

BROADCOM SDN SOLUTIONS OF-DPA (OPENFLOW DATA PLANE ABSTRACTION) SOFTWARE



Network Switch Business Unit
Infrastructure and Networking Group

- **SDN Principles**
- **OpenFlow Switch Options**
- **Introducing OF-DPA 1.0**
- **Use Case Example**
- **How to Download and Use**

- **SDN as Defined by the Open Networking Foundation is Based on Three Principles**
 - Centralization
 - Traditional network devices independently determine how to forward packets using distributed algorithms
 - SDN provides a system-wide view of the network to enable centralized route determination
 - Programmability
 - Applications can program rather than just configure network elements
 - SDN requires open APIs to create application ecosystems
 - Disaggregation
 - Traditional network devices use monolithic forwarding and control plane functions from a single vendor
 - SDN separates data plane packet forwarding from control plane functions such as route determination

- **Use a Single Table**
 - Treat OpenFlow rules like ACLs
 - Implement in TCAMs – most expensive resource

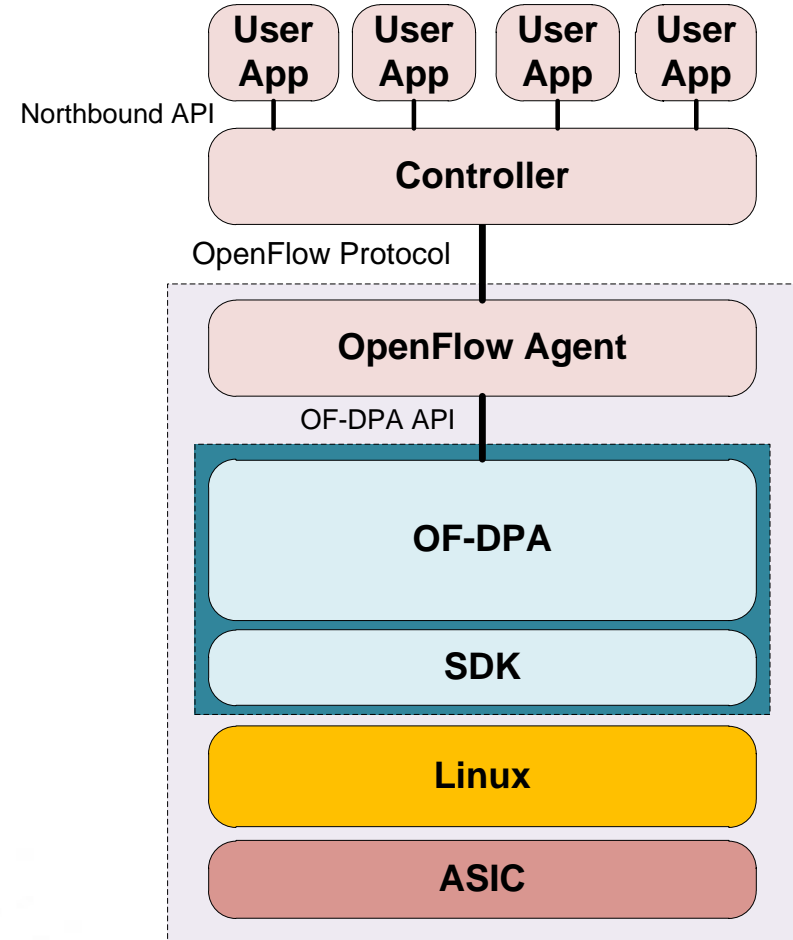
- **Use a Software Switch**
 - Trades off flexibility for performance and cost
 - Suitable for vSwitches but not for infrastructure

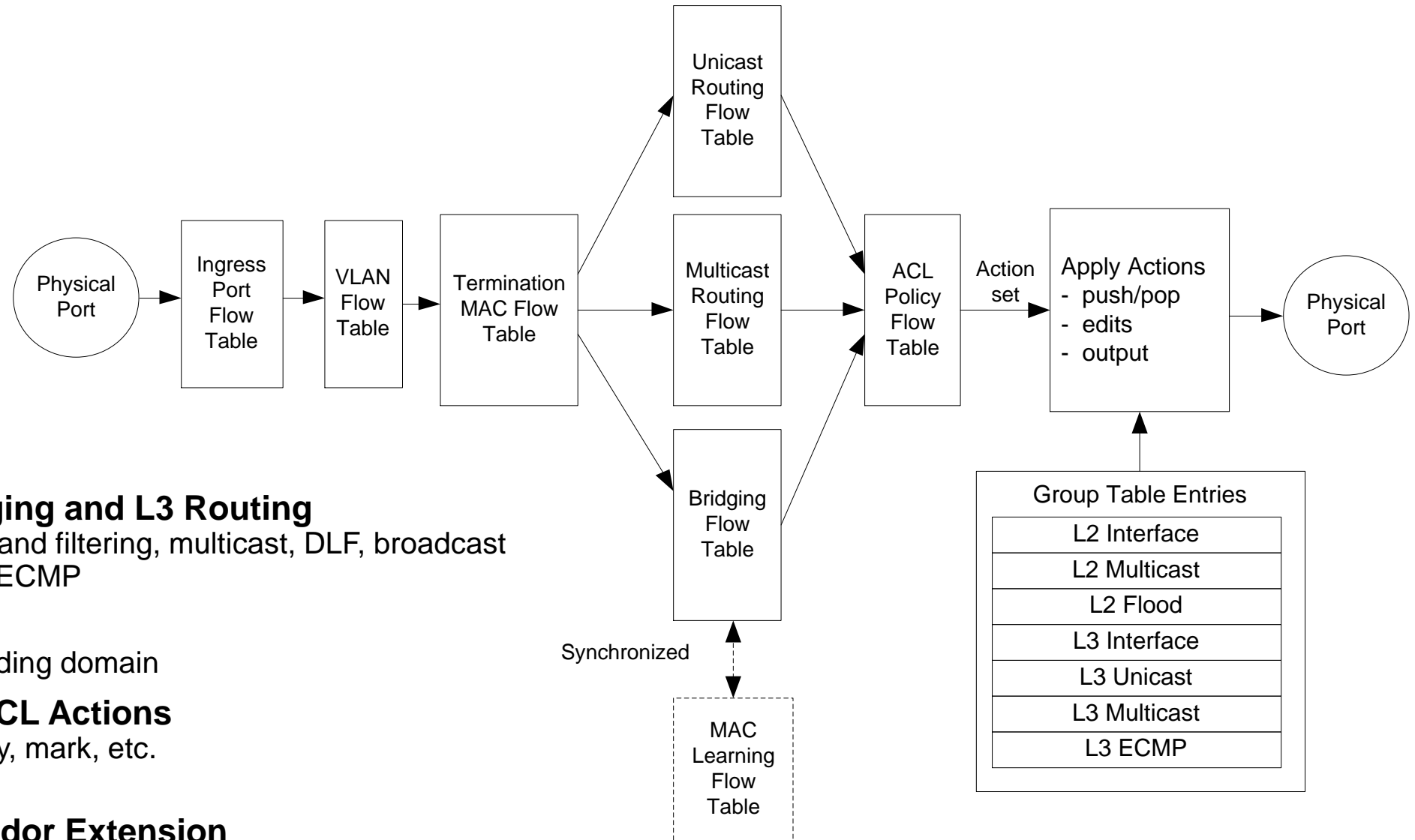
- **Wait for a New Architecture**
 - Interesting research and implementation issues
 - Hardware will not be widely available for a few years

- **Use Existing Hardware**
 - Makes hardware pipeline programmable in terms of OpenFlow objects
 - Enables OpenFlow on widely available network devices today

OpenFlow Data Plane Abstraction (OF-DPA)

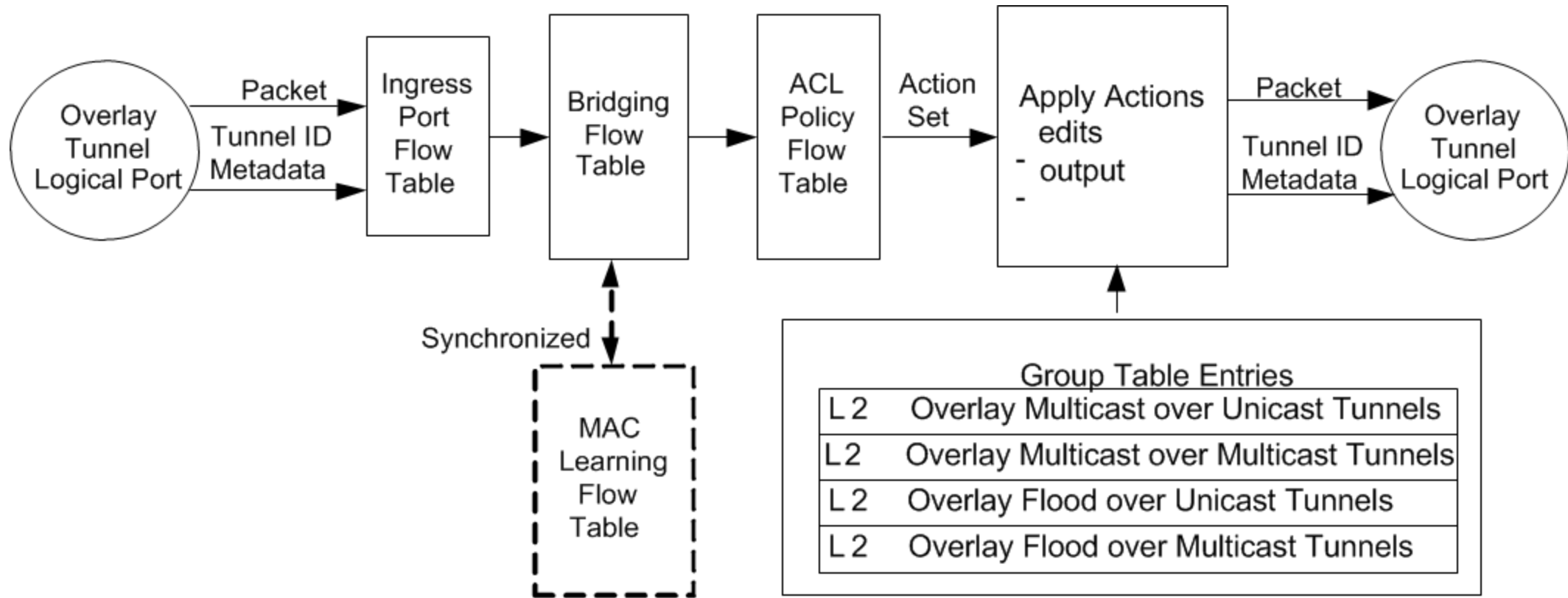
- **OpenFlow 1.3.1 Compliant Programmability for Broadcom ASICs**
- **Enables an OpenFlow 1.3.1 Controller to Provide Centralized Control**
- **Leverages Emerging ONF Forwarding Abstractions Working Group Concepts**
- **Supports Programmability for SDN Use Cases, such as:**
 - Automated Provisioning
 - Network Virtualization
 - Traffic Engineering
 - Monitoring and Analytics Tap Aggregation
 - Service Chaining
- **Layered Application on Broadcom SDK**
 - Portable across multiple ASICs





- **Full-Feature L2 Bridging and L3 Routing**
 - L2 VLAN assignment and filtering, multicast, DLF, broadcast
 - L3 unicast, multicast, ECMP
- **VXLAN Gateway**
 - Isolated tenant forwarding domain
- **Wide-Match Policy ACL Actions**
 - Redirect, drop, classify, mark, etc.
 - L2 header rewrite
- **Source Learning Vendor Extension**

OF-DPA VXLAN OVERLAY PIPELINE



- **Overlays Configured on Logical Ports**
- **Tenant Traffic Identified using Tunnel Id Metadata**
- **Tenant Traffic in Isolated Forwarding Domains**

OF-DPA USE CASE: TRAFFIC ENGINEERING ELEPHANT FLOWS IN LARGE SCALE CLOS NETWORK

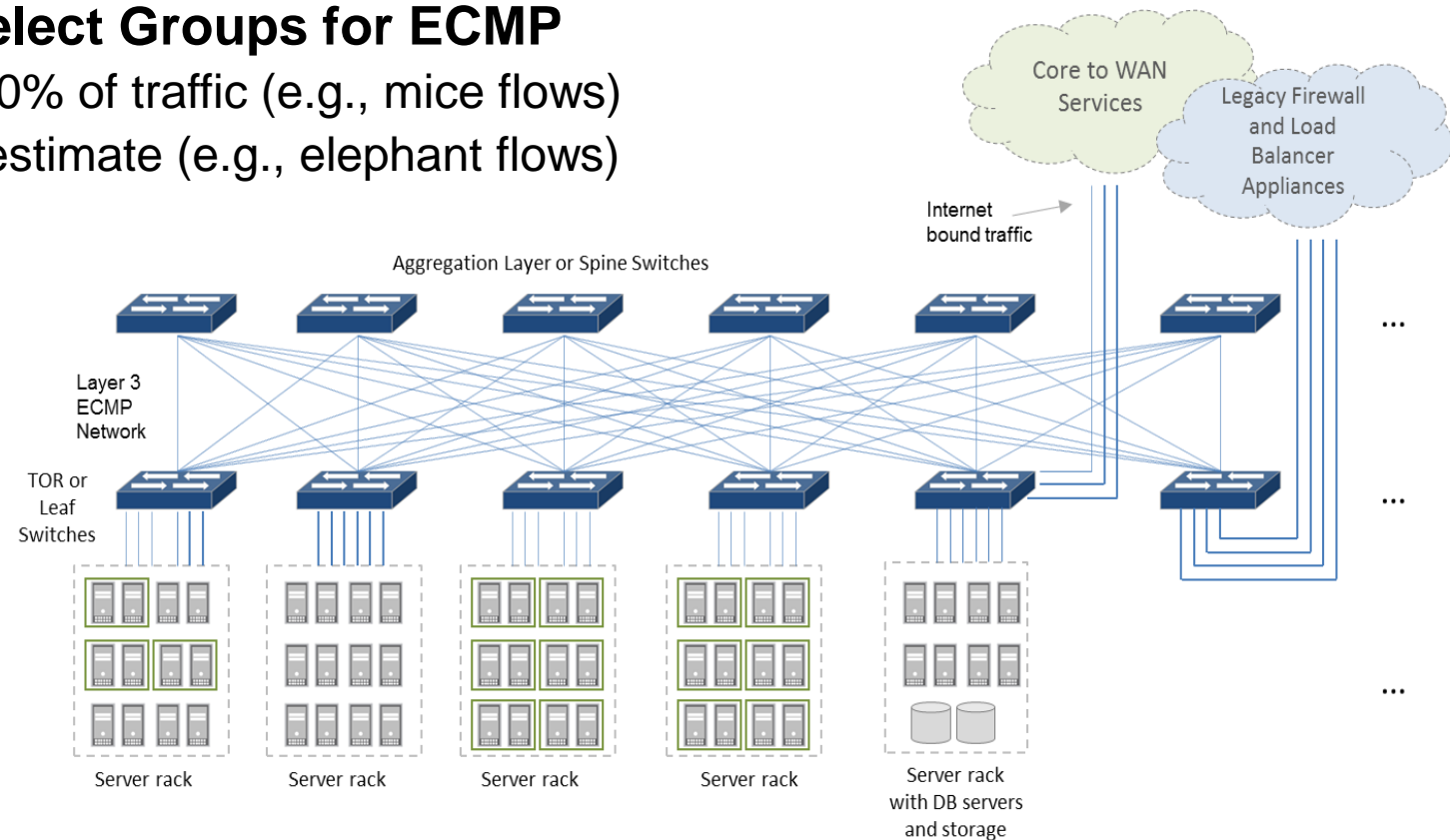
- **Datacenter CLOS Network Designs**
 - IP routed infrastructure, multipath load balancing for resiliency
 - 48x10G+6x40G Leaf Switch, 32x40G Spine Switch
- **OpenFlow 1.0 Requires Many Flows**
- **OpenFlow 1.3.1 Can Use L3 Tables, Select Groups for ECMP**
 - Prescriptive subnet forwarding works for 90% of traffic (e.g., mice flows)
 - Controller places 10% based on demand estimate (e.g., elephant flows)

Engineered Elephant Flows for Boosting Application Performance in Large Scale CLOS Networks

Using OpenFlow 1.3.1-compliant OpenFlow Data Plane Abstraction (OF-DPA) Specification & Software

A White Paper from Broadcom Corporation

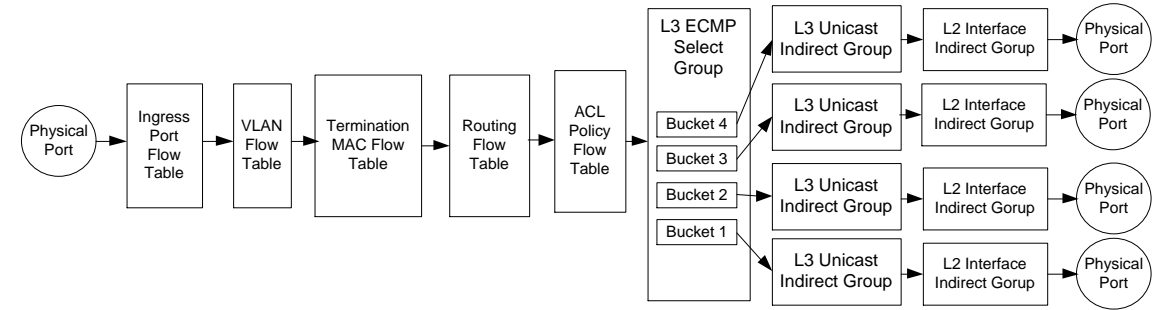
February, 2014



ORDER OF MAGNITUDE ASIC TABLE UTILIZATION AND SCALE IMPROVEMENTS

OF-DPA 1.0 Solution

- Large Routing Flow Table with subnet routes
- Custom L3 ECMP select group entries for mice
- Policy ACL Table redirects for elephants

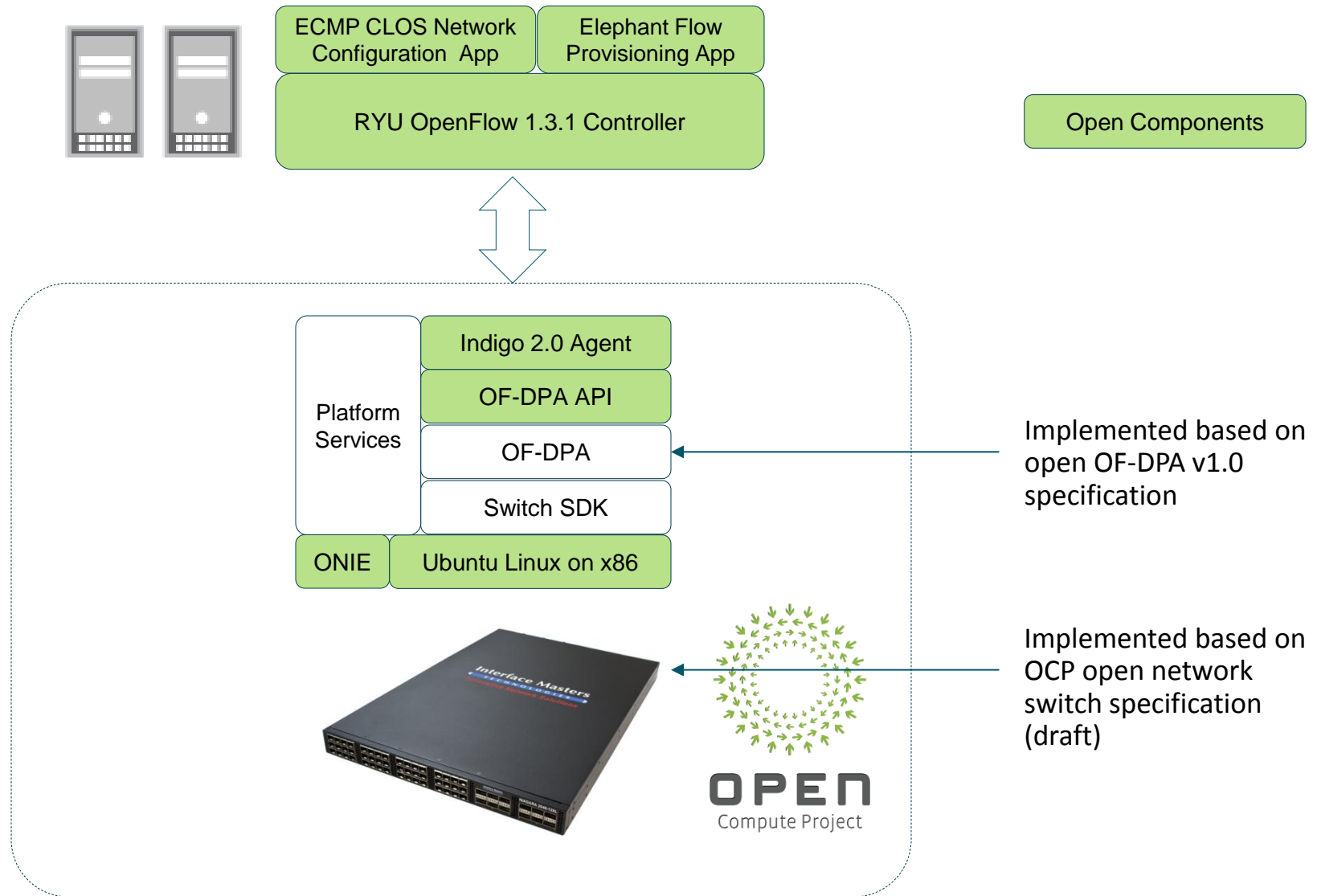


# of hosts, Ingress Ports	# of IP DEST	# of L4 SRC Ports	# of connections (flows)	# of elephant flows (10%)	# of Egress (Uplink) Ports	VLAN Flow Table Entries Needed	Term MAC Table Entries Needed	ECMP Group Entries Needed	L3 Unicast Group Entries Needed	L2 Interface Group Entries Needed	L3 Routing Table Flow Entries Needed	# of ACL Flow Table Entries (TCAM)	# of OpenFlow 1.0 Table Entries (TCAM)	# of OpenFlow 1.3.1 Table Entries
48	1	1024	1024	102	6	48	1	1	6	6	1	102	1126	165
48	2	2048	4096	410	6	48	1	1	6	6	2	410	4506	474
48	4	4096	16384	1638	6	48	1	1	6	6	4	1638	18022	1704
48	6	6144	36864	3686	6	48	1	1	6	6	6	3686	40550	3754
48	8	8912	71296	7130	6	48	1	1	6	6	8	7130	78426	7200
48	10	12288	122880	12288	6	48	1	1	6	6	10	12288	135168	12360
48	12	16384	196608	19661	6	48	1	1	6	6	12	19661	216269	19735

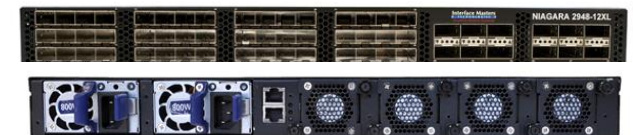
37K total flows, 3.7K are elephant flows –

OpenFlow v1.3.1 with OF-DPA: ~ 3.7K entries using both TCAM and SRAM-based tables
 OpenFlow v1.0: ~41K entries using TCAM-based tables.

OPEN HARDWARE AND SOFTWARE IMPLEMENTATION



- **Specification Developed by Broadcom in Collaboration with HW Vendor**
 - Open Rack 1U form factor (21 inches) and Enterprise Rack 1U form factor (19 inches)
 - Base configuration is bare-bones and power optimized
- **Uses Broadcom StrataXGS® Trident Series Switch and AMD x86 G-Series SO**
 - Trident II BCM56850 with 1.28 Tbps switching capacity and large unified tables
 - AMD x86 GX-416RA with 4 cores, running at 1.6 GHz
 - Supports server-class Linux OS
 - Supports multiple off-the-shelf Network OS options including ICOS 2.3
- **High-Density Configurations**
 - Leaf: 48 x 10 GbE + 6 x 40 GbE and 48 x 10 GbE + 12 x 40 GbE
 - Spine: 32 x 40 GbE
- **Optional Data Plane Processor for L4-L7 Network Functions/Applications**
 - Data Plane Processor Module (DPP): BCM XLP432 with 8 cores, running at 1.5 GHz
 - With DPI, IP Forwarding and IPSEC acceleration



- **Maps OpenFlow Objects to Hardware**
 - Maintains object database with counters, expiration, etc.
 - Implements objects using SDK calls

- **Call-Through (Synchronous) Operation**
 - Changes committed to hardware on successful return
 - Statistics returned from database

- **Agent Implements OpenFlow Protocol-Specific Features**
 - Message processing
 - Asynchronous operation and barrier messages (local state)
 - Object search (e.g., by cookie mask)
 - Message bundling
 - Primary and backup controller interaction

- **Extensible for Future Enhancements**

- **Initialization**
 - Get version, platform, etc.
- **Flow Table APIs**
 - Add/modify/delete flow entries
 - Statistics get
 - Walk flow table
 - Get flows by cookie
- **Group Table APIs**
 - Add/delete group table entries, walk group table
 - Add/modify/delete buckets, walk buckets
 - Get status
- **Port APIs**
 - Configure
 - Walk port table
 - Get status
- **Queue APIs**
 - Configure rates, get status
- **Packet APIs**
 - Packet send, receive
- **Events**
 - Event receive
 - Port, flow removed, flow added

- **Download from GitHub**
 - <https://github.com/Broadcom-Switch/of-dpa>
- **Support via GitHub Forum**
- **Broadcom OF-DPA Page**
 - <http://www.broadcom.com/products/Switching/Software-Defined-Networking-Solutions/OF-DPA-Software>

THANK YOU

