

# Multigrid Modeling and Simulation

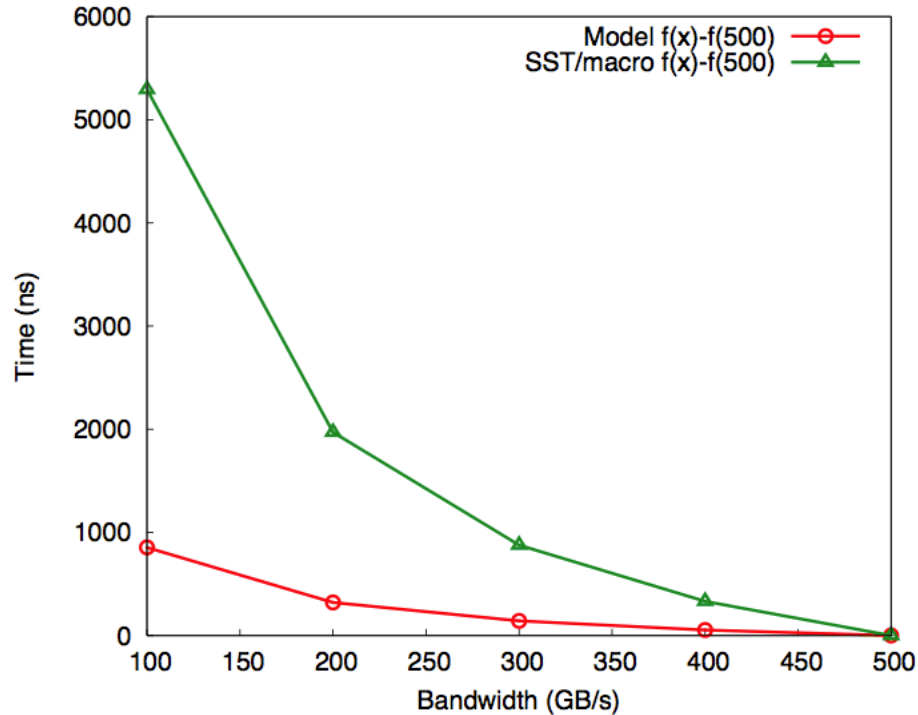
Rob Falgout, Ulrike Meier Yang  
Joe Kenny, Jeremy Wilke, Gilbert Hendry

# Modeling and Simulation: Multigrid

- Exascale (million-way) macro simulations are not routine
  - 100k network endpoints is probably reasonable
  - Parallelization of simulators (not funded by ExaCT)
  - Be creative...extrapolation,replication,???
  - Use scalable analytic models where possible
- ExaCT multigrid analysis
  - Analytic model for PFMG communication on 3D Torus and Fat Tree
  - New SST/macro skeleton for multigrid (PFMG) communication
    - Collaborative effort between SST/macro and PFMG developers



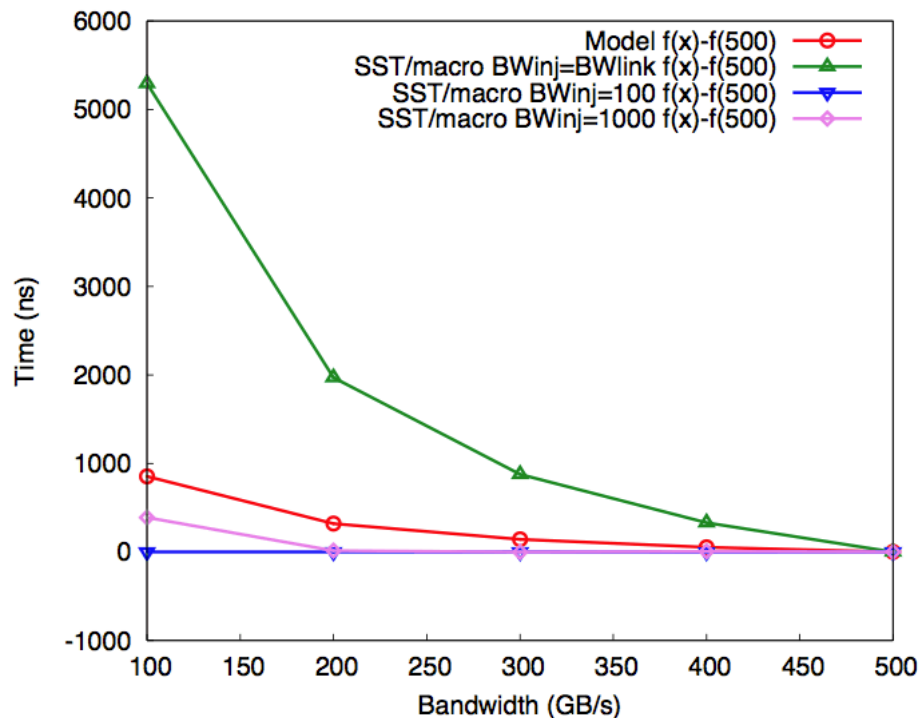
# Model vs Simulation, 3D Torus



- Simulated time slower due to network congestion?
- Or, an injection effect?
  - Current model ignores message injection
  - Chose  $BW_{inj} = BW_{link}$  for simulation

125 nodes  
n=100  
400ns latency

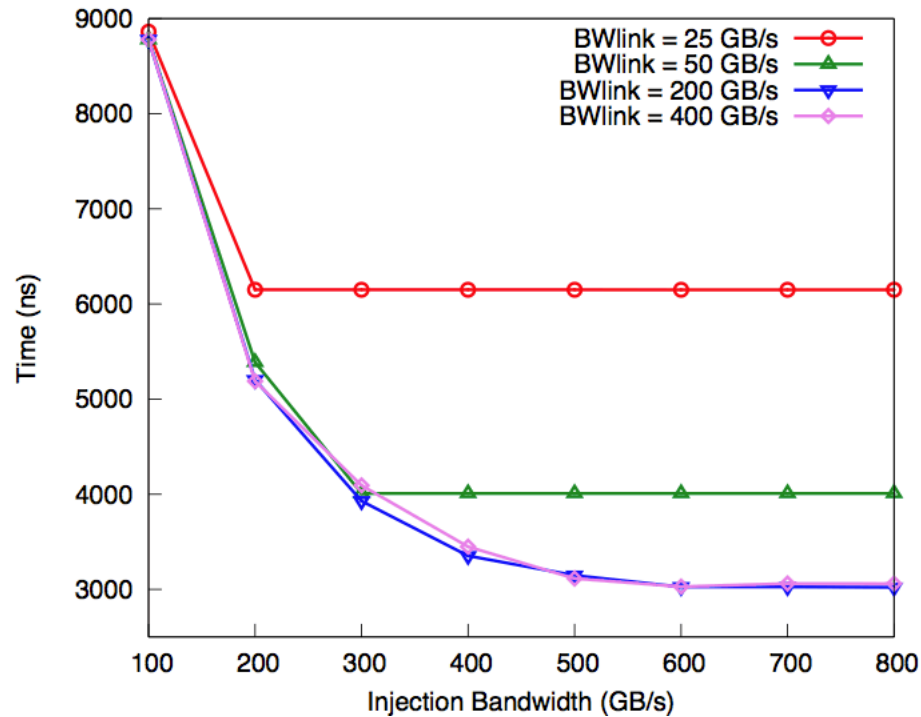
# Probing Link Bandwidth Effects



- Vary link bandwidth with constant injection bandwidth
  - Simulation time not dependent on link bandwidth
  - Network links aren't congested. Congestion is on injection links.

125 nodes  
n=100  
400ns latency

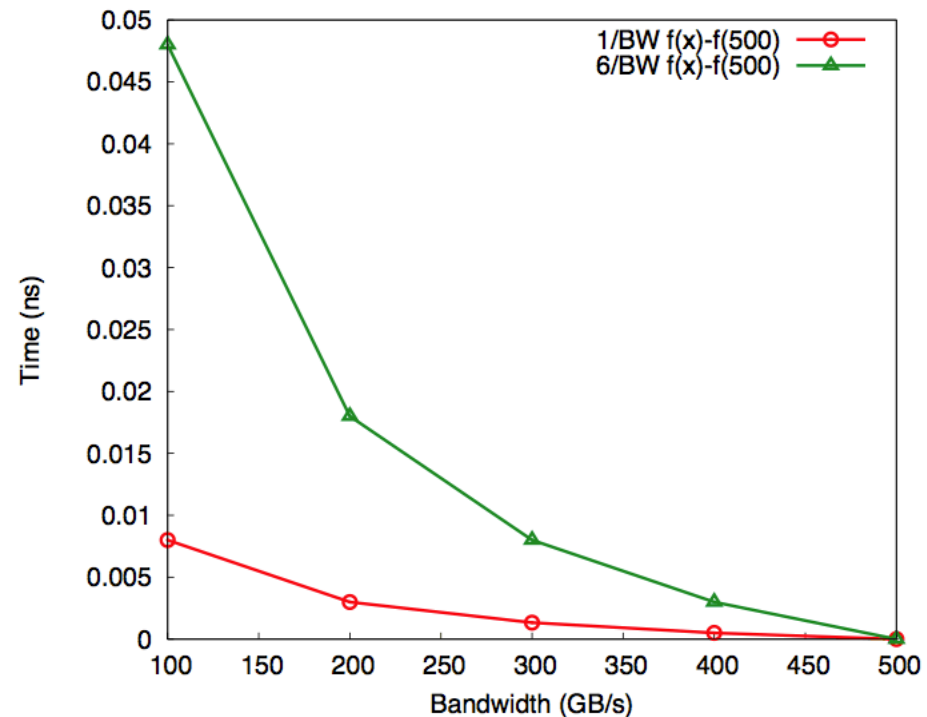
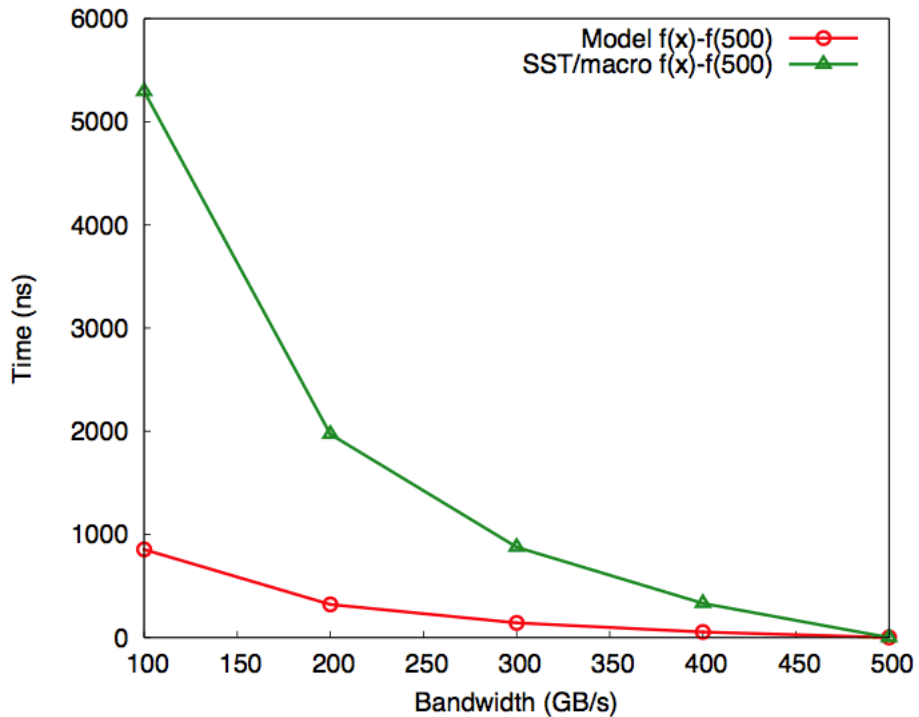
# Probing Injection Bandwidth Effects



- Link bandwidth limited only if  $BW_{link}$  is low relative to  $BW_{inj}$ 
  - Switch from link to injection limited when  $BW_{inj}$  is 6x  $BW_{link}$

125 nodes  
n=100  
400ns latency

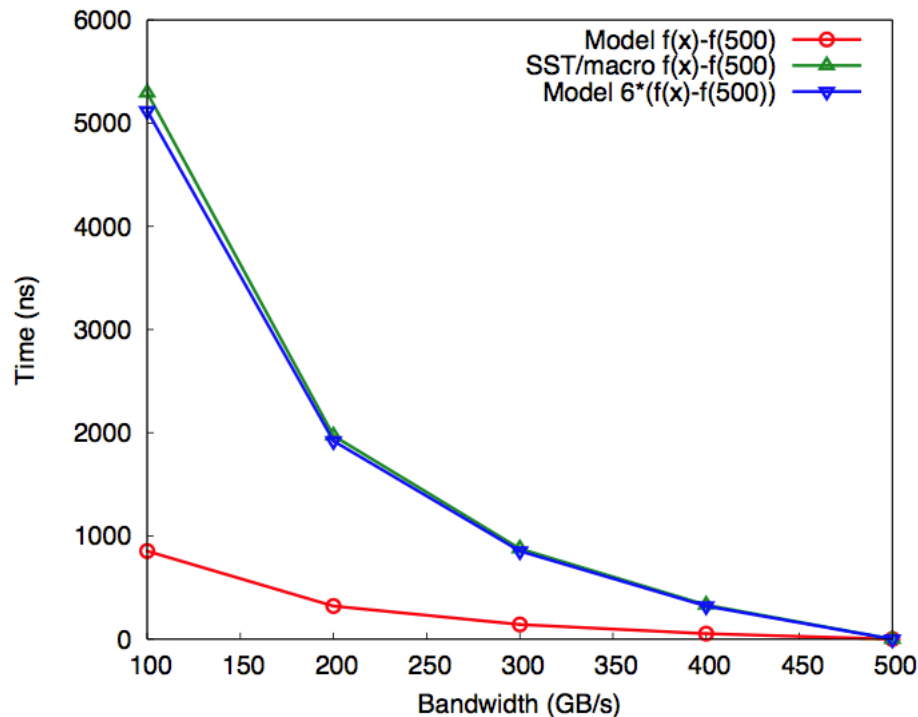
# 3D Torus: Injection Bandwidth



- Injection link carries 6 messages (+x,-x,+y,-y,+z,-z) for each send in stencil operation, network link only carries 1
  - Analytic model doesn't include injection, allocates 6x too much bandwidth to each message

125 nodes  
n=100  
400ns latency

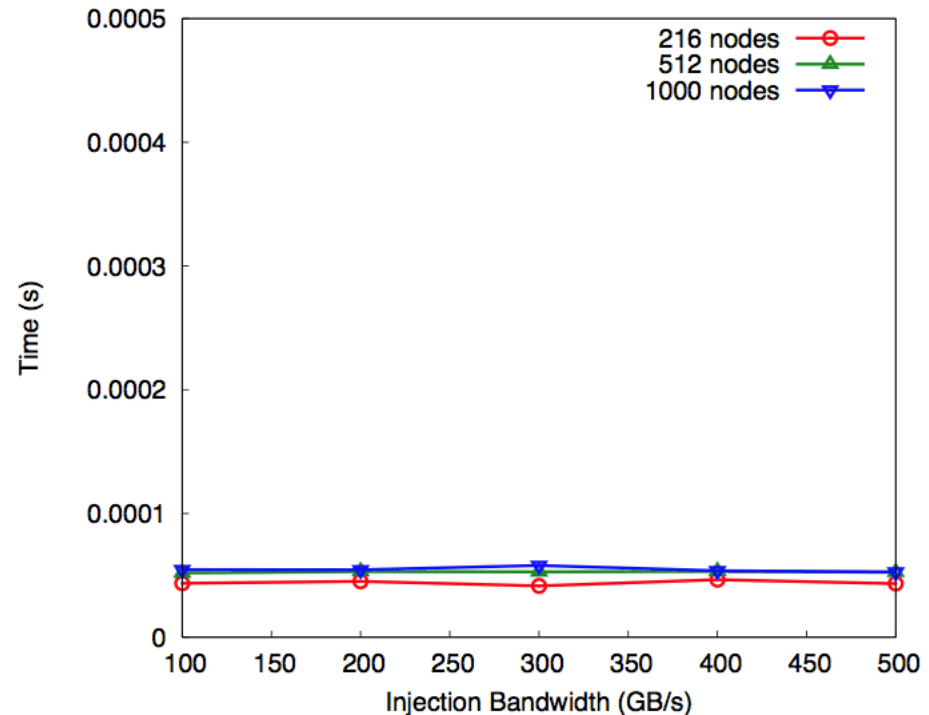
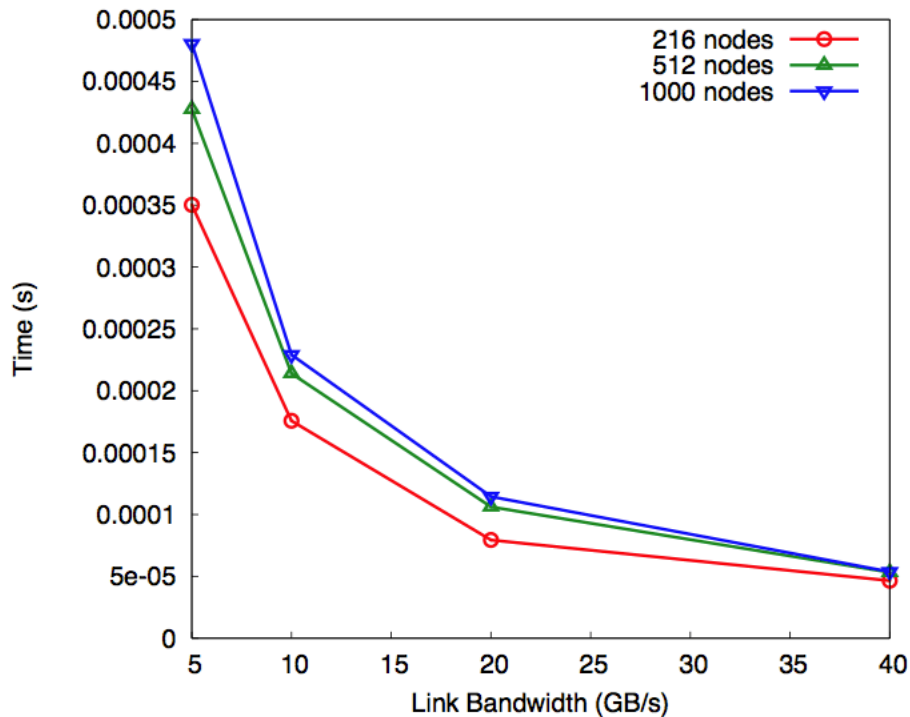
# 3D Torus: Injection Bandwidth



- 7-point stencil maps perfectly to 3D Torus, no link congestion for ideal machine
  - Accurate analytic model possible if injection effects are included

125 nodes  
n=100  
400ns latency

# Fat Tree Simulation: Machine Scale



- Regular stencil operations, conclusions from small scale simulations likely hold at large scale
  - Faults/machine noise are possible complications

n=100  
20ns latency  
400 GB/s injection BW

n=100  
20ns latency  
40 GB/s link BW

# Conclusions

- Important to balance injection and link bandwidths for optimal performance
  - Regular behavior of multigrid makes this easy for ideal machine
  - Analytic models need to account for injection effects
- Conclusions not strongly dependent on simulation size

# Next Steps

- Iterate on analytic models
- Examine additional topologies, job mapping
- Add compute modeling