

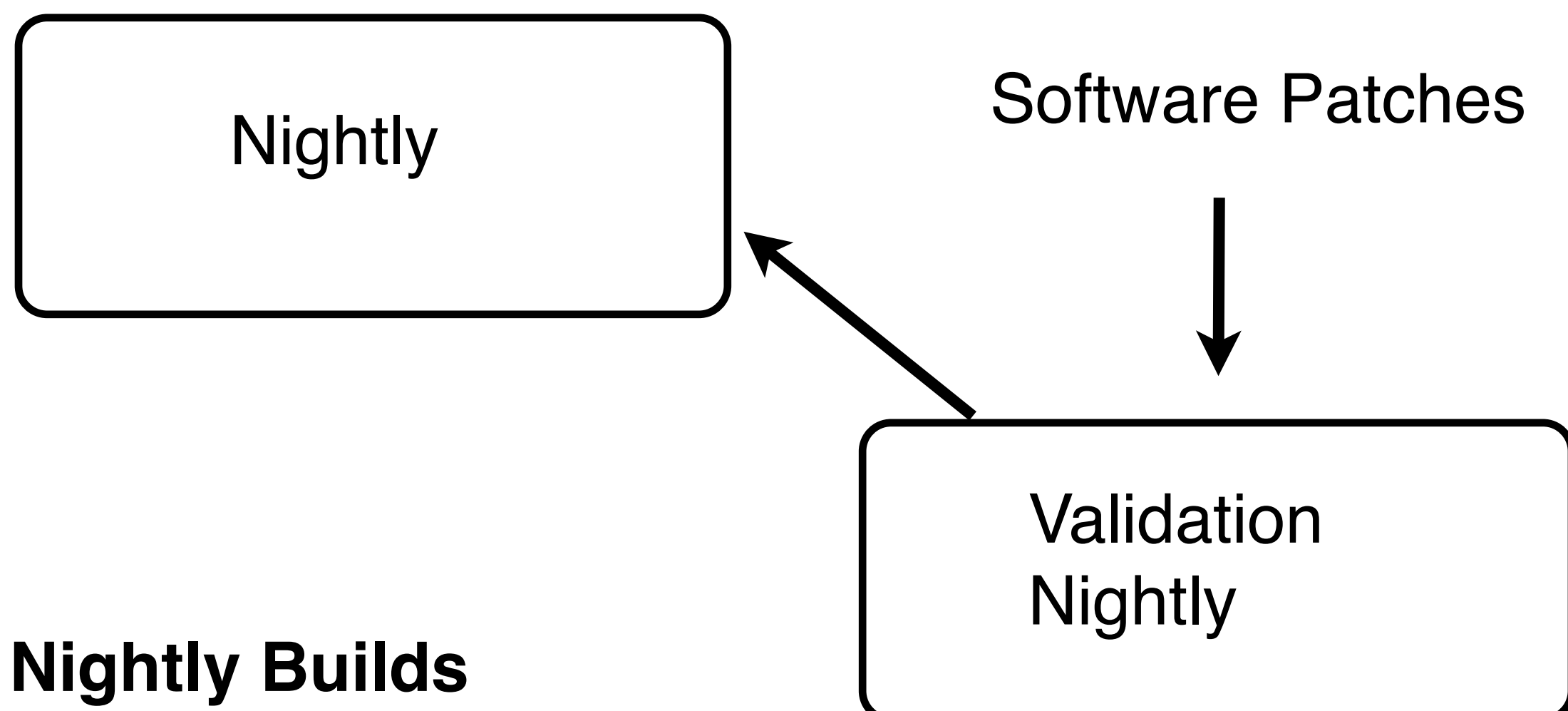
Software Validation in ATLAS

M. Hodgkinson, R. Seuster, B. Simmons, D. Rousseau, P. Sherwood on behalf of the ATLAS

Collaboration

Introduction

The ATLAS experiment processes continuously during data taking the collected data at its Tier0 center using the ATLAS Offline Software; furthermore much of the time is spent producing Monte Carlo simulations of many physics processes in the ATLAS detector. In order to maintain a high quality of software, to enable smooth running of the Tier0 center and timely production of Monte Carlo, a number of strategies are deployed within ATLAS. The software is run in many different configurations on nightly builds to identify problems - these tests are monitored by teams of rotating shifters, who also monitor compilation of the software and follow up on bug reports to ensure progress is made deploying required patches to the software.

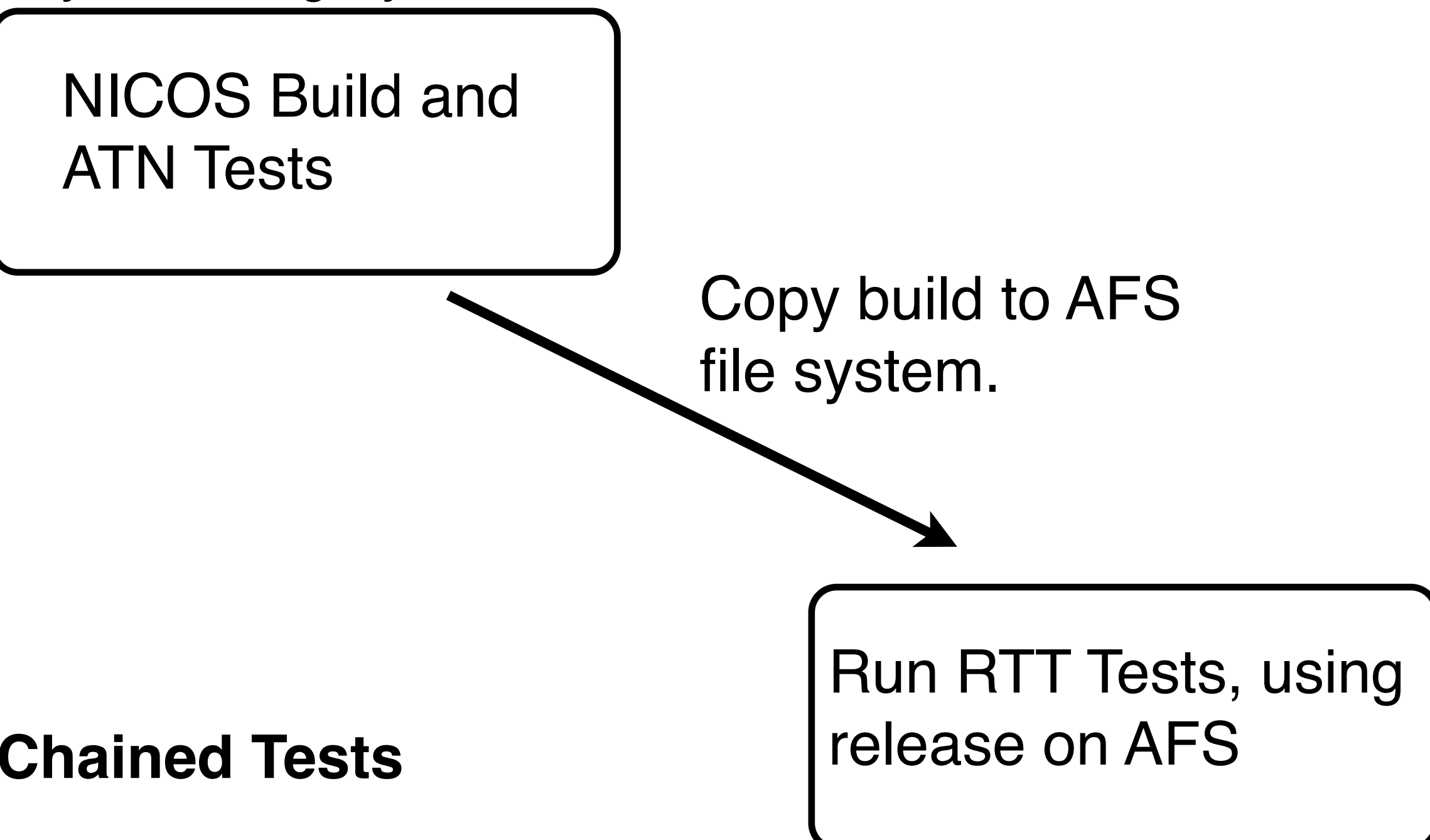


Nightly Builds

New software patches are first added to a validation nightly build. Only in the case the known problem is seen to be fixed, or the new feature is deemed to not cause new problems, can the patch migrate to the regular nightly build. This allows developers to always have a stable build each day to work with. The nightly builds are numerous - the main ones being a development nightly build where all the latest features are added, a data taking build which corresponds to the release used at the Tier0 center and a Monte Carlo build which corresponds to the release used to run the Monte Carlo simulations. Each of these have both a validation and a regular version. Furthermore each build is built with and without debug symbols and in both 32 and 64 bit flavours.

Nightly Tests

Each day a set of tests are run on each nightly build (built by the NICOS¹ system) to assess new problems and changes in output seen. Two software frameworks are used to run these tests: the AtNight (ATN¹) and RunTimeTester (RTT²) frameworks. The ATN framework allows us to run short tests (of order 5 minutes) on the worker nodes where releases are built. The RTT framework runs much longer tests on a dedicated batch farm hosted at CERN - these tests include running of the reconstruction software in numerous configurations and dedicated tests to monitor the memory usage of reconstruction software. Furthermore the FullChainTests (FCT) and Tier0ChainTests (TCT) which emulate the exact sequence of software jobs used for Monte Carlo and data processing respectively are run each day. The RTT allows tests to run for a maximum of 24 hours - after 24 hours all running and pending jobs are cancelled, and the RTT starts a new run for the next days set of nightly builds.



Chained Tests

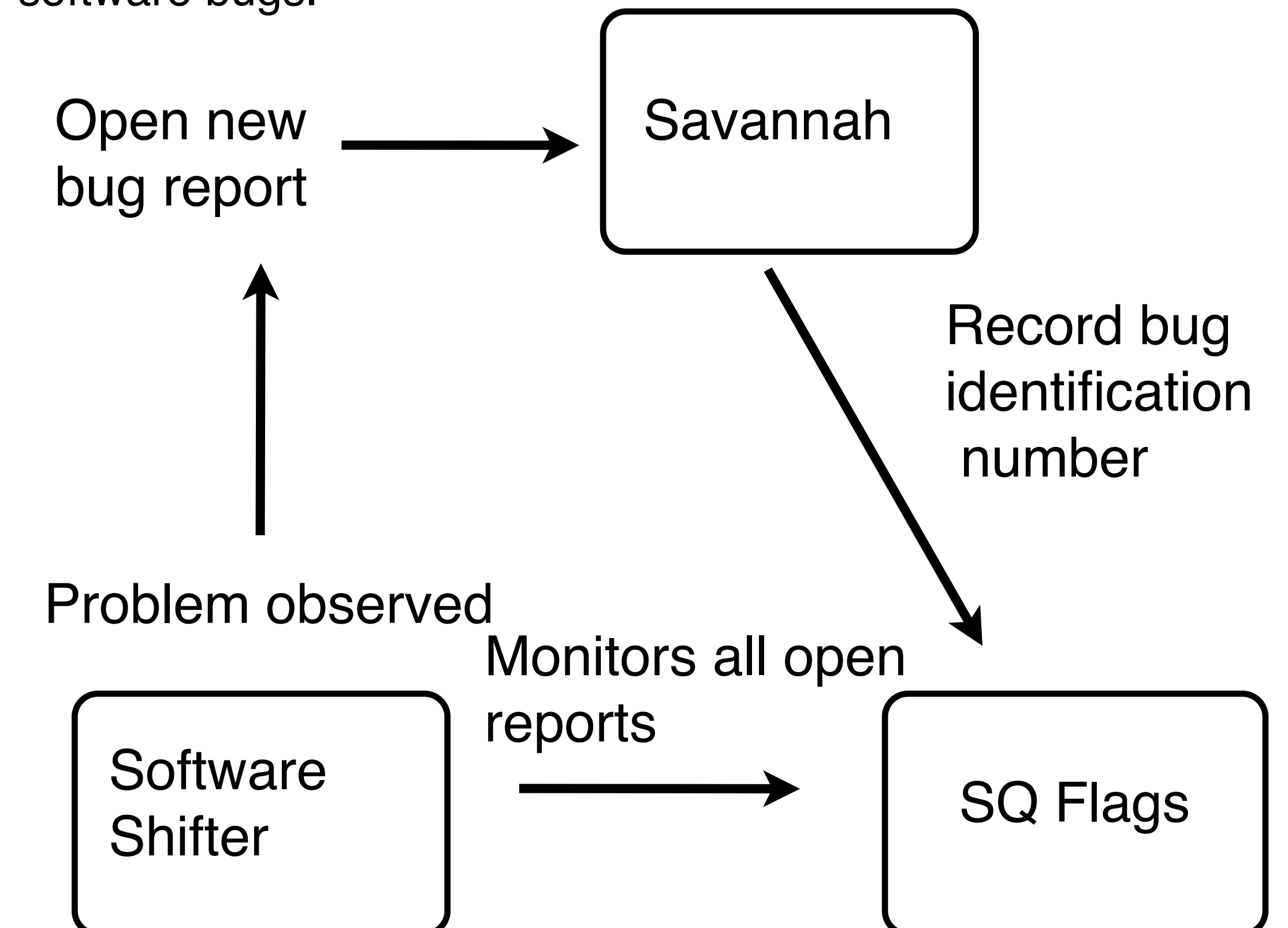
Both the FCT and TCT tests run using the RTT framework. These chain tests typically run on larger event samples than other RTT tests, allowing to not only validate the standard software sets of jobs used at the Tier0 centre and for Monte Carlo production, but to spot rarer crashes in software. Once a given nightly is built into a fixed release and deployed to Grid sites a larger set of tests, the Big Chain Tests (BCT), can be run using the grid. Far larger event samples can be processed to spot even rarer bugs and give a first indication of the quality of the physics output of the software. This is followed up by the Physics Validation³ team who submit a set of physics samples representative of typical physics production jobs (called Sample A) to study the physics data output - for example checking a suite of histograms relevant to physics studies in a particular area against reference histograms.

1. <http://www.usatlas.bnl.gov/computing/software/nicos/>, A. Undrus
2. The ATLAS RunTimeTester Software, B. Simmons et al, 2010 J. Phys.: Conf. Ser. 219 042023
3. Physics and Software Validation in ATLAS, D. Costanzo et al, 2010 J. Phys.: Conf. Ser. 219 032064
4. <http://savannah.cern.ch>
5. <http://ami.in2p3.fr>

Software Shifts

Two main sets of software shifts are undertaken - shifts to monitor the simulation and digitization software, and shifts to monitor the reconstruction software. In the event of a software package failing to compile or generating a compiler warning NICOS sends an automated warning to software developers. The shifters follow up with the developers and nightly release coordinators to ensure a new svn tag of the said package is produced in time for the next days nightly. Furthermore if a new problem is seen in an ATN or RTT tests a bug report is submitted. Existing bug reports are monitored for progress by the shifters, with reminders added if none is forthcoming.

In addition to these sets of shifts ATLAS has specialised shifters monitoring for example dedicated trigger or muon software tests. Finally an expert reconstruction shifter has recently been introduced whose job it is to improve the information in bug reports if necessary, verify the recipe posted to produce the bug works and give advice or active help to solving reconstruction software bugs.



Bug Tracking

Bug reports may be submitted by software shifters, Tier0 shifters (for problems seen in data processing jobs at the Tier0), Grid shifters (for problems seen in Monte Carlo jobs running on the Grid for example) or interested third-parties. Bug reports are submitted into the Savannah⁴ bug tracking system - bugs are categorised according to software area so that one can see all open bugs in the offline reconstruction software for example. An orthogonal system, called Software Quality (SQ) Flags has been developed by the TagCollector⁵ software team, which allows to list all bugs for a given release on one page (with links to the bug in savannah). Software shifters use this page to monitor bugs for all nightly releases they are responsible to monitor.

Shift Communications

At the end of each shift a detailed report is submitted into the ATLAS Electronic Log (ELOG) system for perusal by software experts. Furthermore short emails are sent to a list stating whether the nightly build for a given release is useable or not useable - this allows software developers to understand whether that days nightly is something they can work with, or if they should fall back to usage of a previous days nightly. Furthermore it alerts other shifters to any serious problems present to avoid duplication of effort.

Cleaning Weeks

Periodically we run software cleaning weeks whereby less critical problems are worked on to further improve the software quality. This has included in the recent past cleaning up old bug reports from savannah, looking at less frequently checked software tests for any problems not seen in the standard suite of tests and checking log files for superfluous output that could be removed.

Conclusions

The ATLAS collaboration builds a set of nightly releases each day, on which numerous ATN and RTT based tests are run. Shifters monitor both these tests and compilation statuses of a subset of the nightly releases. They then communicate the findings each day so that appropriate action can be taken. Once a nightly build is considered to have high quality software a fixed release can be built which can then be validated on the grid via the BCT and finally the SampleA production to be used by the Physics Validation teams.

