Contribution ID: **239**   Type: **not specified**

Linux Networking with MANA: RX Path Optimization and Netshaper Bandwidth Control.

Saturday 13 December 2025 17:35 (40 minutes)

Azure VMs have historically run on x86-64 with a 4 KB base page size, but newer ARM64-based SKUs such as ND GB200-v6 introduce configurable base page sizes (4 KB, 16 KB, 64 KB). While this architectural flexibility is great for workloads like model training, it exposed a major inefficiency in the MANA (MICROSOFT AZURE NETWORK ADAPTER) driver's RX path: RX buffers were allocated per page, so a single 1500-byte packet could consume an entire 64 KB page. This led to severe memory waste and left performance on the table. In this talk, we will present a recent Linux kernel optimization for MANA's RX path that uses page_pool fragments and pre-DMA-mapped page pools per RX queue. This allows multiple packets to share a single page, cutting per-packet memory overhead by over 97% and improving throughput up to ~200% on 64 KB ARM64 systems, while also delivering ~50% memory savings and ~15% throughput gains on x86-64.

In the second part of the talk, we will shift focus from raw RX efficiency to fine-grained bandwidth control using the new netshaper infrastructure. Existing tools like ethtool are too static, and tc operates at a higher abstraction level, making it difficult to express driver-specific shaping policies. Netshaper fills this gap with a lightweight, YAML-defined Netlink interface for driver-level bandwidth clamping. We will walk through integrating netshaper into the MANA driver, how this effort uncovered and resolved multiple potential deadlocks and missing netdev_ops, and how new iproute2 commands now make netshaper usable without kernel-tree tools. Attendees will gain both a deep dive into high-performance RX design for large PAGE_SIZE systems and a practical guide to modern, maintainable bandwidth control in Linux networking.

Primary authors: Mr ROY, Dipayan (Microsoft); ERNI, Sri Satya Vennela

Presenters: Mr ROY, Dipayan (Microsoft); ERNI, Sri Satya Vennela

Session Classification: Networking Track

Track Classification: Networking Track