CLLM 4.0

Making reliable low latency multicast messaging even more predictable, deterministic and platform-agnostic

Presenter: Stefan Ott

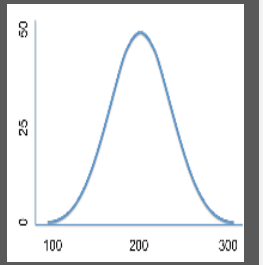
CONFINITY LOW LATENCY MESSAGING (CLLM)

Confinity LLM is the successor product for IBM® WebSphere® MQ Low Latency Messaging.

"It is designed for financial institutions and other organizations that require **near instantaneous and reliable delivery of (extremely large volumes) of data**.

CLLM provides flexible message delivery options like RDMA, MBU, IPoIB and RoCE combined with high system availability and congestion control."

Based on **Publish / Subscribe** paradigm, no queuing mechanism, no broker / dispatcher



Key differentiator in low latency messaging is the **predictability of latency** (steep Gauss curve)

Unique features for **reliable unicast and multicast** messaging (RUM/RMM) allow to create deterministic, stateful applications in distributed environments

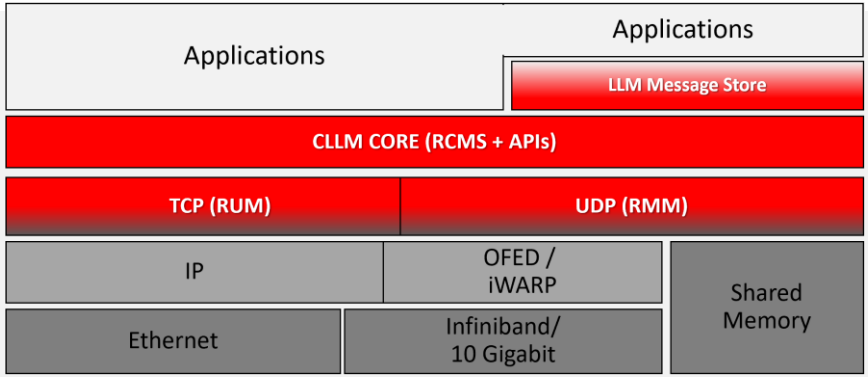
High performance and availability that help maintain high-quality service levels and protect the integrity of the data stream

Message control and filtering that make efficient use of system resources.

System monitoring and congestion controls that deliver messages with improved speed and reliability.

CLLM Architecture Redefined

- Confinity LLM consists of the following layers
  - CLLM Core (RCMS* + APIs)
  - Reliable Multicast Messaging (UDP Based)
  - Reliable Unicast Messaging (TCP Based)
  - Optional: Persistent Message Store



- A client application sends data through Ethernet or InfiniBand by using CLLM reliable mechanism.
- In case of Ethernet, an application's data path is traversing through following components:
  - CLLM Core
  - RMM or RUM
  - Kernel SCI
  - Kernel Network Stack
  - Ethernet Driver and 10G MAC

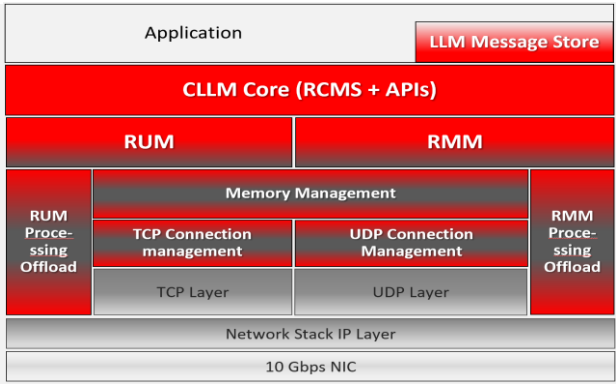
This Long Path adds Latency

*Reliable and Consistent Message Streaming

CLLM 4.0 - Hardware Accelerated Low Latency Messaging

CONFINITY SOLUTIONS

- FPGA Accelleration through Deep Pipelining
- RMM/RUM Offload on free running kernels
- CLLM Message processing at Ethernet line rates (clock cycles)
- Reduction of CPU kernel system calls
- Optimized and customized UDP/TCP kernels
- TCP processing latency speedup up to 20 times (*)
  (CPU 10us vs. FPGA 500ns)
- Ethernet interface efficiency improvements up to 4-fold (*)



- Designed for Alveo U250 or U50 NIC cards from Xilinx
- Xilinx no-DMA Interface Support
  (Slave Bridge Memory mapped)
- Fast access to msg history cache via direct host DDR memory map
- Dynamic function eXchange (DFX) for update of DMA and utility functions without reboot; in-Band FPGA partial reconfig

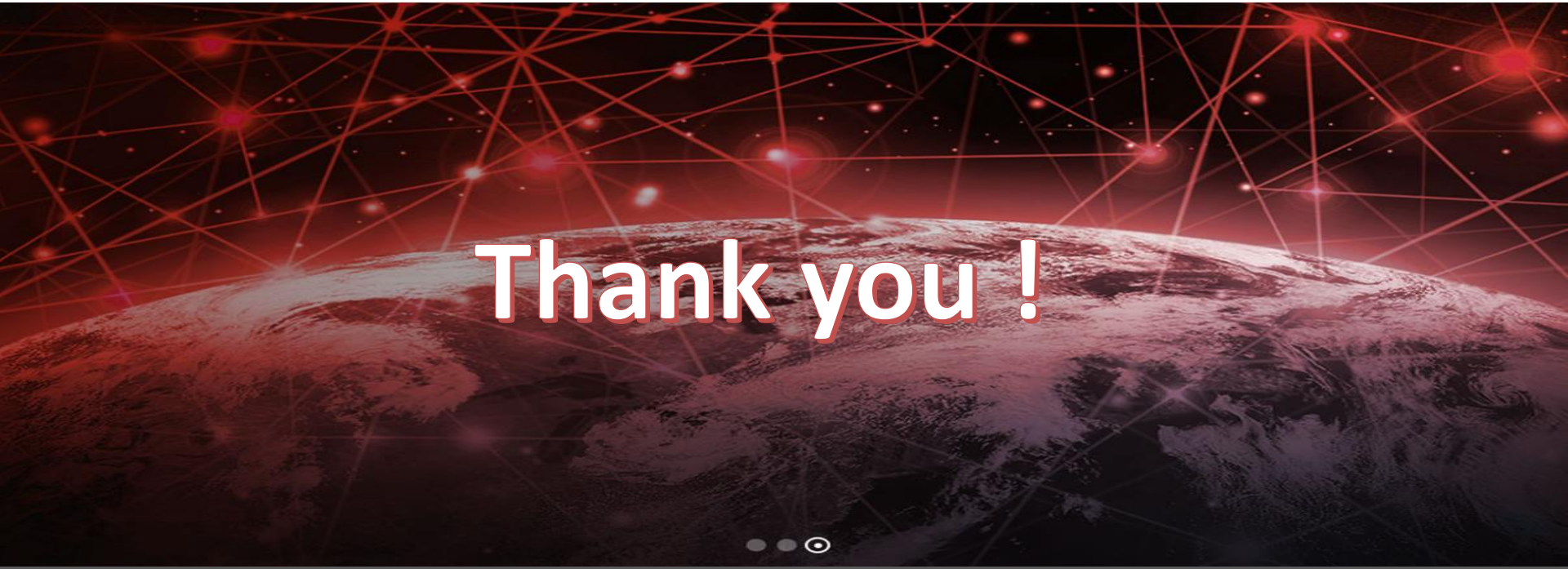
Alveo U50 NIC Board
- ✓ 33.3 Peak INT8 TOPS
- ✓ 872,000 LUTs
- ✓ 8GB HBM2 memory
- ✓ 1x40 GbE or 1x 100GbE (QSFP) or 4x10 GbE or 4x 25GbE (SFP)

(*) *Not a STAC benchmark*

ABOUT CONFINITY SOLUTIONS

- **Confinity Solutions GmbH** was established in 2016 to acquire source code of IBM Software:
  - Websphere MQ Low Latency Messaging and
  - Websphere Front Office

- **Confinity Solutions** is a FINTECH startup with mature products (12 years+), a strong partner ECO system and an established customer base worldwide



Office ESCHBORN
Mergenthalerallee 45-47
65760 Eschborn
Germany
+49 06196 97350 - 0

REGISTER
Registered in Frankfurt
HRB 106501
VAT ID DE309149885

CONTACTS
Stefan Ott, Managing Director
stefan.ott@confinity-solutions.com

WEB & E-MAIL
www.confinity-solutions.com
info@confinity-solutions.com