

# Grid Network Performance Monitoring for e-Science

Mark Leese - Daresbury Laboratory

Wednesday 3/4/2003

[m.j.leese@dl.ac.uk](mailto:m.j.leese@dl.ac.uk)

<http://gridmon.dl.ac.uk/~mjl>



# Contents

- Background & purpose of the work
- Monitoring: why, what, how
- Progress & examples
- The future
- Conclusion
- Questions?



## And Introducing the Grid...

- “global computing resources available for transient heavy users” Ian Marshall, BT
- Essentially a means for sharing resources (processing power, databases, radio telescopes...) via a distributed network.
- Grid enabled projects expected to be high bandwidth users (e.g. splitting of large datasets from HEP experiments for processing at multiple sites).



# Purpose of the work

---

“...design and deploy an infrastructure for network performance monitoring within the UK e-Science community.”

## Key aspects:

- Publish results → Grid middleware
  - Visualisation for humans
  - End-to-end
  - ability of TCP wrt high b/w networks
-



# Monitoring: Why?

- Two main drivers:
  - fault detection
  - determine expected performance
- Why is the Grid special?
  - it's different
  - addition of non-human consumption:
    - middleware
    - applications } Adaptive!

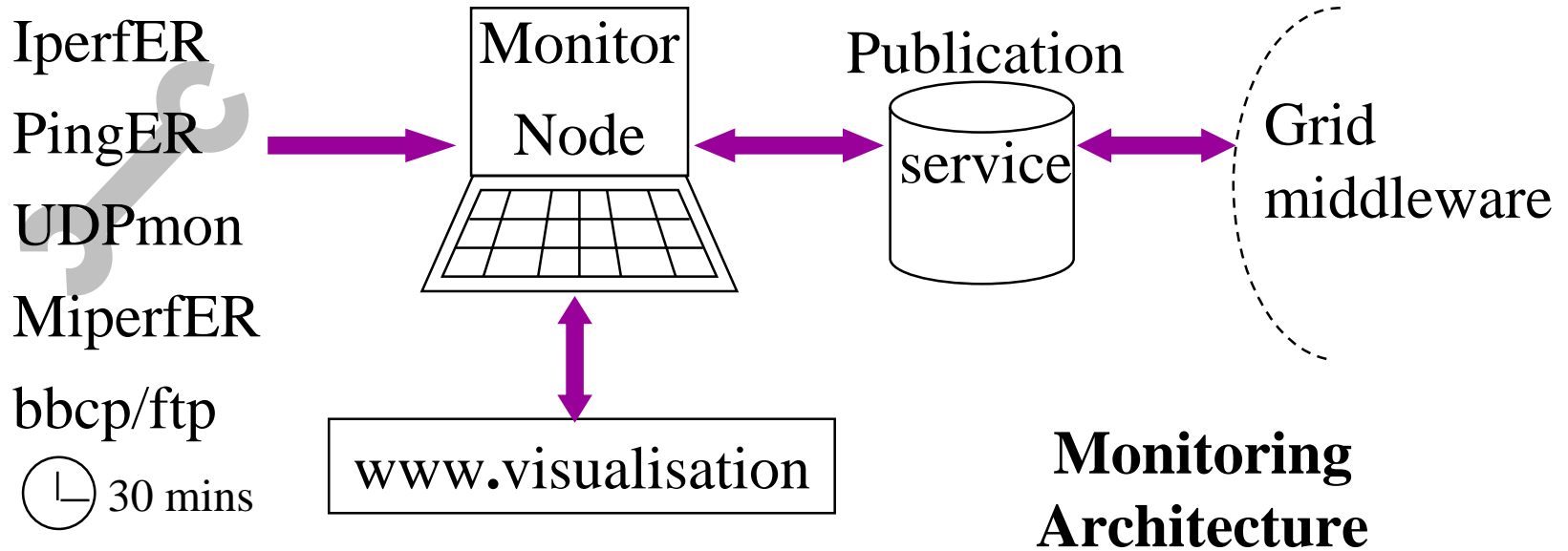


# Monitoring: What?

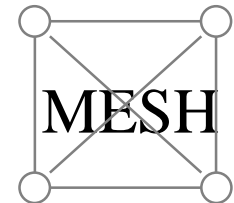
Metric	Relevance
Packet loss	Lost packets re-Tx $\Rightarrow$ traffic jams
RTT	TCP = send-acknowledge protocol delayed acknowledge = delayed traffic
Connectivity	Can you connect?
TCP/ UDP Thru'put	Network view? Application view?
Jitter	Variation in delay - UDP & multicast (UDP) only



# Monitoring: How(1)?



Tools installed on **dedicated** & **similar** node at each centre





# Monitoring: How(2)?

<b>Metric</b>	<b>Tool</b>	<b>Origin</b>
RTT packet loss connectivity	PingER	Ping SLAC et al
TCP thru'put	IperfER	NCSA's iperf SLAC and UCL
	bbcp	SLAC
	bbftp	Tool = IN2P3 Monitoring = SLAC
UDP thru'put	UDPmon	HEP@man.ac.uk, EDG
Multicast thru'put packet loss jitter	MiperfER	IperfER Manchester Computing

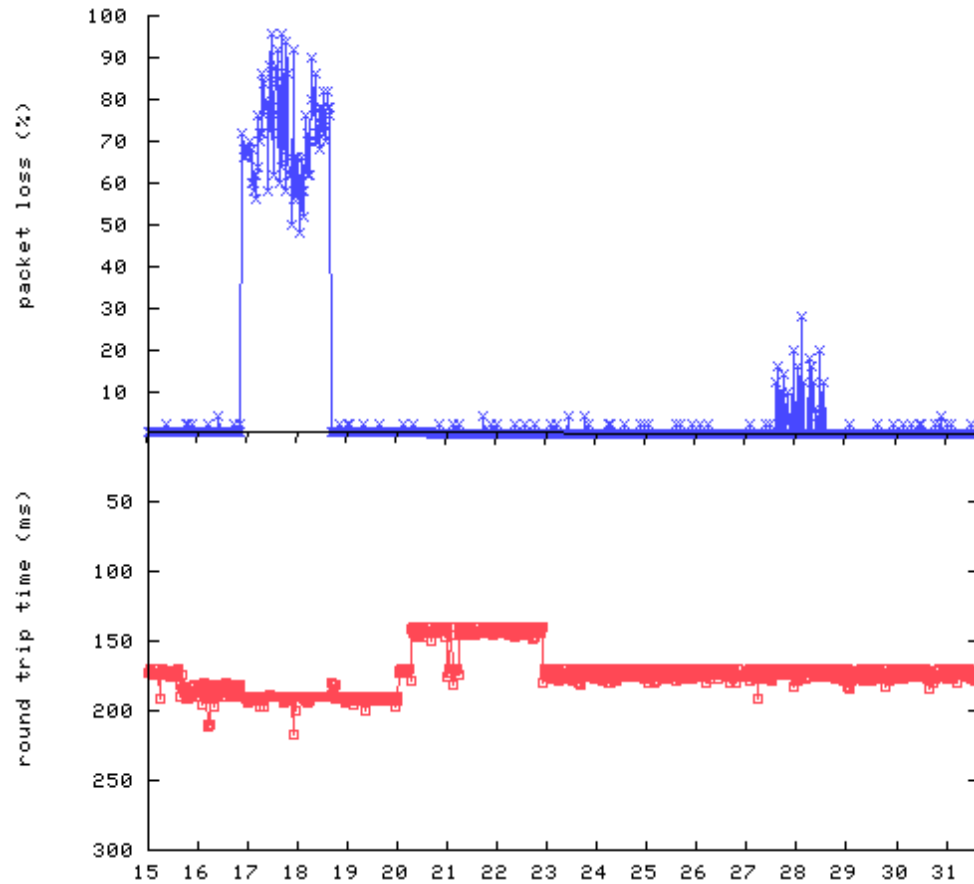


# Progress

- Initial toolkit (IperfER, PingER, UDPmon) at 10 of 12 e-Science centres. Other useful sites being added.
- MiperfER on beta trial at Cambridge, Cardiff, MCC and Newcastle.
- Active map also on trial.

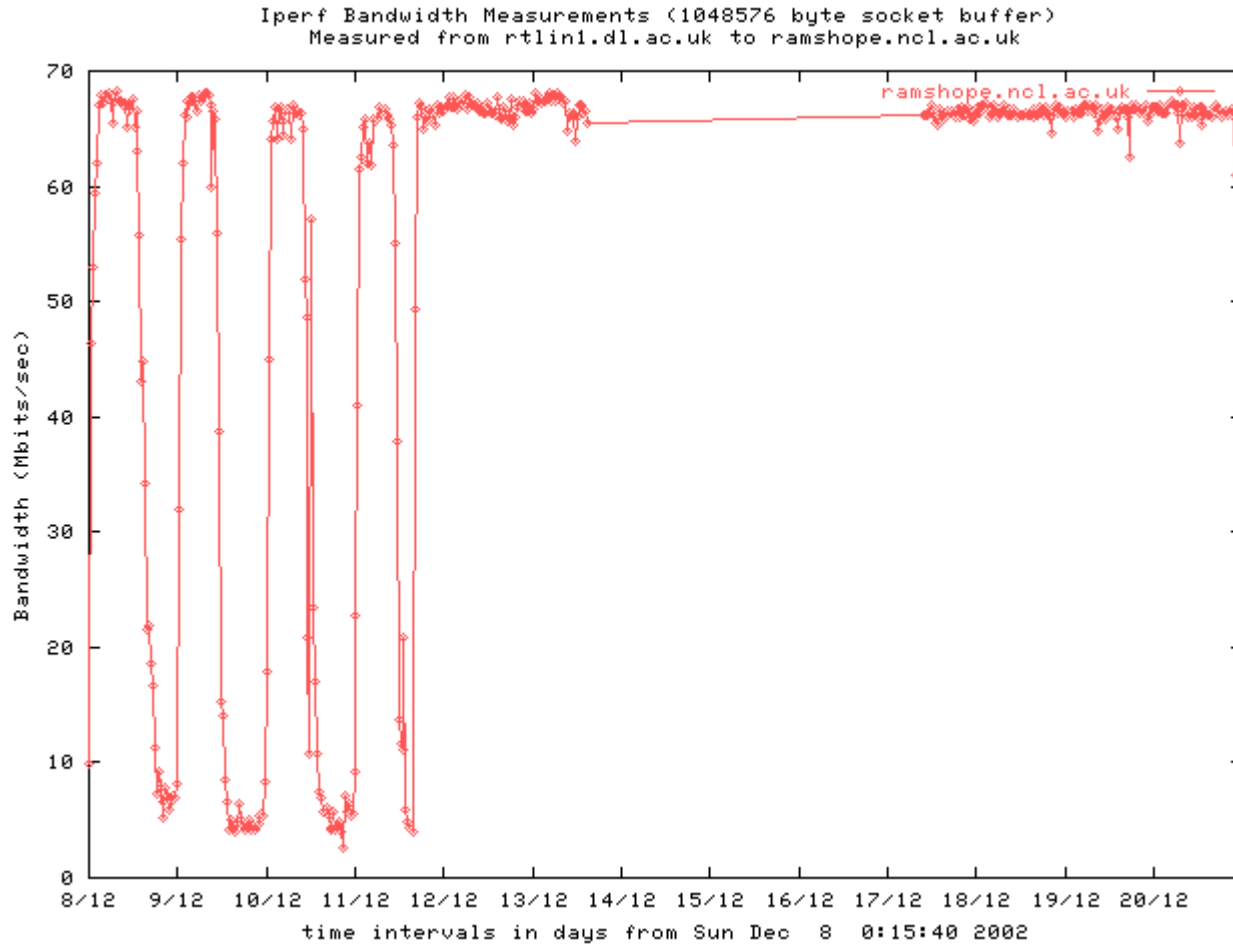


# Examples: PingER



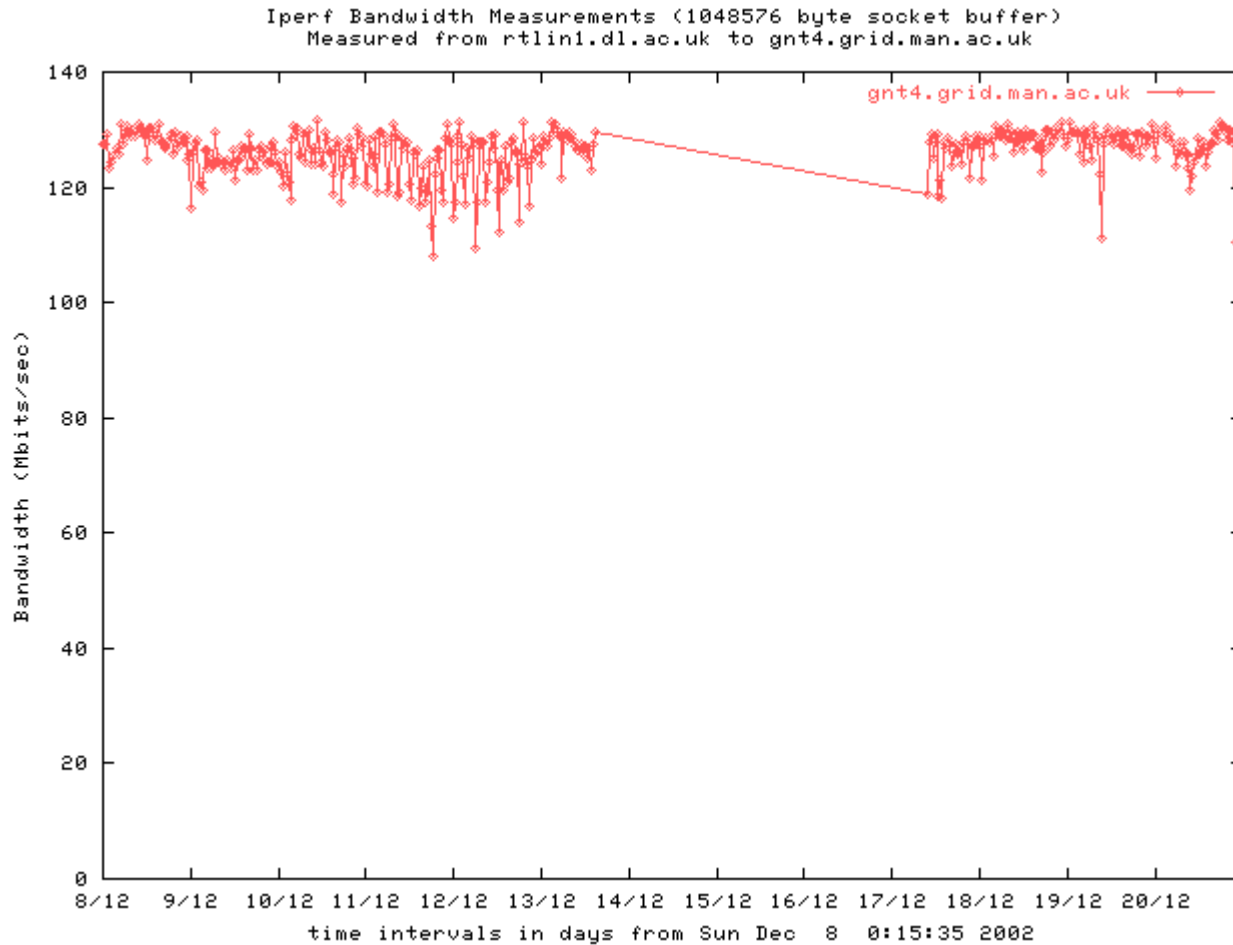


# Examples: IperfER(1)



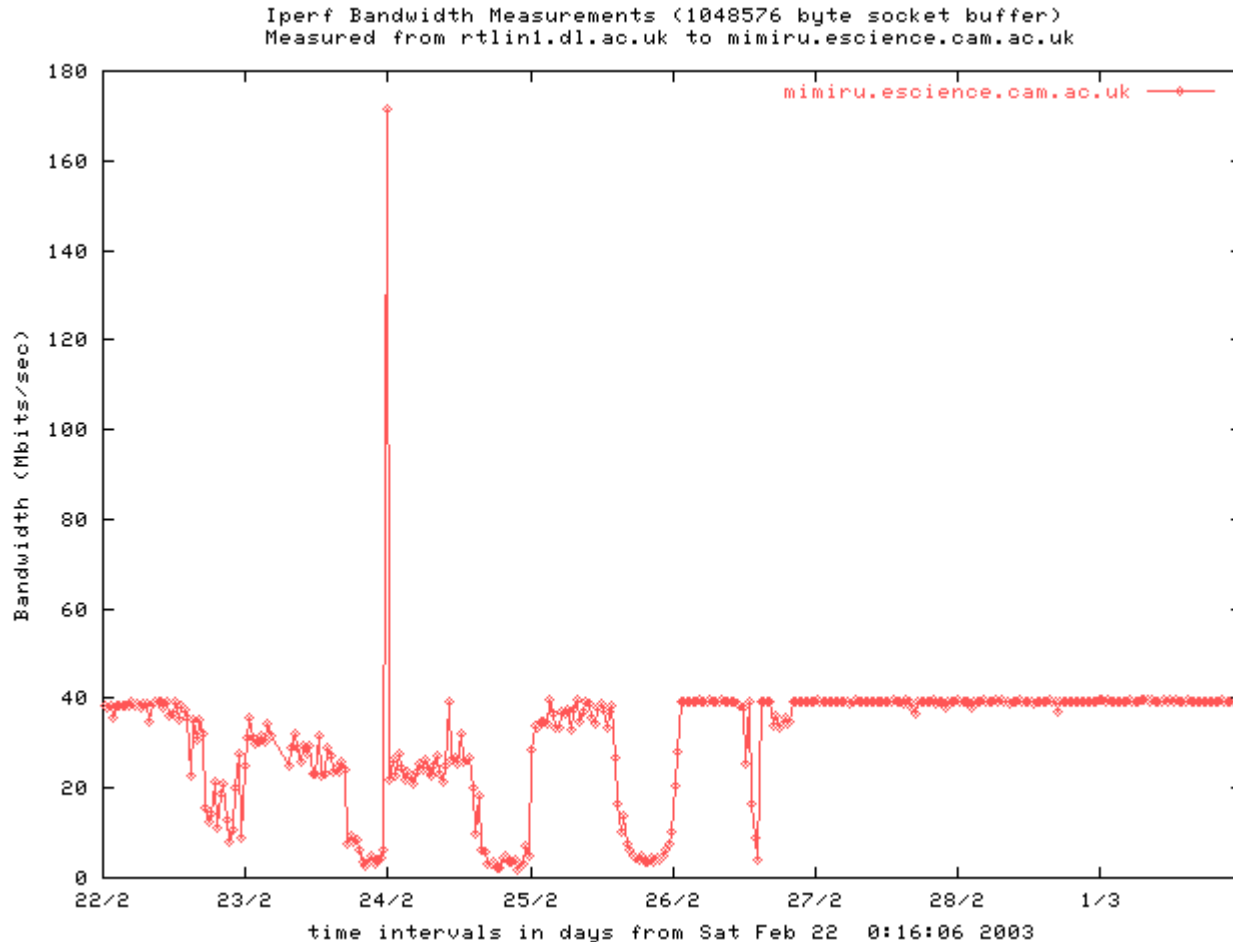


# Examples: IperfER(1)



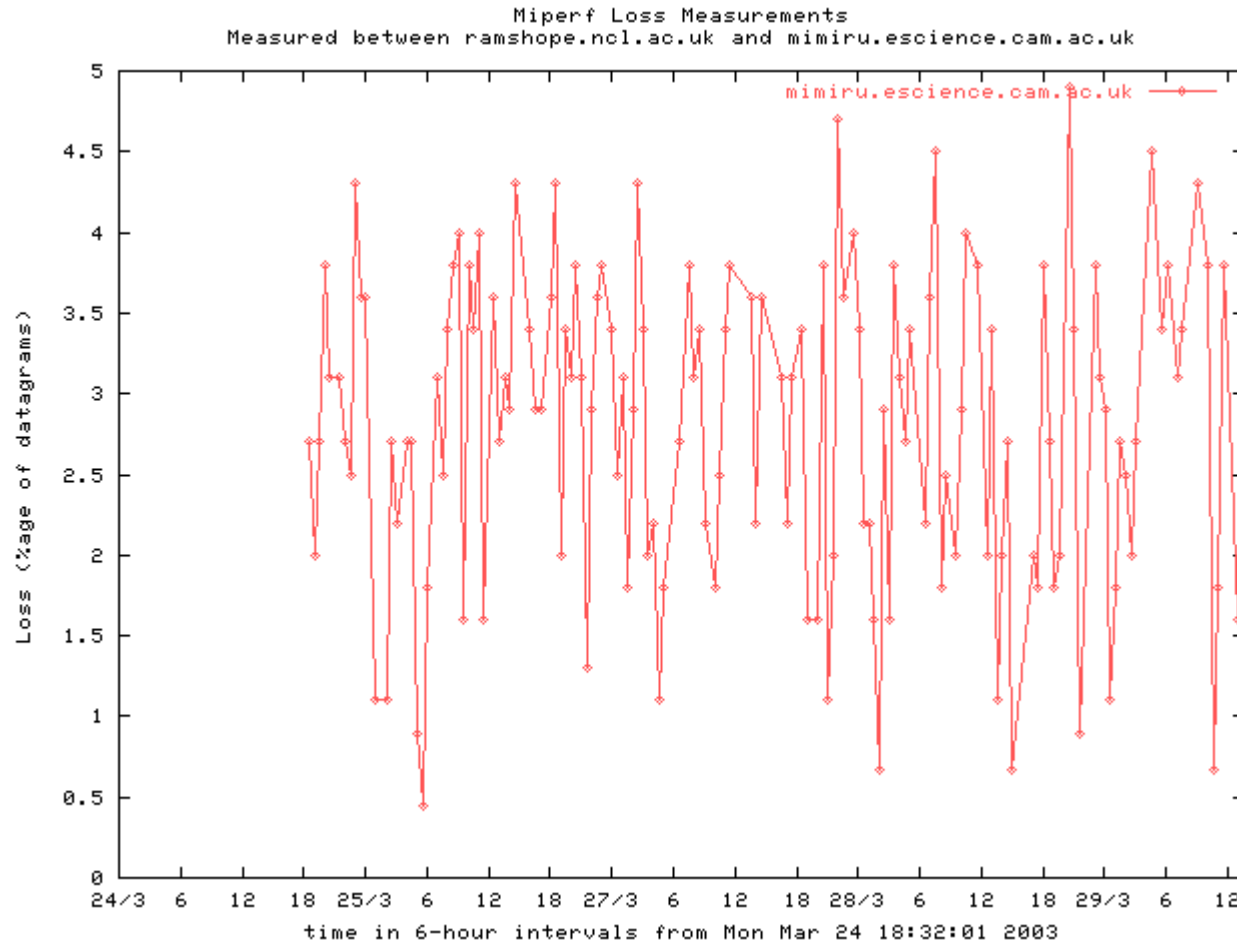


# Examples: IperfER(2)



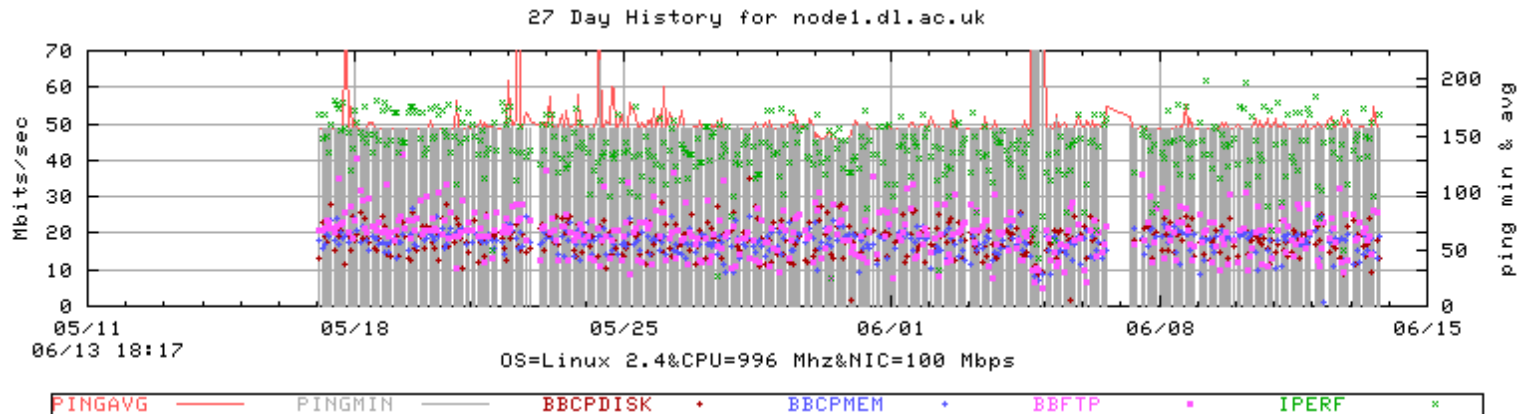
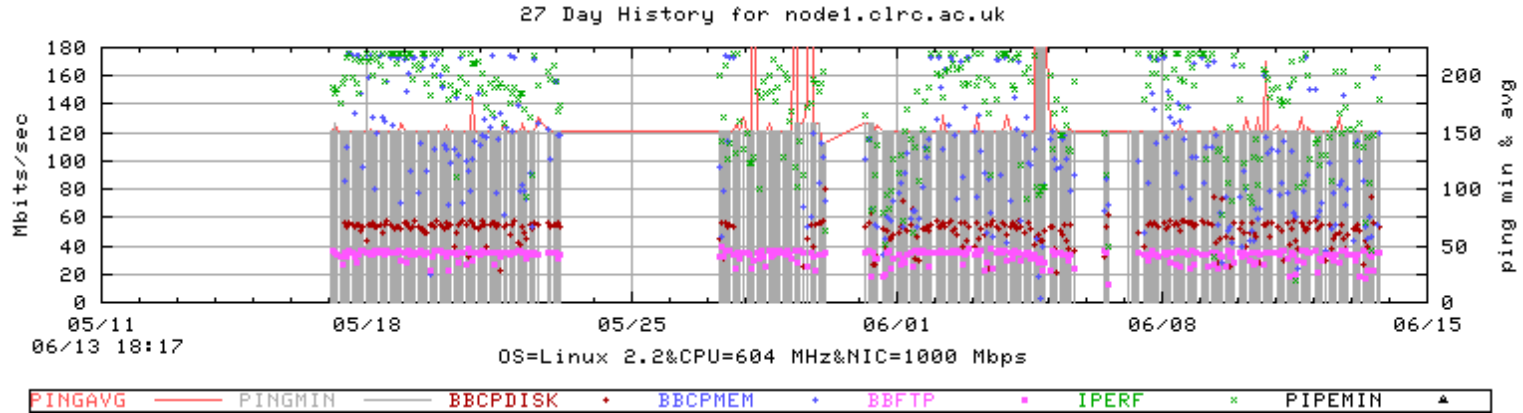


# Examples: MiperfER





# Examples: bbcp & bbftp



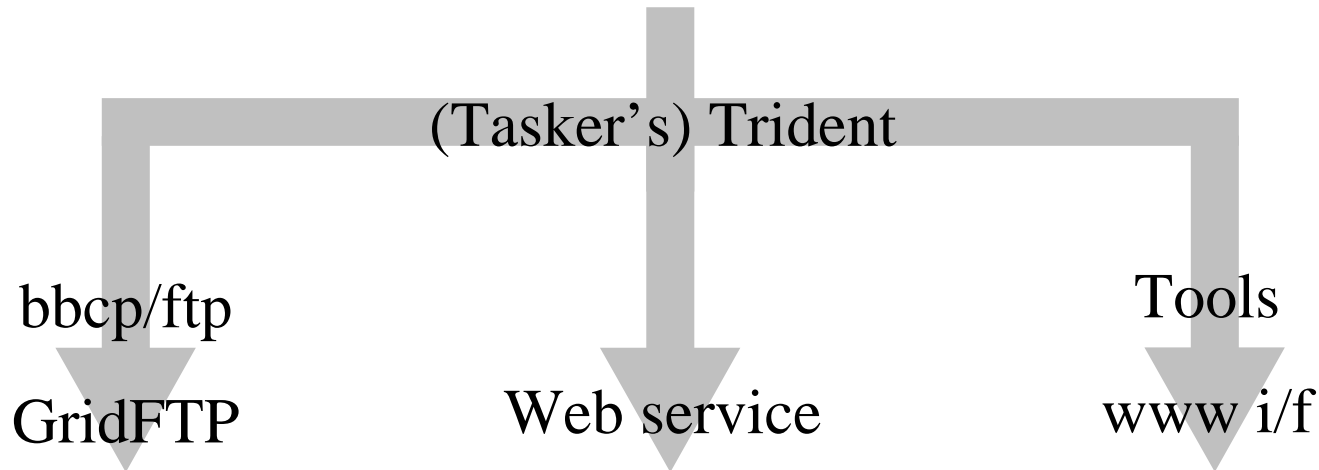


# Benefits vrs Costs

<b>Costs</b>	<b>Benefits</b>
Bandwidth testing is intrusive	<ul style="list-style-type: none"><li>• Grid s/w makes decisions based on network performance</li><li>• Increased use → pronounced consequences</li><li>• "right" metrics ⇒ better decisions</li></ul>
Monitoring host (PC)	Invaluable tool in detecting problems, e.g. bottlenecks
	Users can see performance for themselves



# The Future



Longer term:

- Wishlist features (but perhaps not all 3,000,000)
- Investigate other issues: window sizes, QoS...

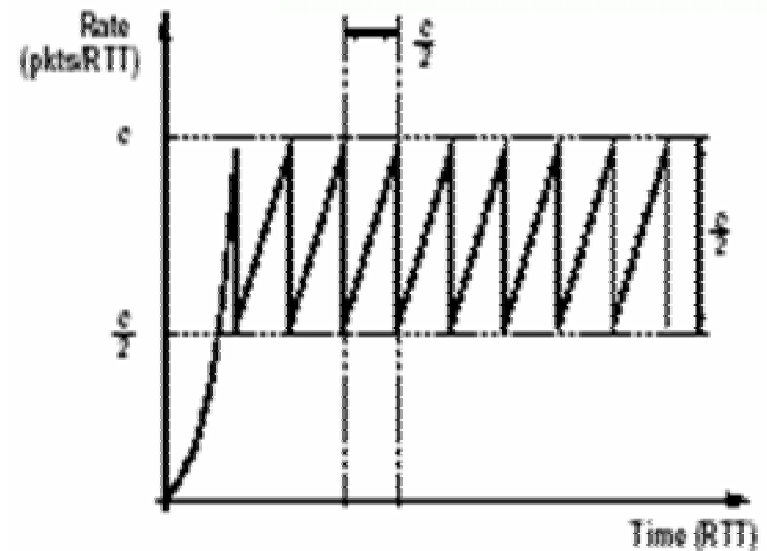
**Err, is that all?**



# The Gandhi Idea

TCP underperforms @ 7,000 revs

- Slow start, rapid fall off



- “High capacity” also a problem
  - Long pipes (high RTT) fill with data. Problems if data needs discarding
  - Same affect with high b/w (large) pipes
  - Not all research is applicable: UK RTT relatively small



# TCP Best Practice

- Demo TCP best practice to users – “this is what you can achieve, if you ‘tune’ your machine like this...”
- Possibles “variables”:
  - Kernel versions and patches
  - MTU
  - Interrupt handling
- However:
  - “...performance from copper-based GigE cards are intimately connected with judicious use and understanding of the corresponding driver(s)...”  
Gray and Betz
  - Bus speed, bus width...

**No new fangled gizmos – using what we already have!**



# Conclusion

- We monitor to:
  - find faults, predict performance
- Important for the grid:
  - the Grid is different, behaviour altering
- The toolkit:
  - not a panacea, but a good start
- What do YOU do next?.....
  - <http://gridmon.dl.ac.uk/>





# Network Monitoring for e-Science

? Questions ?

[m.j.leese@dl.ac.uk](mailto:m.j.leese@dl.ac.uk)

<http://gridmon.dl.ac.uk/~mjl>