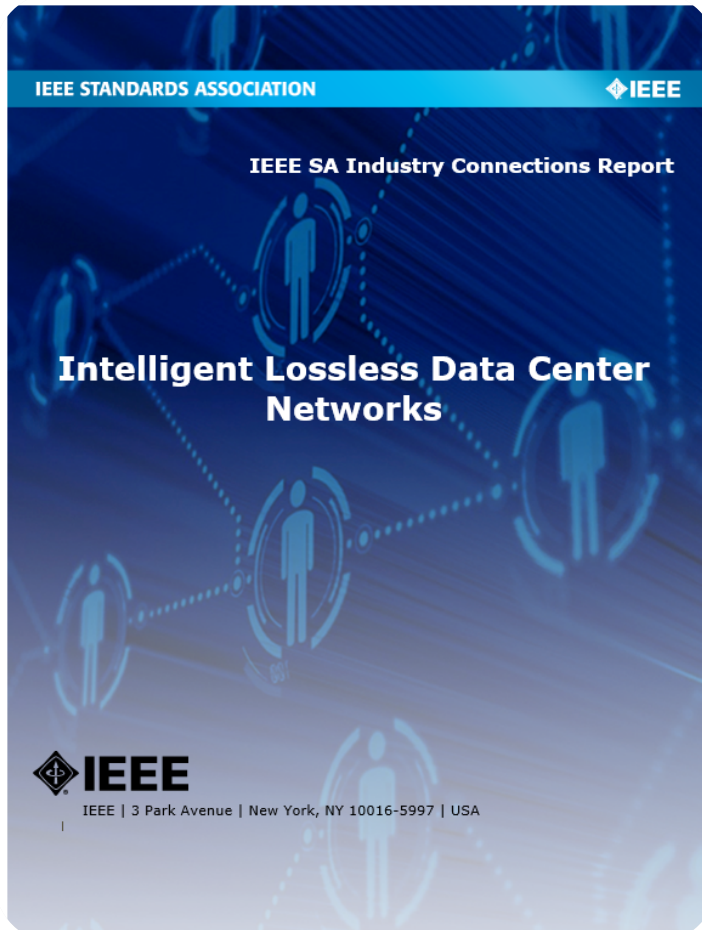


# Lossless Network in Meituan Data Center

---

Meituan Network Engineer : HAO QUAN



1

**Meituan Data Center Network Architecture**

---

2

**Lossless Network in Meituan**

---

3

**Challenges and Requirements for Lossless Network**

---

**/01**

## **Meituan Data Center Network Architecture**

---



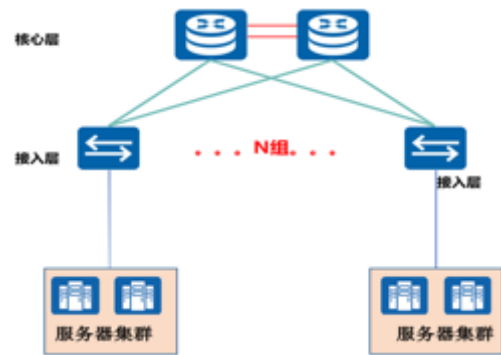
# Meituan Data Center Network Overview

---



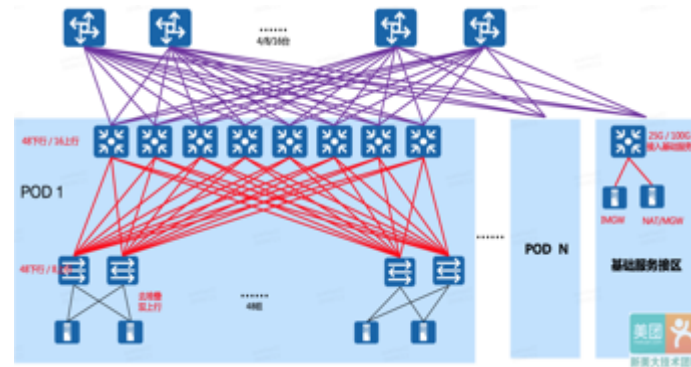
# Meituan Data Center Network Evolution

**MTDP NET 1.0&2.0**  
before 2018



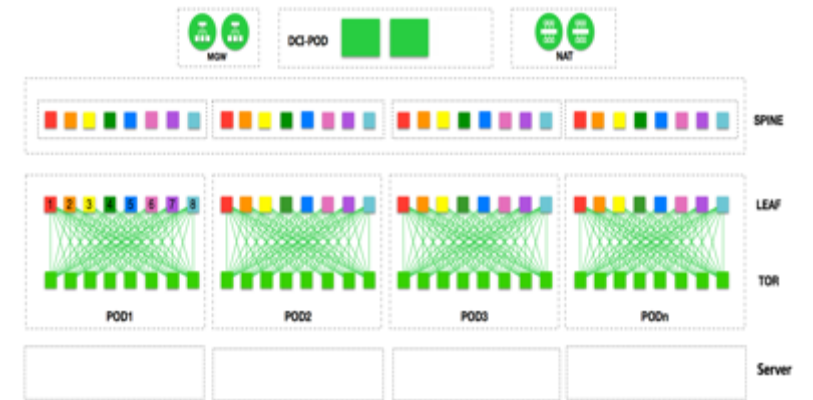
- Physical Connect : Copper
- Link bandwidth : 10G/40G
- Server Bandwidth : 1G/10G
- Cluster Size : 1K-3K

**MTDP NET 3.0**  
2018-2020



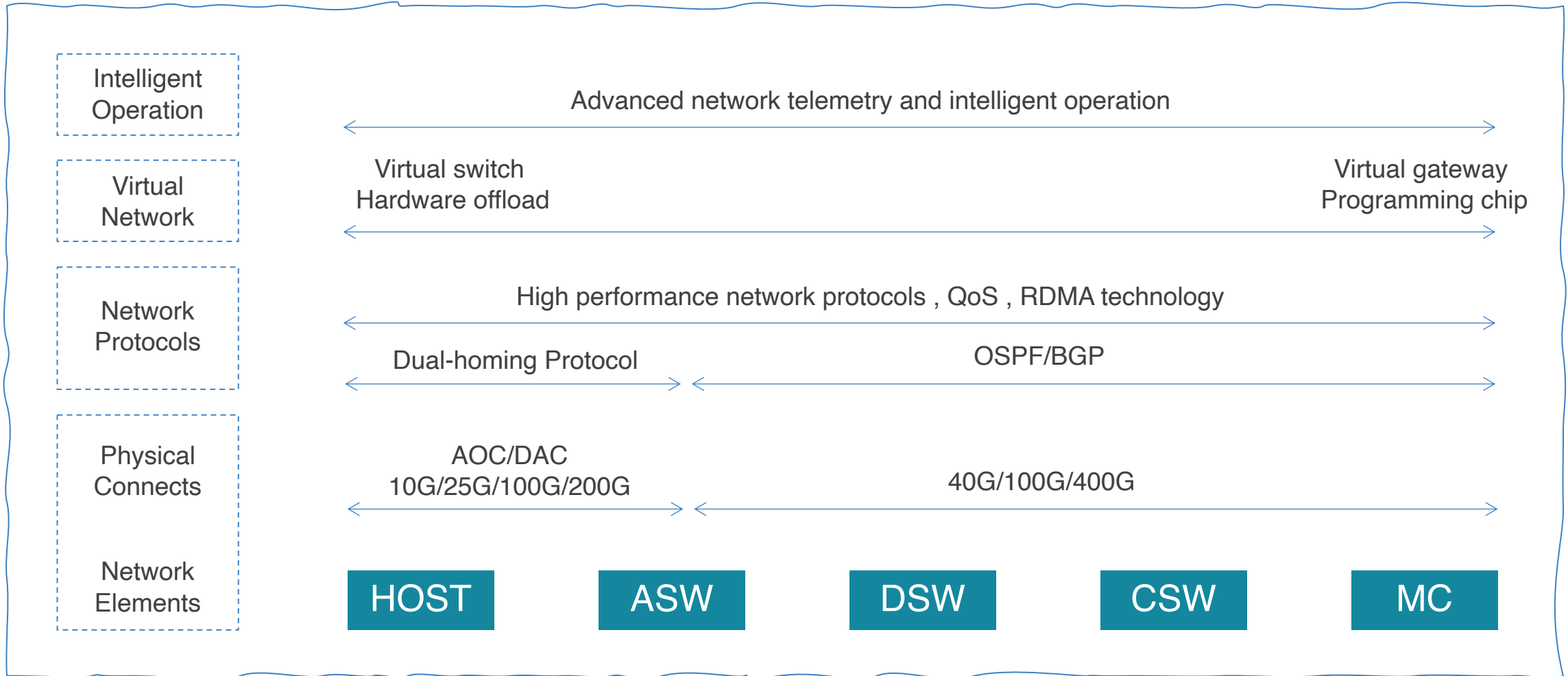
- Physical Connect : AOC
- Link bandwidth : 100G
- Server Bandwidth : 25G/100G
- Cluster Size : 20K-30K

**MTDP NET 4.0**  
2020-



- Physical Connect : AOC/DAC
- Link bandwidth : 400G
- Server Bandwidth : 100G/200G
- Cluster Size : 30K-100K

# Meituan Data Center Architecture Landscape



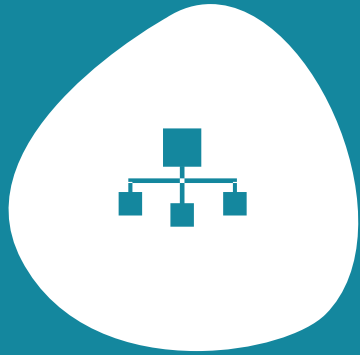
# /02

## Lossless Network in Meituan

---



# Meituan Team Participating in Lossless Network Construction



## Network Team

Research network-related lossless network technology, monitor and maintain lossless network



## Big Data/ML/AI Team

End customers who enjoy the benefits of lossless network



## Server Team

Research server-related lossless network technology, build and maintain servers



# Current Meituan Business using Lossless Network



## Unmanned Vehicle

- Image identification
- Path calculation
- Efficiency optimization



## NLP

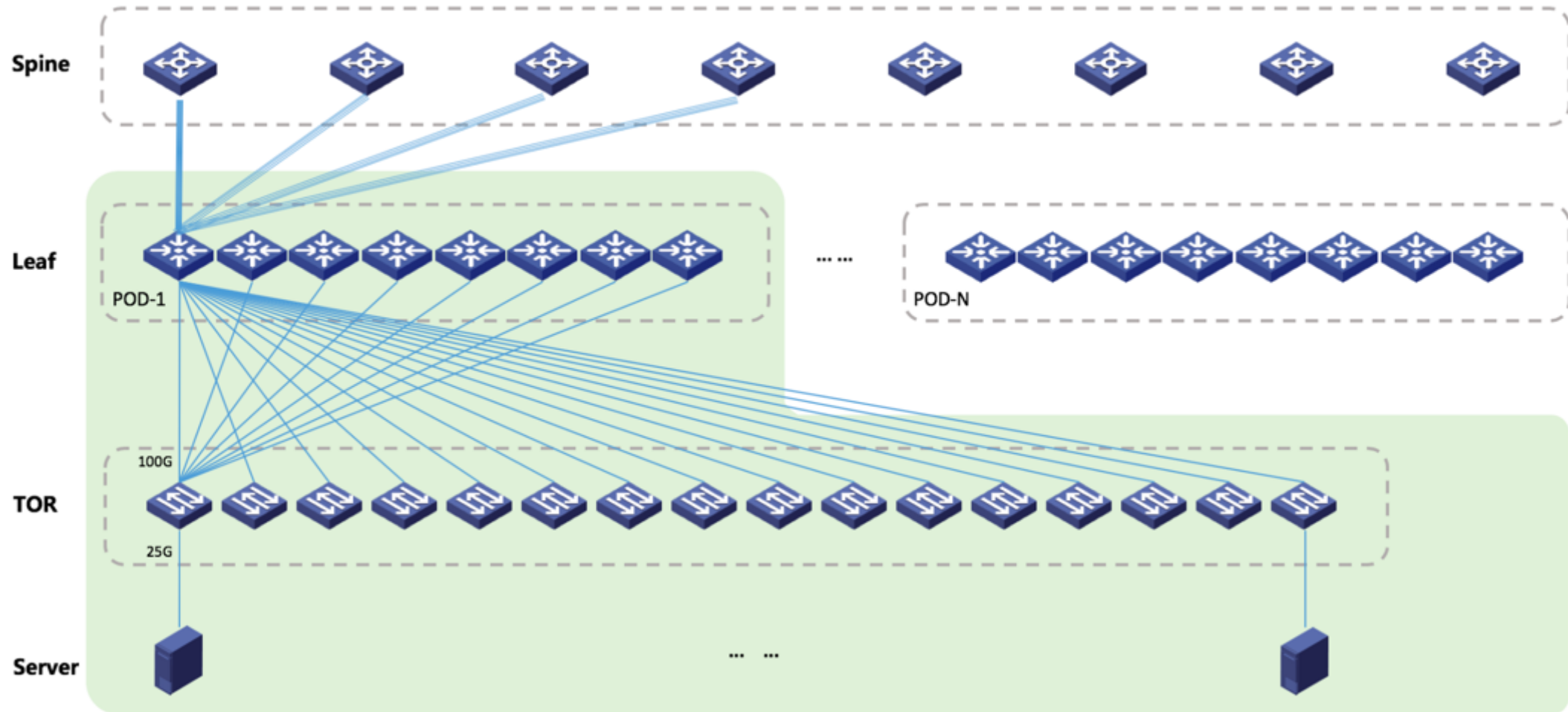
- Word segmentation
- Information extraction
- Textual entailment



## Speech Recognition

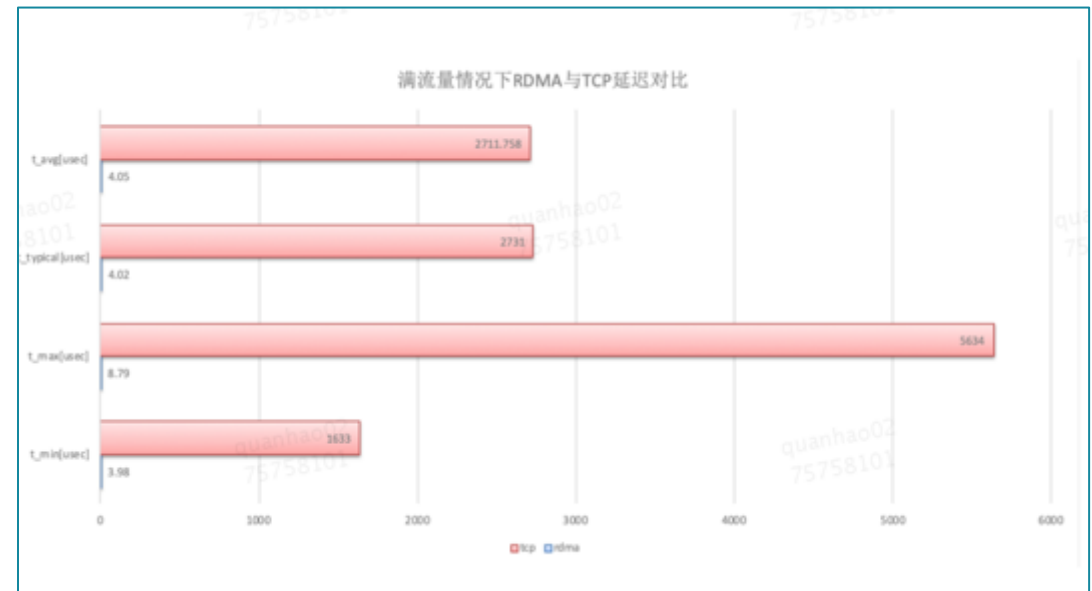
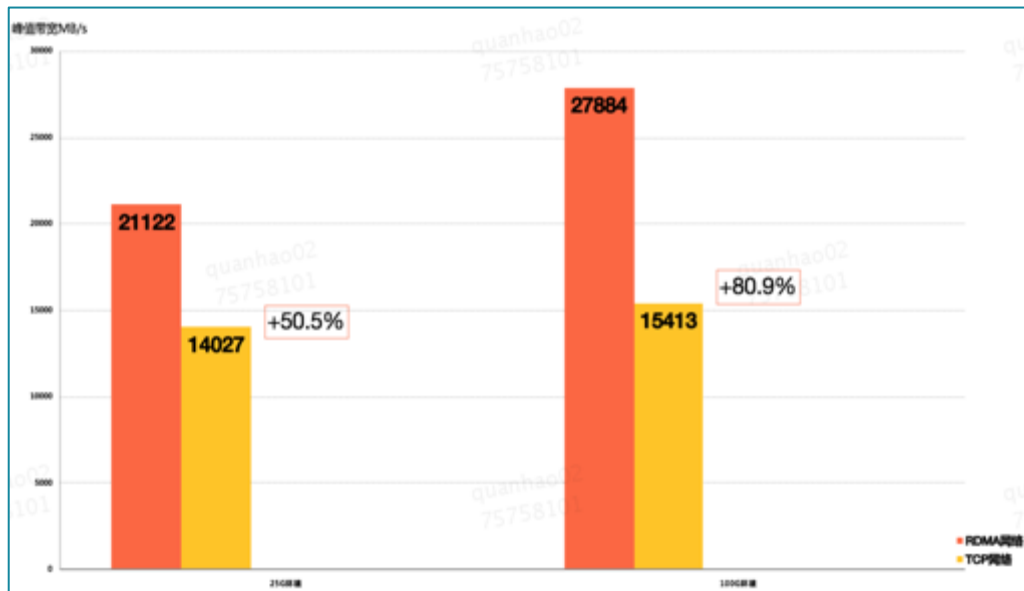
- Parsing
- Question answering
- Automatic summarization

# Meituan Lossless Network Topology



# Lossless Network Benefits Meituan Business

- ✓ More and more businesses will turn to RDMA: delay-sensitive businesses, deep learning, distributed storage, etc.
- ✓ The bandwidth utilization rate under 100G network is increased by 80%, and the delay is reduced by 640%
- ✓ 0.2% packet loss will affect 50% calculation efficiency



# Precautions For Deploying RDMA Networks

---

- ✓ If possible, just use ECN . Give up PFC
- ✓ If need to use pfc for the bottom line, because the use of pfc will cause problems such as unfair speed reduction, PFC storm, deadlock, etc. Therefore, to enable PFC in the data center, you need to strictly monitor and manage pause frames to ensure network connectivity reliability
- ✓ Avoid designing and using network with a convergence ratio of less than 1
- ✓ Understand what training model is using RDMA network

**/03**

## **Challenges and Requirements for Lossless Network**

---



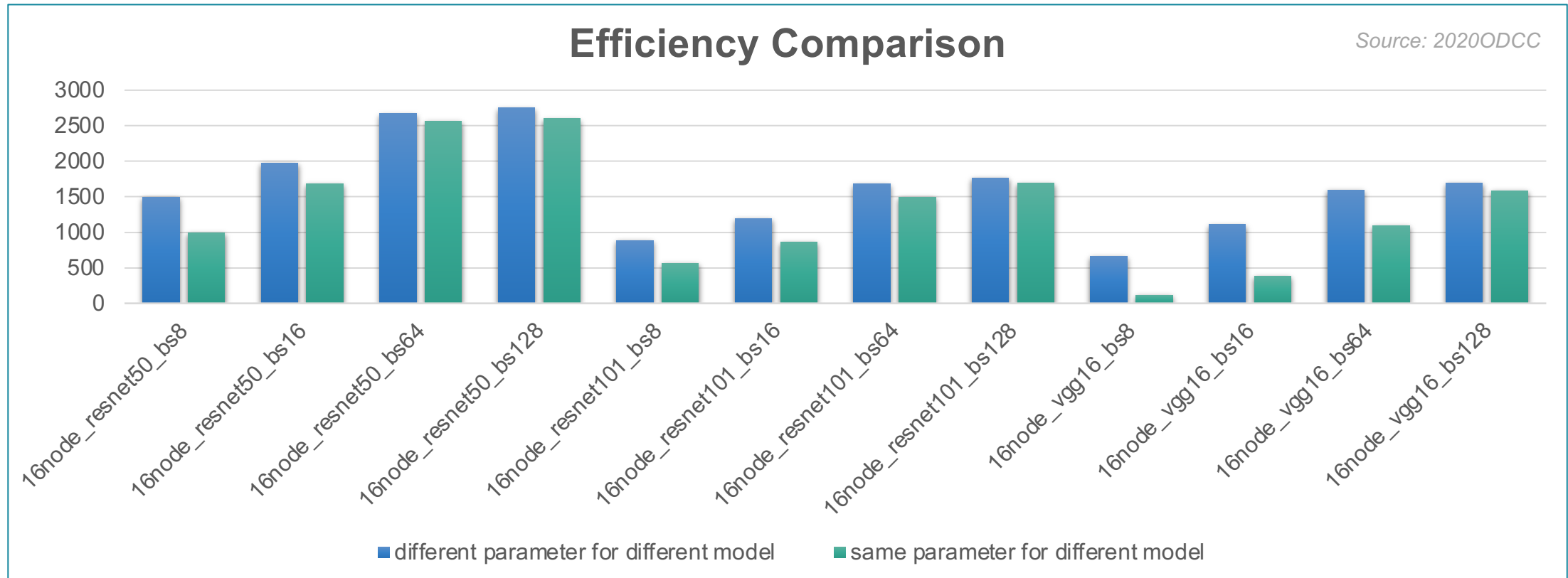
# Best or not ?

- ✓ This set of parameters for this model , could it be better ?
- ✓ When is best ?

Test Result	Performance (3 nodes cluster)		
	TCP Networking	RoCEv2	RoCEv2 + Lossless tech
VDI scenario 16K, Random, R:W=4:6	112006.56	200631.3	208283.9
VDI scenario 8K-64K, Random, R:W=2:8	65770.60	82023.5	108946.5
OLTP scenario 8K, Random, R:W=7:3	190967.73	447664.5	445081.1
OLTP scenario 64K, Random, R:W=5:5	42241.87	59702.6	75776.4
OLAP scenario 512K, Sequential, R:W=9:1	11201.08	11922.1	13266.3
OLAP scenario 1M, Sequential, R:W=9:1	5992.50	6384.6	6895.0
Large file writing 1M, Sequential, R:W=1:9	6058.11	5709.4	6199.7

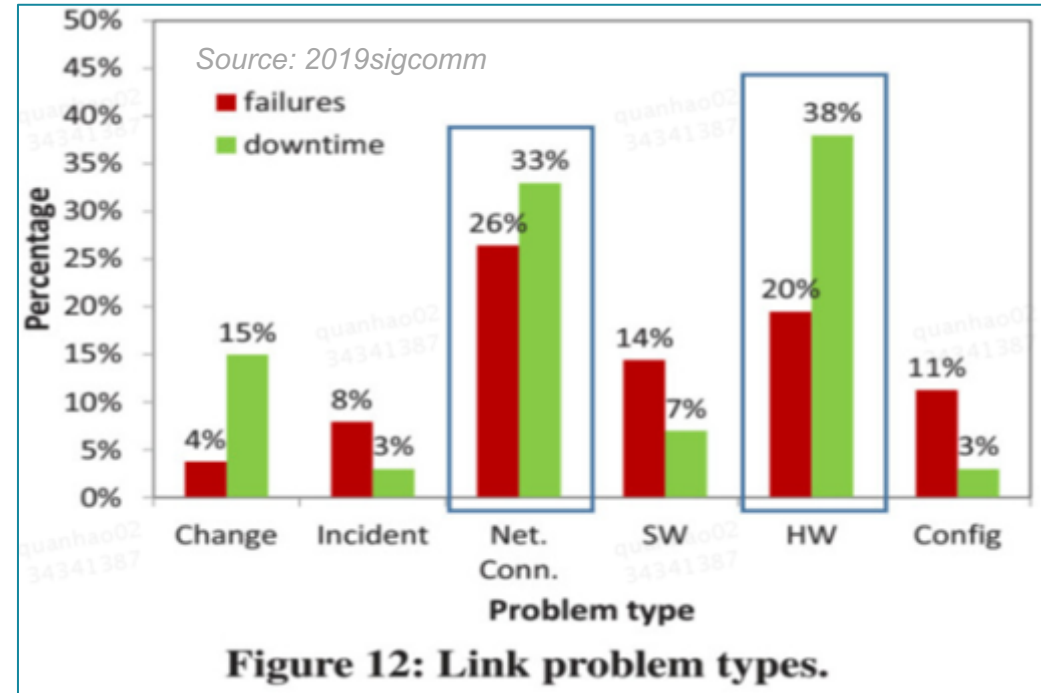
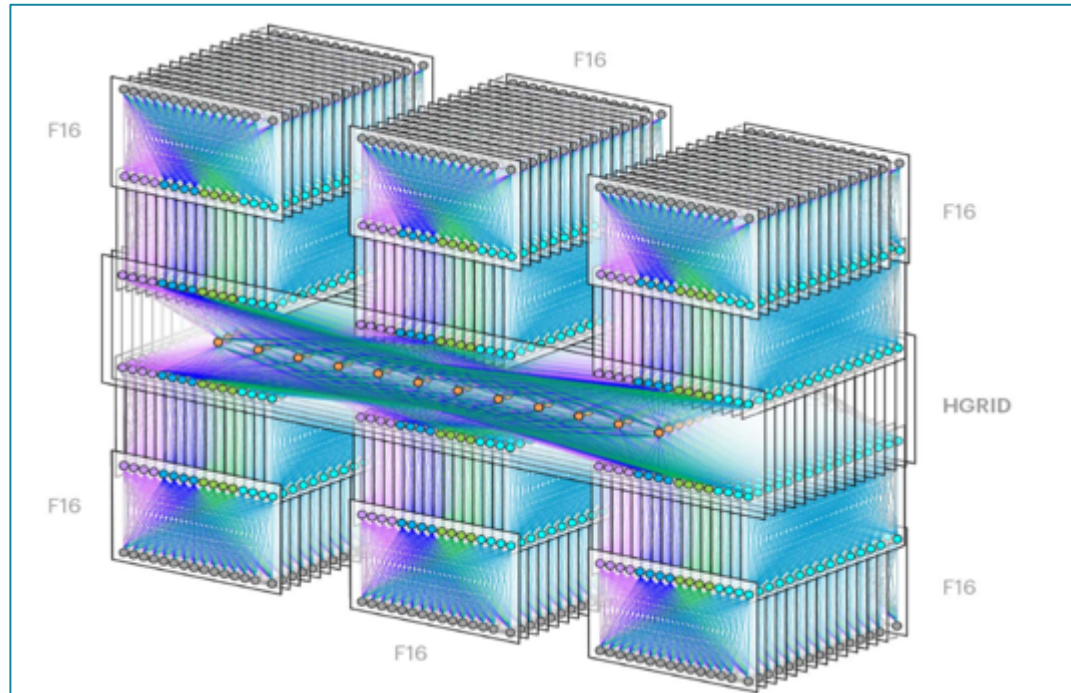
# Artificiality or AI

- ✓ Whether each model uses the same set of parameters or each model is adjusted separately
- ✓ Is there a set of network parameters applicable to all models



# How to deal with link failure ?

- ✓ Most of the faults in the data center network are link-related faults and cannot be eliminated
- ✓ Once link jitter occurs, there will be packet loss, and it is no longer a lossless network







**Thanks**

---