

# Tier-1 centre at NRC «Kurchatov Institute» between LHC Run2 and Run3

Eygene Ryabinkin

Alexander Rogovskiy

Ilya Lyalin

**Igor Tkachenko**

# Evolution of hardware part of Tier-1 site at Kurchatov Institute

- 2013 year – prototype at 10% computing power, about 80 nodes;
- 2015 year – all hardware is ready, about 400 nodes;
- 2015-2018 – upgrading computing and storage facilities, more than 500 nodes.

# How does Tier-1 cluster management system work

- Uses Puppet for cluster management;
- Syncs puppet modules between master node and all nodes in the cluster;
- Applies local modules via **puppet apply**;
- Uses **pdsh** for mass modules applying at cluster nodes.
- Every site administrator have own puppet-environment.

# Problems with original cluster management system

- We have to keep modules in synchronized state: we need special module for modules synchronization and we must not forget to call it before every configuration change at node by updated module
- It is very difficult to notice errors during mass module applying
- High chance to miss node (that was switched off for any reason, for example) and it is very difficult to notice it
- High load of master node during applying modules at all nodes

# How **puppet agent** may solve problems of **puppet apply**

- **apply**: have to keep modules in synchronized state (we need special module for modules synchronization and we must not forget to call it before every configuration change at node by updated module)
- **agent**: does not need modules sync at all. We may still use site administrator environment for tests.

# How **puppet agent** may solve problems of **puppet apply**

- **apply**: it is very difficult to notice errors during mass module applying
- **agent**: puppet agent is writing to log-file results of all operations. We may analyze log by monitoring tools.



# How **puppet agent** may solve problems of **puppet apply**

- **apply**: high chance to miss node (that was switched off for any reason, for example) and it is very difficult to notice it
- **agent**: automatically apply modules at nodes

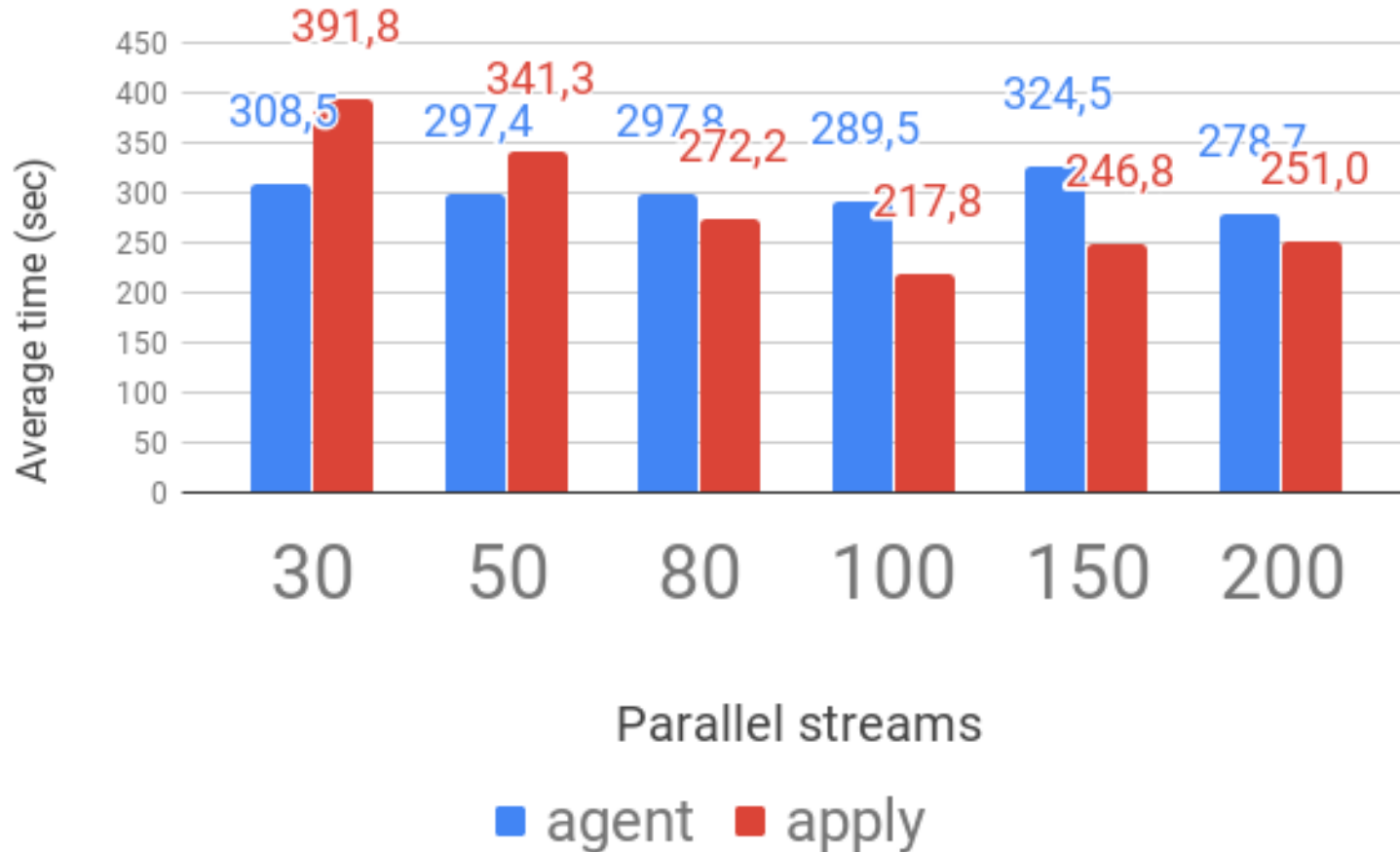


```
lhcb-sdpd11: Notice: /Stage[main]/Sync/File[/etc/puppetlabs/puppet/environments/tia/modules/dcache/manifests/dumps.pp]/ensure: defined content as '{md5}b6c2981bc15b130028fad156e5f1ebdd'
lhcb-sdpd11: Notice: /Stage[main]/Sync/File[/etc/puppetlabs/puppet/environments/tia/modules/dcache/manifests/init.pp]/ensure: defined content as '{md5}12a50930289fa4ad209db91c16b72252'
lhcb-sdpd24: Notice: /Stage[main]/Sync/File[/etc/puppetlabs/puppet/environments/tia/modules/dcache/manifests/config/head.pp]/ensure: defined content as '{md5}b002e3922a6475a750e39f3422b919df'
lhcb-sdpd19: Notice: /Stage[main]/Sync/File[/etc/puppetlabs/puppet/environments/tia/modules/bonding/templates/ifcfg-eth.erb]/ensure: defined content as '{md5}b9cee036e16e46a305da5b505e86f55e'
lhcb-sdpd19: Notice: /Stage[main]/Sync/File[/etc/puppetlabs/puppet/environments/tia/modules/carp]/ensure: created
lhcb-sdpd9: Notice: /Stage[main]/Sync/File[/etc/puppetlabs/puppet/environments/tia/modules/dcache/manifests/install_pgsql_la.pp]/ensure: defined content as '{md5}22f0242ba8114b278f2b191c9f09afde'
lhcb-sdpd13: Notice: /Stage[main]/Sync/File[/etc/puppetlabs/puppet/environments/tia/modules/dcache/manifests/db/billing.pp]/ensure: defined content as '{md5}1cccb29f5e66dclc155cc92a1b29627f'
lhcb-sdpd19: Notice: /Stage[main]/Sync/File[/etc/puppetlabs/puppet/environments/tia/modules/carp/manifests]/ensure: created
lhcb-sdpd10: Notice: /Stage[main]/Sync/File[/etc/puppetlabs/puppet/environments/tia/modules/carp/manifests/install.pp]/ensure: defined content as '{md5}26daa4a37b494a913c04219b1d17ea37'
lhcb-sdpd13: Notice: /Stage[main]/Sync/File[/etc/puppetlabs/puppet/environments/tia/modules/dcache/manifests/db/checking.pp]/ensure: defined content as '{md5}d2256cd77cfe39babb03e57afea643a2'
lhcb-sdpd9: Notice: /Stage[main]/Sync/File[/etc/puppetlabs/puppet/environments/tia/modules/dcache/manifests/logs.pp]/ensure: defined content as '{md5}7df8c730f8e2a04467e8e7b4fb62d853'
lhcb-sdpd9: Notice: /Stage[main]/Sync/File[/etc/puppetlabs/puppet/environments/tia/modules/dcache/manifests/pool]/ensure: created
lhcb-sdpd19: Notice: /Stage[main]/Sync/File[/etc/puppetlabs/puppet/environments/tia/modules/carp/manifests/config.pp]/ensure: defined content as '{md5}a9547617ddeefcc6521e05ff262cf332'
lhcb-sdpd10: Notice: /Stage[main]/Sync/File[/etc/puppetlabs/puppet/environments/tia/modules/carp/manifests/install_rpm.pp]/ensure: defined content as '{md5}845aafb555d53e892657cc4ef460ae3a'
lhcb-sdpd8: Notice: /Stage[main]/Sync/File[/etc/puppetlabs/puppet/environments/tia/modules/batchsys/manifests/create_users.pp]/ensure: defined content as '{md5}2525991f19fc494b2fb8302e8f93abd4'
lhcb-sdpd4: Notice: /Stage[main]/Sync/File[/etc/puppetlabs/puppet/environments/tia/modules/bonding/manifests/create.pp]/ensure: defined content as '{md5}1f71748ccf7857b40bea763567987afb'
lhcb-sdpd23: Notice: /Stage[main]/Sync/File[/etc/puppetlabs/puppet/environments/tia/modules/dcache/manifests/create_srm_dirs.pp]/ensure: defined content as '{md5}489ddf54aeb03f6cde5fd0d05162b1fb'
pdsh@tamer: ch1: ssh exited with exit code 255
lhcb-sdpd19: Notice: /Stage[main]/Sync/File[/etc/puppetlabs/puppet/environments/tia/modules/cvmfs/manifests/config.pp]/ensure: defined content as '{md5}2430eeb04240350d498fd19ebf8f6be2'
atlas-sdpd12: Notice: /Stage[main]/Sync/File[/etc/puppetlabs/puppet/environments/tia/modules/vcsrepo/spec/acceptance/beaker/git/compression/negative/compression_7_checkout.rb]/ensure: defined content as '{md5}8bcf4f06b5394d81887013eb6e6a0d12'
T1.SRVC012: Notice: /Stage[main]/Sync/File[/etc/puppetlabs/puppet/environments/tia/modules/yum/manifests/init.pp]/ensure: defined content as '{md5}62a1468d09163d098d70573b373080e2'
lhcb-sdpd19: Notice: /Stage[main]/Sync/File[/etc/puppetlabs/puppet/environments/tia/modules]/ensure: created
```

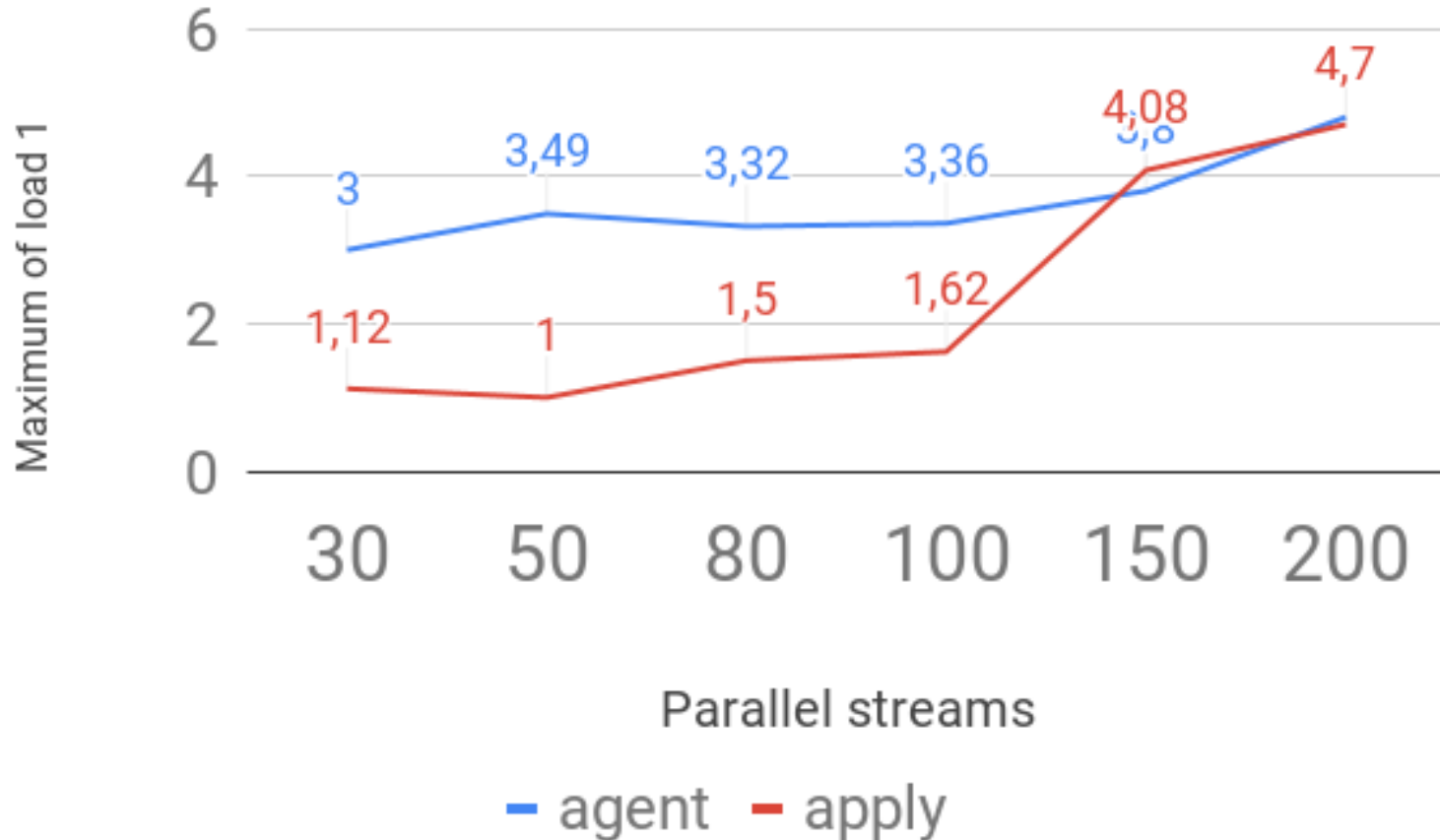
# How **puppet agent** may solve problems of **puppet apply**

- **apply**: high load of master node during applying modules at all nodes; nodes needs synchronization and module applying took a lot of time, despite of local catalog compilation
- **agent**: automatic modules applying is distributed in time, only compiled catalog goes to nodes.

# Compare the efficiency of sync+apply and agent for applying same module at all cluster nodes



# Compare the efficiency of sync+apply and agent for applying same module at all cluster nodes



# Potential problems of puppet agent and how we are going to solve them

- **Problem:** not all modules should be applied automatically:
  - We want more control for cluster critical services
  - Services that work with users should be managed in manual mode (*no one likes when storage pool with one's data suddenly restarts*)
- **Possible solution:** find groups of modules that are critical for cluster: for now it is modules that are related to storage system, job management system and WLCG software.
  - Every module is critical before audit: how it may harm critical cluster components described above. After this audit we may take a decision if this module is safe enough for automatic apply.

# Potential problems of **puppet agent** and how we are going to solve them

- **Problem:** all our modules were written for "one-by-one" mode. In agent-mode we have a lot of duplicate resource declaration and circular dependencies.
- **Solution:** audit of all our puppet code (about 47 modules and 8000 lines of code)

# Potential problems of **puppet agent** and how we are going to solve them

- **Problem:** who watches the watchmen? We have to be sure that agent works in the right way.
- **Possible solution:** use monitoring tools for looking after puppet activity at nodes.

# Potential problems of **puppet agent** and how we are going to solve them

- **Problem:** have to control changes in management code carefully to avoid massive applying of wrong settings at nodes.
- **Solution:** just use code review!



# Our experience of using Phabricator as a code review system

- Tier-1 at KIAE is using code review since prototype stage
- Initially, for code review we used mail list:
  - It is difficult to watch for change status (waiting for update, waiting for review, etc)
  - It is difficult to watch for review progress (what review requests are close to finish)
  - It is difficult to see what was changed since previous iteration
- Since the middle of 2016 we are using Phabricator
  - More than 400 changes was reviewed
  - More than 1500 comments
  - Added automatic codestyle checks
- Plans:
  - Unit-tests
  - Isolate puppet master from manual changes: automatically update puppet-master area after commit to version control system

Thank you for your attention!