

ADC Requirements & Recommendations for Sites

OSG All Hands
March 14, 2016

Mark Sosebee

March 14, 2016



UNIVERSITY OF
TEXAS
ARLINGTON

Outline

- ADC Recommendations/Guidelines
- AGIS
- Memory Requirements
- cgroups



ADC Recommendations for Sites

<https://twiki.cern.ch/twiki/bin/view/AtlasComputing/SitesSetupAndConfiguration>

Log In
AtlasComputing

ATLAS Collaboration

ATLAS TWiki
ATLAS Protected
ATLAS Computing
Public Results

Create a LeftBar for this page

Index
Changes
Notifications

TWiki > AtlasComputing Web > AtlasComputing > AtlasDistributedComputing > ADCOperations > SitesSetupAndConfiguration (2016-01-25, AlessandraForti)

SitesSetupAndConfiguration

- Help and Support
- Recommendations and mandatory services
 - Storage
 - Computing Element
 - Choice of Batch system
 - Batch system shares and limits
 - Worker Node hardware resources
 - Worker Node logical configuration
 - Squids
 - Network
 - Recommended CPU, Storage and Network capacity
- Configuration within ADC
- Typical ATLAS jobs
- References to already existing information
- Opportunistic Resources

Help and Support

- The first entry point for sites is their cloud squad support: atlas-adc-cloud-XX at cern.ch .
- In case of urgent matters please contact the ATLAS Computing Run Coordinator through atlas-adc-crc AT cern.ch .
- For information that you believe is worth being discussed within the whole ATLAS distributed computing community use (don't abuse) atlas-project-adc-operations AT cern.ch .

Recommendations and mandatory services

Storage

- dcache is used at 80% of the T1s. dCache supports tape storage.
- DPM and dCache are good choices for use at T2s DPM is usually chosen by smaller T2s but there are some sizeable T2s (>2PB storage) that use it. While DPM is considered simpler, Dcache has more functionalities for data management.
- There are sites running other storage implementations like EOS, CASTOR, STORM and xroot
- More info about space tokens and protocols in [StorageSetUp](#)

Computing Element

- There are mainly 3 types of CEs, [ARC-CE](#), [CREAM-CE](#) and now [HTCondorCE](#)
- ARC-CE is now the one most commonly installed at new sites in Europe, HTCondorCE is typical in the US

March 14, 2016



UNIVERSITY OF
TEXAS
ARLINGTON

AGIS (I)

("ATLAS Grid Information System")

atlas-agis.cern.ch/agis/

ATLAS Grid Information System

[ATLASSite](#) [DDMEndpoint](#) [PANDA Queue](#) [Service](#) [Central Services](#) [DDM Groups](#)

[Docs](#) [TWiki](#) [OLD](#)

TOPOLOGY MANAGEMENT	SERVICE MANAGEMENT	OPERATIONS	DOCUMENTATION
<ul style="list-style-type: none"> Define RC site Define Experiment site Define DDM endpoint Define RSE endpoint (new implementation) Define PANDA site Define PANDA queue RC pledges Find DDM endpoints links Find TransferMatrix links 	<ul style="list-style-type: none"> Define OS service Define LFC service Define SE service Define CE service Define Redirector service Define PerfSonar service Define Frontier service Define Squid service Define Central service 	<ul style="list-style-type: none"> Crons list ADMINS list Changes log Request ADMIN privileges 	<ul style="list-style-type: none"> Main TWiki TWiki WEBUI instructions API Docs
DOWNTIMES	TOACACHE EXPORT	COMPARISON & VALIDATION TOOLS	
<ul style="list-style-type: none"> Downtime calendar DDM Blacklisting data PANDA Blacklisting data 	<ul style="list-style-type: none"> dynamic ToACache (changes are immediately propagated): http://atlas-agis-api.cern.ch/request/toacache/TiersOfATLASCache.py static ToACache: http://atlas-agis-api.cern.ch/ToACache/TiersOfATLASCache.py previous caches: http://atlas-agis-api.cern.ch/ToACache/cache/ View/Modify ToACache ExtraData (RSE integration) <ul style="list-style-type: none"> ToACache with Extra data 	<ul style="list-style-type: none"> Consistency checker ToAComparator AGIS-BDII CE comparison AGIS-Schedconf-PF mon CE comparison AGIS-OIMGOCDB sites+services comparison AGIS-PANDA PandaResource+SWReleases comparison AGIS-Schedconfig (topology) comparison AGIS-Schedconfig JSON comparison AGIS-GSR services comparison 	

March 14, 2016



UNIVERSITY OF
TEXAS
ARLINGTON

AGIS (II)

(“ATLAS Grid Information System”)

atlas-agis.cern.ch/agis/

- Of recent interest: setting values for **minrss**, **maxrss**
- **maxrss**: default value should be the value of RAM per single core
- This value can then be “overridden” per physical queue in AGIS
- Used by PanDA brokerage to match jobs/tasks to sites

Memory Requirements (I)

- One version of the ATLAS statement:
 - “At least 2 GB of RAM, but 3-4 GB would be beneficial”
- Most sites are probably a mix of hardware
- Can setup “himem” PanDA queues to support jobs/tasks with higher memory requirements
 - Allows PanDA to match jobs that require more memory efficiently

Memory Requirements (II)

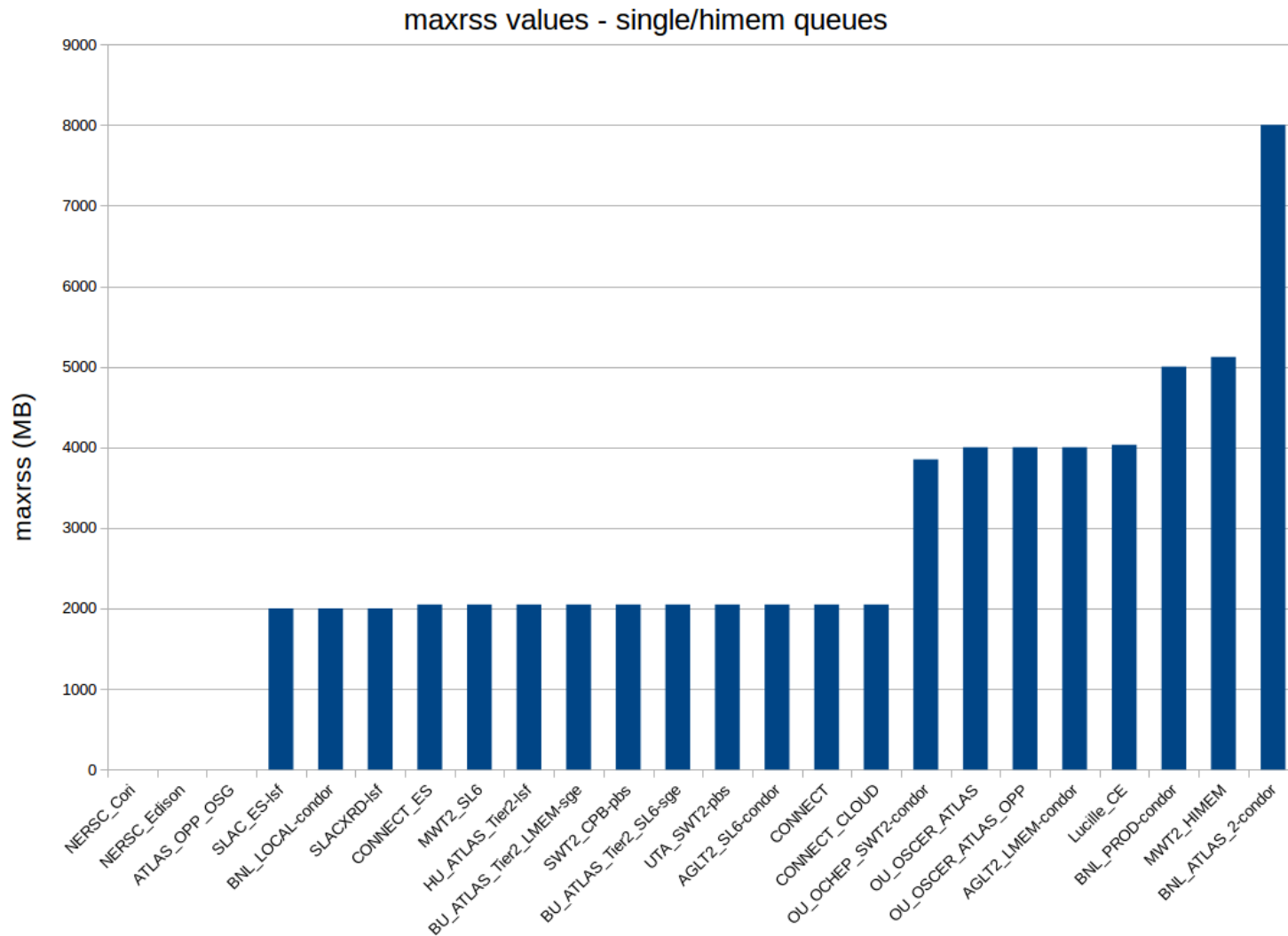
- Currently these PanDA queues are enabled:

AGLT2_LMEM, BNL_PROD_MCOREHIMEM*,
BU_ATLAS_Tier2_LMEM, MWT2_HIMEM,
MWT2_HIMEM_MCORE, SLACXRD_LMEM

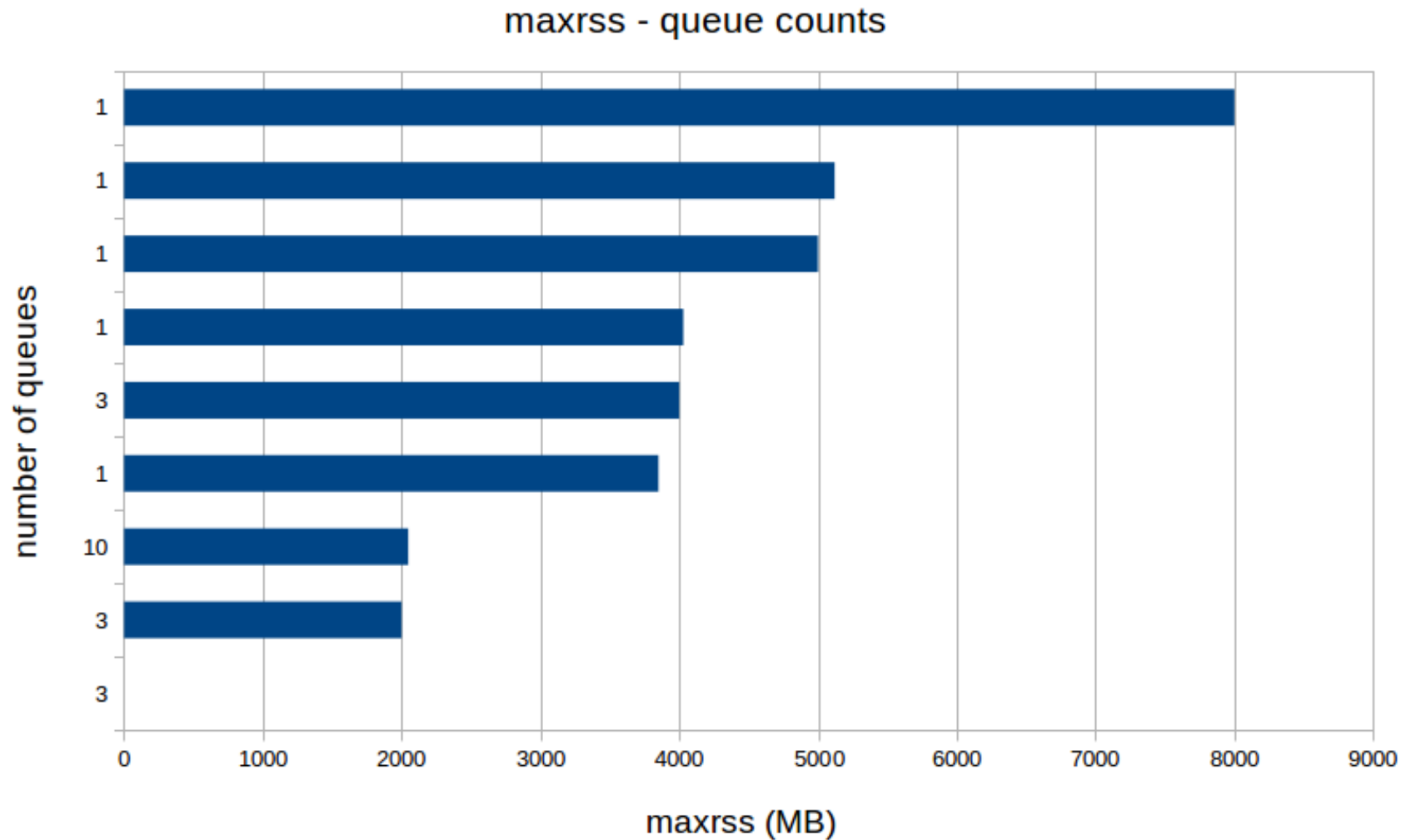
(*generally set to 'brokeroff' – jobs assigned by GDM)

- maxrss, minrss varies among the sites

maxrss – single core/himem



maxrss – single core/himem

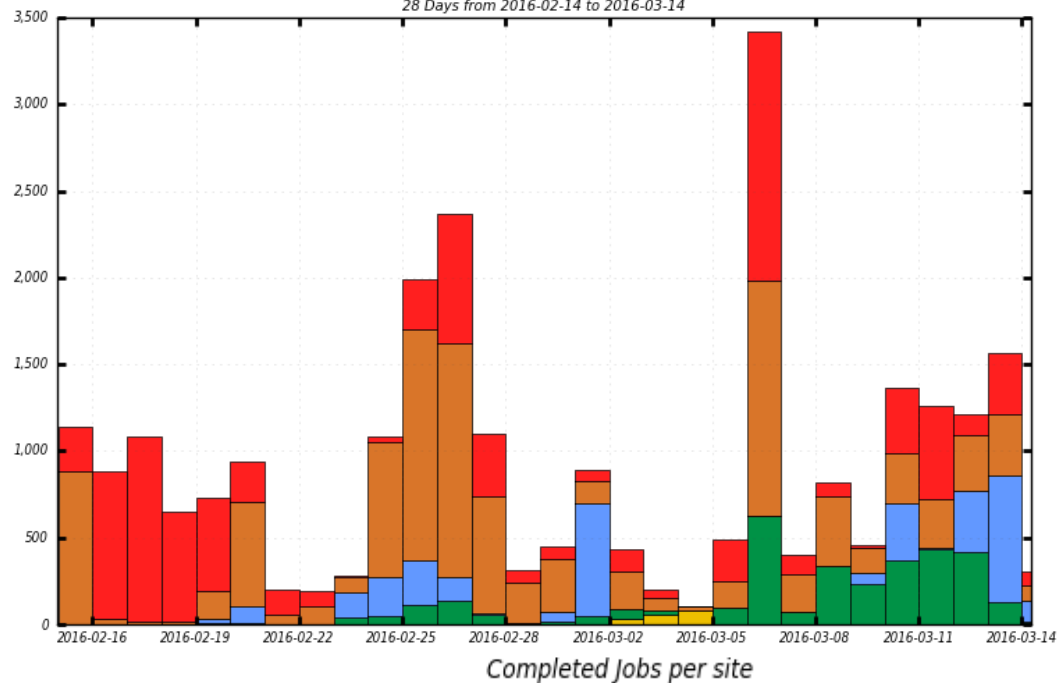


himem usage

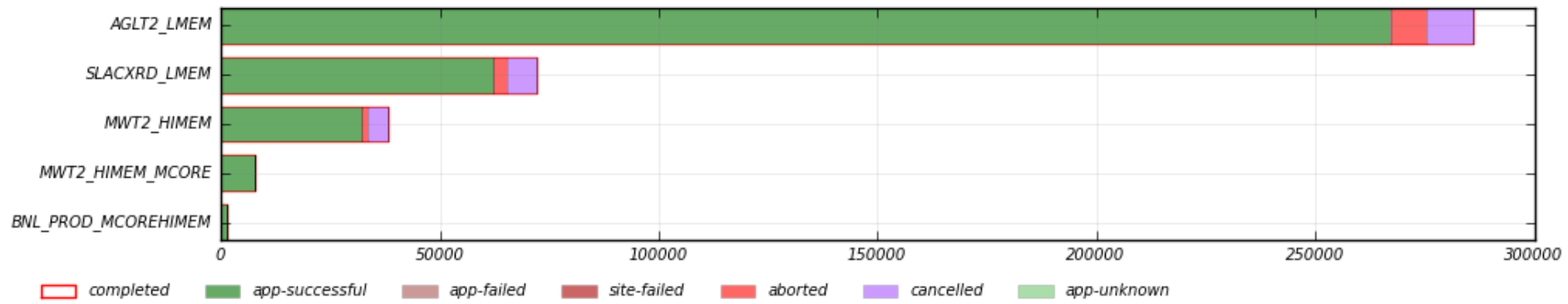


Slots of Running Jobs

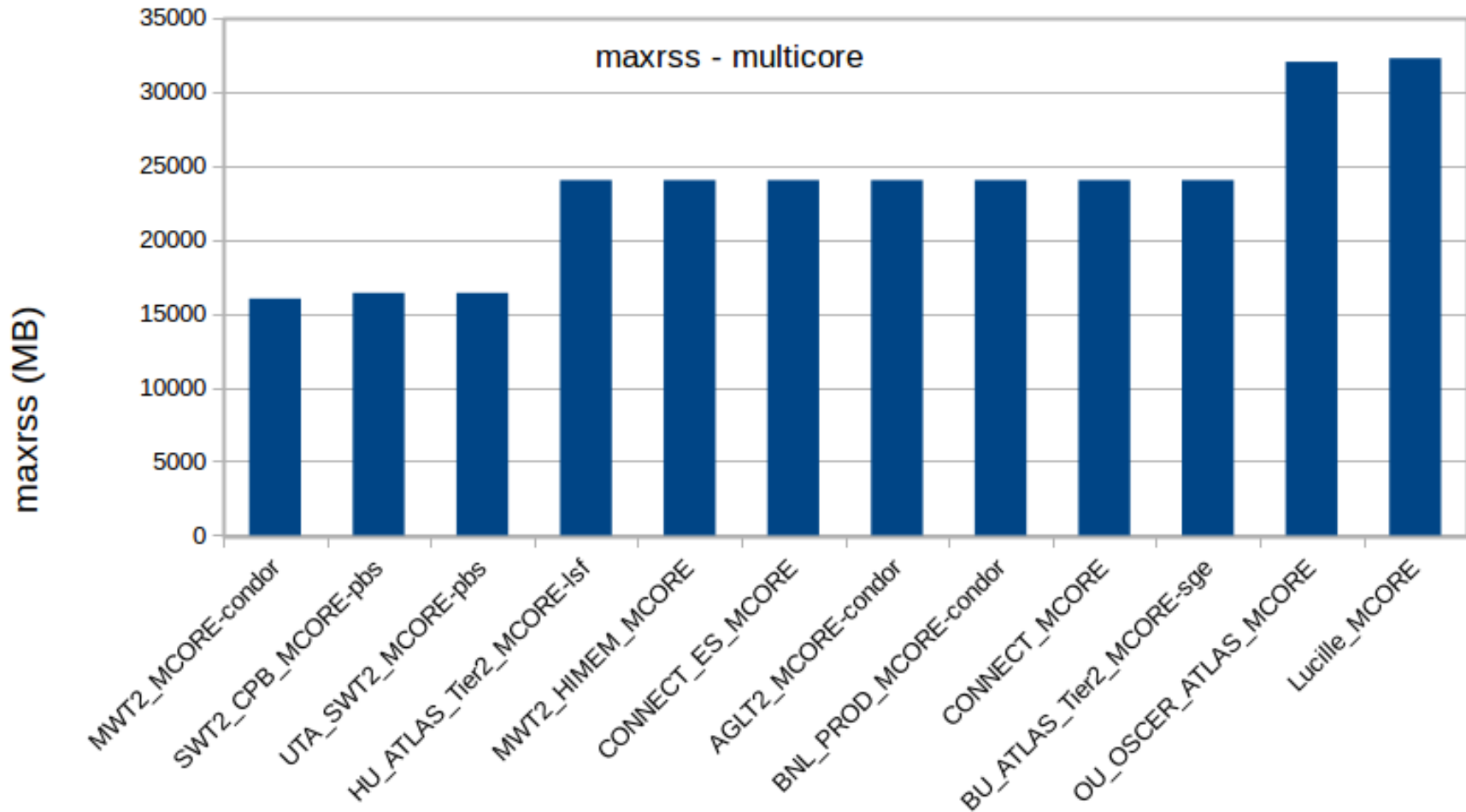
28 Days from 2016-02-14 to 2016-03-14



Completed Jobs per site



maxrss – multicore



Linux cgroups (I)

("control groups")

- Kernel-level control over resource usage by processes
- Recommended by ATLAS for site configuration
- Most batch systems have support for cgroups (Torque/Maui?)

Linux cgroups (II)

("control groups")

- US sites already using cgroups:
 - AGLT2 (deployed 12/14)
 - MWT2 (deployed 2/15 - **notes**)
 - SLAC (**notes**)
 - SWT2 (OSCER)
- Not yet:
 - SWT2 (UTA, OCHEP), NET2 (**notes**)



Sites Jamboree/ Software & Computing Week

More info available in these talks/meetings:

➤ Jamboree:

- [Alessandra Forti's talk](#)
- [Rob's talk \(US report\)](#)
- [Shifts & Communication w/sites](#)

➤ S&C Week:

- [Jamboree summary](#)

