# Modeling Communication Latency in High-Speed Interconnection Networks

Archit Patke[1], Saurabh Jha[1], James Brandt[2], Ann Gentile[2], Zbigniew Kalbarczyk[1], Ravishankar Iyer[1]
[1]*University of Illinois at Urbana-Champaign,* [2]*Sandia National Laboratories*

Sandia National Lab contract #1951381
NSF CNS 15- 13051

## 1 Motivation

### Problem Statement

Contention for network resources leads to application performance degradation in modern data centers.

Modeling and estimating communication primitive (e.g., send and receive) latency from congestion measurements are required to:

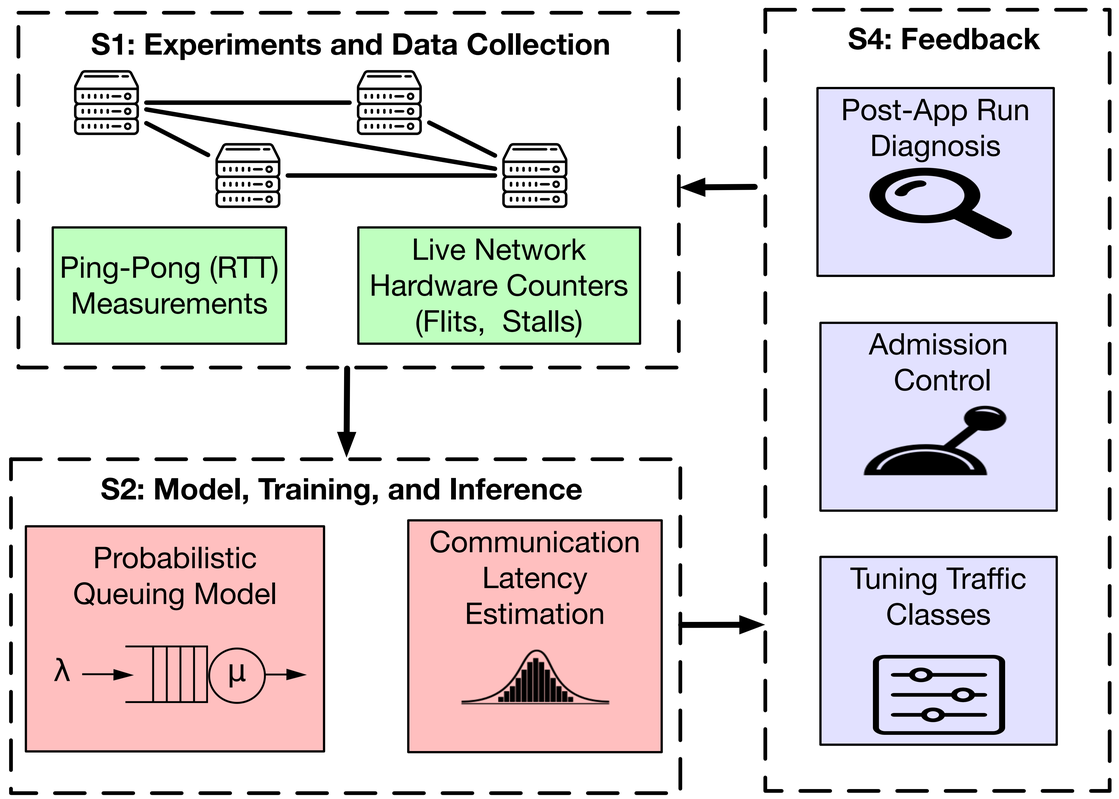
- Pinpoint sources of congestion
- Identify impact of congestion on applications
- Mitigate congestion impact

### Research Challenges

Estimating communication latency is difficult due to :

- Noisy measurements
- Adaptive routing and congestion control mechanisms
- Spatial and temporal congestion variation

## 2 Actionable Feedback from Monitoring




## 3 Latency Model

### Network Architecture

**Flow Control:** Lossless credit-based mechanism
**Topology:** Modified dragonfly
**Flits:** 48 bit data units
**Stalls:** The time links wait to send data

### Model Overview

**Input:** Stalls (s) and Flits (f) counters over a measurement interval, Shortest network paths (P)
**Output:** Communication latency estimate (L)
**Assumptions:**

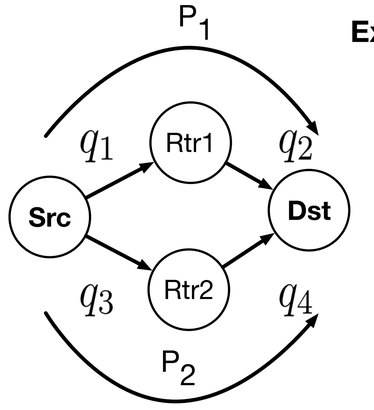
- Communication is iterative (repeats with interval ~s)
- Message sizes are small (~1kB)

### Queuing Model


$$\lambda = f \qquad \mu = f_{max}(1 - \frac{s}{s_{max}})$$

Arrival (λ) and Service (μ) rates for link buffers

### Example Case




Network paths (P) consists of paths $P_1$ and $P_2$.
P1 consists of queues $q_1$ and $q_2$
P2 consists of queues $q_3$ and $q_4$

### Latency Estimation

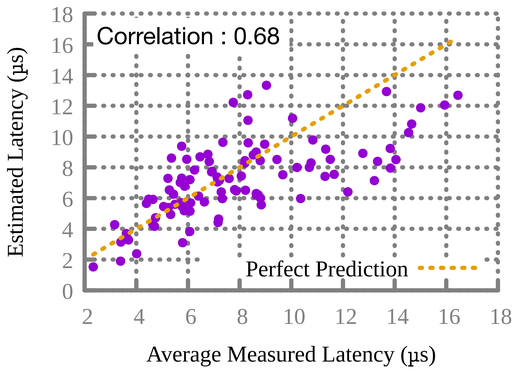
Expected waiting time $W_q$

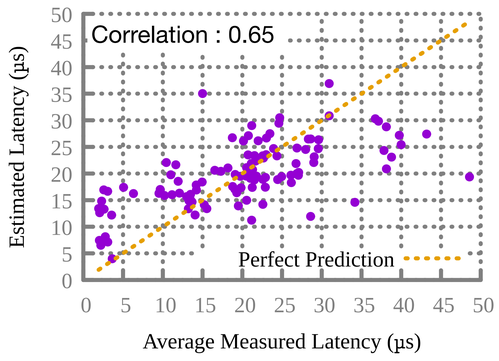
$$E[W_q] = \frac{\lambda}{\mu(\mu - \lambda)}$$

Estimating latency using individual queuing delays

$$L = c_1 \sum_{q \in P} W_q + c_2$$

## 4 Model Evaluation



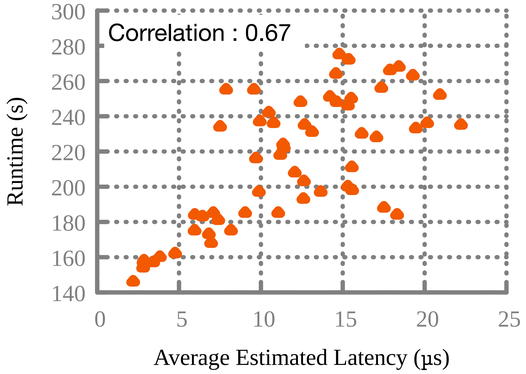




**Platforms:** Cray Aries high performance systems (48 and 236 compute nodes)

**Experiments:** 100 latency tests run along with congestors

**Network counters can be used to estimate communication latency.**

## 5 Impact on Applications




**Objective:** Relating latency estimates and app runtime
**Experiments:** 60 runs of a latency-sensitive application (MILC)

**Latency estimates correlate with application runtime.**