2021 中国智能网卡研讨会

CHINA SMARTNIC WORKSHOP

阿里高性能网络探索与实践

张彭城
阿里云网络架构师

数据中心网络发展阶段

Internet
面向连接

互联网应用
超大规模

AI/Big Data/HPC
云计算社会基础设施

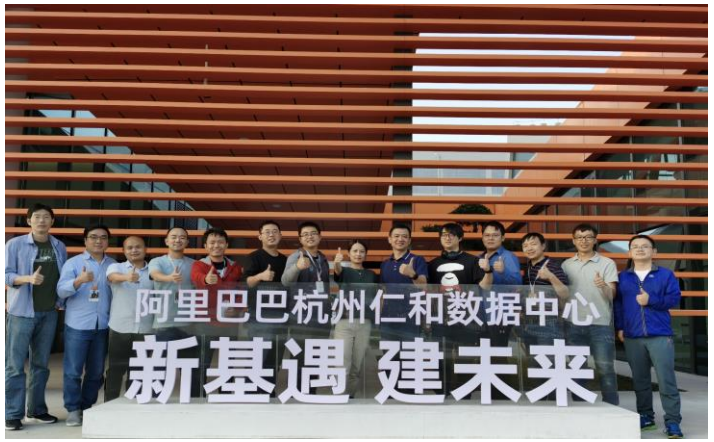
1.0

2.0

3.0

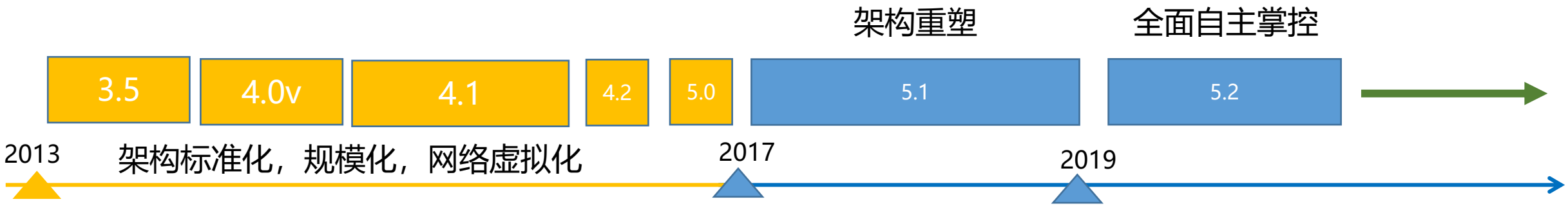
功能丰富，但日趋复杂
厂商主导，设备为王
人肉运营

Scale out 超大规模
软件定义网络，虚拟化
设备解耦，白盒化，运营自动化



算力普惠，规模新台阶
资源极致弹性供给和利用

数据中心网络架构演进：Network 1.0 → 2.0



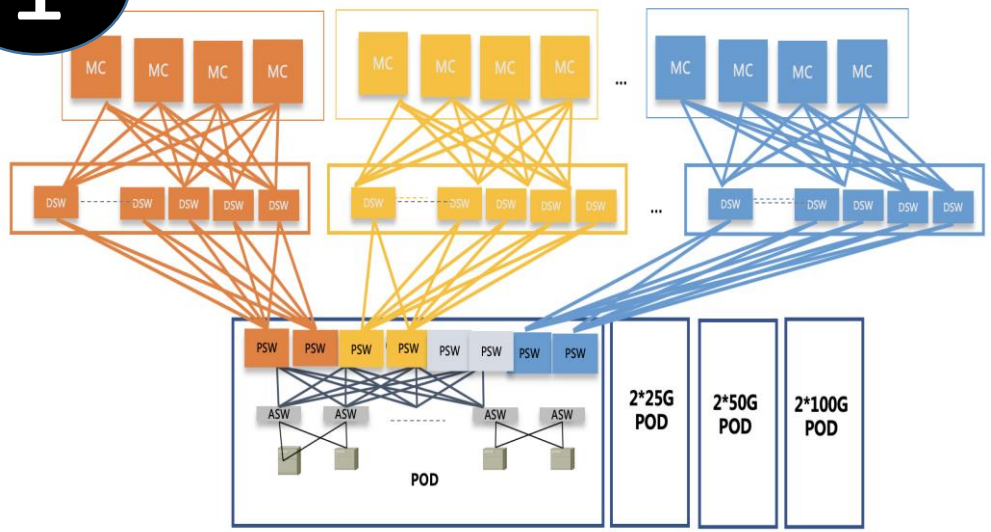
Network 1.0
企业级网络架构

Network 2.0
互联网级网络架构
超大规模，高可靠，高性能，成本优化

阿里数据中心与AWS/Gooogle/Azure/Facebook处于同一水平线
全面完成了 Network 2.0 的架构升级

阿里巴巴 Network 2.0

1 弹性 Scale-out 架构



+

2 自主研发软硬件设备

自主研发 软件协议	自主研发 运营监管控
基于开源网络 OS	
交换机白盒，硬件自定义	
商业化芯片，功能定制/可编程	

SONiC

3 基于大数据的自动化智能化 **运营监控管控系统**

运营经验积累，自主研发设备无缝结合，集中和分布监控有机结合

* High Availability, Intelligence and Low Latency

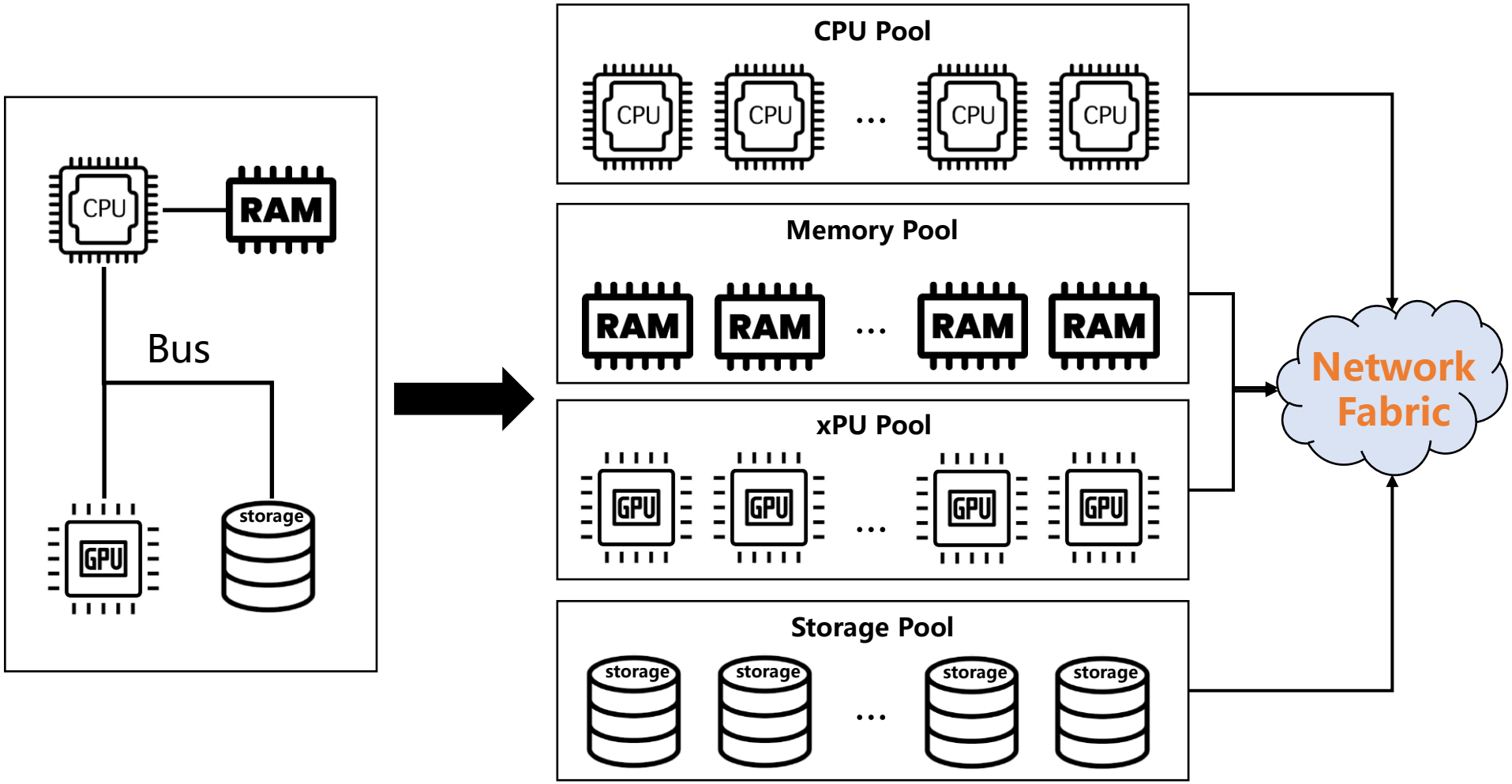
阿里巴巴 Network 3.0

社会基础设施，算力普惠
资源极致弹性 供给和利用

云计算
池化数据中心网络

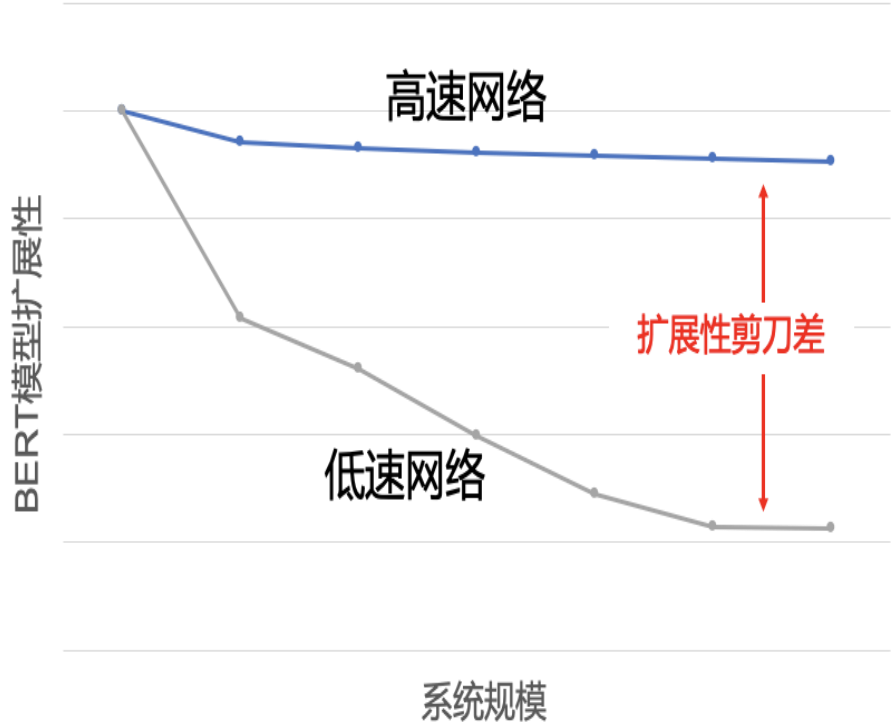
阿里巴巴 Network 3.0：池化数据中心网络

DC as computer， Network as Bus



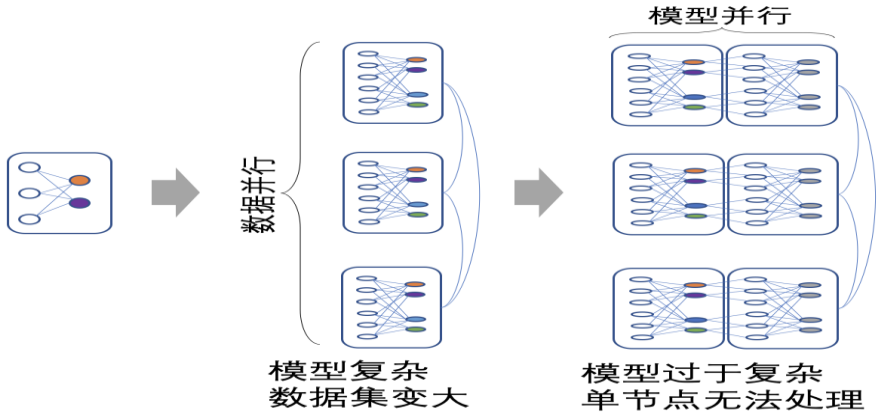
池化数据中心依赖高性能网络

以AI应用为例：



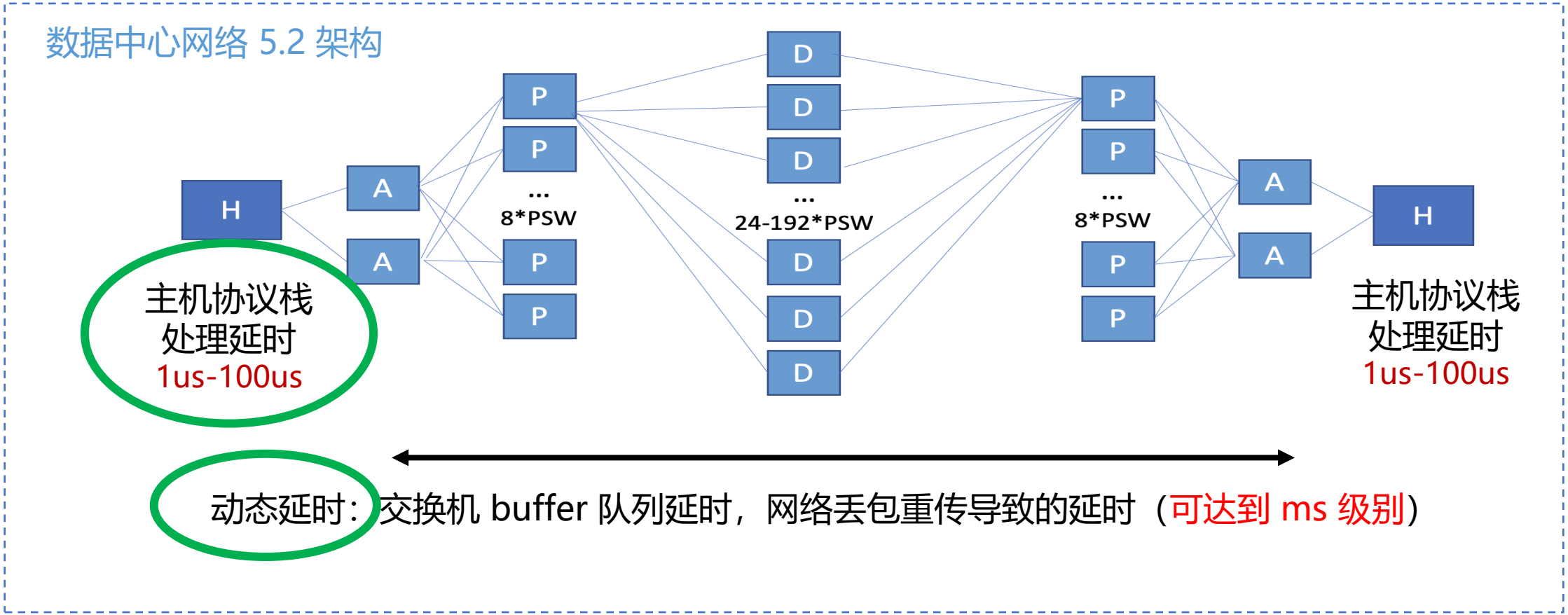
稠密计算类AI：图像处理-拍立淘

512 V100 GPU，通信占比超过50%



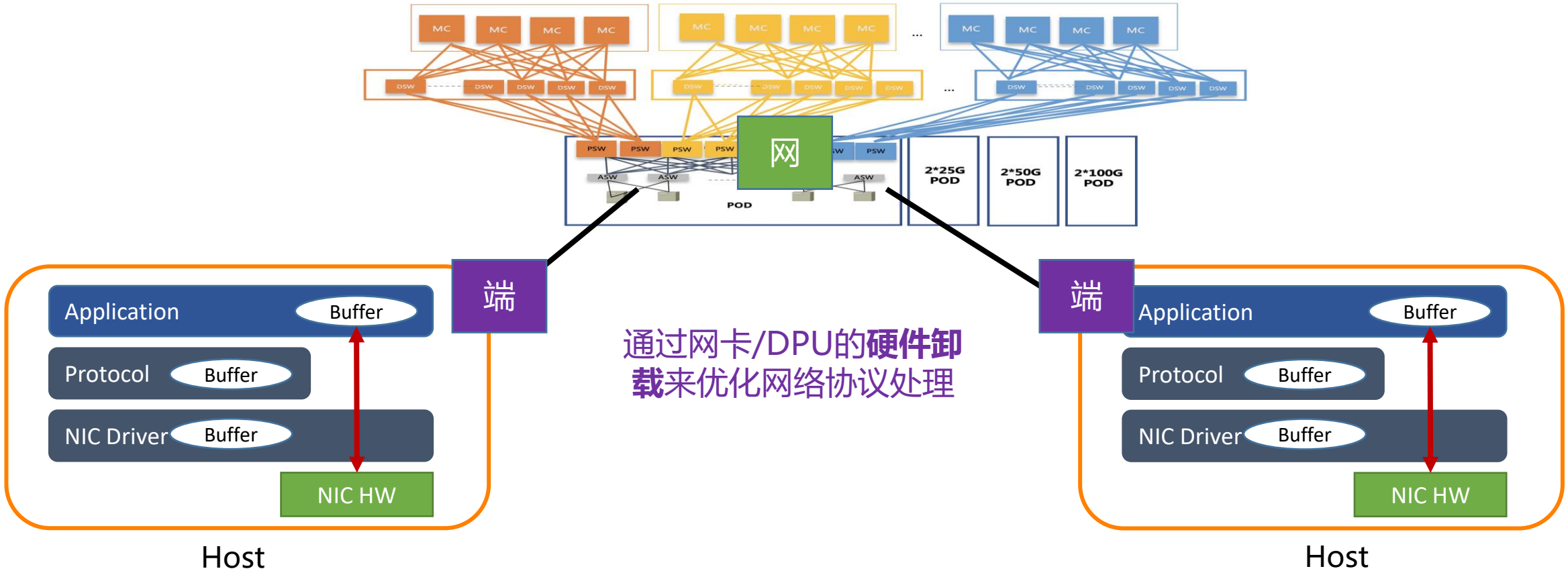
高性能网络挑战：如何将时延再降低1个数量级？

优化关键点：主机协议栈延时优化，网络动态延时优化



高性能网络核心思想：端网协同

通过**端网协同**的**网络流控，多路径优化和全链路网络 QoS** 等机制来降低网络动态延时
减少网络拥塞，故障时快速切换，保障大小流之间的公平性，有效处理网络 incast

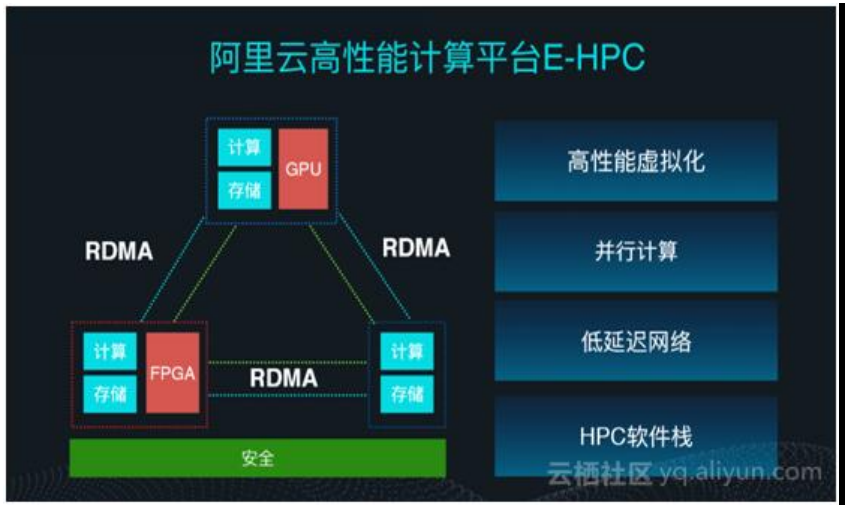


阿里高性能网络演进

第一阶段（2017~2018）：RDMA大规模落地

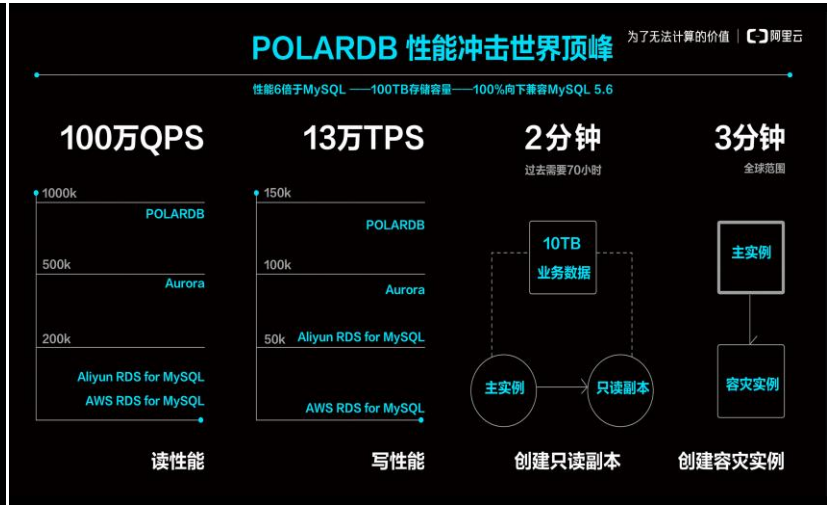
1、RDMA在阿里云多个产品实现大规模落地；
2、端到端的流控，应用性能调优；
3、消除RDMA稳定性风险：建立端网协同的运营体系，消除PFC风险，平稳支撑双11大促；

高性能计算



高性能存储

高性能数据库



阿里高性能网络演进

第二阶段（2019~2020）：自研之路

1、自研高性能网络协议：HPCC拥塞控制算法、Multi-path、xRD传输方式，去PFC实现Lossy RDMA，进一步优化长尾延时，增加RDMA扩展性；

2、自研高性能网卡：实现自研高性能网络协议卸载，目前已经落地云存储；

3、自研高性能通信库：针对大规模AI集群，优化集合通信功能，利用端网融合技术深度结合大规模AI集群的互联架构和多网卡特性，实现创新的无拥塞算法和高速网络传输；

Sigcomm’19

HPCC: High Precision Congestion Control

Yuliang Li[♣♡], Rui Miao[♣], Hongqiang Harry Liu[♣], Yan Zhuang[♣], Fei Feng[♣], Lingbo Tang[♣], Zheng Cao[♣], Ming Zhang[♣], Frank Kelly[♢], Mohammad Alizadeh[♣], Minlan Yu[♡]
Alibaba Group[♣], Harvard University[♡], University of Cambridge[♢], Massachusetts Institute of Technology[♣]

ABSTRACT

Congestion control (CC) is the key to achieving ultra-low latency, high bandwidth and network stability in high-speed networks. From years of experience operating large-scale and high-speed RDMA networks, we find the existing high-speed CC schemes have inherent limitations for reaching these goals. In this paper, we present HPCC (High Precision Congestion Control), a new high-speed CC mechanism which achieves the three goals simultaneously. HPCC leverages in-network telemetry (INT) to obtain precise link load information and controls traffic precisely. By addressing challenges such as delayed INT information during congestion and overreac-

demand on high-speed networks. The first trend is new data center architectures like resource disaggregation and heterogeneous computing. In resource disaggregation, CPUs need high-speed networking with remote resources like GPU, memory and disk. According to a recent study [17], resource disaggregation requires 3-5μs network latency and 40-100Gbps network bandwidth to maintain good application-level performance. In heterogeneous computing environments, different computing chips, *e.g.* CPU, FPGA, and GPU, also need high-speed interconnections, and the lower the latency, the better. The second trend is new applications like storage on high I/O speed media, *e.g.* NVMe (non-volatile memory express)

HPCA’20

EFLOPS: Algorithm and System Co-design for a High Performance Distributed Training Platform

Jianbo Dong, Zheng Cao, Tao Zhang, Jianxi Ye, Shaochuang Wang, Fei Feng
Li Zhao, Xiaoyong Liu, Liuyihan Song, Liwei Peng, Yiqun Guo, Xiaowei Jiang
Lingbo Tang, Yin Du, Yingya Zhang, Pan Pan, Yuan Xie
Alibaba Group
{jianbo.djb, zhengzhi.cz, t.zhang, jianxi.ye}@alibaba-inc.com

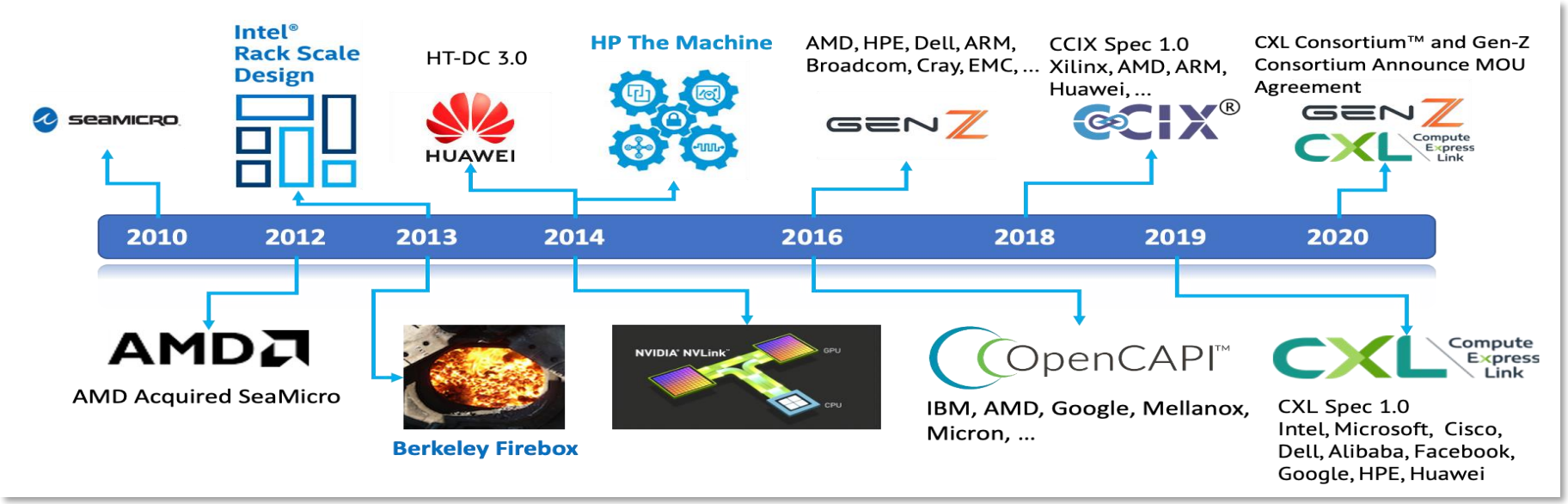
***Abstract*—Deep neural networks (DNNs) have gained tremendous attractions as compelling solutions for applications such as image classification, object detection, speech recognition, and so forth. Its great success comes with excessive trainings to make**

In general, the contemporary techniques for better training performance could be classified into two categories, *scale-up* and *scale-out*. The scale-up techniques focus on the performance of a single device. For instance, various special

阿里高性能网络演进

第三阶段（2021~）：总线+网络融合，仍有很多待突破难题

1、协议融合：PCIE/CXL over Ethernet？如何解决性能和扩展性问题？
2、硬件载体：SmartNIC？DPU？Switch？
3、网络架构：Fat-tree是否还适用？



THANKS