Evolution of SAM in an enhanced model for monitoring the WLCG grid

David Collados

James Casey, Emir Imamagic, John Shade, Konstantin Skaburskas, Steve Traylen

Overview

- Why modify the architecture?
- What do we have today?
- What are the limitations & proposed solutions?
- What have we implemented already?
- What lies ahead?
- Conclusion
- Links

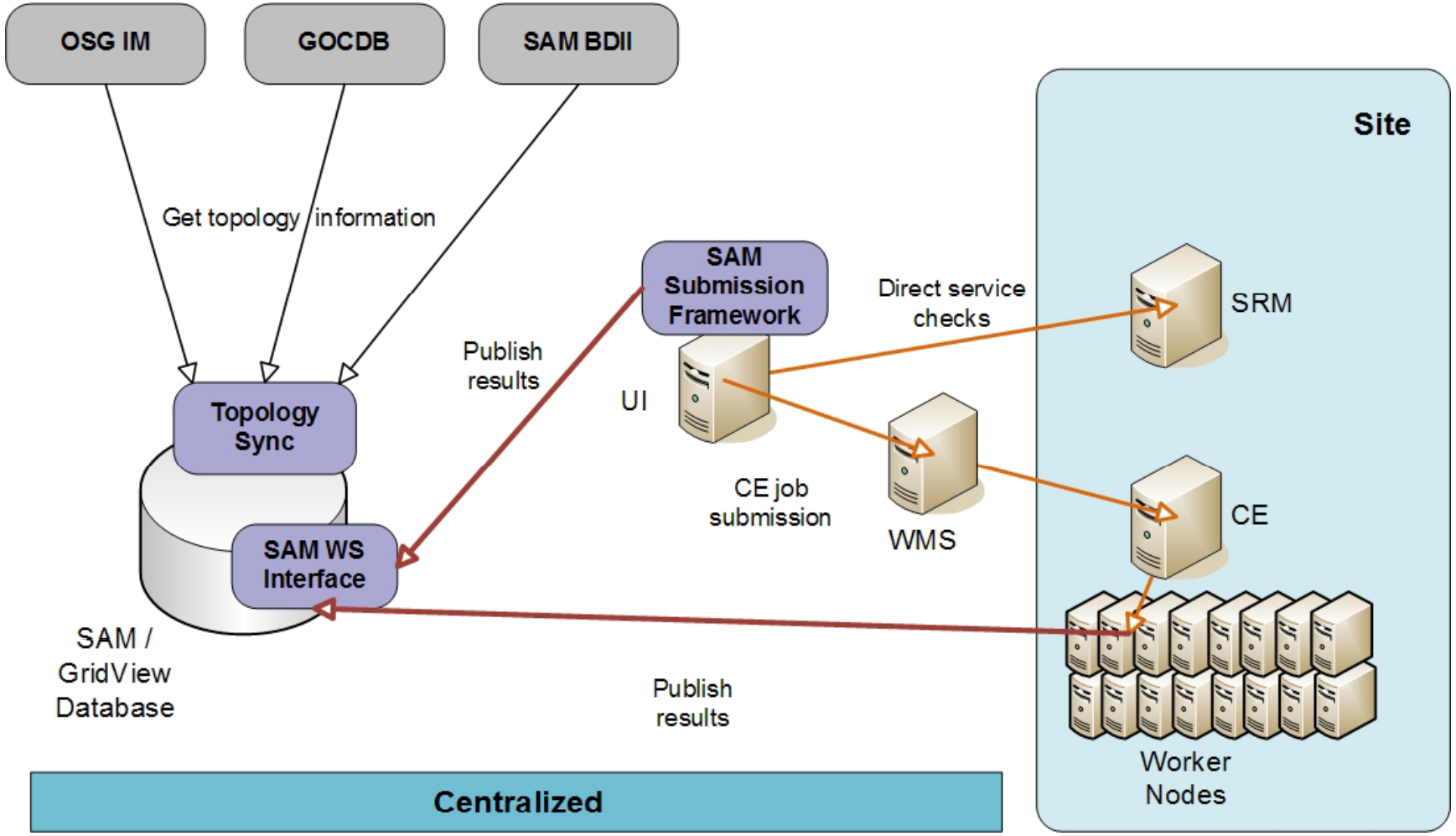
Why modify the architecture?

- **Continuation of monitoring strategy**
  - see James Casey talk @ last CHEP
- **Goal: improve reliability of the WLCG grid**
  - while providing current & historical views of the infrastructure to site admins, service managers, VOs, project managers
- **Monitoring from regions and sites.**
  - reduce manpower needed centrally
- **Mitigate existing limitations**
  - while providing better tools to sites and regions

What do we have today?

- **A central SAM infrastructure**
  - that tests all grid services once an hour
- **A consistent set of tests**
  - under central control
- **OPS and LHC Experiments-specific tests**
- **One single algorithm for calculating availability**

What do we have today?



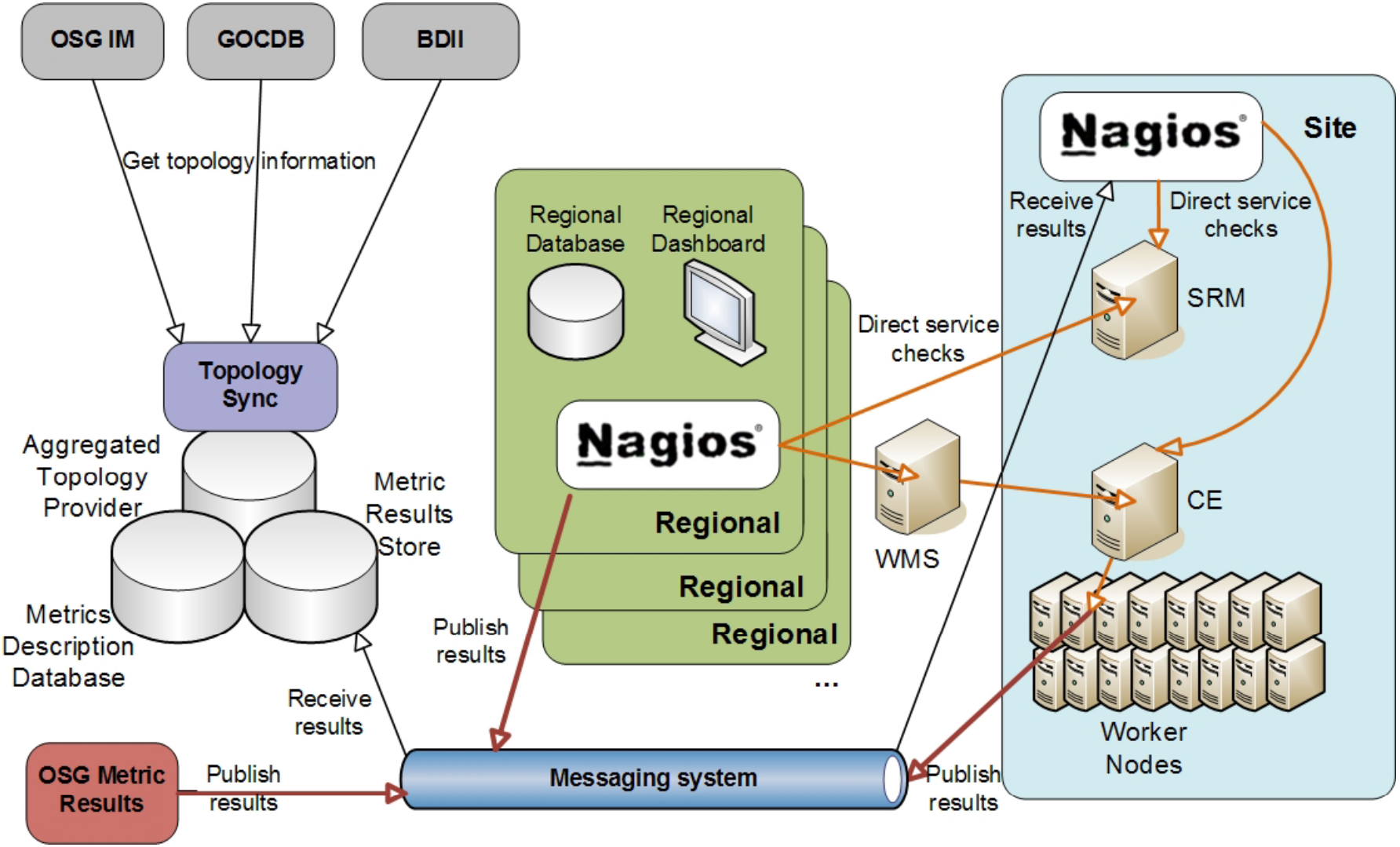
What are the limitations & proposed solutions?

- **Long time before site admins are alerted to problems**
  - Run tests locally, with local alarms
- **Sites are blind if central monitoring fails**
  - Provide sites with local monitoring tools
- **Possible scaling issues if number of services increases significantly**
  - Introduce regional monitoring instances

What are the limitations & proposed solutions?

- **During SAM outages, test results are lost**
  - use store & forward messaging technology
- **No history of grid topology**
  - new component to aggregate and cache topology information
- **Non-flexible availability calculations**
  - project-level component that contains metadata about metrics and how those metrics can be combined to calculate different availabilities, blacklisting, etc.
  - maintain history of which metrics and calculations were valid at which point in time

Enhanced architecture



What have we implemented already?

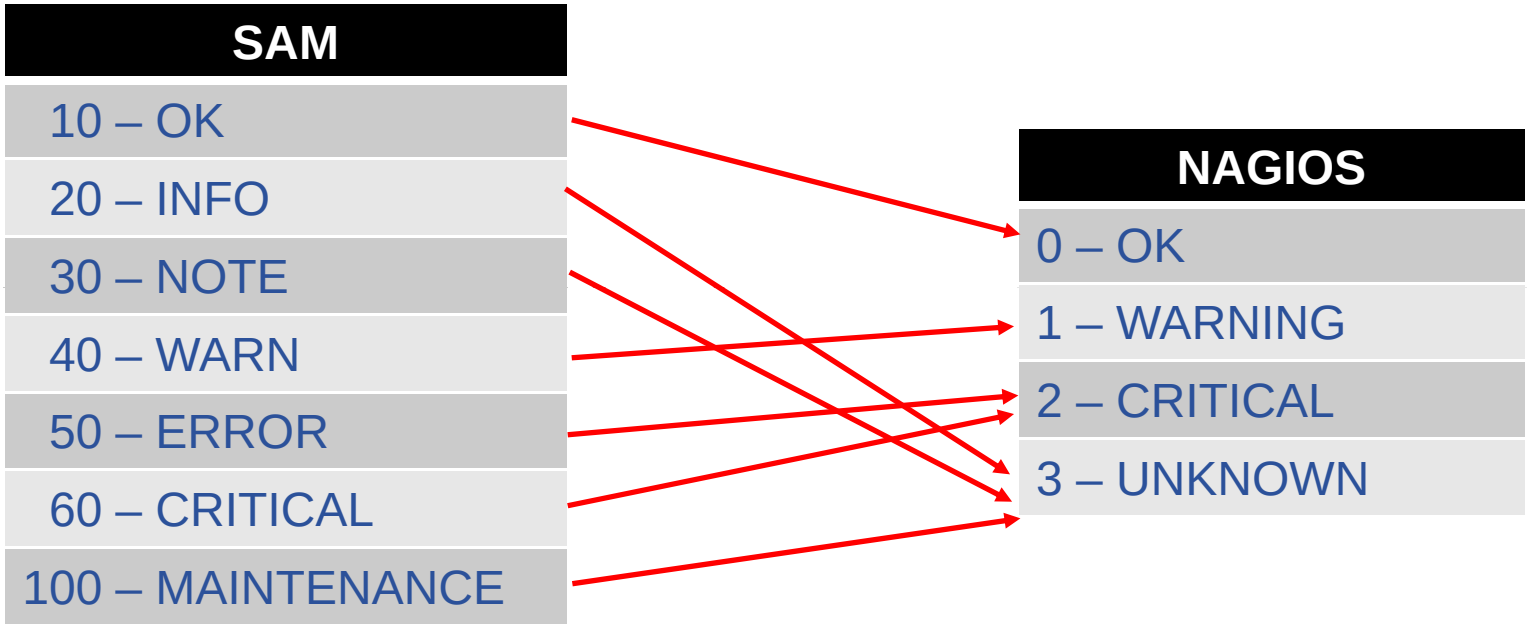
- **We've adopted Nagios for fabric and grid monitoring**
  - YAIM-installable Nagios package
    - Autoconfigures (NCG)
    - Standard set of probes provided (CE, SRMv2 & BDII)
  - Pnp4nagios (visualization) and NDOUtils (local DB) added to the Nagios package for sites and regions
  - SAM tests available through site Nagios (passive checks)
    - Site admins notified with alarms
- **Test-bed of 11 Regional Nagios servers at CERN**
  - Feeding a new instance of current SAM DB (already there)
- **Nagios deployed in 40 boxes monitoring 150 sites**

What have we implemented already?

- **We're using Active-MQ messaging technology**
  - Common components provided to enable end-to-end publication and consumption of results
    - Regional Nagios re-publishes test results on Message Bus
- **SAM Database exposed to Nagios for use by NCG**
  - (VO mappings, BDII and GOCDB)
- **Downtimes & user roles from GOCDB fed to Nagios instances**

What have we implemented already?

- **Mapping of SAM & Nagios status codes**



- **Wrapper executing SAM tests as Nagios checks**

  https://cern.ch/twiki/bin/view/LCG/SAMToNagios

What lies ahead?

- **By April 2009:**
  - ROC level Nagios based monitoring available
    - configured from Metric Description DB and ATP
    - 'SAM Portal' level of visualization complete
  - Full Nagios testing of all resources in grid running
    - At CERN – Central system, simulating 11 ROCs
    - Used to validate equivalence to SAM
    - Availability calculation using current algorithm but with new metrics

- **By July 2009:**
  - Feeding a new central metric store
  - Central metric store result visualization (SAM Portal/gridview)
  - Availability calculations using Metric Store and Metric Description DB

Conclusion

- Multi-level monitoring strategy to improve reliability
- Provides a solution that scales from the site-level upwards, giving the means for sites and ROCs to better monitor their services
- By using commodity software and interfaces we avoid possible issues of maintainability and ownership
- Original architecture realigned to encompass organizational changes

Links

- **OAT web page**
  https://twiki.cern.ch/twiki/bin/view/EGEE/OAT_EGEE_III

- **OAT Multi-level monitoring architecture**
  https://twiki.cern.ch/twiki/bin/view/EGEE/MultiLevelMonitoringOverview

- **OAT Milestones**
  https://twiki.cern.ch/twiki/bin/view/EGEE/MultiLevelMonitoringMilestones

- **Operations Automation Strategy**
  https://edms.cern.ch/document/927171

Thank You!

Questions?