# Automated Testing

### What is automated testing?

Automated testing involves having mechanisms that can tell us if our system or proposed system is behaving the way we want. We normally also have a coordination mechanism that initiates the tests at the appropriate times.

For business intelligence infrastructure, we normally consider two main categories of tests:

| | Test Type | |
|---|---|---|
| | Unit | Integration/ Acceptance |
| Concerns | Granular aspects of functionality | Business rules being followed |
| Example test | Function A returns X when given input Z | Transform T removes duplicates |
| Environment | Limited as possible | Typically uses databases |
| Data | None or trivial | Can be substantial |
| Framework | xUnit style | DbFit/Fitnesse style |

Jenkins is the leading open source Continuous Integration server, and is a safe choice for coordinating the tests required for BI work. It integrates well with xUnit style frameworks such as Nose as well as with Fitnesse.



### Why should I use it?

Unit tests allow you to see that the fine grained functional operation of your infrastructure is correct. It's quicker and easier to specify checking this automatically once than to do so manually every time you make a change. In business intelligence infrastructure, there are typically some core components that are common to all operation, and it's essential to be sure these work correctly.

Integration/Acceptance tests check the processing logic being applied by the business intelligence system. For a simple example, there may be a rule that the number of raw records loaded in a particular batch must match the number of records reported for that batch in a fact table. If this doesn't happen, then records are being filtered out or duplicated. Being able to retest this functionality automatically allows us to detect problems more thoroughly, without using any extra effort.

The coordination mechanism adds significant value by coordinating the running of the tests and automatically reporting the results. This allows certain tests to be run on every push to the central source code repository, giving much more timely alerting than would otherwise be possible, and also allows tests to be run when people are not working, for example during the night. Deployments can also be made conditional on test results, allowing Continuous Deployment.

### Can you help us work out how to use it?

Yes - and we have extensive and specific experience in doing this in an enterprise data warehouse context, so we know a lot about what really does and does not work in practice.

Our approach to automated testing allows you to get additional benefits, for example, the framework we typically recommend for the integration/acceptance tests (DbFit/Fitnesse) saves the developer significant time when doing the initial development, because it can do the test setup and data load as well as running the process and inspecting the results. This means the developer doesn't have to do these processes manually, which improves the cycle time and consistency, and improves quality of code when transitioning to test.

We have worked out an architecture for testing which allows the assurance provided by automated testing to be extended to the production environment, without unnecessary duplication of effort. This control point based architecture simplifies the test definitions and means that acceptance test criteria can be assured in production as well as in development and test environments of various sizes. Because of the complexity of the interactions in a data warehouse, and the potential for bottlenecking testing on data load, this architecture allows real robust assurance in production while remaining scalable for test operation. This includes automatic inspections of expected results on production and automatic alarming and remedial actions.

### Can you help us implement and operate it?

Yes - we built all these things at a major communications provider for an Amazon Redshift data warehouse project. We can also advise and coach on the people and processes aspects.

### What extra value does Cloud BI Limited add?

We have done these things before with great success. We want to share the benefits of a devops influenced approach to data warehousing. We're easy to work with and want to see you succeed.

*To explore how we can help you further, please contact info@cloudbilimited.com*