INAF ISTITUTO NAZIONALE DI ASTROFISICA
OSSERVATORIO ASTRONOMICO DI ROMA

The Heart of the Data Center: A Guide to Designing Scalable and Reliable Networks

A Practical Approach

Federico Fiordoliva
Osservatorio Astronomico di Roma

Coauthors: Stefano Gallozzi(OAR), Fulvio Gianotti(OAS), Ismam Abu(OAS)
Revision: 1/2026

federico.fiordoliva@inaf.it

document created on: 8 gennaio 2026
created from file: `template.tex`
cover page automatically created with `CoverPage.sty`
(available at your favourite CTAN mirror)

LaTeX

Abstract

My work is a very serious game.

Anonymous

A Data Center (DC) network requires a design tailored to accommodate a high density of devices physically distributed across medium-to-large environments.
These devices may need to communicate with other hardware within the same DC, often involving high-volume data throughput. Consequently, the network must fulfill specific requirements, which we will analyze, to achieve a scalable, reliable, and redundant configuration.
The premise of this article assumes that the DC is owned by a single organization, such as the Osservatorio Astronomico di Roma (OAR), which hosts several international projects but does not provide colocation (housing) services.

Keywords — `Data Center, Juniper Switches, Network Design, Redundancy, LACP, VCP, RSTP, Layer 2, MTU, High Availability.`

Contents

List of Acronyms

TOR Top Of the Rack

VC Virtual Chassis

DC Data Center

CS Core Switchs

OAR Osservatorio Astronomico di Roma

HA High Availability

JTAC Juniper Technical Assistance Center

STP Spanning Tree Protocol

RSTP Rapid Spanning Tree Protocol

BPDU Bridge Protocol Data Unit

SCP Secure Copy Protocol

UAC User Account Controll

LLDP Local Link Discovery Protocol

MTU Maximum Transmission Unit

ISSU In-Service Software Upgrades

DAC Direct Attached Copper

LACP Link Aggregation Control Protocol

List of Figures

1 Project Overview and Scope

Our DC requirements involve hosting all the servers for the various OAR research projects.
Since the different research groups manage their hardware infrastructures independently, they have consistently purchased servers based on specific project needs. Consequently, we are dealing with a heterogeneous environment of NAS/SAN storage systems, servers, and network switches.

While it was necessary to maintain their autonomy in managing their respective infrastructures, there was a pressing need to centralize the network management across all projects.
Given that a High Availability (HA) firewall already handles Layer 3 operations for all DC networks, our objective was to unify the management of Layer 2 of the TCP/IP stack only.

The following requirements had to be guaranteed:

- Link redundancy for devices (BONDING).
- Mixed fiber and copper network ports.
- Adequate link speeds (1, 10, and 40 Gbps).
- The ability to perform non-disruptive hardware upgrades In-Service Software Upgrades (ISSU).
- Robust Layer 2 security protocols.
- Maximum internal switching backplane capacity.
- Ensuring connectivity in the event of a device failure within the distribution layer.
- Ensuring connectivity in the event of a device failure within the aggregation layer.

2 Hardware

The choice of network hardware fell on Juniper Networks©, devices, as they provided the necessary features to meet all our requirements.

Note

The configurations described below are widely available on hardware from other manufacturers; while technology names may vary, the underlying principles remain the same.

We selected the **EX** Series, distinguishing between distribution and aggregation devices. The Juniper ©, **EX** Series switches are high-performance switching platforms designed for enterprise (campus), branch office, and data center networks.
They represent the core line for access and aggregation, combining carrier-grade reliability with the simplicity of cloud management.

Here are the fundamental characteristics that define the **EX** family:

Virtual Chassis Architecture

This is Juniper's signature technology, allowing the interconnection of up to 10 switches (depending on the model) to be managed as a single logical device.

- **Centralized Management:** A single IP address and a unified configuration file manage the entire cluster.
- **Redundancy:** If a member switch fails, the remaining Virtual Chassis continues to operate without service disruption.
- **Pay-as-you-grow Scalability:** Network port density can be increased simply by adding a new switch to the existing cluster.

Junos Operating System

All EX Series switches run the same operating system as Juniper's core routers and firewalls (Junos OS).

- **Modularity:** Each process (routing, management, interfaces) runs in a separate protected memory space; if a single process fails, it does not crash the entire switch.
- **Consistency:** The command syntax and automation tools are identical across the entire product line, significantly reducing the learning curve for administrators.

Advanced Security

- **Micro-segmentation:** Leveraging Group-Based Policies (GBP) and support for EVPN-VXLAN architectures to achieve granular traffic isolation.
- **MACsec::** Hardware-based, link-by-link data encryption (AES-256) to secure traffic between switches or between switches and servers.
- **Unified Access Control (UAC):** Deep integration with 802.1X protocols and dynamic user management.

2.1 EX4600-40F

The Juniper **EX4600-40F** is an enterprise and data center-class switch, designed to deliver high performance in aggregation or small-scale core environments. It is a compact 1U platform (note: despite the 2 expansion slots, it is typically a 1U form factor) that focuses on low latency and high-density 10GbE and 40GbE port connectivity.

Here are the technical specifications and key features:

Hardware Specifications

- **Base Port Configuration:** * 24 fixed **10GbE SFP+** ports (also capable of 1GbE operation).
  4 fixed **40GbE QSFP+** ports.
- **Expandability:** Equipped with 2 expansion slots that can accommodate:
  8-port 10GbE SFP+ modules.
  4-port 40GbE QSFP+ modules.
- **Total Capacity:** Supports up to 72 10GbE ports (utilizing breakout cables on 40GbE ports and expansion modules).
- **Performance:**
  Switching Capacity: 1.44 Tbps (bidirectional).
  Throughput: 1.071 Bpps (note: typically expressed as Billions of packets per second for Tbps-class switches, or 1,071 Mpps).
- **Resiliency:** Dual redundant power supplies (AC or DC) and 5 fan modules, all of which are hot-swappable.

2.2 **EX4300-48P**

The Juniper **EX4300-48P** is a high-performance access switch designed for enterprise networks requiring dense connectivity and high reliability.

Here are the technical specifications and key features:

Hardware Specifications

- **Main Ports:** 48 10/100/1000BASE-T RJ-45 ports.
- **Fixed Uplinks:** Equipped with 4 40GbE QSFP+ ports on the rear panel (usable as standard uplinks or as Virtual Chassis Ports - VCPs).
- **Expandability:** One front-facing slot for optional modules (e.g., 4x10GbE SFP+ or 2x40GbE QSFP+).
- **Performance:** Switching capacity: 496 Gbps.

  Forwarding rate: 369 Mpps.
- **Resiliency:** Dual redundant power supplies (AC or DC).

2.3 Bill of Materials

The following devices and modules were purchased to provide full coverage for the entire DC.

- Three (3) **EX4600-40F**
  - Six (6) **EX4600-EM-8F** modules: 8-port SFP/SFP+ expansion modules.
  - Four (4) **QSFP+-40G-ACU7M**: 7-meter 40GbE QSFP+ Direct Attached Copper (DAC) cables.
  - Three (3) **JPSU-650W-AC-AFO**: Additional redundant AC power supplies (Front-to-Back airflow).
- Eight (8) **EX4300-48P** units.
  - Eight (8) **4x 1G/10G SFP/SFP+** modules: 4-port SFP/SFP+ expansion modules.
  - Two (2) **QSFP+-40G-ACU7M** 7-meter 40GbE QSFP+ DAC cables.
  - Six (6) **QSFP+-40G-ACU5M** 5-meter 40GbE QSFP+ DAC cables.
  - Three (3) **JPSU-1100-AC-AFO**: Additional redundant AC power supplies.
- **Tranceiver**:
  - SFP, SFP+, and QSFP+ optical modules (populated as needed).
  - SFP Copper (RJ-45) modules.

3 Configurations

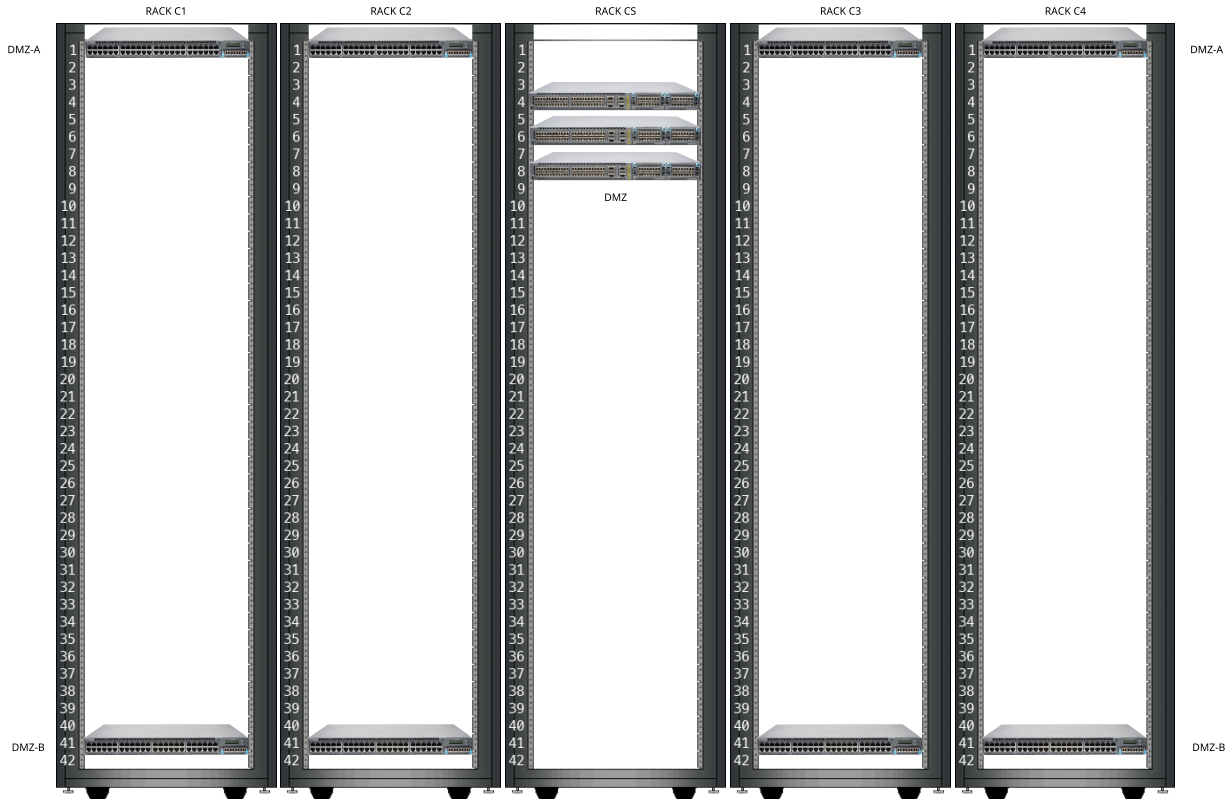
By leveraging Juniper's **Virtual Chassis** technology, three distinct stacks were created: one Core Switchs (CS) and two Top Of the Rack (TOR) stacks.
The TOR stacks consist of the **EX4300-48P** switches, while the CS stack is composed of the **EX4600-40F** units.

The physical layout of the equipment was designed to maximize available connectivity in each rack and to ensure full redundancy. The configuration spans a total of five racks, with the Core Rack positioned in the center and the Server Racks, housing the TOR units, on either side.
To facilitate identification, the stacks have been named as follows:

- **DMZ**: Central Rack (CS)
- **DMZ-A**: Lateral Rack, TOR
- **DMZ-B**: Lateral Rack, TOR

The Junos OS version installed is the one currently recommended by the Juniper Technical Assistance Center (JTAC) for these specific devices.

Figure 1: **Rack**



3.0.1 **UAC**

For device management, a dedicated management VLAN was created, along with users associated with three distinct authorization profiles: `super-user`, `operator`, and `read-only`. User authentication is restricted exclusively to SSH key-based authentication.

3.0.2 **RSTP**

In a DC) environment where only servers are hosted and hardware deployments are strictly controlled, the Spanning Tree Protocol (STP) approach must differ from that used for workstations.
The goal is to ensure that servers—whether newly installed or following a reboot—gain immediate network connectivity without waiting for standard STP negotiation timers. At the same time, the network must be protected against potential server-side misconfigurations.
The adopted strategy is to define these network interfaces as `edge` ports. As soon as the link status becomes 'up', the port transitions immediately to `forwarding mode`.
To maintain security against misconfigurations (such as accidental bridging on the server side), these edge ports must be configured with Bridge Protocol Data Unit (BPDU) protection to drop any incoming BPDUs.

Regarding the ROOT BRIDGE election, it is best practice to manually configure the priority for each stack. In this scenario, the DMZ stack is assigned a priority of 4096, while for DMZ-A and DMZ-B, a priority of 16384 has been selected to ensure it remains lower than the default value of 32768.

3.0.3 **DHCP Snooping**

Another essential protocol for both workstation and DC environments is `DHCP Snooping`.
To prevent unauthorized DHCP servers from assigning IP addresses within the network, this protocol is enabled on all `edge ports`. Any packet identified as part of a DHCP server handshake arriving on an untrusted edge port will trigger a port block.
Crucially, DHCP Snooping must be disabled (or the ports must be configured as trusted) on interfaces hosting legitimate DHCP servers; otherwise, the service will be blocked.

3.0.4 **LLDP**

To maintain constant visibility of what is connected to the network, even within a DC, LLDP should be enabled.

It is a lightweight yet powerful protocol. By enabling it on the server side as well, you gain a complete map of the physical interconnections associated with basic server information, such as the hostname and the specific network interface in use.

3.0.5 **MTU**

On all inter-switch links, the MTU has been set to **9216**, while it will be managed on an ad-hoc basis for specific ports requiring this configuration.

3.0.6 **Backup**

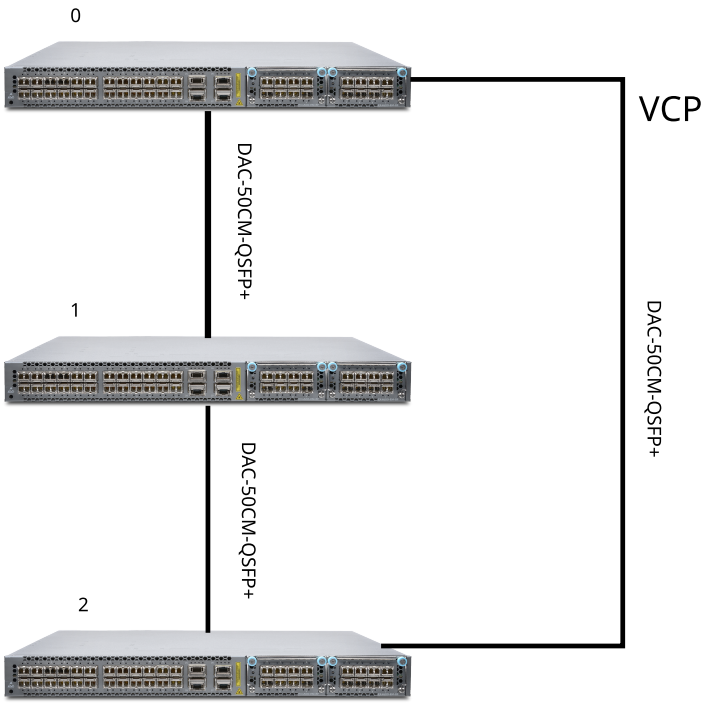
Juniper switches feature an automatic configuration backup mechanism that, when enabled, creates a backup file and transfers it to a remote system. There are two primary backup approaches: **Transfer on Commit** or **Transfer Interval**.
We have chosen the former method; upon every commit, the switch transmits the configuration backup to an internal server using Secure Copy Protocol (SCP).

3.1 Stack - DMZ

The three **EX4600-40F** units were joined into a single Virtual Chassis (VC) using three QSFP+ DAC cables.

This setup creates a ring topology to enhance redundancy; in the event of a device failure, there is always a secondary path for packet forwarding.
The backplane switching speed between the three units in the VC is determined by the link speed used—in this specific deployment, it provides a bandwidth of 40 Gbps.

Figure 2: **DMZ - VC**



This stack acts as both an aggregation and distribution layer, as nearly all servers equipped with fiber interfaces will be connected directly to it.

This approach allows for the full use of all NIC bonding modes on the server side.

Regarding the inter-stack connectivity, to maximize throughput and reliability, a dual 40GbE uplink was implemented using the 802.3ad Link Aggregation Control Protocol (LACP).

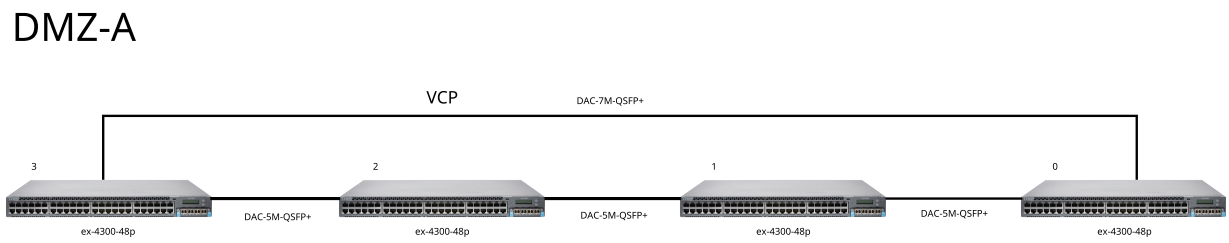
3.2 **Stack - DMZ-A**

The four **EX4300-48P** units were integrated into a single VC using four QSFP+ DAC cables.

This setup establishes a ring topology to maximize redundancy; in the event of a hardware failure, a secondary path is always available for packet forwarding.

The backplane switching throughput between the units within the Virtual Chassis matches the speed of the physical links used—in this deployment, 40 Gbps.

Figure 3: **DMZ-A - VC**



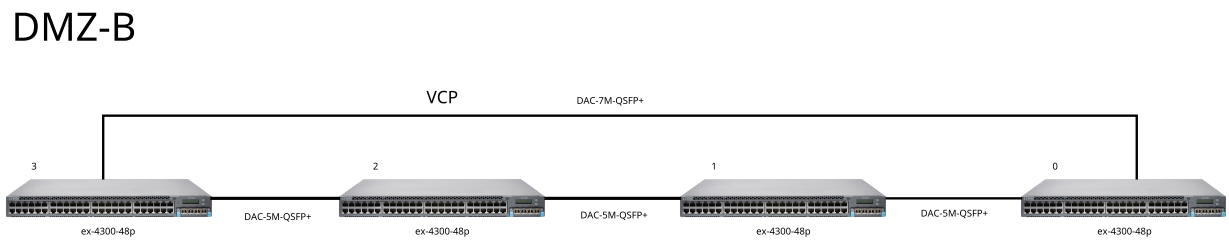
3.3 **Stack - DMZ-B**

The four **EX4300-48P** units were integrated into a single VC using four QSFP+ DAC cables.

This setup establishes a ring topology to maximize redundancy; in the event of a hardware failure, a secondary path is always available for packet forwarding.

The backplane switching throughput between the units within the Virtual Chassis matches the speed of the physical links used—in this deployment, 40 Gbps.

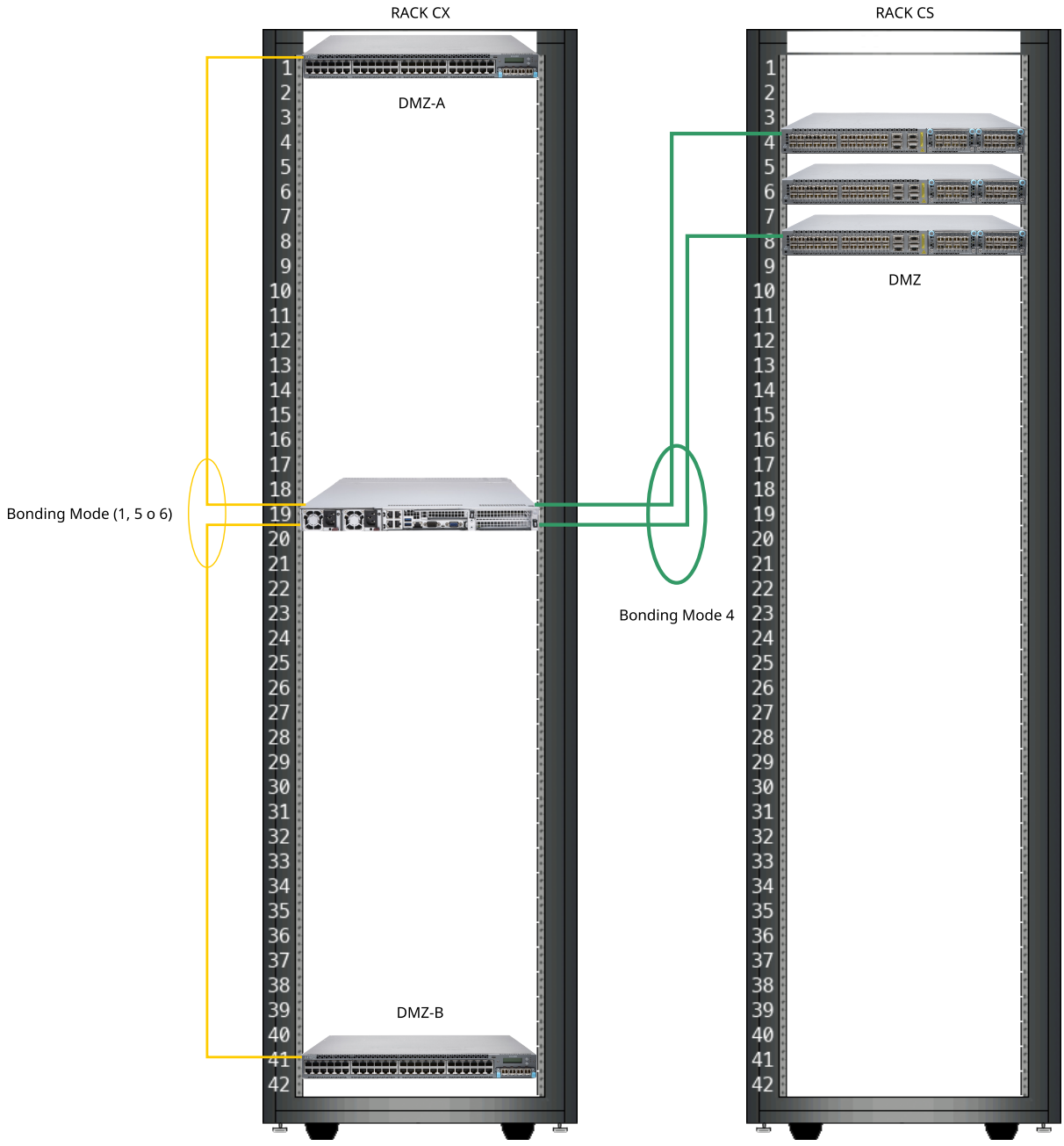
Figure 4: **DMZ-B - VC**



3.4 Server Connectivity

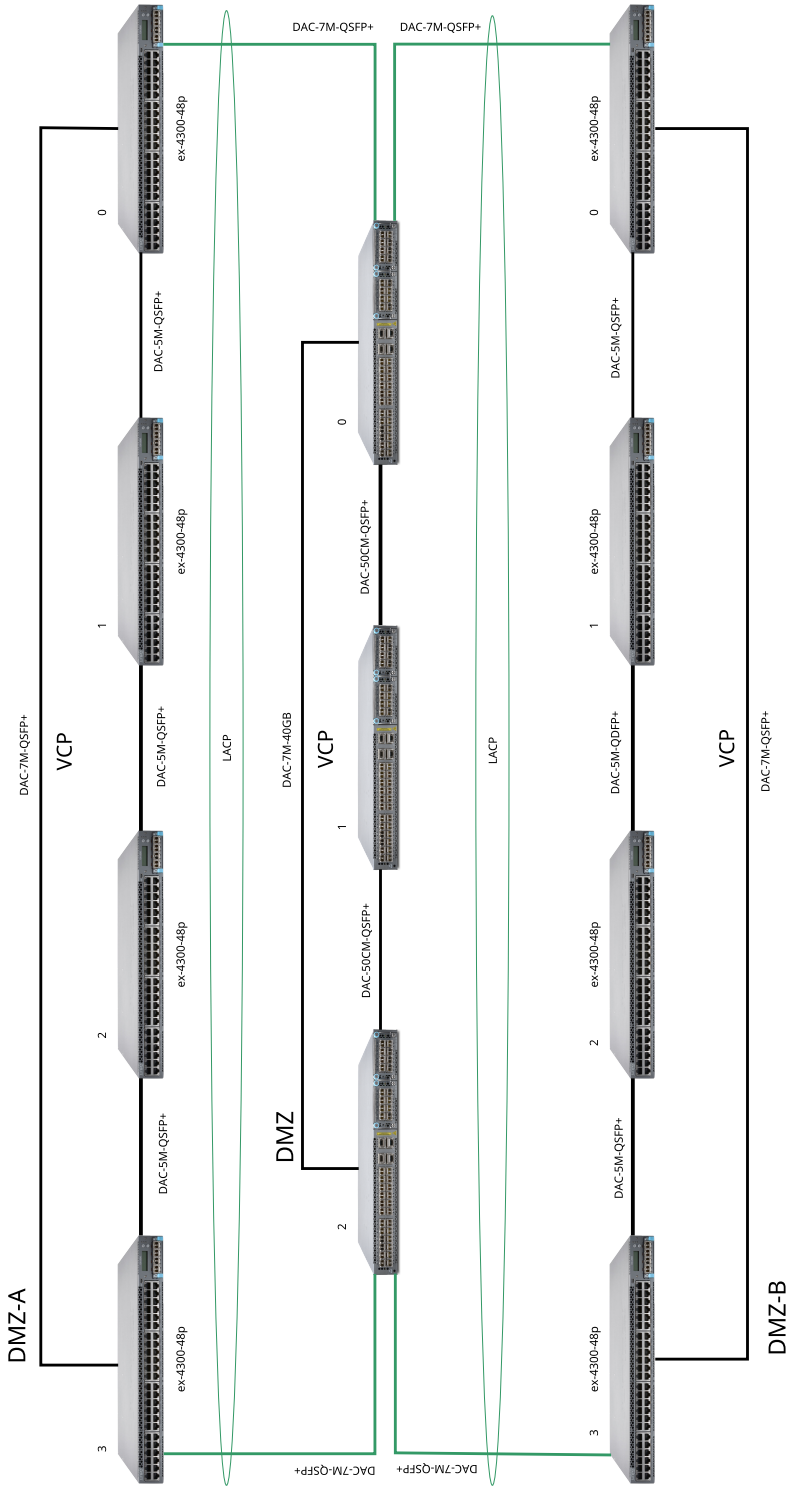
Servers must be equipped with at least two network ports to be used in a bonding configuration. If using copper interfaces, one port will be connected to the **DMZ-A** stack and the other to the **DMZ-B** stack, utilizing a Bonding Mode that does not require the 802.3ad protocol (specifically Modes 1, 5, or 6).
In the case of fiber connectivity, both ports will be connected directly to the **DMZ** (Core) stack, distributed across different stack members. For this setup, the 802.3ad protocol will be used with Bonding Mode 4.

Figure 5: **Server Connectivity**



3.5 **Network Diagram**

Figure 6: **Network Diagram**



4 Conclusions

The infrastructure design, based on Juniper Virtual Chassis technology, provides a robust, redundant, and highly scalable network environment. By strategically positioning the CS and the TOR stacks, we have achieved an optimal balance between physical cabling efficiency and logical management simplicity.

The implementation of security protocols, such as DHCP Snooping, BPDU Protection, and SSH key-based authentication, ensures a hardened environment against misconfigurations and unauthorized access. Furthermore, the adoption of JTAC-recommended software and the automation of configuration backups guarantee long-term stability and rapid recovery capabilities.

Finally, the chosen topology not only meets current connectivity requirements but also offers a clear expansion path. With the ability to scale both the core and access layers up to 10 units per stack, the network is fully prepared to support future growth within the DC.