

# MPI: Scalable Network Requirements

**LEADERSHIP  
COMPUTING FACILITY**  
NATIONAL CENTER FOR COMPUTATIONAL SCIENCES



*presented by*  
Richard L. Graham

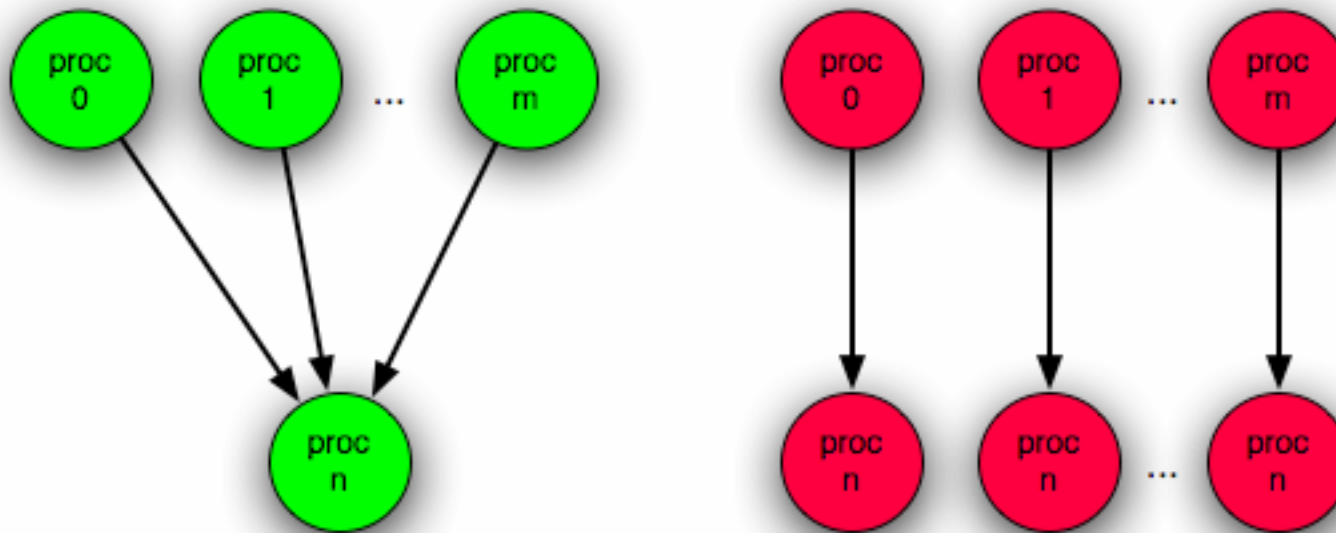
Oak Ridge National Laboratory  
U.S. Department of Energy

# MPI Communication Types

- Point-to-point two sided
- Collective
- Remote Memory Access (put/get)

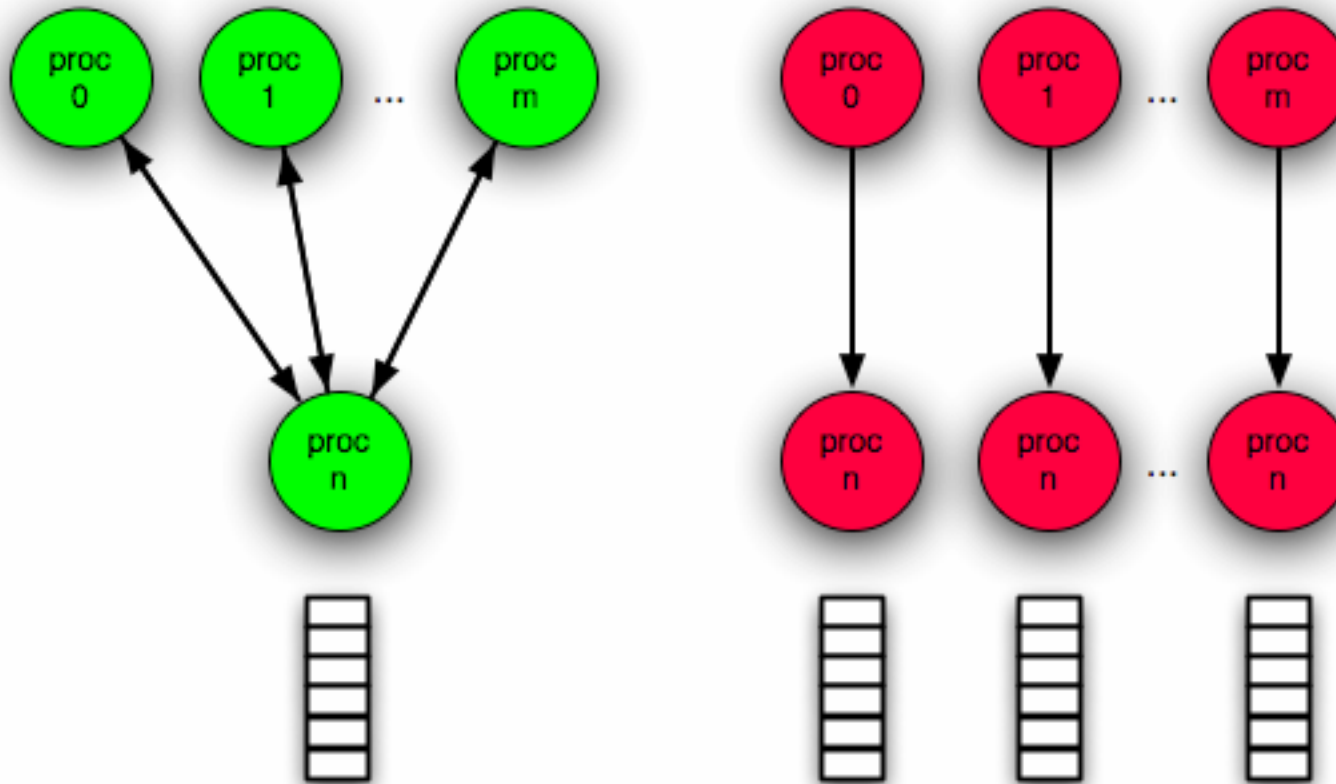
# Point-To-Point Communications

## Scalable Latency



# Point-To-Point Communications

## Network-Level Flow Control

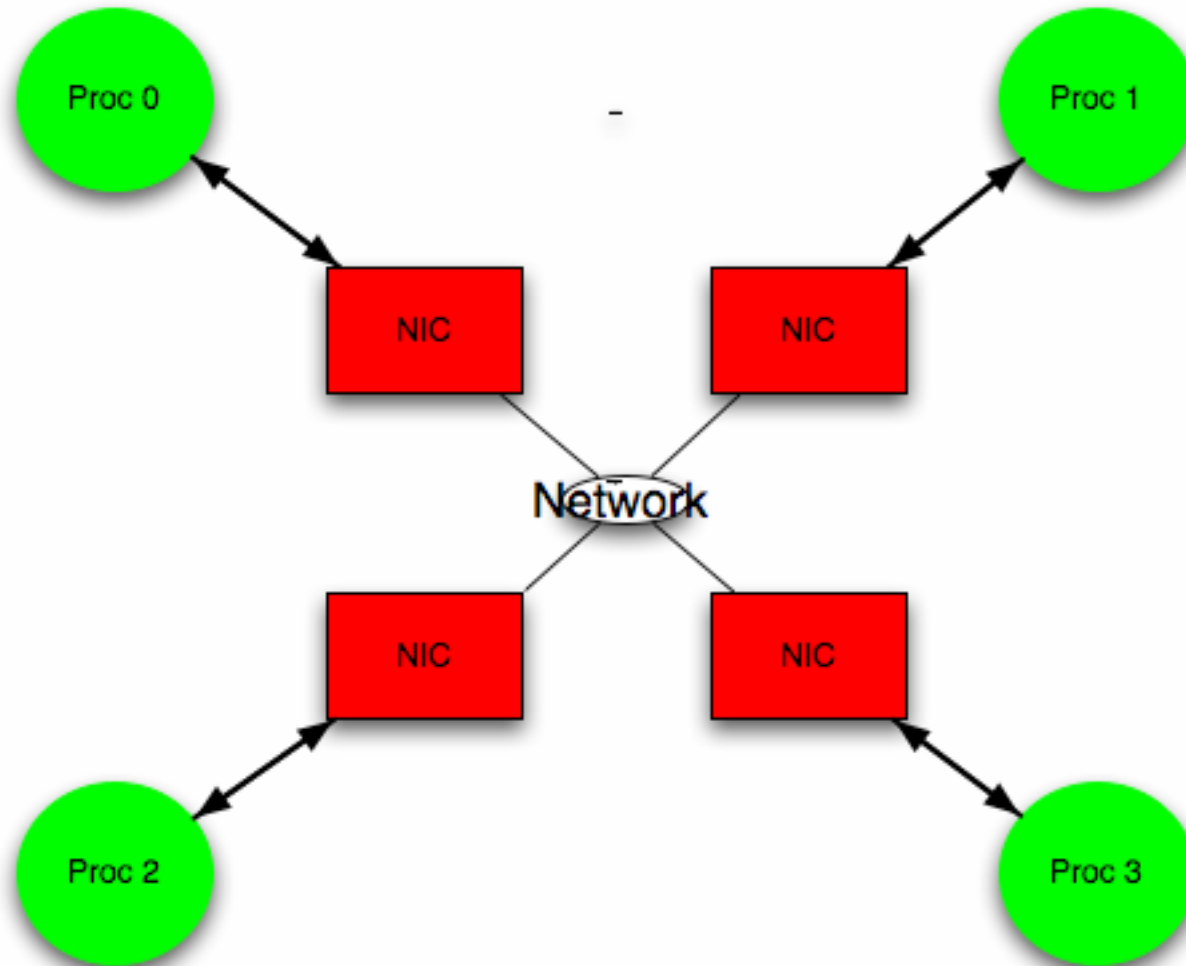


# Point-To-Point Communications

- Efficient Asynchronous progress
- High Bandwidth
  - Network Striping
- Matching on the NIC
  - Matching on multiple NIC's
  - True zero copy
- Adaptive routing
  - Selectable routes on a per data-segment basis
- Congestion control
- Reliable

# Collective Communications

## Network Offload



# Remote Memory Access

## Chained Transactions

- Cheap/Fast Memory registration

