A Novel Scheduling and Queue Management Scheme for Multi-band Mobile Routers

Mohammed Atiquzzaman
Md. Shohrab Hossain
Husnu Saner Narman

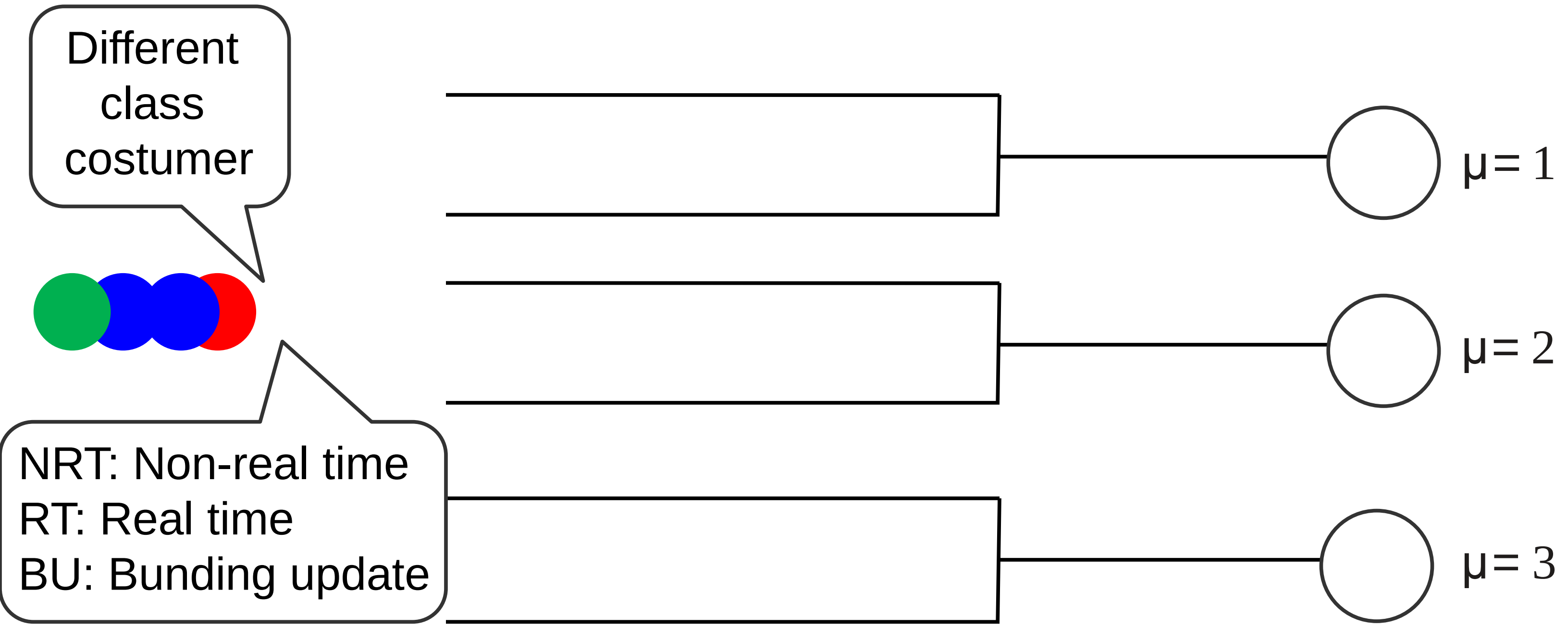
Telecommunications and Networking Research Lab
The University of Oklahoma
Norman, OK

IEEE ICC @ Budapest
June 12, 2013

Outline of the Presentation

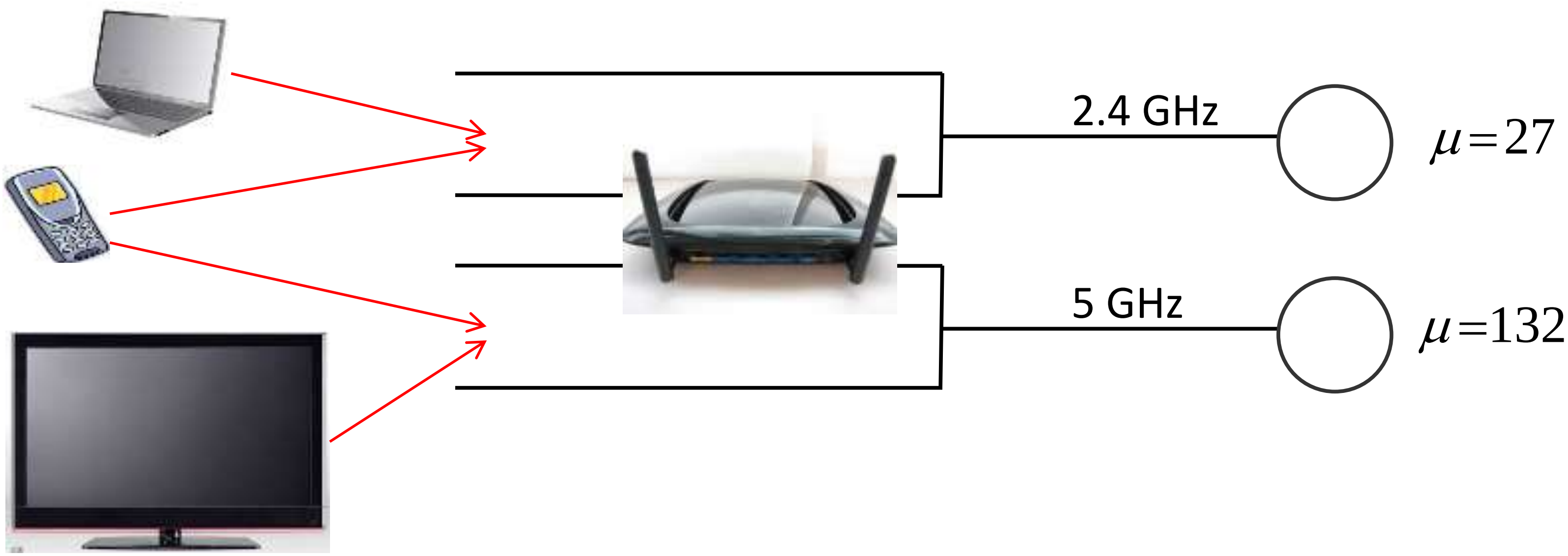
- Typical Multi-band Router Architecture
- Proposed Multi-band Router Architecture
- Analytical Model
- Results
- Conclusion

Heterogeneous Multi Server with Multi Class and Multi Queuing System



1) Priority = Which type of customer is served first?
2) Flexibility = Which type of customer will be served by which server?
3) Performance metrics for queues and classes

Current Multi Band Router System

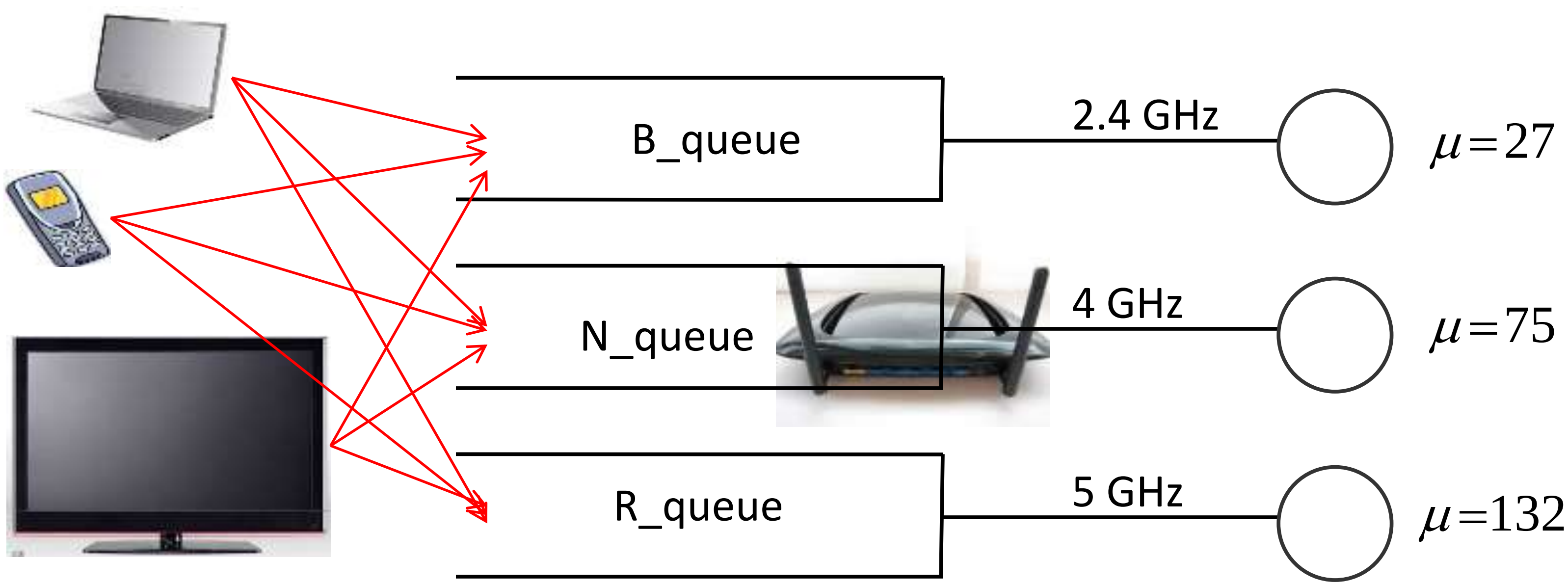


- Each device only can use one band at a time to send and recieve data.
- No sharing of traffic classes among the queues

Our Contribution

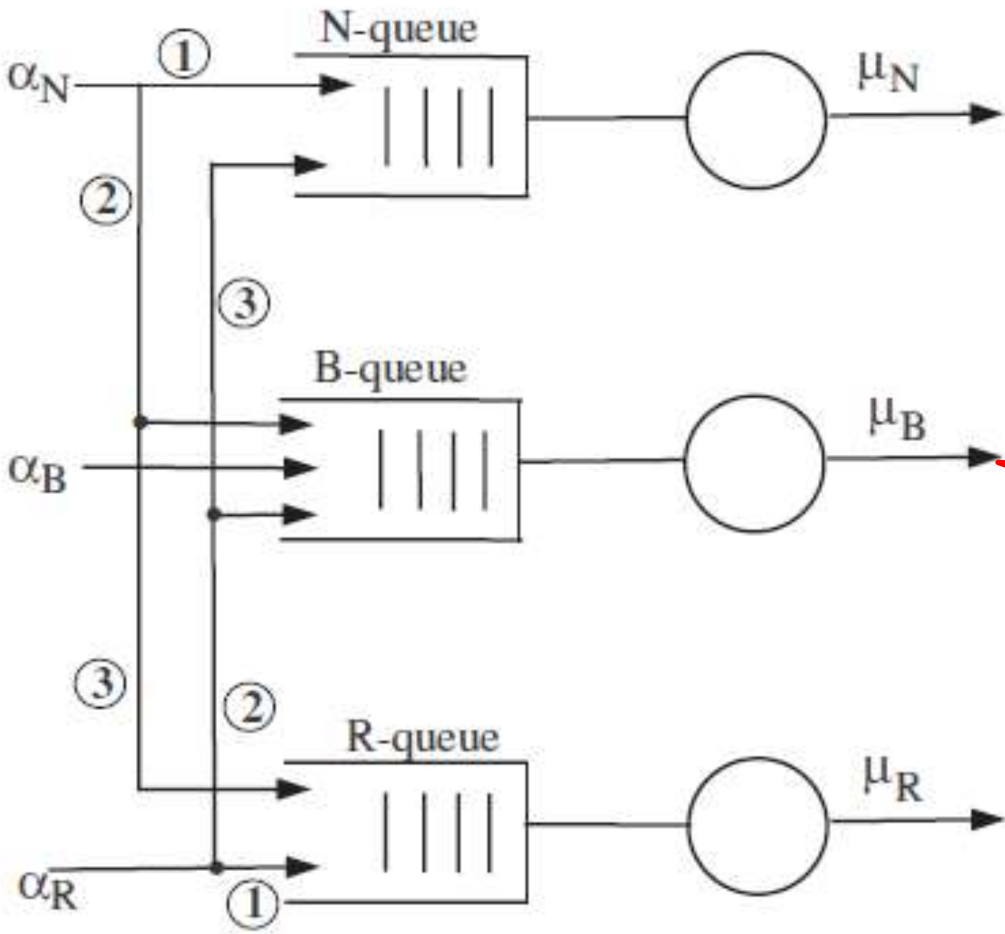
- Proposing a band-sharing router architecture and a novel scheduling algorithm to ensure maximum possible utilization of the system.
- Analytical model of the proposed multi-band system performance.
- Comparing the proposed router architecture with the typical one.

Proposed Multi Band Router System



- Each device can use all bands at the same time to send and recieve data.
- Each band carries only one class data.
- If needed, multiple class data can be transfered over one band.

Proposed Multi Band Queuing System



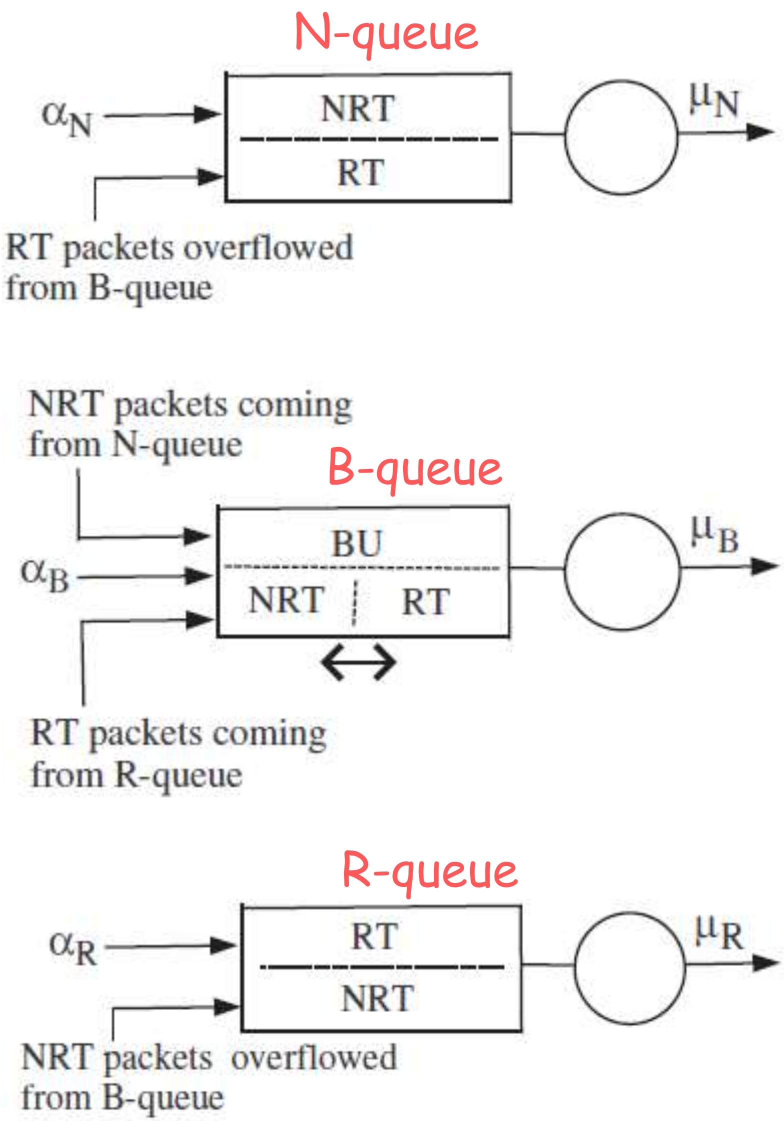
Each class has higher prioriity in its queue

Priory of NRT and RT packets in B-queue, dynamicaly calculated

$$\sigma_{N_{BQ}} = \frac{\alpha_R}{\alpha_B + \alpha_R + \alpha_N}$$

$$\sigma_{R_{BQ}} = \frac{\alpha_N}{\alpha_B + \alpha_R + \alpha_N}$$

Queue Management



Scheduling Algorithm

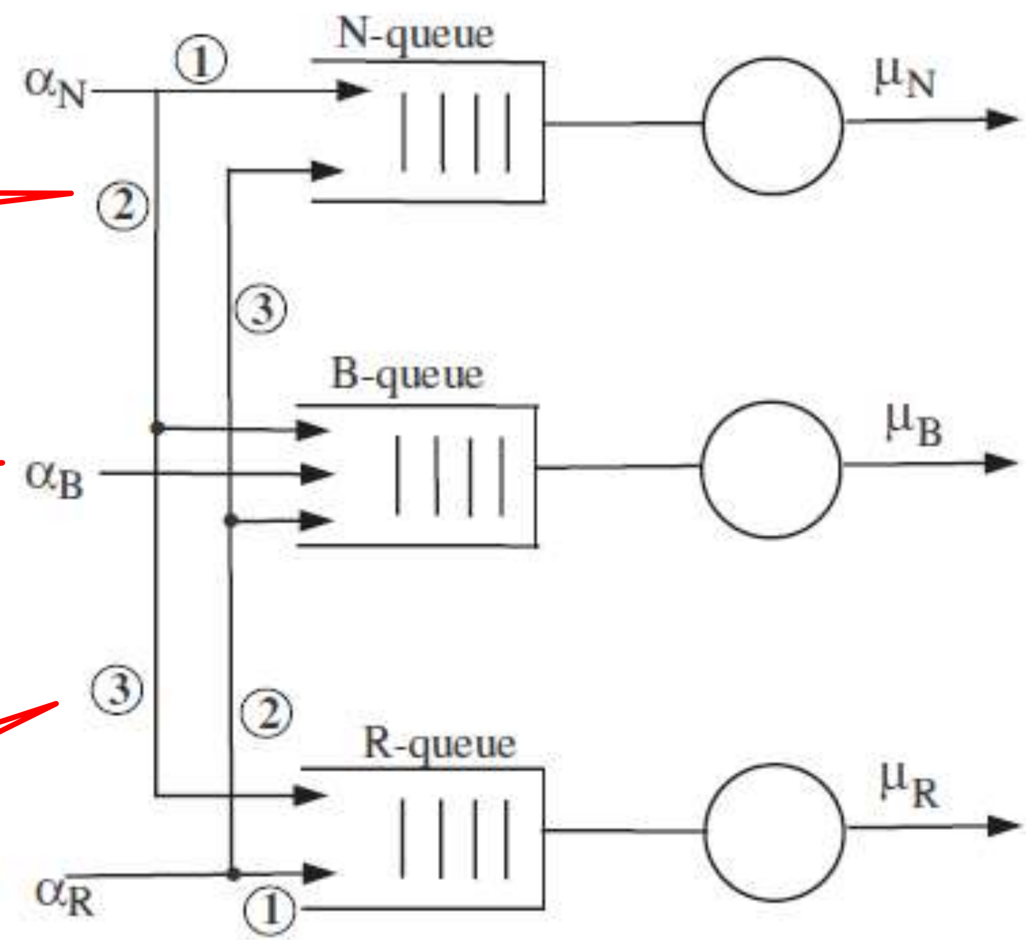
- Attempts are first made to queue different class of traffic in their corresponding buffer.
- If N-queue (or R-queue) overflows, traffic is forwarded to B-queue.
  - If insufficient space in B-queue, then overflowed NRT and RT packets compete in B-queue based on priority
- If overflowed RT packets cannot be accommodated in B-queue, they are queued in N-queue (if space)
  - If the RT packets cannot even be accommodated in N-queue, they are dropped from the system.
- Similar policy is enforced when dealing with NRT packets in the B-queue followed by R-queue.

Analysis of Queuing System

Average waiting time of each class and queue

Drop Rate of each class and queue

Average queue length of each class and queue



- Assumptions:
  - Packet arrivals are Poisson.
  - Queue discipline: FIFO with non-preemptive priority among various traffic classes.

- Notations

N_T	Queue size of T-queue in the MR,
α_T	Total packet arrival rate at T-queue of i-th MN,
μ_T	Service rate at T-queue of i-th MN,
$\sigma_{T_{BQ}}$	Priority of class-T traffic in B-queue,
$P_{dT_{XQ}}$	T-type packet drop probability in X-queue, where $X \in \{B, N, R\}$,
$E(D_T)$	Average queuing delay of class T packets,
$E(n_T)$	Average queue occupancy of class T packets,
P_{dT}^{sys}	Final packet drop probability of class T packets.

Analytical Model: Performance Metrics

- We have derived various performance metrics for the proposed multi-band MR architecture.
  - Packet drop probability
  - Average queue length
  - Average queue occupancy
  - Throughput
  - Average packet delay

- For example, Packet drop prob. of RT packets in B-queue

$$P_{dR_{BQ}} = \frac{(1-\rho_{BR})}{(1-\rho_{BR}^{N_B+2})}\rho_{BR}^{N_B+1} + \frac{\alpha_B}{\alpha_R'}\left(\frac{(1-\rho_{BR})}{(1-\rho_{BR}^{N_B+2})}\rho_{BR}^{N_B+1} - P_{dB_{BQ}}\right)$$

Analytical Model: Performance Metrics

- Average queue length of T-type packet:

$$E(n_T) = \begin{cases} \frac{\rho_T - (N_T+1)\rho_T^{N_T+1} + N_T \rho_T^{(N_T+2)}}{\left(1-\rho_T\right)\left(1-\rho_T^{N_T+1}\right)} & \text{, if } \rho_T \neq 1 \\ \frac{N_T}{2} & \text{, if } \rho_T = 1 \end{cases}$$

- Average queue occupancy of RT packets in the system:

$$E(n_R^{sys}) = E(n_R^{RQ}) + E(n_R^{BQ}) + E(n_R^{NQ})$$
$$= E(n_R^{RQ}) + \left(E(n_{B+R}^{BQ}) - E(n_B^{BQ})\right) + \left(E(n_{N+R}^{NQ}) - E(n_N^{NQ})\right)$$

- Throughput of T-type packets:

$$\gamma_T^{sys} = (1 - P_{dT}^{sys})\alpha_T$$

- Average packet delay of T-type packet:

$$E(D_T^{sys}) = \frac{E\left(n_T^{sys}\right)}{(1 - P_{dT}^{sys})\alpha_T}$$

Results

- We have used discrete event simulation in MATLAB following M/M/3/N procedures.
- Equal buffer length (of 50 packets) for each queue.
- RT and NRT packets: 512 bytes, BU packets:64 bytes.
- We ran each simulation for 20 trials having different traffic class arrival rates.

Traffic class arrival rates

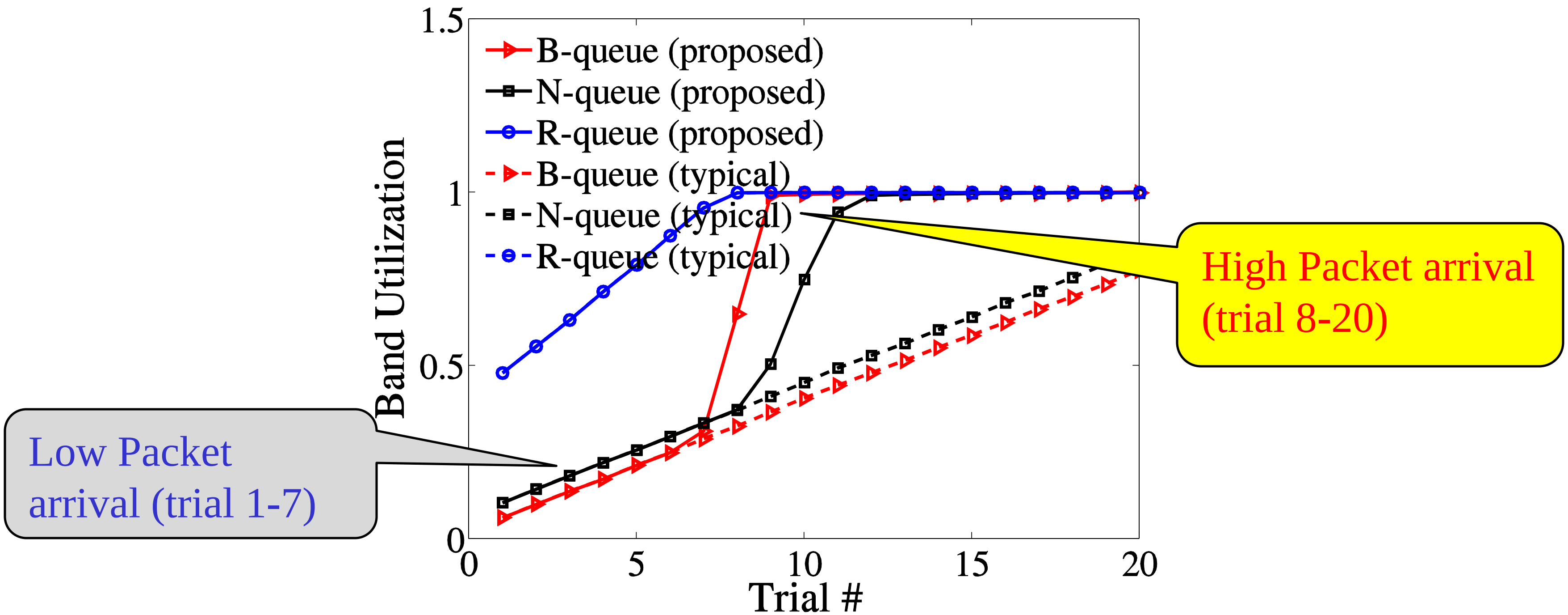
- Simulations were with increased arrival rates of all types of traffic to observe the impact of heavy traffic on the multi-band system.
- Traffic class arrival rates at different trials:

$$\lambda_B(i) = \{\ \mathrm{i}\ \},\ \lambda_N(i) = \{\ 3\mathrm{i}\ \},\ \lambda_R(i) = \{\ 18\mathrm{i}\ \}$$

where i = 1,2, …, 20

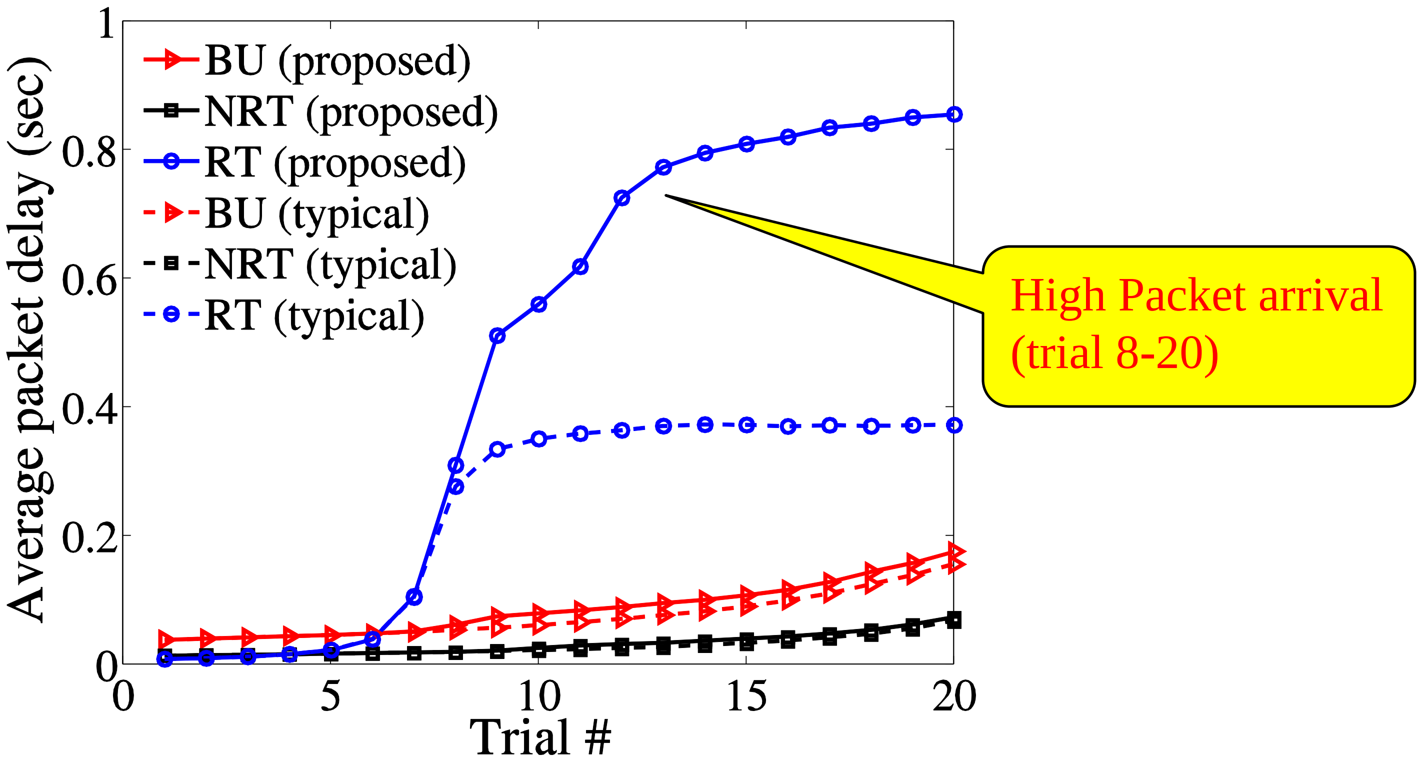
- The arrival rate of B-queue and N-queue are increased slowly in each trial whereas the RT traffic arrival rate are increased at a much higher rate.
  - This eventually saturates the R-queue and we explain the impact of this overflow on different performance parameters of our proposed system and typical multiband router.

Band Utilization



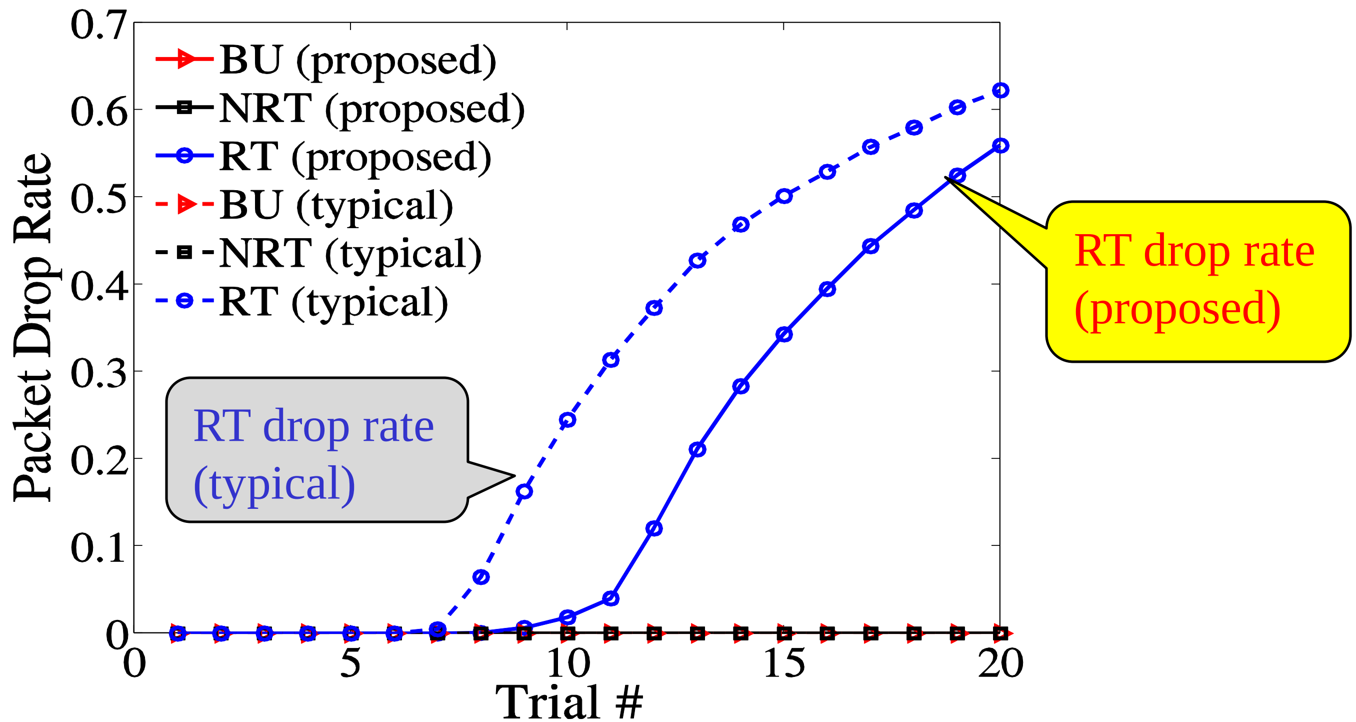
- <u>Low arrival rate</u>: Both architecture have similar utilization.
- <u>High arrival rates</u>: B-queue and N-queue utilizations are much higher for proposed architecture than for typical one
  - Reason: Increased number of RT packets are dropped in typical architecture whereas in proposed one, they are accommodated in B-queue and N-queue, thereby improving their utilizations and maximizing system performance.

Average Packet Delay



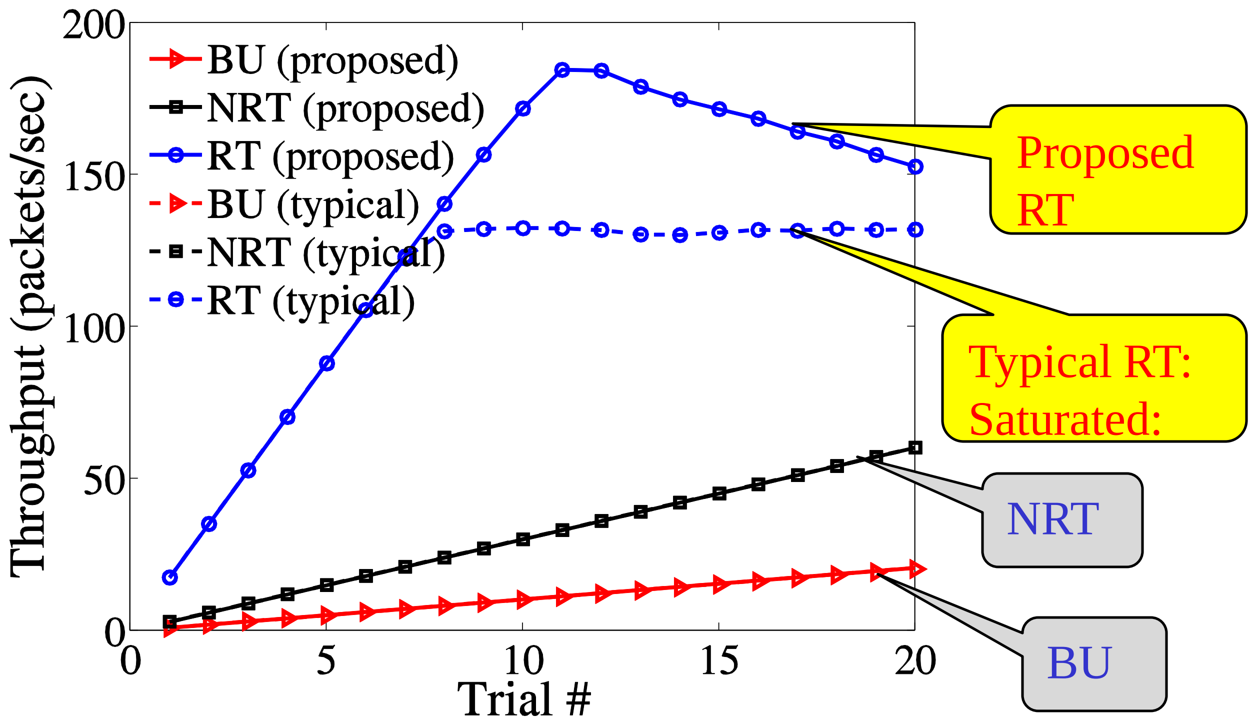
- During high arrival rates, the delay for RT traffic in proposed architecture is higher than the typical one.
  - Reason: excessive RT packets are immediately dropped in typical architecture and they are NOT considered in delay calculations.
  - In proposed architecture overflowed RT packets are queued in B and N-queues before being dropped, thereby increasing the delay.

Packet Drop rate



- For high RT arrival rates, RT packet drop rate gradually goes up for typical (non-shared) architecture.
- However, proposed architecture does not allow RT traffic to drop as long as they can be queued in B- and N-queues.

Throughput



- Throughput of NRT and BU class are increased with the increase of their arrival rates
- For RT class and for the typical architecture, the throughput is saturated
- RT class throughput (proposed architecture) go much higher (due to sharing of other under-utilized bands) and reaches its peak.
- After that it starts to decrease slowly due to the impact of increased arrival rates of other queues (B and N-queue).

Summary of Results

- Proposed architecture maximizes utilization through band sharing.

- For RT traffic:
  - Average queue occupancy and delay of RT traffic affected.
  - Packet drop and throughput significantly improved.

Conclusion

- Proposed scheduling algorithm for multi-band mobile routers that exploits band sharing.
- Developed analytical model of proposed multi-band system and validated by extensive simulations.
- Proposed architecture maximizes utilization through sharing of capacities among the bands
- Proposed scheduling algorithm can help network engineers build next generation mobile routers with higher throughput and utilization

Thank You

http://cs.ou.edu/~atiq

atiq@ou.edu