

# **Understanding Application Quality's Interaction with Resource Management**

Or  
Achieving Predictable  
Application Quality through  
Resource Management



# Issue

## ***Current measurement and reporting approaches are not sufficient to achieve quality management***

- Operate over too large a timescale
- Increasing reporting/analysis rate is not a scalable answer
  - VoIP is a 50Hz quality problem, SLA's operate at somewhere between 0.003Hz to 0.00001Hz

*Nyquist and Shannon had something to say about this!*

# Problem Space:

## What are the constraints?

- Centralized control mechanisms are not going to be able operate at much higher than 0.01Hz to 0.001Hz
- At frequencies higher than this distributed control is all you going to ever have
- But on what basis are you going to perform that control?
  - How to evaluate the effect of the local, distributed decisions on the global quality

# What should the common model be?

- Likely effect on the user
  - Just need to define ‘effect’ and ‘user’!
- These concepts can be more formally defined e.g.
  - Effect of quality on VoIP
    - MOS/PQOS
  - Effect on Long lived (TCP) transfer
    - Time to Finish Transfer
  - Effect on Web RPC
    - Time to return information to user
- Permits the exploration of which real-time quality effects the user experience of applications

# What else can this approach help with?

- Resource contention issues
  - When resource begins to run out, how should it be assigned to achieve ‘business goals’
- Permits (a-priori) evaluation of normal ranges for measurements, i.e.
  - What is ‘significant’
- The above are needed to permit the successful overbooking of a resource.
  - Permits the evaluation of risks of such overbooking