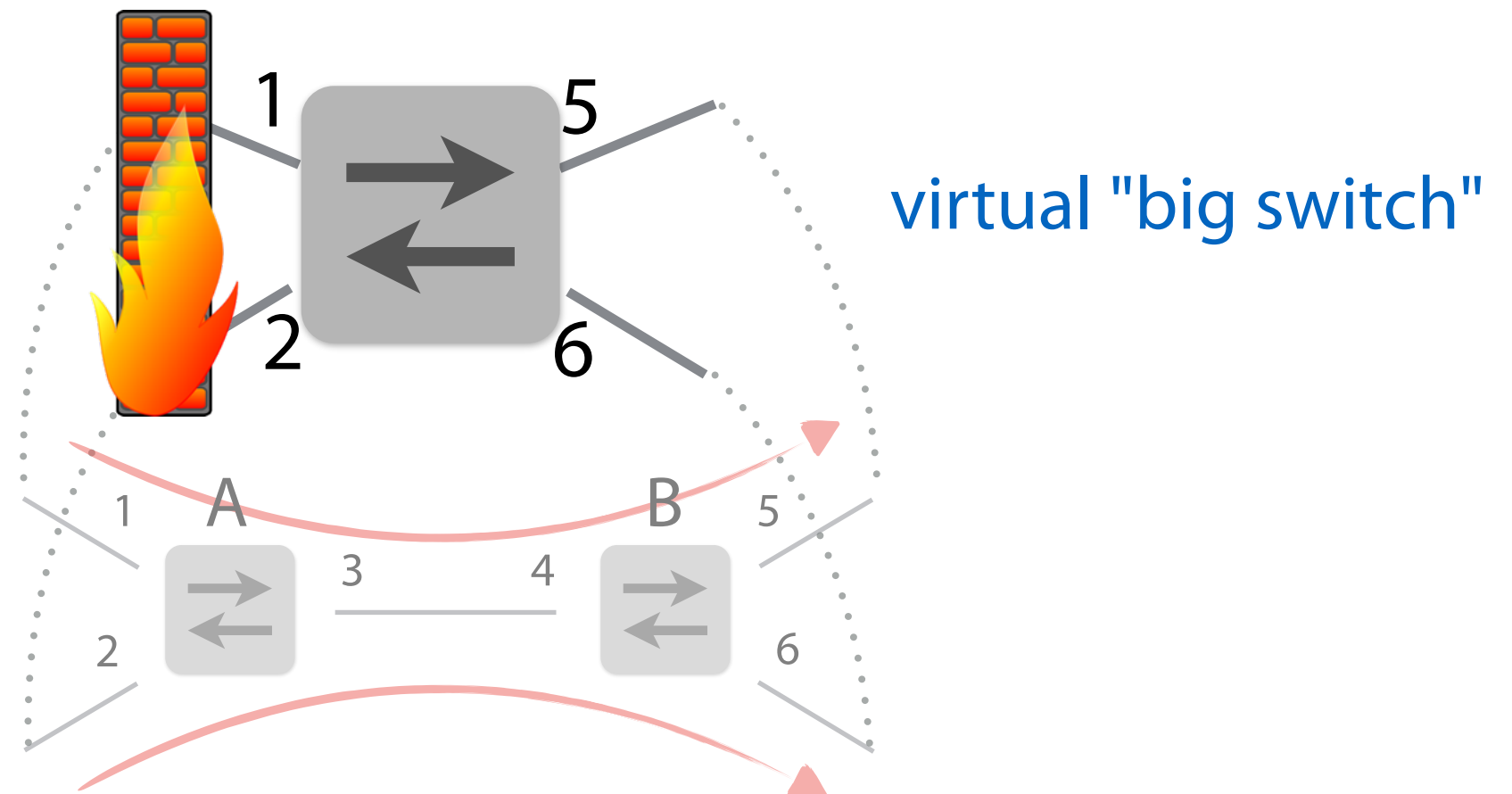
Virtual Program



virtual "big switch"

port=1 • port:=5
+
port=2 • port:=6

Virtual Program



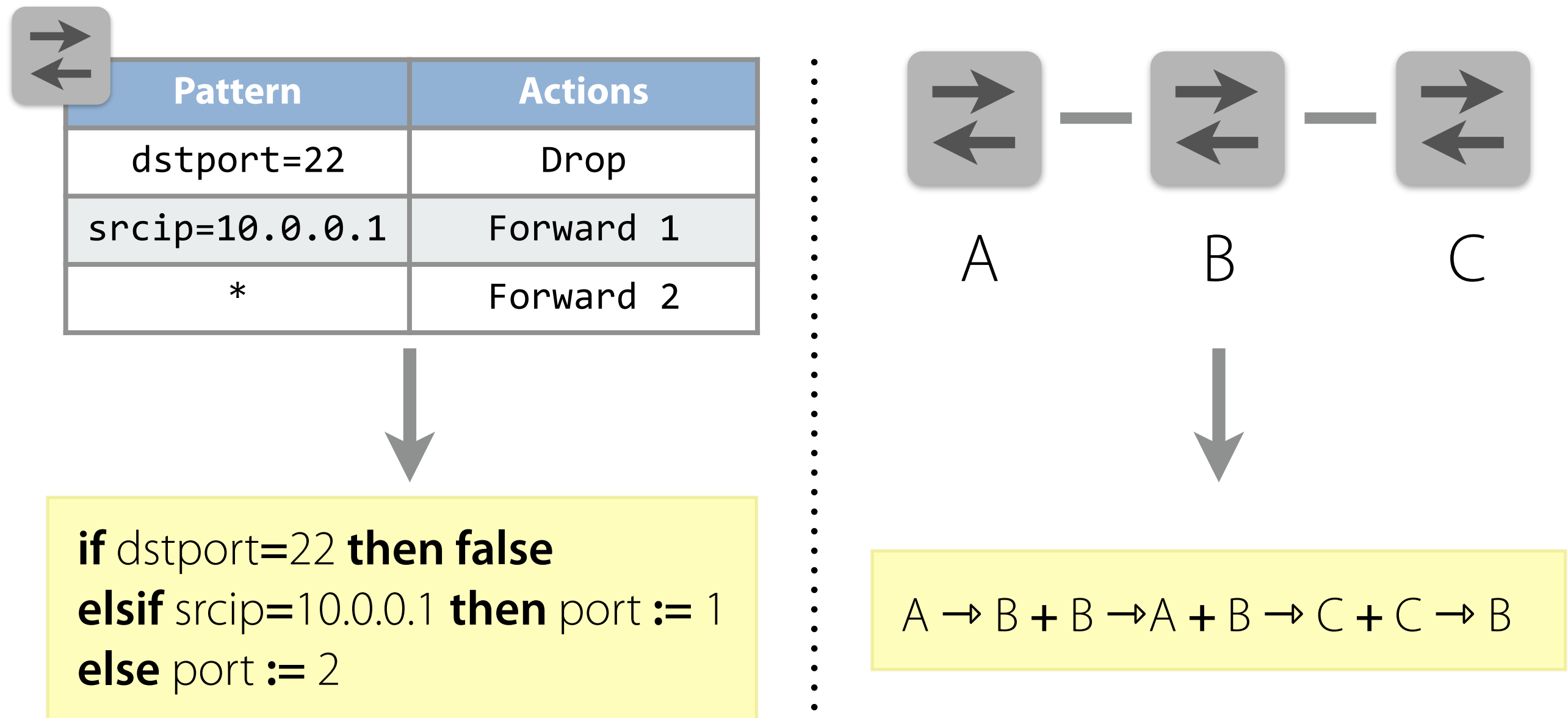
! (tcpSrcPort = 22)

•

port=1 • port:=5
+
port=2 • port:=6

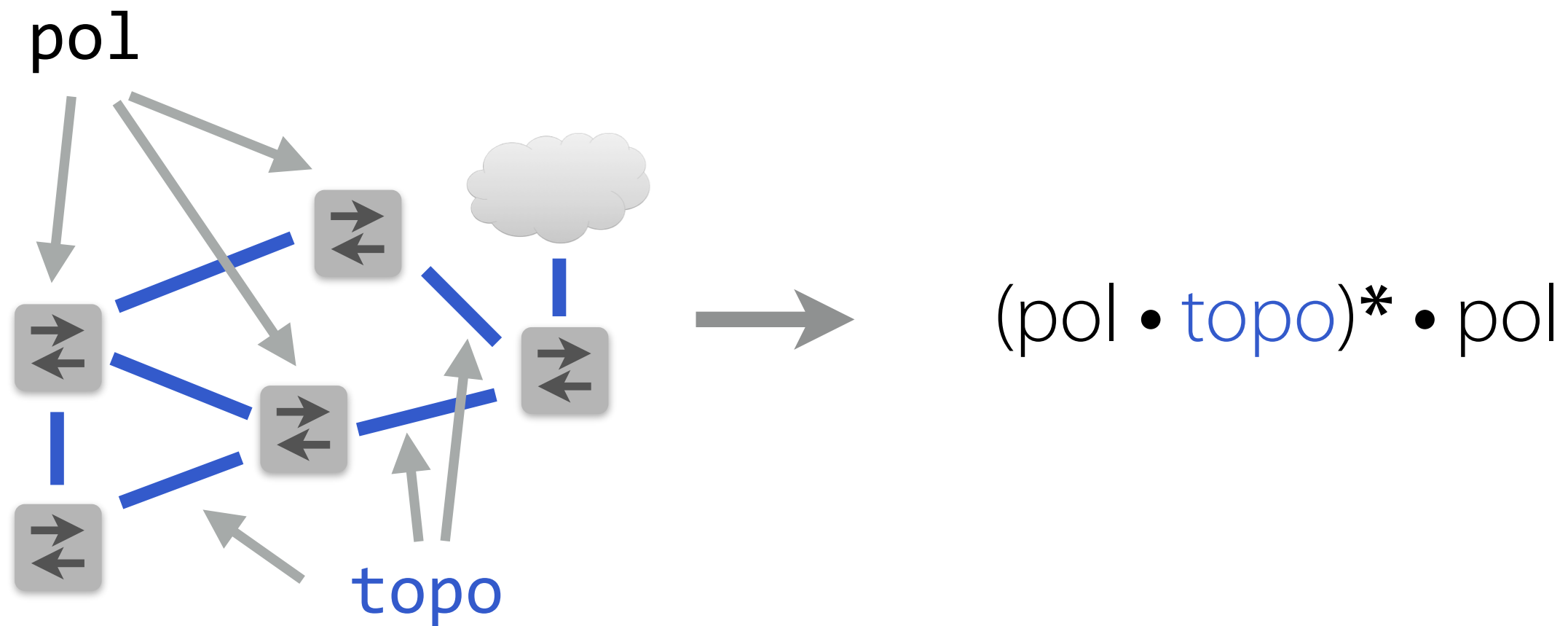
Encoding Networks

Switch forwarding tables and network topologies can be represented in NetKAT using straightforward encodings



Encoding Networks

A network can be encoded in NetKAT by interleaving steps of processing by switches and topology



NetKAT Semantics

NetKAT Semantics



Denotational

NetKAT Semantics



Denotational

Axiomatic

$\vdash p \equiv q$

NetKAT Semantics

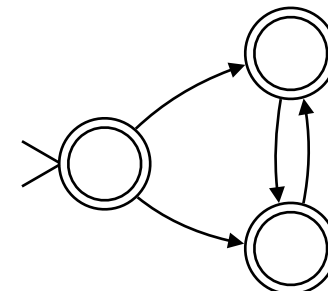


Denotational

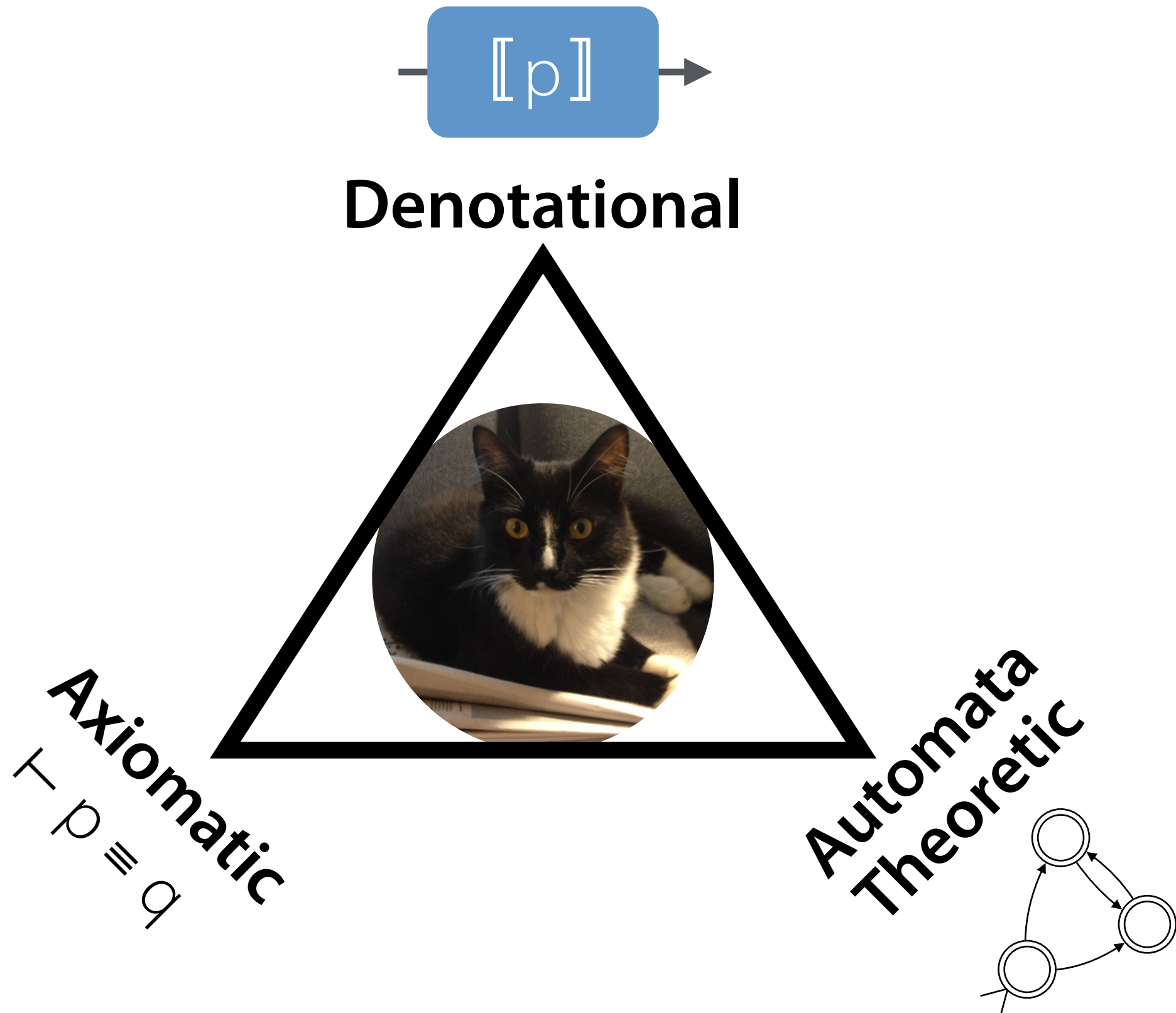
Axiomatic

$\vdash p \equiv q$

**Automata
Theoretic**



NetKAT Semantics

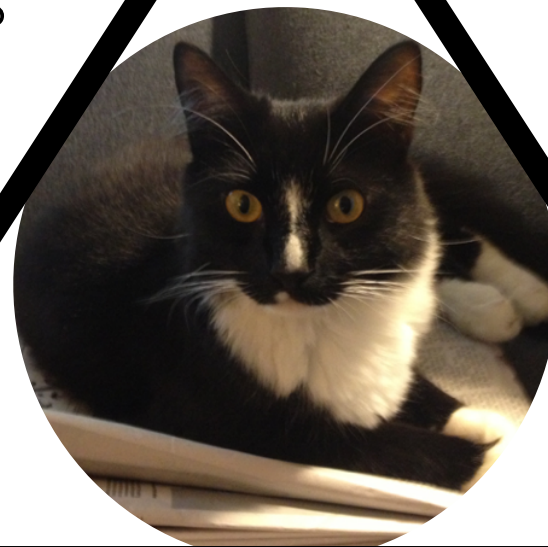


NetKAT Semantics



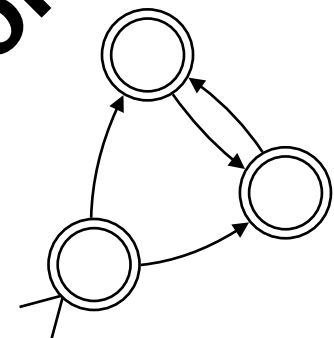
Denotational

Soundness+Completeness
[POPL'14]



Axiomatic
 $\vdash p \equiv q$

**Automata
Theoretic**



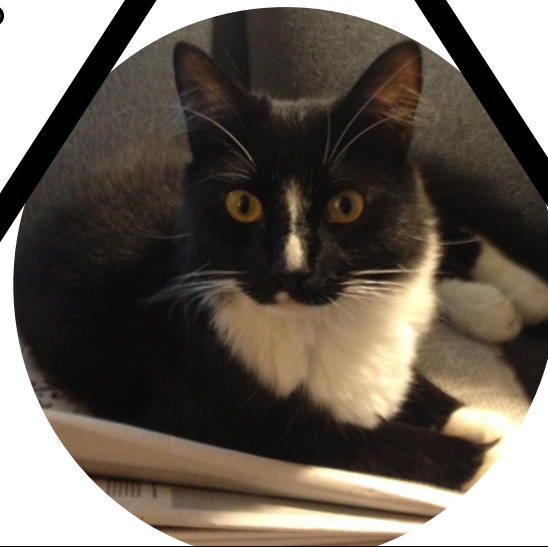
NetKAT Semantics



Denotational

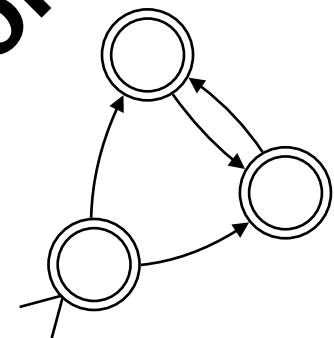
Soundness+Completeness
[POPL'14]

Kleene's Theorem
[POPL'15]



Axiomatic
 $\vdash p \equiv q$

**Automata
Theoretic**



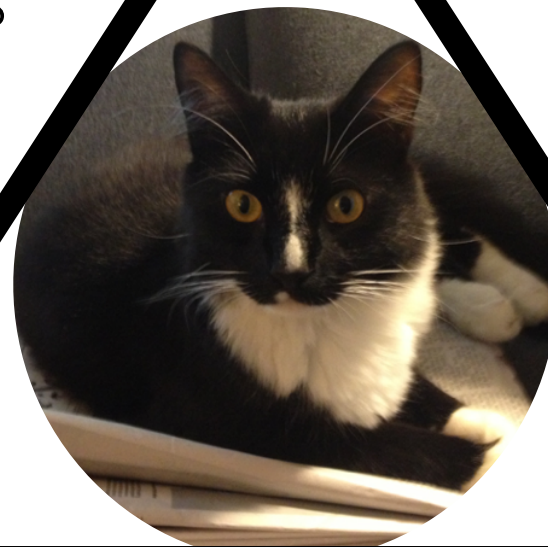
NetKAT Semantics



Denotational

Soundness+Completeness
[POPL '14]

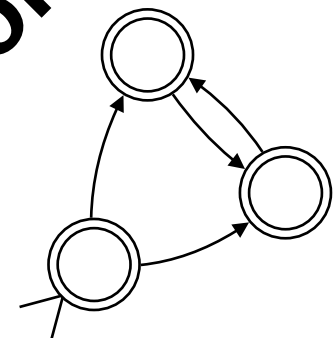
Kleene's Theorem
[POPL '15]



Axiomatic
 $\vdash p \equiv q$

Proof-Carrying
Code
[Ongoing Work]

**Automata
Theoretic**



Denotational Semantics

$\text{pol} ::=$

| **false**

| **true**

| $\text{field} = \text{val}$

| $\text{field} ::= \text{val}$

| $\text{pol}_1 + \text{pol}_2$

| $\text{pol}_1 \bullet \text{pol}_2$

| $!\text{pol}$

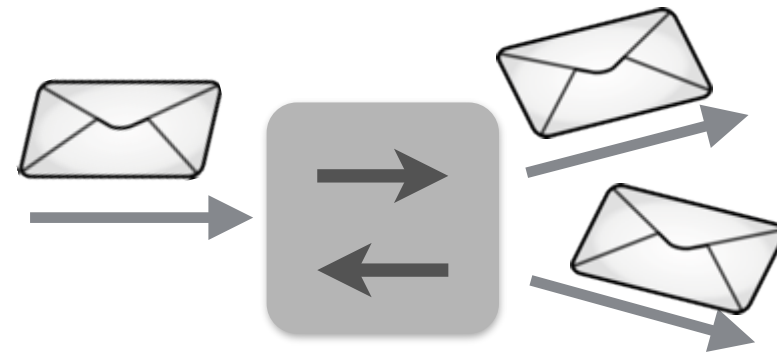
| pol^*

| **dup**

Denotational Semantics

```
pol ::=  
| false  
| true  
| field = val  
| field ::= val  
| pol1 + pol2  
| pol1 • pol2  
| !pol  
| pol*  
| dup
```

Local NetKAT: input-output behavior of switches



$\llbracket \text{pol} \rrbracket \in \text{Packet} \rightarrow \text{Packet Set}$

Denotational Semantics

$\text{pol} ::=$

| **false**

| **true**

| $\text{field} = \text{val}$

| $\text{field} := \text{val}$

| $\text{pol}_1 + \text{pol}_2$

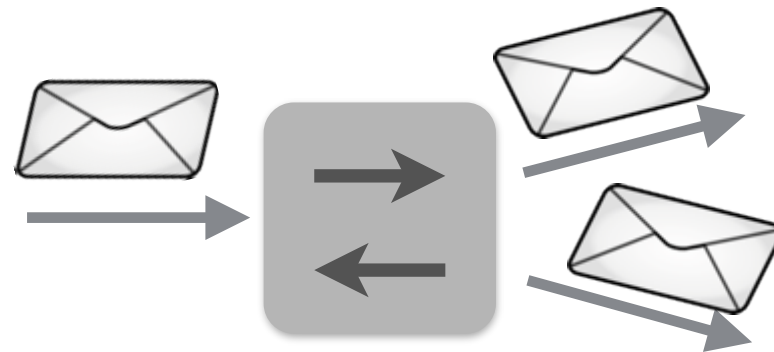
| $\text{pol}_1 \bullet \text{pol}_2$

| $!\text{pol}$

| pol^*

| **dup**

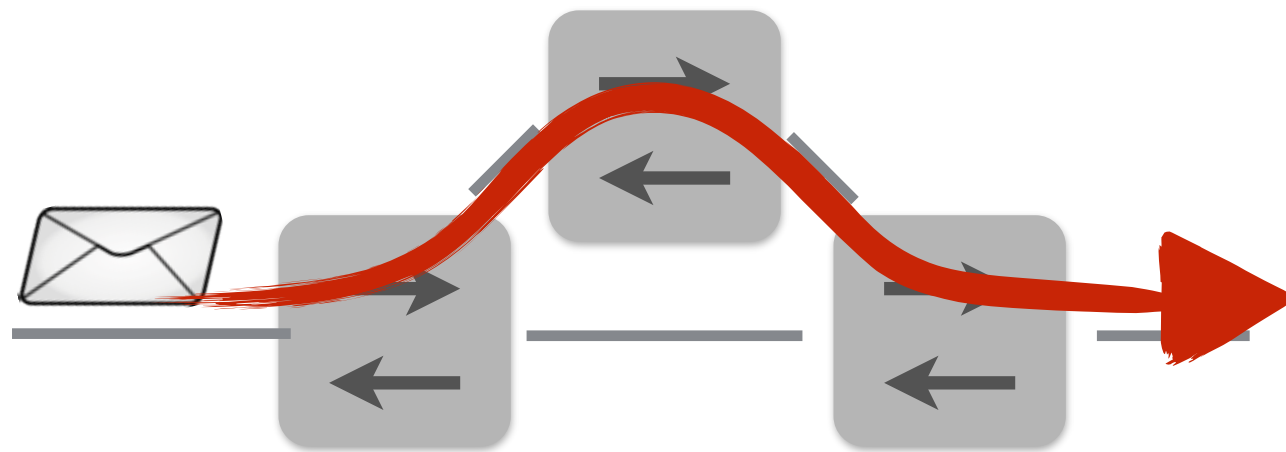
Local NetKAT: input-output behavior of switches



$\llbracket \text{pol} \rrbracket \in \text{Packet} \rightarrow \text{Packet Set}$

.....

Global NetKAT: network-wide paths



$\llbracket \text{pol} \rrbracket \in \text{History} \rightarrow \text{History Set}$

NetKAT Semantics

$\llbracket \text{pol} \rrbracket \in \mathbf{History} \rightarrow \mathbf{History Set}$

$\llbracket \mathbf{true} \rrbracket h = \{ h \}$

$\llbracket \mathbf{false} \rrbracket h = \{ \}$

$\llbracket f = n \rrbracket pk :: h = \begin{cases} \{ pk :: h \} & \text{if } pk.f = n \\ \{ \} & \text{otherwise} \end{cases}$

$\llbracket ! \text{pol} \rrbracket h = \{ h \} \setminus \llbracket \text{pol} \rrbracket$

$\llbracket f := n \rrbracket pk :: h = \{ pk[f:=n] :: h \}$

$\llbracket \text{pol}_1 + \text{pol}_2 \rrbracket h = \llbracket \text{pol}_1 \rrbracket h \cup \llbracket \text{pol}_2 \rrbracket h$

$\llbracket \text{pol}_1 \bullet \text{pol}_2 \rrbracket h = (\llbracket \text{pol}_1 \rrbracket \bullet \llbracket \text{pol}_2 \rrbracket) h$

$\llbracket \text{pol}^* \rrbracket h = (\bigcup_i \llbracket \text{pol} \rrbracket^i h)$

$\llbracket \mathbf{dup} \rrbracket pk :: h = \{ pk :: pk :: h \}$

$f, g \in \mathbf{History} \rightarrow \mathbf{History Set}$

$(f \bullet g) h = \bigcup \{ g h' \mid h' \in f h \}$

Axiomatic Semantics

NetKAT's design is based upon canonical structures:

- Regular operators ($+$, \cdot , and $*$) encode network paths
- Boolean operators ($+$, \cdot , and $!$) encode switch tables

The combination of a Boolean Algebra and a Kleene Algebra is called a ***Kleene Algebra with Tests (KAT)***

[Kozen '96]

KAT has an accompanying proof system for showing equivalences of the form $p \equiv q$

NetKAT Axioms

Kleene Algebra Axioms

$p + (q + r) \equiv (p + q) + r$
 $p + q \equiv q + p$
 $p + \mathbf{false} \equiv p$
 $p + p \equiv p$
 $p \cdot (q \cdot r) \equiv (p \cdot q) \cdot r$
 $p \cdot (q + r) \equiv p \cdot q + p \cdot r$
 $(p + q) \cdot r \equiv p \cdot r + q \cdot r$
 $\mathbf{true} \cdot p \equiv p$
 $p \equiv p \cdot \mathbf{true}$
 $\mathbf{false} \cdot p \equiv \mathbf{false}$
 $p \cdot \mathbf{false} \equiv \mathbf{false}$
 $\mathbf{true} + p \cdot p^* \equiv p^*$
 $\mathbf{true} + p^* \cdot p \equiv p^*$
 $p + q \cdot r + r \equiv r \Rightarrow p^* \cdot q + r \equiv r$
 $p + q \cdot r + q \equiv q \Rightarrow p \cdot r^* + q \equiv q$

Boolean Algebra Axioms

$a + (b \cdot c) \equiv (a + b) \cdot (a + c)$
 $a + \mathbf{true} \equiv \mathbf{true}$
 $a + !a \equiv \mathbf{true}$
 $a \cdot b \equiv b \cdot a$
 $a \cdot !a \equiv \mathbf{false}$
 $a \cdot a \equiv a$

Packet Axioms

$f := n \cdot f' := n' \equiv f' := n' \cdot f := n \quad \text{if } f \neq f'$
 $f := n \cdot f' = n' \equiv f' = n' \cdot f := n \quad \text{if } f \neq f'$
 $f := n \cdot f = n \equiv f := n$
 $f = n \cdot f := n \equiv f = n$
 $f := n \cdot f := n' \equiv f := n'$
 $f = n \cdot f = n' \equiv \mathbf{false} \quad \text{if } n \neq n'$
 $\text{dup} \cdot f = n \equiv f = n \cdot \text{dup}$
 $\Sigma_i f = n_i \equiv \mathbf{true}$

NetKAT Axioms

Kleene Algebra Axioms

$$p + (q + r) \equiv (p + q) + r$$

$$p + q \equiv q + p$$

$$p + \mathbf{false} \equiv p$$

$$p + p \equiv p$$

$$p \bullet (q \bullet r) \equiv (p \bullet q) \bullet r$$

$$p \bullet (q + r) \equiv (p \bullet q) + (p \bullet r)$$

$$(p + q) \bullet r \equiv (p \bullet r) + (q \bullet r)$$

$$\mathbf{true} \bullet p \equiv p$$

$$p \equiv p \bullet \mathbf{true}$$

$$\mathbf{false} \bullet p \equiv \mathbf{false}$$

$$p \bullet \mathbf{false} \equiv \mathbf{false}$$

$$\mathbf{true} + p \bullet p^* \equiv p^*$$

$$\mathbf{true} + p^* \bullet p \equiv p^*$$

$$p + q \bullet r + r \equiv r \Rightarrow p^* \bullet q + r \equiv r$$

$$p + q \bullet r + q \equiv q \Rightarrow p \bullet r^* + q \equiv q$$

Boolean Algebra Axioms

$$a + (b \bullet c) \equiv (a + b) \bullet (a + c)$$

$$a + \mathbf{true} \equiv \mathbf{true}$$

$$a + !a \equiv \mathbf{true}$$

$$a \bullet b = b \bullet a$$

Soundness: If $\vdash p \equiv q$, then $\llbracket p \rrbracket = \llbracket q \rrbracket$

Completeness: If $\llbracket p \rrbracket = \llbracket q \rrbracket$, then $\vdash p \equiv q$

$$f := n \bullet f' = n' \equiv f' = n' \bullet f := n$$

if $f \neq f'$

if $f \neq f'$

$$f := n \bullet f = n \equiv f := n$$

$$f = n \bullet f := n \equiv f = n$$

$$f := n \bullet f := n' \equiv f := n'$$

$$f = n \bullet f = n' \equiv \mathbf{false}$$

if $n \neq n'$

$$\text{dup} \bullet f = n \equiv f = n \bullet \text{dup}$$

$$\sum_i f = n_i \equiv \mathbf{true}$$

NetKAT Automata

A *NetKAT automaton* $M = (S, s_0, \varepsilon, \delta)$ is a tuple where:

- S is a finite set of states,
- $s_0 \in S$ is the start state,
- $\varepsilon \in S \rightarrow \text{Packet} \rightarrow \text{Packet Set}$ is the “observation” function
- $\delta \in S \rightarrow \text{Packet} \rightarrow (\text{State} * \text{Packet}) \text{ Set}$ is the “continuation” function

NetKAT Automata

A *NetKAT automaton* $M = (S, s_0, \varepsilon, \delta)$ is a tuple where:

- S is a finite set of states,
- $s_0 \in S$ is the start state,
- $\varepsilon \in S \rightarrow \text{Packet} \rightarrow \text{Packet Set}$ is the “observation” function
- $\delta \in S \rightarrow \text{Packet} \rightarrow (\text{State} * \text{Packet}) \text{ Set}$ is the “continuation” function

Inputs: $\text{pkt}_{\text{in}} \cdot \text{pkt}_1 \cdot \mathbf{dup} \cdot \dots \cdot \mathbf{dup} \cdot \text{pkt}_n \mathbf{dup} \cdot \text{pkt}_{\text{out}}$

NetKAT Automata

A NetKAT automaton $M = (S, s_0, \varepsilon, \delta)$ is a tuple where:

- S is a finite set of states,
- $s_0 \in S$ is the start state,
- $\varepsilon \in S \rightarrow \text{Packet} \rightarrow \text{Packet Set}$ is the “observation” function
- $\delta \in S \rightarrow \text{Packet} \rightarrow (\text{State} * \text{Packet}) \text{ Set}$ is the “continuation” function


Inputs: $\text{pkt}_{\text{in}} \cdot \text{pkt}_1 \cdot \mathbf{dup} \cdot \dots \cdot \mathbf{dup} \cdot \text{pkt}_n \mathbf{dup} \cdot \text{pkt}_{\text{out}}$

A NetKAT automaton *accepts* an input in state s if:

- $\text{accept } s (\text{pkt}_{\text{in}} \cdot \text{pkt}_{\text{out}}) \Leftrightarrow \text{pkt}_{\text{out}} \in \varepsilon s \text{ pat}_{\text{in}}$
- $\text{accept } s (\text{pkt}_{\text{in}} \cdot \text{pkt}_1 \cdot \mathbf{dup} \cdot w) \Leftrightarrow$
 $\exists s'. (\text{pkt}_1, s') \in \delta s \text{ pkt}_{\text{in}} \text{ and } \text{accept } s' (\text{pkt}_1 \cdot w)$

NetKAT Derivatives

*Labels, one per
occurrence of **dup***



$$E(\text{pol}) \in \text{Pol}$$

$$E(\mathbf{false}) = \mathbf{false}$$

$$E(\mathbf{true}) = \mathbf{true}$$

$$E(f = n) = f = n$$

$$E(f := n) = f := n$$

$$E(!\text{pol}) = !\text{pol}$$

$$E(\text{dup}^l) = \mathbf{false}$$

$$E(\text{pol}_1 + \text{pol}_2) = E(\text{pol}_1) + E(\text{pol}_2)$$

$$E(\text{pol}_1 \bullet \text{pol}_2) = E(\text{pol}_1) \bullet E(\text{pol}_2)$$

$$E(\text{pol}^*) = E(\text{pol})^*$$

$$D(\text{pol}) \in (\text{Pol} * L * \text{Pol}) \text{ Set}$$

$$D(\mathbf{false}) = \{\}$$

$$D(\mathbf{true}) = \{\}$$

$$D(f = n) = \{\}$$

$$D(f := n) = \{\}$$

$$D(!\text{pol}) = \{\}$$

$$D(\text{dup}^l) = \{ (\mathbf{true}, l, \mathbf{true}) \}$$

$$D(\text{pol}_1 + \text{pol}_2) = D(\text{pol}_1) + D(\text{pol}_2)$$

$$D(\text{pol}_1 \bullet \text{pol}_2) = D(\text{pol}_1) \bullet \text{pol}_2 + \\ E(\text{pol}_1) \bullet D(\text{pol}_2)$$

$$D(\text{pol}^*) = E(\text{pol})^* \bullet D(\text{pol}) \bullet \text{pol}^*$$

NetKAT Automata

We can build an automaton using derivatives as follows:

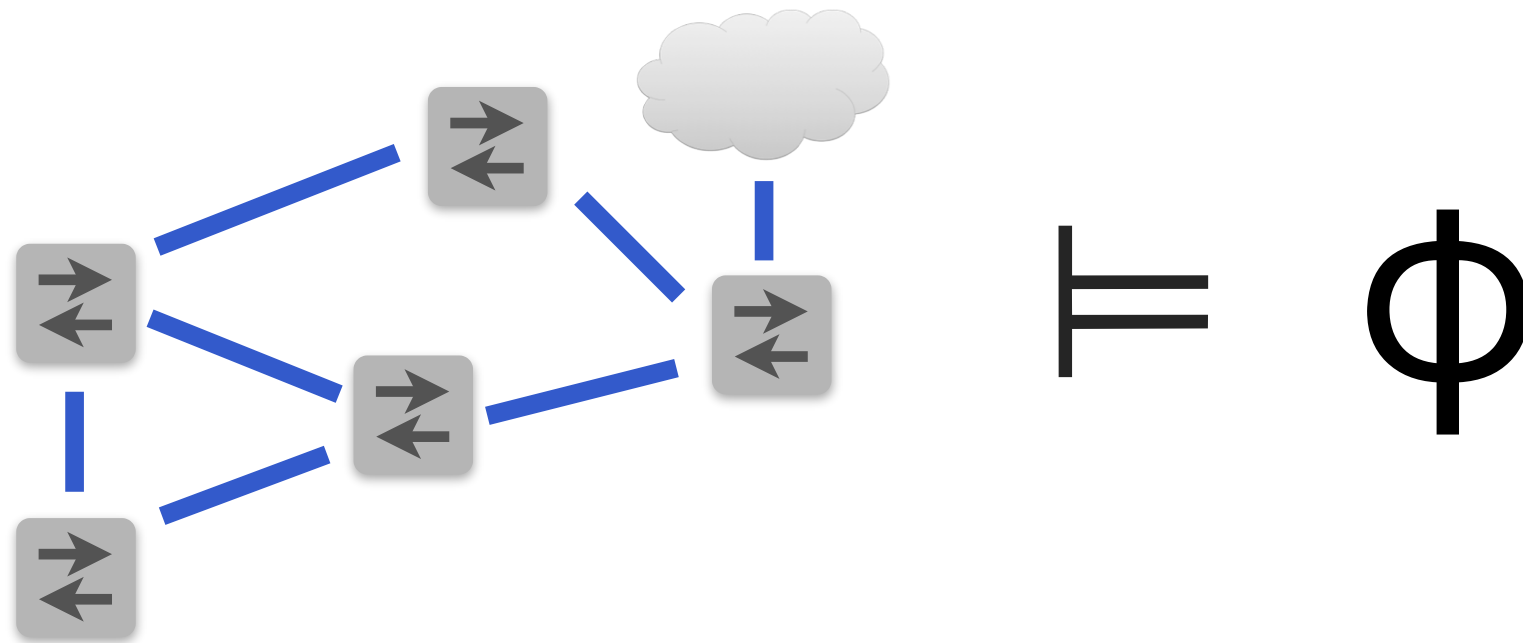
- S is the set of labels in pol , plus a fresh start state 0 ,
- $\varepsilon \mid \text{pkt} = \{\text{pkt}' \mid \langle \text{pkt}' \rangle \in \llbracket E(k_l) \rrbracket \langle \text{pkt} \rangle\}$
- $\delta \mid \text{pkt} = \{(\text{pkt}', l') \mid (d, l', k) \in \llbracket D(k_l) \rrbracket \wedge \langle \text{pkt}' \rangle \in \llbracket d \rrbracket \langle \text{pkt} \rangle\}$

Notation: k_l denotes the unique continuation of dup^l

Of course, in practice, it is important to use symbolic representations (e.g., FDDs [ICFP '15]) to keep the size of the automata manageable.

Applications

Reachability [POPL '14]



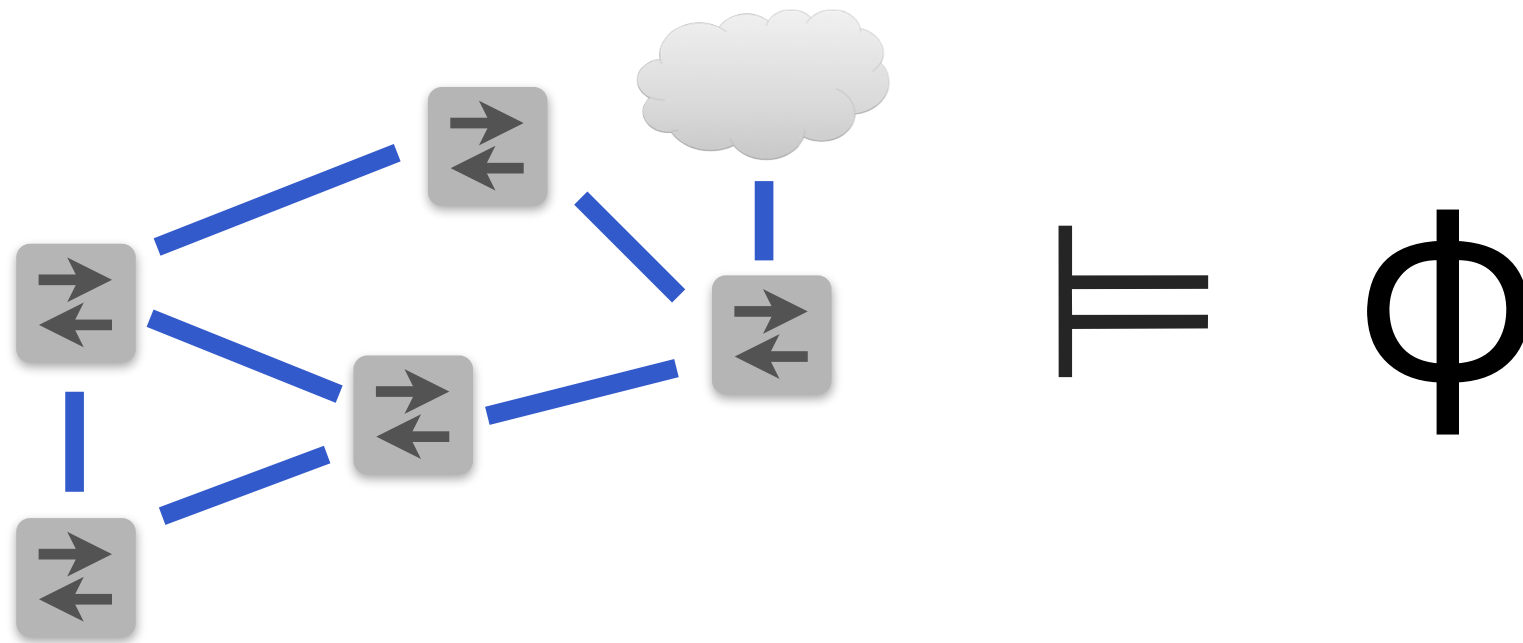
We'd like to be able to answer questions like:

“Does the network forward from ingress to egress?”

Can reduce this question (and others) to (in)equivalence

$\text{in} \bullet (\text{pol} \bullet \text{topo})^* \bullet \text{pol} \bullet \text{out} \neq \mathbf{false}$

Loop Freedom [POPL '15]



Can use automata to check if a network is loop free

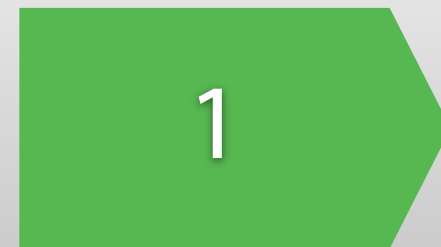
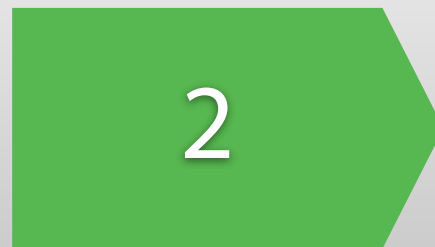
Intuition: $\forall a. \text{in} \bullet (p \bullet t)^* \bullet a \bullet (p \bullet t)^+ \bullet a \equiv \mathbf{false}$

- $\forall \text{pkt}, \text{pkt}'. \text{pkt}' \in \llbracket E(\Phi(\text{in} \bullet (p \bullet t)^*)) \rrbracket \text{pkt}$
- Check whether $\text{pkt}' \in \llbracket E(\Phi(p \bullet t)^+) \rrbracket \text{pkt}'$

Notation: $\Phi(p)$ denotes replacing **dup** with **true**

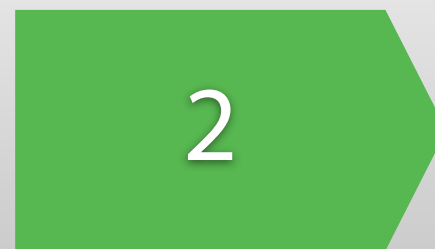
Compilation [ICFP '15]

NetKAT Compiler Pipeline



Compilation [ICFP '15]

NetKAT Compiler Pipeline



local
policy



| Pattern | Actions |
|---------|---------|
| dstpt=2 | drop |
| srcpt=7 | fwd 1 |
| * | fwd 2 |



~ 100x faster
than competitors

Compilation [ICFP '15]

NetKAT Compiler Pipeline



global
policy

**Global
Compiler**

local
policy

**Local
Compiler**

| Pattern | Actions |
|---------|---------|
| dstpt=2 | drop |
| srcpt=7 | fwd 1 |
| * | fwd 2 |



network-wide
behavior



~ 100x faster
than competitors

Compilation [ICFP '15]

NetKAT Compiler Pipeline

virtual
policy

**Virtual
Compiler**

global
policy

**Global
Compiler**

local
policy

**Local
Compiler**

| Pattern | Actions |
|---------|---------|
| dstpt=2 | drop |
| srcpt=7 | fwd 1 |
| * | fwd 2 |

abstract
topologies

network-wide
behavior

~ 100x faster
than competitors

Compilation [ICFP '15]

NetKAT Compiler Pipeline

virtual
policy

**Virtual
Compiler**

global
policy

**Global
Compiler**

local
policy

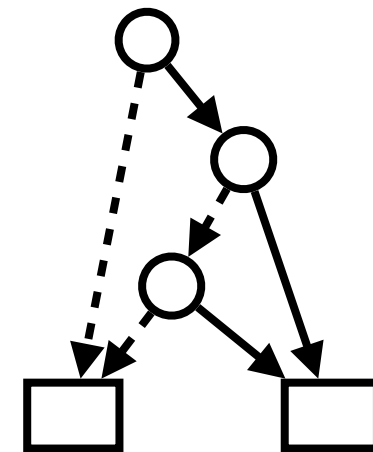
**Local
Compiler**

| Pattern | Actions |
|---------|---------|
| dstpt=2 | drop |
| srcpt=7 | fwd 1 |
| * | fwd 2 |

abstract
topologies

network-wide
behavior

~ 100x faster
than competitors



Compilation [ICFP '15]

NetKAT Compiler Pipeline

virtual
policy

**Virtual
Compiler**

global
policy

**Global
Compiler**

local
policy

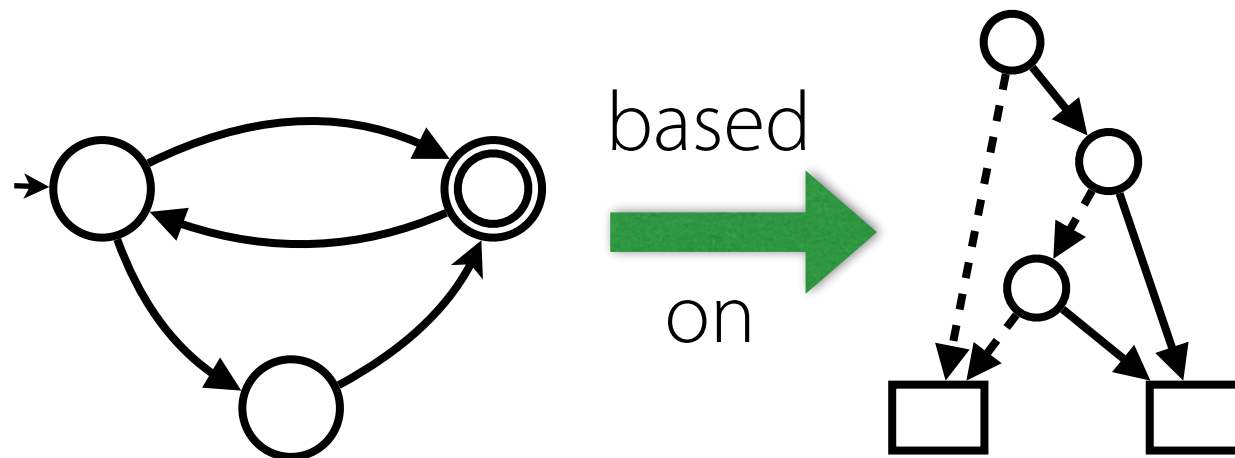
**Local
Compiler**

| Pattern | Actions |
|---------|---------|
| dstpt=2 | drop |
| srcpt=7 | fwd 1 |
| * | fwd 2 |

abstract
topologies

network-wide
behavior

~ 100x faster
than competitors



Compilation [ICFP '15]

NetKAT Compiler Pipeline

virtual
policy

**Virtual
Compiler**

global
policy

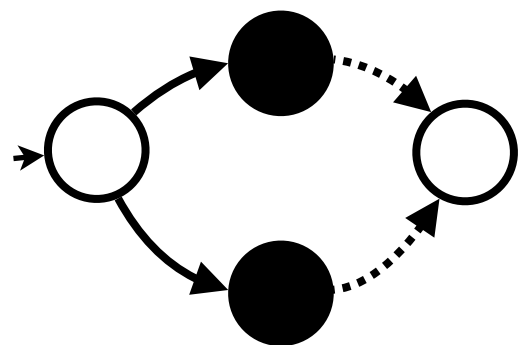
**Global
Compiler**

local
policy

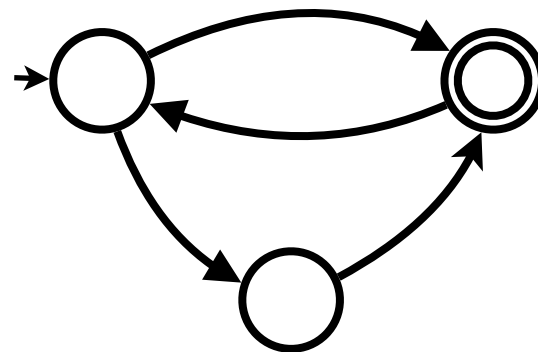
**Local
Compiler**

| Pattern | Actions |
|---------|---------|
| dstpt=2 | drop |
| srcpt=7 | fwd 1 |
| * | fwd 2 |

abstract
topologies

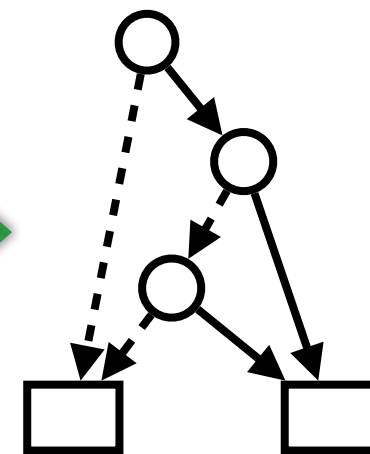


network-wide
behavior

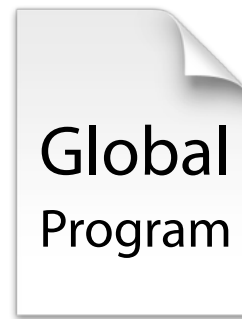


~ 100x faster
than competitors

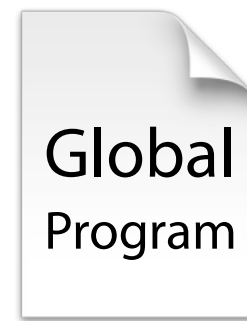
based
on



Global Compilation

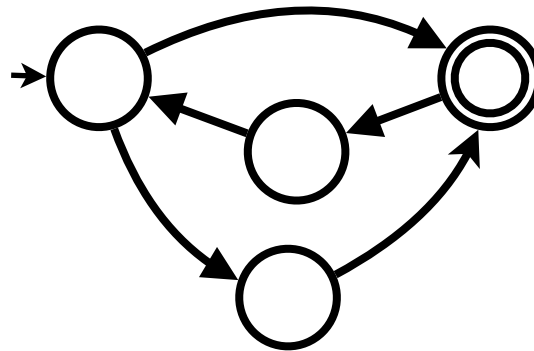


Global Compilation

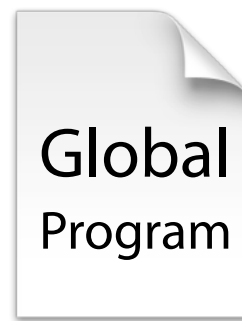


Adding Extra State
= Translation to Automaton

NetKAT NFA

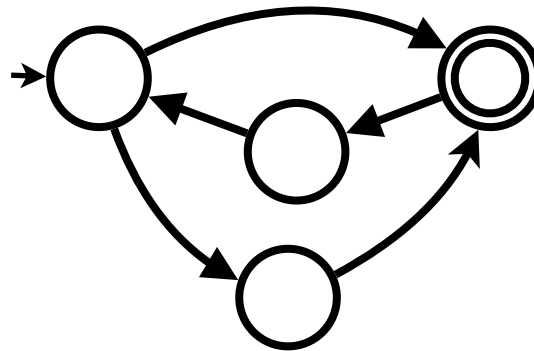


Global Compilation



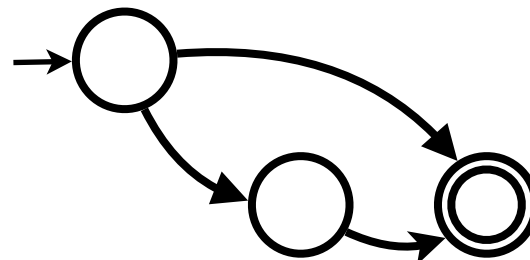
Adding Extra State
= Translation to Automaton

NetKAT NFA

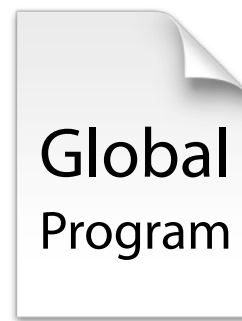


Avoiding Duplication
= Determinization

NetKAT DFA

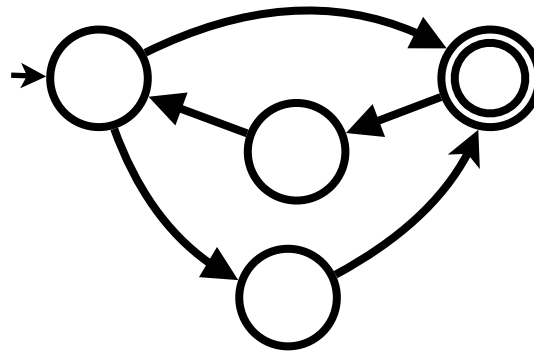


Global Compilation



Adding Extra State
= Translation to Automaton

NetKAT NFA

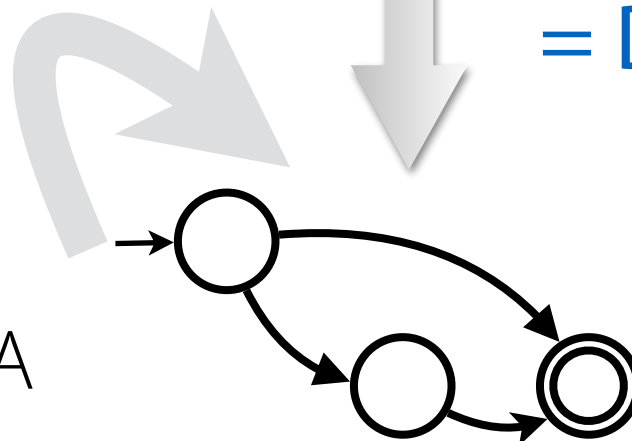


Automaton Minimization
= Tag Elimination

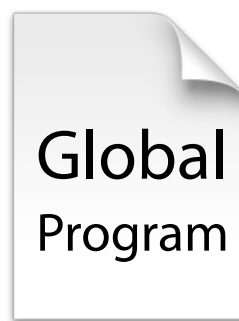


Avoiding Duplication
= Determinization

NetKAT DFA

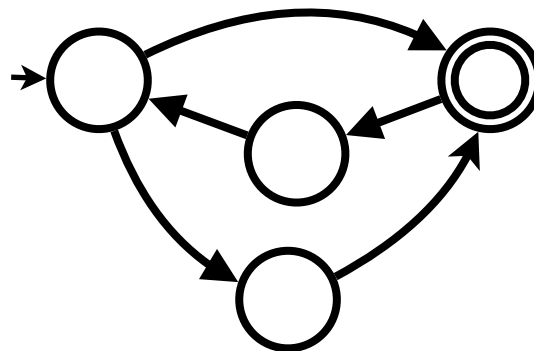


Global Compilation



Adding Extra State
= Translation to Automaton

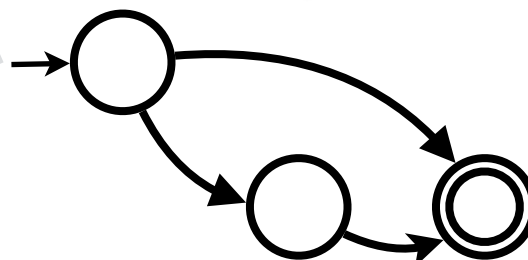
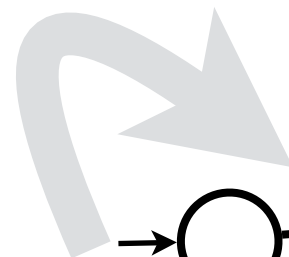
NetKAT NFA



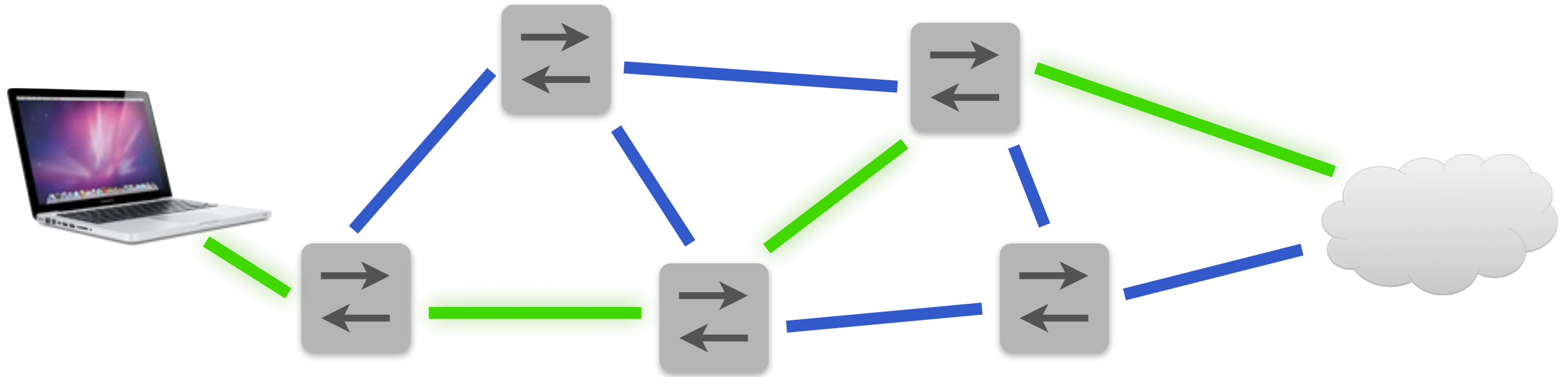
Automaton Minimization
= Tag Elimination

Avoiding Duplication
= Determinization

NetKAT DFA

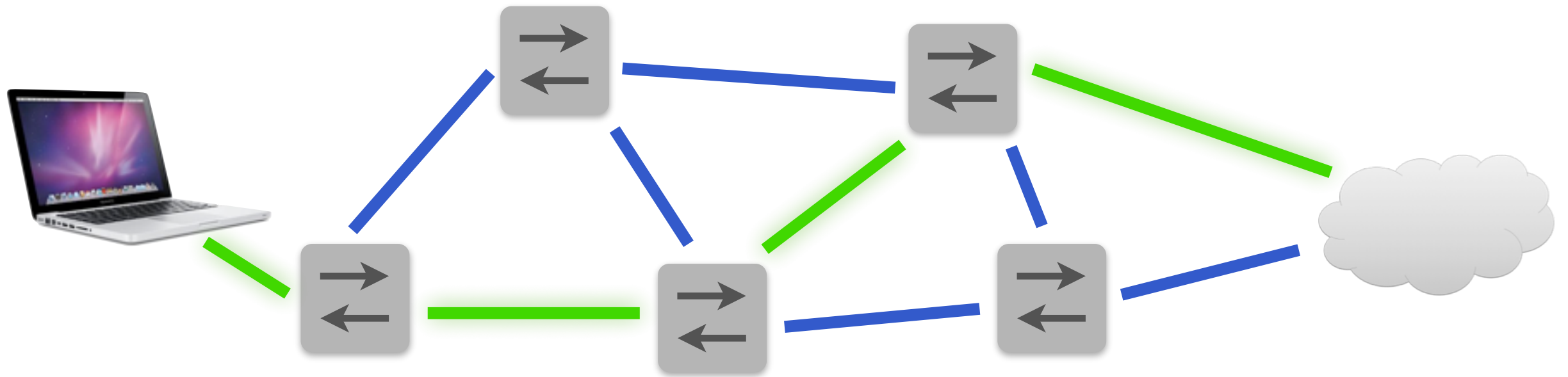


Measurement [SOSR '16]



Can implement path queries at the network edge!

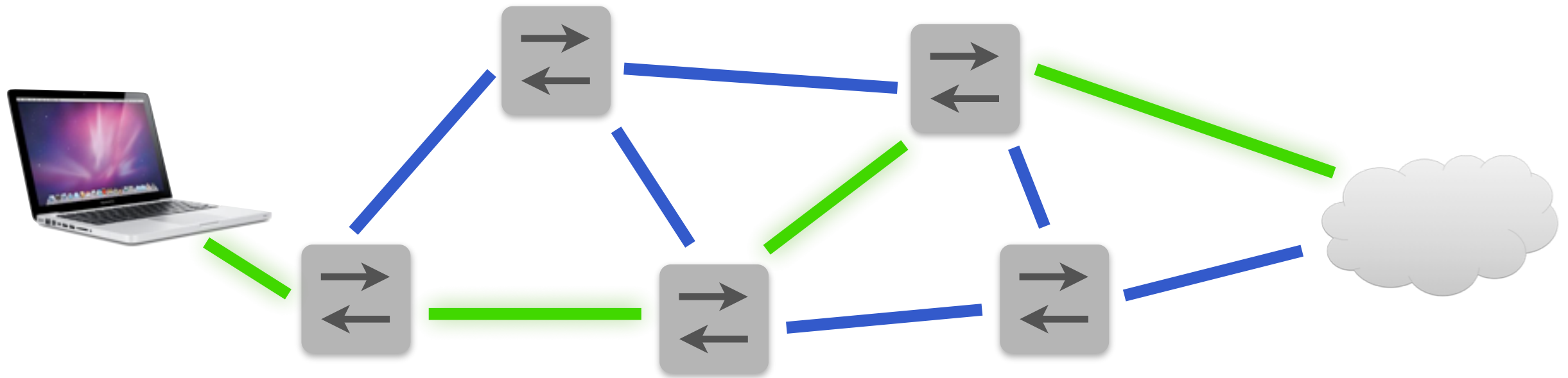
Measurement [SOSR '16]



Can implement path queries at the network edge!

- Combine policy p and query q via a simple translation

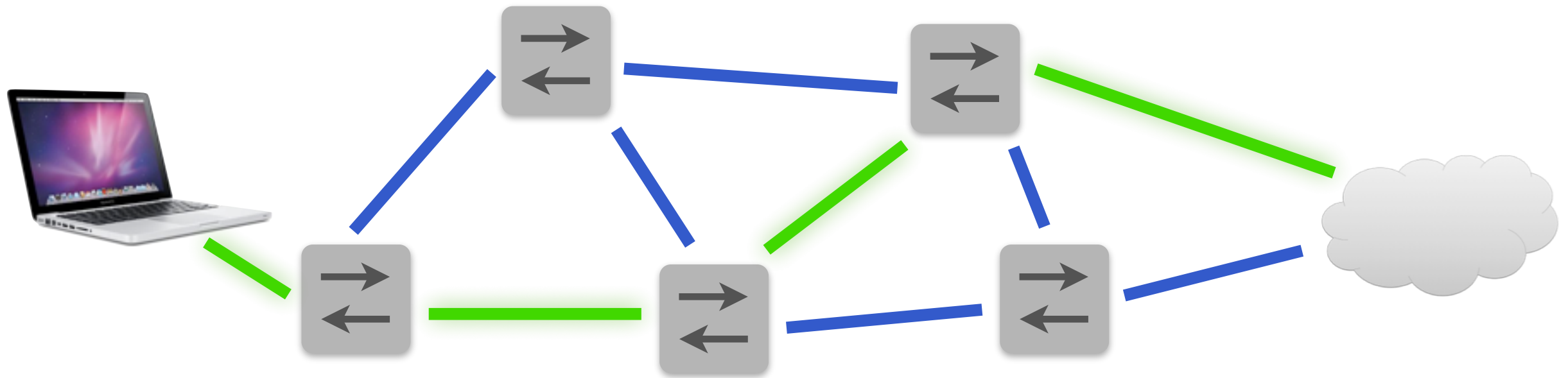
Measurement [SOSR '16]



Can implement path queries at the network edge!

- Combine policy p and query q via a simple translation
- Use Φ to replace all occurrences of **dup** with **true**

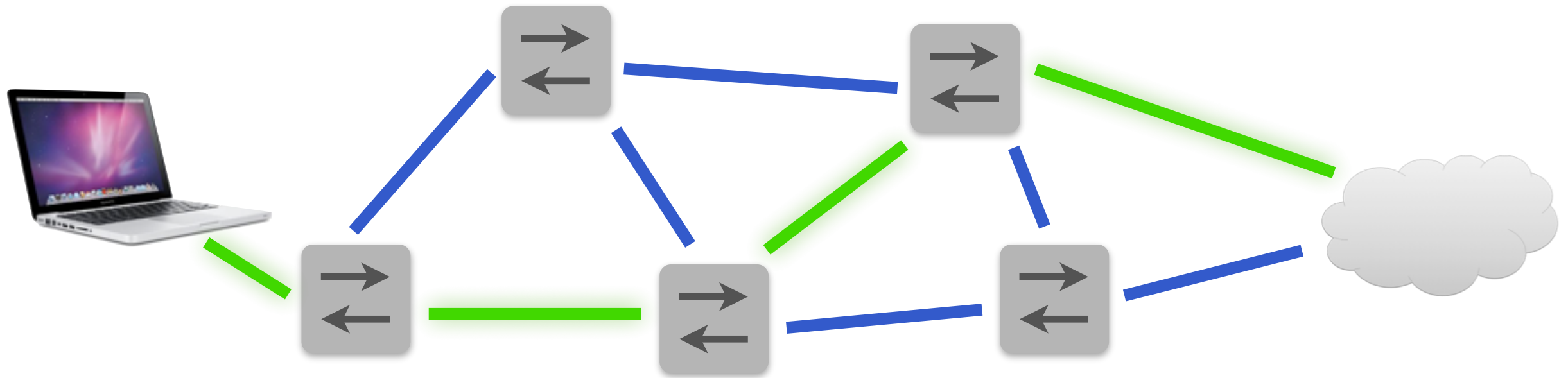
Measurement [SOSR '16]



Can implement path queries at the network edge!

- Combine policy p and query q via a simple translation
- Use Φ to replace all occurrences of **dup** with **true**
- Compute predicates from observation map E

Measurement [SOSR '16]



Can implement path queries at the network edge!

- Combine policy p and query q via a simple translation
- Use Φ to replace all occurrences of **dup** with **true**
- Compute predicates from observation map E
- Tabulate packets matching predicates on end hosts

Wrapping Up...

Ongoing Work

Stateful NetKAT [PLDI '16]

- Enriches the language with new features for programming stateful data planes
- Semantics is based on causal consistency and Winskel's "event structures"

Probabilistic NetKAT [ESOP '16]

- Enriches the language with random choice
- Can be used to model uncertainty about demands and failures, and randomized algorithms (e.g., ECMP)
- Semantics is based on Markov Kernels

Conclusion

- NetKAT offers a rich foundation for network programming
- Key features include boolean predicates, regular paths, and modular composition operators
- Semantics has been extensively studied and offers powerful mathematical tools for transforming and reasoning about programs, as well as guidance when designing extensions
- Many practical applications can be built using NetKAT including verification tools, compilers, and analysis tools

Reading

- Carolyn Jane Anderson, Nate Foster, Arjun Guha, Jean-Baptiste Jeannin, Dexter Kozen, Cole Schlesinger, and David Walker.
NetKAT: Semantic Foundations for Networks. In *ACM SIGPLAN--SIGACT Symposium on Principles of Programming Languages (POPL)*, January 2014.
- Nate Foster, Dexter Kozen, Matthew Milano, Alexandra Silva, and Laure Thompson. **A Coalgebraic Decision Procedure for NetKAT.** In *ACM SIGPLAN--SIGACT Symposium on Principles of Programming Languages (POPL)*, January 2015.
- Steffen Smolka, Spiridon Eliopoulos, Nate Foster, and Arjun Guha.
A Fast Compiler for NetKAT. In *ACM SIGPLAN International Conference on Functional Programming (ICFP)*, September 2015.

Collaborators

- Carolyn Anderson (UMass)
- Spiros Eliopoulos (Inhabited Type)
- **Arjun Guha** (UMass)
- Jean-Baptiste Jeannin (Samsung Labs)
- **Dexter Kozen** (Cornell)
- Matthew Milano (Cornell)
- Mark Reitblatt (Facebook)
- Cole Schlesinger (Samsung Labs)
- **Alexandra Silva** (UCL)
- **Steffen Smolka** (Cornell)
- Laure Thompson (Cornell)
- David Walker (Princeton)

Questions?



Questions?



<http://github.com/frenetic-lang/frenetic/>

frenetic >>



Pyretic

SDX

FUJITSU