The Internet is (unfortunately) not yet the computer

Roch Guérin
Washington University in St. Louis

NSF Future Wireless Cities Workshop

February 2-3, 2016

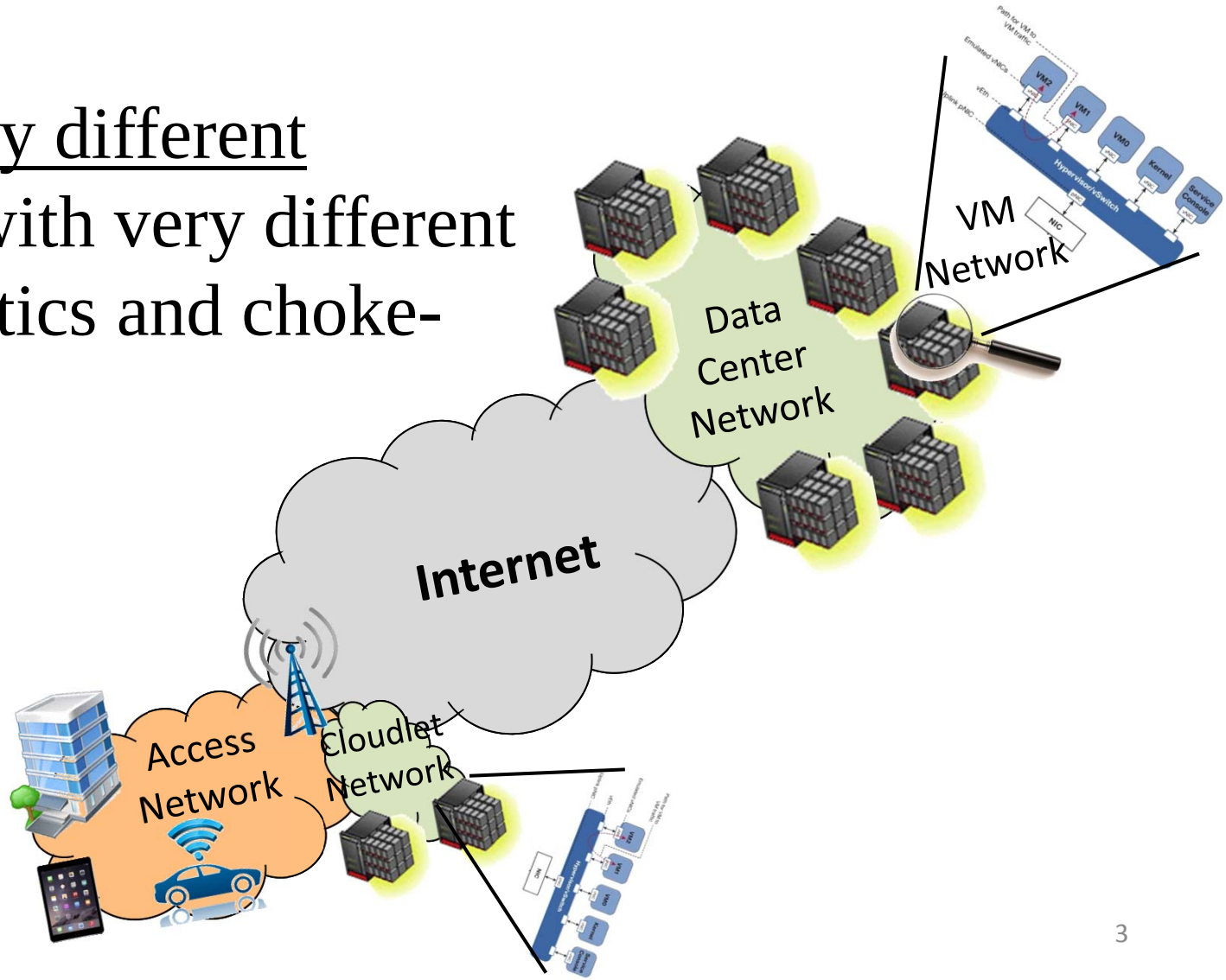
Acknowledgments

- The ideas outlined in this talk are based on joint projects with Chris Gill and Chenyang Lu at Washington University

[1] C. Li, S. Xi, C. Lu, C. Gill and R. Guerin, "Prioritizing Soft Real-Time Network Traffic in Virtualized Hosts Based on Xen." Proc. RTAS 2015, Seattle, WA, April 2015.

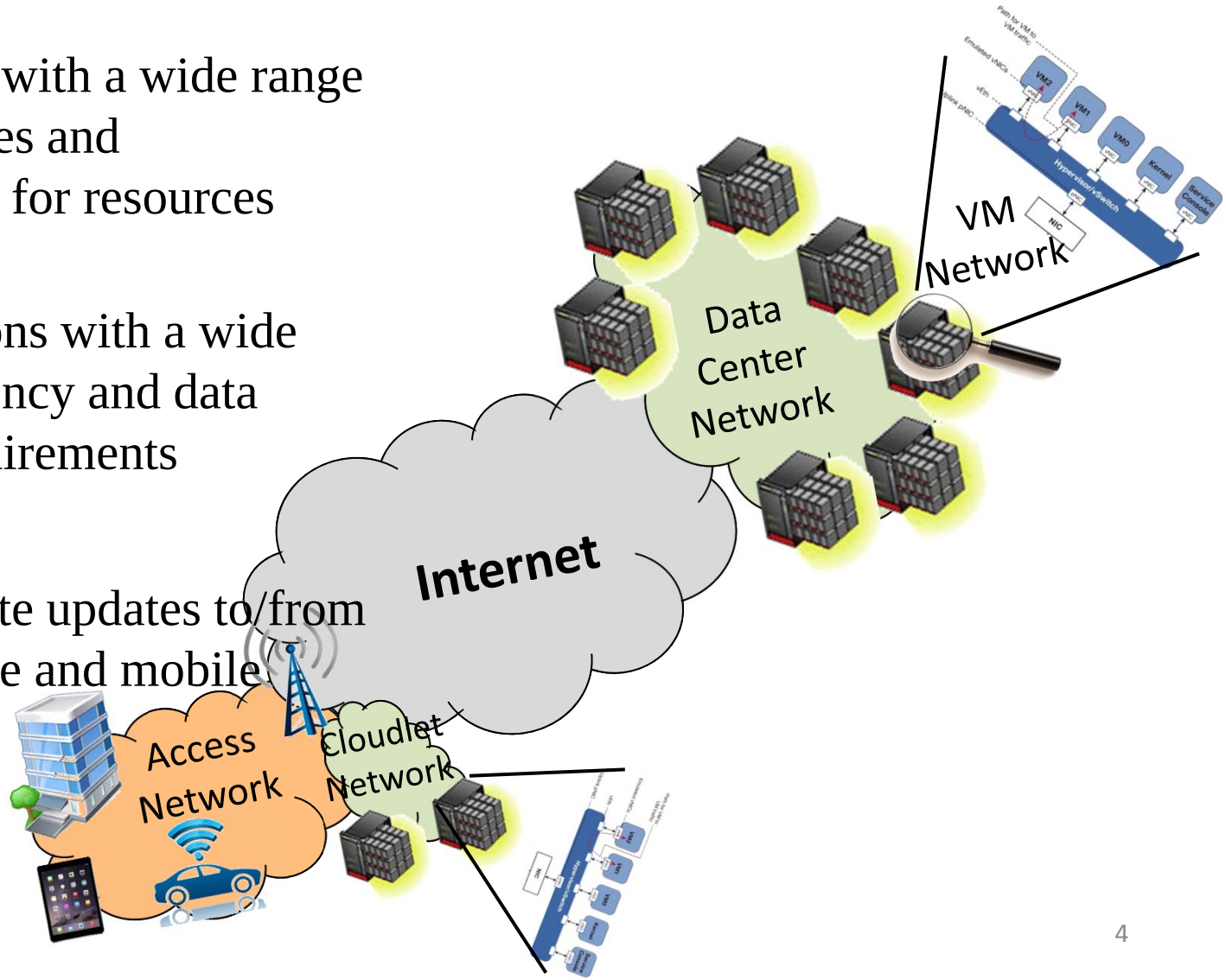
A High-Level View

- Lots of very different networks with very different characteristics and choke-points



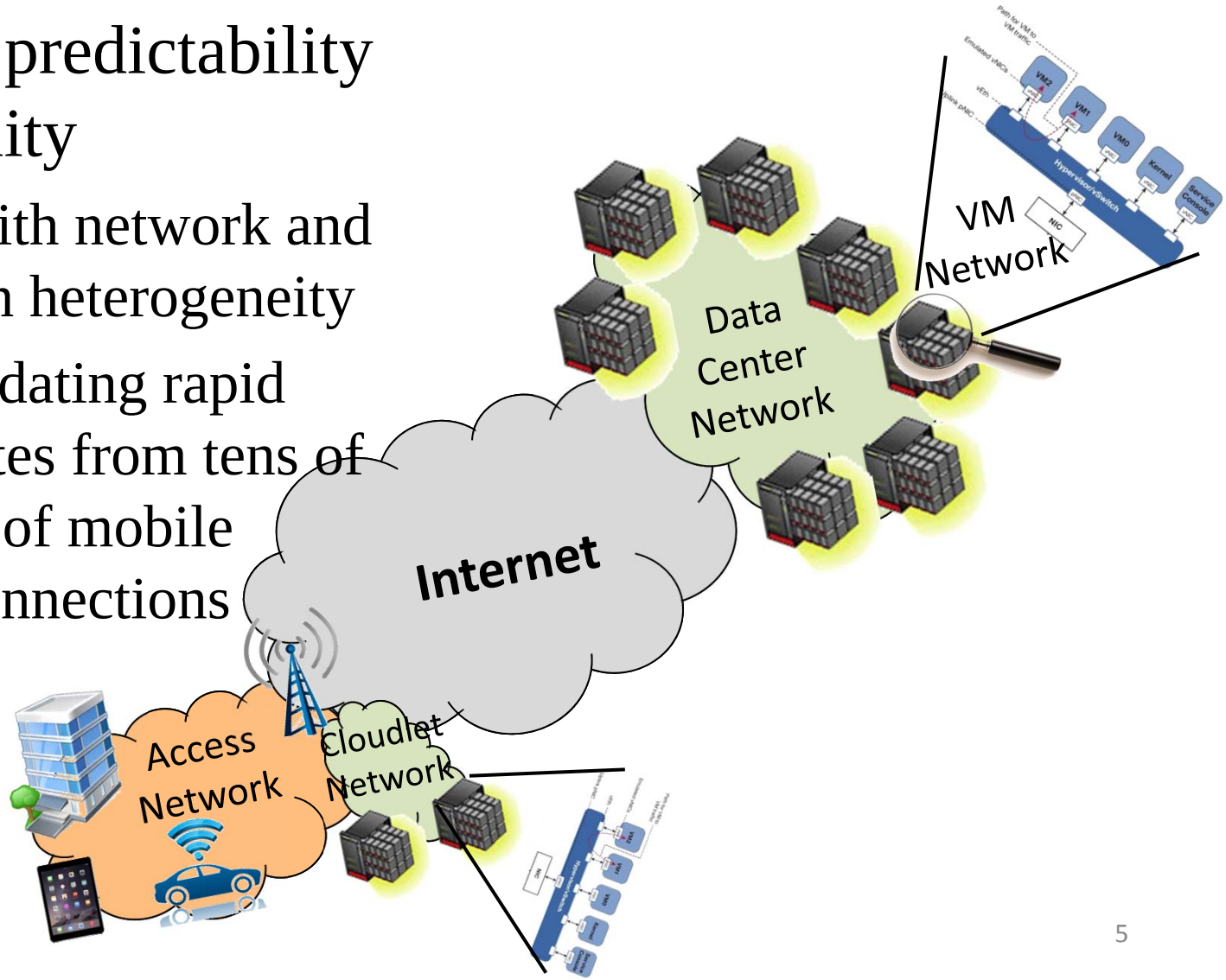
Premises & Pain Points

- Diversity
  - of networks with a wide range of capabilities and mechanisms for resources arbitration
  - of applications with a wide range of latency and data volume requirements
- Mobility
  - Frequent state updates to/from infrastructure and mobile devices



Goals

- End-to-end predictability and scalability
  - Dealing with network and application heterogeneity
  - Accommodating rapid state updates from tens of thousands of mobile devices/connections



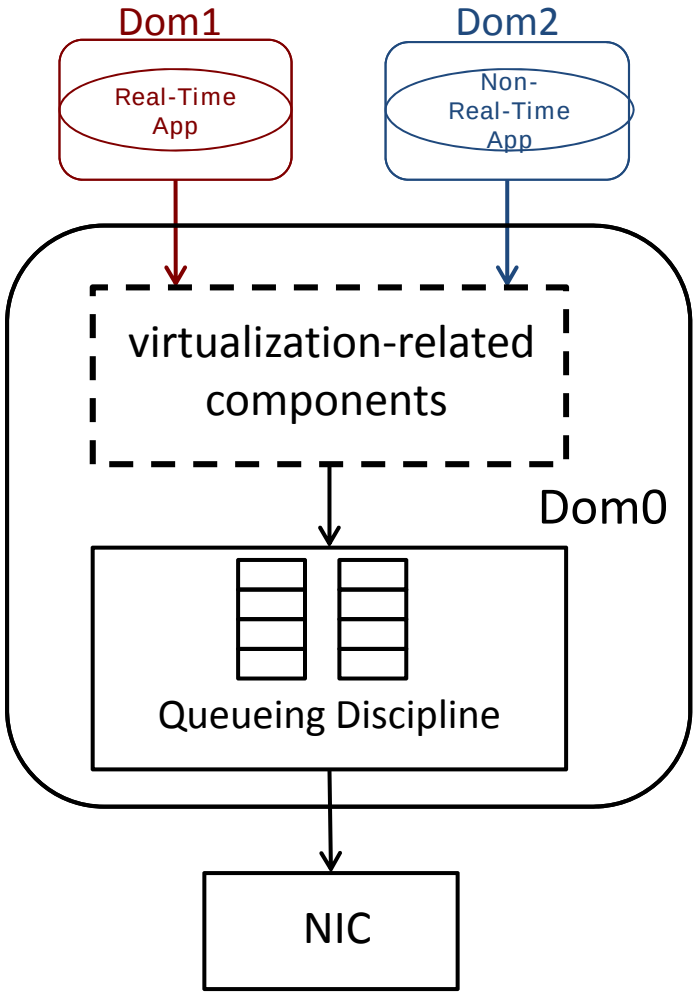
An Example & an Ongoing Investigation

- Latency guarantees across heterogeneous VMs
  - Focus on Xen, but similar issues in other virtualization systems
- Real-time messaging middleware
- Testbed implications
  - Easy to realistically test individual pieces
  - Hart to do holistic system-level tests

Network Components in Xen Virtualized Hosts

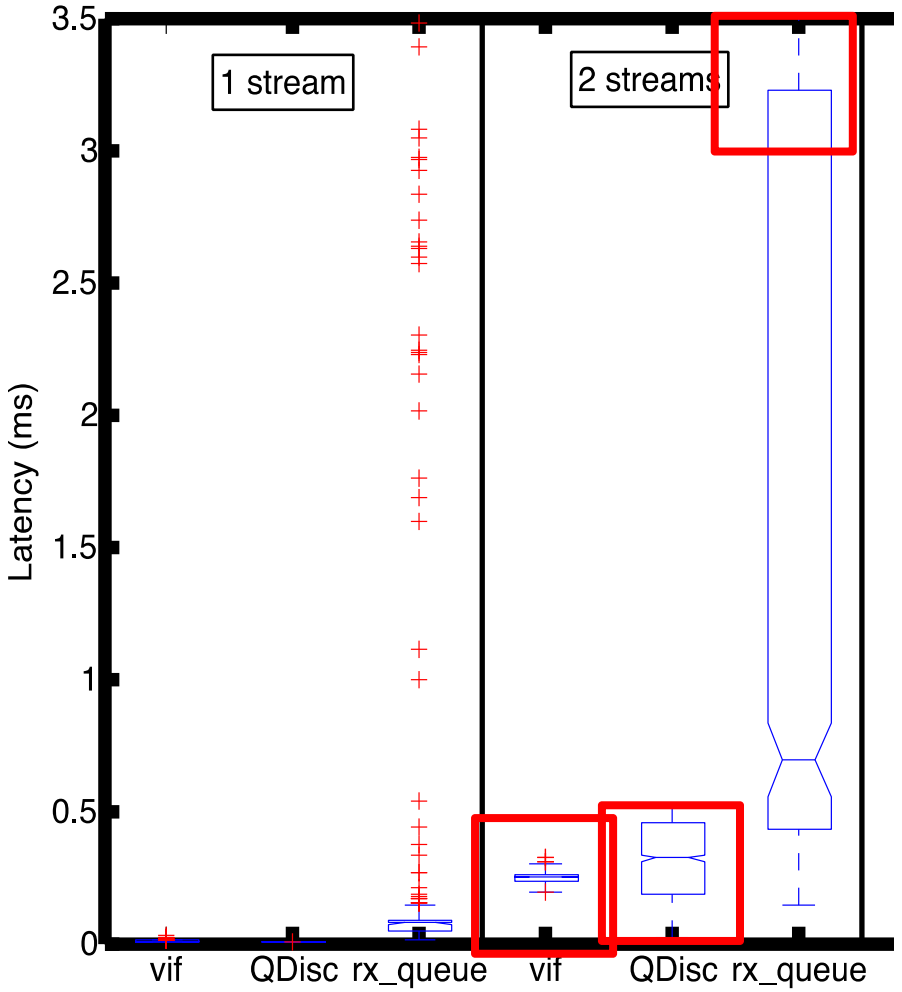
- Dom0: customized Linux VM

- Network components in dom0:
  - Queueing Discipline
  - Virtualization-related components
- Queueing Discipline (Qdisc) implements traffic control
  - Rate-limit and shape each flow
  - Prioritization or fair packet scheduling
- Problem: virtualization-related components
  - A variety of places where priority inversion can arise



Latency of Real-Time Traffic

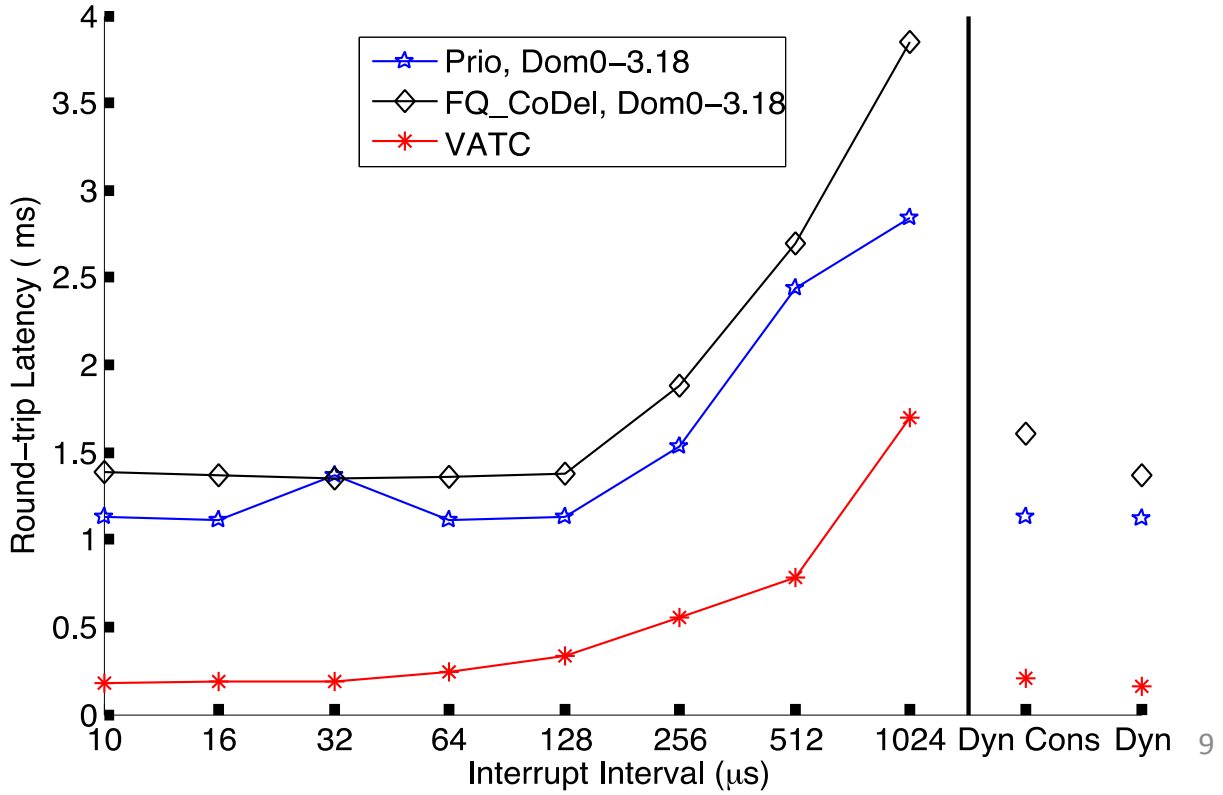
- Measure real-time traffic latency, with increasing number of interfering streams
- Characterize delay contributions of individual components
  - Virtual interface (vif) – impact of vif polling mechanism (poll_list)
  - QDisc layer – handling of interrupt handler controlling NIC buffer cleanup (NET_RX_SOFTIRQ) can create NIC congestion
  - rx_queue – rx_kthread delayed by NET_RX_SOFTIRQ handler



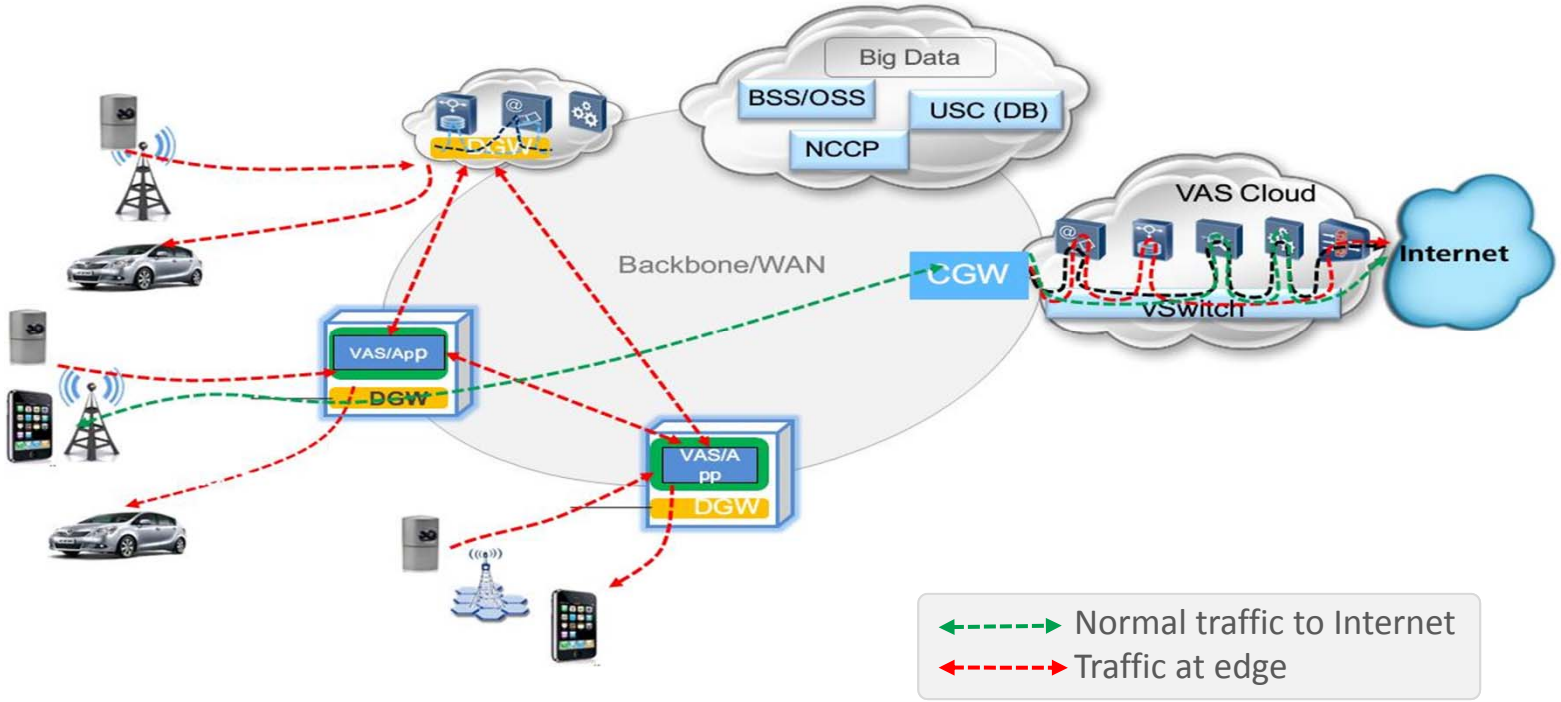
Mitigating Priority Inversion

- A thread-based solution
  - TX/RX handled by prioritized kernel threads
    - Dedicated tx_queue and rx_queue for each priority

Same scenario as before with one real-time stream and two interfering streams



A Typical 5G Scenario



- Distributed gateways (VNFs) and APPs are deployed at the edge
  - Close to mobile users → shorten transmission path → reduce latency
- Low-latency & light-weight messaging middleware is key for application coordination at the edge

Requirements & "Solutions"

- Key properties
  - Real-time
    - Millisecond latency
  - Lightweight
    - < 100 Mbytes memory consumption
    - < 10% of one CPU core
  - Distributed
    - Broker vs. brokerless vs. peer-to-peer
  - Fault-tolerance (persistency)
- Delivery options
  - Compiled in application
  - Micro-service

- A plethora of contenders
  - RabbitMQ
  - ZeroMQ (Nanomsg)
  - NSQ
  - Kafka
  - Aeron

- No systematic understanding of impact of architectural choices on performance, functionality, flexibility trade-off