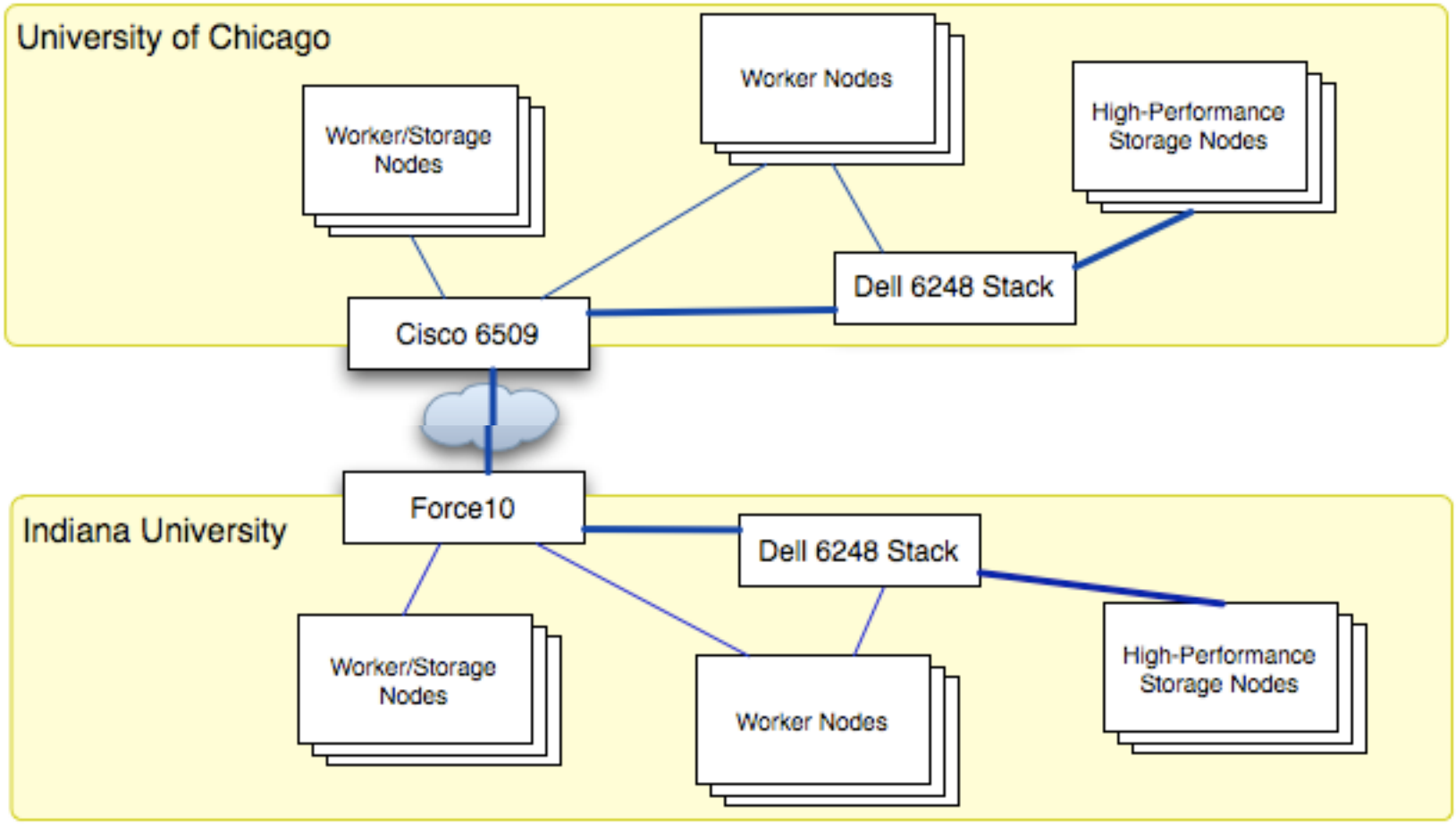
Wide-Area Xrootd testing at MWT2

Sarah Williams
Indiana University

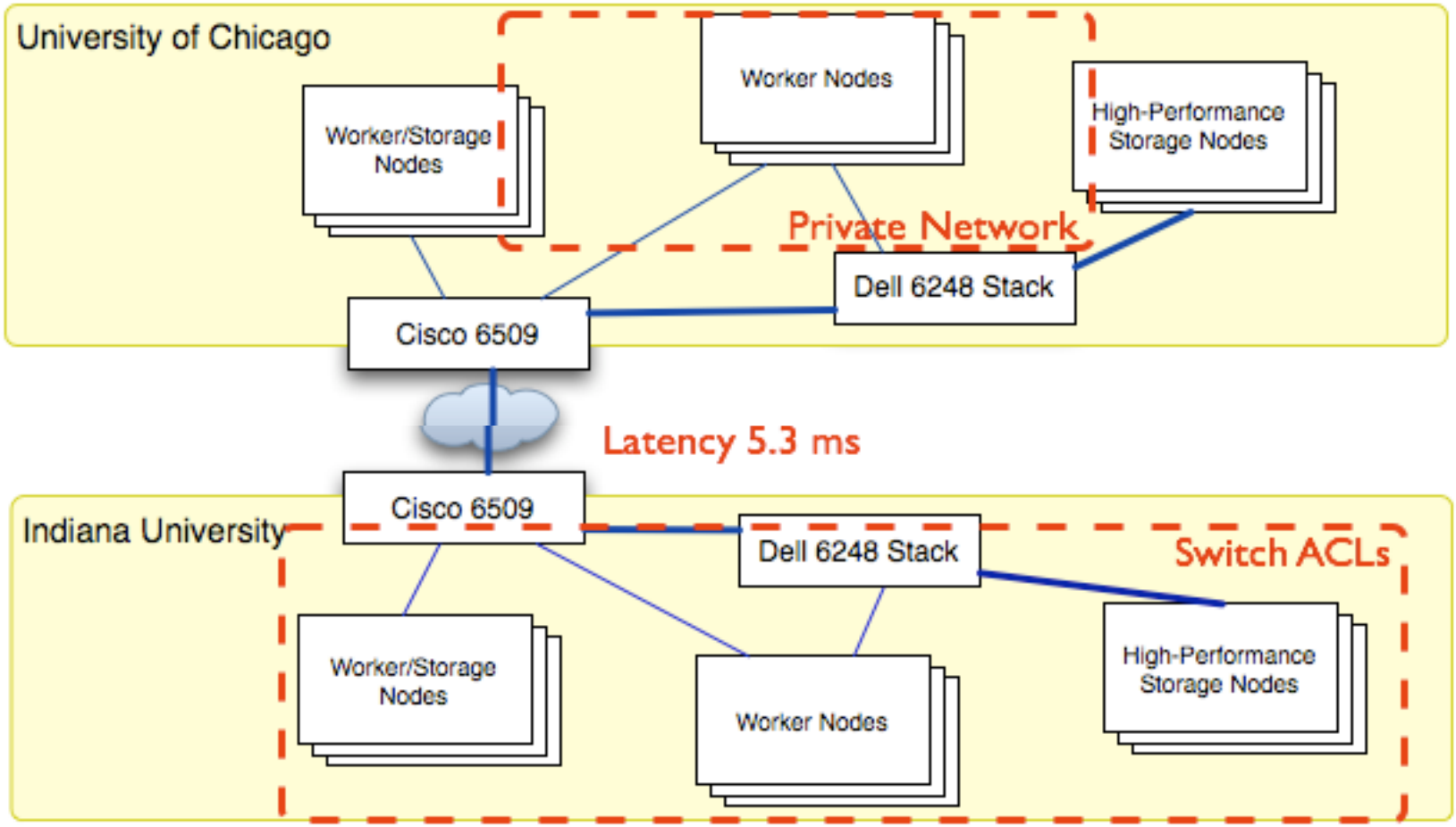
Background & Motivation

- Currently MWT2_IU and UC function as separate clusters, using dCache on the SE
- Cluster Unification – more storage for analysis, more capacity available, one set of configurations to manage
- Looking for something easier to manage
- Good read, write and analysis performance
- Doesn't introduce new stability or reliability issues

The Current Situation



Obstacles to a Cross-site Solution



Test Cases

- Can a private-ip worker node read & write to a public-only dataserver? a public-only redirector?
- Does a dual-homed public/private ip dataserver handle requests from public and private ips correctly? Does a public/private ip redirector?
- What is read & write speed compared to dCache with dccp?
  - When connecting to a high-performance storage node? To a combination worker/storage node?
  - When connecting between sites with xrdcp vs within the site with dccp

Test Setup 1

Role	Hostname	Location	Network	Speed
Redirector	Iut2-s1	IU	Public	10G
DataServer	Iut2-s3	IU	Public	10G
	Iut2-c034	IU	Public	1G
	Uct2-s3	UC	Public/ Private	10G
	Uct2-c034	UC	Public/ Private	1G
Worker node	Iut2-c042	IU	Public	1G
	Uct2-c035	UC	Private	1G

Test 1 Conclusions

- Can a private-ip worker node read & write to a public-only dataserver? a public-only redirector? YES and YES. Switch ACLs must allow traffic from all nodes to port 1094
- Does a public & private ip dataserver handle requests from public and private ips correctly? YES

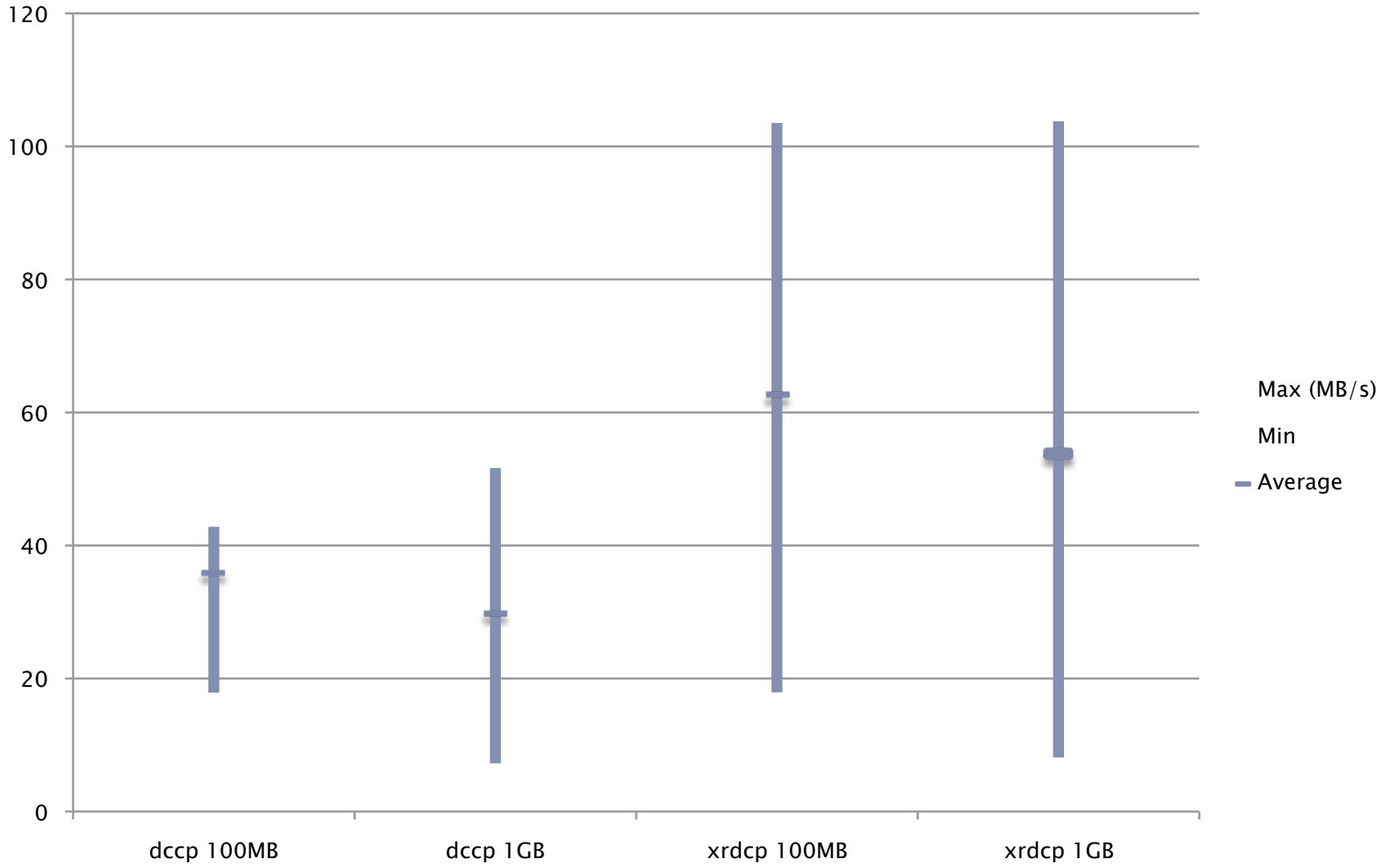
Test Setup 2

Role	Hostname	Location	Network	Speed
Redirector	Uct2-s1	UC	Public/ Private	10G
DataServer	Iut2-s3	IU	Public	10G
	Iut2-c034	IU	Public	1G
	Uct2-s3	UC	Public/ Private	10G
	Uct2-c034	UC	Public/ Private	1G
Worker node	Iut2-c042	IU	Public	1G
	Uct2-c035	UC	Private	1G

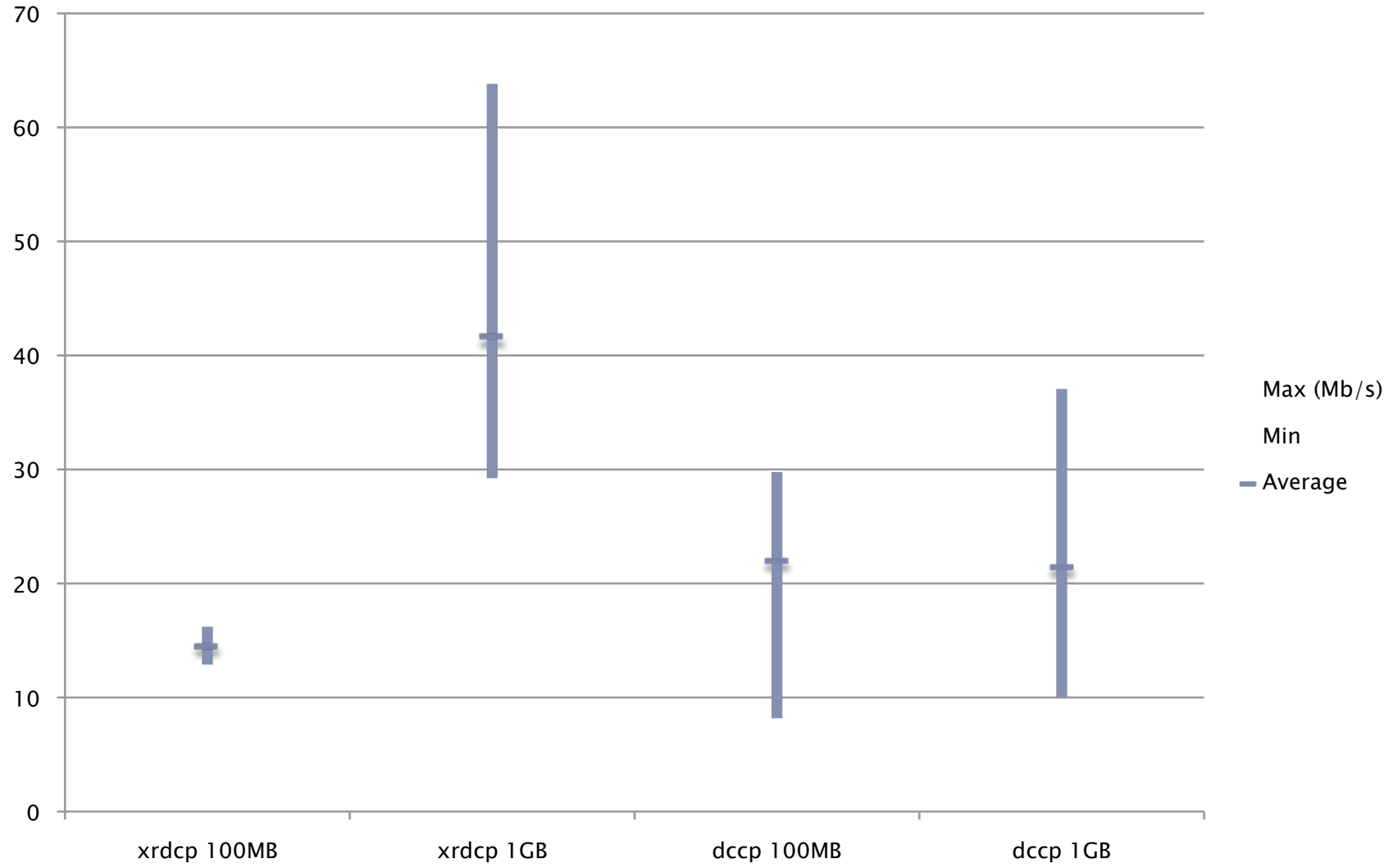
Test 2 Conclusions

- Does a public & private redirector handle requests from public and private ips correctly? YES
- Switch ACLS must allow traffic from the redirector to all high-numbered ports on the data servers

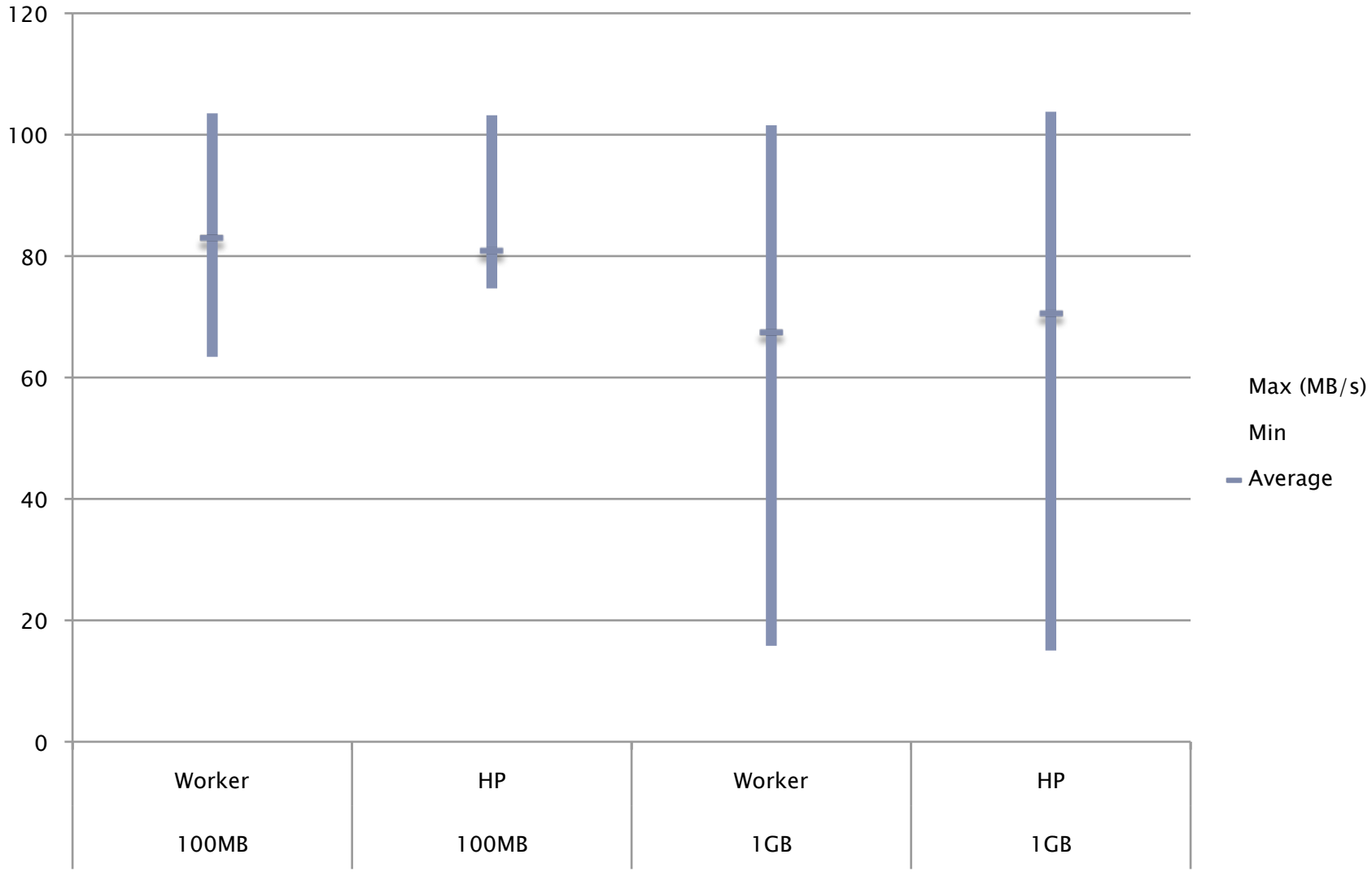
Read Measurements



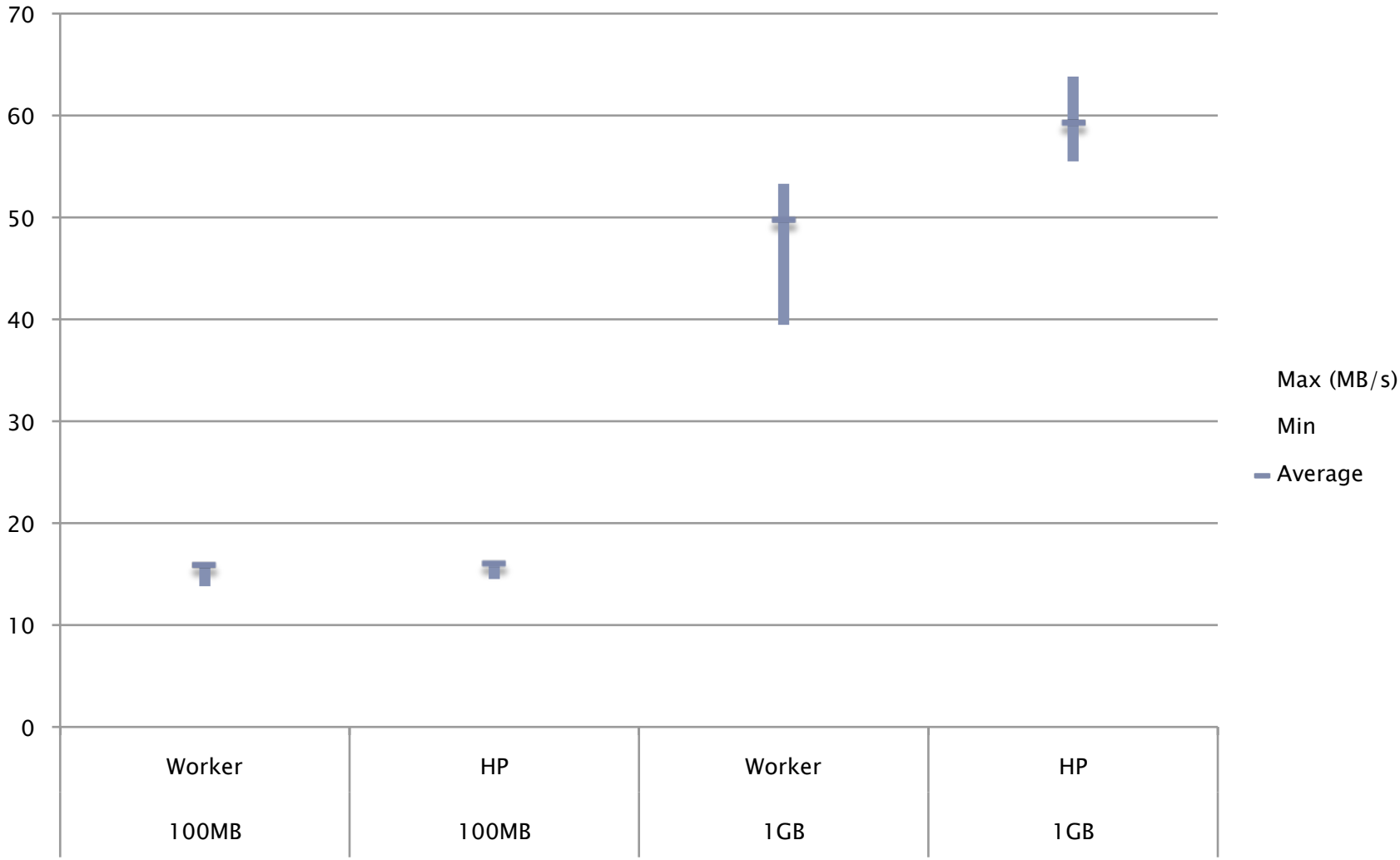
Write Measurements



Read Performance by Hardware



Write Performance by Hardware



Conclusions

- Xrootd can be configured to work with our network constraints
- Worker nodes need to be tuned for WAN performance
- Write performance is acceptable
- Read performance is good

Proposed Production Configuration

- Peered redirectors, one at each site, in a load-balanced configuration. Provides fail-over if one dies or if the link between the sites fails.
- The link between the sites has never (yet) failed.
- In addition, a supervisor node at each site. Xrootd requires supervisors when the number of data servers >64. MWT2 has 87.

Questions?