Extending SDN to Large-Scale Networks

VMware
Martin Casado, Teemu Koponen

UC Berkeley
Murphy McCauley, Aurojit Panda, Scott Shenker

The Problem

Policy / Behavior
(Routing, ACLs, TE, etc.)

OpenFlow
Controller ??

**Geographically distributed
10,000 or 100,000 node
Physical Network**
(Switches, links, etc.)

Given a desired policy or behavior, how do you realize it on the physical network?
really large physical network?

Possible Approaches

- Beefy Central Controller
  - Eventually breaks down
  - Bad failure locality
  - Unclear where local policy fits

- Shard and Distribute
  - Requires independent shards

- Fully Distribute
  - Every new problem requires new distributed algorithm
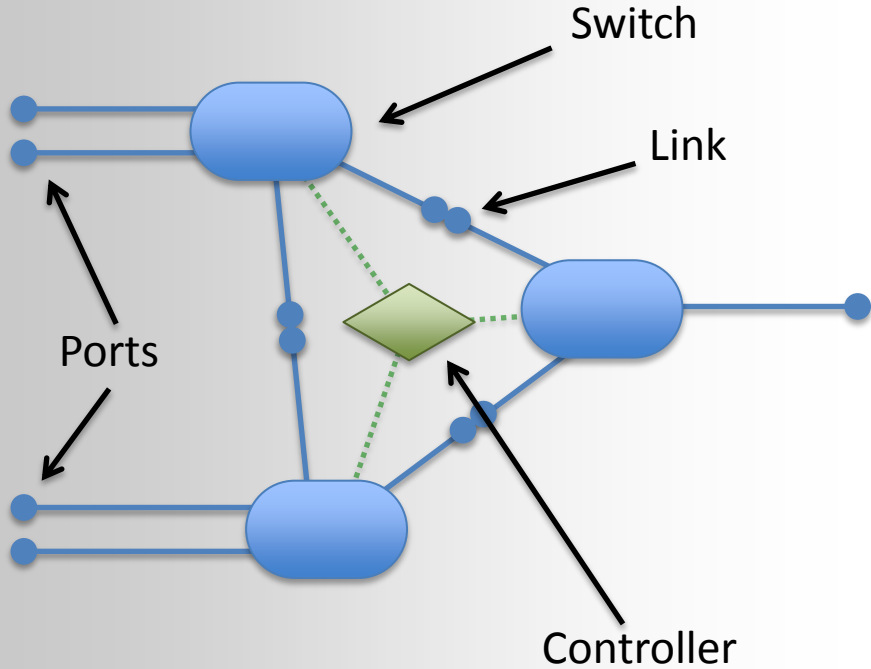 - Abandons benefits of SDN!

Where does that leave us?

- None of the standard approaches allow SDN to scale to arbitrarily large networks

- Combine classic network scaling approaches with SDN in a new way:
  - Utilize aggregation and hierarchy
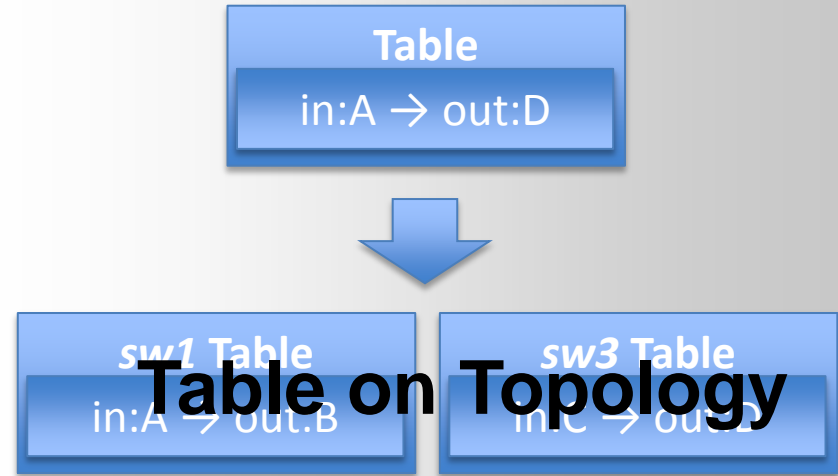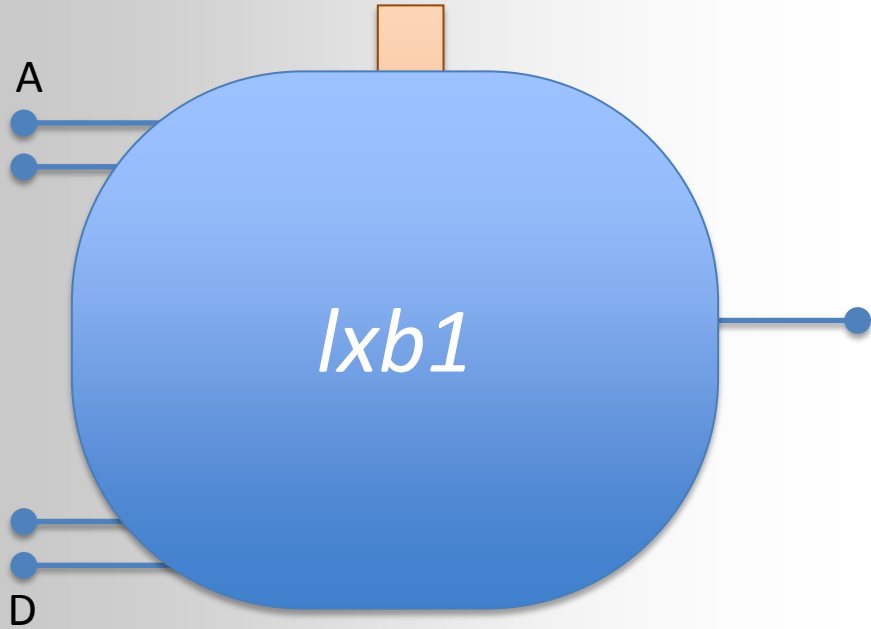 - Structure using a recursive abstraction

Old Techniques in New Setting

- Design built around Logical xBars (LXBs)
  - A programmable unit of forwarding ("switch")
  - Implemented in terms of itself

- Results in design with:
  - Scalability to extremely large networks
  - Straightforward convergence/failure behavior
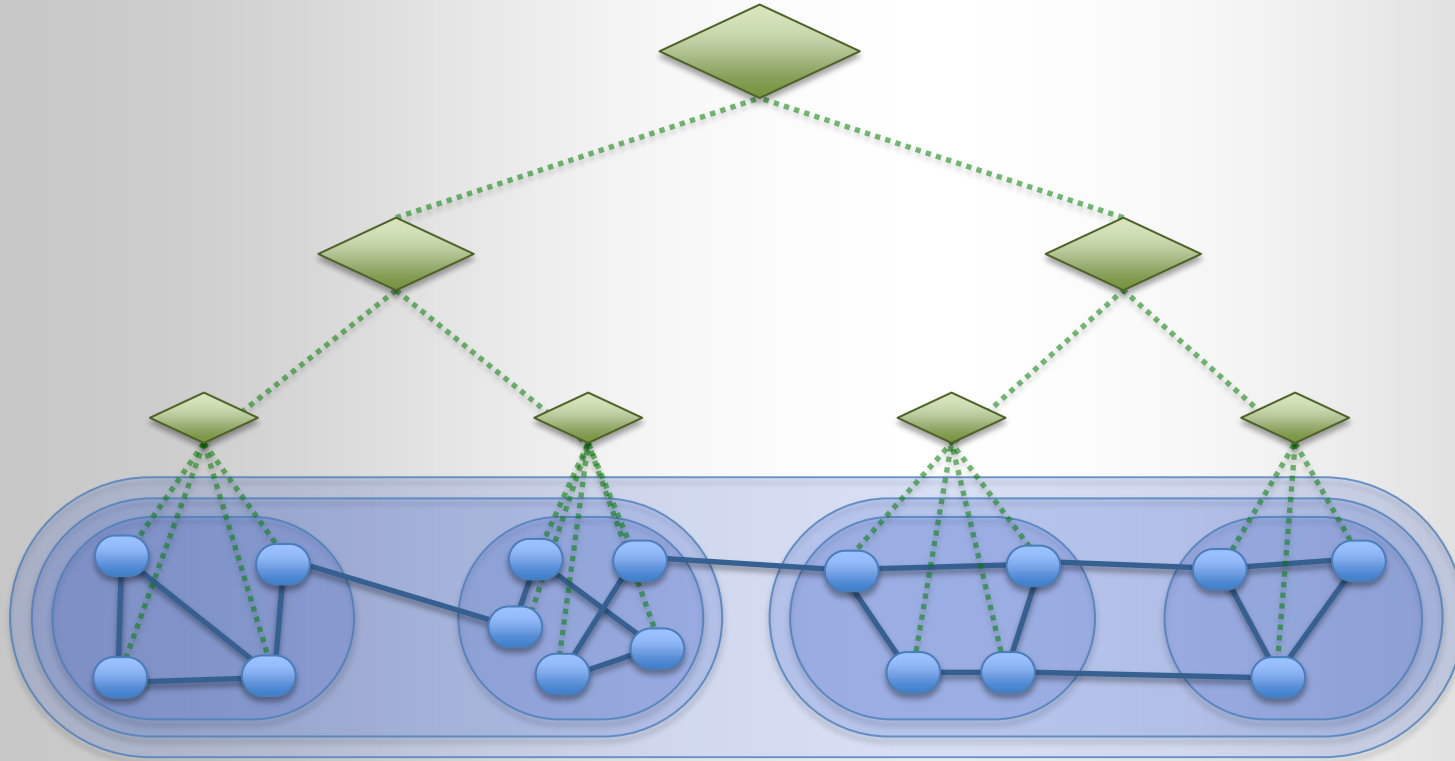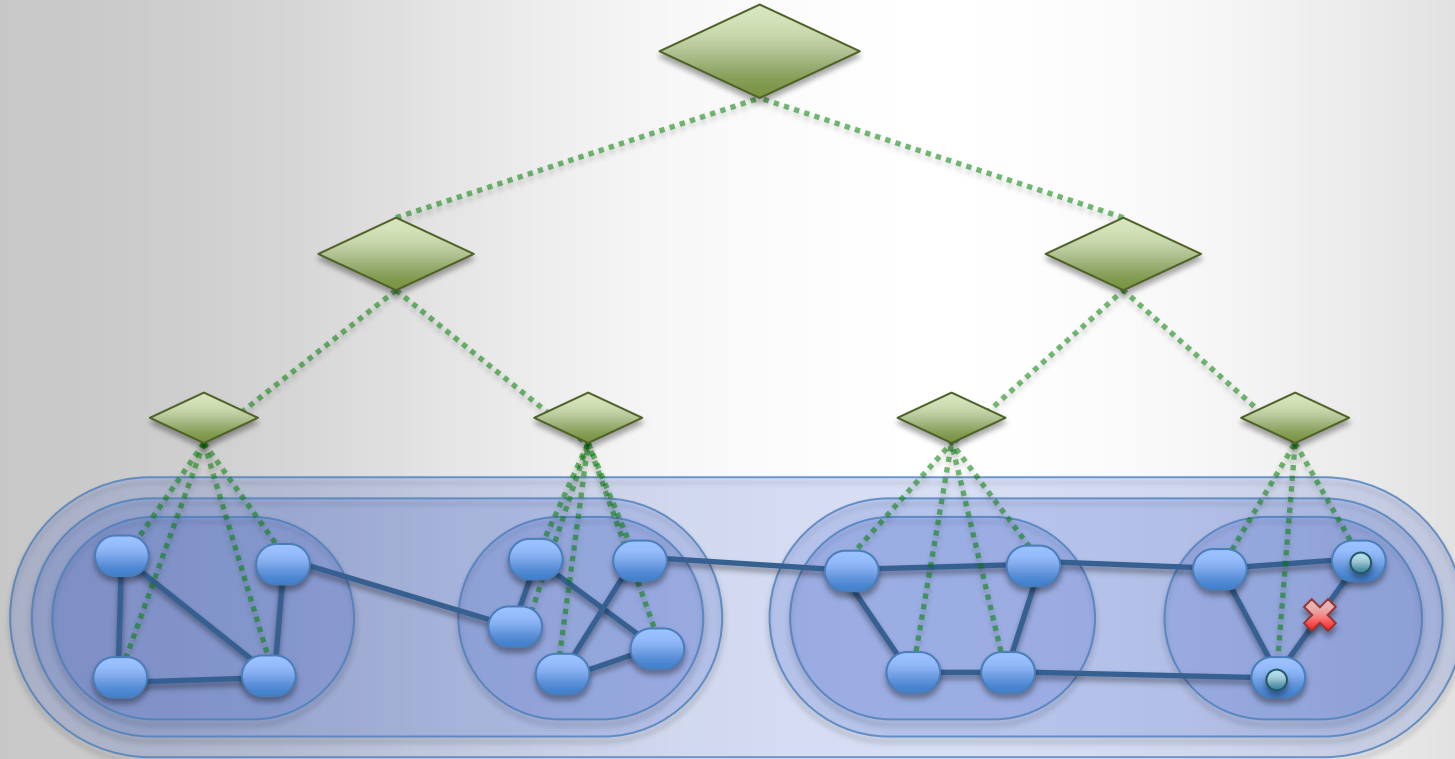 - Transactional network changes

xBar Basics

Switch

Link

Ports

Controller

xBar Basics

A

D

lxb1

Table
in:A → out:D

↓

sw1 Table
in:A → out:B

sw3 Table
in:C → out:D
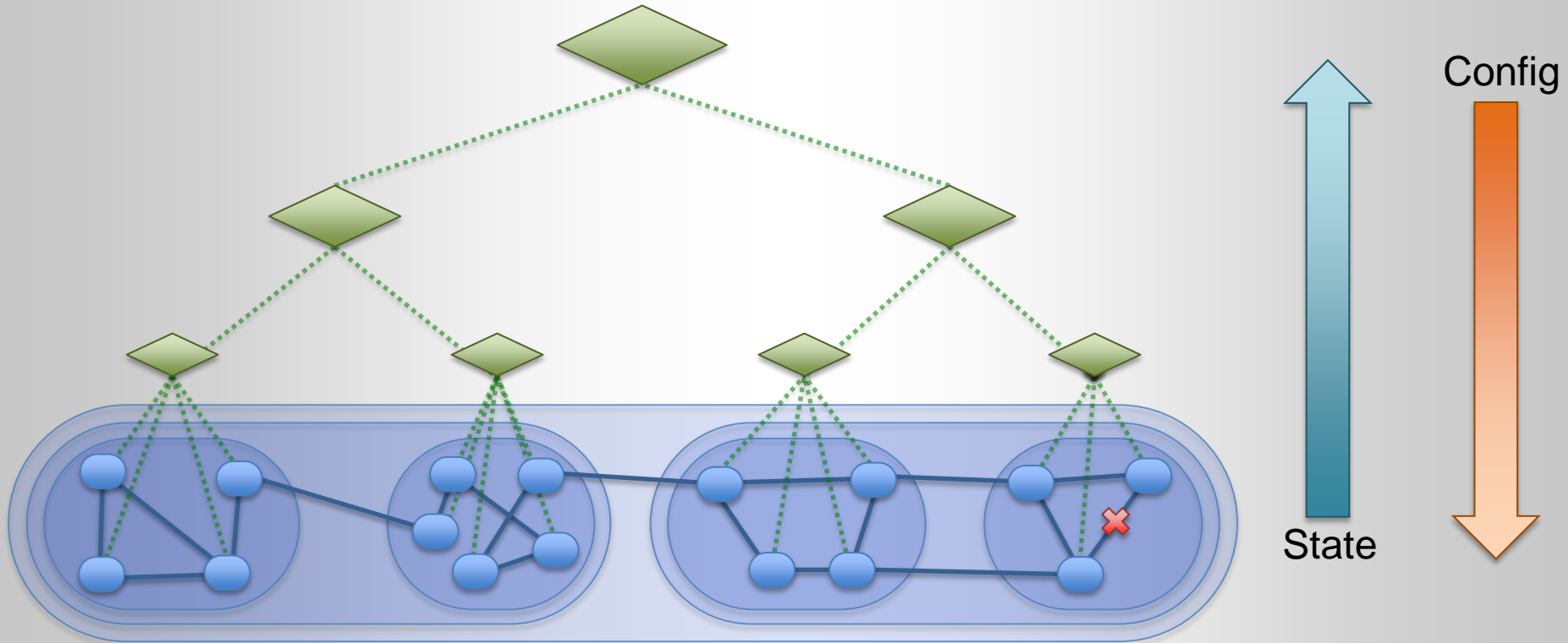
Table on Topology

xBar Basics

lxb3

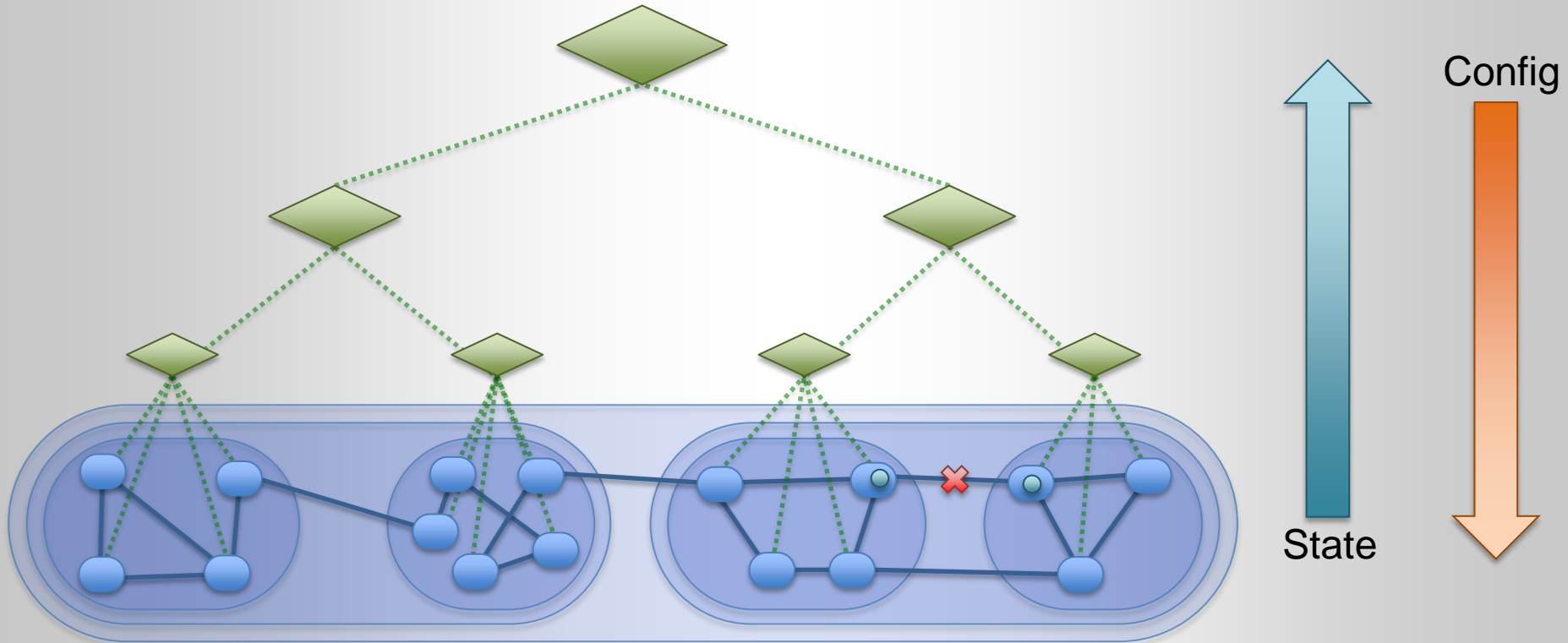# Hierarchy & Recursion

Hierarchy & Recursion

Hierarchy & Recursion

Hierarchy & Recursion

LXB Properties

- Hierarchy cleanly stacks arbitrarily high

  – Can match LXBs to physical/organizational boundaries

  – You can express local policies within "your" LXB as long as the externally-visible policies conform to those given by your parent

LXB Properties

- Computation kept in check

  - 10,000 nodes in four tiers: individual LXBs only have 10 node graphs

  - Multiple levels of aggregation means that computations can be kept towards lower (most distributed) portion of controller hierarchy

LXB Properties

- "SDN-Friendly"

  - Agnostic to choice of switch control mechanism (e.g., OpenFlow) and "bottom tier" controller (e.g., Floodlight, NOX, etc.)

Status

- We've looked at:
  - Routing (several variations), ACLs, TE

- Stuff we haven't looked at yet:
  - Multicast, anycast, … ?

- What we'd love to know:
  - Characteristics of actual large networks
  - Pain points for large networks – what are you hoping SDN can solve for you?

Thanks!

murphy@berkeley.edu