Cisco ASR 9000 Architecture

Mustafa Bayramov, Consulting System Engineer.
Приз за знания

Принимайте активное участие в Cisco Expo и получите в подарок Linksys E900.

Как получить подарок:

- внимательно слушать лекции по технологиям Cisco
- посещать демонстрации, включенные в основную программу
- пройти тесты на проверку знаний

Тесты будут открыты:

с 15:00 25 октября по 16:30 26 октября

www.ceq.com.ua
Session Goal

- To provide you with a thorough understanding of the Cisco ASR 9000 Router architecture, RSP, fabric, and line card design, packet flows, and ASR 9000 nV architecture

- This session will not examine baseline IOS-XR, for example, IOS-XR control plane and management plane protection, modular OS design, software package, SMU installation, IOS-XR routing configuration, etc

- This session will not examine the baseline configuration of ASR 9000, such as L3, L2 forwarding and feature configurations.
Agenda

- **HW Overview**
  - Chassis, RSP, line card, service modules

- **System Architecture**
  - Fabric architecture
  - Line card architecture
  - ASR 9001 architecture
  - Packet flow, control plane and data plane
  - Internal multicast replication
  - System scale

- **nV (network virtualization) Architecture**
  - nV edge
  - nV satellite
Cisco ASR 9000 Overview

Next Generation Service Provider Edge/Aggregation
Large Data Center Inter-Connect and Cloud Gateway

- Designed for Longevity & TCO: Scalable up to 440Gbps/1.2Tbps of Bandwidth per Slot
- Based on IOS-XR & Cisco PRIME for Nonstop Availability & Manageability
- Full L2 and L3 feature set, enables Network Convergence of Business & Residential Services for both Fixed & Mobile Networks
- Advanced Video & Mobility DNA
- Universal data center inter-connect and cloud gateway Router
- Next Generation Broadband Aggregation.
- nV (network virtualization) for operational saving.
- Full HW portfolio: 1RU chassis to full rack 48Tbps system, T1/E1 to nx100G port and LC
ASR 9000 Chassis Overview

Same HW&SW architecture, identical features across different chassis type → one ASR 9000 family

<table>
<thead>
<tr>
<th>Max Capacity (bi-directional)</th>
<th>ASR 9001 (Ironman)</th>
<th>ASR 9006</th>
<th>ASR 9010</th>
<th>ASR 9922 (Megatron)</th>
</tr>
</thead>
<tbody>
<tr>
<td>120Gbps</td>
<td>440G/slot 4 I/O slots</td>
<td>440G/slot 8 I/O slots</td>
<td>1.2T/slot 20 I/O slot</td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>2RU</td>
<td>10RU</td>
<td>21RU</td>
<td>44RU</td>
</tr>
<tr>
<td>Air Flow</td>
<td>Side to side</td>
<td>Side to back</td>
<td>Front to back</td>
<td>Front to back</td>
</tr>
<tr>
<td>FCS</td>
<td>shipping</td>
<td>Shipping</td>
<td>Shipping</td>
<td>Shipping</td>
</tr>
</tbody>
</table>

240 Gbps

3.5 Tbps

7 Tbps

48 Tbps
ASR 9010 and ASR 9006 Chassis

- Integrated cable management with cover
- Front-to-back airflow
- System fan trays
- Side-to-back airflow
- RSP (0-1)
- Line Card (0-3, 4-7)
- System fan trays
- Air draw
- 6 or 8 Modular Power Supplies
- 3 or 4 Modular Power Supplies
- Line Card (0-3)
- RSP (0-1)
- cable management
ASR 9001 “Iron Man” Chassis

Two Modular bays
Supported MPA: 20xGE, 2/4x10GE, 1x40GE (2HCY12)

- Redundant (AC or DC) Power Supplies
- Field Replaceable
- GPS, 1588
- Console, Aux, Management
- BITS
- EOBC ports for nV Cluster (2xSFP)
- Fixed 4x10G SFP+ ports
- Fan Tray Field Replaceable
ASR 9922 “Megatron” Chassis

**Slots**
- 20 Line Card Slots
- 2 dedicated RP slots*
- multi-plane, multi-stage fabric
- 4+1 Switch Fabric Redundancy

**Dimensions**
- Height: 44 RU (AC & DC)
- Depth: 28.65” (800mm)
- Width: 17.75” (fits 19” rack)

**Power**
- AC & DC power supplies
- Pay As You Grow Modular Power
- 24KW max power, ~30W per 10GE

**Bandwidth**
- efficient, scalable fabric silicon
- 550G with 4+1 fabric
- higher BW fabrics in development

* ASR 9922 RP and switch fabric ASIC are separated on different physical card. The RP portion is the same as RSP440.
Power and Cooling

### DC Supplies

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1/1.5 kW</td>
<td>2.1 kW</td>
</tr>
</tbody>
</table>

### AC Supplies

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 kW</td>
<td>3 kW</td>
</tr>
</tbody>
</table>

- Fans unique to chassis
- Variable speed for ambient temperature variation
- Redundant fan-tray
- Low noise, NEBS and OSHA compliant

### Features

- Single power zone
- All power supplies run in active mode
- Power draw shared evenly
- 50 Amp DC Input or 16 Amp AC for Easy CO Install
## ASR 9006/9010 RSP (Route/Switch Processors)

<table>
<thead>
<tr>
<th></th>
<th>RSP2</th>
<th>RSP440</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processors</td>
<td>2 x 1.5GHz Freescale 8641D CPU</td>
<td>Intel x86 Jasper Forest 4 Core 2.27 GHz</td>
</tr>
<tr>
<td>RAM (user expandable)</td>
<td>4GB @133MHz SDR</td>
<td>6GB (RSP440-TR) and 12GB (RSP440-SE) version @1066MHz DDR3</td>
</tr>
<tr>
<td>Cache</td>
<td>L1: 32KB, L2: 1MB</td>
<td>L1: 32KB per Core, L2: 8MB shared</td>
</tr>
<tr>
<td>Primary persistent storage</td>
<td>4GB</td>
<td>16GB - SDD</td>
</tr>
<tr>
<td>Secondary persistent storage (HD/SSD)</td>
<td>30GB - HDD</td>
<td>16GB - SDD</td>
</tr>
<tr>
<td>USB 2.0 port</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>HW assisted CPU queues</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>nV Cluster – EOBC ports</td>
<td>No</td>
<td>Yes, 2 x 1G/10G SFP+</td>
</tr>
<tr>
<td>Switch fabric bandwidth</td>
<td>184G/slot (with dual RSP)</td>
<td>440G/slot (with dual RSP)</td>
</tr>
</tbody>
</table>
RSP440 – Front Ports

1G/10G SFP+ *
EOBC ports for nV Cluster

IEEE 1588, GPS
SyncE, IEEE1588 master and slave
10/100M Copper Ethernet

BITS/J.211
Sync 0, Sync 1
RJ45

Sync 0, Sync 1

USB Type A

LEDs
Status, Alarm

Aux

Console

Management Ethernet

* nV EOBC port only support 1G with SFP optics in the current shipping release. 10G (SFP+ optics) support plan for the future release

Note, red color is the new functional ports, which is supported on RSP440 only, not RSP2
RSP Engine Architecture

Front Panel:
- BITS
- CF card or USB
- Mgt Eth
- Mgt Eth
- Console
- Aux
- Alarm

CPU Complex:
- CPU
- MEM
- HDD
- 4G disk
- I/O FPGA
- NVRAM
- Boot Flash

Timing Domain:
- Clock
- Time FPGA

EOBC/Internal GE switch:
- Ether Switch

Switch fabric:
- FIA
  - Punt FPGA
  - Crossbar Fabric ASIC

Arbitration:
- Arbiter
ASR 9000 Ethernet Line Card Overview

First-generation LC (Trident NP)
- \( -L \), \( -B \), \( -E \)
- A9K-40G
- A9K-4T
- A9K-8T/4
- A9K-2T20G
- A9K-8T
- A9K-16T/8

Second-generation LC (Typhoon NP)
- \( -TR \), \( -SE \)
- A9K-24x10GE
- A9K-2x100GE
- A9K-MOD80
- A9K-MOD160
- MPAs
  - 20x1GE
  - 2x10GE
  - 4x10GE
  - 1x40GE
  - 2x40GE
ASR 9000 SIP-700 and SPA

SPA support list:

Powerful and Flexible QFP Processor
- Flexible ucode Architecture for Feature Richness
- L2 + L3 Services: FR, PPP, HDLC, MLPPP, LFI, L3VPN, MPLS, Netflow, 6PE/VPE

Strong QoS
- 128+K Queues,
- 128+K Policers
- H-QoS
- Color Policing

Scalability
- Distributed control and data plane; 20G, 4 Bays
- L3 if, route, session, protocol – scaled up for MSE needs

High Availability
- IC-Stateful Switch Over capability,
- MR-APS
- IOS-XR base for high scale and reliability

A9K-SIP-700

SPA Support:
- ChOC-12: T1 / T3, MLPPP, LFI, IC-SSO, MR-APS
- ChOC-3/STM1 + ChOC-12: DS0 / STM1 + ChOC-48: DS3 / E3 / STM4
- POS: OC3/STM1, OC-12/STM4, OC-48/STM16, OC-192/STM64
- Ch T1/E1, Ch T3/E3, CEoPs, ATM
ASR 9000 ISM (Integrated Service Module)

- Integrated application intelligence into the ASR 9000 System
- Reduced footprint, power and cooling
- Use with existing CDS deployments
- CGN Integration: NAT44, DS-lite AFTR
- Pay as you grow modular investment
- Carrier class resiliency

Cisco ASR 9000 ISM

Feature | ASR 9000 ISM Capabilities
--- | ---
Applications | Ultra-Dense VoD, TV, Internet Streaming, Error Repair, CGv6
Bandwidth | 30-40 Gbps streaming capacity
| ~3 Gbps cache fill rate
Compatibility | Works with all CDS appliances
Concurrent Streams | Up to 8,000 SD equivalent
Content Cache | 3.2 TBytes at FCS - Modular Design
Video Formats | MPEG2 & AVC/H.264
Transport | MPEG over UDP / RTP
Session Protocols | RTSP / SDP
Environmental | NEBS / ETSI compliant

CDS: Manage 8,000 streams up to 40G per second
CGv6: 20M translations, 1M translations/sec., ~15Gbps throughput / ISM
Agenda

- HW Overview
- System Architecture
  - Fabric architecture
  - LC architecture
  - ASR 9001 architecture
  - Packet flow, control plane and data plane
  - Internal multicast replication
  - System scale
- nV Architecture
ASR 9000 Switch Fabric Overview

3-Stage Fabric

8x55Gbps = 440Gbps with dual RSP
4x55Gbps = 220Gbps with single RSP

8x7.5G = 60G raw bandwidth
55G available use bandwidth per fabric channel
Fabric Back-compatible: Mix New and Existing LC

- **FIA0**
  - Dual-FIA 8xNPs 1\textsuperscript{st} gen Linecard
  - 8x23G bi-directional

- **FIA1**
  - 8x55G bi-directional

- **FIA**
  - Single-FIA 4xNPs 1\textsuperscript{st} gen Linecard
  - 4x23G bi-directional

- **RSP0**
  - Arbiter
  - fabric

- **RSP1**
  - Arbiter
  - fabric

- **2\textsuperscript{nd} gen Line Card**
  - FIA
Fabric Forward-compatible: Mix New LC and Existing Switch Fabric

- **FIA0**: Dual-FIA 8xNPs, 1st gen Linecard
- **FIA1**: Dual-FIA 8xNPs, 1st gen Linecard
- **FIA**: Single-FIA 4xNPs, 1st gen Linecard

Each FIA is connected to an Arbiter and an RSP (RSP0 and RSP1). The connections are 4x23G bi-directional for FIA and 8x23G bi-directional for FIA0 and FIA1.
ASR 9922 Fabric Architecture: 5-plane System

- 2nd gen Line Cards
- FIA
- Fabric cards

All fabric connections are fully-meshed, non-blocking

- 2x55G links
- 110G per fabric plane

550Gbps/LC or 440Gbps/LC with fabric redundancy

Fabric cards

2nd gen Line Card

2nd gen Line Card

Fabric

FIA

FIA

FIA

2nd gen Line Card

RSP 3 Switch Fabric 7-plane (6+1)
LC Architecture Example – 24x10G

- 3x10GE SFP +
- NP
- FIA
- LC Switch Fabric
- ASIC
- 8x55G
- CPU

Switch Fabric
- RSP0
- RSP1

Original packet format
Super-frame format (unicast only) between switch fabric and FIA, fabric and fabric
LC Architecture Example – 2x100G
Each NPU has Four Main Associated memories TCAM, Search/Lookup memory, Frame/buffer memory and statistics memory:
- TCAM is used for VLAN tag, QoS and ACL classification
- Lookup Memory is used for storing FIB tables, Mac address table and Adjacencies
- Stats memory is used for all interface statistics, forwarding statistics etc
- Frame memory is buffer memory for Queues

- E/B/L or –SE/-TR line card have different TCAM, Stats and Frame Memory size, which give different scale number such as ACL, QoS queues, L2/L3 sub-interfaces, etc per line card
- However, lookup Memory is the same across line card s→ why?
  - To support mix of the line cards without impacting the system wide scale including FIB table, MAC address table, MPLS label space
ASR 9001 Architecture Overview

- NP
  - MPAs: 2,4x10GE, 20xGE, 1x40GE
  - SFP+ 10GE
  - SFP+ 10GE
  - SFP+ 10GE
  - SFP+ 10GE
- FIA
- Switch Fabric ASIC
- On-board 4x10 SFP+ ports
- Internal EOBC
- LC CPU
- RP CPU
Agenda

- HW Overview
- System Architecture
  - Fabric architecture
  - LC architecture
  - ASR 9001 architecture
    - Packet flow, control plane and data plane
    - Internal multicast replication
    - System scale
- nV Architecture
ASR 9000 Fully Distributed Control Plane

**LPTS:** control plane policing

EOBC (1G internal switch)

- **LC1:** Switch Fabric
- **RP:** Switch Fabric
  - **CPU:**
  - **PU:** FPGA
  - **FIA**

- **RP CPU:** Routing, MPLS, Multicast, HSRP/VRRP, etc
- **LC CPU:** ARP, ICMP, BFD, Netflow, OAM, etc

LPTS: control plane policing
Packet Flow Overview – 2-stage Forwarding

1. **Ingress NP look up**
   - Get egress NP information (added into the NP/fabric header), apply ingress features

2. **Egress NP look up**
   - Get egress logical port, VLAN, MAC, ADJ information, etc for packet rewrite, apply egress features
Internal Multicast Replication Overview

1. Fabric to LC Replication
2. LC fabric to FIA Replication
3. FIA to NP Replication
4. NP to egress port Replication

Efficiency: replicate if required
Line rate: for fully loaded chassis
Simple and clean architecture
Predictable performance

FGID – Fabric Group ID
MGID – Multicast Group ID
MFIB – Multicast Forwarding Information Base

FGID/FPOE
MGID/FPOE

Switch Fabric
FIA
NP
3x10GE SFP +

Switch Fabric
FIA
NP
3x10GE SFP +

Switch Fabric
FIA
NP
3x10GE SFP +

Switch Fabric
FIA
NP
3x10GE SFP +

Switch Fabric
FIA
NP
3x10GE SFP +

Switch Fabric
FIA
NP
3x10GE SFP +

Switch Fabric
FIA
NP
3x10GE SFP +

Switch Fabric
FIA
NP
3x10GE SFP +

Switch Fabric
FIA
NP
3x10GE SFP +

Switch Fabric
FIA
NP
3x10GE SFP +

Switch Fabric
FIA
NP
3x10GE SFP +
## ASR 9000 System Scale Overview

<table>
<thead>
<tr>
<th>Feature</th>
<th>RSP2 and Trident line card</th>
<th>RSP440 and Typhoon line card</th>
<th>Comments</th>
</tr>
</thead>
</table>
| FIB (V4+V6)     | 1.3M                       | 4M                          | V4 and V6 share the same table  
V6 uses two FIB entries  
Support per-VRF FIB table download per LC  |
| Multicast FIB   | 32K                        | 128K                        |                                                                          |
| MAC             | 512K                       | 2M                          |                                                                          |
| L3 VRF          | 4K                         | 8K                          |                                                                          |
| BD/VFI          | 8K                         | 64K                         |                                                                          |
| PW              | 64K                        | 128K                        |                                                                          |
| L3 interface    | 20K                        | 20K                         | -SE card: 20K/LC, -TR card: 8K/LC,  
-E: 20K/LC, -L/-B: 4K/LC                                                   |
| L2 interface    | 64K                        | 128K                        |                                                                          |
Agenda

- HW Overview
- System Architecture
  - Fabric architecture
  - LC architecture
  - ASR 9001 architecture
  - Packet flow, control plane and data plane
- nV Architecture
Before: nV Technology

Each device managed separately: different CLI experience, different image upgrade, different release cycle

Manual and complex protocols configuration between edge and aggregation

Inconsistent features and potential inter-operability issue between edge and aggregation

Port scale limited to physical chassis

After: nV Technology

Edge and aggregation managed as one virtual system through Cisco Prime IP NGN

Plug-N-Play for the satellite: Reduced protocol complexity between edge and aggregation

Simplified image upgrade

Single release vehicle offering feature consistency.

Scale the GE port by adding more satellite
ASR 9000 nV Satellite Overview
Self-managed Access

- Install special satellite image on the selected access device to make it ASR9K nv satellite
- Satellite and ASR 9000 Host run satellite protocol for auto-discovery, provisioning and management
- Satellite and Host could co-locate or in different location. There is no distance limit between satellite and Host
- The connection between satellite and host is called "nv fabric link", which could be L1 or over L2 virtual circuit (future)

- From end user point of view, satellite looks/feels/works like a ASR9K "remote or virtual" line card. The interfaces on the satellite looks/feels/works the same as the interfaces on the local ASR9K line cards
- From end user point of view, ASR9K Host and associated satellites is one virtual Router system. Satellite is plug-n-play, zero touch configuration/management

Satellite is plug-n-play, zero configuration
ASR 9000v Hardware Overview

**Power Feeds**
- Redundant -48vDC Power Feeds
- Single AC power feed

**Pluggables**
- 44x10/100/1000 Mbps Pluggables
  - Full Line Rate Packet Processing and Traffic Management
  - Copper and fiber SFP optics

Max Power 210 Watts
Nominal Power 159 Watts

**Field Replaceable Fan Tray**
- Redundant Fans
- ToD/PSS Output
- Bits Out

**4x10G SFP+**
- Initially used as Fabric Ports ONLY (could be used as access port in the future)
- Copper and fiber SFP+ optics

**Industrial Temp Rated**
- -40C to +65C Operational Temperature
- -40C to +70C Storage Temperature
Satellite – Host Control Plane
Satellite discovery and control protocol

**Discovery Phase**
- A CDP-like link-level protocol that discovers satellites and maintains a periodic heartbeat
- Heartbeat sent once every second, used to detect satellite or fabric link failures. BFD based fast failure detection plan for future release

**Control Phase**
- Used for Inter-Process Communication between Host and Satellite
- Cisco proprietary protocol over TCP socket for the time being. It could move to standard in the future
- Get/ Set style messages to provision the satellites and also to retrieve notifications from the satellite

Standardization is considered for future
Satellite – Host Data Plane Encapsulation

On Satellite

- Satellite receive Ethernet frame on its access port
- Satellite add special nV-tag, then local xconnect packet to its fabric port
- Put packet into fabric port egress queue, transmit packet out

On Host

- Host receive the packet on its satellite fabric port
- Check the nV tag, then map the frame to the corresponding satellite virtual access port
- From there, process packet just as local port, apply potential L2/L3 features, qos, ACL, etc
- Packet is forwarded out of local port, or satellite fabric port to same or different satellite
Satellite Operation (1) – End User View

- Satellite uplink port is treated as internal “fabric” port
- Satellite access port is represented by virtual “nv” interface on the Host. User configure this virtual interface just as regular local L2/L3 interface or sub-interface on the Host
- All satellite configuration is done on the Host
- If real access port goes down, then the “nv” interface will go down as well. If shut down the “nv” interface, then the real satellite access port will shut down as well

“nv” Ethernet interface sample CLIs

```plaintext
interface GigabitEthernet 100/0/0/1
  ipv4 address 1.1.1.1 255.255.255.0
interface GigabitEthernet 100/0/0/2.100 12transport
  encapsulation dot1q 100
  rewrite ingress tag push dot1q 2
```
Satellite Operation (2) – Packet Flow

- No local switching/routing on satellite, all forwarding is via Host
- Satellite ONLY does local connect between access port and fabric, NOT between access ports. No MAC learning involved
- Advanced features are processed on the Host chassis satellite virtual port
- Very few features could be offloaded to satellite directly, including basic QoS, multicast replication, OAM performance measurement, SyncE, 1588*. However, the configuration is still done on the Host

* Only QoS is offloaded to satellite in the initial release
ASR 9000 nV Edge Overview

Simple Resiliency and More Capacity

Leverage existing IOS-XR CRS multi-chassis SW infrastructure
Simplified/Enhanced for ASR 9000 nV Edge

Single control plane, single management plane, fully distributed data plane across two physical chassis → one virtual nV system

Super, Simple network resiliency, and scalable node capacity
nV Edge Architecture Overview

Control Plane EOBC Extension (L1 or L2 connection)

- **Control plane extension**: Active RSP and standby RSP are on the different chassis, they sync up via external EOBC links “AS IF” they are in the same physical chassis.

- **Data plane extension**: bundle regular data links into special “nV fabric link” to simulate switch fabric function between two physical chassis for data packet across.

- Doesn’t require dedicated fabric chassis → flexible co-located or different location deployment.

One Virtual ASR 9000 nV System

- Special external EOBC 1G/10G ports on RSP
- Internal EOBC
- Inter-chassis data link (L1 connection)
  - 10G bundle (up to 32 ports)

- Regular 10G data ports
- LC
nV Edge Configuration

1. Configure nV Edge globally

```text
nv

edge-system
serial FOX1437GC1R rack 1
serial FOX1439G63M rack 0
```

2. Configure the inter-chassis fabric (data plane) links

```text
interface TenGigE1/2/0/0
  nv edge interface

interface TenGigE0/2/0/0
  nv edge interface
```

NO need to configure the inter-chassis control plane EOBC ports. It’s plug-and-play 😊

After this configuration, rack 1 will reload and then join cluster after it boot up
Now you successfully convert two standalone ASR 9000 into one ASR 9000 nV Edge
As simple as this !!!
Supported Topologies in initial release

Satellite must have direct Ethernet over Fiber/copper or Ethernet over Optical transport system (such as Ethernet over a SONET/SDH/CWDM/DWDM network: ring, mesh topology, etc)

- Single home, static pinning
- Single home, fabric link bundle
- Dual home to nV edge, static pinning
- Dual home to nV Edge with fabric link bundle
Satellite Configuration Examples

```
v
satellite 101 \(\leftarrow\) define satellite ID range <100-65534>
  type asr9000v
  ipv4 address 10.0.0.101 \(\leftarrow\) internal communication IP address between host and satellite. This configuration will be optional in the future release with the "auto-IP" feature

satellite 102 \(\leftarrow\) define satellite
  ipv4 address 10.0.0.102
  type asr9000v

interface TenGigE 0/2/0/2
  ipv4 point-to-point
  ipv4 unnumbered Loopback0
nv
  satellite-fabric-link satellite 101
    remote-ports
      GigabitEthernet 0/0/0-9

interface bundle-ethernet 1
  ipv4 point-to-point
  ipv4 unnumbered Loopback0
nv
  satellite-fabric-link satellite 102
    remote-ports
      GigabitEthernet 0/0/0-43
```
BUILT FOR THE HUMAN NETWORK