# A min-plus system interpretation of available bandwidth estimation

Markus Fidler

Multimedia Communications Lab
Technical University Darmstadt

Joint work with
Jörg Liebeherr and Shahrokh Valaee
University of Toronto



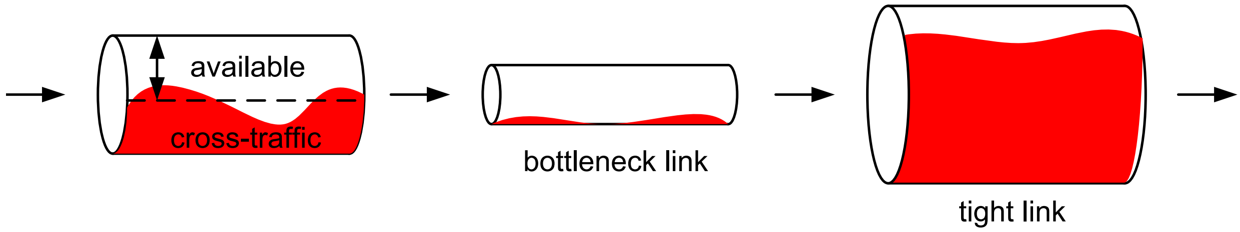


## Outline

# The task of available bandwidth estimation

Available bandwidth estimation seeks to infer the residual capacity that is leftover by cross-traffic along a network path from traffic measurements at the network ingress and at the network egress:

- Passive measurements monitor life traffic
- Active measurements inject artificial probing traffic




- Bottleneck link: Link that has the minimum capacity
- Tight link: Link that has the minimum available bandwidth





# Definition of available bandwidth

Utilization of link $i$ in the interval $[t_0, t_0 + \tau]$

$$u_i(t_0, t_0 + \tau) = \frac{1}{\tau} \int_{t_0}^{t_0+\tau} u_i(t) dt$$

Available bandwidth of link $i$ with capacity $C_i$

$$AvBw_i(t_0, t_0 + \tau) = C_i(1 - u_i(t_0, t_0 + \tau))$$

End-to-end available bandwidth of a network path

$$AvBw(t_0, t_0 + \tau) = \min_{i=1\dots n} \{AvBw_i(t_0, t_0 + \tau)\}$$

Common assumption: Traffic is viewed as constant rate fluid.

$$AvBw_i = C_i(1 - u_i)$$

Common assumption: There exists only a single bottleneck link.

$$AvBw = \min_{i=1\dots n} \{AvBw_i\}$$

## Available bandwidth estimation methods

Areas of application of available bandwidth estimation include:

- congestion control, e.g. TCP
- quality of service
  - measurement-based admission control
  - service level agreement verification
- network monitoring
- capacity provisioning
- traffic engineering

A variety of methods for available bandwidth estimation exist, which, however, resort to characteristic types of probing traffic:

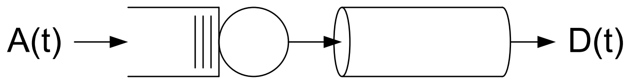
- packet pairs, e.g. Spruce [Strauss, Katabi, Kaashoek, IMC'03]
- packet trains, e.g. Pathload [Jain, Dovrolis, SIGCOMM'02]
- packet chirps, e.g. Pathchirp [Ribeiro et al., PAM'03]



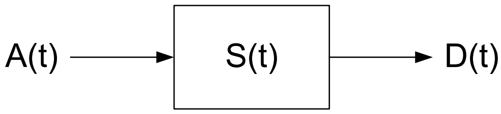


## Outline

# Definition of service curve




Network calculus abstracts queues, schedulers, and links as systems that are characterized by a service curve.




A system has a lower service curve $S(t)$ if it holds for all pairs of arrivals and departures $(A, D)$ of the system and all $t \geq 0$ that

$$D(t) \geq \inf_{\tau \in [0,t]} \{A(\tau) + S(t-\tau)\} = A \otimes S(t)$$

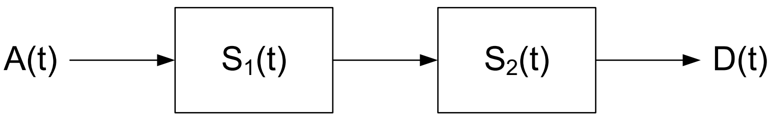
where the operator $\otimes$ is referred to as the min-plus convolution [Baccelli et al., Cruz et al., Chang, LeBoudec, Thiran].





# Systems in series

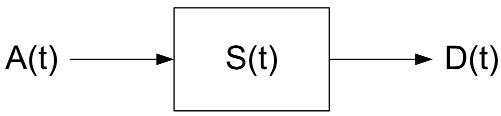
This view is particularly advantageous in case of tandem systems.




Iterating the definition of lower service curve yields

$$D(t) \geq (A \otimes S_1) \otimes S_2(t) = A \otimes (S_1 \otimes S_2)(t)$$

due to the associativity of convolution. Thus, the tandem system is equivalent to a single system with service curve $S(t) = S_1 \otimes S_2(t)$.




I.e. results obtained for single systems extend to tandem systems.

# Bandwidth estimation problem

From the point of view of network calculus bandwidth estimation seeks to find an unknown network service curve $S_u(t)$ from measurements of the traffic arrivals and departures of a network.

The task of bandwidth estimation can be phrased as finding the largest function $S_u(t)$ that satisfies $D(t) \geq A \otimes S_u(t)$ for all $t \geq 0$ and for all pairs of arrivals and departures $(A, D)$ of the network.

$$\begin{aligned} &\text{maximize} \quad S \\ &\text{subject to } D(t) \geq \inf_{\tau \in [0,t]} \{A(\tau) + S(t-\tau)\} \\ &\qquad\qquad \forall t \geq 0, \text{ for all pairs}(A, D) \end{aligned}$$

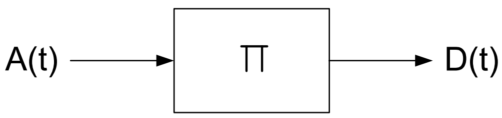
I.e. measurement-based available bandwidth estimation can be viewed as seeking to solve a (difficult) max-min optimization.





# Min-plus linearity

The service curve approach is related to linear systems theory.




Generally, a (linear or nonlinear) system implements a mapping Π

$$D(t) = \Pi(A(t))$$

The system is min-plus linear, if

- $\Pi(c + A_i(t)) = c + D_i(t)$
- $\Pi(\inf\{A_i(t), A_j(t)\}) = \inf\{D_i(t), D_j(t)\}$

It is time-invariant, if

- $\Pi(A_i(t-\tau)) = D_i(t-\tau)$

$\forall t \geq 0$, $\forall \tau \in [0, t]$, any constant $c$, and all pairs $(A_i, D_i)$, $(A_j, D_j)$.

# Impulse response

It can be shown that min-plus linear, time-invariant systems have an exact service curve $S(t)$, i.e. for any pair of arrivals and departures of a system $(A, D)$ and all $t \geq 0$ it holds that

$$D(t) = A \otimes S(t)$$

The service curve $S(t)$ is the impulse response of the system, i.e. given an impulse as arrivals, the departures correspond to the service curve

$$S(t) = \Pi(\delta(t))$$

The impulse function under the min-plus algebra is defined as

$$\delta(t) = \begin{cases} \infty & \text{, for } t > 0 \\ 0 & \text{, for } t \leq 0 \end{cases}$$

The impulse is an infinite burst of arrivals.





# Nonlinearities

Given a min-plus linear, time-invariant system, the impulse function, i.e. probing with $\delta(t)$, reveals the service curve.

In practice the impulse function can be emulated using a finite burst of data that is sent at line speed, i.e. a train of back-to-back packets. This method has been used in early bandwidth estimation tools such as CProbe [Carter, Crovella, PEVA'96].

The approach is, however, intrusive. Large bursts of data cause congestion and interfere with existing traffic. This causes certain systems, e.g. FIFO multiplexer, to become nonlinear.

The challenge is to select probing traffic $A_p(t)$ which permits an inversion of the min-plus convolution, i.e. which permits solving

$$D_p(t) = A_p \otimes S_u(t)$$

for $S_u(t)$ without making the system nonlinear.

# Legendre transform

In classical systems theory similar estimation problems can be solved in the frequency domain, i.e after Fourier transformation.

The Legendre transform

$$\mathcal{L}_f(r) = \sup_{t \geq 0}\{rt - f(t)\}$$

plays a similar role in min-plus systems theory. For convex functions $f(t)$ the Legendre transform is its own inverse

$$\mathcal{L}(\mathcal{L}_f)(t) = f(t)$$

It takes the min-plus convolution to a simple addition

$$\mathcal{L}_{f \otimes g}(r) = \mathcal{L}_f(r) + \mathcal{L}_g(r)$$

which permits the desired inversion as long as $\mathcal{L}_g(r)$ is finite.





# Interpretation of the Legendre transform

Consider a min-plus linear, time-invariant system with exact service curve $S(t)$ and arrivals $A(t) = rt$. The system's backlog bound is

$$\begin{aligned} B_{\max} &= \sup_{t \geq 0}\{A(t) - D(t)\} \\ &= \sup_{t \geq 0}\{rt - \inf_{\tau \in [0,t]}\{r\tau + S(t-\tau)\}\} \\ &= \sup_{t \geq 0}\{\sup_{\tau \in [0,t]}\{r(t-\tau) - S(t-\tau)\}\} \\ &= \sup_{u \geq 0}\{ru - S(u)\} \\ &= \mathcal{L}_S(r) \end{aligned}$$

i.e. the backlog bound is the Legendre transform of the system's service curve.

# Rate scanning method

Consider a min-plus linear, time-invariant system with convex service curve $S_u(t)$.

Send constant rate packet train probes $A_p(t) = rt$ into the system. The maximum backlog for each rate can be derived from measurements of the arrivals $A_p(t)$ and departures $D_p(t)$ as

$$B_{\max}(r) = \sup_{t \geq 0}\{A_p(t) - D_p(t)\}$$

The backlog bound $B_{\max}(r)$ is the Legendre transform of $S_u(t)$, i.e. $\mathcal{L}_{S_u}(r)$. Backwards transformation yields the unknown service curve

$$S_u(t) = \mathcal{L}(B_{\max})(t)$$



Available Bandwidth Estimation Min-plus system interpretation Basics Linear systems Legendre transform Methods

# Pathload

Pathload [Jain, Dovrolis, SIGCOMM'02] is a bandwidth estimation method that uses rate scanning. It performs a binary search until the rate of the packet trains converges to the available bandwidth.

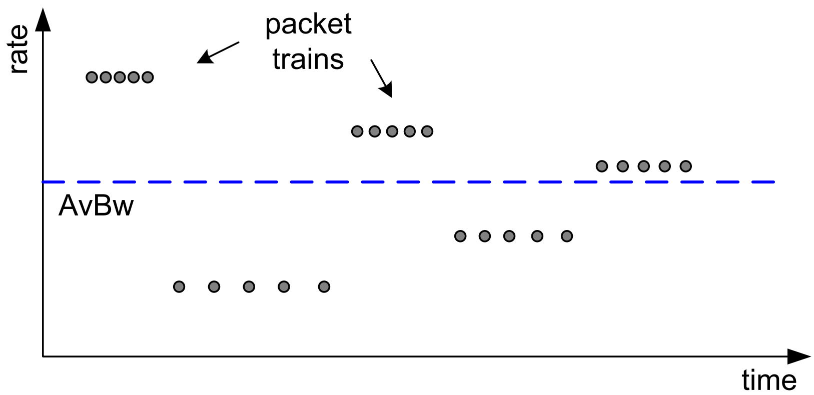


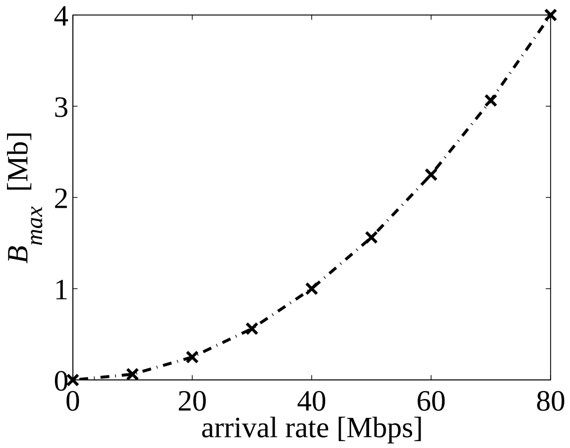

Illustration borrowed from V. Ribeiro

Pathload iterates until a certain resolution is achieved

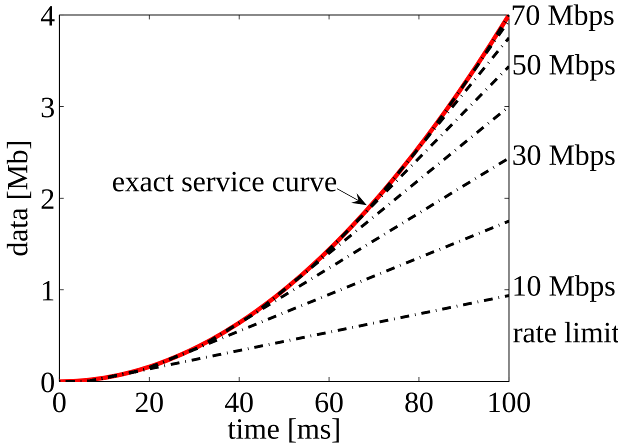
- in case of overload the rate of the next train is reduced
- otherwise the rate of the next train is increased

# Rate scanning example

Backlog measurements from rate scanning and service curve estimates from Legendre transform with different rate limits.




maximum backlog as a function of the rate




service curve estimates with different rate limits





# Rate chirp method

Consider a min-plus linear, time-invariant system with convex service curve $S_u(t)$, i.e. $D(t) = A \otimes S_u(t)$. From the properties of the Legendre transform it follows for all $r \geq 0$ for which $\mathcal{L}_A(r)$ is finite

$$\mathcal{L}_D(r) = \mathcal{L}_A(r) + \mathcal{L}_S(r)$$
$$\Leftrightarrow \mathcal{L}_S(r) = \mathcal{L}_D(r) - \mathcal{L}_A(r)$$

Packet chirps [Ribeiro et al., PAM'03] are packet streams with exponentially increasing rate. This ensures that $\mathcal{L}_A(r)$ stays finite.

Send a packet chirp probe $A_p(t)$ into the system. The Legendre transform of the service curve can be derived from measurements of the arrivals $A_p(t)$ and departures $D_p(t)$ as

$$\mathcal{L}_{S_u}(r) = \mathcal{L}_{D_p}(r) - \mathcal{L}_{A_p}(r)$$

Backwards transformation yields the unknown service curve

$$S_u(t) = \mathcal{L}(\mathcal{L}_{D_p} - \mathcal{L}_{A_p})(t)$$

# Pathchirp

Pathchirp [Ribeiro et al., PAM'03] is a bandwidth estimation method that uses rate chirps. It seeks to detect the rate at which the chirp creates overload. This rate is used as an estimate of the available bandwidth.

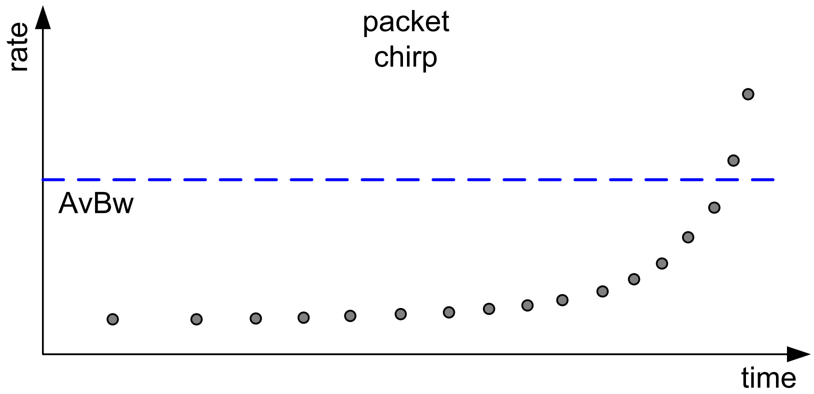



Illustration borrowed from V. Ribeiro

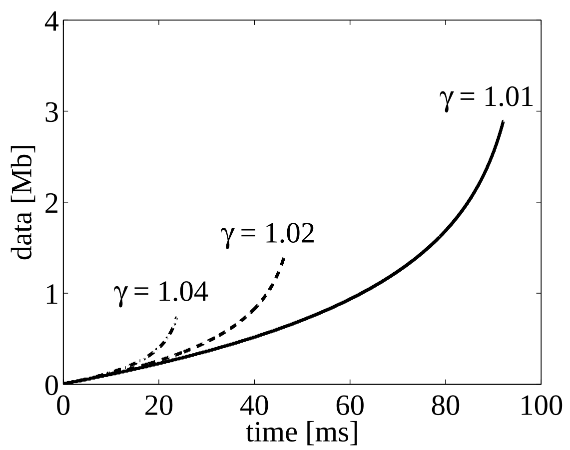
This way an iterative rate scan can be replaced by a single packet chirp that scans over all rates until overload is detected.



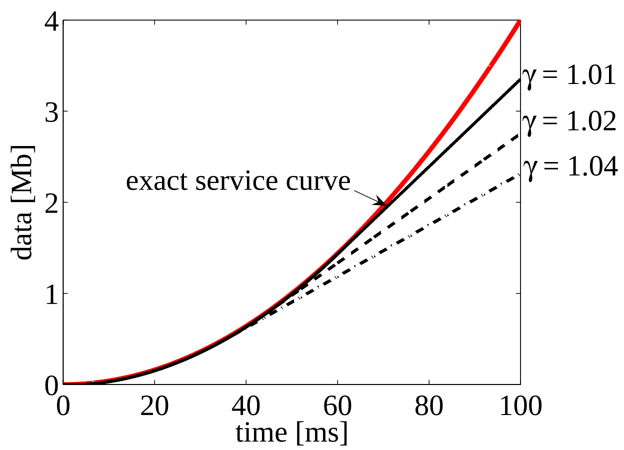


# Rate chirps example

Service curve estimates from rate chirps with different spread factor $\gamma$. The spread factor is the speed with which the chirp grows from its minimum to its maximum rate.



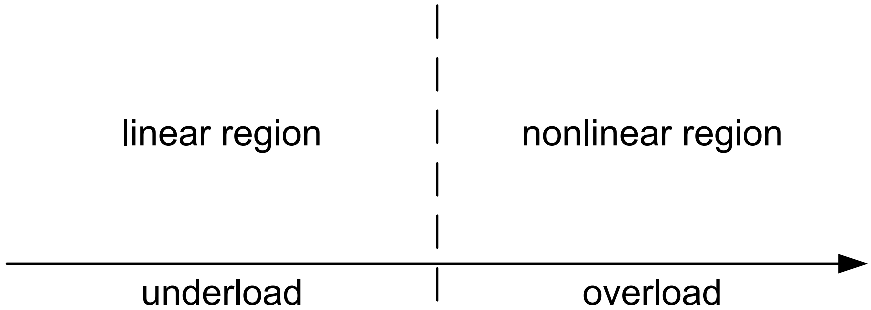

rate chirps with different spread factor $\gamma$




service curve estimates with different rate chirps

# Nonlinearities

The developed methods assume min-plus linearity. There exist, however, systems, such as FIFO multiplexer, which are linear at low load but cross into a nonlinear region once overload occurs.




This permits using the rate scanning and rate chirp methods, if the rate scan respectively rate chirp is stopped once nonlinearity or overload is detected.





# Conclusion

Reference [Liebeherr, Fidler, Valaee: A Min-Plus System Interpretation of Bandwidth Estimation, INFOCOM'07]:

- Expresses available bandwidth estimation in the framework of network calculus
- Shows that rate scanning and rate chirp methods can be derived in min-plus systems theory
- Explains the difficulties and the distinctive treatment of FIFO systems in terms of nonlinearities
- Provides a method for passive measurements (not shown here)

Future work: Measurement-based path selection in overlay networks [Jain, Dovrolis, Networking'07] using service curve-based routing [Recker, Telecommunications Systems'03].