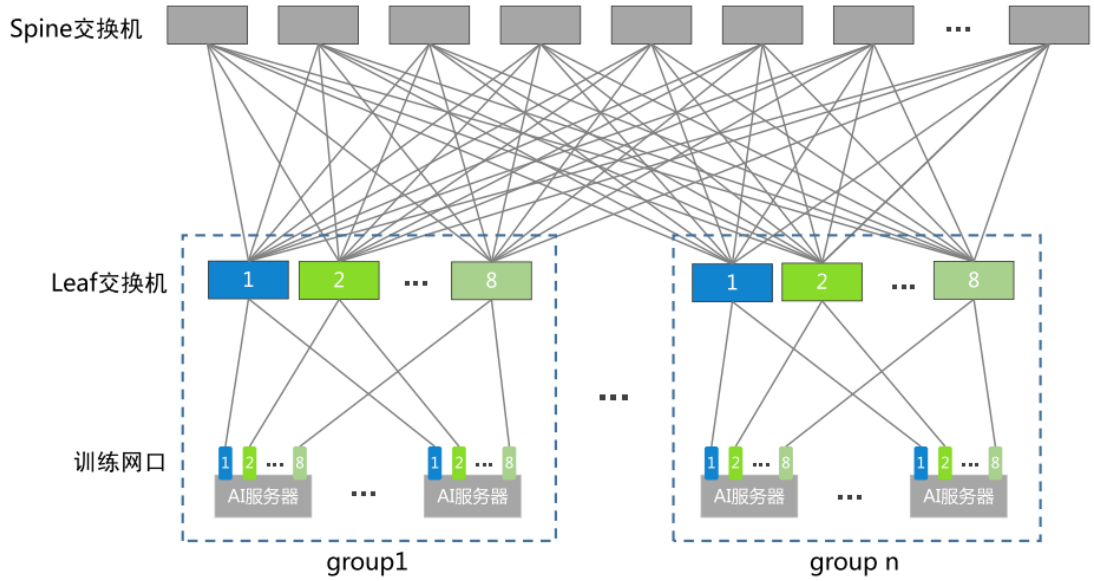
# 算力网络测试平台

自 ChatGPT 发布以来，科技界掀起了一场大模型的竞争热潮。数据成为新生产要素，算力成为新基础能源，大模型则成为新生产工具，各行各业从“+Al”向“Al+”的转变已势不可挡。随着模型参数量从千亿迈向万亿，模型能力更加泛化，大模型对底层算力的诉求进一步升级，超万卡集群成为这一轮大模型基建军备竞赛的标配。超万卡集群将有助于压缩大模型训练时间，实现模型能力的快速迭代，并及时对市场趋势作出应对。然而，如何在超万卡集群中实现高效的训练，并长期保持训练过程的稳定性，是将大模型训练扩展到数万张 GPU 卡上所要面临的双重挑战。超万卡集群运行过程中涉及到集群有效算力发挥、超大规模互联网络稳定性保障、故障的快速排查和修复等关键问题，目前都是业内关注的焦点。

超万卡集群网络包括参数面网络、数据面网络、业务面网络、管理面网络。业务面网络、管理面网络一般采用传统的 TCP 方式部署，参数面网络用于计算节点之间参数交换，要求具备高带宽无损能力。数据面网络用于计算节点访问存储节点，也有高带宽无损网络的诉求。超万卡集群对参数面网络要求最高，主要体现在四个方面：大规模，零丢包，高吞吐，高可靠。

目前业界成熟的参数面主要包括 IB(InfiniBand)和 RoCE 两种技术。面向未来 Al 大模型演进对网络提出的大规模组网和高性能节点通信需求，业界也在探索基于更新的算力网络协议技术规范。

随着智算集群规模不断扩大，集群运维管控与集群应用之间的矛盾日益凸显。随着单集群的 Al 加速卡数量从从千级增长到万级，相应的故障范围扩大超过 10 倍。典型故障范围从单服务器单卡迅速扩散到算网存多域全栈。为彻底改变这一运维困境，亟需引入新的运维理念和技术，以集群全链路可视化监控、故障快速定位和运维侧快速修复为原则来建设新的集群计算智能运维系统。




XproAINetwork 算力网络测试工具是触点互动在此 AI 大模型大发展的背景下推出的算力网络仿真测试工具，通过仿真大量 AI 服务器的跨节点集合通信过程在算力网络开通之前或者网络部署调整之前进行仿真压力测试，测试算力网络能够达到的总带宽容量，延时，丢包等核心通信性能指标。可以有效帮助算力数据中心的运营者在大模型上线前充分优化其算力网络的部署。

XproAInetwork 测试工具采用 X86 平台，以软件化形式部署在 X86 服务器之上来仿真 GPU 节点之间的集合通信数据交换过程。其支持的网络接口带宽为：100GE/200GE/400GE，同时无需更换硬件测试平台，在平台接口上更改网卡模式可以自由切换为 InfiniBand 或者 ROCE 模式。

XproAInetwork 测试工具可以模拟如下的典型 GPU 集合通信过程：

1. AllReduce
2. ReduceScatter
3. AllGather
4. Broadcast
5. Reduce
6. RingReduce

并且随着部署的服务器数量增加，模拟的 GPU 服务器节点数量也线性增加。通过多台服务器的线性叠加部署来模拟真实万卡模型的海量通信数据量。为了节省服务器成本，单台 X86 服务器最多可支持 4.8Tbps 的流量模拟能力，单台 X86 服务器可以模拟到 12 台 AI 服务器的通信流量。通过这种有效的收敛比可以显著降低测试时所需的 AI 服务器数量。同时 XproAInetwork 测试工具完成算力网络开局测试后，其占用的服务器可以释放出来继续做为 AI 服务器使用。这一灵活的部署方式，显著降低了算力网络在大规模集群情况下的测试成本。

XproAInetwork 测试工具具备如下针对算力网络的测试能力：

- 支持 100 纳秒精度的单向转发延时测量精度
- 支持测试工具跨服务器机箱的高精度单向转发延时测量
- 支持在高带宽情况下的精确延时测量
- 支持 ECN 协商
- 支持集合通信各种动作的性能指标：
    - Reduce 的指令完成时延
    - Broadcast 的指令完成时延
    - Gather 的指令完成时延
    - Scatter 的指令完成时延
    - All to all 的指令完成时延
    - 针对不同传输数据块大小的时延
- 支持统计每个网口收到的带 ECN 标识的报文数量
- 支持统计每个网口发送的 CNP 控制流量报文的数量
- 支持统计每个网口接收并处理的 CNP 控制流量报文的数量
- 支持统计每个网口收到的 PFC 暂停帧数量
- 支持统计每个网口的 PFC 暂停时间长度