Proposal: A Cost-Effective AI Computing Network Solution with Flexilink and FPGA-Based RDMA Integration

1. Introduction: AI Computing Network and the Need for Two Networks

AI data centers require **high-performance computing (HPC) networks** to support large-scale deep learning training and inference tasks. These networks must handle two critical types of data traffic:

1. **Computing Network** – High-speed, low-latency inter-GPU communication for distributed AI training.
2. **Data Network** – Large-scale data storage and retrieval, supporting datasets, model weights, and inference results.

Current solutions include:

- **NVIDIA InfiniBand**: The de facto high-performance computing network, offering ultra-low latency and high bandwidth with **proprietary RDMA technology**.
- **Open RDMA over Converged Ethernet (RoCE)**: An alternative to InfiniBand that runs on standard Ethernet infrastructure, reducing cost while maintaining high performance.

Both networks play distinct roles:

- **Computing Network (InfiniBand or RoCE)** handles AI workloads, where GPUs must synchronize model parameters at high speeds.
- **Data Network (Ethernet or NVMe-oF)** ensures efficient data retrieval and model storage.

Our goal is to provide a **cost-effective alternative** that maintains low latency while reducing deployment costs, making AI computing more accessible.

---

2. RDMA Standards, Open Protocols, and Chip Suppliers

Remote Direct Memory Access (RDMA) enables direct memory-to-memory communication between devices **without CPU intervention**, significantly reducing latency and increasing bandwidth efficiency.

RDMA Protocols & Standards:

- **InfiniBand (IBTA Standard)**: Proprietary solution controlled by **NVIDIA (Mellanox)**.
- **RoCE (RDMA over Converged Ethernet)**:
  - **RoCE v1** – Layer 2 Ethernet-based RDMA, limited to local networks.
  - **RoCE v2** – Layer 3 (IP-based) RDMA, enabling long-distance scalability.
- **iWARP**: An alternative RDMA implementation over TCP/IP but with higher latency.

Major RDMA Chip Suppliers:

Supplier	RDMA Solutions	Notes
NVIDIA (Mellanox)	InfiniBand, RoCE (ConnectX-7, Quantum-2)	Market leader, but high cost
Broadcom	RoCE-based NICs	Open alternative, lower cost

Supplier	RDMA Solutions	Notes
Intel	Ethernet RDMA (E810 series)	Moderate performance, open-standard
Marvell	RoCE & iWARP NICs	Emerging player
AMD/Xilinx	FPGA-based RDMA acceleration	Programmable and customizable

By leveraging **open RDMA standards like RoCE**, we can build an alternative to InfiniBand using **FPGA-based networking solutions** with Flexilink's dynamic Time-Division Multiplexing (TDM) approach.

---

3. Flexilink's FPGA-Based Solution for AI Computing Networks

Flexilink, as a **low-latency networking protocol**, was originally developed for real-time multimedia streaming. It integrates a dynamic **TDM-based scheduling mechanism**, which provides guaranteed quality for time-sensitive traffic while also accommodating best-effort traffic.

How Flexilink Works for AI Networks

- **TDM-like scheduling**: Ensures predictable latency and bandwidth allocation for critical AI workloads.
- **Hybrid traffic management**: Time-critical AI synchronization traffic is prioritized, while other background tasks (e.g., storage access) are handled in available slots.
- **FPGA-based acceleration**: Our **FPGA network processing unit (NPU)** provides hardware-optimized packet handling with deterministic timing.

Our Proposed Solution

1. **FPGA-Based RoCE NICs**:
    - Implement **RDMA acceleration** with custom Flexilink scheduling.
    - Reduce latency **by optimizing memory access patterns** in AI workloads.
    - Offer **low-cost, high-performance alternatives** to NVIDIA's proprietary network cards.
2. **FPGA-Based RoCE Switches**:
    - **Custom soft-core switching** for efficient packet forwarding in AI clusters.
    - Supports **100GbE / 200GbE RoCE v2**, integrated with Flexilink's scheduling.
    - **Lower cost vs. traditional RoCE switches** (e.g., Broadcom Tomahawk).
3. **Integration with AI Compute Infrastructure**:
    - Works with **ICUBE AI servers** and other AI compute nodes.
    - Provides **NCCL-compatible networking** for deep learning frameworks like TensorFlow and PyTorch.

---

4. Advantages: Performance, Cost, and Market Differentiation

Performance Benefits

- **Lower latency than standard Ethernet**: Flexilink ensures **sub-microsecond timing precision**, improving GPU-to-GPU synchronization.
- **Better resource utilization**: Dynamic time slot allocation prevents congestion and packet loss.
- **Deterministic data transfer**: Unlike traditional Ethernet, which suffers from jitter, Flexilink ensures **consistent end-to-end latency**.

Cost Savings

- **30-50% cheaper than NVIDIA's InfiniBand** by leveraging **FPGA-based RDMA NICs & switches**.
- **Standardized Ethernet infrastructure** means compatibility with existing data center hardware, avoiding costly vendor lock-in.

Market Differentiation

- **Designed for mid-sized AI data centers** that need high-performance networking but cannot afford the premium cost of InfiniBand.
- **Flexible architecture** that supports **both RDMA traffic and best-effort traffic** in a unified network.
- **Integration with ICUBE AI compute servers**, providing a complete AI cluster solution.

---

5. Conclusion: A New Path for Cost-Effective AI Networking

The AI computing industry requires high-performance yet **cost-effective** networking solutions. While **InfiniBand dominates the high-end market**, its proprietary nature and high costs create a gap in the market.

Our **FPGA-based Flexilink solution** offers: ✅ **Lower-cost AI networking alternative** to InfiniBand.
✅ **Optimized RDMA performance via RoCE** (100GbE / 200GbE).
✅ **Seamless integration with AI frameworks** (NCCL, GPUDirect).
✅ **Dynamic traffic allocation using TDM principles**, reducing congestion.

By combining **Flexilink's low-latency protocol**, **FPGA-accelerated RoCE**, and **ICUBE AI compute servers**, we can deliver a **high-performance AI computing network** at a **fraction of the cost** of existing solutions.

🚀 **Next Steps**:

- **Work with ICUBE** to integrate RoCE RDMA support.
- **Develop FPGA-based RoCE NICs and switches**.
- **Benchmark Flexilink against InfiniBand and standard RoCE** to validate performance.

Appendix: Key Terms

- **RoCE (RDMA over Converged Ethernet)**: An RDMA protocol that runs over Ethernet, allowing direct memory access between servers without CPU intervention.
- **RDMA (Remote Direct Memory Access)**: A networking technology that enables fast data transfer with low latency.
- **NCCL (NVIDIA Collective Communications Library)**: A GPU communication library used for AI training, requiring RDMA compatibility.

- **Flexilink**: A TDM-based network protocol designed for low-latency multimedia and AI applications.
- **FPGA (Field-Programmable Gate Array)**: A customizable hardware platform for accelerating networking tasks.

💡 **Proposal Summary**

Goal: Develop a **low-cost, high-performance AI network** using **Flexilink + FPGA-based RoCE**.
Impact: Reduce AI cluster networking costs while maintaining **RDMA-grade performance**.
Next Steps: Collaborate with ICUBE, prototype FPGA RoCE NICs, and benchmark against existing AI networking solutions.

Would you like me to refine this further or add technical diagrams for clarity? 🚀