Ripcord: A Modular Platform for Data Center Networking

Brandon Heller, David Erickson, Nick McKeown
Stanford University
Stanford, CA, USA
{brandonh, derickso, nickm}@stanford.edu

Rean Griffith, Igor Ganichev, Scott Shenker
University of California
Berkeley, CA, USA
{rean, igor}@eecs.berkeley.edu
shenker@icsi.berkeley.edu

Kyriakos Zarifis
International Computer Science Institute
Berkeley, CA, USA
zarifis@icsi.berkeley.edu

Daekyeong Moon
Nicira Networks
Palo Alto, CA, USA
dkmoon@nicira.com

Scott Whyte, Stephen Stuart
Google, Inc.
Mountain View, CA, USA
{shwhyte, sstuart}@google.com

ABSTRACT

In this demo, we present Ripcord, a modular platform for rapidly prototyping scale-out data center networks. Ripcord enables researchers to build and evaluate new network features and topologies, using only commercially available hardware and open-source software. The Ripcord demo will show three examples of custom network functions, operating together, on top of a 160-node cluster. The first is a routing engine that isolates classes of traffic. The second is a dynamic network manager that adjusts links and switch power states to reduce energy. The third is a statistics aggregator that supports network health monitoring and automatic alerts. The demo will be interactive, with a visualization of live parameters for each link and switch, such as bandwidth, drops, and power status, as well as a control panel to modify the traffic load. We feel that an interactive demo is the best way to introduce the research community to Ripcord and get their feedback.

Categories and Subject Descriptors: C.2.1 [Network Architecture and Design]: Network communications; C.2.2 [Network Protocols]: Routing protocols

General Terms: Design, Experimentation, Management

Keywords: Data center network, Ripcord, OpenFlow

1. DETAILS

This section describes each major component of the proposed demo, including the base Ripcord platform, interactive dashboard, custom modules, and hardware setup. Note that we have a different goal from the SIGCOMM 2009 FlowVisor demo [6]. Instead of multiple controllers sharing an unstructured enterprise network, our goal is to show a single modular controller operating on structured data center networks.

1.1 Ripcord

See Figure 1 for an overview of Ripcord. The architecture includes a number of primitives for building new data center routing algorithms and management tools, and is intended to help researchers address key design challenges in the data center, including scalability, server migration, and forwarding state. Ripcord leverages NOX [2], an OpenFlow controller platform, to pass messages between modules and to modify and view switch state (such as flow entries and statistics).

The Ripcord prototype implements multiple data center routing engines, including ones with similar elements to VL2 [1] and PortLand [4], as described in [7]. Ripcord can simultaneously run multiple routing schemes on the same network, enabling side-to-side comparisons as well as distinct routing engines for different services. Each routing engine uses the structured topology representation; for example, PortLand-style routing can run on a VL2-style aggregated topology, with no code changes.

Figure 1: Ripcord Architecture
1.4 ElasticTree

ElasticTree is a dynamic optimizer that tries to shut off as many unneeded network elements as possible, while respecting performance and fault tolerance constraints \cite{3}. Given a traffic matrix and network topology, ElasticTree generates the set of switches and links that need to be active. A range of optimizers have been implemented, which vary in optimality, generality, and solution time. This demo will extend the ElasticTree paper to a larger system with more realizability, generality, and solution time. This demo will extend the ElasticTree paper to a larger system with more realizability, generality, and solution time.

1.5 Aggregated Statistics

The Aggregated Statistics application collects flow, switch and link statistics from the network and provides the underlying data to both ElasticTree and the visualization. It also enables interactive queries of displayed network elements. For example, users can obtain detailed metrics on individual (or aggregated) flows in the network. They can gather and track individual port measurements from switches, or track link utilization trends to assess the impact of changing traffic engineering and power management decisions.

1.6 Data Center Platform

The expected platform is a 160-node cluster. The network is organized as a three-layer fat tree with four-port switches, except that instead of using two 10 Gbps downlinks, each edge switch uses 20 1 Gbps downlinks to hosts. Each switch runs a port of OpenFlow, an open-source, vendor-neutral, flow-based API added to commercial switches, routers and access points \cite{6}. When an OpenFlow-enabled switch receives a packet for which there is no matching flow entry, it is sent to a controller that makes a decision to add flow entries in switches as needed to set up a path. Alternatively, as is done by some Ripcord routing engines, paths can be set up in advance.

Even though the prototype is not production quality, we believe that Ripcord presents a compelling framework for researchers to implement, evaluate, and (eventually) deploy new ideas. The three modules created for this demo show its flexibility, and we expect the visualization to be useful both in debugging new Ripcord modules, as well as understanding the traffic patterns of data center applications.

2. REFERENCES


