On the Throughput Limits for Communication Networks

Kashif Mahmood*, Yuming Jiang *, Jeong-woo Cho *

*Centre for Centre for Quantifiable Quality of Service
in Communication Systems (Q2S), NTNU, Norway

EuroNF Workshop on Traffic Management and Traffic Engineering for the Future Internet.
December 7-8, 2009

Outline

- Problem Statement
  - Network Information Theory (NIT): The Grand Challenge
  - NIT: What and Where
  - Classical Information Theory (CIT): and its Limitations
- Modus Operandi
- Target Areas
  - Wired Networks under deterministic guarantees
  - Wired Networks under stochastic guarantees
  - Wireless Networks under stochastic guarantees
- Conclusions with a Look ahead...

NIT: Network Information Theory
CIT: Classical Information Theory

NIT: The Grand Challenge

- Internet as well as wireless services were initially designed to provide "Best Effort Service" with no guarantee on QOS.
- Emerging applications such as VOIP, multimedia conferencing, visualization, and virtual reality require guaranteed QOS.
- Before these real-time applications can be broadly used, a lot of challenges have to be addressed.

NIT: Network Information Theory

NIT: The Grand Challenge

1) The infrastructure must be modified to support real-time QoS, which provides some control over end-to-end packet delays.

2) Supporting mechanisms should be upgraded.

3) Network Information Theory should be developed to find out the capacity limits of the networks (wireless in particular) under QoS constraints.

- A lot of work has been done on 1 & 2
- 3 is a Grand Challenge [J.Andrew et.al, '03].

NIT: Network Information Theory

NIT: The Grand Challenge

1) The infrastructure must be modified to support real-time QoS, which provides some control over end-to-end packet delays.

2) Supporting mechanisms should be upgraded.

3) Network Information Theory should be developed to find out the capacity limits of the networks (wireless in particular) under QoS constraints.

- A lot of work has been done on 1 & 2
- 3 is a Grand Challenge [J.Andrew et.al, '03].

NIT: Network Information Theory

NIT: What and Where

- What?
  - Information Theory for the communication networks taking into account bursty nature and delay requirements of real sources.

- Where?
  - To study the queuing behavior of packet switched networks
  - To understand fundamental throughput and delay performance limits of communication networks.
  - In designing resource allocation algorithms (Power control, Medium Access, Routing ….)
  - In sizing up resources of the networks to meet the throughput limits.

- Why NIT or What are the limitations of Classical Information Theory (CIT)? …

NIT: Network Information Theory

CIT and its Limitations

- CIT* focuses on point-to-point, source-channel-destination model of communication, ignoring the bursty nature of real sources.
- Idle periods of source silence / inactivity are not considered.
- However, in networks source burstiness is the central phenomenon that underlies the process of resource sharing for communication.

- CIT has focused on asymptotic limits of the tradeoff b/w accuracy and rate of communication, ignoring the role of delay as a parameter that may effect this tradeoff.
- In networking, delay is a fundamental quantity that must be take into account.

*CIT: Classical Information Theory

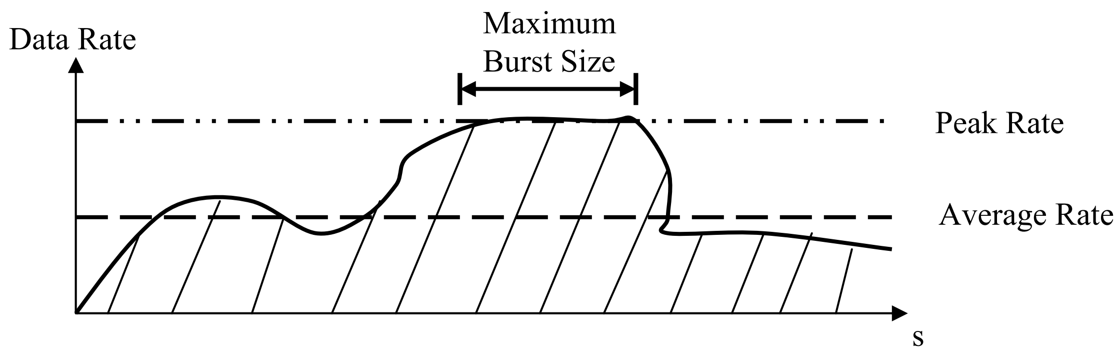
CIT and its Limitations

- CIT* focuses on point-to-point, source-channel-destination model of communication, ignoring the bursty nature of real sources.
- Idle periods of source silence / inactivity are not considered.
- However, in networks source burstiness is the central phenomenon that underlies the process of resource sharing for communication.

- CIT has focused on asymptotic limits of the tradeoff b/w accuracy and rate of communication, ignoring the role of delay as a parameter that may effect this tradeoff.
- In networking, delay is a fundamental quantity that must be take into account.

*CIT: Classical Information Theory

CIT and its Limitations



- Shannon capacity assumes infinite delay, so only average rate is important. Hence burstiness (peak rate) has been neglected.

$$C = \lim_{T \to \infty} \frac{\log_b M(T)}{T}$$

*CIT: Classical Information Theory

CIT and its Limitations

- So far the study of wireless networks has been conducted in the restricted framework of non-bursty and delay-insensitive sources.

- CIT has ignored networking concerns like bursty traffic, finite flows and sessions, queuing delay and routing [J.Andrew et.al, '03].

- NIT considering bursty traffic and delay constraints should be developed.

- What should be the modus operandi?

CIT: Classical Information Theory

CIT and its Limitations

- So far the study of wireless networks has been conducted in the restricted framework of non-bursty and delay-insensitive sources.
- CIT has ignored networking concerns like bursty traffic, finite flows and sessions, queuing delay and routing [J.Andrew et.al, '03].
- NIT considering bursty traffic and delay constraints should be developed.
- What should be the modus operandi?

CIT: Classical Information Theory

Outline

- Problem Statement
  - NIT the Grand Challenge
  - NIT, What and Where
  - CIT and its Limitations
- Modus Operandi
- Target Areas
  - Wired Networks under deterministic guarantees
  - Wired Networks under stochastic guarantees
  - Wireless Networks under stochastic guarantees
- Conclusions with a Look ahead...

Modus Operandi

- Method for quantifying the effective data rate of a bursty source should be identified.

- The following two have been proposed in literature:

1. Statistical Method based on theory of effective Bandwidth.

2. Calculus of Constraints (Deterministic or Stochastic).

Statistical Method

- Statistical Method is based on theory of effective bandwidth [Kelly,'96].

- Effective bandwidth of a source gives the amount of resources that must be reserved depending on the statistical properties and QoS requirements of a source.

$$\alpha_i(s,t) = \frac{1}{st} \log E\left[e^{sX_i[0,t]} \right]$$

- So, a variable data stream has an effective bandwidth (between peak and mean rate of stream) and characterizes the resource requirement of flow.

Calculus of Constraints

- Impose constraints on the data admitted into the Network.
- Network should then be able to provide guaranteed QOS.
- Constraints should fulfill the following requirements.

✓ Flexibility
✓ Easy to Enforce or Monitor.
✓ Operational significance to the Network.

Calculus of Constraints

- Constraints on the data admitted into the Network.
- Flexibility
  - Should Allow Controlled degree of Burstiness on part of data sources.
- Easy to Enforce or Monitor
  - Should be easy to police a data stream (through dropping or delaying part of stream) to produce output stream satisfying the constraints.
- Operational significance to the Network
  - Should be easy for the network to exploit constraints on admitted data streams in order to deliver performance guarantees.

Calculus of Constraints

- Basic Popular DataStream constraint is the (σ,ρ) constraint.
- Flexibility
  - Allows a stream to contain an occasional burst of size σ.
- Easy to Enforce or Monitor
  - Leaky bucket regulator is used in which tokens arrive at rate ρ
- Operational significance to the Network
  - When (σ,ρ) stream passes through a buffered link with constant service rate C, then the delay at the buffer is D and the output satisfies the (σ',ρ) constraint for $\sigma'=\sigma+\rho D$.

- Since seminal work on (σ, ρ)-constraint by Cruz, a lot of advancement has been made on QOS guarantees in the form of:
  - Deterministic Network Calculus which provides deterministic service guarantees
  - Stochastic Network Calculus which provides stochastic service guarantees.

Calculus of Constraints

- Basic Popular DataStream constraint is the (σ,ρ) constraint.
- Flexibility
  - Allows a stream to contain an occasional burst of size σ.
- Easy to Enforce or Monitor
  - Leaky bucket regulator is used in which tokens arrive at rate ρ
- Operational significance to the Network
  - When (σ,ρ) stream passes through a buffered link with constant service rate C, then the delay at the buffer is D and the output satisfies the (σ´,ρ) constraint for σ´=σ+ρD.

- Since seminal work on (σ, ρ)-constraint by Cruz, a lot of advancement has been made on QOS guarantees in the form of:

  - Deterministic Network Calculus which provides deterministic service guarantees
  - Stochastic Network Calculus which provides stochastic service guarantees.

Calculus of Constraints

- Communication network consists of data flows and network (service) elements.

- Correspondingly, a theory for network analysis is typically built on two fundamental concepts: traffic model and server model.

  – Traffic model characterizes the traffic behavior of a flow.
  – Server model characterizes the service behavior of a network element.

- Arrival Curve α : Flow satisfies, for any $0 \le s \le t$,

- Server provides Service Curve β iff for all $t \ge 0$, output satisfies

Calculus of Constraints

- Communication network consists of data flows and network (service) elements.
- Correspondingly, a theory for network analysis is typically built on two fundamental concepts: traffic model and server model.
  - Traffic model characterizes the traffic behavior of a flow.
  - Server model characterizes the service behavior of a network element.
- Arrival Curve α : Flow satisfies, for any $0 \le s \le t$,

$$A(s,t) \le \alpha(t-s) \quad \Leftrightarrow \quad A(t) \le A \otimes \alpha(t)$$

- Server provides Service Curve β iff for all $t \ge 0$, output satisfies

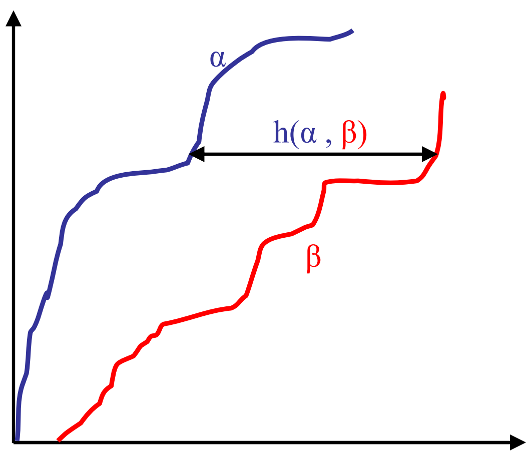
$$A^*(t) \ge A \otimes \beta(t)$$

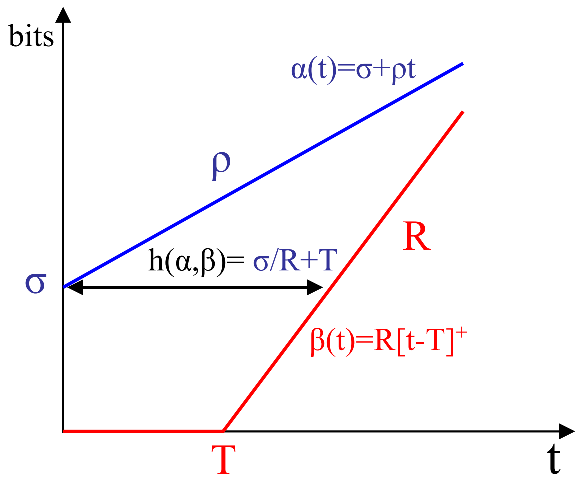
Min-plus convolution

$$(f \otimes g)(t) = \inf_{0 \le s \le t} \{f(t-s) + g(s)\}$$

Network Calculus: A Primer

- Given the arrival and service curve
  - Maximum packet delay experienced by a packet is the
    - Maximum horizontal distance





$$d(t) \leq h(\alpha, \beta)$$

$$h(\alpha,\beta) = \sup_{s \geq 0}\left\{\inf\left\{\tau \geq 0 : \alpha(s) \leq \beta(s+\tau)\right\}\right\}$$

Outline

- Problem Statement
  - NIT the Grand Challenge
  - NIT, What and Where
  - CIT and its Limitations
- Modus Operandi
- Target Areas
  - Wired Networks under deterministic guarantees
  - Wired Networks under stochastic guarantees
  - Wireless Networks under stochastic guarantees
- Conclusions with a Look ahead...

Throughput Limits for Wired Networks Under Deterministic Service Guarantees

Throughput Limits for Wired Networks Under Deterministic Service Guarantees

- A DSG guarantees that all packets of a flow arrive at the destination within their required performance measures such as throughput, delay and loss bounds in the network

- Pr [Experienced Performance ≥ Desired Performance] = 1

Problem Definition

- Given the
  - Arrival process
  - Service Process*
  - Delay Guarantee (D_g)
- What is the achievable throughput?

- Definition:
  - Maximum arrival rate which can be supported by the node for a given delay guarantee.
- Crucial Step:
  - Finding the maximum delay experienced by a packet.
  - D_g can only be provided if $D \leq D_g$

Throughput of a Node

- Arrival process (σ,ρ)-constrained
- Service Process* (Constant rate server $\beta(t) = Rt$)
- Delay Guarantee (D_g)

- What is the achievable throughput?
- Steps:
  - Maximum Delay (D) experienced by a packet of a flow
    - Tool: Deterministic NetCal or GR analysis
  - For n-flows
  - D_g can only be provided if $D \leq D_g$
  - Combining with stability constraint
  - ($n^*\rho$) gives us throughput

*Fluid Model

Throughput of a Node

- Arrival process (σ,ρ)-constrained
- Service Process* (Constant rate server $\beta(t) = Rt$)
- Delay Guarantee (D_g)

- What is the achievable throughput?
- Steps:
  - Maximum Delay (D) experienced by a packet of a flow $D = \frac{\sigma}{R}$
    - Tool: Deterministic NetCal or GR analysis
  - For n-flows $D = \frac{n\sigma}{R}$
  - D_g can only be provided if $D \le D_g$
  - Combining with stability constraint
  - (n*ρ) gives us throughput

$$\frac{n\sigma}{R} \le D_g \Rightarrow n \le \frac{D_g}{\sigma / R}$$

$$n \le \min\left\{\frac{D_g}{\sigma / R}, \frac{R}{\rho}\right\}$$

*Fluid Model

Throughput of a Network Under Deterministic Service Guarantees*

- Given
  - ➢ EF class in which each flow is (σ,ρ)-constrained at Ingress Node.
  - ➢ Service process (Guaranteed rate Scheduling).
  - ➢ FIFO aggregation.
  - ➢ Delay Guarantee.
  - ➢ Network of Arbitrary topology
- What is the achievable throughput?
- Intermediate Steps:
  - Arrival process for the EF aggregate.
  - Finding the maximum delay experienced by a packet.
    - Deterministic Network Calculus (DNC).
    - GR analysis
  - $D \leq D_g$
- What if we allow some packets to violate the D_g?

*Submitted to IEEE Sarnoff Symposium'10

Throughput of a Network Under Deterministic Service Guarantees*

- Given
  - EF class in which each flow is (σ,ρ)-constrained at Ingress Node.
  - Service process (Guaranteed rate Scheduling).
  - FIFO aggregation.
  - Delay Guarantee.
  - Network of Arbitrary topology
- What is the achievable throughput?
- Intermediate Steps:
  - Arrival process for the EF aggregate.
  - Finding the maximum delay experienced by a packet.
    - Deterministic Network Calculus (DNC).
    - GR analysis
  - $D \leq D_g$
- What if we allow some packets to violate the D_g?

*Submitted to IEEE Sarnoff Symposium'10

Throughput Limits
Under Stochastic Service Guarantees

- A SSG allows the QoS objectives specified by a flow to be guaranteed with a probability smaller than one
  - Pr [Experienced Performance < Desired Performance] = ε

Stochastic Service Guarantees

- Sources of Multiplexing Gain
  1. Scheduling
  2. Traffic characterization and conditioning
  3. Statistical Multiplexing
- Statistical Multiplexing outperforms 1 & 2 at high data rates [Ciucu, Burchard, Liebeherr,'05].
- Deterministic NetCal (DNC) has no statistical multiplexing gain pessimistic.
- Stochastic NetCal = Statistical Multiplexing gain + DNC.

Stochastic Service Guarantees

- Sources of Multiplexing Gain
  1. Scheduling
  2. Traffic characterization and conditioning
  3. Statistical Multiplexing

- Statistical Multiplexing outperforms 1 & 2 at high data rates [Ciucu, Burchard, Liebeherr,'05].

- Deterministic NetCal (DNC) has no statistical multiplexing gain ➡ pessimistic.

- Stochastic NetCal = Statistical Multiplexing gain + DNC.

Stochastic Network Calculus

- What?
  - SNC is essentially a method based on deriving the distribution functions for stochastic service guarantee.

- Where?
  - For flows with laxer requirements (Multimedia flows) stochastic service guarantees are more important, as they make better use of multiplexing gain and network resources.
  - For CSMA/CD; allocated BW to a host is highly affected by load from other hosts.
  - For wireless networks as they provide stochastic service guarantee, hence forth SNC is the best option.

Stochastic Network Calculus

- What?
  - SNC is essentially a method based on deriving the distribution functions for stochastic service guarantee.

- Where?
  - For flows with laxer requirements (Multimedia flows)
  - For CSMA/CD; allocated BW to a host is highly affected by load from other hosts.
  - For wireless networks as they provide stochastic service guarantee, hence forth SNC is the best option.

Stochastic Network Calculus: Primer

- Min-plus stochastic Arrival Curve

$$P\{A(t) - A \otimes \alpha(t) > x\} \le f(x)$$

- Min-plus stochastic Service Curve

$$P\{A \otimes \beta(t) - A^*(t) > x\} \le g(x)$$

- Delay Bound

$$P\{d(t) > h(\alpha + x, \beta)\} \le f \otimes g(x)$$

DNC

$$A(t) \le A \otimes \alpha(t)$$

$$A^*(t) \ge A \otimes \beta(t)$$

$$d(t) = \inf\{\tau \ge 0 : A(t) \le A^*(t+\tau)\}$$

$$\{d(t) > x\} \subset \{A(t) > A^*(t+x)\}$$

$$(f \otimes g)(t) = \inf_{0 \le s \le t} \{f(t-s) + g(s)\}$$

Throughput Limits for Wireless Networks

Wireless Networks

- Future wireless systems need to handle diverse multimedia traffic.
  - Packet switching
    - Queuing Behavior (QB)
  - Current physical models donot capture QB
- Early Work
  - For wireless network with n stationary $\Theta(W/\sqrt{n \log n})$ nodes transmiting at W bits/sec. [P:Gupta & P.R Kumar,'00]
  - Delay and throughput with n stationary nodes [A.E.Gamal $D(n) = \Theta(nT(n))$
  - Both these works donot consider queuing delay
  - [Kim & Krunz,00] consider QB but for limited traffic sources
- Also, wireless domain is challenging
  - Due to the time varying nature of wireless channels.
  - So, wireless channel provides stochastic service guarantees.

$f(n) = \Theta(g(n)) \Rightarrow \exists c_1, c_2, n_o > 0$ s.t. $c_1 g(n) \leqslant f(n) \leqslant c_2 g(n) \forall n \geqslant n_0$.

Throughput Limits for wireless Networks

- SNC takes into account
  - the queuing behavior
  - The stochastic nature of wireless channel
- Many traffic sources can be mapped to existing models [SNC Book; Jiang & Liu,'08].
- Provides decoupling of arrival and service processes
- Application of stochastic network calculus to wireless channels.
  - Network Calculus with MGF [Fidler, '06].
  - Impairment process based [Jiang & Emstad,'05].

SNC: Stochastic Network Calculus

Throughput Limits for wireless Networks

- SNC takes into account
  - the queuing behavior
  - The stochastic nature of wireless channel
- Many traffic sources can be mapped to existing models [SNC Book; Jiang & Liu,'08].
- Provides decoupling of arrival and service processes
- Application of stochastic network calculus to wireless channels.
  - Network Calculus with MGF [Fidler, '06].
  - Impairment process based [Jiang & Emstad,'05].

SNC: Stochastic Network Calculus

Network Calculus with MGF

- Network Calculus with MGF [Fidler, '06].
- Delay bound for FIFO scheduling in wireless networks.

$$d = \inf_{\theta>0}\{\inf[\tau : \frac{1}{\theta}(\ln \sum_{s=\tau}^{\infty} \mathrm{M}_A(\theta, s-\tau)\overline{\mathrm{M}}_S(\theta, s) - \ln\varepsilon) \leq 0]\}$$

$$\mathrm{M}_A(\theta, t) = \mathbb{E}[e^{\theta A(0,t)}]$$

$$\mathrm{M}_S(\theta, t) = \mathbb{E}[e^{\theta S(0,t)}]$$

$$\overline{\mathrm{M}}_S(\theta, t) = \mathrm{M}_S(-\theta, t)$$

- Challenge: Finding MGF for service process by modeling wireless link as finite state Markov channel specially for multihop networks.

Impairment process based

- Impairment process based [Jiang & Emstad,'05].

- Service behavior of server is characterized by two stochastic processes: an ideal service process and an impairment process.

$$S(s,t) \geq \hat{\beta}(t-s) - I(s,t)$$

- Find the stochastic nature of the impaired service caused by the bad link condition.
- Then map it to stochastic service model

- Alternate: Time domain modeling: [Jing & Jiang,'09]
  - Description in terms of time.
  - Avoids intermediate conversion.
- MGF method may be used to find the accurate bounding function for time domain modeling.

Impairment process based

- Impairment process based [Jiang & Emstad,'05].

- Service behavior of server is characterized by two stochastic processes: an ideal service process and an impairment process.

$$S(s,t) \geq \hat{\beta}(t-s) - I(s,t)$$

- Find the stochastic nature of the impaired service caused by the bad link condition.
- Then map it to stochastic service model

- Alternate: Time domain modeling: [Jing & Jiang,'09]
  - Description in terms of time.
  - Avoids intermediate conversion.
- MGF method may be used to find the accurate bounding function for time domain modeling.

Outline- - FinalLine

- Problem Statement
  - NIT the Grand Challenge
  - NIT, What and Where
  - CIT and its Limitations
- Modus Operandi
- Target Areas
  - Wired Networks under deterministic guarantees
  - Wired Networks under stochastic guarantees
  - Wireless Networks under stochastic guarantees

Sincerely Yours

Suggestions are most welcome