Conservative Cascades: an Invariant of Internet Traffic

Steve **UHLIG**

E-mail: suh@info.ucl.ac.be
URL: http://www.info.ucl.ac.be/~suh/

Computer Science and Engineering Dept.
Université catholique de Louvain, Belgium

Motivation

- More than 95 % of all Internet traffic is carried by the TCP protocol

⇒ Understanding impact of TCP on Internet traffic is important

- Conservative cascade model proposed since '98 is known to be sound but...

Question : How much is it sound ?

The scaling framework

$$E|d(j,k)|^q \propto \exp(f(q)\, g(2^j))$$

- **separability** of the moments (**q**) and the timescales (**j**)

- higher-order moments **q** emphasize on larger irregularities

- larger timescales **j** focus on smoothed versions of the process (zooming out)

⇒ **framework to study irregularities and timescales independently**

Scaling types

Self-similarity :

$$E|d(j,k)|^q \propto \exp(q\, H \ln(2^j))$$

Multiscaling :

$$E|d(j,k)|^q \propto \exp(H(q) \ln(2^j))$$

Infinitely divisible cascades :

$$E|d(j,k)|^q \propto \exp(H(q)\, n(2^j))$$

Probabilistic interlude

- Cascade : multiplicative process that breaks a process into smaller and smaller fragments according to some (deterministic or random) rule.

- Name "Infinitely divisible cascades" comes from "infinitely divisible distributions" (Feller vol. 2) :

F is infinitely divisible if for every n there exists a distribution F_n such that $F = F_n^{n*}$ where "*" denotes the convolution operator.

or equivalently

F is infinitely divisible iff for each n it can be represented as the distribution of the sum $S_n = X_{1,n} + ... + X_{n,n}$ of n independent random variables with a common distribution F_n.

Conservative cascades

<u>2 objectives :</u>

- preservation of total mass of the process

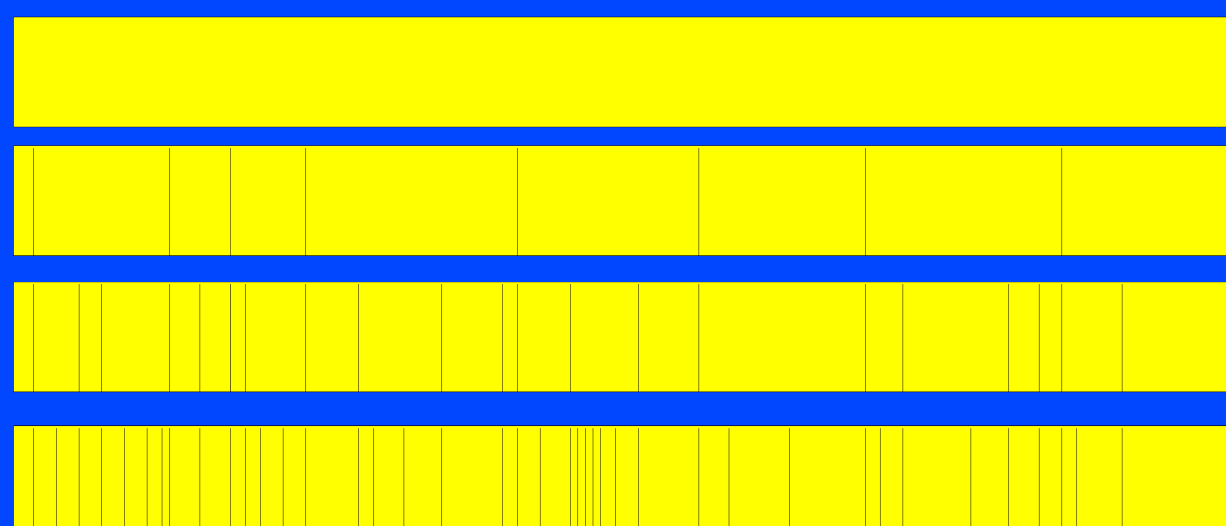
- randomness in the splitting of the mass of the process

<u>Example :</u>

Step 1 :

Step 2 :

Step 3 :



Cascade parameters

H(q) : describes cascade generator (1 step of the breaking of the data)

n(2^j) : how many times the generator has been applied at timescale **j** (how many times the convolution operator has been applied)

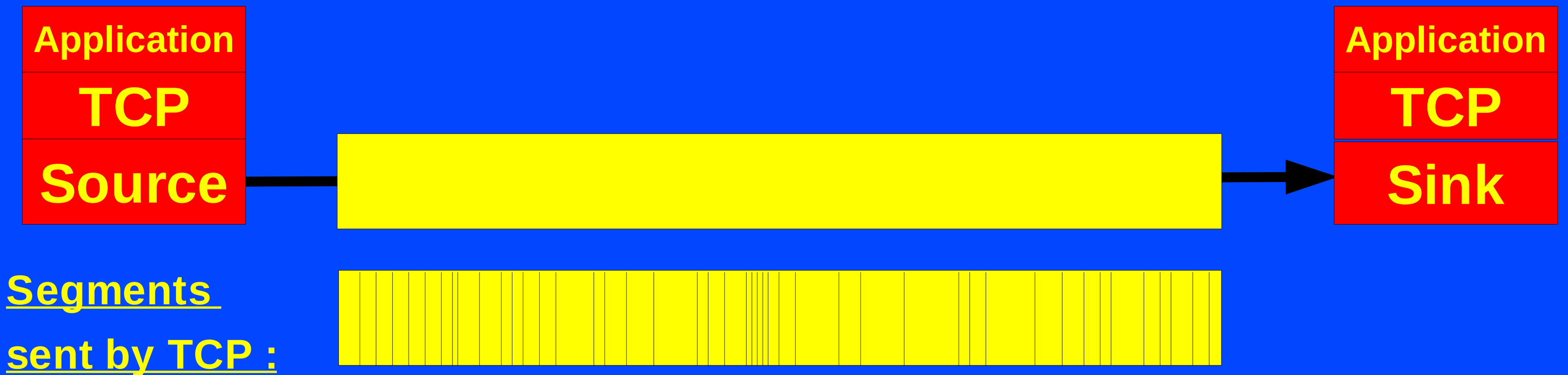
⚠ **H()** and **n()** "*summarize*" the numerical behavior of the cascade

Conservative cascades and TCP traffic

TCP :

- breaks the total mass of the data to be sent into segments

- application sending data, TCP state machine and random network conditions drive how data segments are broken and when they are sent over time



Wavelet analysis

- process increments **d(j,k)** replaced by wavelet coefficients $\delta(j,k)$

- <u>Wavelet-based partition function :</u>

$$\mathbf{S(q,j)} = \sum_k |2^{-j/2}\delta(j,k)|^q \qquad (\propto E|d(j,k)|^q)$$

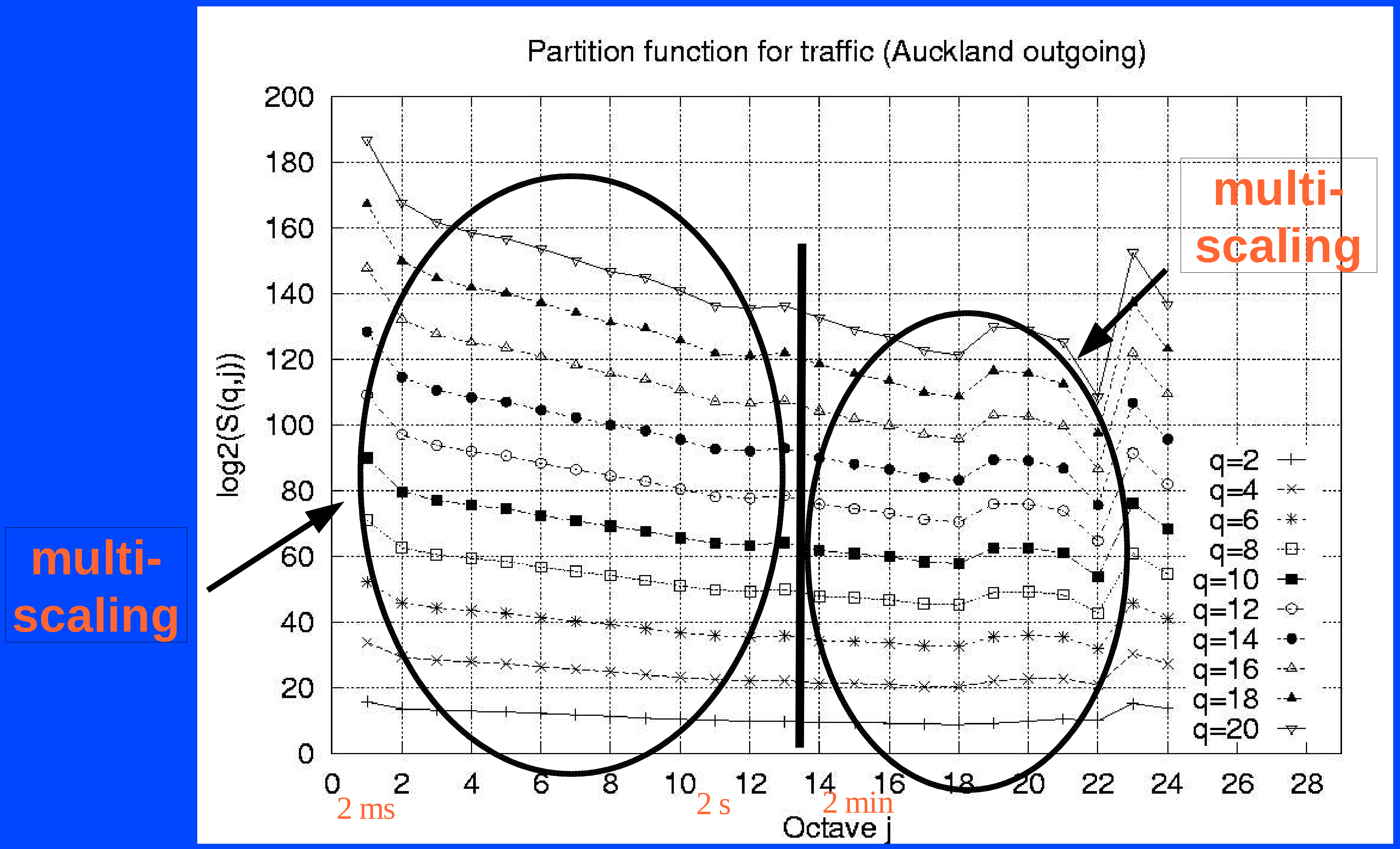
- <u>Estimation of **H(q)** and **n(2^j)** via regression :</u>

$$\ln(S(q,j)) \cong -H(q)\, n(2^j) + K_q$$

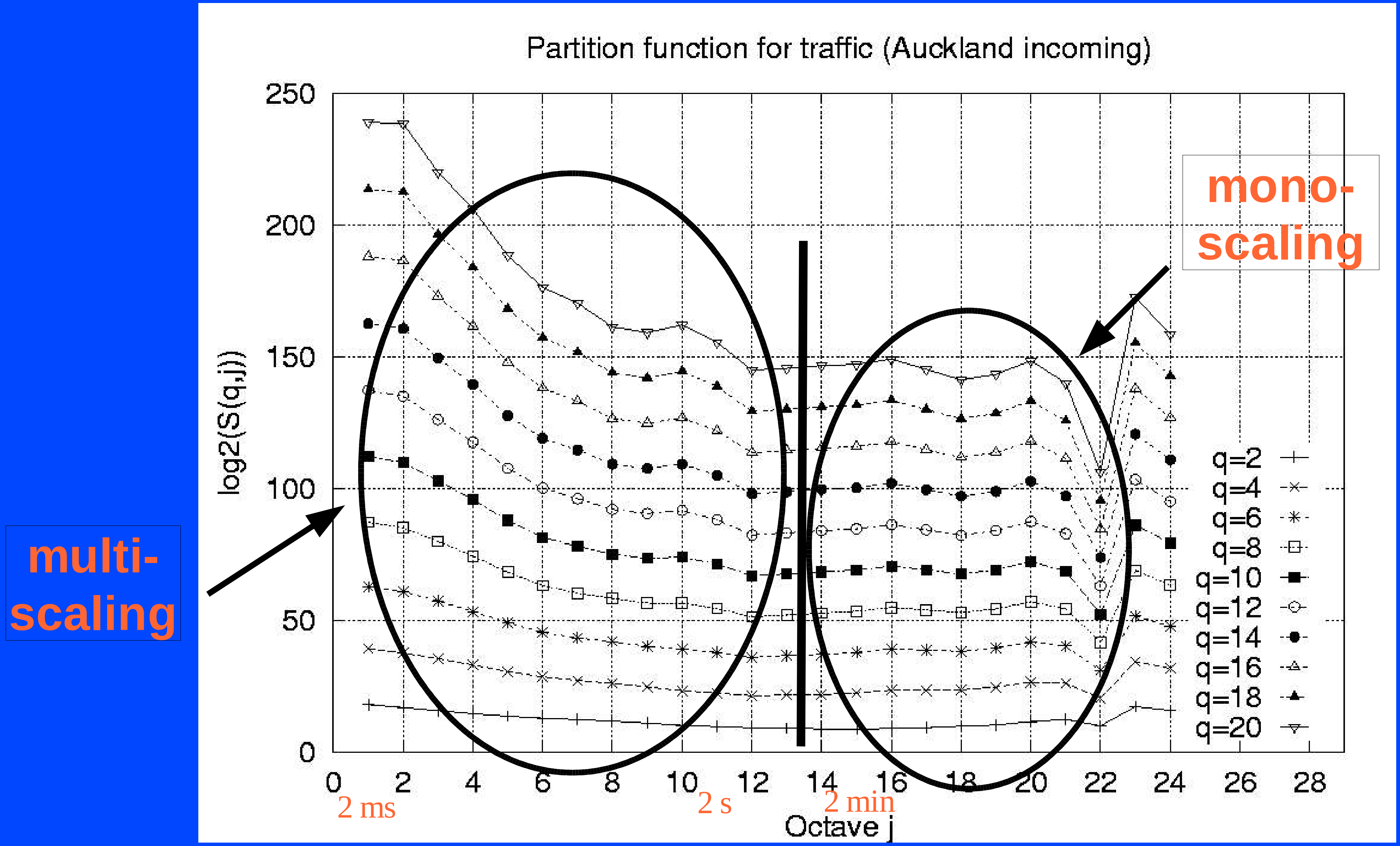
Internet traffic analysis

- **15** hours of IP packets from Auckland
- both traffic sent to the Internet and received from the Internet
- **1,629,069** incoming TCP flows, **1,613,976** outgoing
- more than **13** GBytes of incoming traffic, **10** GBytes outgoing
- 1 μs precision, 1 ms time granularity used.

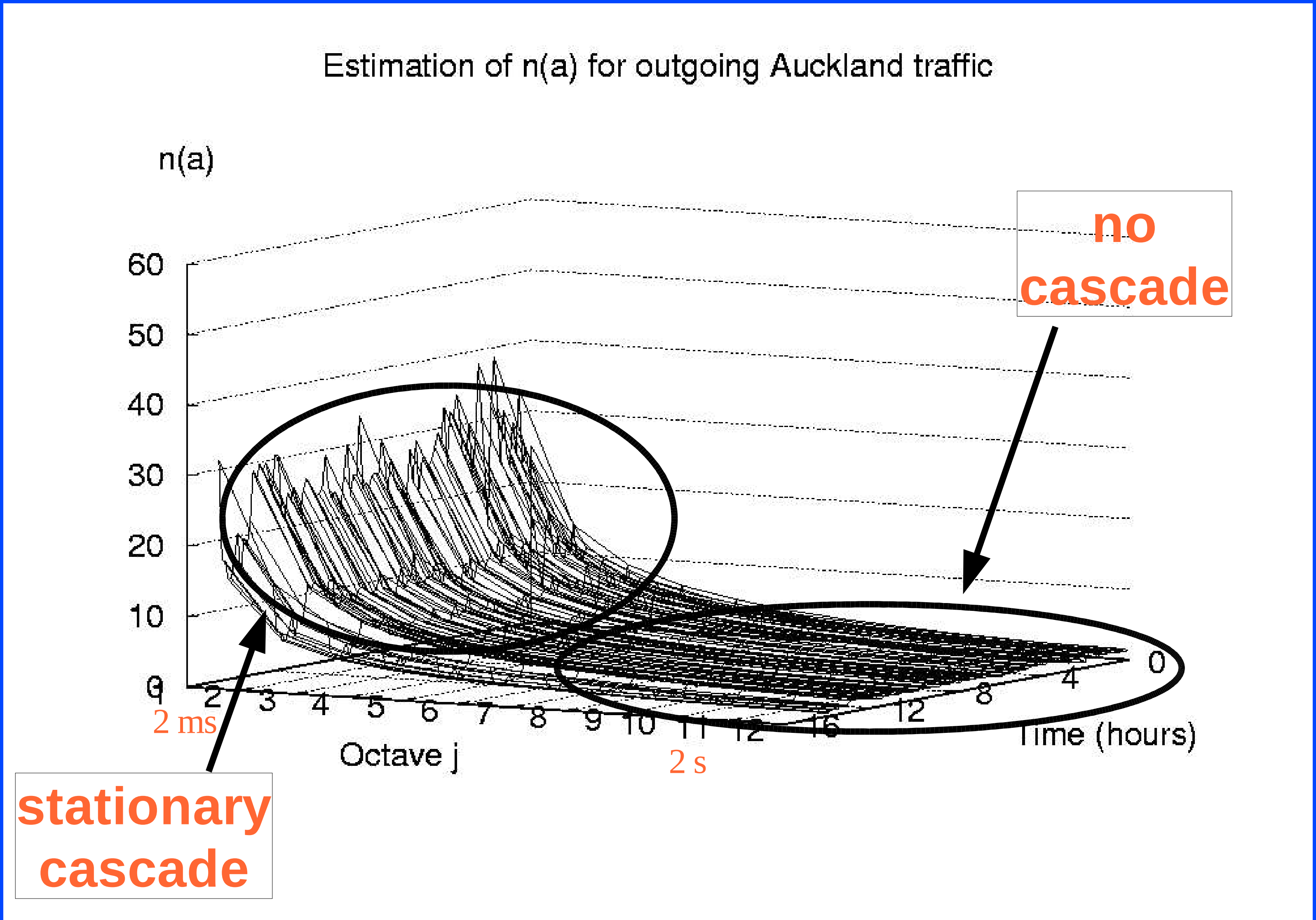
Multiscaling analysis (out)



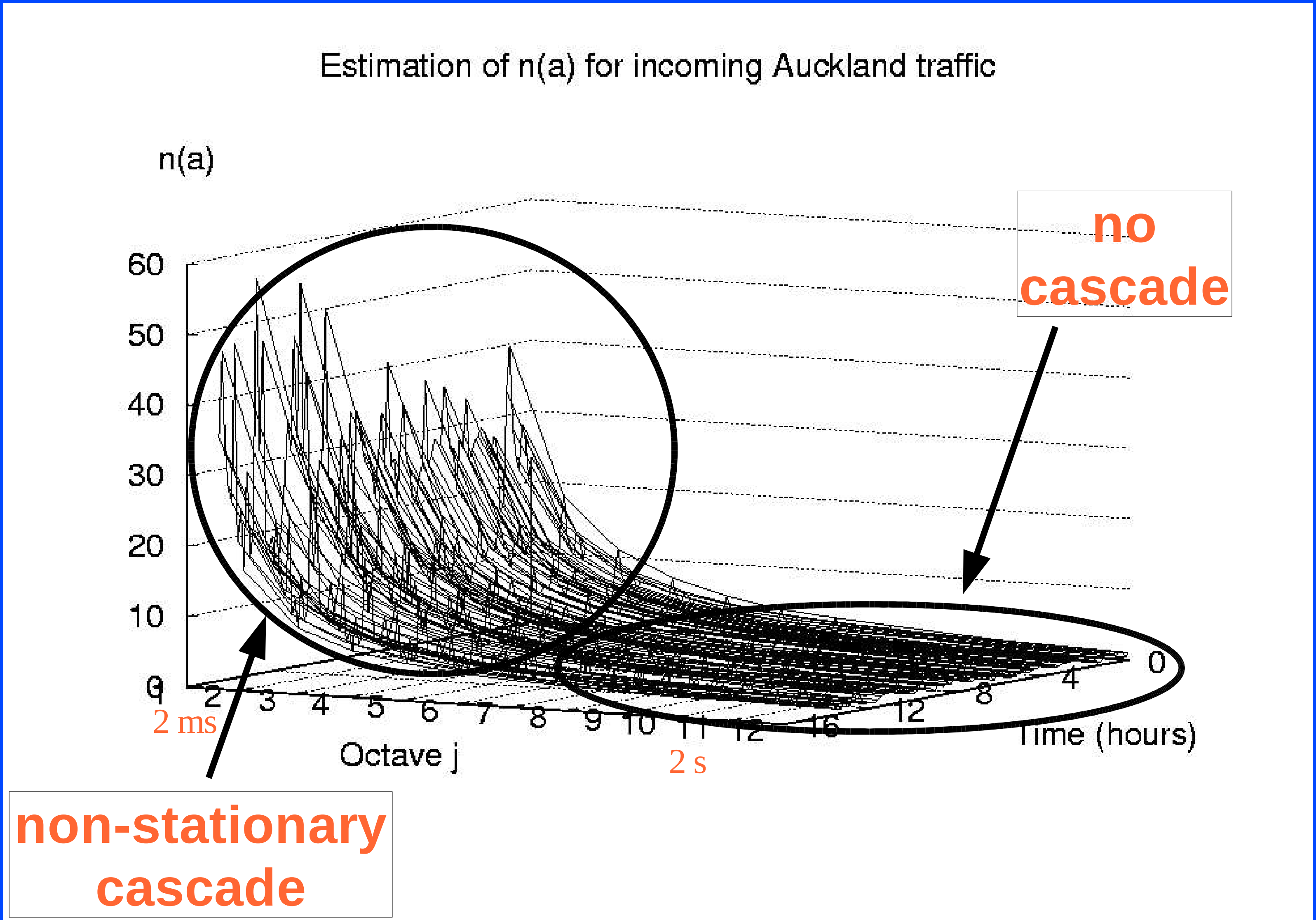
Multiscaling analysis (in)



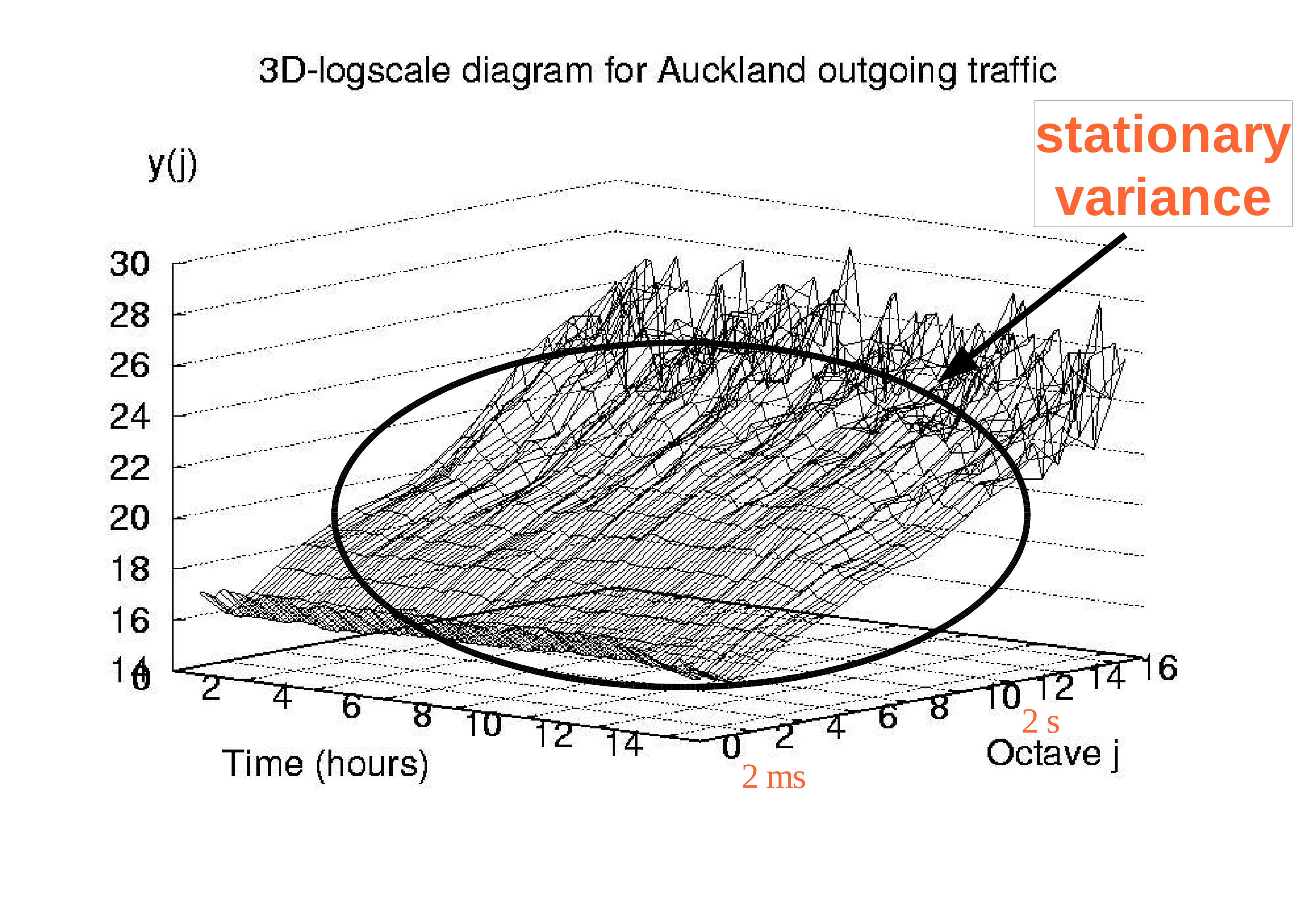
Cascade analysis (out)



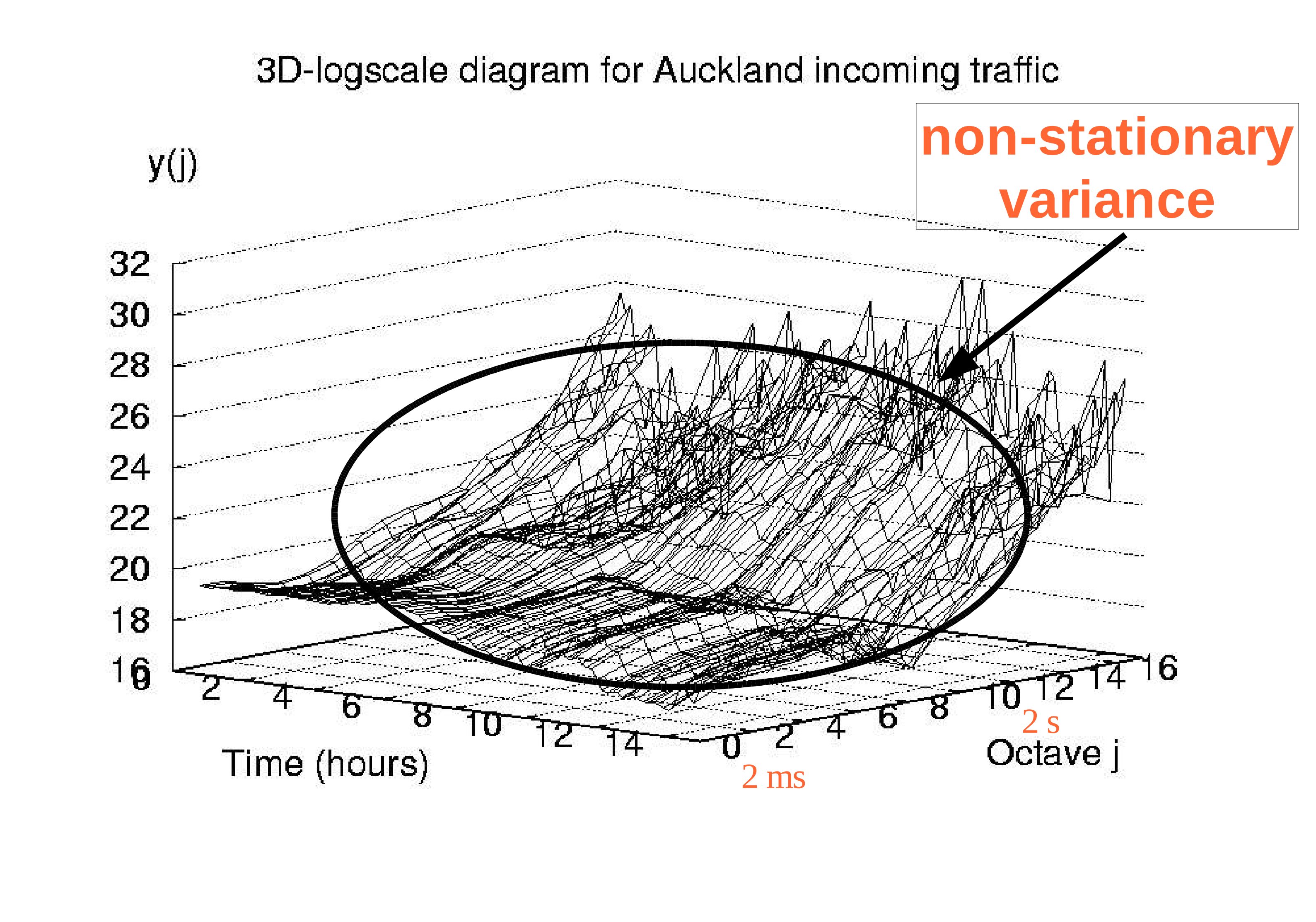
Cascade analysis (in)



Variance analysis (out)



Variance analysis (in)



Conclusions

- Cascade model captures well invariance in TCP behavior
- Cascade model does not fully capture traffic dynamics, only TCP traffic segmentation
- Cascade parameters do not show effect of cross-traffic, 2nd order properties do