> Anonymized portfolio / case-study version. Product/customer name, domains, hostnames, and public addresses have been removed; the architecture patterns are unchanged. Work sample, IT systems integration.

Case Study: Compute Layer on Proxmox VE (Multi-Tenant SaaS Platform)

> Compute-layer documentation for a multi-tenant SaaS platform (product anonymized). Single-server virtualization on a dedicated bare-metal server, cleanly structured along a VMID convention, with an isolated backup layer and a disaster-recovery runbook.

1. Executive Summary

Single-server setup on a dedicated bare-metal server (Hetzner) running Proxmox VE 9.1, installed natively as the hypervisor. **35+ guests in operation** — a deliberate mix of VMs (QEMU/KVM) and LXC containers, structured along a clearly defined VMID convention plus a dedicated, network-isolated backup layer.

VMs are used where genuine kernel isolation and hardware features are required (firewalls, database, workload hosts); LXC containers are used for lightweight services (backup server, CI runner, cert / docs / notify services) that do not need their own kernel and run with a small resource footprint.

The platform is self-operated (pre-launch): the infrastructure is in full production, while the SaaS offering built on top of it is not yet serving customers.

Quick Facts

Metric	Value
Dedicated server	Bare-metal at Hetzner, German data center (hostname anonymized)
Hypervisor	Proxmox VE 9.1 (kernel 7.0.0-3-pve), native
Active guests	35+ in operation (29 VMs + 8 LXC)
Public IPv4 addresses	4 total: 3 on the host (each its own Hetzner MAC) + 1 on the separate Cloud VPS Shield
Storage	ZFS (local pool for VMs + containers)
Host firewall	firewalld (default zone `public` = default-deny); pve-firewall NOT in use
Backup	Proxmox Backup Server (PBS) (LXC, isolated bridge), off-site, encrypted, restore-tested
Out-of-band access	Provider Robot KVM + out-of-band WireGuard to the host
Automation	Ansible (Infrastructure as Code)

2. Hardware Platform

Dedicated Server

Field	Value
Server type	Dedicated bare-metal server at Hetzner (hostname anonymized)
Location	German data center (location anonymized)
Hypervisor software	Proxmox VE 9.1 (kernel 7.0.0-3-pve), native as the hypervisor
CPU features	VT-x + nested virtualization enabled
Physical NIC	One physical NIC (designation anonymized)

Host Network Connectivity

The host has **one physical NIC**. Two logical paths run on top of it:

- a **main bridge for public traffic** (carries the three host-side public IPs via MAC-bound vNICs)
- a **Hetzner vSwitch** (VLAN 4000, MTU 1400) as a private Layer 2 to the cloud, connecting the sites/services across a provider-internal, segregated segment.

Each firewall transit additionally gets its **own dedicated Proxmox bridge** — there is deliberately **no shared L2** between the firewall VMs.

Public IP Assignment with MAC Filtering

Hetzner binds each additional public IP to a **separate MAC address** ("additional IP with dedicated MAC", anti-spoofing at the upstream-router level). When the VM is created, the MAC is hard-set on the vNIC. **Three** public IPs sit on the host this way; a **fourth** sits on the separate Cloud VPS Shield (different provider, its own MAC there) — **4 public IPv4** in total:

Public IP	MAC assignment	Assigned to
`<HOST-PUBLIC-IP>`	Host's own MAC	Proxmox VE host — **out-of-band management only**, isolated from the data path
`<EDGE-LAN-PUBLIC-IP>`	Separate MAC (provider-assigned)	Edge LAN firewall (VMID 100, OPNsense)
`<EDGE-DMZ-PUBLIC-IP>`	Separate MAC (provider-assigned)	Edge DMZ firewall (VMID 102, OPNsense)
`<SHIELD-PUBLIC-IP>`	Separate MAC (different provider)	Cloud VPS Shield — reverse ingress to the DMZ (separate provider location)

Important: the **host's own public IP serves out-of-band management exclusively** and is separated from the production data path. The hypervisor does not sit in the workloads' data flow.

Off-Site Backup Target

Field	Value
Backup target	Provider off-site storage (dedicated sub-account, dedicated password)
Protocol	CIFS/SMB mount over the backup path
Encryption	Encrypted PBS datastore repository (off-site)
Path PBS datastore	`/<sub-account>/pbs-datastore/` (on the off-site storage)
Path PVE mount	`/mnt/pbs-storagebox/` (on the PVE host)
Path LXC bind mount	`/mnt/datastore/storagebox` (inside the backup LXC)

Security Rationale for MAC Filtering

MAC filtering is the first line of defense ahead of our own firewall. Without the assigned MAC, a packet never even reaches the VM. This prevents: - IP spoofing from other customers in the same provider subnet - MAC flooding in the provider backbone - failover attacks on IP takeovers without a change of control

---

3. VMID Convention

Instead of numbering VMs in their order of creation, the platform follows a functional, dependency-based VMID convention. The VMID reveals at a glance which layer a machine serves.

Scheme

```
100-199    Firewall zone (production)
200-299    Pentest / security tools
300-399    Database layer
400-499    Development (dev hosts, test VMs)
500-599    Reserve
600-699    Backup / monitoring layer
700-899    Reserve (queue, storage, identity, once concrete)
900-999    Reserve / inactive
1000-1999  Application layer (HA-capable, broadly scalable)
2000-2999  Docker host cluster (HA-capable)
3000+      Future / additional tenant ranges
```

Rationale for the Ordering

Bottom-up by dependency: - Firewalls / security = network foundation - Database = data foundation (apps need databases) - Backup = cross-cutting, its own range, isolated - Application = on top, can scale - Docker hosts = container runtime, parallel to classic app VMs

Deliberate scaling decision: - 100-blocks for static components (FW/DB/Dev/Backup) - 1000-blocks for application + Docker: apps grow horizontally in HA setups and need plenty of room

---

4. Inventory (anonymized, grouped)

Excerpt from the inventory, representative per range. Service names are kept generic; product- and host-specific identifiers have been removed. In total, 35+ guests are running (29 VMs + 8 LXC).

Firewalls (100-103)

VMID	Role	Firewall engine	Public IP	Zone
100	Edge LAN	OPNsense	`<EDGE-LAN-PUBLIC-IP>`	Edge

VMID	Role	Firewall engine	Public IP	Zone
101	Core LAN, admin hub	pfSense	internal (RFC1918)	Core LAN
102	Edge DMZ	OPNsense	`<EDGE-DMZ-PUBLIC-IP>`	Edge
103	Core DMZ	pfSense	internal (RFC1918)	Core DMZ

Naming convention by function: Edge LAN = OPNsense, Edge DMZ = OPNsense, Core LAN = pfSense, Core DMZ = pfSense. A total of **4 firewall VMs** segment the inbound and internal traffic.

Pentest / Security Tools (200)

VMID	Purpose
200	Pentest VM, sits in the DEV VLAN

Database Layer (300-301)

VMID	Service	Network
300	Redis (cache, sessions, queue) on Debian, native VM instead of a container	internal /29
301	PostgreSQL (platform + internal apps), with row-level security	internal /29

Note on the architecture decision: critical infrastructure (PostgreSQL, Redis) deliberately runs as a native systemd VM, not as a container. This eliminates a container-daemon SPOF and makes VM snapshots in the backup trivial. PostgreSQL uses row-level security for tenant isolation.

Development (400-402)

VMID	Purpose
400	Development server 1
401	Development server 2 (main dev)
402	Generic test / sandbox VM

Backup Layer (600)

CT/VMID	Service	Network / Bridge
CT 600	Proxmox Backup Server (PBS) (LXC, unprivileged, Debian), datastore on off-site storage, backup target for all PVE guests	internal backup /29 (isolated bridge)

Architecture detail: - Container type: **unprivileged** (cleaner than privileged for a backup workload) - Features: nesting, FUSE, keyctl - Protection: enabled (no accidental deletion) - Resources: 2 vCPU, 2 GB RAM, 10 GB root disk on the ZFS pool - Mountpoint: `mp0` = bind mount, host `/mnt/pbs-storagebox/pbs-datastore` to LXC `/mnt/datastore/storagebox` - Web UI: reachable only via the isolated backup bridge, access via SSH tunnel or out-of-band WireGuard - SSH: active with key auth (LXC access via PVE host or VPN)

LXC was chosen deliberately here: a backup service does not need its own kernel, runs frugally on 2 GB RAM, and can be redeployed within minutes.

Reserve (900-901)

VMID	Status
900	Windows Server with the Hyper-V role, currently inactive (provisioned for possible later Windows testing). Note: on the first PBS backup a Microsoft UEFI 2011 cert warning appeared (expires mid-2026). On activation, run `qm enroll-efi-keys 900`.
901	Test firewall (experimental, for config tests before going into production)

Application Layer (1000-1003)

VMID	Service	VLAN assignment
1000	Backend (NestJS API)	APP_API (internal /29)
1001	Core service (control / coordination layer)	APP_CORE (internal /29)
1002	Frontend / landing	APP_WEB
1003	Office collaboration (Collabora Online)	OFFICE-WORKSPACE

Docker Host Cluster (2000)

VMID	Service	Container count
2000	Docker host (microservices via macvlan)	several dozen containers (reverse proxy, mail stack, SIEM, log aggregation, monitoring, messaging, tunnel daemon, and others)

LXC Services (lightweight)

Alongside the backup LXC, further LXC containers run lightweight services that do not need their own kernel: **CI runner, cert service, docs service, and notify service**. This separation (8 LXC alongside 29 VMs) keeps the setup resource-efficient without giving up the isolation of the critical workloads.

---

5. Proxmox Bridge Layout (overview)

No shared network between firewalls. Each firewall-to-firewall link is its own Proxmox bridge. In addition, there are dedicated admin bridges for a star-shaped management topology around the Core LAN firewall, as well as an isolated backup bridge for PBS.

WAN Connectivity + Physical

Bridge	Type	VLAN-aware	Purpose
(physical NIC)	Physical NIC	n/a	Provider uplink (one physical NIC)
`vmbr0`	Linux bridge	Yes	Main public bridge, carries the three host-side public IPs via MAC filtering
Hetzner vSwitch	VLAN interface	VLAN 4000	Private L2 to the cloud, MTU 1400

Backup / Management Bridge (isolated from production VLANs)

Bridge	Type	VLAN-aware	Purpose
`pbs01`	Linux bridge	No	Internal-only bridge for the PBS backup path (internal backup /29). Outbound via firewalld MASQUERADE through `vmbr0` (to the off-site storage). No L2 contact with production VLANs.

Security rationale: PBS sits at the same isolation level as PVE itself (management tool, not a production VM). A compromise of the PBS LXC does not yield L2 access into the production network, since `pbs01` has only a Layer 3 path through the PVE host (firewalld).

Rescue / Maintenance Bridges

Each firewall has a dedicated rescue bridge to which, in an emergency, a maintenance VM can be attached directly. Analogous to the physical setup, where you plug a laptop directly into a switch/router when there is a network problem.

Bridge	Purpose
`edge_lan`	Rescue port for the Edge LAN firewall
`edge_dmz`	Rescue port for the Edge DMZ firewall
`core_lan`	Rescue port for the Core LAN firewall
`core_dmz`	Rescue port for the Core DMZ firewall

Main Transits + VLAN Trunks

Bridge	Members	Purpose	Transit net
`br_lan01`	Edge LAN + Core LAN	Transit Edge LAN to Core LAN	dedicated /30
`br_dmz01`	Edge DMZ + Core DMZ	Transit Edge DMZ to Core DMZ	dedicated /30
`vmbr2_VLAN`	LAN service VMs	802.1Q trunk, Core LAN routes, service VMs with VLAN tag	n/a
`dmz_vlan01`	DMZ service VMs	802.1Q trunk, Core DMZ routes	n/a

Cross-Zone + Admin Bridges (star-shaped around the Core LAN firewall)

Bridge	Members	Purpose
`inter_fw`	Core LAN to Core DMZ	Cross-zone bridge for admin path + mail stream
`lan_opn`	Core LAN to Edge LAN	Admin bridge: Core LAN administers Edge LAN GUI/SSH
`dmz_opn`	Core LAN to Edge DMZ	Admin bridge: Core LAN administers Edge DMZ GUI/SSH

6. vNIC Assignment of the Firewall VMs

VM	vNICs / bridges
100 Edge LAN	`vmbr0` (WAN via MAC), `br_lan01` (transit), `lan_opn` (admin), `edge_lan` (rescue, inactive)
101 Core LAN	`br_lan01` (transit), `vmbr2_VLAN` (LAN trunk), `inter_fw` (cross-zone), `lan_opn` + `dmz_opn` (admin), `core_lan` (rescue, inactive)
102 Edge DMZ	`vmbr0` (WAN via MAC), `br_dmz01` (transit), `dmz_opn` (admin), `edge_dmz` (rescue, inactive)
103 Core DMZ	`br_dmz01` (transit), `dmz_vlan01` (DMZ trunk), `inter_fw` (cross-zone), `core_dmz` (rescue, inactive)
CT 600 (PBS)	`pbs01` only (internal backup /29)

7. Storage / ZFS

The host uses **ZFS** as the storage backend for VMs and containers — with snapshots, replication capability, and data integrity at the block level.

Pool	Type	Purpose
`local`	Directory	Templates, local snippets, dump directory
`local-zfs`	ZFS	ZFS storage pool for VMs + containers, primarily for production guests
`pbs-storagebox`	Proxmox Backup Server	Active backup target (chunk-based dedup, incremental, off-site, encrypted)

Recommendations

- **Primary for VMs/CTs:** the ZFS pool (snapshots, replication-capable)
- **Backups:** `pbs-storagebox` (PBS on CT 600), in production and restore-tested

8. Backup / Disaster Recovery (Proxmox Backup Server)

Current State (production)

Primary backup stack: - **CT 600:** Proxmox Backup Server as an unprivileged LXC - **Datastore:** chunk-based, encrypted dedup repository on the off-site storage - **PVE storage `pbs-storagebox`:** API-token-based push from PVE to PBS (token with the `DatastoreBackup` role)

Architecture:

```
[PVE host] --pbs01 (isolated backup bridge)-- [CT 600 PBS]
                                        |
                                        v (CIFS via vmbr0 + firewalld MASQ)
                                  [off-site storage]
                                        |
                                        v
                                  /pbs-datastore/.chunks/  (encrypted)
```

There is **no L2 path** from the backup LXC to the production VLANs — the backup layer is physically and logically separated from the data path.

Schedules:

Job	Schedule	Run	What
Backup	`02:00` daily	PVE datacenter job	vzdump of all production VMs/CTs to PBS
Prune	`*-*-* 03:00` daily	PBS datastore-internal	delete old snapshots per retention
Garbage collection	`sat 04:00` weekly	PBS datastore-internal	clean up orphaned chunks

Retention (configured at the PBS datastore level, not in the PVE backup job): - `keep-daily: 7` - `keep-weekly: 4` - `keep-monthly: 6` - `keep-yearly: 1` (optional)

Backup job selection: - All firewalls (100-103) - Databases (300, 301) - Dev VMs (400-402) - Application layer (1000-1003) - Docker host (2000) - Optional: pentest VM (200), reserve VM (900) - **Not** the PBS LXC itself: a self-backup is pointless, the datastore lives externally anyway, and the LXC can be redeployed in about 10 minutes.

Per-Service Backups (additional, defense-in-depth)

Component	Tool	Frequency	Retention
PostgreSQL	pg_dump + WAL archiving	hourly / continuous	RPO 15 min
Redis	AOF snapshots	continuous	RPO < 1 min
Docker volumes	Restic	daily	14 daily + 6 weekly
Mail stack	Native + WORM archive	daily + real-time	10 years (GoBD)

Off-Site Status

Currently: the off-site storage target resides in a physically separate provider data center; the backups are encrypted and restore-tested. This covers the 3-2-1 pattern for the current single-server setup. True provider redundancy (a second, independent provider) is not yet established and is earmarked as an expansion stage (e.g., a second PBS target via an S3 endpoint or rclone sync; alternatively, a self-hosted PBS at a second physical site, connected via a WireGuard tunnel).

Disaster-Recovery Runbook

Scenario A: PVE host completely lost, off-site storage intact 1. Set up a new PVE host (Robot rescue to reinstall) 2. Rebuild the `pbs01` bridge (see Section 5) 3. Redeploy CT 600 4. CIFS mount + bind mount as in the original setup 5. PBS finds the existing chunks on the off-site storage, datastore re-import via `proxmox-backup-manager datastore add ...` 6. Restore the VMs from PBS

Scenario B: PBS LXC broken, off-site storage intact - Steps 3-5 above. The datastore chunks remain on the off-site storage.

Scenario C: Off-site storage lost - Currently the painful case (no second, independent off-site yet). This is exactly why provider redundancy is on the expansion list.

RTO / RPO

Category	RTO	RPO
Full system (all VMs, from PBS)	< 4 hours	24 hours
Single-VM restore (from PBS)	< 30 minutes	24 hours
File-level restore from a VM backup (PBS feature)	< 10 minutes	24 hours
PostgreSQL (point-in-time)	< 1 hour	15 minutes
Mail stack	< 2 hours	24 hours
Firewall configuration (XML export)	< 15 minutes	before every change (manual export)

---

9. Host Firewall + Out-of-Band Management (one-way trust)

The host protects itself — independently of the firewall VMs

The host **does host 4 firewall VMs (OPNsense/pfSense)**, but it does **not** rely on them for its own protection. This is a deliberate **one-way trust**:

- **Host firewall: firewalld**, default zone `public` = **default-deny**. The only publicly reachable service is **WireGuard UDP**.
- **Management exclusively over the WireGuard tunnel** (zone `trusted`) — out-of-band, independent of the production firewall VMs.
- `pve-firewall` **is NOT in use** — the host is secured entirely via firewalld, a deliberate reduction to a single, clearly auditable firewall layer at the host.

This way the hypervisor stays reachable and protected even if a firewall VM fails or is compromised. The host's protection does not depend on the VMs it hosts.

Three Separate Management Paths

Path	Function	When to use
Provider Robot KVM	Hardware console, rescue system, BIOS	Server crash, boot problems, disk recovery
PVE Web UI (only over WireGuard)	Hypervisor lifecycle, VM console	Normal admin access
Out-of-band WireGuard	Encrypted remote access to the PVE Web UI and PBS Web UI	Admin access on the go

PBS Access (via VPN or SSH tunnel)

The PBS Web UI sits on the isolated backup bridge and is not directly reachable from the internet. Two access paths:

Option A, via out-of-band WireGuard (`wg0`): - Configure a WireGuard client with a route to the internal backup /29 through the host VPN gateway - Browser: `https://<pbs-internal-ip>:8007` - Prerequisite: firewalld `wg0` in the trusted zone

Option B, SSH tunnel forwarding (ad hoc):

```
ssh -L 8007:<pbs-internal-ip>:8007 root@<pve-host>
# Browser: https://localhost:8007
```

Security Principle

PVE has no network path to the Core LAN firewall or to the production VMs. It can lifecycle-control the VMs (start/stop/migrate) but cannot interfere with their production traffic. In a PVE compromise the VMs can be shut down, but the production data is not directly accessible. The same applies to the PBS LXC: it sits on the separate `pbs01` bridge and has no L2 path to the production VMs.

---

10. Nested Virtualization (Reserve VM 900)

VMID 900, currently inactive, provisioned for possible later use (e.g., Windows agent testing, domain-join testing, Windows-specific software testing).

Prerequisites / TODOs on activation: - PVE CPU flag `+vmx` or `host` for Intel VT-x passthrough (active) - Windows Server with the Hyper-V role (license required) - Nested-virtualization setup in the PVE VM config - EFI cert update before activation: `qm enroll-efi-keys 900` (the Microsoft UEFI 2011 cert expires mid-2026, a new cert is required) - The EFI/TPM disks still carry old disk designations from the VMID migration; on activation, optionally normalize them via `zfs rename` (cosmetic)

---

11. VMID Migration Pattern

In the course of introducing the new VMID convention, several VMs were migrated, bottom-up by dependency (firewall, then DB, then dev, then app, then Docker host). The migration ran script-driven on the PVE host.

Standard pattern per VM:

```
# 1. Shut down the VM
qm shutdown <old-id>

# 2. Rename the config file
mv /etc/pve/qemu-server/<old-id>.conf /etc/pve/qemu-server/<new-id>.conf

# 3. Rename the ZFS disk volumes
zfs rename rpool/data/vm-<old-id>-disk-0 rpool/data/vm-<new-id>-disk-0

# 4. Update the disk references in the config
sed -i "s/vm-<old-id>-disk/vm-<new-id>-disk/g" /etc/pve/qemu-server/<new-id>.conf

# 5. Start the VM + verify
qm start <new-id>
qm status <new-id>
ping <vm-ip>
```

IPs stayed unchanged, so no application connection strings had to be adjusted. Only the PVE-internal administration changed.

---

12. PVE CLI Cheat Sheet (excerpt)

VM and LXC Management

```
qm list                                 # status of all VMs
qm start <vmid> / qm shutdown <vmid>    # start / ACPI shutdown
qm clone <vmid> <new-vmid> --name <name> --full
qm resize <vmid> <disk> +<size>G        # grow disk

pct list                                # status of all LXC
pct enter <ctid>                        # enter LXC
pct exec <ctid> -- <command>            # run a command in the LXC without logging in
pct set <ctid> -protection 1            # enable anti-delete
```

Storage / ZFS

```
pvesm status                            # list storage pools
zfs list                                # ZFS datasets
```

Backup (PBS-based)

```
# Send a backup to PBS manually
vzdump <vmid> --storage pbs-storagebox --mode snapshot --compress zstd

# PBS datastore status (from the PVE host, delegated into the LXC)
pct exec 600 -- proxmox-backup-manager datastore show <datastore>

# Restore a VM from PBS via the PVE CLI
qmrestore <pbs-storage>:<backup-id> <new-vmid>
```

Network / Bridges / Host Firewall

```
bridge vlan show                        # check the bridge VLAN filter
firewall-cmd --get-active-zones         # firewalld zones (public / trusted)
firewall-cmd --zone=public --list-all   # default-deny + allowed services (WireGuard UDP)
firewall-cmd --zone=public --query-masquerade
```

CIFS Mount Diagnostics (for the off-site storage)

```
mount | grep cifs                       # mount status
umount /mnt/pbs-storagebox && mount -a   # re-mount after maintenance
ls -la /mnt/pbs-storagebox/pbs-datastore/ # verify UID/GID mapping
```

---

13. Compliance References (BSI IT-Grundschutz)

Compute-layer-relevant building blocks:

BSI building block	Description	Implementation
SYS.1.5	Virtualization	Proxmox VE 9.1 native, dedicated bridges per transit, MAC binding, isolated backup bridge
SYS.1.3	Servers under Linux	Debian on all VMs/LXC
NET.1.1	Network segmentation	firewalld default-deny at the host, no shared L2 between firewalls, out-of-band WireGuard management
OPS.1.1.2	Backup/Restore	PBS in production (CT 600), encrypted + off-site + restore-tested, DR drill quarterly (planned)
OPS.1.1.3	Naming convention	VMID scheme (bottom-up by function) + functional names
CON.3	Data backup	Multi-tier backup (PBS off-site encrypted + per-service tools + WORM for mail), provider redundancy earmarked as an expansion stage

---

14. Skills Applied (summary)

- **Proxmox VE 9.1** as the native hypervisor: VM and LXC lifecycle, ZFS storage, bridge layout, nested virtualization. Deliberate VM/LXC split (29 VMs + 8 LXC = 35+ guests) according to isolation and resource needs.
- **Network segmentation at the hypervisor level:** dedicated bridges per firewall transit (no shared L2), 802.1Q VLAN trunks, Hetzner vSwitch (VLAN 4000, MTU 1400) as a private L2, an isolated backup bridge with no L2 path into the production network, a star-shaped admin topology.
- **Host hardening with one-way trust:** firewalld (default-deny, public; pve-firewall deliberately not used), management exclusively over out-of-band WireGuard — the host protects itself independently of the 4 firewall VMs it hosts.
- **Backup / DR with Proxmox Backup Server:** unprivileged LXC, chunk-based dedup, encrypted off-site datastore, retention policy, a tested disaster-recovery runbook with defined RTO/RPO targets.
- **Infrastructure as Code:** Ansible for reproducible, automated provisioning and configuration.
- **Structured inventory management:** dependency-based VMID convention, script-driven migration without changing the application connection strings.
- **Out-of-band management:** three separate management paths, VPN-only access to sensitive UIs.
- **Compliance awareness:** mapping to BSI IT-Grundschutz building blocks, GoBD-compliant mail retention.

Anonymized work sample, IT systems integration. Product-, customer-, host-, and provider-specific identifiers, as well as public and internal addresses, have been removed; the architecture patterns are unchanged.