



Digital Institute

---

# Simulating Condor

Stephen McGough, Clive Gerrard & Jonathan Noble  
Newcastle University

Paul Robinson, Stuart Wheater  
Arjuna Technologies Limited

Condor Week 2012





Digital Institute

---

# Overview

- Motivation and Background
- Condor Simulation
- Power Management Evaluation
- Conclusion



Digital Institute

---

# Overview

- Motivation and Background
- Condor Simulation
- Power Management Evaluation
- Conclusion

# Motivation

---

- Newcastle University has strong desire to reduce energy consumption
  - Currently powering down computer & buying low power PCs
  - “If a computer is not ‘working’ it should be powered down”
- Can we go further to reduce wasted time?
  - Reduce computer idle time
  - Identify wasteful work sooner?
- We have a number of policies we’d like to evaluate
  - Difficult on running system, measuring power
- Aims
  - Investigate policy for reducing energy consumption
  - Determine the impact on high-throughput users

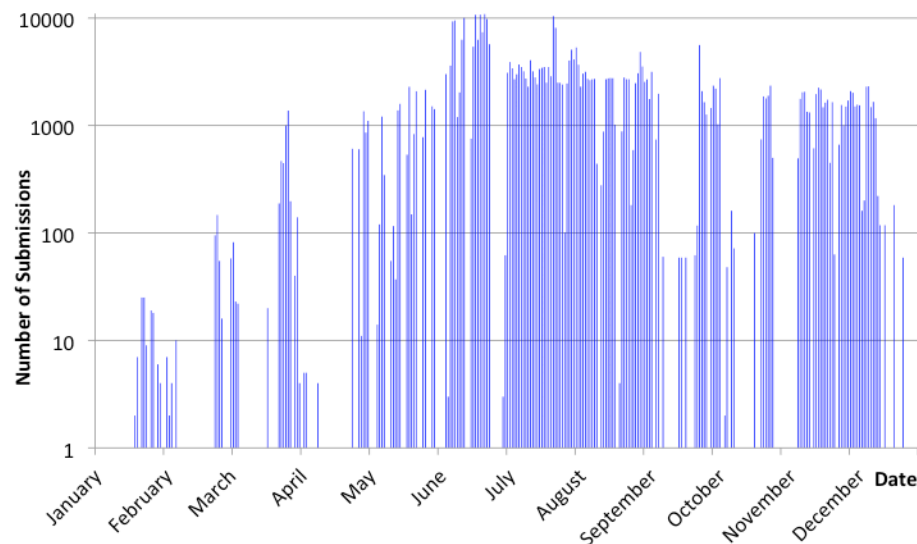


# Condor At Newcastle

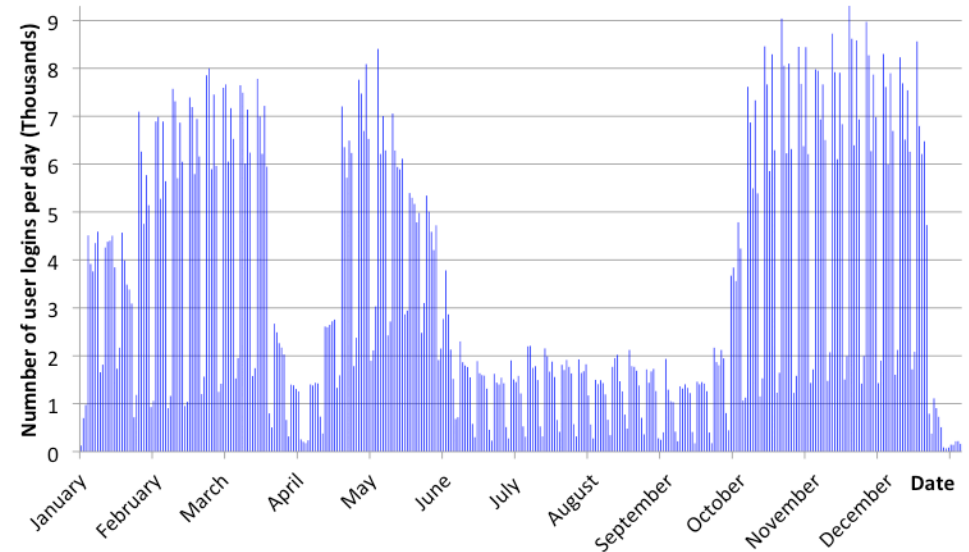
Digital Institute

- Comprises of ~1300 open-access computers based around campus in 35 'clusters'
- All computers at least dual core, moving to quad / 8 core

Job Submissions



User Logins

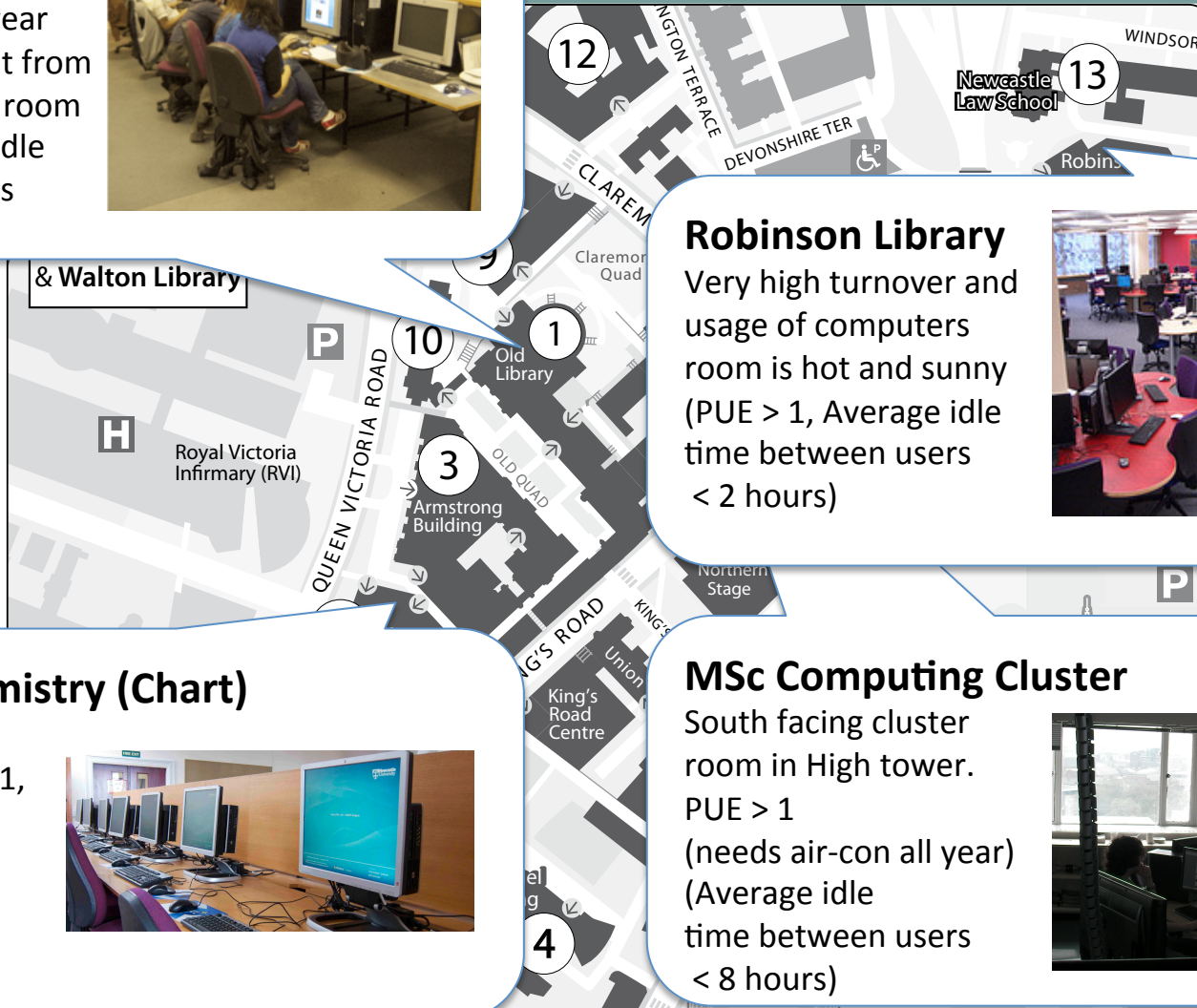




# ter Locations

## Old Library

Basement Cluster room  
Needs heating all year  
PUE < 1 (offset heat from computers against room heating) (Average idle time between users < 5 hours)



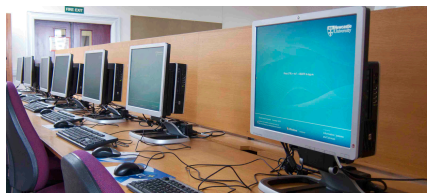
## Robinson Library

Very high turnover and usage of computers  
room is hot and sunny (PUE > 1, Average idle time between users < 2 hours)



## School of Chemistry (Chart)

Very low usage of Computers (PUE ~ 1, Average idle time between users ~23 hours)



## MSc Computing Cluster

South facing cluster room in High tower.  
PUE > 1 (needs air-con all year) (Average idle time between users < 8 hours)



Power Usage Effectiveness (PUE) – depends on location of computer (and time)  
Power Efficiency: efficiency = flops/(PUE \* watts)



Digital Institute

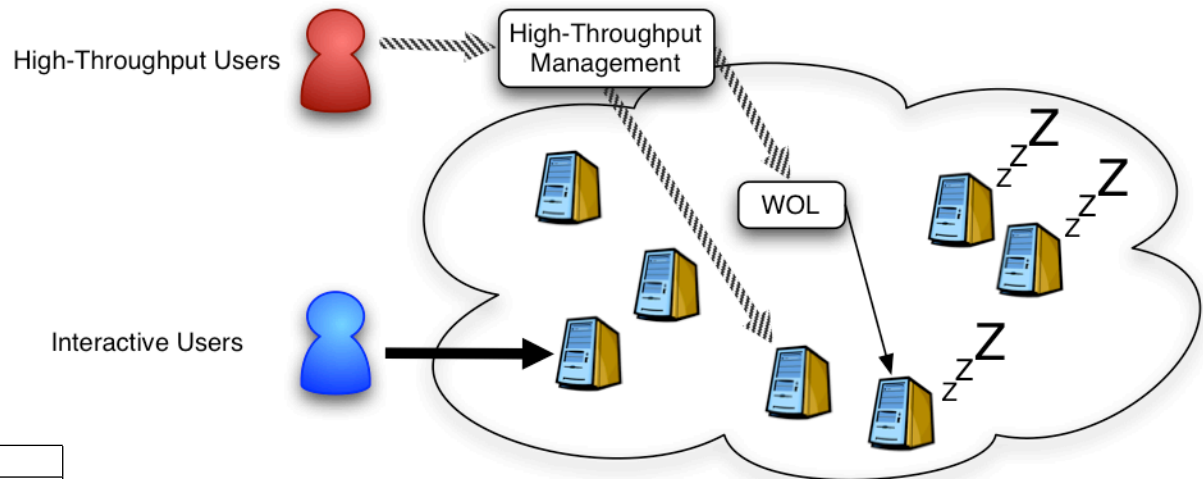
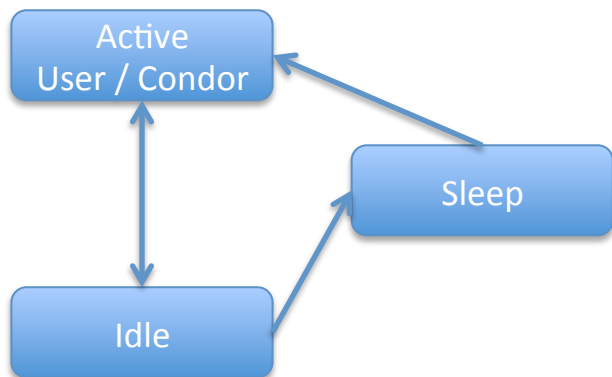
---

# Overview

- Motivation and Background
- **Condor Simulation**
- Power Management Evaluation
- Conclusion

# Condor Simulation

- High Level Simulation of Condor
  - Trace logs from the last year are used as input
    - User Logins / Logouts (computer used)
    - Condor Job Submission times (and duration)
    - Cluster open times and and policy



Type	Cores	Speed	Power Consumption		
			Active	Idle	Sleep
Normal	2	~3Ghz	57W	40W	2W
High End	4	~3Ghz	114W	67W	3W
Legacy	2	~2Ghz	100-180W	50-80W	4W





Digital Institute

---

# Overview

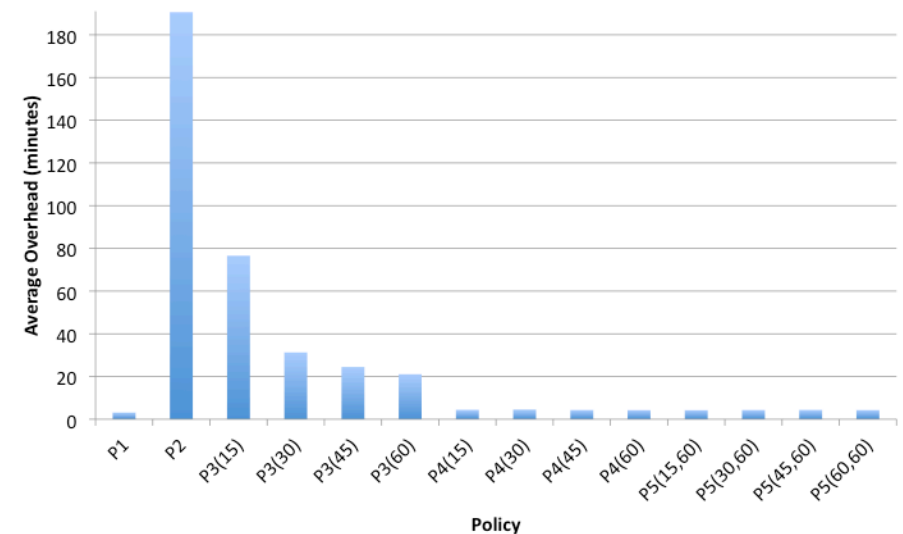
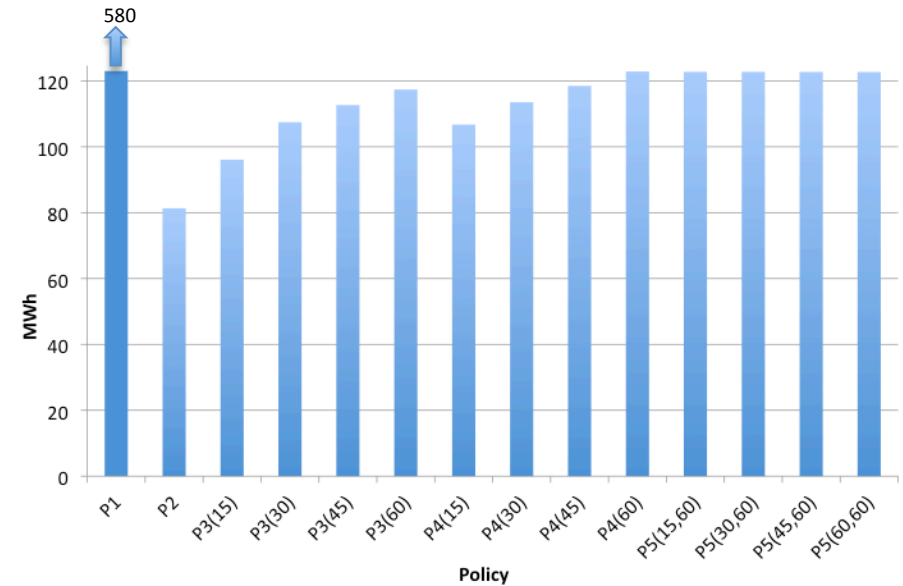
- Motivation and Background
- Condor Simulation
- **Power Management Evaluation**
- Conclusion



# Power State Policy

Digital Institute

- P1: Computers are always on
- P2: On during cluster open hours and off otherwise, no mechanism to wake up
- P3: Computers sleep after n minutes of inactivity with no remote wake up
- P4: Sleep after n minutes of inactivity but can be remotely woken up
- P5: Sleep after n mins of inactivity but Condor is only informed every m mins

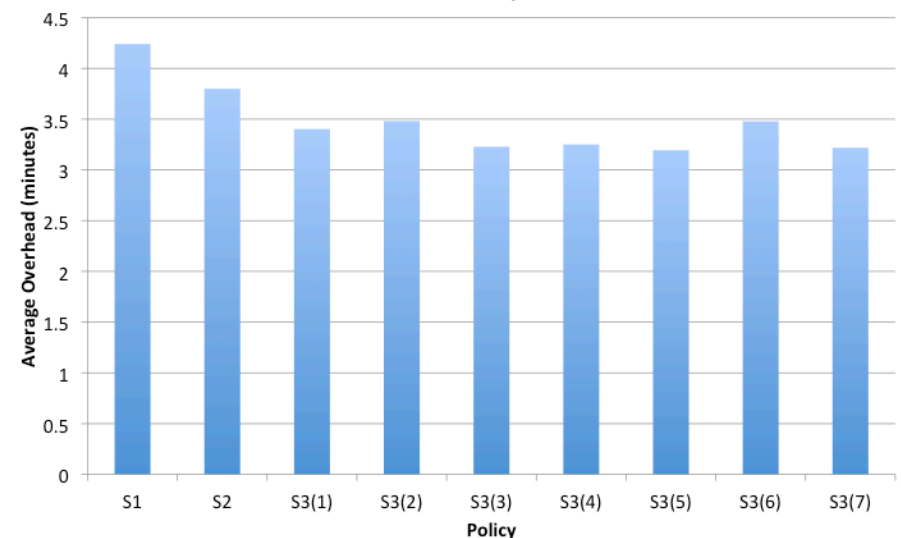
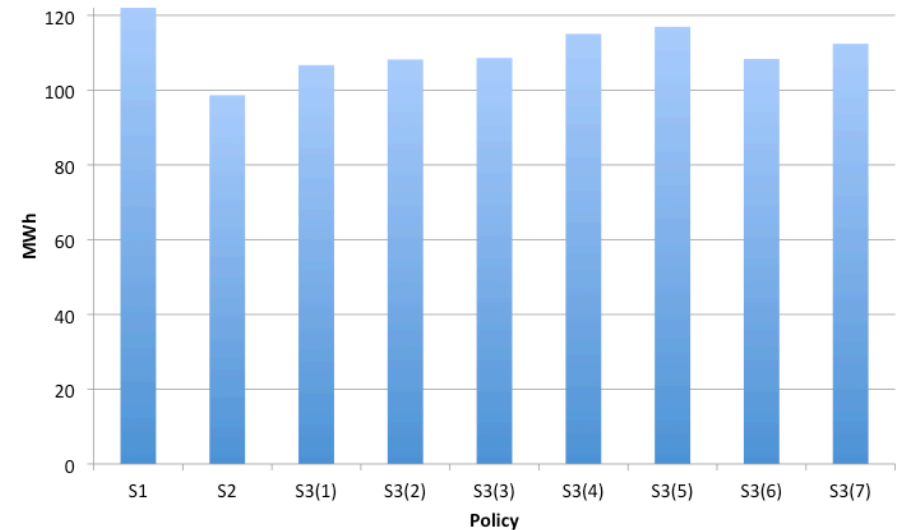




# Computer Selection Policy

Digital Institute

- S1: No preference (random)
- S2: Target most energy efficient computers
- S3: Target least used computers
  - Least number of interactive logins
  - Largest intervals between logouts and logins

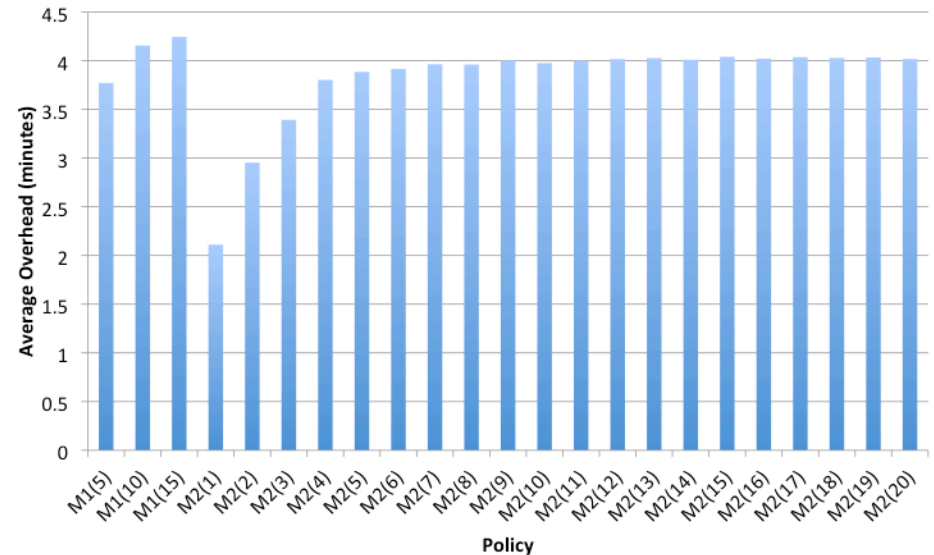
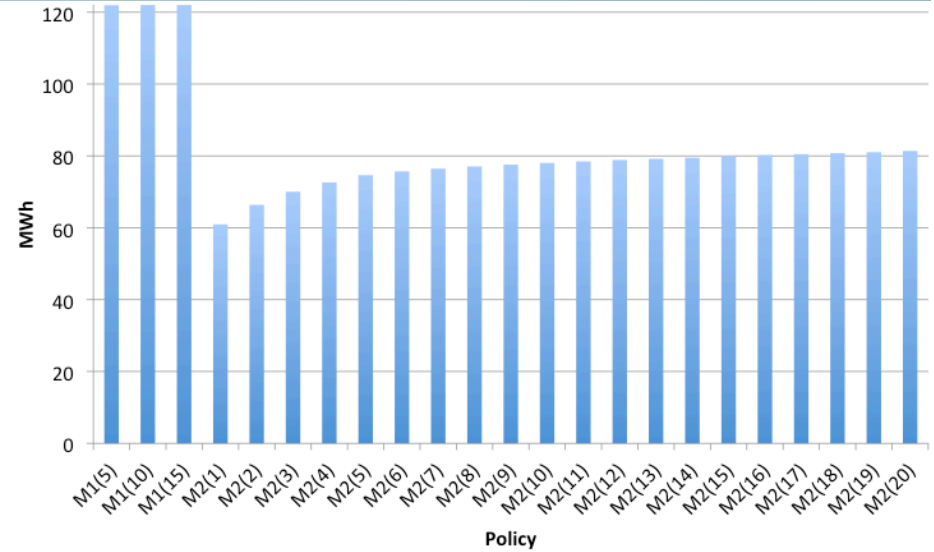
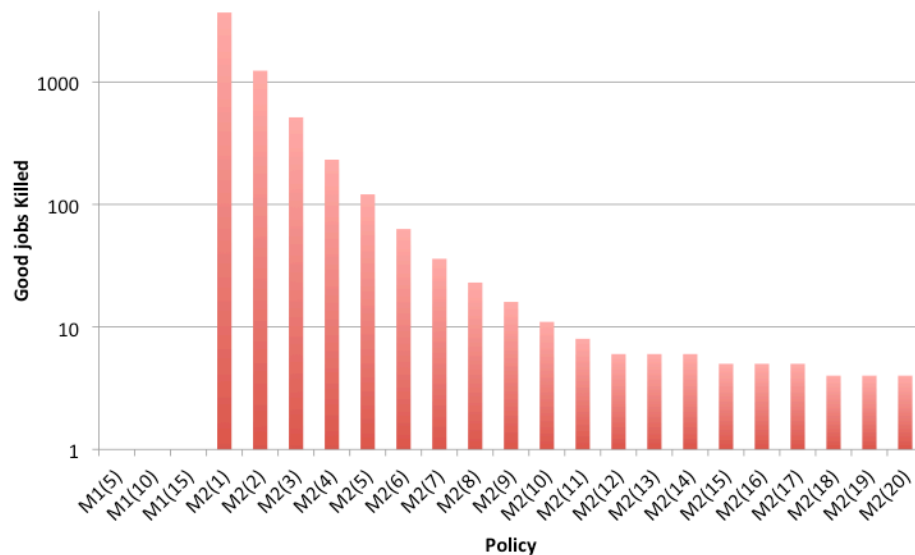




# Management Policy

Digital Institute

- M1: Computer is idle for at least n minutes before a Condor job can run on it
- M2: If a job is started more than n times mark it as 'miscreant' and don't re-start

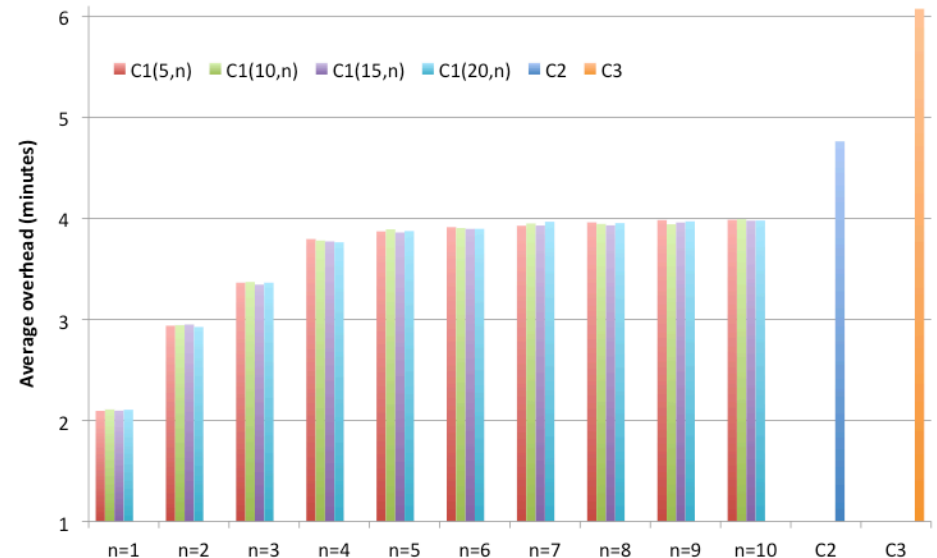
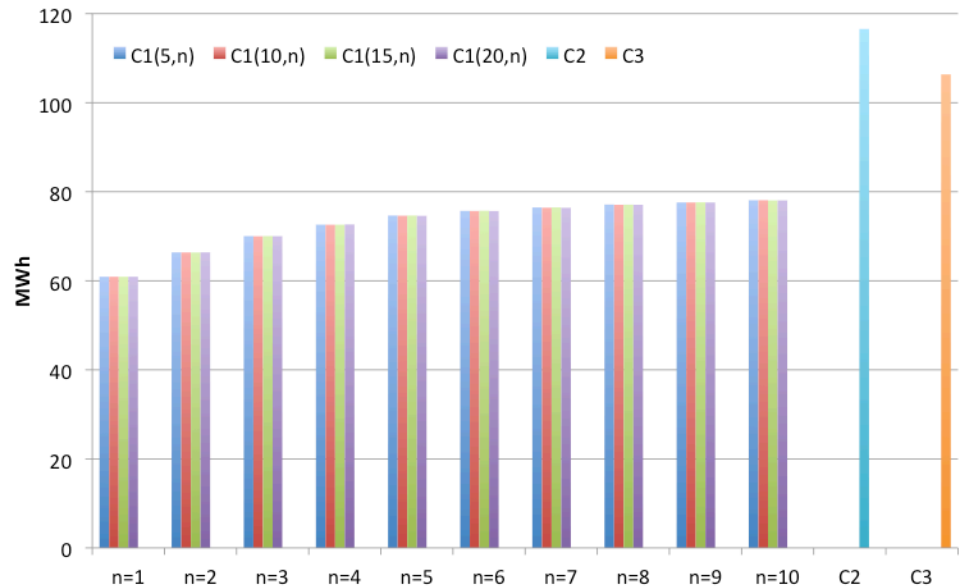




# Cluster Change Policy

Digital Institute

- C1: Dedicated computers for 'miscreant' jobs
  - Run these jobs on computers where they can't be evicted
- C2: High-throughput jobs defer nightly reboots
- C3: High-throughput jobs use computers at the same time as interactive users





Digital Institute

---

# Overview

- Motivation and Background
- Simulating Condor
- Power Management Evaluation
- **Conclusion**



# Conclusion

---

Digital Institute

- We can save energy (with minimal user impact)
  - P4 is the most optimal policy
  - S3 – greater impact on overhead
  - S2 – greater impact on power consumption
    - These could be merged
  - M2 can kill off lots of good jobs
    - Fix this by using C1
  - Benefits of C2 and C3 lost due to number of miscreant jobs
    - Need a better way to identify these
  - Policies are not mutually exclusive
    - could save ~70MWh (~60% of current usage) without significant impact on high-throughput user
  - Powering down cluster saves the most energy
- Looking for other uses
  - Already simulated running jobs on Cloud
  - Do others have data we could use?



Digital Institute

---

# Questions?

[stephen.mcgough@ncl.ac.uk](mailto:stephen.mcgough@ncl.ac.uk)