Capacity Planning for Internet Service Networks

Geoff Huston

NTW Track4

Issues

TCP/IP Protocol Behavior

Issues

Usage Profile

Capacity Guidelines

Growth Levels

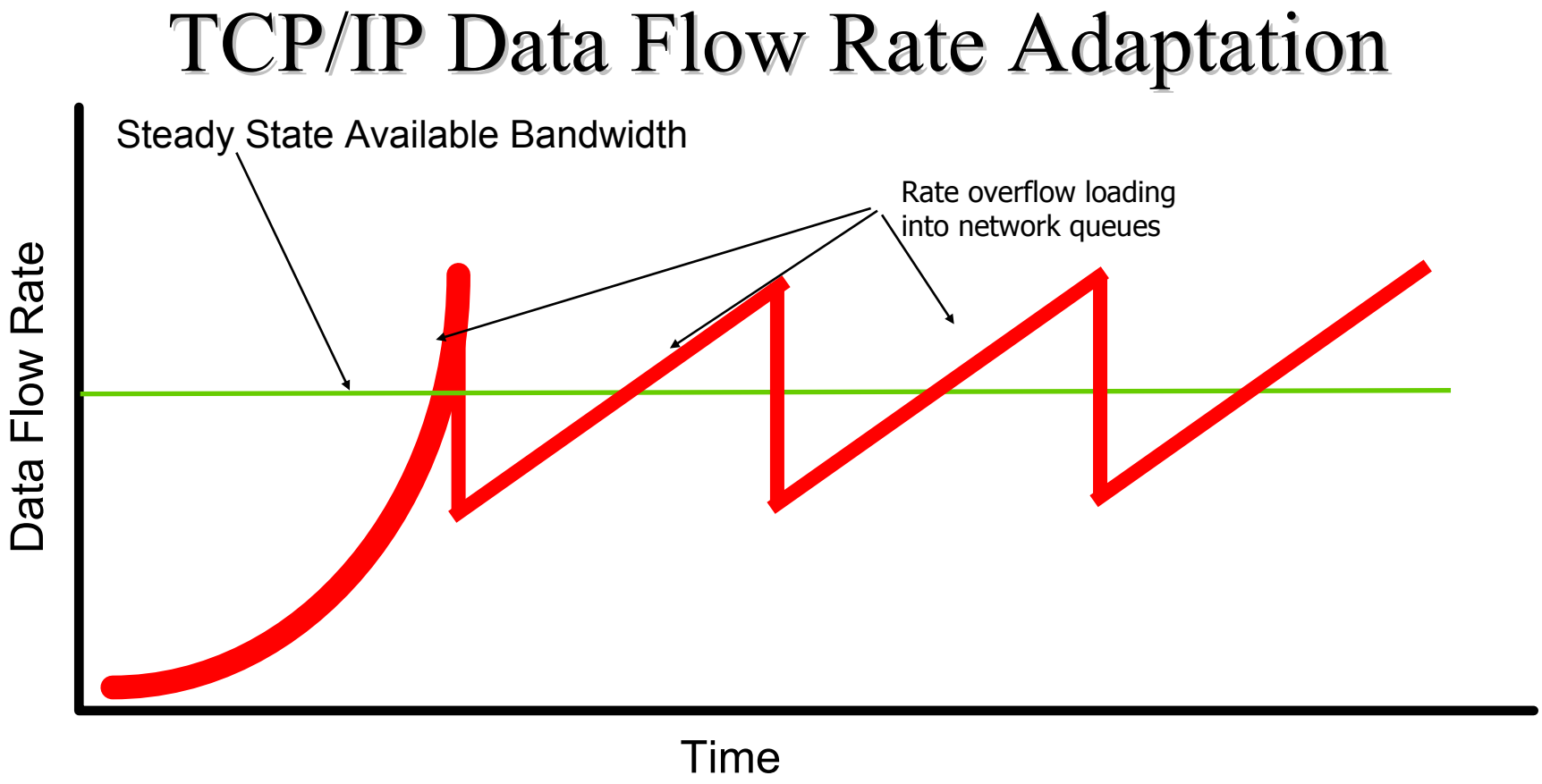
Planning Issues

- Understand the domain of operation
  - technical issues
  - market issues
  - competitive issues
  - regulatory issues

TCP/IP Protocol Issues

- TCP/IP is NOT a flow damped protocol
  - end to end flow management
  - sliding window protocol
  - adaptive flow rate designed to probe and use max available end to end bandwidth
  - only limited by end system buffering size
    - bandwidth x delay
  - system buffers are getting larger as OS vendors come to understand the problem

TCP/IP Protocol Issues



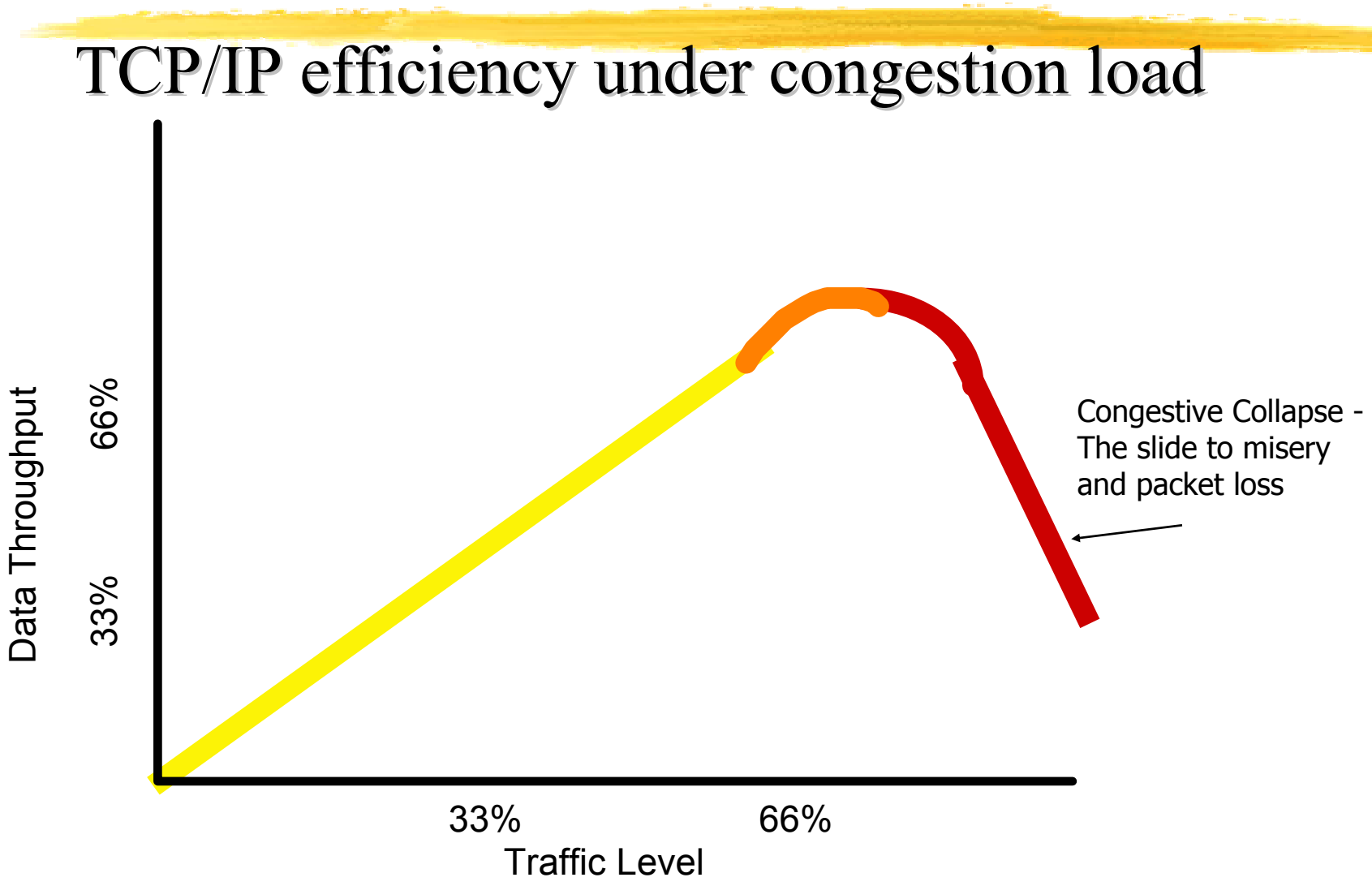
TCP/IP Protocol Issues

- No network-based flow control mechanism
- Network-based packet loss signals end systems to collapse transmission window size
- Varying window size allows adaptive flow metrics to adapt to changing maximum available capacity
- Sustained insufficient capacity leads to congestion induced collapse of data throughput

TCP/IP Protocol Issues

- Many simultaneous TCP sessions interact with non-predictive non-uniform load (ftp://thumper.bellcore.com/pub/dvw/sigcom93.ps.Z)
- Peaks start to synchronize with each other
- Buffering evens out individual flows, but buffers themselves behave with fluctuating load
- Buffering adds latency

TCP/IP Protocol Issues



TCP/IP Protocol Issues

- TCP vs UDP
- UDP-based applications
  - Internet Phone, Video, Workgroup
- UDP Issues
  - no flow control mechanism
  - sustained use forces precedence over TCP flows
  - increasing use of flow bandwidth negotiated protocols for these applications (RSVP)

TCP/IP Protocol Issues

- Damping network capacity is not a demand management tool
- Network capacity must be available to meet peak demand levels without congestion loss

Usage Profile

- Two major Internet use profiles:
  - Business use profile
    - peak at 1500 - 1600
    - plateau 1000 - 1730
  - Residential dial profile
    - peak at 2030 - 2330
    - plateau 1900 - 2400

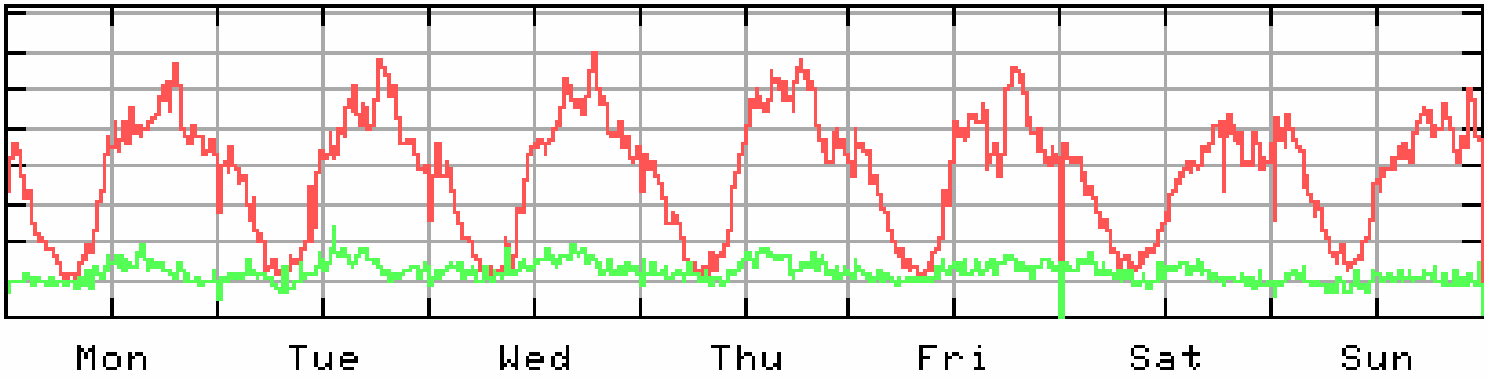
Usage Profile

- Distance profiles
  - 12% Local
  - 18% Domestic Trunk
  - 70% International
- Traffic mix due to:
  - Distance invisible applications without user control
  - Distance independent user tariff

Capacity Guidelines

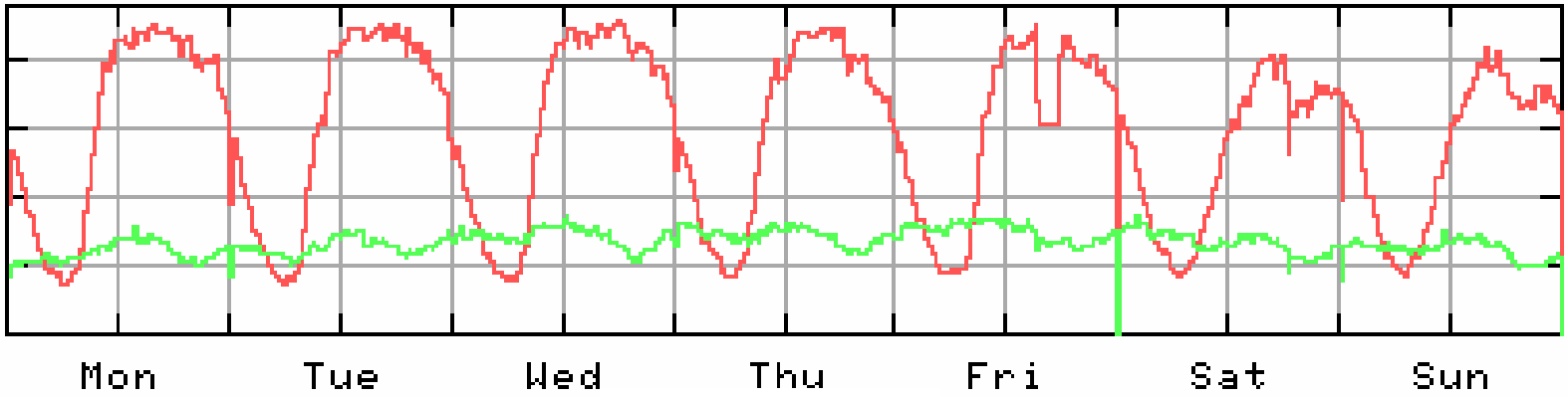
- Link Utilisation
  - Average weekly traffic level set to 50% of available bandwidth.
- Core network capacity should be dimensioned according to aggregate access bandwidth

Link Usage Profile - optimal



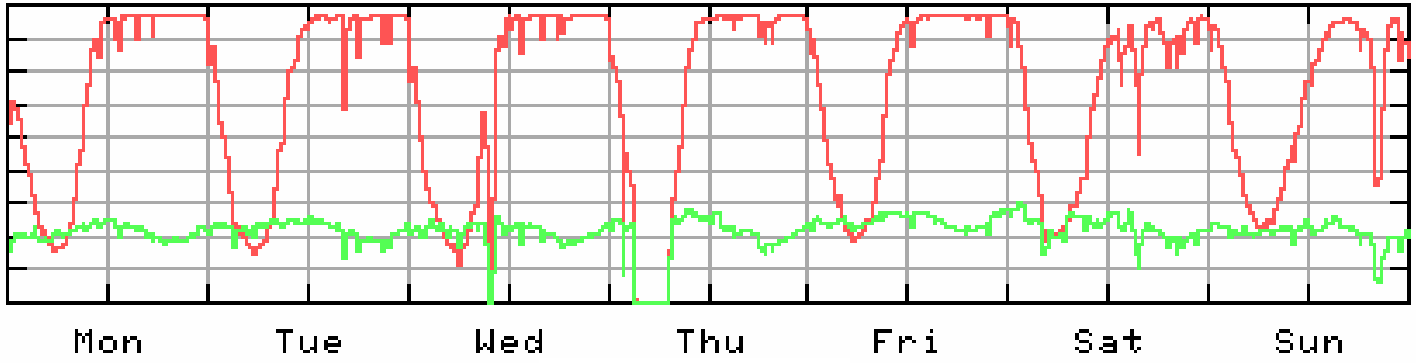
- peak loading less than 10% time
- greater than 50% loading for 50% time
- traffic bursting visible

Link Usage Profile - overloaded



- 90% peak loading for 45% time
- 60% peak loading for 60% time
- no burst profile at peak loads
- imbalanced traffic (import based)

Link Usage Profile - saturated



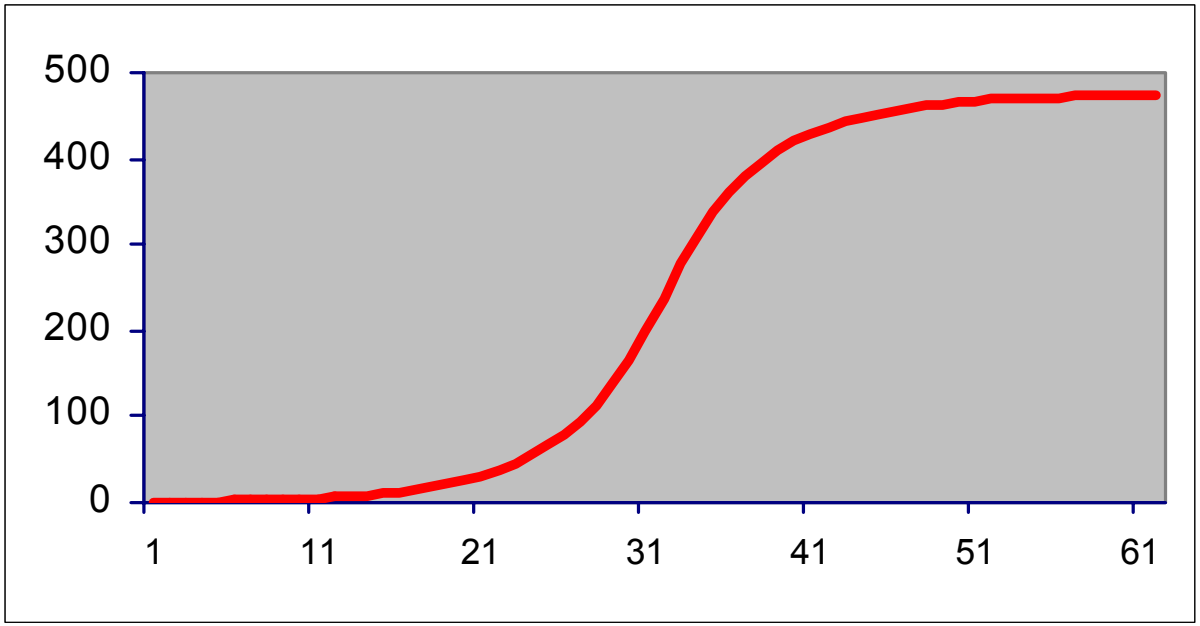
- visible plateau traffic load signature
- small load increases cause widening plateau

Overall Growth Levels

- Two growth pressures:
  - serviced population
  - the changing Internet service model
    - more network-capable applications
    - using more bandwidth

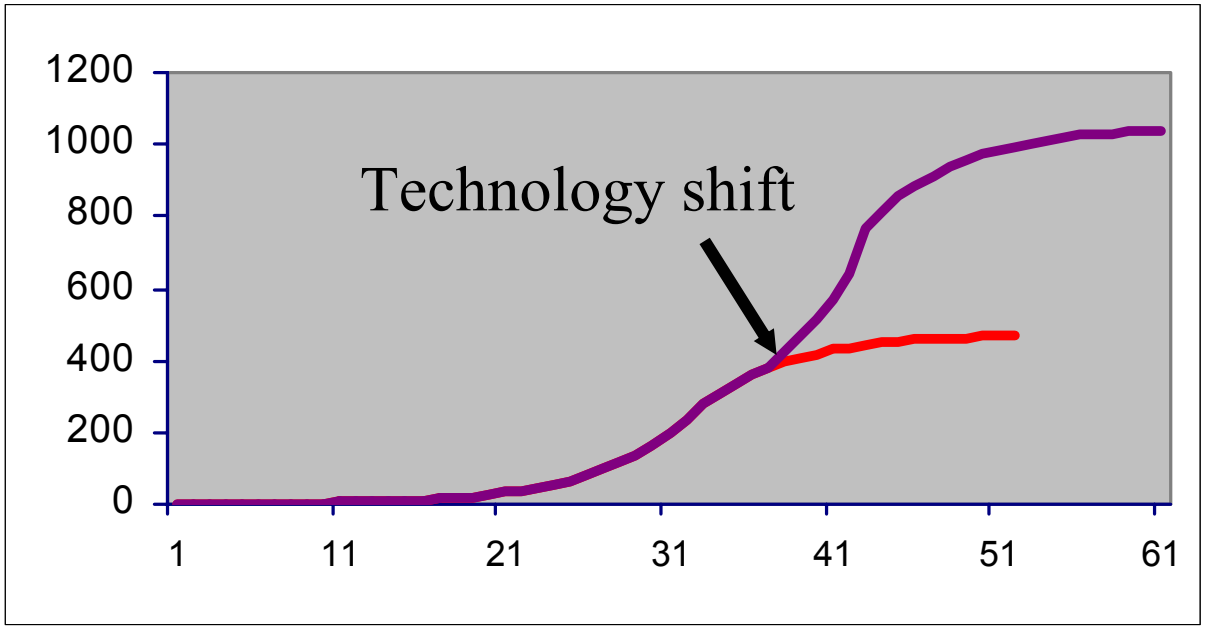
Overall Growth Levels

- For a constant service model the growth curve will exhibit demand saturation



Overall Growth Levels

- For a changing service model the saturation point will move
- More intense network use by increasingly sophisticated applications



How to plan

- Generate a market demand model
  - forecast the number of services in operation
    - existing services
    - growth rate
    - market capture level (competitive position)
  - forecast the average demand per service
    - dial access, leased line
    - web, ftp, usenet
    - caching trends
    - new Internet services

How to plan

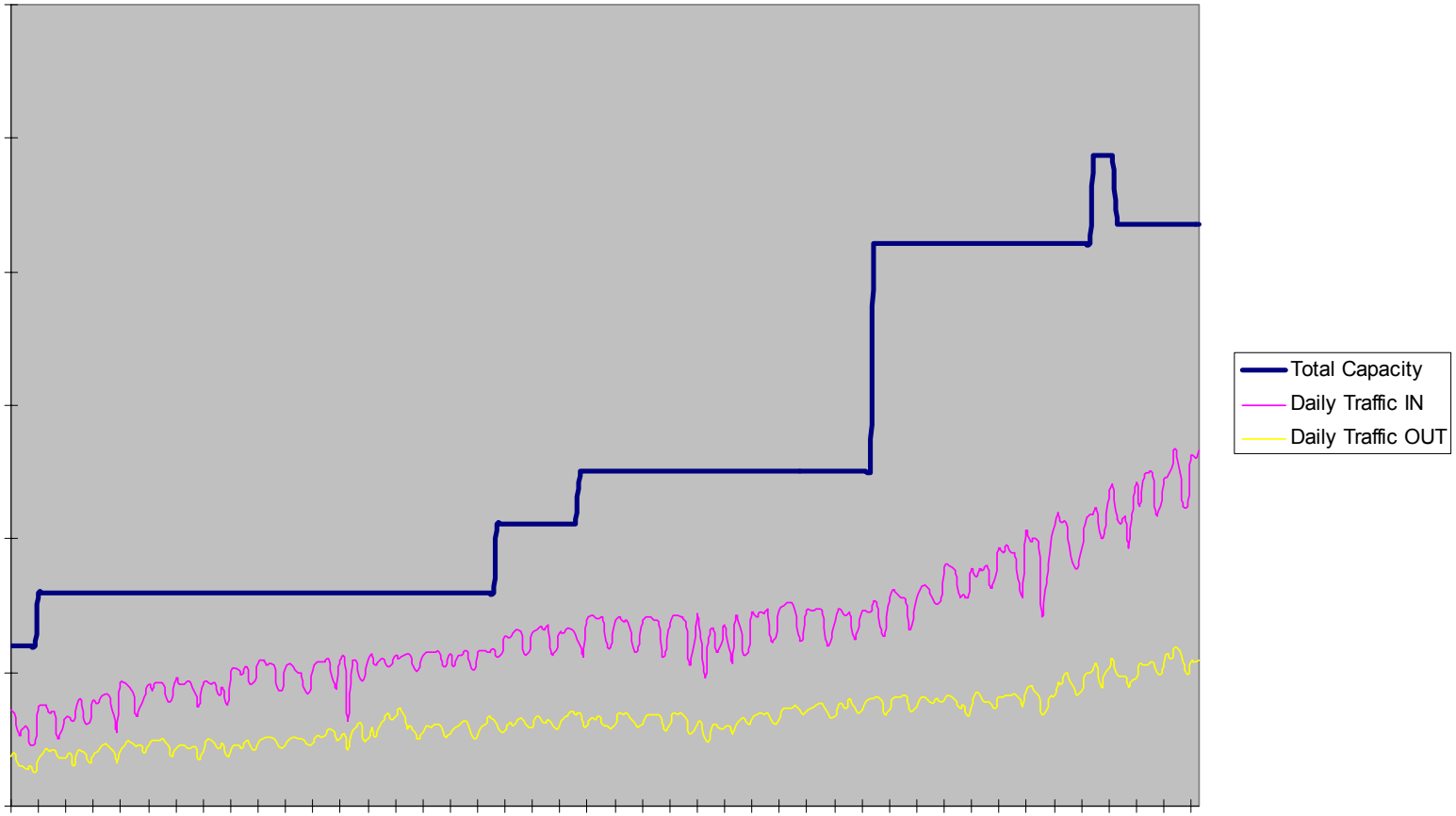
- demand models are typically very uncertain indicators
  - high level of uncertainty of externalities
  - highly dynamic competitive position
  - poorly understood (and changing) service demand model

How to plan

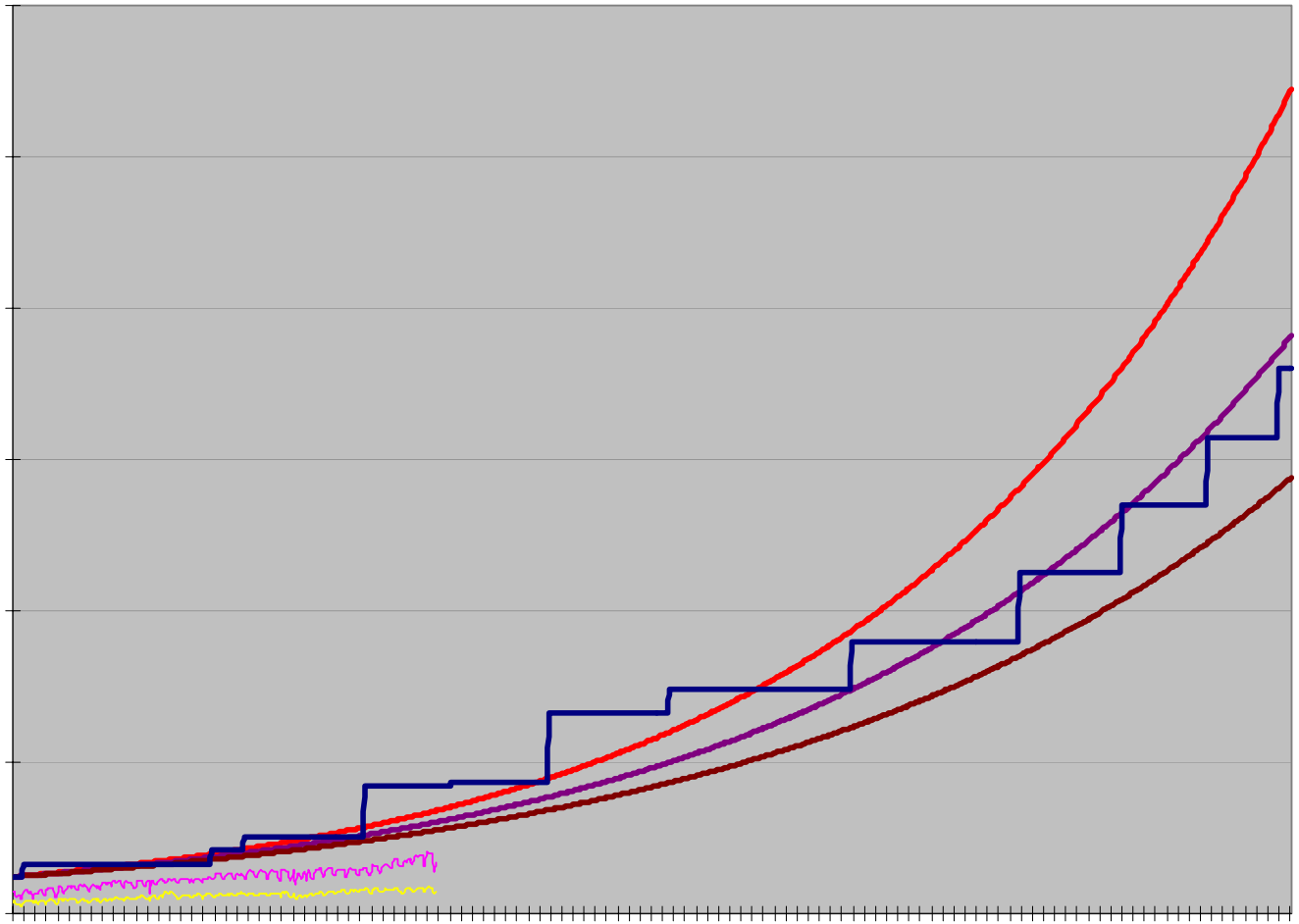
- Forward extrapolation
  - assume existing traffic follows a general growth model
  - forward extrapolat the growth model
- Good for short term planning (12 months)
- Cannot factor
  - latent demand
  - market price sensivity

Trend forecasting

- historical usage vs capacity data



Growth Trends



Planning

- undertake demand and trend forecast models
- constantly review the model against generated usage data
- recognise that the larger the capacity you need the longer the lead time to purchase it
- recognise that the bigger the purchase the greater the requirement for capital