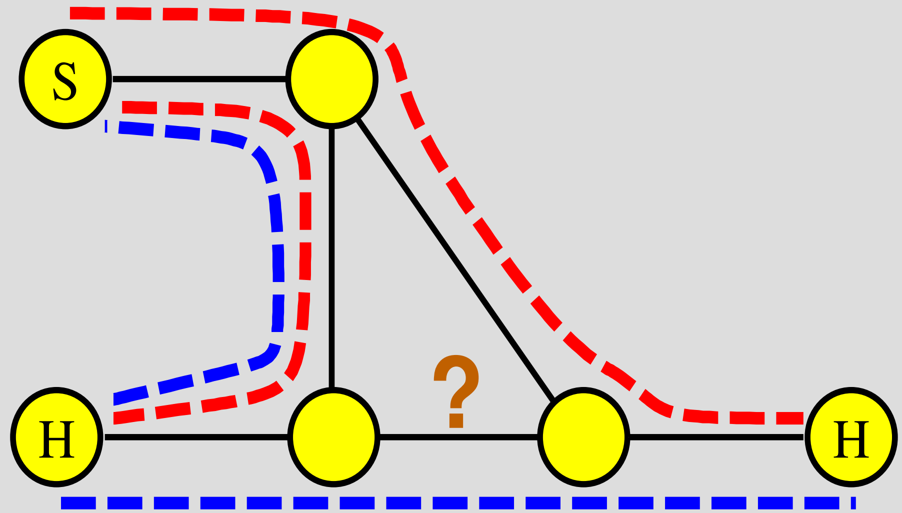
DAR
Active measurement in the large

Tony McGregor

RIPE NCC Visiting Researcher
tony.mcgregor@ripe.net

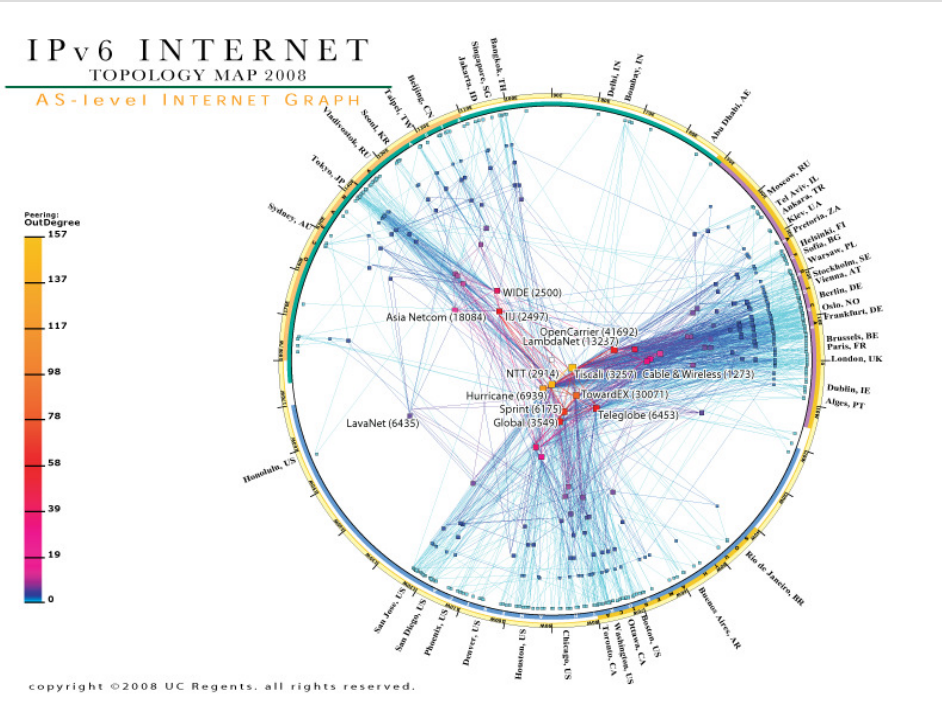
The University of Waikato
tonym@cs.waikato.ac.nz



Challenges in Active Measurement

Topology

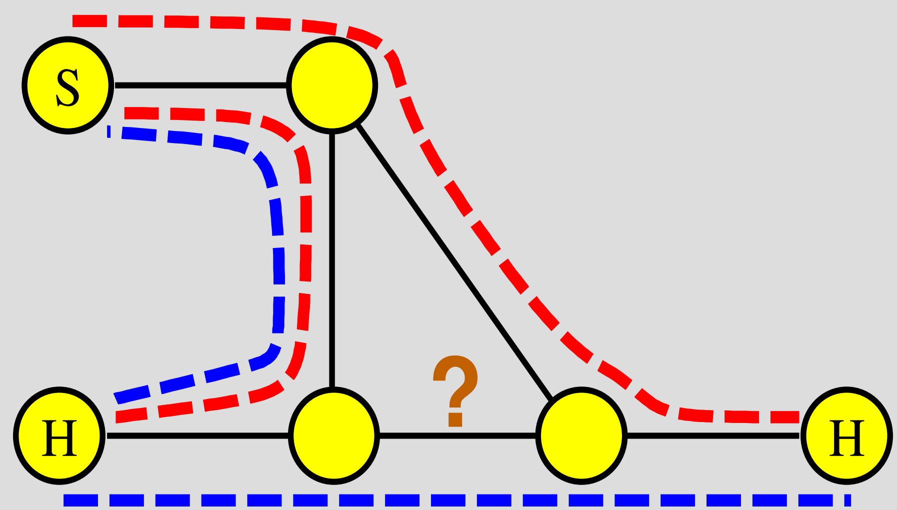
- Can measure topology from a small (~100s) number of sources to many destinations
  - e.g. ARC/scamper (CAIDA)
  - PlannetLab
- Probe perspective bias
  - Academic
  - Well connected
- Selected destinations
  - May not be active
- Asymmetry
  - Peer to peer
- Cycle time
- NATs



Challenges in Active Measurement

Routing Failures

- Can discover many failures as seen from available perspectives
  - Hubble
- Missed Failures
- Masked failures
  - A failure close to a monitor masks others
- Accurate location
  - Direction of failure
    - Limits of spoofing
  - Extent of failure
  - Path asymmetry



Challenges in Active Measurement

Summary

- Limited perspectives
  - Roughly in the order of
    - 0.001% of end-hosts
    - 0.2% of Autonomous Systems
  - Won't have a probe that sees many events
  - Asymmetry

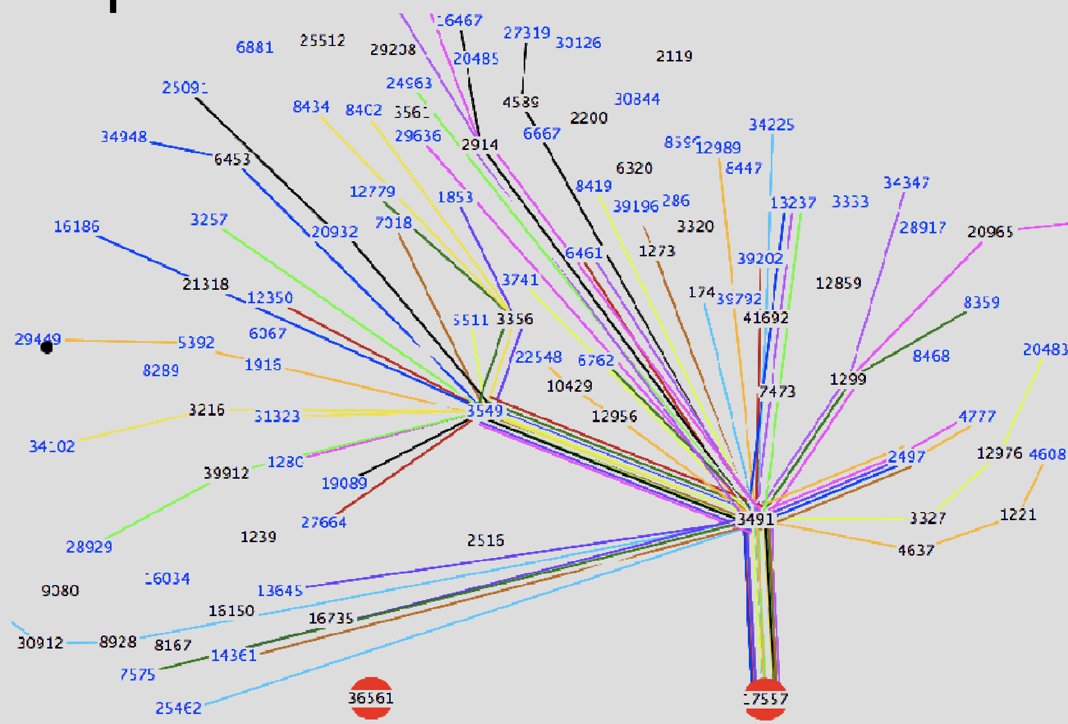
- Probing to third party destinations
  - Responsiveness
    - Timely response
    - Any response
    - Loading

- NAT

Example Application

Is my network globally reachable?

- Notification service for reachability events like the YouTube hijack
  - Or smaller event affecting just one network
- Current data (e.g. RIS) useful after the event
  - Path changes are normal operation
  - Need real time reachability
    - Hubble like
    - wider range of vantage points
      - Non-academic
      - Leaf-nodes
      - More
  - Possibly combined with BGP data



Other Applications

- Is there a routing hole in my network
  - between particular source/destination pairs
- Bidirectional topology
  - How asymmetric is the Internet?
  - What is the path from X **to** me?
  - For testing of new protocols and applications
    - simulation
- Overlay network routing
- What is the performance to my network?
  - on average
  - from a particular network?

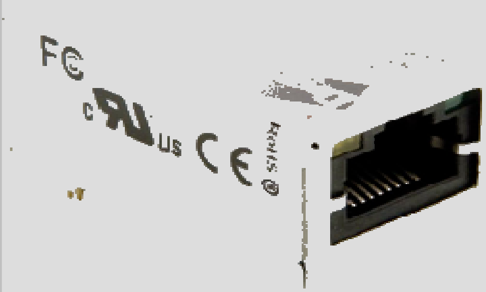
DAR

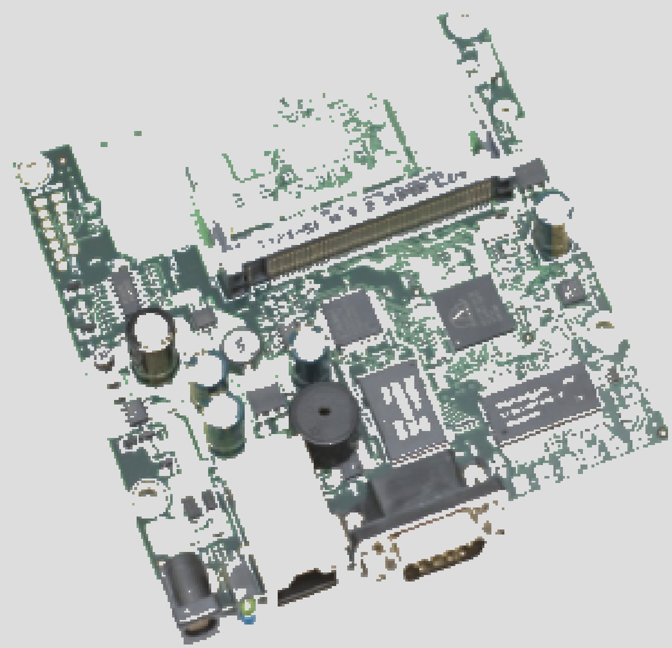
Diverse Aspect Resource

- Can we design, build, maintain and make good use of an active measurement system with in the order of 100.000 active probes?
- What might it look like?
- What are the key challenges?

Hardware Probes

- Hardware must be cheap and robust
- Token or single board computer
- Specs in the ballpark of:
  - 300MHz processor
  - 64MB Flash
  - 64MB SDRAM
  - 10/100 Mbit/s Ethernet
- Heterogeneous deployment

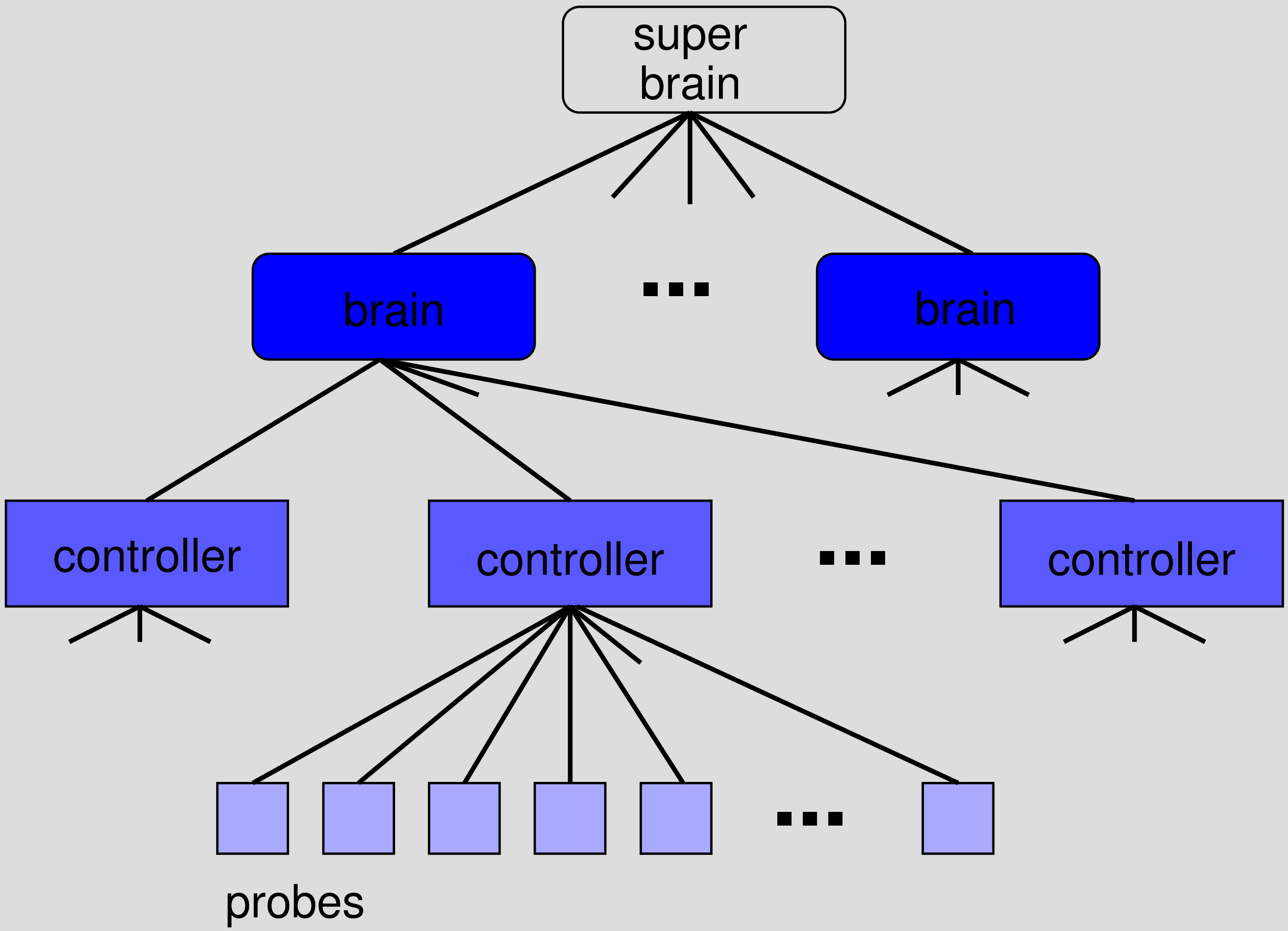




Software Probes

- DAR should also support software only probes.
  - Package downloaded and run on a host
- More volatile than hardware probes
- Different performance characteristics

Hierarchy

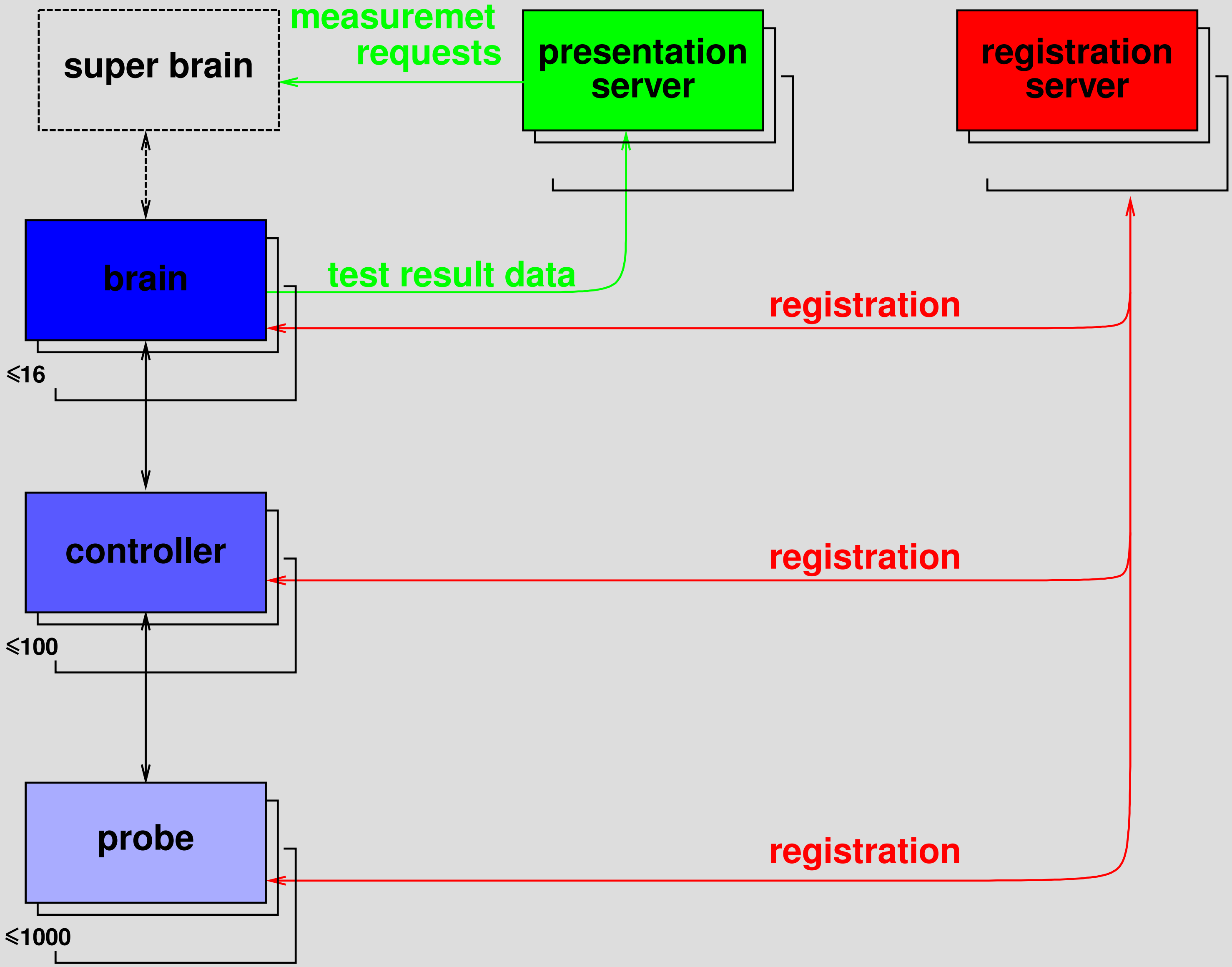


super brain
...
brain
brain
controller
controller
...
controller
...
probes

Architecture

- Still very fluid
- Presented here to give overall impression
- Numbers are possibilities

Overview of an Architecture



Probe

- Token
- Performs low level measurements
  - ping, traceroute, send packet
- On boot registers with a controller
  - Finds suitable controller via registration server
- Software remotely upgradeable
- Resources will be limited
  - Hardware
  - User limits

- Low reliability
  - The set of available probes is always in flux
- In the order of 100,000 probes

Controller

- Manages a set of probes
- Keeps track of what probes are available
- Can answer questions about what resources each probe has
  - Location (ip, as)
  - Bandwidth available
  - Memory for result storage
- Accepts work requests from brain
- Aggregates results

Controller

- Medium reliability
  - Shouldn't go down but system must continue operation if one or more controllers have failed
- Up to 1000 controllers with up to 1000 probes each

Brain

- Manages a set of controllers
- “Implements” a measurement application
  - May involve many low level tests
- Knows or can discover what resources each controller has.
  - Allocates work to controllers

- Very reliable. Measurement fails if a brain fails.

- 1 – 16 brains each controlling up to 256 controllers

Super brain

- Not clear that there will be a super brain
- If there is
  - Overall supervision of brains
    - Allocation of work between brains
    - Maintaining state of brains
  - Location of resources that only some brains may support

- Only ever a single super brain
- Hardened against failure
  - If the super brain fails brains continue to operate but new measurements may not be possible

Presentation Service

- Interface with users
  - Presents data (e.g. via web)
  - Accepts requests for new work from users
- Store data
- May be multiple servers cooperating to provide enough resources and stability.
  - Standard approaches

- High availability but data collection should continue (for a while) if service fails
- 1 – 10 servers

Registration Service

- Contacted by probes and controllers when the boot
  - Exists at well know location (DNS and/or IP)
- Very simple service
  - Highly reliable and can handle many requests
- Very stable
  - Replicated for reliability
- 1 – 5 identical instances, up to 100,000 probes per instance

Major Challenge

- It is not obvious how to design measurements from a very large number of probes
  - Probably can't do full mesh measurements
    - 100,000 pings + 100,000 replies + 100,000 other nodes pinging + replies = full capacity of 256Kb link for ~10 min. => long cycle time
  - Even investigating a routing failure to a single destination a traceroute from every source to target creates a hot spot at the target
- Optimised measurement techniques needed
  - e.g. doubletree for traceroute
  - Optimised ping?
- Focus of current work

Other Questions

- What principles should guide the choice of which controllers to associate a probes with.
  - Function
  - Location
- Similarly for controller/bring and brain/super brain association

- How generic should we be
  - More generic more likely to meet future needs
  - Less efficient
  - More complex

Other Questions

- How to encourage users to deploy probes
  - Hardware or software
- How to respond to a failed probe
  - Automated
- “Abuse” notifications
- And lots more!

Conclusion

Thoughts and comments are very welcome

sg@ripe.net