



CnuasSwitch, Datasheet

Virtual top of rack switch, RoCEv2 and native InfiniBand, in switch Subnet Manager

Document DS-CNU-002 • **Revision** C • **Issued** 12 August 2026 • **Status** Published

Item	Value
Part	cnuas-vswitchd
Type	Soft top of rack switch, RoCEv2 and native InfiniBand, runs as a host daemon
Version	v0.1.0-8-gd630c1f
Repo	PacketFive/CnuasSwitch

1. Overview

CnuasSwitch is the emulated top-of-rack switch of the Cnuas fabric. The `cnuas-vswitchd` daemon presents a 10-port switch that forwards both **RoCEv2** (UDP/4791 over emulated Ethernet) and **native InfiniBand** (LRH/BTH) traffic between attached CnuasNIC endpoints. It includes an **in-switch Subnet Manager (SM)** and Subnet Management Agent (SMA), so an InfiniBand fabric becomes operational without an external `opensm`.

Deployment modes

`cnuas-vswitchd` is an **ordinary host process**. It is not a QEMU device and never runs inside a guest. Each port is an `AF_UNIX SOCK_SEQPACKET` listening socket, so **any** host process that can open a UNIX socket can attach to a port and exchange frames. A QEMU guest running CnuasNIC is one such client, not a requirement.

Mode	What attaches to a port	Needs a guest	Typical use
Soft switch (host only)	Any host process speaking the port framing	No	Protocol development, forwarding and DCB work, continuous integration
Rack fabric	A guest with CnuasNIC, over the QEMU device	Yes, for the endpoints	Full rack deployments with unmodified verbs applications

This makes CnuasSwitch a **hybrid soft RoCE and soft InfiniBand RDMA switch**: one daemon serves both link layers, with the personality selectable per port. The host-only mode is what the switch test suite uses. It drives forwarding, classification, DCB and the InfiniBand pipeline from host processes with no virtual machine present.

Key features

- Runs entirely as a host daemon; no hypervisor, guest or kernel module.
- 10-port fixed switch: 8 fabric ports + uplink + console/observability.
- Dual-personality forwarding: RoCEv2 and native InfiniBand on the fabric ports.
- Per-port personality: RoCE, InfiniBand, hybrid, or stacked.
- Data Center Bridging: PFC (802.1Qbb), ETS (802.1Qaz), ECN (RFC 3168).
- Ethernet forwarding database (FDB) and InfiniBand Linear Forwarding Table (LFT).
- Built-in Subnet Manager with directed-route SMPs, LID assignment, and sweeps.
- Single-threaded `epoll` event loop for deterministic, teachable behaviour.
- JSON management interface with 18+ commands; OpenTelemetry export.



2. Port model

Diagram available in the online documentation at <https://github.com/PacketFive/cnuas>.

3. Functional specifications

Parameter	Value
Total ports	10 (CNUAS_TOTAL_PORTS)
Fabric ports (RoCEv2 + IB)	8 (ports 0, 7)
Uplink port	1 (port 8, Ethernet)
Console / OTel port	1 (port 9, observability)
Per-port personality	RoCE / InfiniBand / hybrid / stacked
Supported protocols	RoCEv2 (UDP dst 4791), native InfiniBand
Max frame / path MTU	Up to 9216 B (jumbo); IB path MTU up to 4096 B
DCB features	PFC (IEEE 802.1Qbb), ETS (IEEE 802.1Qaz), ECN (RFC 3168)
Ethernet forwarding	MAC FDB (add/del/dump)
InfiniBand forwarding	LID Forwarding Table (LFT), ~256 entries
LID addressing	16-bit LIDs
Subnet Manager	In-switch SM + SMA (GET/SET, directed-route SMPs, sweeps)
Concurrency model	Single-threaded epoll loop, MAX_EVENTS = 32, 1000 ms tick

4. Daemon internals

Subsystem	Source	Responsibility
Event loop	epoll_loop.c	epoll_loop_init() / epoll_loop_run(); per-port and mgmt FDs
Management	mgmt.c	JSON command parse/dispatch, responses
Subnet Manager	ib_sm.c	ib_sm_init(), ib_sm_trigger(), sm_build_dr_smp(), sweeps
Subnet Mgmt Agent	ib_sma.c	sma_set_node_info/port_info/switch_info/pkey_table/guid_info
Telemetry	otel.c	OpenTelemetry counter export

Four file descriptor context tags drive the epoll loop: FD_PORT_LISTENER, FD_PORT_DATA, FD_MGMT_LISTENER, and FD_MGMT_CLIENT.

Each port carries its own state, holding the MAC address, personality and mode, link state, VLAN, priority flow control queues, enhanced transmission selection configuration, explicit congestion notification configuration, and counters for frames and bytes in each direction, drops, pause frames sent and received, and frames marked for congestion.

5. Management command set (JSON)

Group	Commands
Port	port_status, set_port_mode, set_link_state
QoS / DCB	set_pfc, set_ecn, set_ets
Ethernet FDB	fdb_add, fdb_del, fdb_dump



Group	Commands
InfiniBand LFT	<code>lft_add</code> , <code>lft_del</code> , <code>lft_dump</code>
Subnet Manager	<code>sm_enable</code> , <code>sm_disable</code> , <code>sm_trigger</code> , <code>sm_status</code>
Telemetry	<code>telemetry_dump</code> , <code>telemetry_clear</code>

Response format: `{"status": "ok" | "error", ...}`.

6. Interfaces

Interface	Description
Port transport	UNIX domain sockets (SOCK_SEQPACKET), one per port
Management	JSON over UNIX socket at <code>/var/run/cnuas/mgmt.sock</code>
Run directory	<code>/var/run/cnuas/</code>
Observability	OpenTelemetry export on console port

7. Software support

Item	Value
Daemon	<code>cnuas-vswitchd</code> (C / GNU C)
Management CLI	<code>cnuas-cli</code> (port status, SM, DCB configuration)
Dashboard	<code>cnuas-webui</code>
Design reference	Virtual Switch Design

8. Typical usage

```
# Run the switch daemon
./build/cnuas-vswitchd --run-dir /var/run/cnuas --foreground

# Inspect ports and drive the Subnet Manager
cnuas-cli port status
cnuas-cli sm trigger
```

9. Ordering / integration information

Item	Value
Repo	PacketFive/CnuasSwitch
Submodule path	<code>src/cnuasswitch</code> (in <code>PacketFive/cnuas</code>)
Version	v0.1.0
Language	C (GNU C)
Build	Make / Bazel

10. Revision history



Revision	Date	Notes
A	2026-07-05	Initial datasheet
B	2026-07-05	Added port diagram, daemon internals, full JSON command set (through I5e/I5f SM)
C	2026-08-12	Recorded that the switch daemon runs on the host with no guest, as a soft RoCE and InfiniBand switch

PacketFive, Packet Five Networks Ltd., Dublin, Ireland.

Cnuas Virtual AI/HPC Infrastructure is published at <https://github.com/PacketFive/cnuas> under the Apache License 2.0; read it at <https://github.com/PacketFive/cnuas/blob/main/LICENSE>.

Cnuas Virtual AI/HPC Infrastructure is emulation software. It is not affiliated with, endorsed by, or derived from any hardware vendor, and every device it models is a software artefact. The work is published by its authors in a personal capacity and is not sponsored or endorsed by any employer.

Specifications describe the referenced revision of the software and may change without notice. Contact info@packetfive.com.