## A Scalable, Commodity Data Center Network Architecture

传统的数据中心网络架构采用树形拓扑，即使是采用高端的网络设备，顶层的设备都难以支撑网络边缘的数据总流量。面临着流量的非均匀化，软件协议设计的复杂性，使得成本大大提高。在这篇论文中，作者展示了利用传统的网络交换设备并结合胖树网络架构可以换来高性能的传输效率，并且完全兼容现有的网络协议和硬件接口。

在以往的数据中心里，通常具备如下几个特点：

1. 数据中心有成千上万的计算机构成，且带宽需求量巨大。
2. 网络架构由路由和交换机组成树形结构，越来越多的高端设备的加入使得网络层次结构越发的明晰。
3. 即使顶层采用最高端的网络设备也只能支撑起整个网络50%的带宽。
4. 流量大小的不一致性导致传统的算法很难适配数据中心，或者说需要设计很复杂的算法。

而在新型的数据中心架构中，应当具备的特点是：

1. 通过大规模使用传统的网络设备就支撑起大量的网络流量。
2. 采取新型架构可以以较低的成本换取高性能。
3. 完全向后兼容，不必更改现有网络设备和协议。

**问题分析：**

商用集群的出现使得任何一个组织都可以轻松实现大量的计算和PB级的存储，但该网络的关键瓶颈在于节点间的通讯，譬如在web请求、云计算、并行计算等诸多应用中都存在这样的问题。

作者认为解决这样问题的思路是，一是通过利用新型互联硬件和软件，如infiniBand，Myrinet，实现成本高，且无法兼容现有网络协议，因此在现有网络中进行改进的可扩展性不强。二是采用传统的网络设备和协议，但随着网络节点数量的增加，管理成本和利用率、可扩展性效果都变得极差，这在先行的ip网络下也是如此。本文采用的是第二种方案，此时，网络整体的性能受限于核心层节点的处理能力，因此需要从网络架构和设备、软件协议方面入手进行适当的改进。

**设计目标：**

1. 具有良好的可扩展性，且任意加入主机都能够以本地连接处的全部带宽与其它主机进行通信。
2. 廉价性，能够利用传统设备集群实现大规模部署。
3. 向后兼容，新的架构能够在不修改网络接口和软件协议的基础上与现有网络兼容。

**背景知识介绍：**