IEEE 802 Nendica Report:
Intelligent Lossless Data Center Networks

Editor

Name	Affiliation
Guo, Liang	CIACT/ODCC
Congdon, Paul	Huawei

Nendica Chair

Name	Affiliation
Marks, Roger	Huawei

Contributors/Supporters

Name	Affiliation
Li, Jie	CIACT/ODCC
Gao, Feng	Baidu
Gu, Rong	China Mobile
Zhao, Jizhuang	China Telecom
Chen, Chuansheng	Tencent
Yin, Yue	Huawei
Song, Qingchun	Mellanox
Lui, Jun	Cisco
He, Zongying	Broadcom
Sun, Liyang	Huawei

Trademarks and Disclaimersa

IEEE believes the information in this publication is accurate as of its publication date; such information is subject to change without notice. IEEE is not responsible for any inadvertent errors.

Comments on this report are welcomed by Nendica: the IEEE 802 "Network Enhancements for the Next Decade" Industry Connections Activity: <https://1.ieee802.org/802-nendica>
Comment submission instructions are available at: <https://1.ieee802.org/802-nendica/nendica-dcn>
Comments will be considered for a future revision of this document and may stimulate the development of a new revision.

The Institute of Electrical and Electronics Engineers, Inc.
3 Park Avenue, New York, NY 10016-5997, USA

PDF: ISBN xxx-x-xxxx-xxxx-x XXXXXXXXXX

IEEE prohibits discrimination, harassment, and bullying. For more information, visit http://www.ieee.org/web/aboutus/whatis/policies/p9-26.html*.*

To order IEEE Press Publications, call 1-800-678-IEEE.
Find IEEE standards and standards-related product listings at: http://standards.ieee.org

NOTICE AND DISCLAIMER OF LIABILITY CONCERNING THE USE OF IEEE SA INDUSTRY CONNECTIONS DOCUMENTS

This IEEE Standards Association ("IEEE SA") Industry Connections publication ("Work") is not a consensus standard document. Specifically, this document is NOT AN IEEE STANDARD. Information contained in this Work has been created by, or obtained from, sources believed to be reliable, and reviewed by members of the IEEE SA Industry Connections activity that produced this Work. IEEE and the IEEE SA Industry Connections activity members expressly disclaim all warranties (express, implied, and statutory) related to this Work, including, but not limited to, the warranties of: merchantability; fitness for a particular purpose; non-infringement; quality, accuracy, effectiveness, currency, or completeness of the Work or content within the Work. In addition, IEEE and the IEEE SA Industry Connections activity members disclaim any and all conditions relating to: results; and workmanlike effort. This IEEE SA Industry Connections document is supplied "AS IS" and "WITH ALL FAULTS."

Although the IEEE SA Industry Connections activity members who have created this Work believe that the information and guidance given in this Work serve as an enhancement to users, all persons must rely upon their own skill and judgment when making use of it. IN NO EVENT SHALL IEEE OR IEEE SA INDUSTRY CONNECTIONS ACTIVITY MEMBERS BE LIABLE FOR ANY ERRORS OR OMISSIONS OR DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO: PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS WORK, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE AND REGARDLESS OF WHETHER SUCH DAMAGE WAS FORESEEABLE.

Further, information contained in this Work may be protected by intellectual property rights held by third parties or organizations, and the use of this information may require the user to negotiate with any such rights holders in order to legally acquire the rights to do so, and such rights holders may refuse to grant such rights. Attention is also called to the possibility that implementation of any or all of this Work may require use of subject matter covered by patent rights. By publication of this Work, no position is taken by the IEEE with respect to the existence or validity of any patent rights in connection therewith. The IEEE is not responsible for identifying patent rights for which a license may be required, or for conducting inquiries into the legal validity or scope of patents claims. Users are expressly advised that determination of the validity of any patent rights, and the risk of infringement of such rights, is entirely their own responsibility. No commitment to grant licenses under patent rights on a reasonable or non-discriminatory basis has been sought or received from any rights holder. The policies and procedures under which this document was created can be viewed at http://standards.ieee.org/about/sasb/iccom/.

This Work is published with the understanding that IEEE and the IEEE SA Industry Connections activity members are supplying information through this Work, not attempting to render engineering or other professional services. If such services are required, the assistance of an appropriate professional should be sought. IEEE is not responsible for the statements and opinions advanced in this Work.

TABLE OF CONTENTS

1 Introduction

Bringing the data center to life… or something else? I'm expecting this will be a short intro and the more detailed background intro is section 2.

Scope

The scope of this report includes...

Purpose

The purpose of this report is to ...

2 Bringing the data center to life

A new world with data everywhere

- ✓ Enterprise digital transformation needs more data for using AI
- ✓ Machine translation and search engines need to be able to process huge data simultaneously
- ✓ The era of internet celebrity webcast, all-people online games, data explosion
- ✓ Consumption upgrade in the new era of take-out, online takeout platform schedule and deliver massive orders
- ✓ The XX service of the carrier has higher requirements on data center network
- ✓ Data-based New World Requires Ubiquitous Data Center Technologies.

Today's data center enables the digital real-time world

- ✓ **AI related services**: AI cloud data center improves these applications performance: smart manufacturing/finance/energy/transportation (cloud data centers go to AI era. A cloud data center is more like a service support center. It is application-centric and uses the cloud platform to quickly distribute IT resources. The data center for AI services evolves into a business value center based on the cloud data center. The data center focuses on how to efficiently process data based on AI)

- ✓ **Distributed storage**: Stay ahead of rapid storage growth driven by new data sources and evolving technologies, a flexible storage efficiency is critical for customers to maximize the revenue of every bit. The development of high speed storage technology will help users to access the content more conveniently. Other data center technologies should be evolved together with distributed storage to ensure customers can obtain high input and output speed.

- ✓ **Cloud Database**: A cloud database may be a native service within a public cloud provider, or it may be a database from a cloud agnostic software vendor, designed for cloud architectures and requirements. Data centers make use of new technologies to address distributed cloud databases modern high performance application requirements.

3 Evolving data center requirements and technology

Technology evolution

- ✓ The development of fast storage provides necessary media for big data (distributed storage)
    - Storage performance needs to improve by an order of magnitude to achieve more than 1 million input/output operations per second (IOPS).
    - Communication latency has recently increased from 10% to 60% of storage E2E latency.
- ✓ Computing speed improvement (distributed computing)

AI computing model complexity is exploding

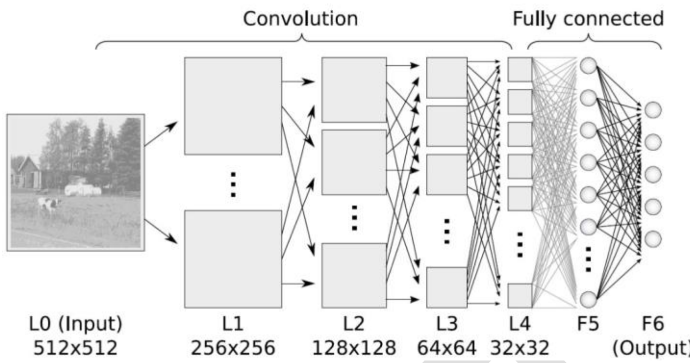
AI training is becoming increasingly complex with the development of services. For example, there are 7 ExaFLOPS and 60 million parameters in the Microsoft Resnet in 2015. The number came to 20 ExaFLOPS and 300 million parameters when Baidu trained their deep speech system in 2016. In 2017, the Google NMT used 105 ExaFLOPS and 8.7 billion parameters.

AI inference is the next great challenge so there must be an explosion of network design. The new characteristics of AI algorithm and huge computing workload require evolution of data center network.

Characteristics of AI computing

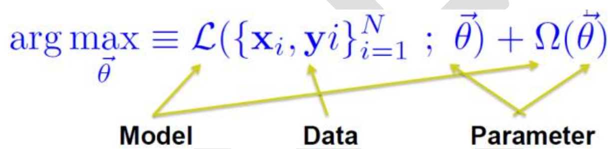
Traditional data center services (web, video, and file storage) are transaction-based and the calculation results are deterministic. For such tasks, there is no correlation or dependency between single calculation and network communication, and the occurrence time and duration of the entire calculation and communication are random. AI computing is based on target optimization and iterative convergence is required in the computing process, which causes high spatial correlation in the computing process of AI services and temporally similar communication modes.

Figure X Iterative machine learning network model



A typical AI algorithm refers to an optimization process for a target. The computing scale and features mainly involve models, input data, and weight parameters.

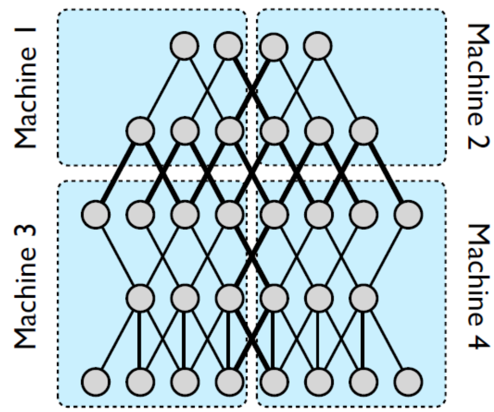
Figure X Machine learning algorithm



To solve the Big Data problem, the computing model and input data need to be large (for a 100 MB node, the AI model for 10K rules requires more than 4 TB memory), for which a single server cannot provide enough storage capacity. In addition, because the computing time needs to be shortened and increasingly concurrent AI computing of multiple nodes is required, DCNs must be used to perform large-scale and concurrent distributed AI computing.

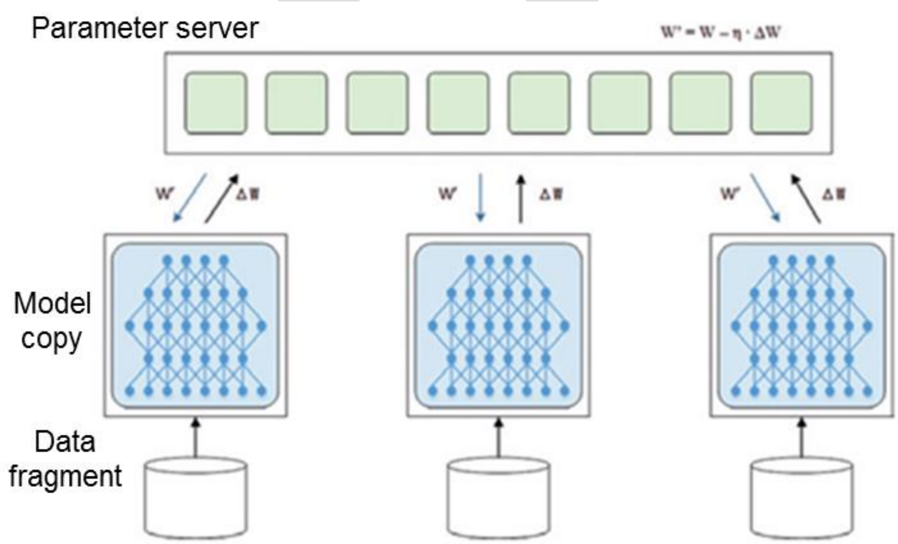
Distributed AI computing has the following two modes: model parallel computing and data parallel computing.
For model parallel computing, each node computes one part of the algorithm. After computing is complete, all data fragmented across models needs to be transferred to other nodes, as shown in Figure X.

Figure X Model parallel training



For parallel data computing, each node loads the entire AI algorithm model. Multiple nodes can calculate the same model at the same time, but only part of the input data is input to each node. When a node completes a round of calculation, all relevant nodes need to aggregate updated information about obtained weight parameters, and then obtain the corresponding globally updated data. Each weight parameter update requires that all nodes upload and obtain the information synchronously.

Figure X Data parallel training



No matter the development of distributed storage or distributed AI training, data center network comes to the communication pressure. The waiting time for GPU communication exceeds 50% of the job completion time [Omar, NANOG 76].

- ✓ Cloud-based AI platforms
  - Combines CPUs, storage and networking to simulate cognitive functions such as problem-solving, learning, reasoning, social intelligence
  - Data Center resource planning and utilization are critical to success
- ✓ SmartNIC become the computer in front of computer

- SmartNIC is a NIC with all NIC functions regardless CPU/FPGA. Host CPU only request to install NIC driver.
- SmartNIC is a computer in front of computer. SmartNIC has independent OS and is able to run some applications independently.
  - SmartNIC can be used to accelerate application
    - Accelerate computing, storage…
  - SmartNIC can be used to offload host CPU to run specific application more efficient
  - SmartNIC is part of computing resource. Participate the application computing together with host CPU and GPU.
    - Complement of CPU and GPU computing resource
    - SmartNIC is not the replacement of CPU and GPU, major applications still run on CPU/GPU
  - SmartNIC can be the independent domain than host domain and protect the host domain
    - Offload OVS to SmartNIC to isolate the data classification from hypervisor
  - SmartNIC can be emulated to other PCIe devices to support more advanced application
    - NVMe emulation
- SmartNIC is programmable and easy use
  - Open source software, major Linux
  - Easy to program, no special request for programmer
- SmartNIC is not proprietary NIC, one NIC fits many applications, easy for user to program

Network requirements

✓ New protocol is widely used for high performance application (Introduction RDMA related technologies, Cisco)

RDMA (Remote Direct Memory Access) is a new technology designed to solve the problem of server-side data processing latency in network applications, which transfers data directly from one computer's memory to another without the intervention of both operating systems. This allows for high bandwidth, low latency network communication and is particularly suitable for use in massively parallel computer environments. By transferring telegrams directly into the storage space of the other computer through the network, data can be quickly transferred from one system to the storage space of another system, reducing or eliminating the need for multiple copies of data telegrams during transmission, thus freeing up memory bandwidth and CPU cycles and greatly improving system performance.

RDMA's development in the transport layer/network layer currently goes through 3 technologies, Infiniband, iWarp and RoCEv1/RoCEv2.

Infiniband

In 2000, the IBTA (InfiniBand Trade Association) released the first RDMA technology, Infiniband, which is a customized network technology for RDMA multi-layered, new design from the hardware perspective to ensure the reliability of data transmission. The InfiniBand technology uses RDMA technology to provide direct read and write access to remote nodes. RDMA used InfiniBand as the transport layer in its early days, so it must use InfiniBan switches and InfiniBan network cards to implement.

iWarp (Internet Wide Area RDMA Protocol)

Internet wide area RDMA protocol, also known as RDMA over TCP protocol, is the IEEE/IETF proposed RDMA technology. It uses the TCP protocol to host the RDMA protocol. This allows RDMA to be used in a standard Ethernet environment (switch) and the network card requirement is an iWARP enabled network card. In fact iWARP can be implemented in software, but this takes away the performance advantage of RDMA.

RoCE (RDMA over Converged Ethernet)

In April 2010, the IBTA released RoCEv1, which was released as an add-on to the Infiniband Architecture Specification, so it is also known as IBoE (InfiniBand over Ethernet).The RoCE standard replaces the TCP/IP network layer with an IB network layer on top of the Ethernet link layer and does not support IP routing. The Ethernet type is 0x8915. in RoCE, the link layer header of the infiniband is removed and the GUID used to represent the address is converted to an Ethernet MAC. infiniband relies on lossless physical transport, and RoCE relies on lossless Ethernet transport.

RoCEv2

Since the RoCEv1 data frame does not have an IP header, it can only communicate within a 2-tier network. To solve this problem, in 2014 IBTA proposed RoCE V2, which extends RoCEv1 by replacing GRH (Global Routing Header) with a UDP header + IP header. Because RoCE v2 packets are routable at Layer 3, they are sometimes referred to as "Routable RoCE" or "RRoCE" for short. As shown in the figure below.

RoCE technology can be implemented through a common Ethernet switch, but the server needs to support RoCE network cards. Since RoCEv2 is a UDP protocol, although the UDP protocol is relatively high efficiency, but unlike the TCP protocol, there is a retransmission mechanism to ensure reliable transmission, once there is a packet loss, must rely on the upper layer of the application found and then do retransmission, which will greatly reduce the transmission efficiency of RDMA. So in order to play out the true effect of RoCE, it is necessary to build a lossless network environment for RDMA without losing packets.

RoCE can run in both lossless and compromised network environments, called Resilient RoCE if running in a compromised network environment, and Lossless RoCE if running in a compromised network environment.

RDMA is more and more widely used in market, especially in OTT companies. There have been tens of thousands of servers supporting RDMA, carrying our databases, cloud storage, data analysis systems, HPC and machine learning applications in production. Applications have reported impressive improvements by adopting RDMA [HPCC, Singcomm2019]. For instance, distributed machine learning training has been accelerated by 100+ times compared with the TCP/IP version, and the I/O speed of SSD-based cloud storage has been boosted by about 50 times compared to the TCP/IP version. These improvements majorly stem from the hardware offloading characteristic of RDMA.

✓ High I/O throughput with low-latency storage network

- As media access speeds increase, network latency becomes the bottleneck
- Storage interface protocols evolve from Serial Attached SCSI (SAS) to Non-Volatile Memory Express (NVMe)
- Reducing dynamic latency (latency from queuing and packet loss) is key to reducing the NVMe over Fabric latency

✓ Ultra-low latency network for distributed computing

DCN Requirement of distributed AI computing
As the number of AI algorithms and AI applications continue to increase, and the distributed AI computing architecture emerges, AI computing has become implemented on a large scale. To ensure enough interaction takes place between such distributed information, there are more stringent requirements regarding communication volume and performance. Facebook recently tested the distributed machine learning platform Caffe2, in which the latest multi-GPU servers are used for parallel acceleration. In the test, computing tasks on eight servers resulted in insufficient resources on the 100 Gbit/s InfiniBand network. As a result, it proved difficult to achieve linear computing acceleration of multiple nodes. The network performance greatly restricts horizontal extension of the AI system.

- Controlling the tail latency of these applications is critical. It must be measured in microseconds, not milliseconds

4 Challenges with today's data center network

High bandwidth and low latency tradeoff

✓ It's difficult to achieve high bandwidth and low latency simultaneously

✓ Experimentation shows the tradeoff still exists after varying algorithms, parameters, traffic patterns and link loads

- ✓ Reason explanation about why tradeoff exists

Deadlock free lossless network

- ✓ High-performance RDMA applications requires lossless network (Zero packet loss and low latency)
- ✓ Lossless Ethernet requires Priority-based Flow Control (PFC, in IEEE802.1Qbb)
- ✓ PFC storm may cause severe deadlock problem in data center
- ✓ Example deadlock problem in a CLOS network

Congestion control issues in large-scale data center networks

- ✓ How large scale today's data center is?
- ✓ Use cases for TCP and RoCE flows mixture
- ✓ Smart-buffer mechanisms in mainstream switch chips
- ✓ SLAs cannot be guarantee when TCP and RoCE traffic coexists

Configuration complexity of congestion control algorithms

- ✓ Tuning RDMA networks is an important factor to achieving high-performance
- ✓ Current method of parameters configuration can be a complex operation
- ✓ Congestion control algorithms usually requires collaboration between the NIC and switch
- ✓ Traditional PFC manual configuration needs complex calculation with lots of parameters
- ✓ Excessive headroom leads to reduce the number of lossless queues while too little headroom leads to packet loss

5 New technologies to address new data center problems

Approaches to PFC storm elimination

- ✓ Tuning RDMA networks is an important factor to achieving high-performance
- ✓ Current method of parameters configuration can be a complex operation

- ✓ Congestion control algorithms usually requires collaboration between the NIC and switch
- ✓ Traditional PFC manual configuration needs complex calculation with lots of parameters
- ✓ Excessive headroom leads to reduce the number of lossless queues while too little headroom leads to packet loss

Improving Congestion Notification

- ✓ Improved Explicit Congestion Notification
- ✓ Enhanced version of Quantized Congestion Notification (originally IEEE 802.1Qau)
- ✓ Intelligent Methods of improving QoS support in mixed traffic environments
- ✓ Test verification (ODCC lossless DCN test specification and result)

Intelligent congestion parameter optimization

- ✓ Intelligent heuristic algorithms for identifying congestion parameters
- ✓ Methods for dynamic optimization based on services
- ✓ Test verification (ODCC lossless DCN test specification and result)

Buffer optimization of lossless queues

- ✓ Intelligent headroom calculation
- ✓ Self-adaptive headroom configuration

6 Standardization considerations

Things for the IEEE 802 and IETF to consider. Possibly others as well – SNIA, IBTA, NVMe, etc..

7 Conclusion

Closing words…

8

Citations