SANDIA REPORT

SAND2008-5991
Unlimited Release
Printed March 2009

The Process for Integrating the NNSA Knowledge Base

Dorthe Carr, Jeff Hampton, Lisa Wilkening, Elaine Martinez, Christopher Young

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under Contract DE-AC04-94AL85000.

Approved for public release; further dissemination unlimited.

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

NOTICE: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, make any warranty, express or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government, any agency thereof, or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof, or any of their contractors.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: http://www.osti.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd.
Springfield, VA 22161

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.fedworld.gov
Online order: http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2008-5991
Unlimited Release
Printed March 2009

The Process for Integrating the NNSA Knowledge Base

Dorthe Carr, Elaine Martinez, Christopher Young
Next Generation Monitoring Systems Department

Lisa Wilkening
Information System Development Department

Sandia National Laboratories
P.O. Box 5800
Albuquerque, NM 87185-0401

Jeff Hampton
Lockheed Martin Mission Services
Orion Avionics and SW SEIT
Houston, TX 77058

ABSTRACT

From 2002 through 2006, the Ground Based Nuclear Explosion Monitoring Research & Engineering (GNEMRE) program at Sandia National Laboratories defined and modified a process for merging different types of integrated research products (IRPs) from various researchers into a cohesive, well-organized collection know as the NNSA Knowledge Base, to support operational treaty monitoring. This process includes defining the KB structure, systematically and logically aggregating IRPs into a complete set, and verifying and validating that the integrated Knowledge Base works as expected.

This page intentionally left blank

CONTENTS

Figures

This page intentionally left blank

Introduction

The National Nuclear Security Administration (NNSA) Ground-based Nuclear Explosion Monitoring Research & Engineering (GNEMRE) program provides research and development to improve the U.S. capability to detect, locate and identify nuclear explosions using ground-based technologies. A large part of the GNEMRE program is to provide products that will be utilized by the Air Force Technical Applications Center (AFTAC) in operation of the United States National Data Center (USNDC), which is responsible for fulfilling the United States' nuclear explosion treaty monitoring mission. Specifically, the GNEMRE program develops a suite of operational products, which include scientific data and the software tools to process and exploit the data.

These individual products developed by GNEMRE researchers at the national labs and other researchers funded by NNSA have great potential impact. However, to be operationally useful, they must be integrated into a comprehensive and cohesive package called the NNSA Knowledge Base (KB). To accomplish this, a process was developed by Sandia National Laboratories (SNL) to integrate products into the KB, validate that the products were usable, and deliver the KB to AFTAC. This document describes the Knowledge Base Integration Process developed and refined by SNL over a series of KB releases from 1998 through 2006.

History of the Knowledge Base

A timeline of deliveries for the Knowledge Base is shown in Figure 1. Through the series of KB releases, the general trend has been towards increasing complexity and formality with each delivery, reflecting the growing volume of information, as well as increased customer expectations for quality and robustness. The final process used for the KB7.0 release included a detailed integration, validation and delivery schedule and defined standards for documentation, data formats and testing methods. In the following sections, we review the major releases and highlight important changes made for each.

KB Version 2.0

The first delivery of the KB to AFTAC, KB2.0, was in 1998. Los Alamos National Laboratory (LANL) and Lawrence Livermore National Laboratory (LLNL) delivered collections of focused data sets for their regions of interest. These collections of data sets were called Information Products (IPs), a term that was used henceforth. Basic guidelines for formats for the various kinds of data were followed, but official formats had not yet been established. Reference event data, station parameter information and waveform information were delivered in database tables that followed the Center for Seismic Studies schema 3.0 (CSS3.0) developed by Anderson et al. (1990). Travel-time corrections (two-dimensional parametric surfaces) and surface-wave group velocities were delivered in flat files. All the contextual data was delivered in either Environmental System Research Institute (ESRI) proprietary or military-standard formats with ArcView based projects for data visualization and querying. Along with the data, documents were written by each of the labs, explaining how the data sets were put together and tested. These documents were called Information Product (IP) documents, and they followed a standard format.

The integration for KB2.0 was minimal; data sets were checked for consistency prior to being sent on to AFTAC. Checks included

- verifying that datasets were comprehensive with no missing data,
- verifying documentation (metadata) was available,
- verifying database tables followed the CSS3.0 schema,
- verifying that no null values existed in the database tables,
- verifying that the primary and unique keys could be set on the database tables
- verifying that flat files, text files and PDF files were readable

Once the datasets were checked, they were sent to AFTAC where AFTAC personnel installed the KB. The delivery and installation were successful, but use of the KB was limited, in large part because no index or software interfaces were provided to assist users with finding and making use of the content.

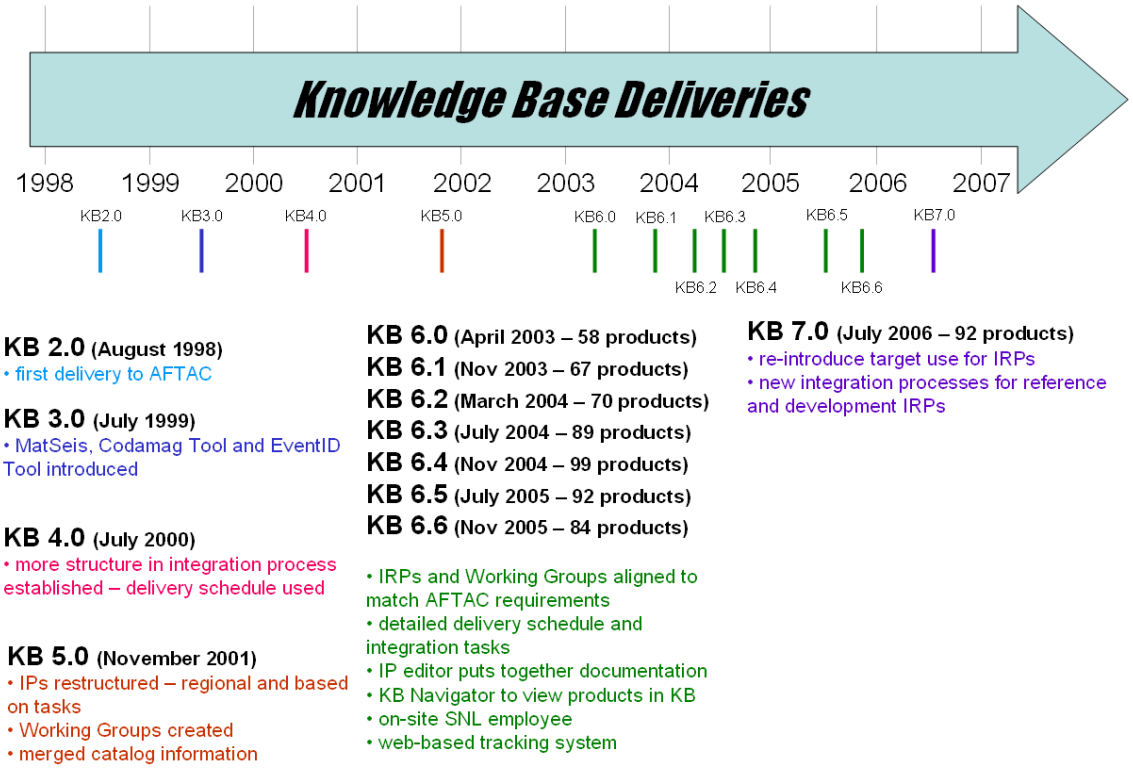


Figure 1. Knowledge Base Deliveries from 1998 through 2006.

KB Version 3.0

In 1999, KB3.0 was delivered to AFTAC. The types of data and formats were very similar to KB2.0, but the volume of data incorporated into the KB had increased. In addition the delivery included three custom-designed seismic regional analysis tools (MatSeis, Codamag Tool and EventID Tool) to help AFTAC make use of specific types of data content of high interest. The integration process at SNL was the same as was used for KB2.0. For KB3.0, personnel from SNL went to AFTAC to install the datasets, and test that they were properly installed.

KB Version 4.0

KB4.0 was delivered to AFTAC in July 2000. The major change in KB4.0 was that more structure was added to the integration and delivery process. While the actual data integration process was not changed from KB3.0, a formal release schedule was developed by SNL and the KB Coordinators at LANL and LLNL that focused on deadlines for delivery of products to SNL. A major benefit of the schedule was much improved communication between the SNL KB Coordinator and designated points of contact (KB Coordinators) at LANL and LLNL during the critical period spanning the integration of the data sets into the KB at SNL.

KB Version 5.0

A number of important changes were made for KB5.0, delivered in November 2001. Five Working Groups were created (each based on a specific capability needed in the KB) in order to tailor the KB to support AFTAC monitoring tasks. The Working Groups were composed of researchers from all the labs and AFTAC, and were responsible for developing IPs for specific capabilities. In addition, the Working Groups provided technical validation of data sets proposed by various researchers for inclusion in the IP. Along with the IPs from the various working groups, LANL and LLNL still delivered IPs based on their regions of interest.

Another major change in the integration process for KB5.0 was that for the first time the database information from LANL and LLNL was merged into one set of tables by the KB Integrator at SNL This meant that there was one set of database tables that contained all the supporting database information instead of having a separate set of database tables for each different region.

KB Version 6.0 through KB Version 6.6

In response to feedback received for KB5.0, major changes were implemented in the KB integration process. These changes affected the process of getting data into the KB, the organization of the content, and the way the content was presented to users at AFTAC. AFTAC and NNSA agreed to tie the KB to AFTAC requirements to try to focus development efforts on the most important needs. To reduce the time required to get useful products to AFTAC, the decision was made to deliver the KB more often than once per year. The redefinition of the process took 18 months to complete, and KB6.0 was delivered in April 2003. Six additional releases of KB Version 6 (KB6.1 through KB6.6) were delivered during the period of November 2003 through November 2005. With each new delivery, changes were made to the integration process to make the process work better for the KB Integrator, lab researchers and KB users at AFTAC.

The first step in the redefinition of the process for KB6.0 was to write a document that described it (Gallegos et al., 2003). This document explained the roles and responsibilities of the KB Integrator, Product Integrators and the Integration Board (IB). It also explained the process of how data sets, now defined as Integrated Research Products (IRPs), were created, validated, integrated and put into an IP. A new role of IP Editor was established to help with the growing volume of documentation associated with the KB. This document was updated as the integration process was modified and made more streamlined. The most recent version of the integration document was written in 2005 (Gallegos et. al., 2005).

To help with installation and use of KB releases, an SNL employee was placed on-site at AFTAC soon after the delivery of KB6.0. He was able to interact daily with AFTAC employees and understand their computing environment at a much more detailed level than had been possible before. This information changed how SNL delivered the various pieces of the KB, and how the KB was organized so that it could be more easily used by the users at AFTAC.

A more detailed integration schedule was also developed. This schedule listed the 40 tasks that needed to be accomplished in order for the IRPs to be created, verified, validated and integrated into the KB. To provide transparency into the integration process for the rest of the GNEMRE program, an on-line internet-based tool was developed to track the IRPs for a specific release through the integration process. Each IRP was assigned a product number based on a three-part numbering scheme: IP#.IRP#.Version#. Each IRP was entered into the on-line tracking tool by product number and tracked through the integration process. Thus, at any time during the integration process, the IB and Product Integrators were able to go to the on-line tool and see the progress of each IRP.

To address complaints about difficulty finding and using content in KB5.0, SNL developed the KB Navigator (Merchant, et al., 2004), a graphical interface into the KB that allows a user to find specific IRPs, check geographic extents, access standard GIS data sets, view metadata and launch the associated KB tools all through a simple, intuitive interface. The KB Navigator displays the content as a hierarchy of folders, showing each IRP (labeled with product number) listed under its IP by the product number. This made it much easier to make sure all IRPs were captured in the KB and to find a specific IRP.

KB Version 7.0

In January 2006, AFTAC provided an updated draft requirements document to NNSA. Since the KB is tied to the AFTAC requirements, this new draft provided an opportunity to reassess the KB integration process. The decision was made to reintroduce the concept of tying each IRP to an AFTAC delivery destination to ensure that an appropriate level of attention is given to each IRP. Four delivery destinations were defined – Ops-NDC, Ops-TT, Reference and Development – and each of them implied important differences in quality and completeness expectations which in turn implied differences in timelines for product delivery, integration level required, the completeness of the documentation, and the level of AFTAC evaluation. Because of the fundamental nature of this change, the IB decided that the version number of the July 2006 release of the KB would be incremented to 7.0.

The Fully Mature KB Integration Process

The final, mature KB integration process, which was used to integrate KB7.0 in July 2006, is described in the following sections.

Overview of KB Integration Process

The process begins at the point where the KB Integrator (SNL) prepares for the contributions from the various contributors, and ends with an integrated KB that has been thoroughly evaluated by AFTAC. A flowchart of the process is found in the Figure 2. The integration process can be broken into the following high-level steps:

1. Prepare Product Tracking Matrix
2. Prepare Directory Structure and Databases at SNL
3. Prepare Directory Structure and Databases at AFTAC
4. Prepare KB Documentation
5. Pre-Validation Preparation at SNL
6. Pre-Validation Preparation at AFTAC
7. Build SNL_Tool_Root and Install at SNL and AFTAC
8. Pre-Integration Product Validation at SNL
9. Database Merging at SNL
10. Pre-Integration Product Validation at AFTAC
11. KB Integrator Assembles Executive Summary
12. IP Editor Assembles IP Documents
13. Integrate Products, Documents and Database Information on Classified System at SNL
14. Post-Integration Product Validation at SNL
15. KB Integrator Prepares Final Executive Summary
16. Install Products, Documents and Database Information at AFTAC
17. Post-Integration Product Validation at AFTAC
18. Demo KB to AFTAC
19. Assist AFTAC in Preparing Preliminary KB Evaluation
20. Post Delivery Activities

The following section describes in detail the different steps. In addition, a detailed flowchart of the entire integration process can be found in Appendix A.

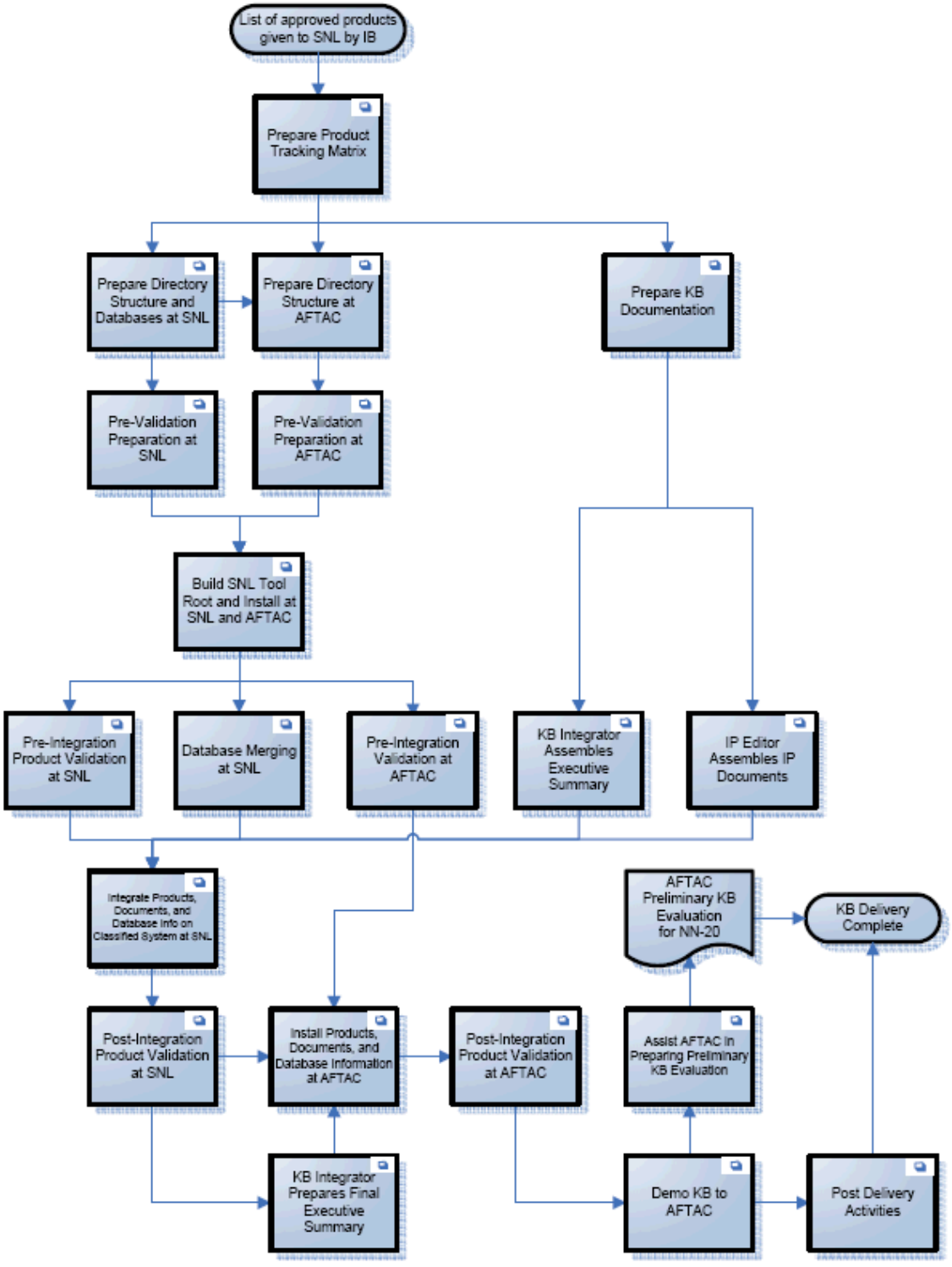


Figure 2. Top level of KB Integration Process Flow

Description of KB Integration Process Steps

1. Prepare Product Tracking Matrix

The Product Tracking Matrix is an on-line tool used to track the products through the integration process. There is a series of forty defined tasks that take an IRP from being proposed through final integration into the KB. In order to use the on-line Product Tracking Matrix, it must be set up for each release of the KB. The first step is to add the due dates for the forty defined tasks. Then all the new and updated IRPs expected in the upcoming KB release are added to the matrix, and the default delivery is set to the upcoming delivery. (See more detail in the Appendix, pages A-2 to A-6).

2. Prepare Directory Structure and Databases at SNL

Before IRPs get to the KB Integrator, the directories and database accounts must be set up. A new directory structure is constructed for each release of the KB, both on the unclassified and classified system. The KB has a specific directory structure that is focused into three major areas – Products, Documents and Tools. Within the tools directory there is a Support directory which holds parameter and configuration files needed for the tools to work properly.

The directory structures are different between the unclassified and classified systems, but the process to set up the directories is similar. Shell scripts are used to create the directory structure. The scripts must be updated to include the new and updated IRPs and to copy over information for IRPs that are not changing from previous releases.

There are three database accounts that hold all incoming database tables for IRPs. In the unclassified database, the accounts are LANLDATA and LLNLDATA, to hold the incoming unclassified data from partner GNEMRE laboratories. In the classified database the account is RPVAL. This account holds all incoming classified database tables. It is also the area where the unclassified database tables from the LANLDATA and LLNLDATA are put when they are first moved to the classified database to be merged. Once the database tables have been integrated, they are put into the classified database account KBASE. All of these accounts are checked to make sure there is enough space for the new database tables that are expected, based on estimates from the Product Integrators. (See more detail in the Appendix, page A-7).

3. Prepare Directory Structure and Databases at AFTAC

The process for creating the KB directory structure at AFTAC is similar to creating the classified directory structure at SNL. In fact, scripts that are used to create the classified directory structure at SNL are sent to AFTAC and run there after being modified for the local environment. An additional step at AFTAC is to take a snapshot of the current *KB Navigator* tree which is sent back to SNL. This will be the starting point for the *KB Navigator* tree for the new release. The *KB Navigator* is the standard software interface to all the IRPs in the KB.

The databases at AFTAC are managed by a Database Administrator (DBA) who has control over amount of space to which each database account has access. The only

database account at AFTAC used by the KB is KBASE. The SNL on-site representative talks with the DBA to make sure there will be enough space in the KBASE account for all the new data coming from SNL. (See more detail in the Appendix, page A-8).

4. Prepare KB Documentation

There are six high level IP documents included in each KB release:

- Detection, Association and Location
- Event Identification, Infrasound and Hydroacoustics
- Yield
- Reporting
- Geophysical Modeling
- Data Management

The IP documents are made up of individual IRP documents which are created by the researcher who is acting as the Product Integrator for the IRP.

To initiate the preparation of the IP documents for the release, the IP Editor sends out IRP templates to the Product Integrators of new IRPs, or if the IRP is being updated, the IP Editor sends the Product Integrator the IRP document from the latest KB release. The Product Integrators use the template to prepare the IRP document for their new IRP, or update the old IRP document. The completed IRP documents are sent back to the IP Editor for integration into the IP documents.

At the same time the IP Editor is working with the Product Integrators to complete the IRP documents, the KB Integrator Coordinator sends out a briefer template to each Product Integrator to get basic information about the IRP that is needed to help the KB Integrator successfully verify, validate and integrate the IRP. These completed documents also include information that the KB Integrator Coordinator uses to write the IP introductions. The KB Integrator Coordinator sends the IP introductions to the IP Editor at the same time the Product Integrators send their IRP documents.

In addition to the six IP documents, there is an Executive Summary which is compiled by the KB Integrator. Each Executive Summary follows a specific format, including an Appendix that lists the history of all the IRPs that have been in the KB. (See more detail in the Appendix, pages A-9 to A-10).

5. Pre-Validation Preparation at SNL

Each IRP has to be verified and validated by the KB Integrator. The KB Integrator assigns “SNL Validators” to all the new and updated IRPs in the KB release to perform these tasks. The SNL Validator, along with the KB Integrator, coordinates with the Product Integrator on how the IRP is supposed to be used and set-up in the KB. This coordination is typically done with a series of emails, so that there is good documentation for what is expected. The Product Integrator can use the email documentation to complete the Files/Installation Testing Procedures document that is sent to the KB Integrator a week before the IRP is due. The SNL Validator for the IRP checks the Files/Installation Testing Procedures document and contacts the Product Integrator to make sure that it is clear and complete. When IRPs arrive at SNL, they are put into the correct unclassified

product directories, unless an IRP is classified, in which case it is put directly onto the SNL classified system. Any database tables in the product are imported into the database, and any configuration or parameter files needed to work with the products are created. As these tasks are completed, the KB Integrator Coordinator updates the Product Tracking Matrix. (See more detail in the Appendix, pages A-11 to A-12).

6. Pre-Validation Preparation at AFTAC

In addition to the IRPs sent to SNL by LANL and LLNL for integration at SNL, there are some IRPs that are prepared at AFTAC and cannot be released outside AFTAC. These IRPs are designated as "AFTAC Local" IRPs. To process these special IRPs, similar pre-validation steps are done at AFTAC that are done at SNL for the other IRPs. The SNL Validator for the AFTAC Local IRPs is the SNL on-site representative. As with the IRPs at SNL, the SNL Validator at AFTAC checks the Files/Installation Testing Procedures document and contacts the Product Integrator if there are things the SNL Validator does not understand. Once these problems have been dealt with, the IRP is copied to the correct product directory at AFTAC, and any database tables in the IRP are imported into the database. Configuration or parameter files needed to work with the KB setup are then prepared. A final step is to create a list of orids (origin identifiers) from the origin database tables for the AFTAC Local IRPs to be used later to merge database information. (See more detail in the Appendix, page A-13).

7. Build SNL_Tool_Root and Install at SNL and AFTAC

The SNL_Tool_Root is the set of tools that are used to create, review and utilize for monitoring the data products in the KB. It includes SNL-developed software as well as all third party software that is required for the tools to run. These tools include Java applications and libraries, C++ libraries, and custom Matlab applications.

A hierarchy of build scripts is used to build the software in the SNL_Tool_Root. Top-level scripts launch release builds and low-level scripts build individual applications. The scripts are written in Apache ANT®, an XML-based scripting language that enables cross-platform builds. ANT® is also capable of building code written in several programming languages, making it a good fit of the SNL_Tool_Root. ANT-Contrib®, an extension to ANT® is also used and enables functional programming and has a platform-independent C++ compiler task.

The software release process is a three-stage process that is handled by the designated release builder and software tester. First the SNL_Tool_Root is built, and then it is installed on the unclassified system at SNL. Testing takes place on the unclassified system in order to validate IRPs delivered there and to test the build prior to moving to the classified system at SNL. If all performs as expected, the SNL_Tool_Root release is transferred to the classified system. More testing of the SNL_Tool_Root occurs on the classified system when it is used to validate classified IRPs and do final KB integration testing. Once the SNL_Tool_Root is successfully tested on the classified system at SNL, it is sent to AFTAC where it is installed and tested for a final time. The SNL_Tool_Root is always released as an integrated tool suite. This ensures that all dependencies are

satisfied and that all the software works together. (See more detail in the Appendix, page A-14).

8. Pre-Integration Product Validation at SNL

Before an IRP can be integrated, the KB Integrator must make sure it can be accessed and used at the KB Integrator site in the same manner it was accessed and used at the Product Integrator's site. This is accomplished by the SNL Validator running the same Installation Tests that the Product Integrator has run at his/her location using the Files/Installation Testing Procedure document that was sent by the Product Integrator in Step #5.

The Files/Installation Testing Procedure document contains a listing of the files and database tables that make up the IRP, and a standard test to run to make sure that the IRP has been installed correctly. The SNL Validator first checks to make sure all the files and database tables listed in the files section of the Files/Installation Testing Procedures document are present. If anything is missing, the SNL Validator contacts the Product Integrator and has them resend the IRP. For products involving database content, the SNL Validator confirms that the database tables for the IRP are present, and then runs the *QCTool* program to check that the tables follow the defined schemas for the KB. If errors are found running *QCTool*, a list of the problems are sent to Product Integrator. The Product Integrator then fixes the tables and resends them. The final step in this part of the process is to prepare any configuration or parameter files that are needed to view the IRP with the tools used by AFTAC. These files should exist because the Product Integrator is expected to test the IRP at their location using the tools that AFTAC will be using, but if they are not present, the SNL Validator will try to create them.

The last Pre-Integration Product Validation step is for the SNL Validator to run the Installation Testing Procedures provided by the Product Integrator. The results achieved at the KB Integrator site are compared to what the results were when the Installation Test was run by the Product Integrator. If the results match, then the IRP is considered to be validated. At this point, the product is ready to be integrated into the KB. (See more detail in the Appendix, pages A-15 to A-17).

9. Database Merging at SNL

The single biggest task for the KB Integrator is to take Oracle database tables from different Product Integrators and merge the information into one set of complete tables. This needs to be done without duplicating records, but such that proper links are maintained between related information and making sure that the information in the tables is accurate. This is done using a tool for quality control checking, *QCTool*, and a suite of tools to manipulate information in the database tables, *DBTools*. For more information on these tools see Carr et al., 2007, and Lewis and Ballard, 2007.

As discussed in Step #8, prior to merging the database tables in the IRP are checked with the *QCTool* to make sure they follow the standard schemas and that the information in the tables is reasonable. A master parameter file for the *QCTool* has been developed that contains the checks the KB Integrator believes are necessary to achieve a reliable and consistent KB. Product Integrators are encouraged to use this parameter file to check their

database tables at their site before the IRP is sent to the KB Integrator. If there are inconsistencies that they do not plan to fix, they can also send a file that explains the reasons for any failed checks. The KB Integrator will run *QCTool* using the parameter file sent by the Product Integrator. If there are problems found running *QCTool* that have not been explained by the Product Integrator, then the tables are sent back to the Product Integrator to be fixed.

Once the database tables have passed the quality control checks in the *QCTool*, the database tables are merged using a number of tools in the *DBTools* suite. The first step is to merge all the database tables that correspond to bulletin information and EventID, Codamag and Location calibration IRPs using the tool *EvLoader*. *EvLoader* is a merging tool that groups together origins (event hypotheses) that are correlated in distance and time as specific events. *EvLoader* is usually run in parallel mode (*parallelyzevloader*) in order to have a number of different processes running at the same time. *EvLoader* checks constraints, relationships and makes sure there are no duplicate records. If any problems are found while merging the data with *EvLoader*, they are analyzed by the KB Integrator and changes are made to tables or the parameter file as necessary, and the events are remerged.

After the bulletin and calibration information has been merged, the tool *WFMerge* is run to merge the waveforms and update information in the two database tables that contain waveform information, **wfdisc** and **wftag**. As with the EvLoader merge, if any problems are found while merging the waveforms, the problems are analyzed by the KB Integrator, fixed, and the waveform data is remerged.

The last set of database tables to be merged is the station information (lookup tables). These tables cannot be merged with *EvLoader* because of issues with their schema design. The *DBTools* software requires non-association tables to be idowner tables with a primary key that is a numeric id. Several of the station information tables do not meet this requirement. So to be able to merge them, the KB Integrator creates custom tables that are copies of the NNSA KB Core station information tables except with an added column that is a unique numeric id. Then the tool *DTX* is used to merge the information. *DTX* (DaTa eXchange) is another tool to merge data, but it doesn't have the time/distance correlation features of *EvLoader*. If any problems are found while merging the station information, the problems are analyzed by the KB Integrator, fixed and the information remerged. Once the station information is merged, the merged tables are converted back to the official station information tables. (See more detail in the Appendix, page A-18).

10. Pre-Integration Product Validation at AFTAC

Pre-Integration Product Validation is also done for the IRPs that were prepared at AFTAC to make sure that these IRPs can be accessed and used successfully. The same tasks done to validate the IRPs before integration at SNL are done at AFTAC. The on-site representative at AFTAC (the SNL Validator) checks to make sure all the files and database tables listed in the files section of the IRP Files/Installation Testing Procedures document are present. If anything is missing, the SNL Validator contacts the Product Integrator and has the Product Integrator fix and update the IRP. The SNL Validator confirms that the database tables for the IRP are present, and then runs the *QCTool*

program to check that the tables follow the defined schemas for the KB. If errors are found running *QCTool*, the Product Integrator is contacted and asked to come to AFTAC to fix the database tables. The final step in this part of the process is to prepare any configuration or parameter files that are needed to view the IRP with the tools used by AFTAC. These files should exist because the Product Integrator is expected to test the IRP at their location using the tools that AFTAC will be using. If they are not present, the SNL Validator creates them.

Finally, the SNL Validator runs the Installation Testing Procedures provided by the Product Integrator. The results achieved at the AFTAC site are compared to what the results were when the Installation Test was run by the Product Integrator. If the results match, then the IRP is considered to be validated. (See more detail in the Appendix, pages A-19 to A-21).

11. KB Integrator Assembles Executive Summary

The Executive Summary from the prior release is copied, and used as the starting point for the new Executive Summary for the current release. The title is changed to be consistent with the current release number and all other references to the old release number in the text and footers are also updated. Appendix A of the Executive Summary, which lists the history of all IRPs that have been in the KB, is updated to include the new and updated IRPs for this release of the KB. Using Appendix A, the IP introductions, and information from the IB, the rest of the document is updated. The only section that is not updated at this point is the "IB Actions" section, since that part of the Executive Summary documents what happens at the IB meeting which occurs at the end of the KB Integration process. The final step is to create hyperlinks from Appendix F which lists all the report products to the individual reports. When the draft Executive Summary is completed, a PDF version is created and sent to the IB members for their review and comments. It is also sent to the IP Editor so that it can be edited for spelling, grammar, and consistency. (See more detail in the Appendix, pages A-22, A-24).

12. IP Editor Assembles IP Documents

Once the IP Editor receives all the new components of the IP documents – new and updated IRP documents and the IP introductions – he starts assembling the documents. The IP documents are compiled using the word processor Framemaker. If any of the IRP documents arrived in a different format, (e.g. Microsoft Word), the IP Editor must convert those documents to Framemaker. The IP Editor then changes the footers to make sure they have the correct KB Release number, and makes sure that the headings have the correct product number. The IP Editor checks for spelling, punctuation and grammar, and adds hyperlinks where needed. IRP document placeholders are made for the AFTAC Local IRPs (see Step#6), because their documentation cannot leave AFTAC.

Once the introduction and IRP documents for a single IP document are edited and complete, then the IP Editor creates a Framemaker book file that contains all the IRP documents in sequential order. A reference section is added at the end of the book, and a table of contents and title page are added at the front. When the IP Editor finishes the IP documents, he creates PDF files of the finished IP documents, plus PDF files of all the

individual IRP documents. The IP Editor sends all the PDF documents and the Framemaker books to the KB Integrator.

The KB Integrator sends out the IP documents (in PDF) to the IB members at the same time as the final draft Executive Summary. The Framemaker books are sent to AFTAC, and the IP Editor travels to AFTAC to edit and incorporate the documentation for the AFTAC Local IRPs into the Framemaker books. At AFTAC the IP Editor creates new PDFs for the IP documents that involve AFTAC Local IRPs. (See more detail in the Appendix, pages A-23, A-25).

13. Integrate Products, Documents and Database Information on Classified System at SNL

As all the new and updated IRPs are validated on the unclassified system, they are moved into the correct product directory on the classified system at SNL. The merged database tables are imported into the KBASE database account in the classified database, and all necessary constraints and indexes are created. Configuration and parameter files associated with the IRPs are moved to the Support directory and updated to point to the newly merged database tables with the IDs from the merged tables. The various documents – Executive Summary from the KB Integrator, IP documents and the individual IRP documents from the IP Editor are placed in the appropriate document directories.

Once all the information is in place, the snapshot of the *KB Navigator* sent from AFTAC is used as a starting point for the new release. A new *KB Navigator* tree flat file structure is created and the nodes from the old release are copied into the new tree. The nodes for the existing IRPs are updated to reflect the new location of the IRP on disk. Nodes are created for the new IRPs, and the nodes for any updated products are updated to reflect the new node name, new location of the IRP on disk and any new files or documents that have been added to the IRP. Once all the nodes have been entered into the *KB Navigator* tree, they are organized by Working Group and product number. (See more detail in the Appendix, pages A-26 to A-27).

14. Post-Integration Product Validation at SNL

Validation processes that were done in the Pre-Integration Product Validation at SNL are repeated, this time using the integrated information - database tables, configuration files, and parameter files. In addition, the *KB Navigator* is used to test every node in the KB.

If there is a document included with the IRP, it is opened through *KB Navigator* and reviewed to make sure the document is readable and that the node name matches the title of the document. If there is a database connection in the IRP, the database query nodes in the *KB Navigator* tree are launched and the results plotted with the *KB Navigator* Map Viewer. The database setup, environment variable information and table names are checked to make sure they are correct. If the IRP uses a KB Tool, or is a KB Tool, the node that launches the tool is tested to make sure it works. Finally, if there is a product validation script node, it is run and checked to make sure the results are what are expected. (See more detail in the Appendix, page A-28).

15. KB Integrator Prepares Final Executive Summary

With the KB assembled and validated at SNL, the Executive Summary needs be completed for delivery as well. The IB members and IP Editor send any comments they have on the draft Executive Summary to the KB Integrator. The KB Integrator compiles the comments and makes any necessary changes to the document. An updated draft of the Executive Summary is sent to the IB members at least one week prior to the IB conference call to approve the KB release.

The final Executive Summary is written by incorporating the results of the IB conference call that approved the KB release into the final draft document. A PDF document of the finalized Executive Summary is created. The last Appendix in the Executive Summary, Appendix F, is a bibliography of all the reference reports in the KB. In the PDF version of the Executive Summary, hyperlinks are created in Appendix F so that the reference reports are opened when the reader clicks on the specific document in the bibliography. Testing is done to make sure that all the reference reports open when clicking on the hyperlinks in Appendix F.

The new PDF version of the Executive Summary replaces the draft Executive Summary that was used to do the Post-Integration product validation at SNL. Using the *KB Navigator*, the hyperlinks in Appendix F in the Executive Summary are checked to ensure they work. A paper copy of the Executive Summary is printed and sent to the NNSA GNEMRE program manager in NA-22. (See more detail in the Appendix, page A-29).

16. Install Products, Documents and Database Information at AFTAC

Once an IRP has been verified and validated, a tar file of the IRP directory structure is created. The information in the tar file includes all files associated with the IRP, including database table exports and a sql script to set constraints and indexes in those tables. Tar files of all the new and updated IRP directories from the classified system at SNL are created. In addition to the IRP tar files, tar files are created for the Support area on the classified system at SNL, which includes all the parameter and configuration files needed for the IRPs, and the documents directory that contains the Executive Summary, IP documents and metadata and all the waveforms and instrument responses. The last tar file created contains the flat file database representing the new *KB Navigator* configuration. All the tar files are sent to AFTAC via a SIPRNet ftp site.

Once at AFTAC, the tar files from SNL are put in a staging area. Systematically each tar file is manually moved into the correct spot in the directory structure and untarred. All database tables are imported into the kbase database account, and the indexes and constraints are created. The AFTAC Local IRPs are also moved into the correct spot in the directory structure, and their database tables imported into the database. There are instances when an IRP has to be resent to AFTAC because minor changes were made at SNL after the IRP was originally sent. In those cases, a copy of the existing product directory is created and any differences that are found between the old information and new information are reconciled. After the IRPs are all in place, the tars of the Support

area, waveforms and instrument responses are all untarred as well. If necessary, the Q Model files (attenuation information used for magnitude and/or discrimination) in the Support area are updated with information that is only at AFTAC.

The database information from the AFTAC Local IRPs needs to be merged into the database tables sent from SNL. The same procedures that were followed at SNL are done at AFTAC. The tables are checked to make sure they follow the standard schemas and that the information in the tables is reasonable using the *QCTool*. The new bulletin information and EventID, Codamag and Location calibration IRP database information are merged into the tables from SNL using *EvLoader*. The new waveform information is merged with *WFMerge* and the new station location information is merged with *DTX*. Updates are made to the *dir* columns in the **wfdisc** and **instrument** tables to point to correct locations for waveforms and instrument responses respectively.

Once all the database tables have been merged, a final step is done at AFTAC that is not done at SNL. The IDs in the newly merged tables are remapped to have the same IDs as they did in the prior version of the KB. This is done to enable AFTAC users to verify changes to specific database content that they may have validated and/or used operationally. The tool *Remap* is used to map the orids from the **origin** table of the new release to the **origin** table of the prior release. The comparison is made by using the unique key from the **origin** table – *lat*, *lon*, *depth*, *time*, *auth*. The mapping information is put into a specified remap table in the database. The script update_ids.sql is run and it uses the information from the remap table to update the evids and orids in the database tables to be the numbers from the previous release. After the evids and orids are remapped to be the prior release numbers, another remapping occurs to remap the events in the KB to the events in the Bulletins account at AFTAC. The remap table that is created in this process makes it easy to compare information from the KB to the other databases at AFTAC.

Once the database information has been merged, the tar file that contains the documents is untarred and the documents are all put in place.

The final step in the delivery and installation process is to update the flat file *KB Navigator* tree sent from SNL. Nodes are created for all the AFTAC Local IRPs. (See more detail in the Appendix, pages A-30 to A-39).

17. Post-Integration Product Validation at AFTAC

Working through the *KB Navigator* new flat file database, all products are checked to make sure they work at AFTAC. If there are problems, the nodes are fixed locally or SNL is contacted to help resolve the problems. All documents are checked to make sure they can be opened, read and the hyperlinks work. Again, this gives validation to both the products and to the *KB Navigator* hierarchy that has been set up to access them. Once everything is working, the new *KB Navigator* nodes are imported from the flat files and saved to the database. The final step is to run a script to make the *KB Navigator* the one that is used by all the researchers at AFTAC. (See more detail in the Appendix, page A-40).

18. Demo KB to AFTAC

Once the KB has been made live, any of the products in it should be ready for demonstration. The Product Integrators send to AFTAC any demonstration files and presentations they need to demonstrate their IRP to the AFTAC researchers. The on-site SNL representative loads them on the system and checks that they work. The overall KB release demo typically starts with the KB Integrator providing an overview of the new KB release. Then each individual Product Integrator presents and demos their IRP to AFTAC personnel. If requested, workshops or one-on-one tutorials are also provided. (See more detail in the Appendix, page A-41).

19. Assist AFTAC in Preparing Preliminary KB Evaluation

AFTAC evaluates and checks each node in the *KB Navigator* tree and documents the status. If there are problems with any of the nodes, AFTAC works with the KB Integrator to fix or explain the issues. Once all the IRPs have been examined, a copy of the evaluation document is given to the on-site SNL representative. He comments on the problems and fixes the issues noted if possible. These comments are incorporated into the final evaluation document which is delivered to the NNSA GNEM program manager at NA-22. (See more detail in the Appendix, page A-42).

20. Post Delivery Activities

At AFTAC the SNL on-site representative goes through the list of any additional steps that were required as the KB was installed, integrated and tested at AFTAC. He prepares a brief document that details the major problems that occurred during the process, such as having to make the same changes to a large number of files, difficulties in importing database tables, etc. Using this report, plus the KB Evaluation document created by AFTAC, the KB Integration team meets and discusses how to change the process to make it work better. The KB Integrator Coordinator writes a short report that details the changes to the process and ensures that they are implemented for the next release. (See more detail in the Appendix, page A-43).

References

Anderson, J., W. E. Farrell, K. Garcia, J. Given and H. Swanger, 1990, Center for Seismic Studies Version 3 Database: Schema Reference Manual, CSS Technical Report C90-01

Carr, D., J. Lewis, S. Ballard, E. Martinez, J. Hampton, B. Merchant, R. Stead, M. Crown, J. Roman-Nieves, and J. Dwyer, 2007, Advances in the Integration of Large Data Sets for Seismic Monitoring of Nuclear Explosions in Proceedings of the 29th Annual Monitoring Research Review Symposium, Denver, Colorado, USA, pp 927-934.

Gallegos, D. P., D. B. Carr, C. J. Young, P. B. Herrington, J. M. Harris, C. L. Edwards, S. R. Taylor, J. C. Aguilar-Chang, J. J. Zucca, D. B. Harris, D. N. Anderson, L. A. Casey, 2003, The Integration Process for Incorporating Nuclear Explosion Monitoring Research Results into the National Nuclear Security Administration Knowledge Base, SAND2002-2772.

Gallegos, D. P., D. B. Carr, C. J. Young, P. B. Herrington, J. M. Harris, C. L. Edwards, S. R. Taylor, J. C. Aguilar-Chang, J. J. Zucca, D. B. Harris, D. N. Anderson, L. A. Casey, 2005, The Integration Process for Incorporating Nuclear Explosion Monitoring Research Results into the National Nuclear Security Administration Knowledge Base, SAND2002-2772 Supercedes SAND2002-2772 dated September 2003.

Lewis, J.E. and S. Ballard, 2007, GNEMRE DBTools: A Suite of Tools for Access, Maintenance and Manipulation of Data in a Relational Database, SAND2007-5992.

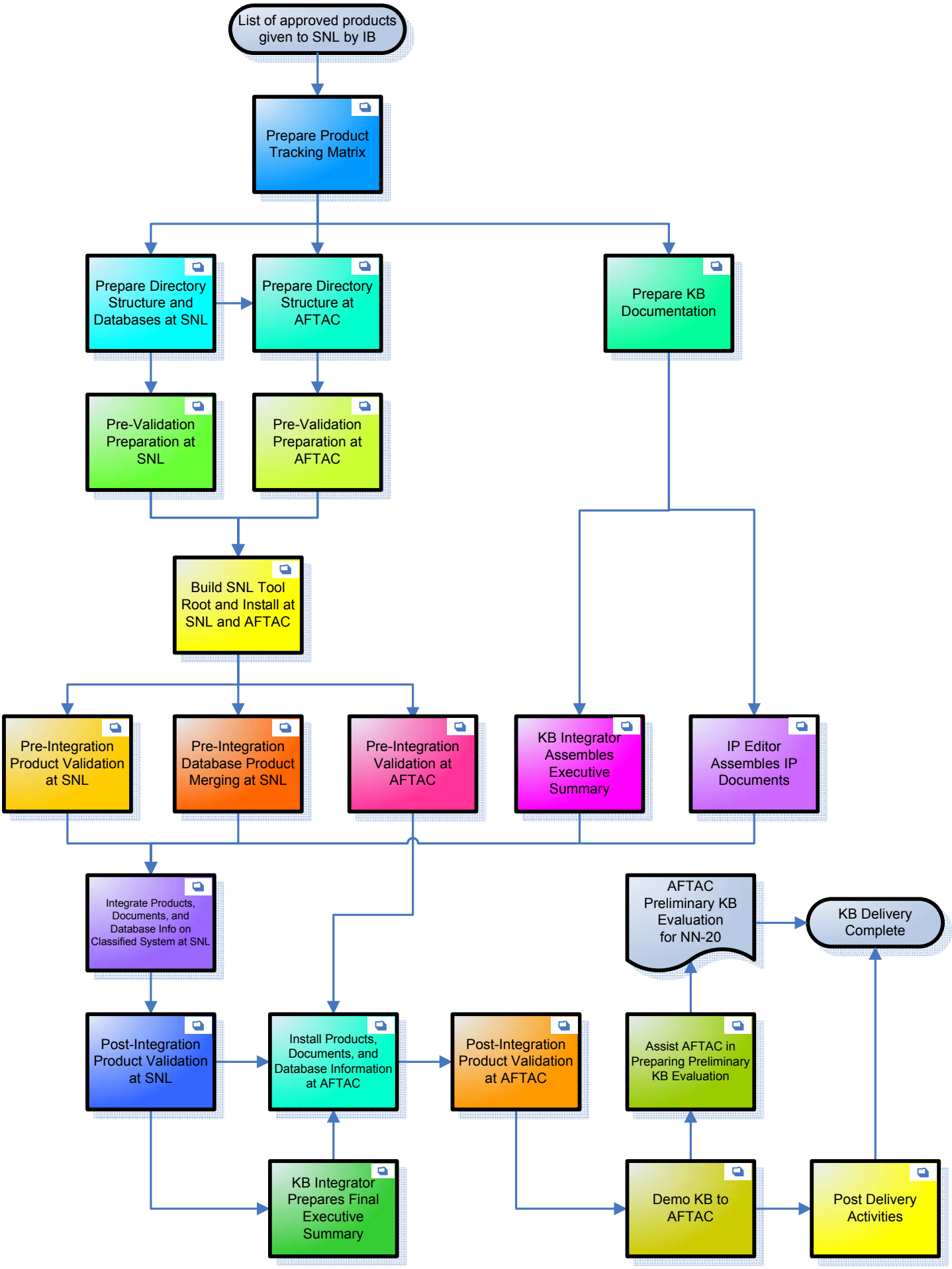
Merchant, B. J., D. Hart, E. P. Chael, M. N. Chown and J. W. Hampton, 2004, Knowledge Base Navigator Facilitating Regional Analysis Inter-Tool Communication in Proceedings of the 26th Annual Seismic Research Review Symposium, Orlando, Florida, USA., pp 760-769

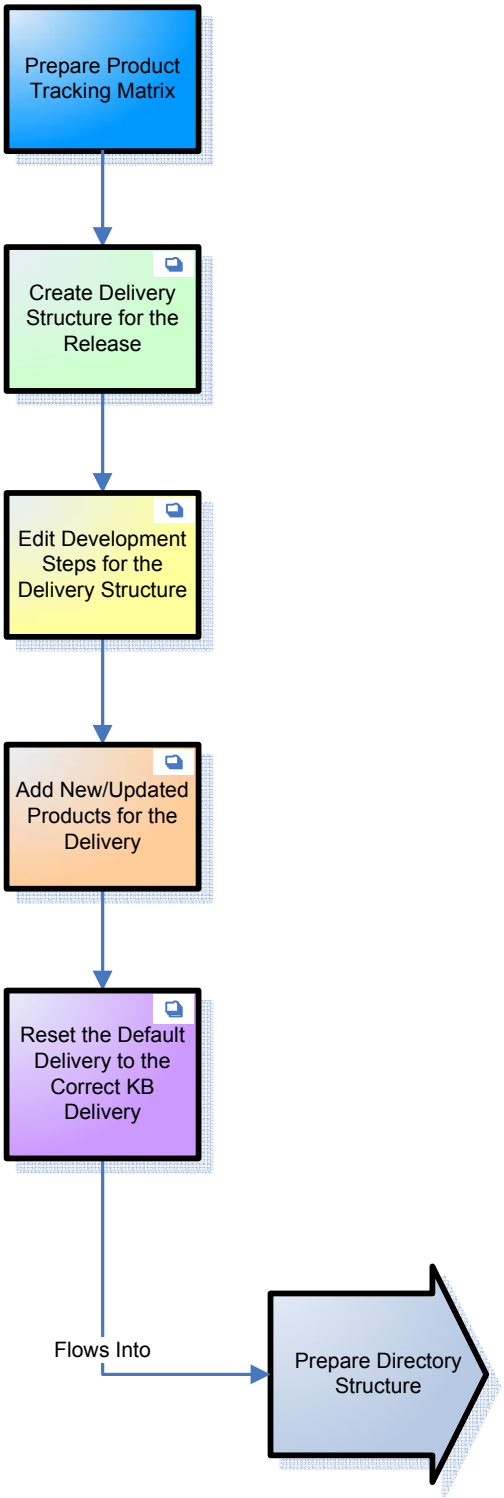
Appendix – Flow Chart of Process

To understand the flow of the Integration process, a flowchart was created to show the details. The first page of the Appendix is the shows the overall flow. Following the first page are drill down flowcharts for all the various processes. If you have an electronic version of this document, click on a box in the first page to get to details for that process. The process step(s) being described on each page are identified with labels at the bottom of each page. For example, the page describing the Verification Testing of Products step done during the Pre-Integration Product Validation at SNL has the label:

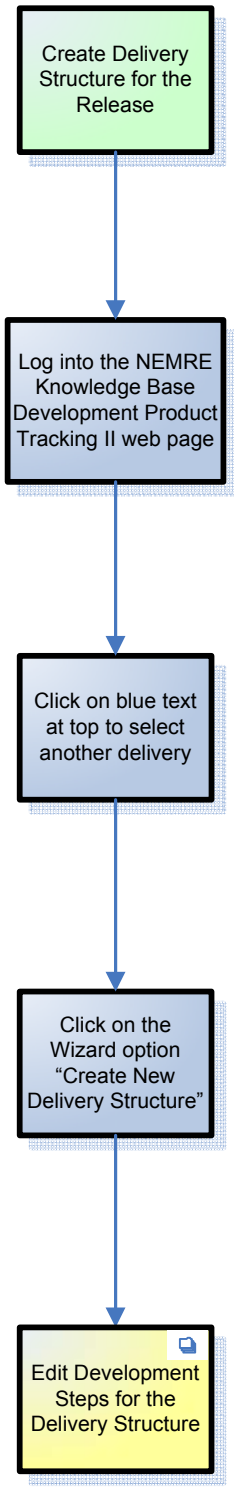
Pre-Integration Product Validation at SNL
Pre-Integration Verification Testing at SNL

Knowledge Base Integration Process Flow

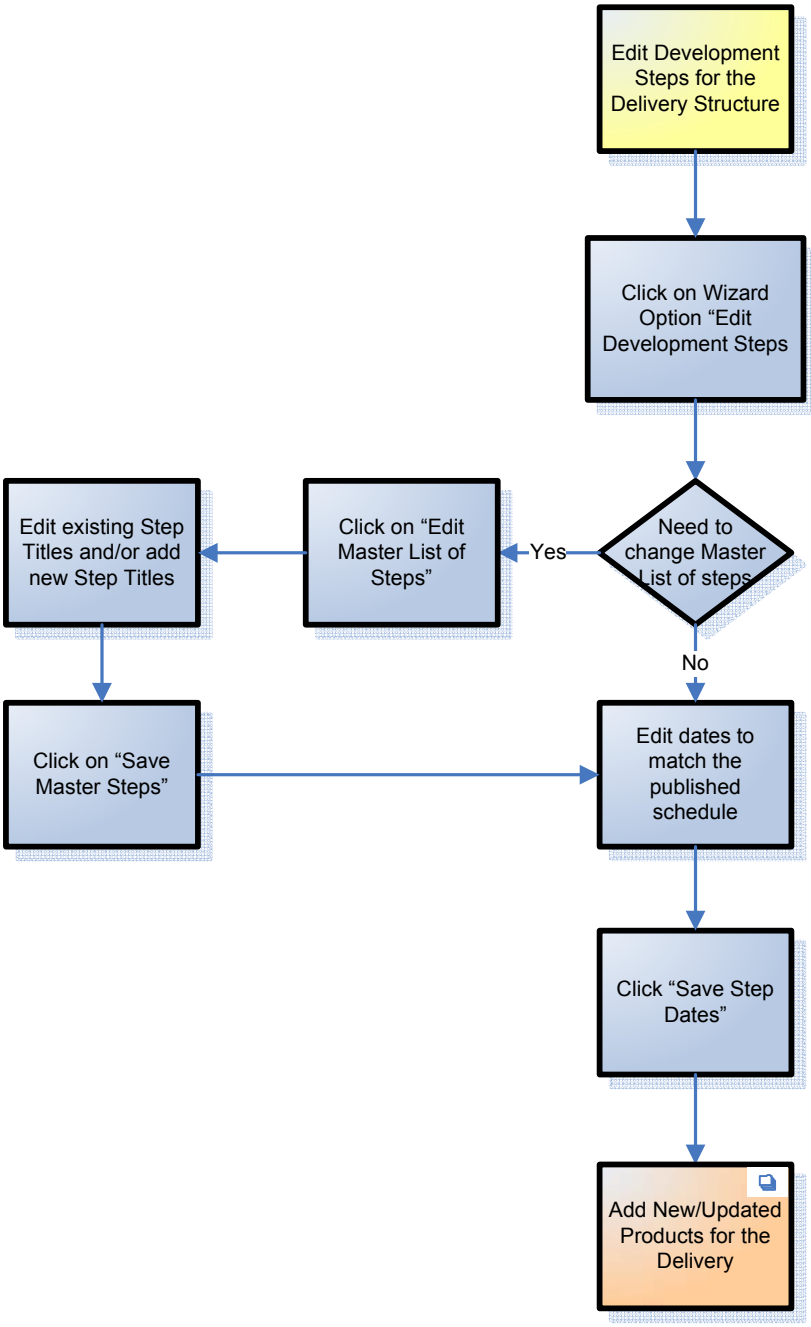




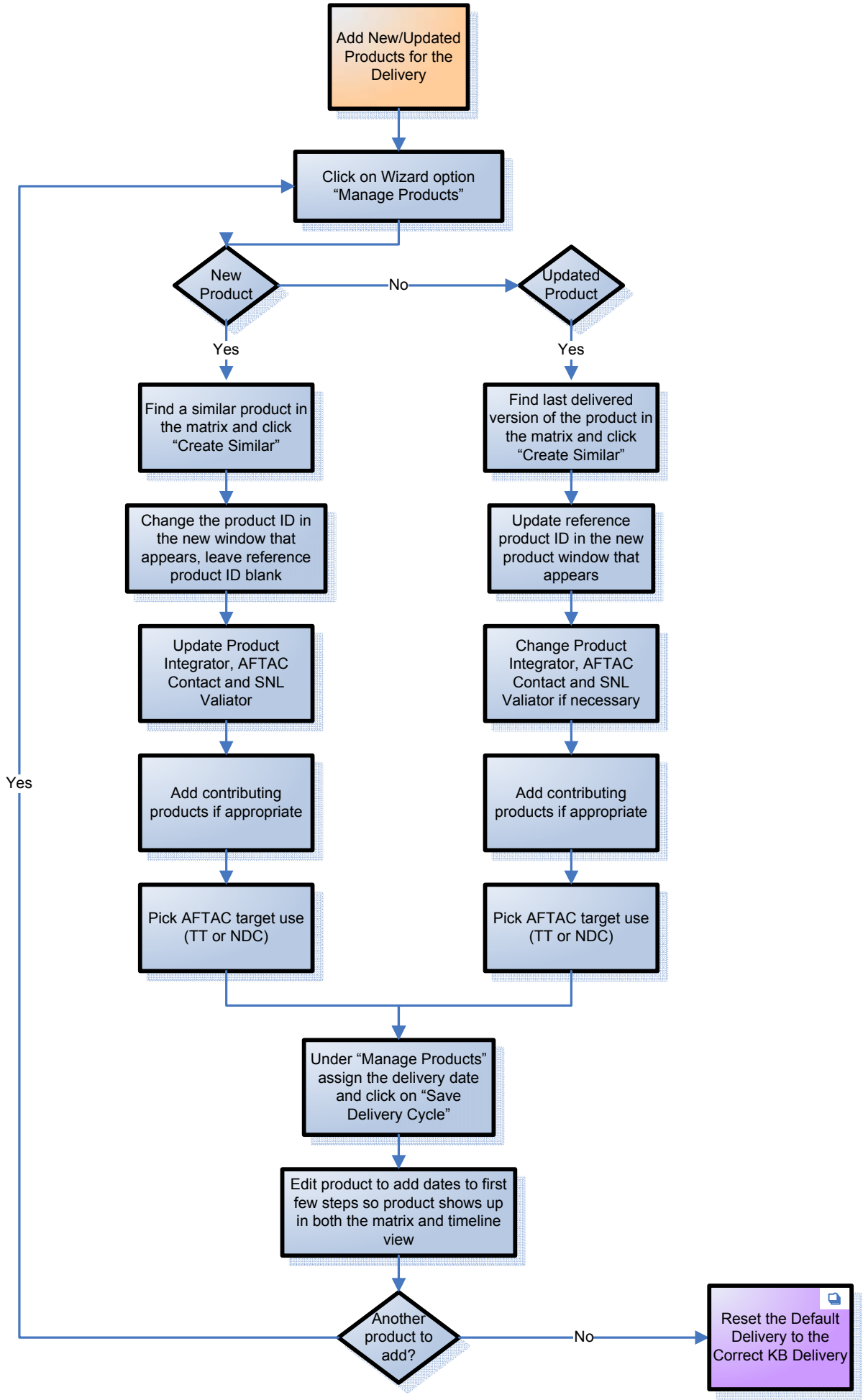
Prepare Product Tracking Matrix



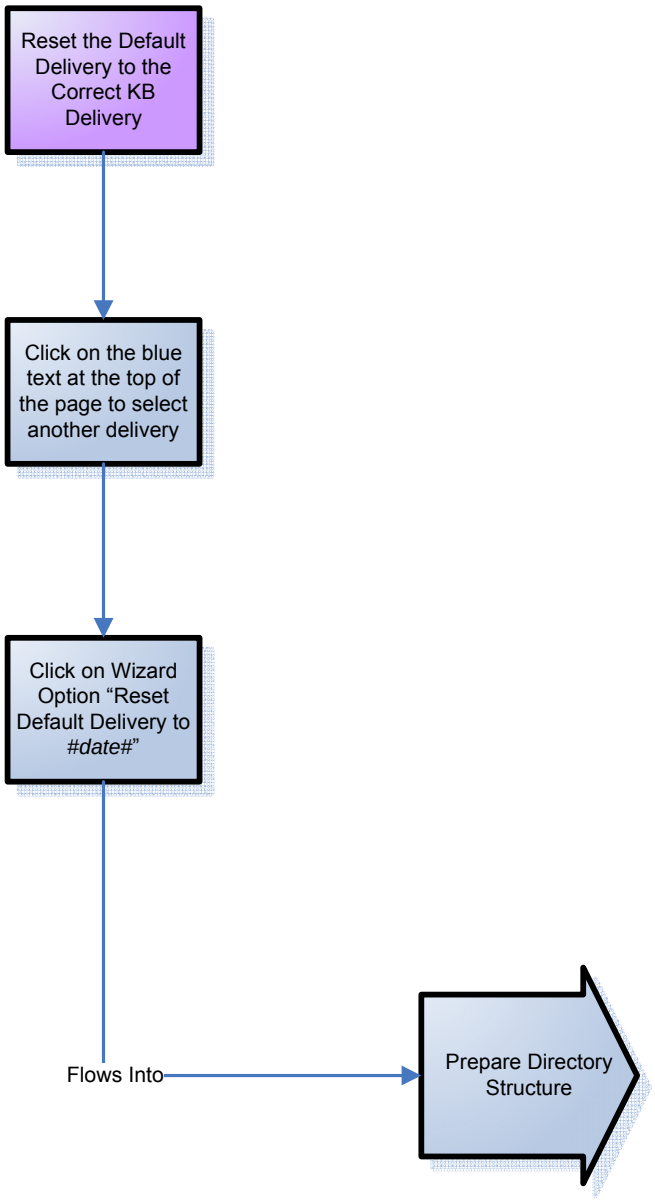
Prepare Product Tracking Matrix
Create Delivery Structure



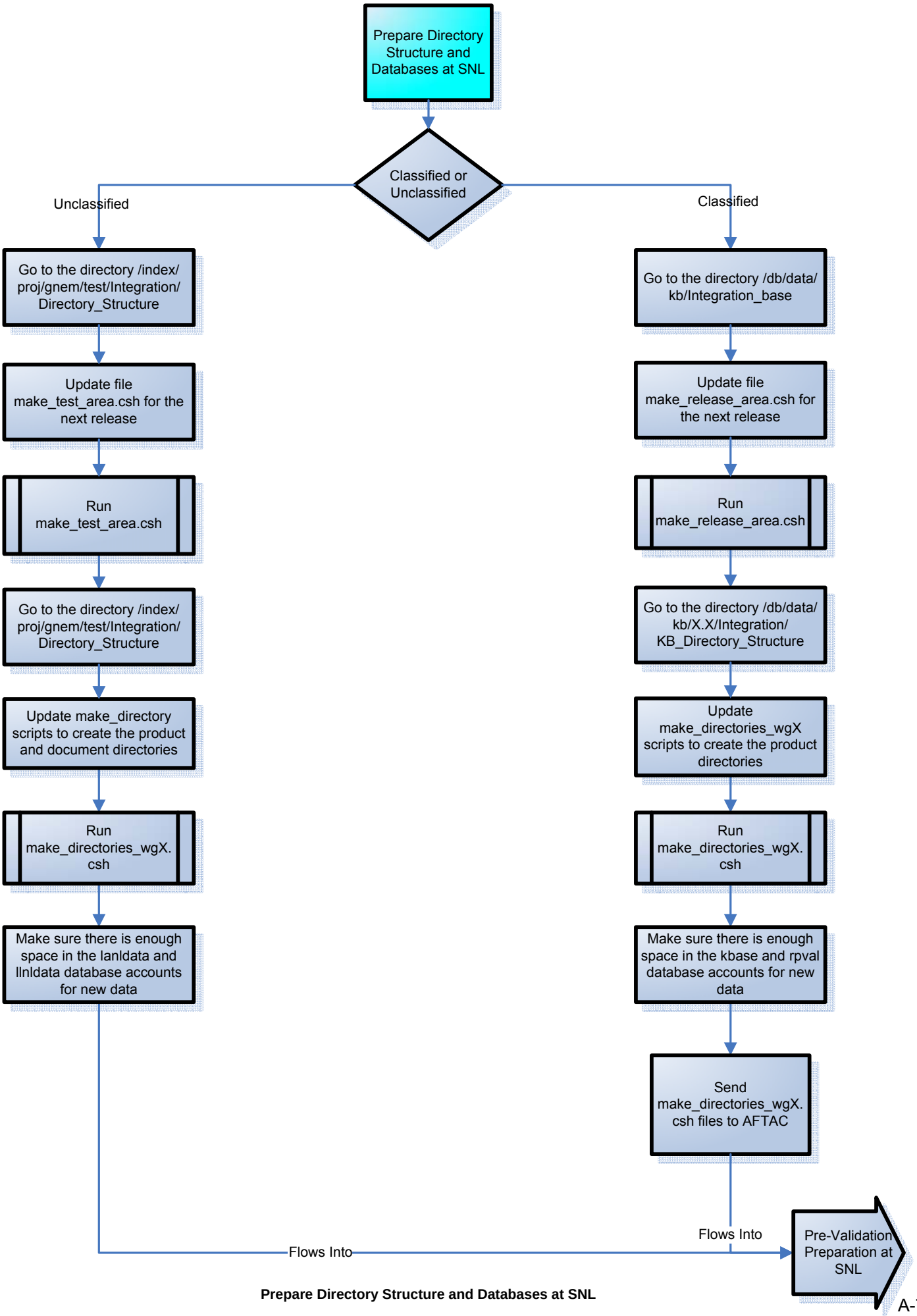
Prepare Product Tracking Matrix
Edit Development Steps



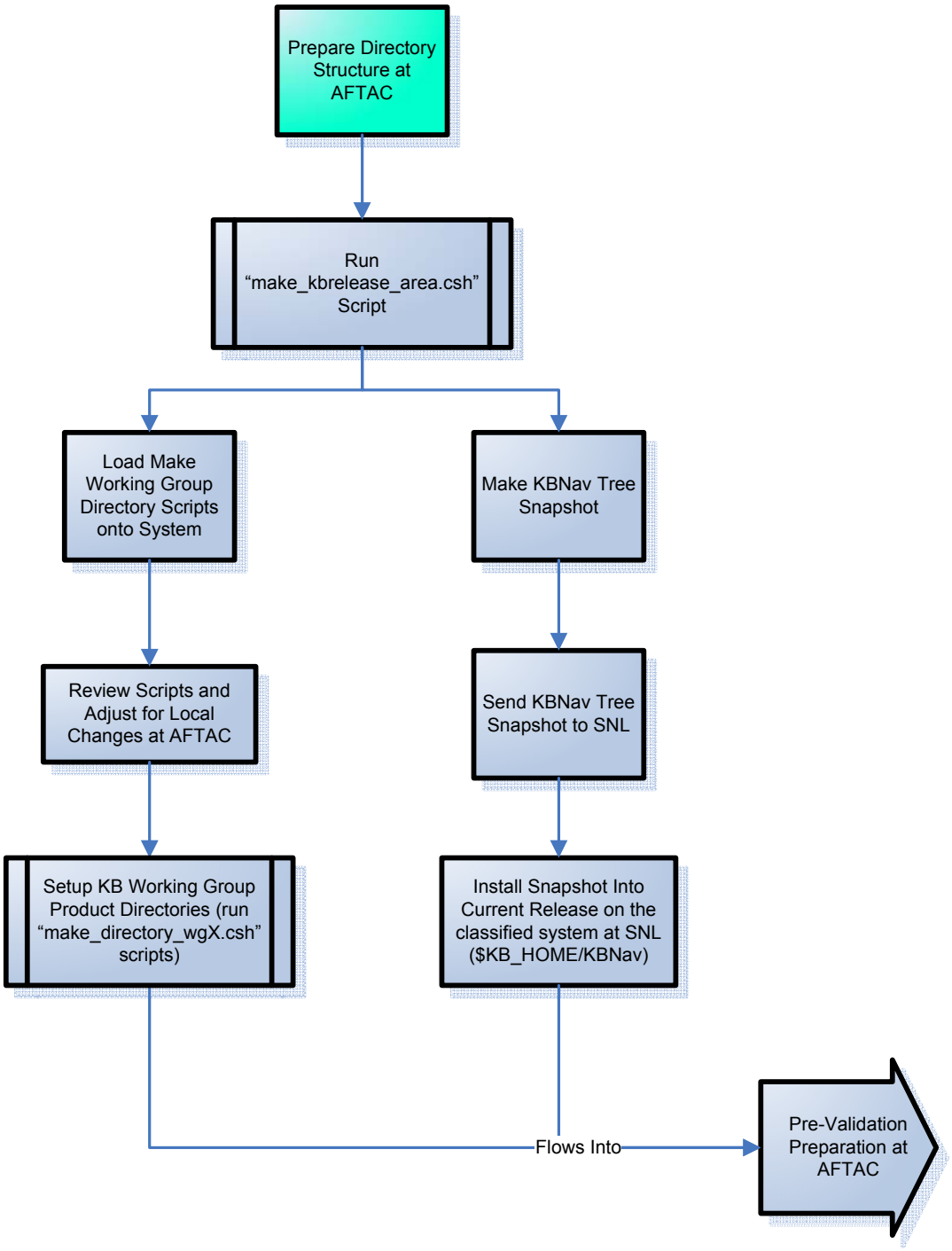
Prepare Product Tracking Matrix
Add Products for Delivery



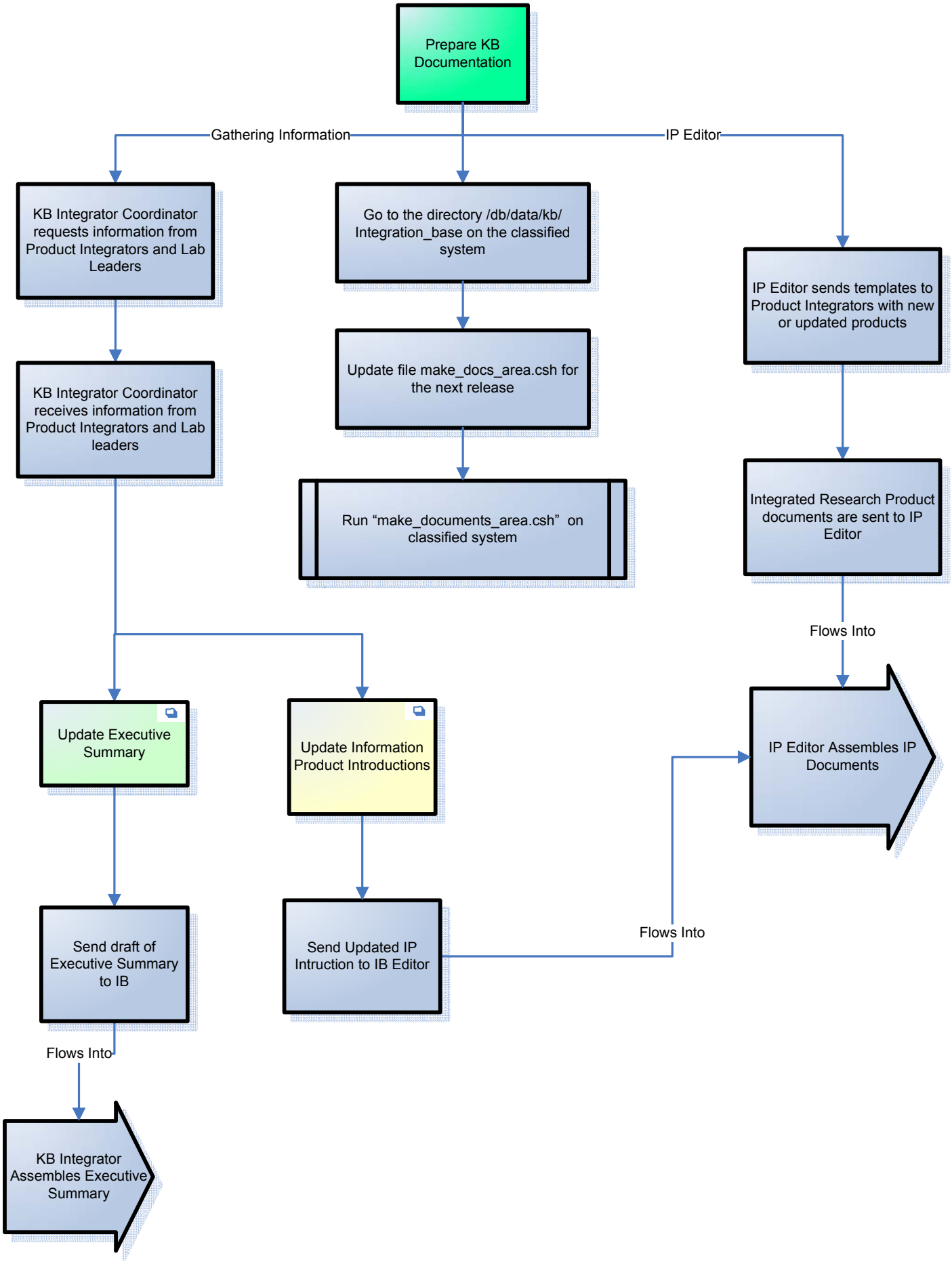
Prepare Product Tracking Matrix
Reset Default Delivery



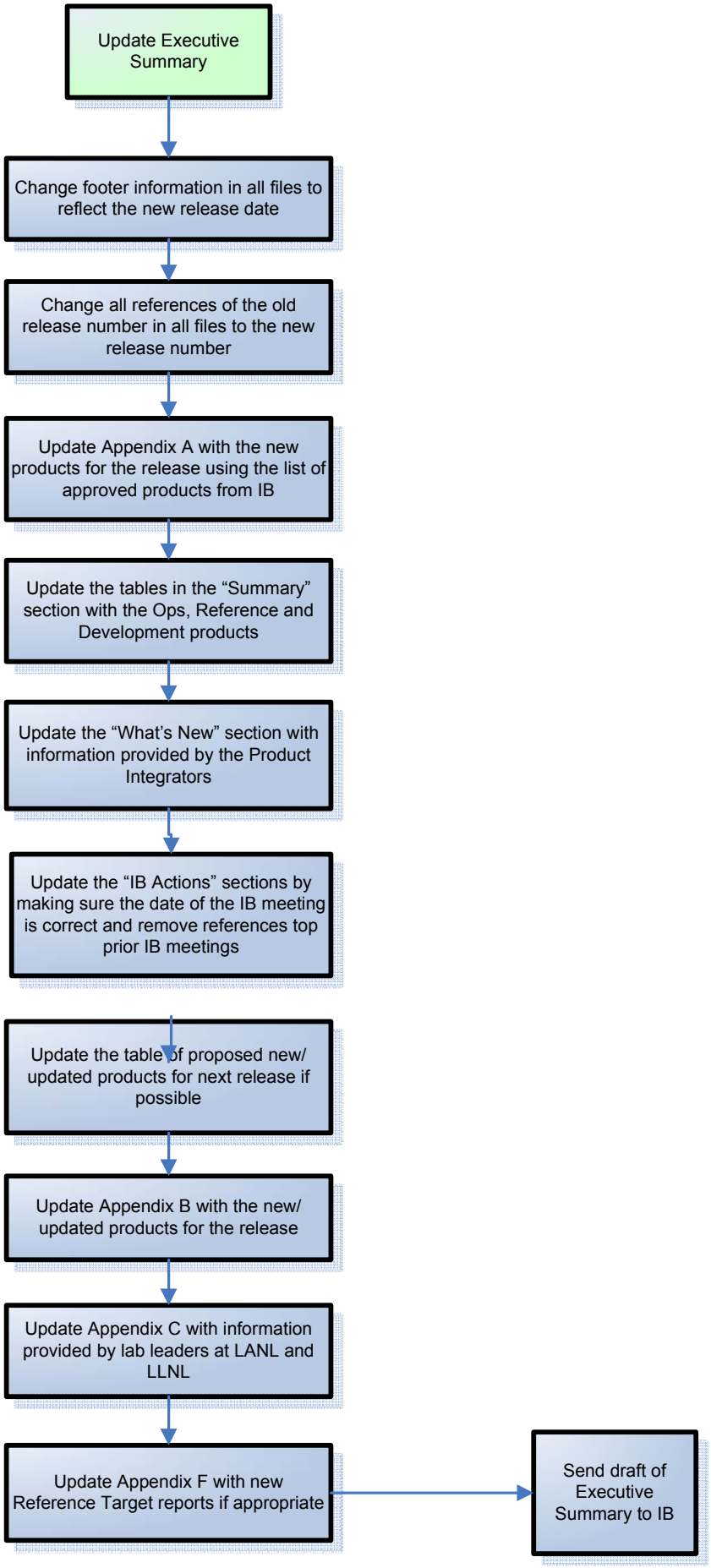
Prepare Directory Structure and Databases at SNL



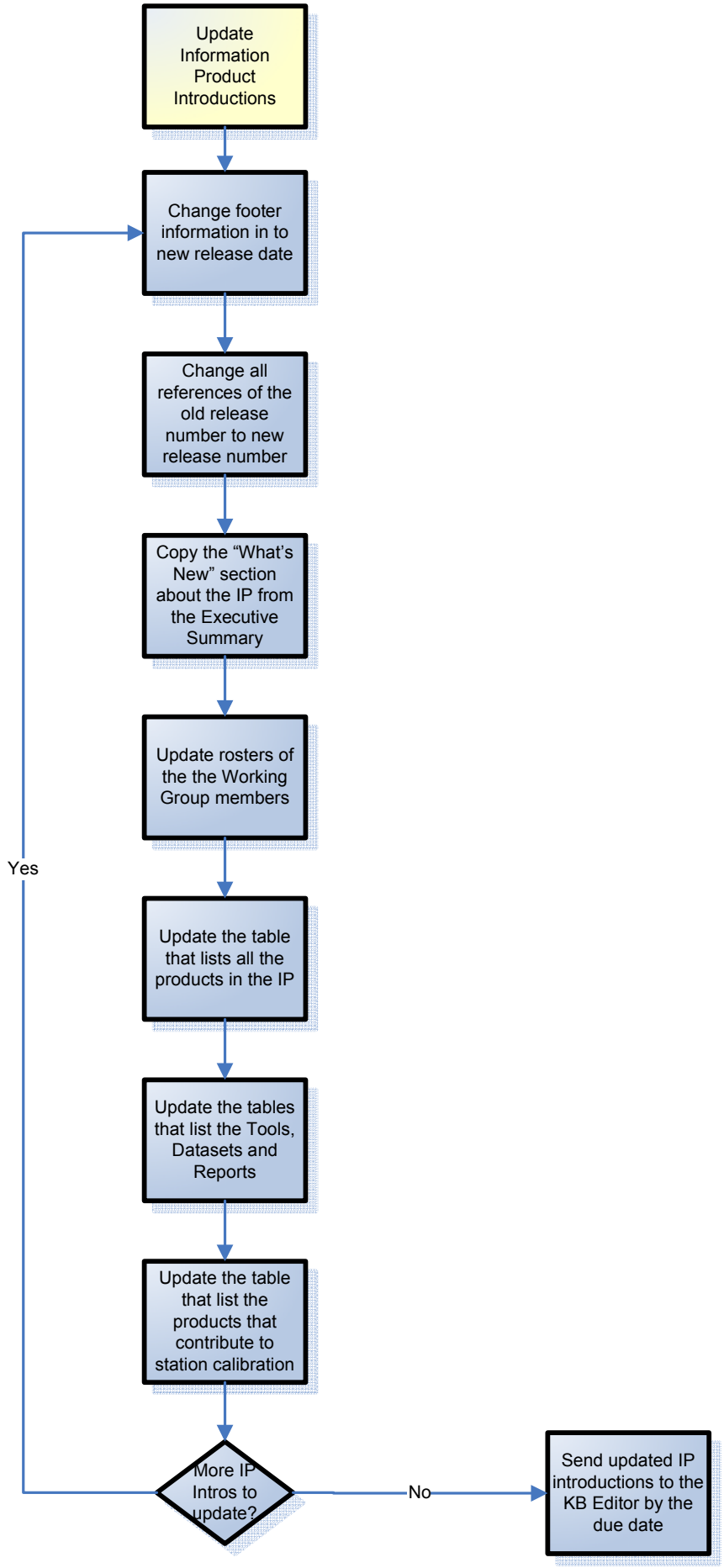
Prepare Directory Structure at AFTAC

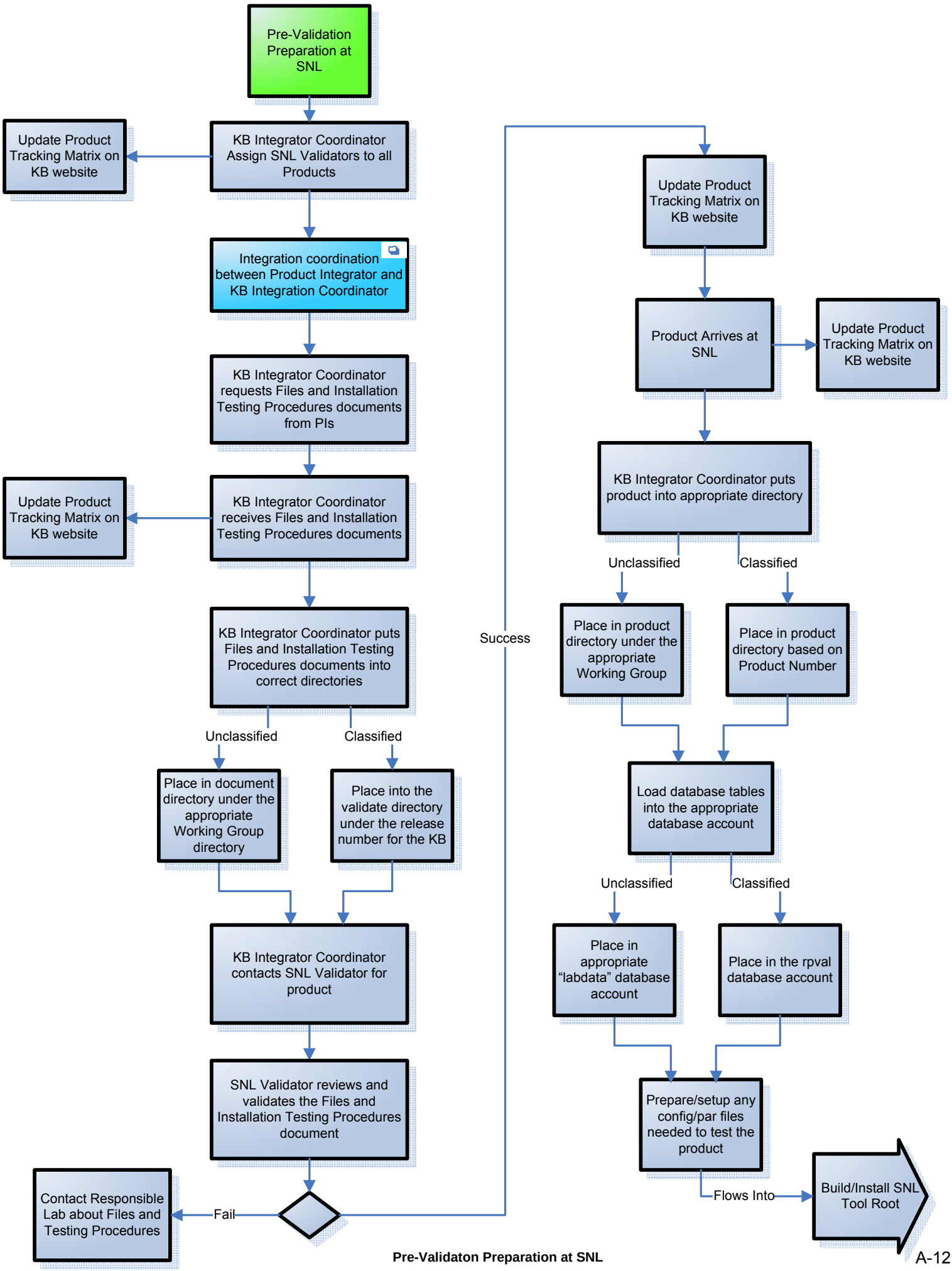


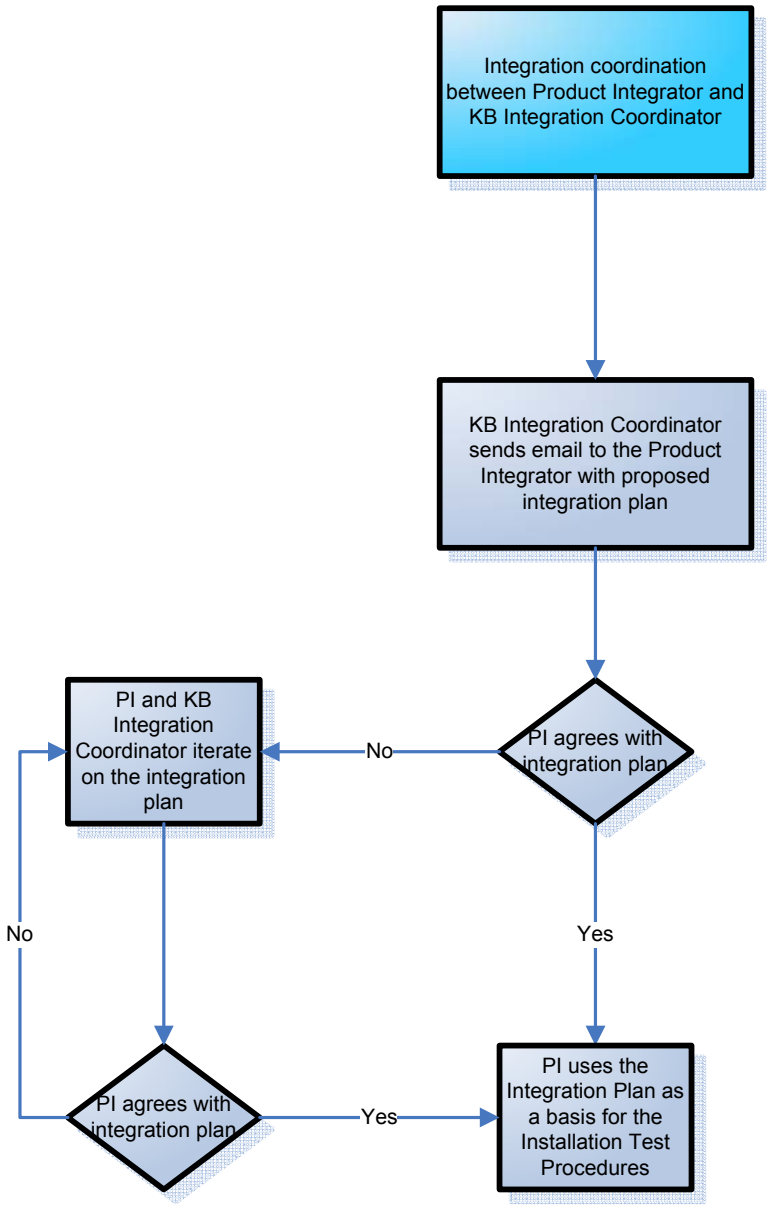
Prepare KB Documentation



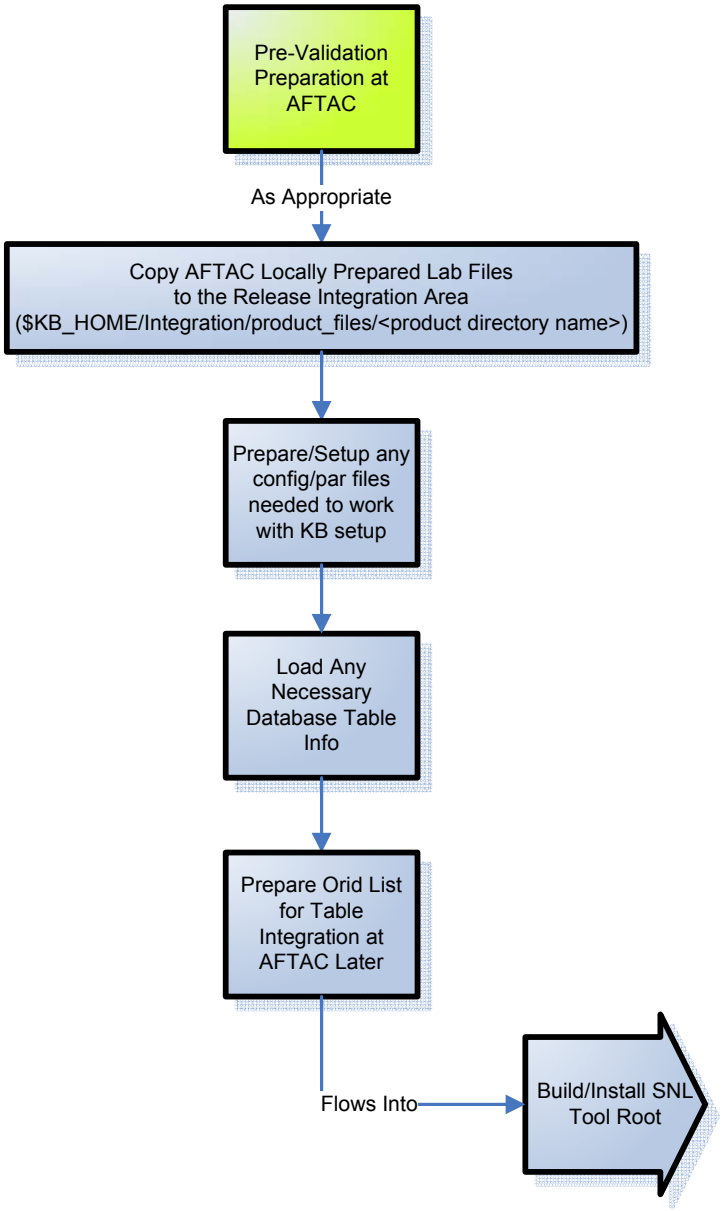
Prepare KB Documentation
Update Executive Summary


Update Information Product Introductions
Change footer information in to new release date
Change all references of the old release number to new release number
Copy the "What's New" section about the IP from the Executive Summary
Update rosters of the the Working Group members
Update the table that lists all the products in the IP
Update the tables that list the Tools, Datasets and Reports
Update the table that list the products that contribute to station calibration
More IP Intros to update?
Yes
No
Send updated IP introductions to the KB Editor by the due date

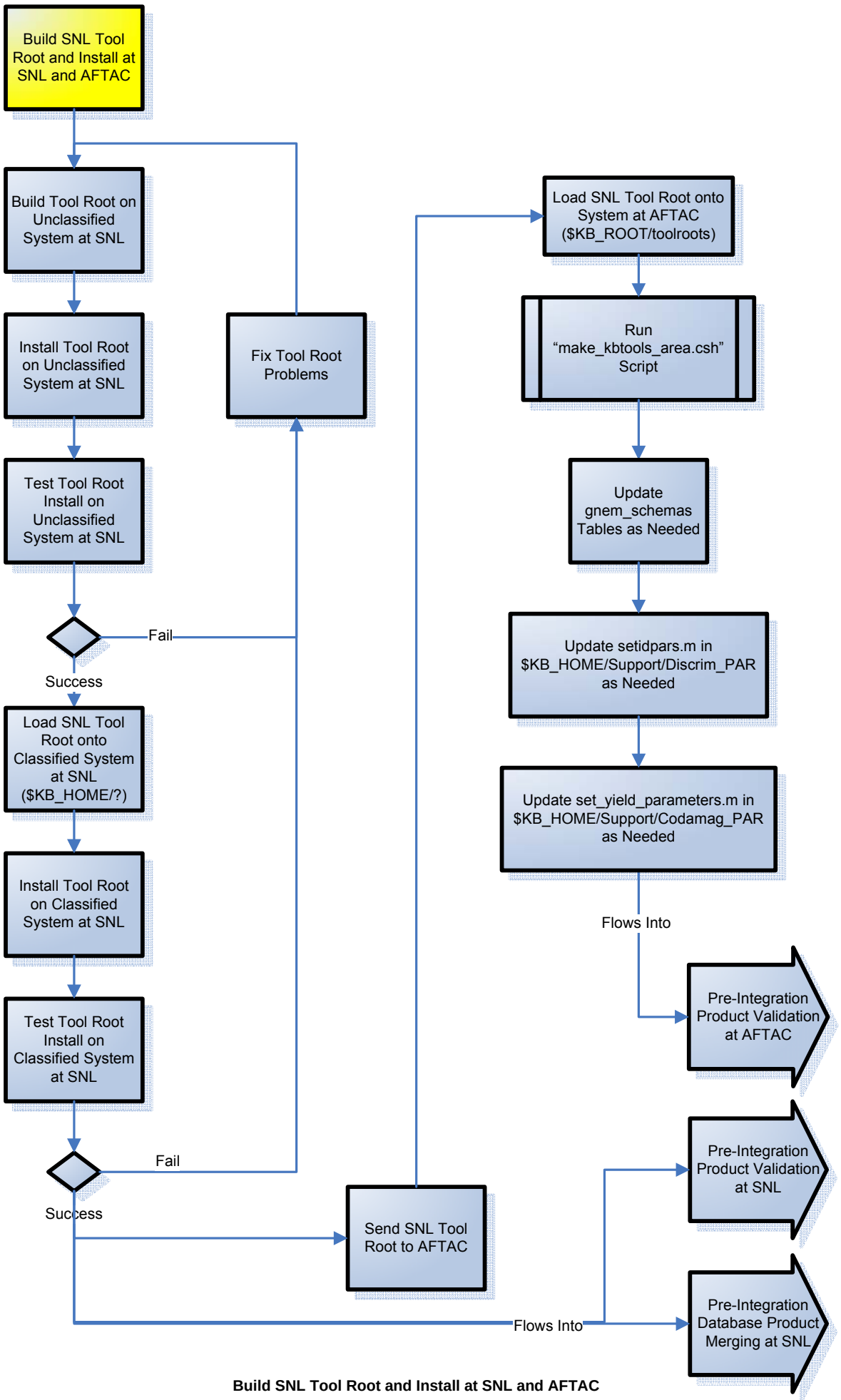

Pre-Validation Preparation at SNL
KB Integrator Coordinator Assign SNL Validators to all Products
Update Product Tracking Matrix on KB website
Integration coordination between Product Integrator and KB Integration Coordinator
KB Integrator Coordinator requests Files and Installation Testing Procedures documents from PIs
KB Integrator Coordinator receives Files and Installation Testing Procedures documents
Update Product Tracking Matrix on KB website
KB Integrator Coordinator puts Files and Installation Testing Procedures documents into correct directories
Unclassified
Classified
Place in document directory under the appropriate Working Group directory
Place into the validate directory under the release number for the KB
KB Integrator Coordinator contacts SNL Validator for product
SNL Validator reviews and validates the Files and Installation Testing Procedures document
Fail
Contact Responsible Lab about Files and Testing Procedures
Success
Update Product Tracking Matrix on KB website
Product Arrives at SNL
Update Product Tracking Matrix on KB website
KB Integrator Coordinator puts product into appropriate directory
Unclassified
Classified
Place in product directory under the appropriate Working Group
Place in product directory based on Product Number
Load database tables into the appropriate database account
Unclassified
Classified
Place in appropriate "labdata" database account
Place in the rpval database account
Prepare/setup any config/par files needed to test the product
Flows Into
Build/Install SNL Tool Root



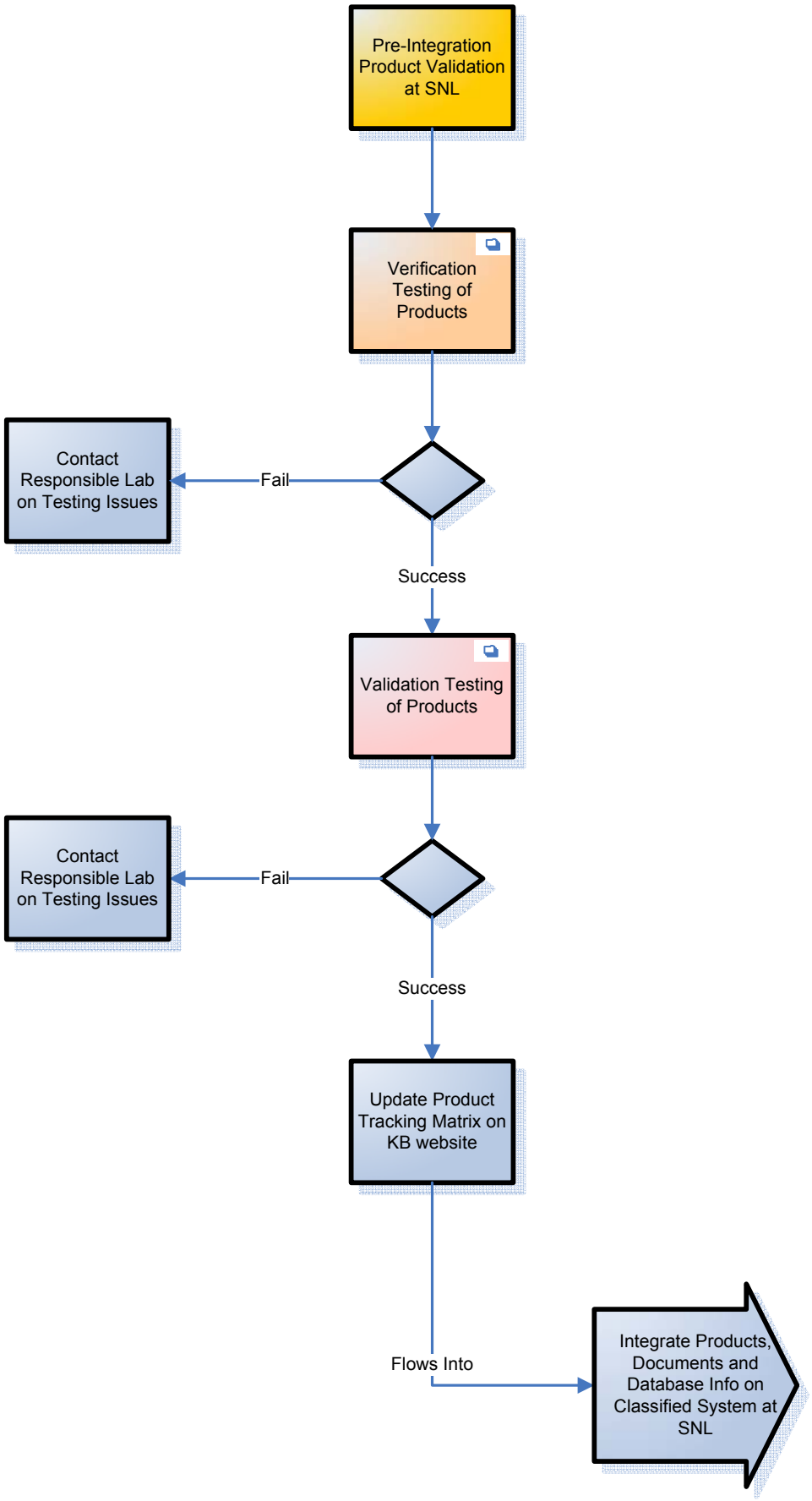
Pre-Validaton Preparation at SNL
Integration Coordination



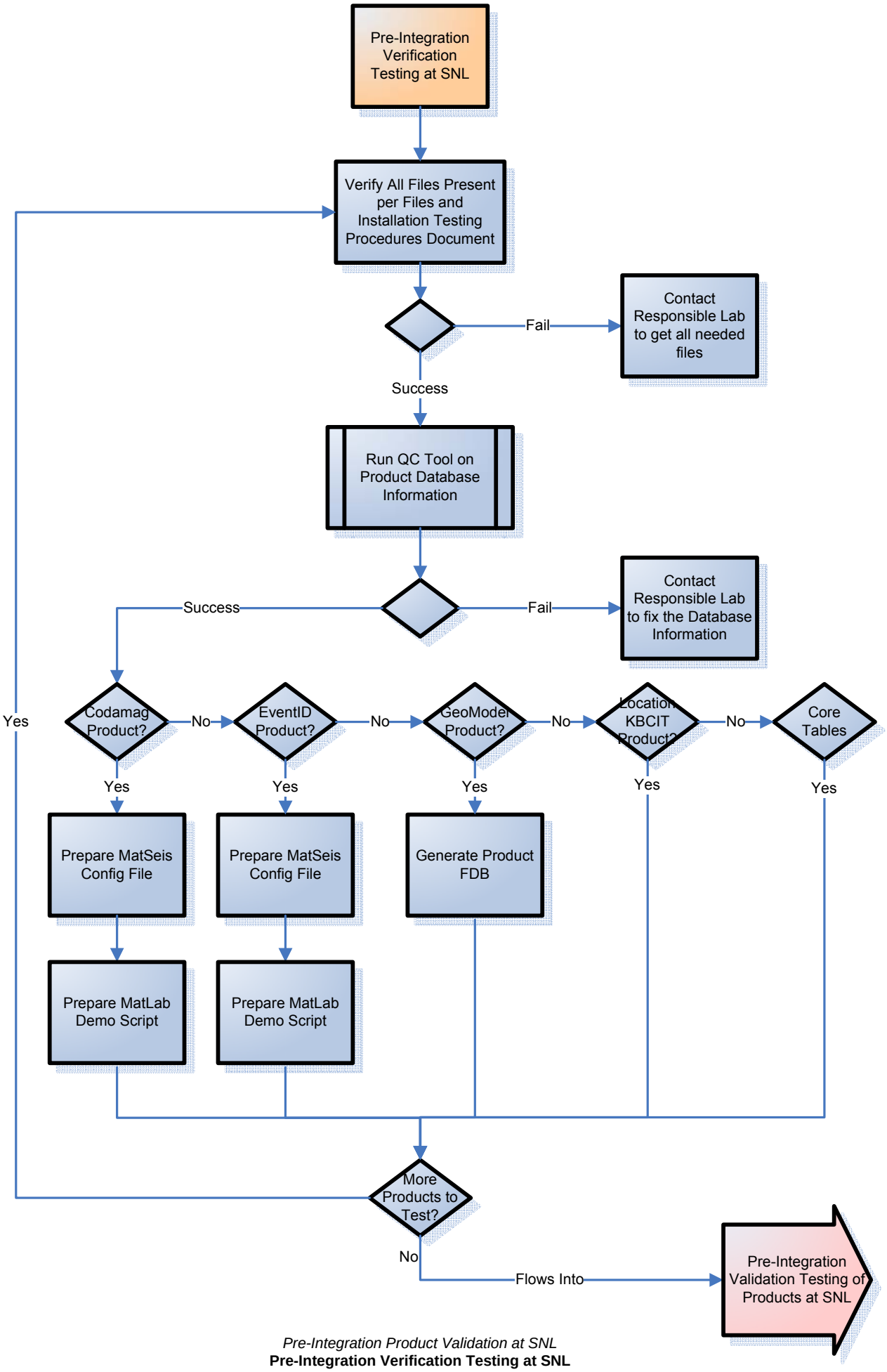
Pre-Validaton Preparation at AFTAC



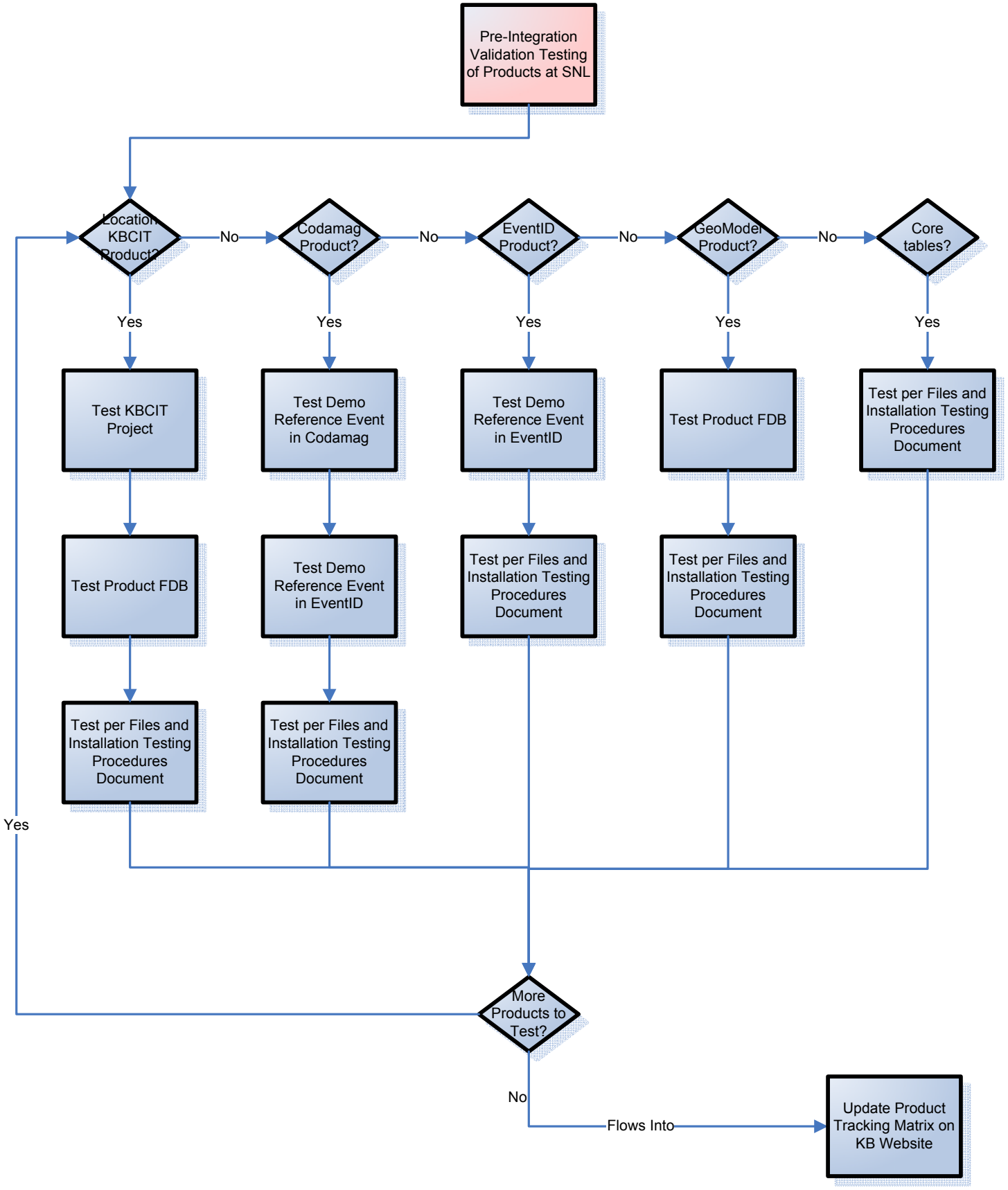
Build SNL Tool Root and Install at SNL and AFTAC

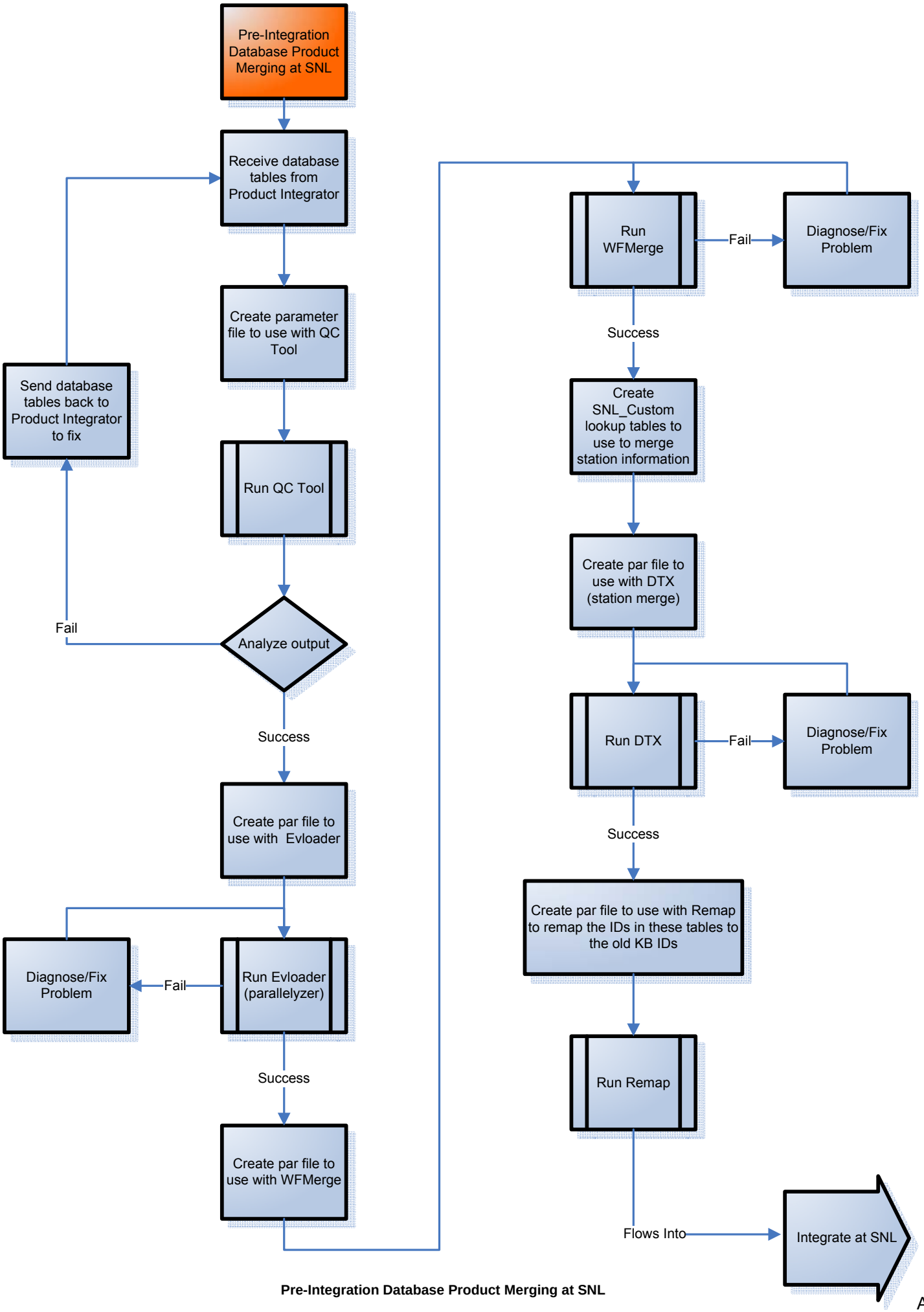


Pre-Integration Product Validation at SNL

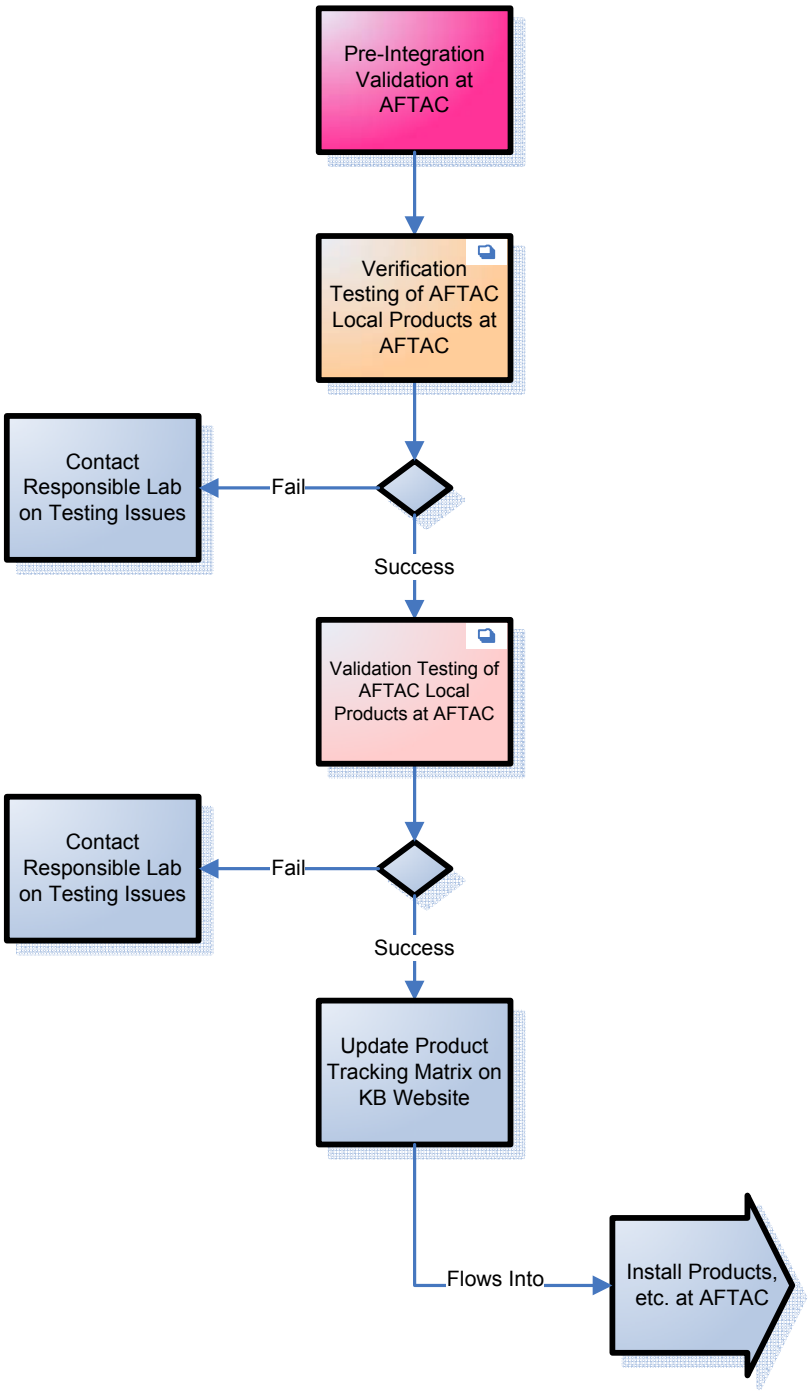


Pre-Integration Product Validation at SNL
Pre-Integration Verification Testing at SNL

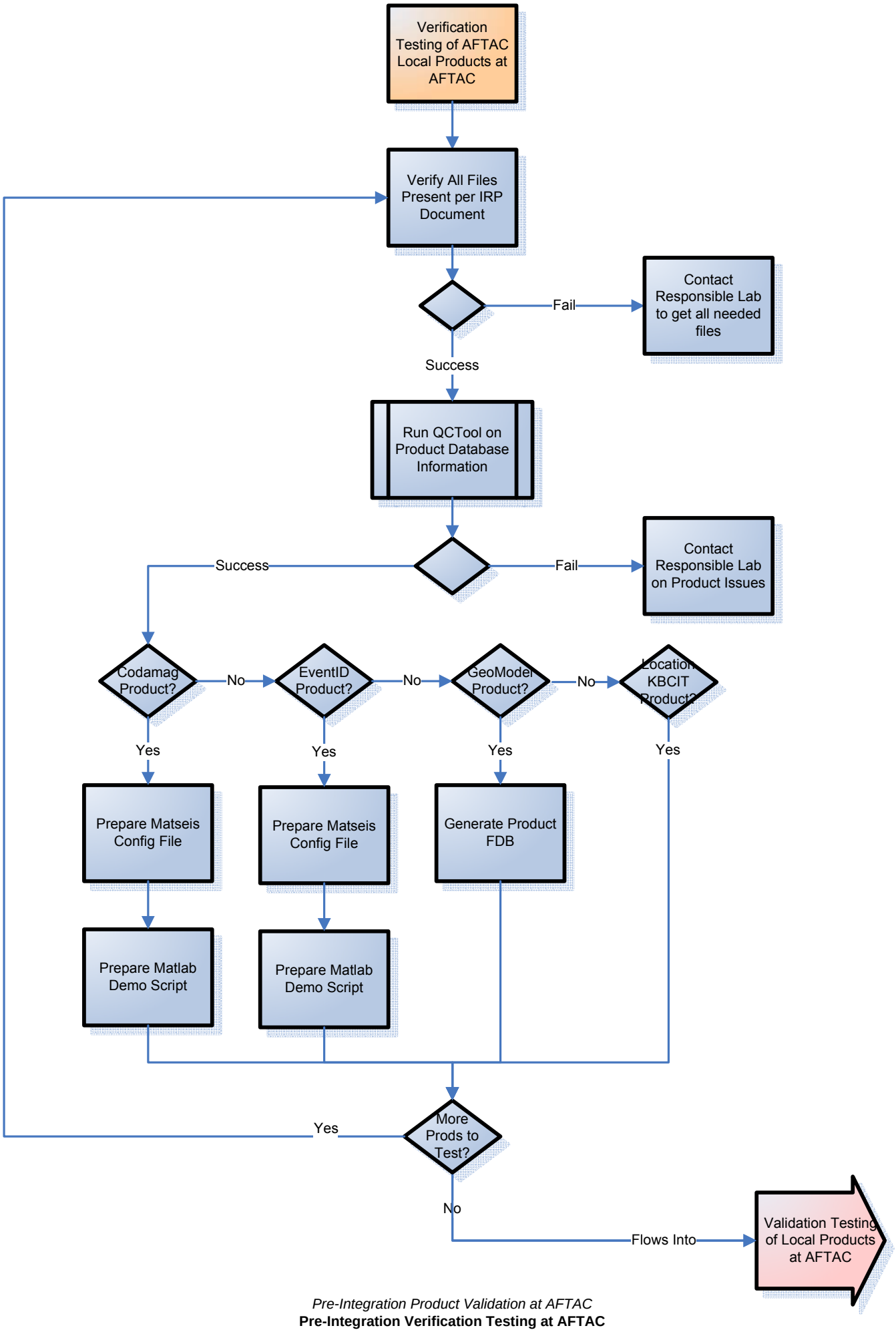

Pre-Integration Validation Testing of Products at SNL
Location KBCIT Product?
No
Codamag Product?
No
EventID Product?
No
GeoModel Product?
No
Core tables?
Yes
Yes
Yes
Yes
Yes
Test KBCIT Project
Test Demo Reference Event in Codamag
Test Demo Reference Event in EventID
Test Product FDB
Test per Files and Installation Testing Procedures Document
Test Product FDB
Test Demo Reference Event in EventID
Test per Files and Installation Testing Procedures Document
Test per Files and Installation Testing Procedures Document
Test per Files and Installation Testing Procedures Document
Test per Files and Installation Testing Procedures Document
Yes
More Products to Test?
No
Flows Into
Update Product Tracking Matrix on KB Website

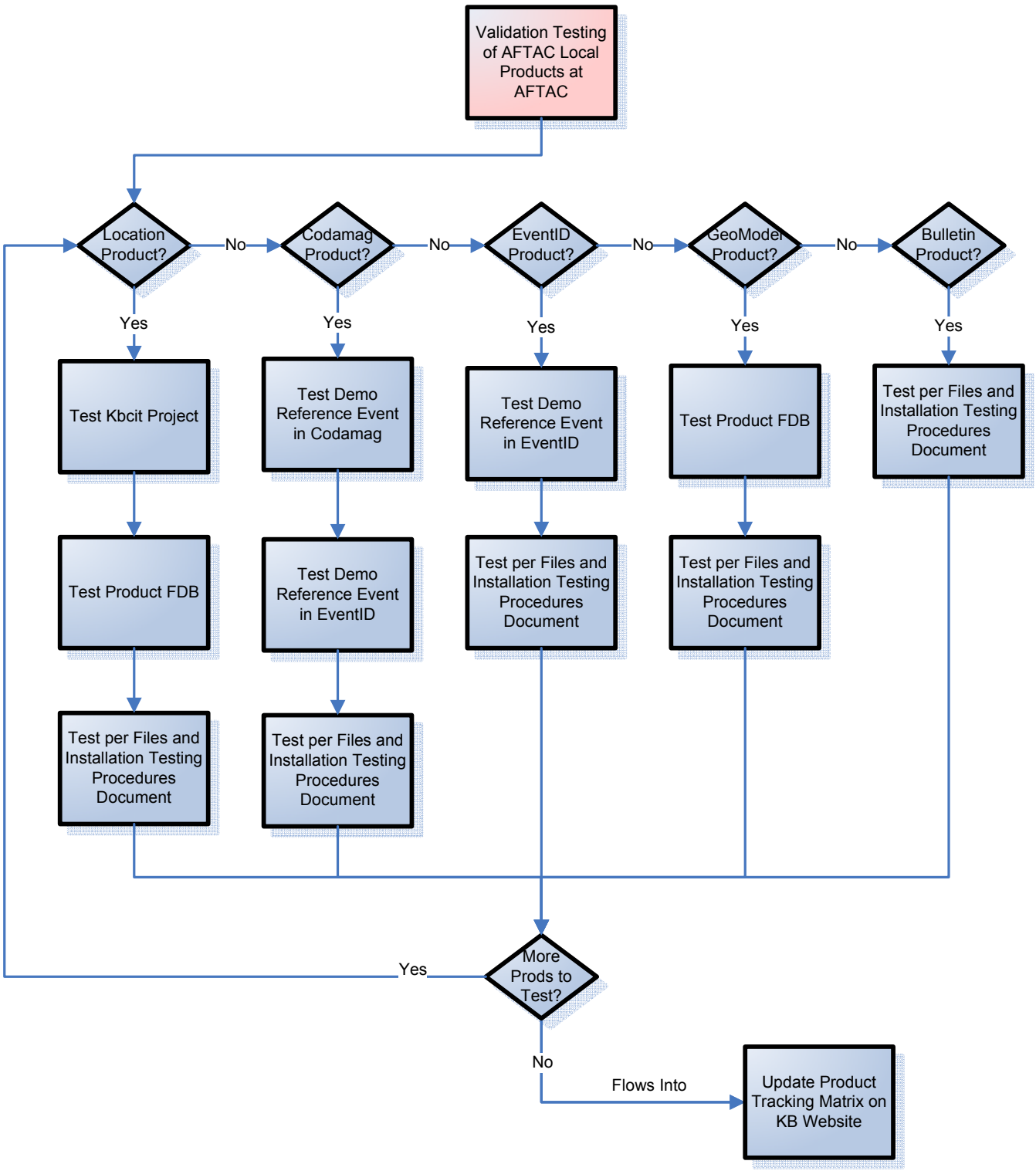


Pre-Integration Database Product Merging at SNL

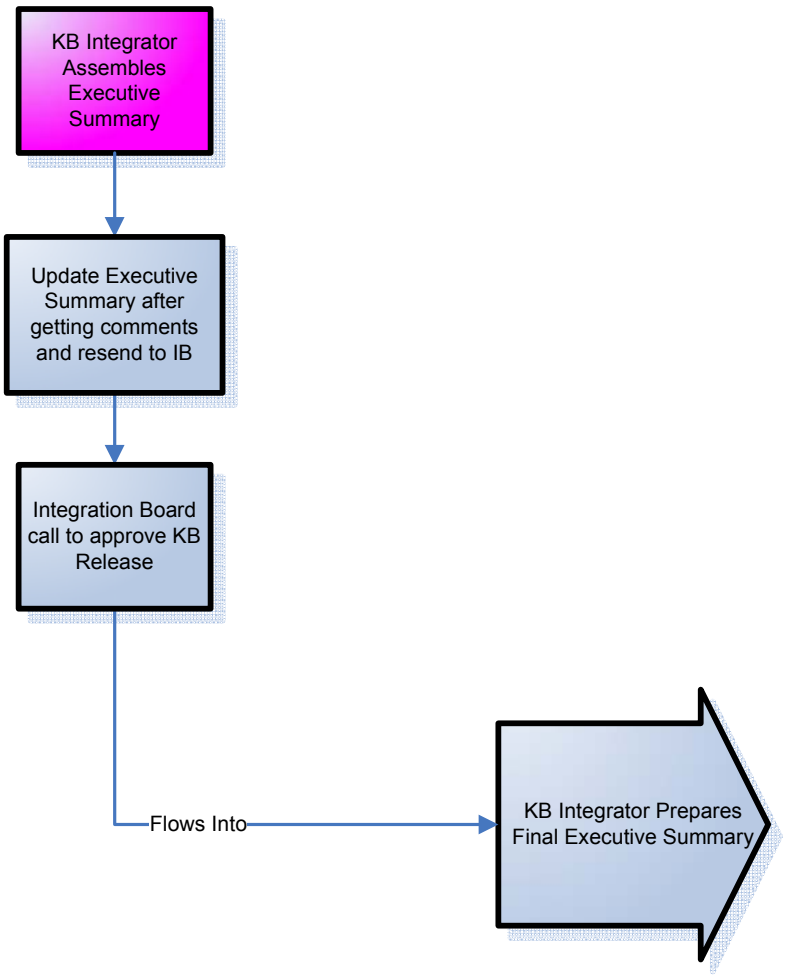


Pre-Integration Product Validation at AFTAC

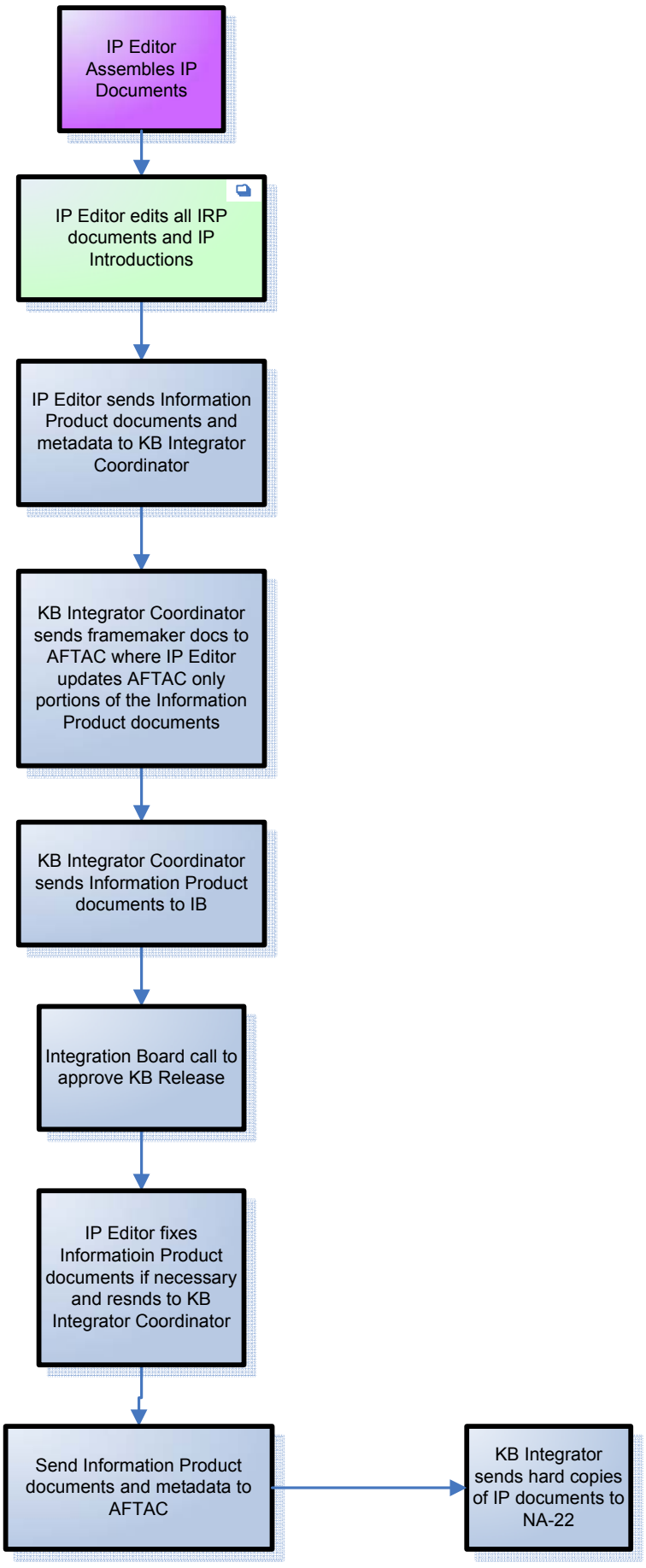

Verification Testing of AFTAC Local Products at AFTAC
Verify All Files Present per IRP Document
Fail
Contact Responsible Lab to get all needed files
Success
Run QCTool on Product Database Information
Success
Fail
Contact Responsible Lab on Product Issues
Codamag Product?
No
EventID Product?
No
GeoModel Product?
No
Location KBCIT Product?
Yes
Yes
Yes
Yes
Prepare Matseis Config File
Prepare Matseis Config File
Generate Product FDB
Prepare Matlab Demo Script
Prepare Matlab Demo Script
Yes
More Prods to Test?
No
Flows Into
Validation Testing of Local Products at AFTAC



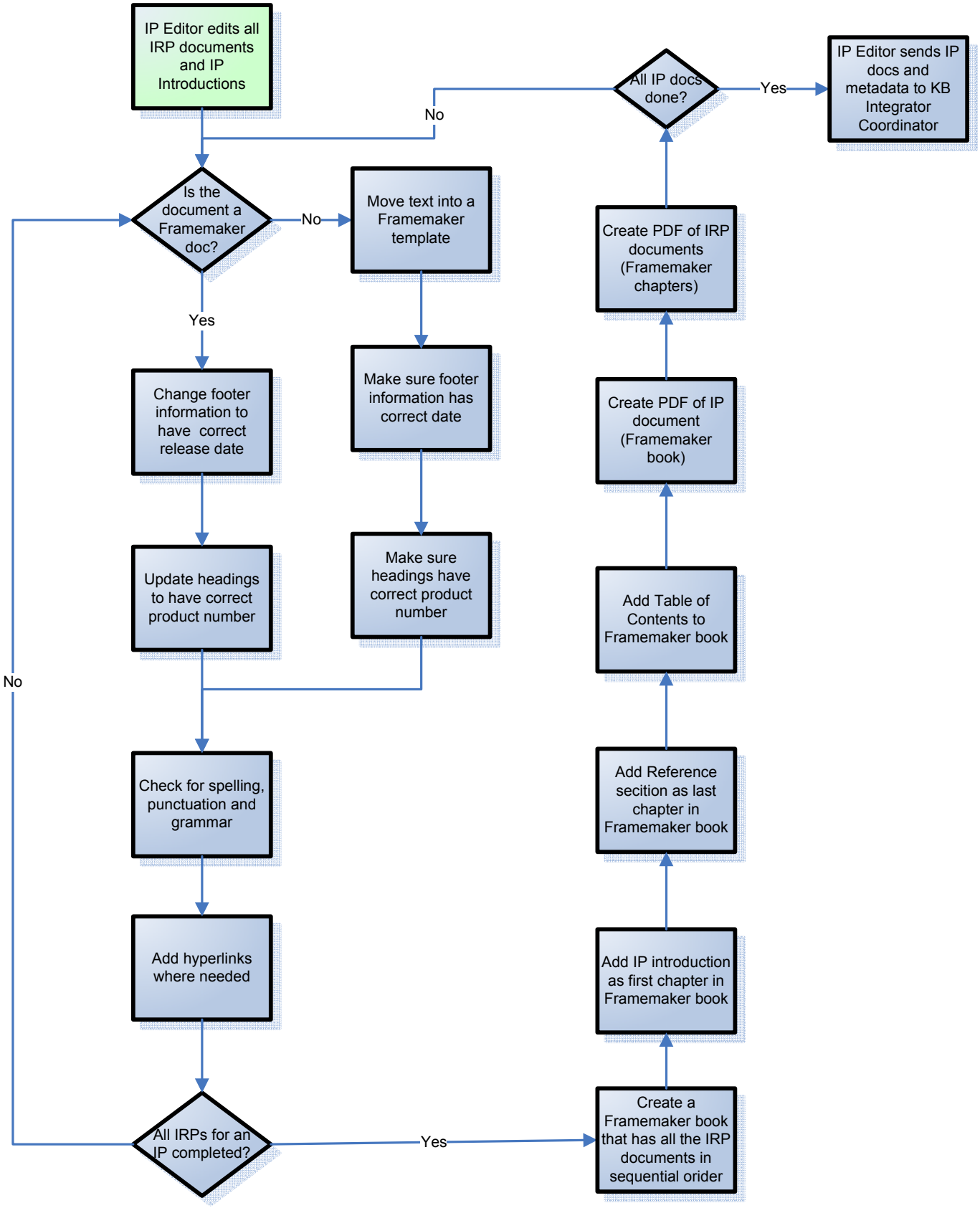
Pre-Integration Product Validation at AFTAC
Pre-Integration Validation Testing at AFTAC



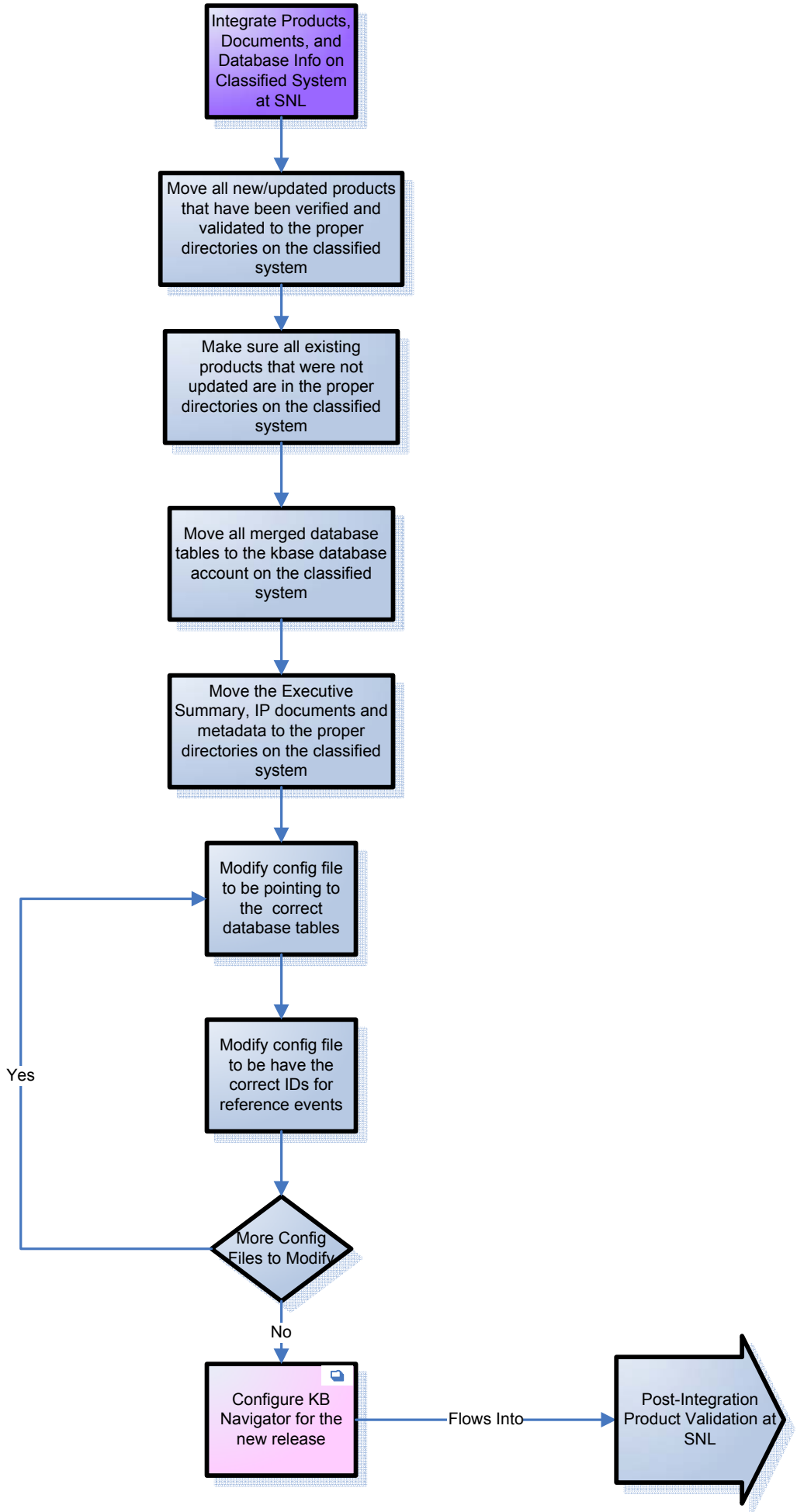
KB Integrator Assembles Executive Summary



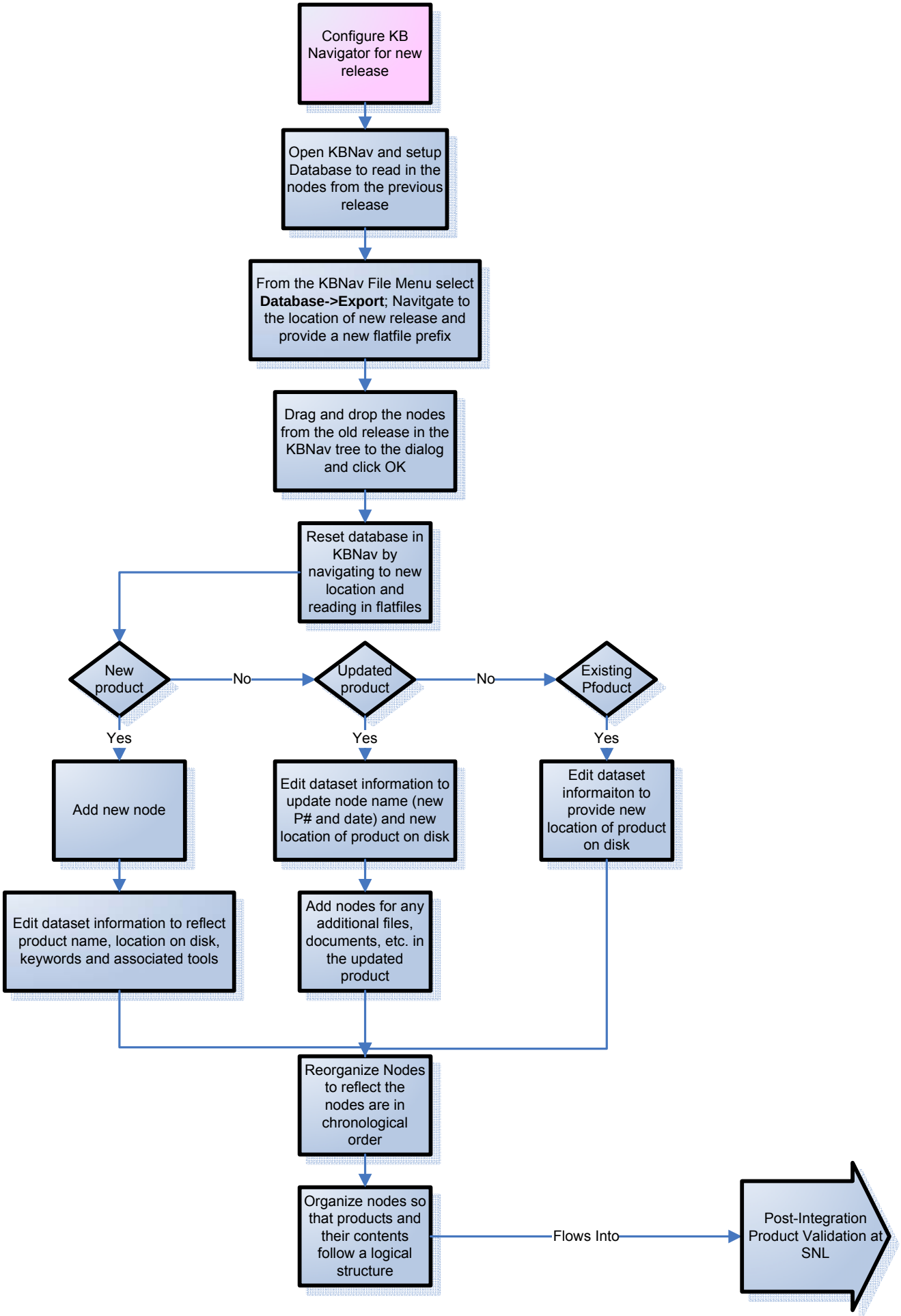
IP Editor Assembles Information Product Documents

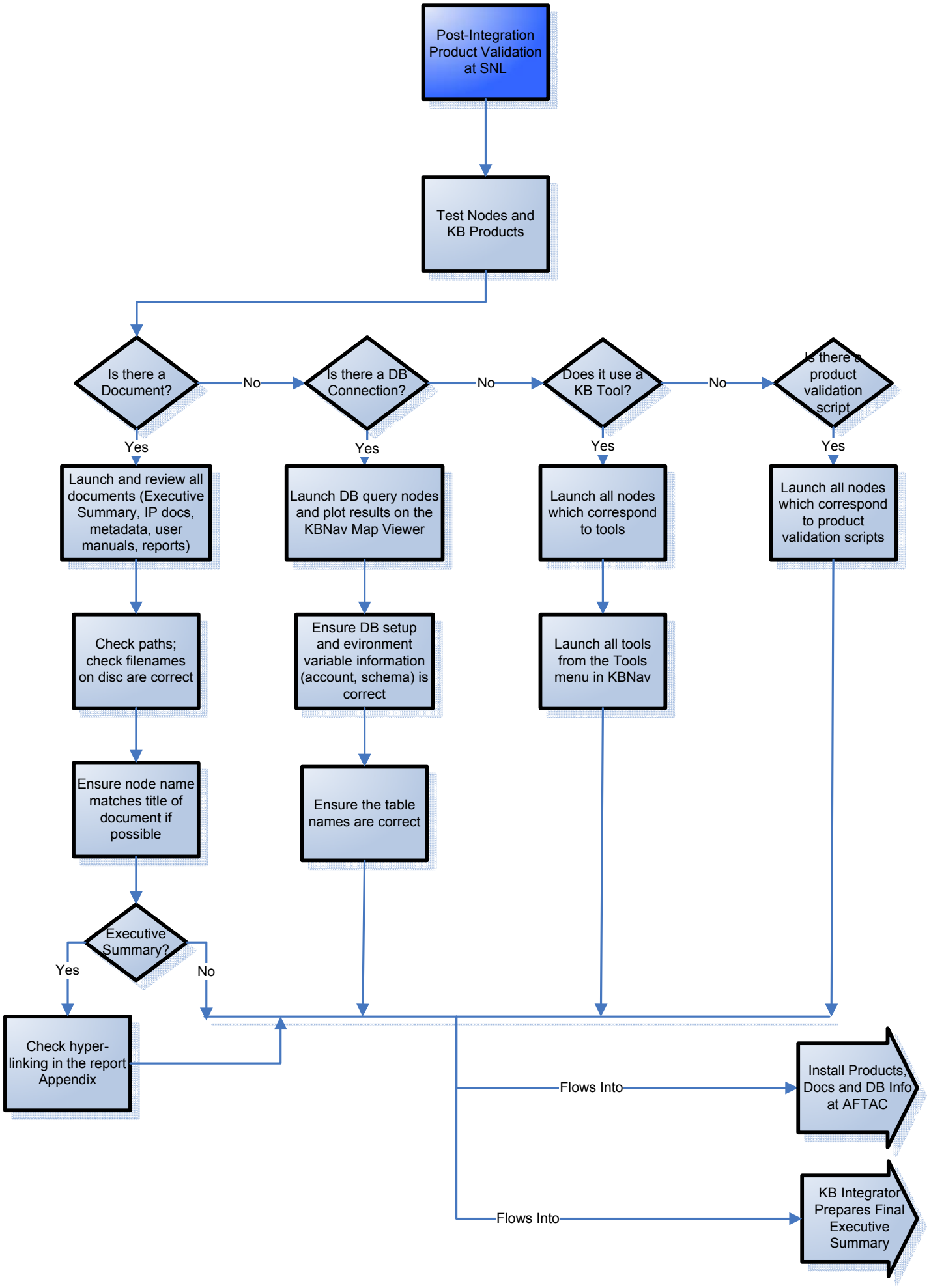


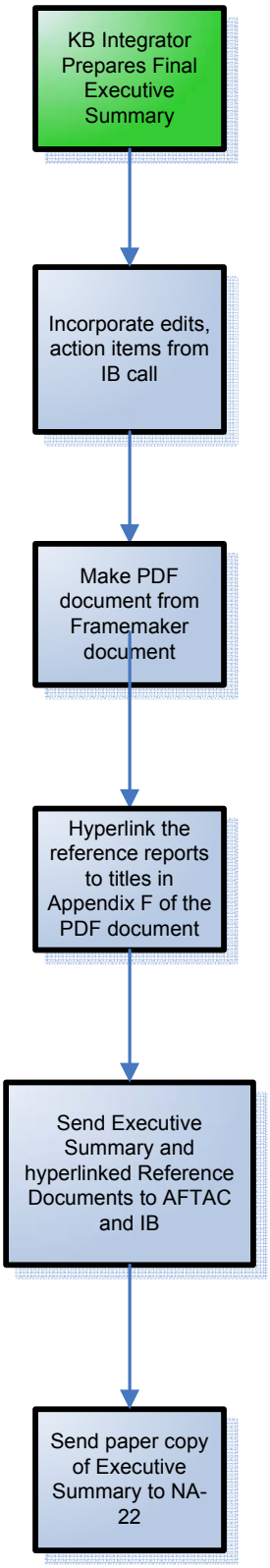
IP Editor Assembles IP Documents
IP Editor edits IRP documents and IP Introductions



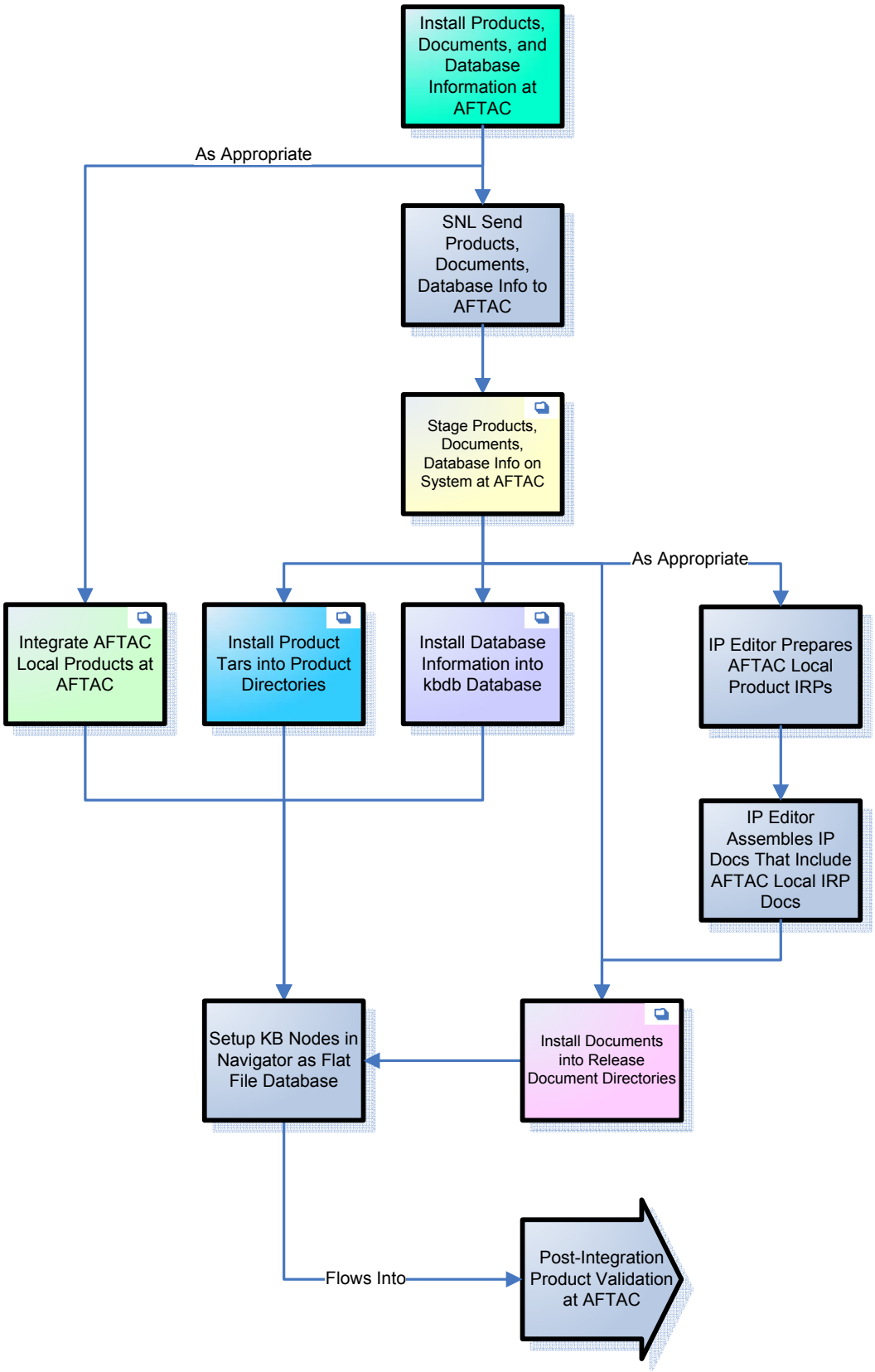
Integrate Products, Documents and Database Info on classified system at SNL


Configure KB Navigator for new release
Open KBNav and setup Database to read in the nodes from the previous release
From the KBNav File Menu select **Database->Export**; Navitgate to the location of new release and provide a new flatfile prefix
Drag and drop the nodes from the old release in the KBNav tree to the dialog and click OK
Reset database in KBNav by navigating to new location and reading in flatfiles
New product
No
Updated product
No
Existing Pfoduct
Yes
Yes
Yes
Add new node
Edit dataset information to update node name (new P# and date) and new location of product on disk
Edit dataset informaiton to provide new location of product on disk
Edit dataset information to reflect product name, location on disk, keywords and associated tools
Add nodes for any additional files, documents, etc. in the updated product
Reorganize Nodes to reflect the nodes are in chronological order
Organize nodes so that products and their contents follow a logical structure
Flows Into
Post-Integration Product Validation at SNL

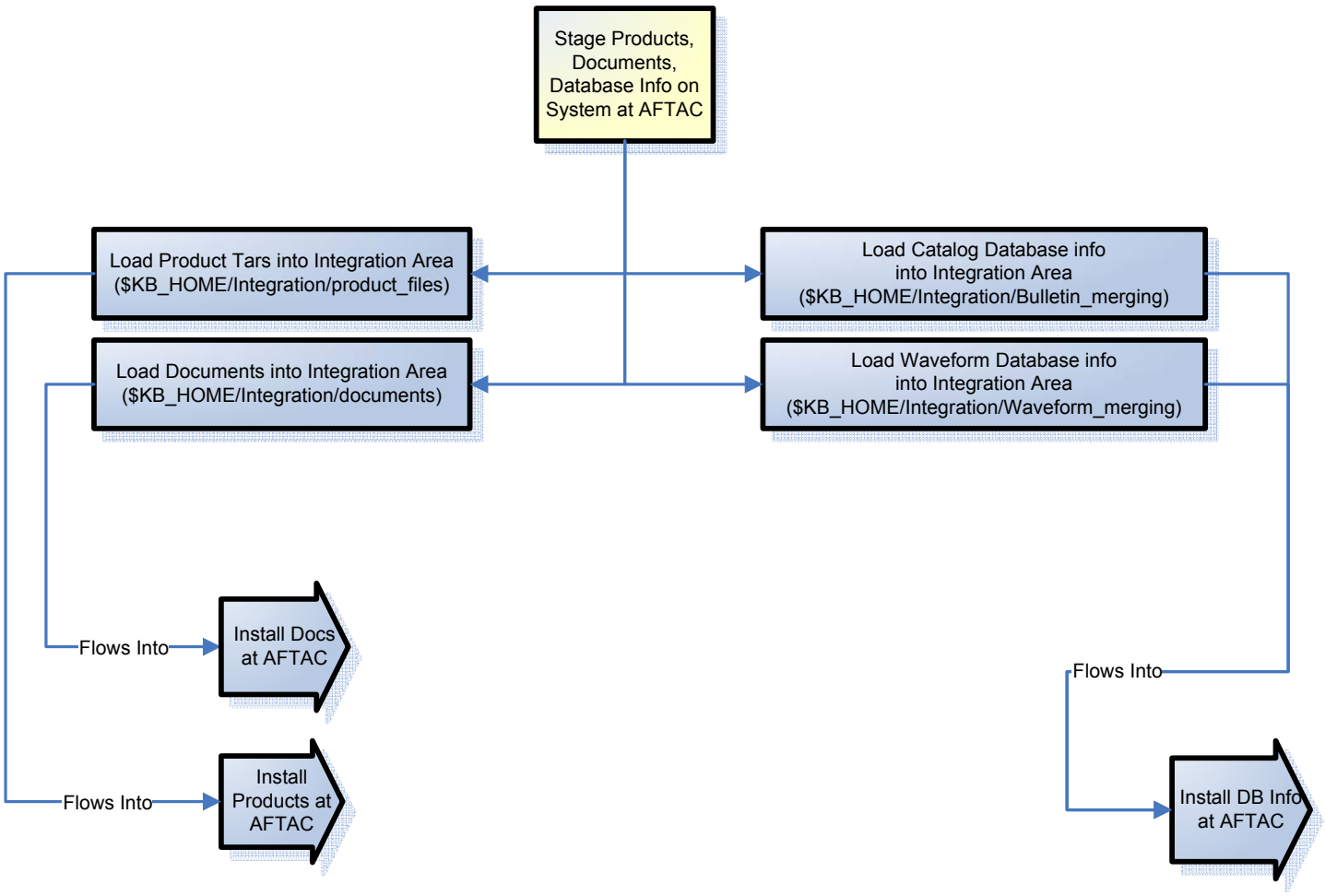

Post-Integration Product Validation at SNL
Test Nodes and KB Products
Is there a Document?
No
Is there a DB Connection?
No
Does it use a KB Tool?
No
Is there a product validation script
Yes
Launch and review all documents (Executive Summary, IP docs, metadata, user manuals, reports)
Check paths; check filenames on disc are correct
Ensure node name matches title of document if possible
Executive Summary?
Yes
No
Check hyper-linking in the report Appendix
Yes
Launch DB query nodes and plot results on the KBNav Map Viewer
Ensure DB setup and evironment variable information (account, schema) is correct
Ensure the table names are correct
Yes
Launch all nodes which correspond to tools
Launch all tools from the Tools menu in KBNav
Yes
Launch all nodes which correspond to product validation scripts
Flows Into
Install Products, Docs and DB Info at AFTAC
Flows Into
KB Integrator Prepares Final Executive Summary



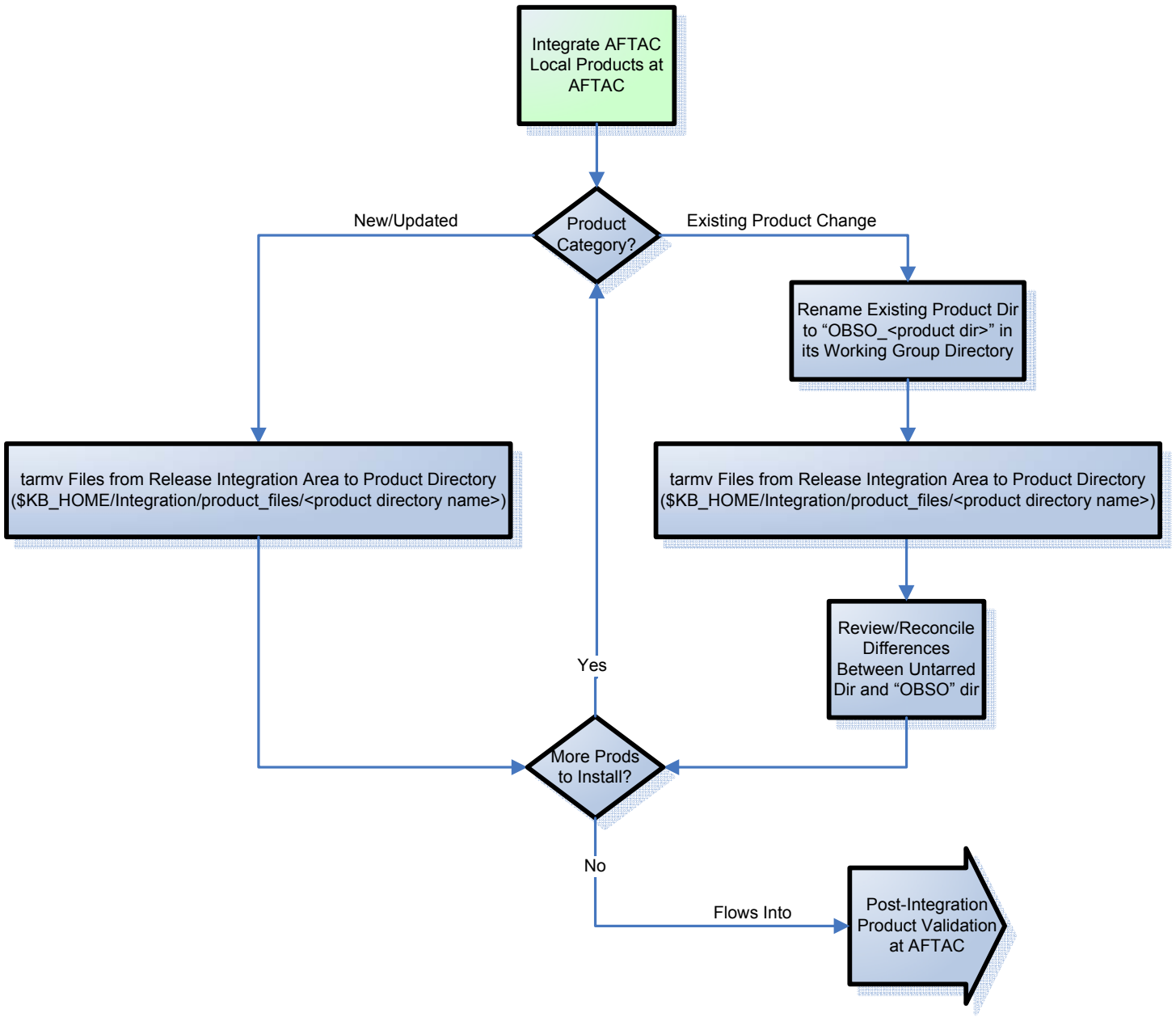
KB Integrator Prepares Final Executive Summary



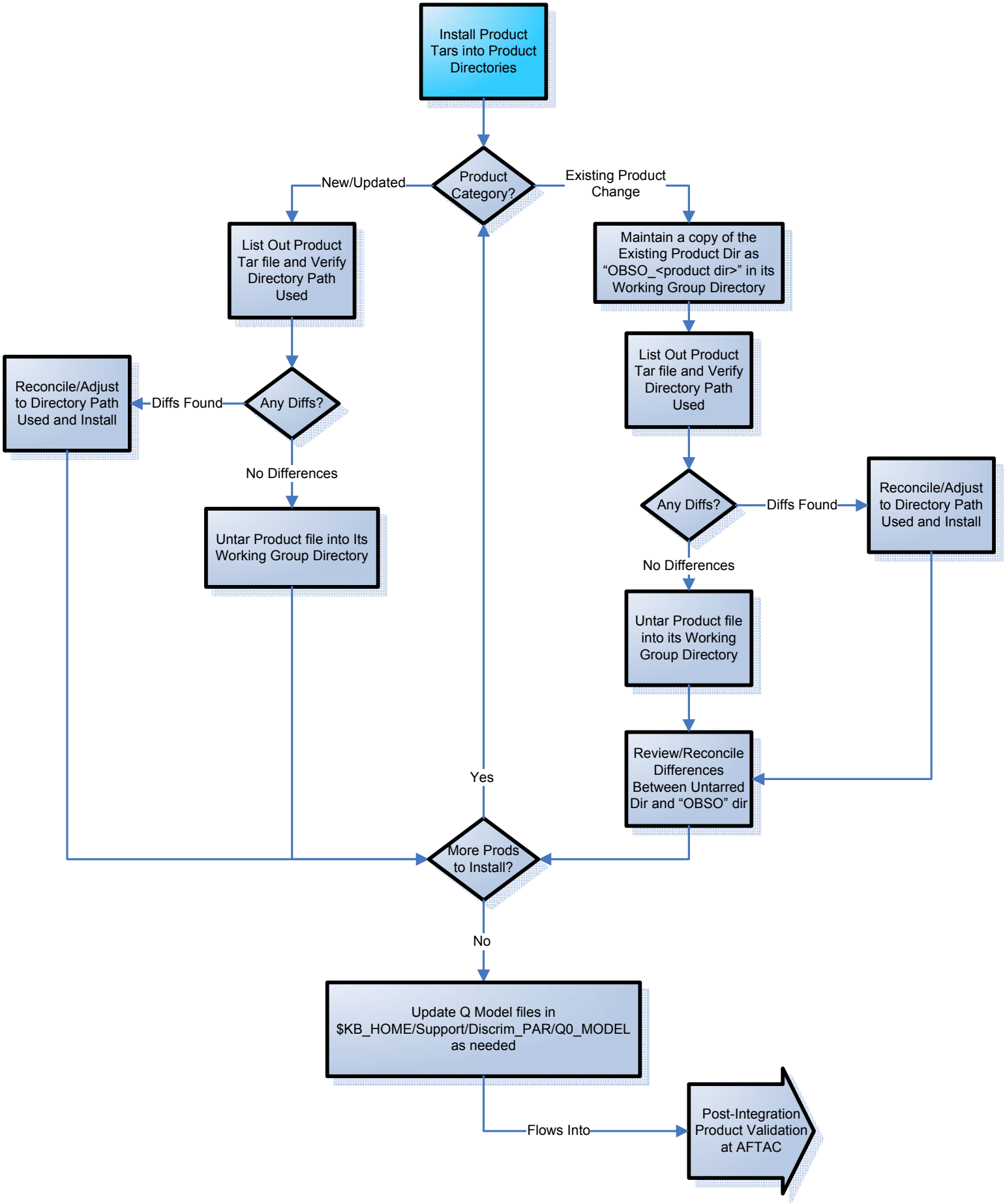
Install Products, Documents and Database Information at AFTAC

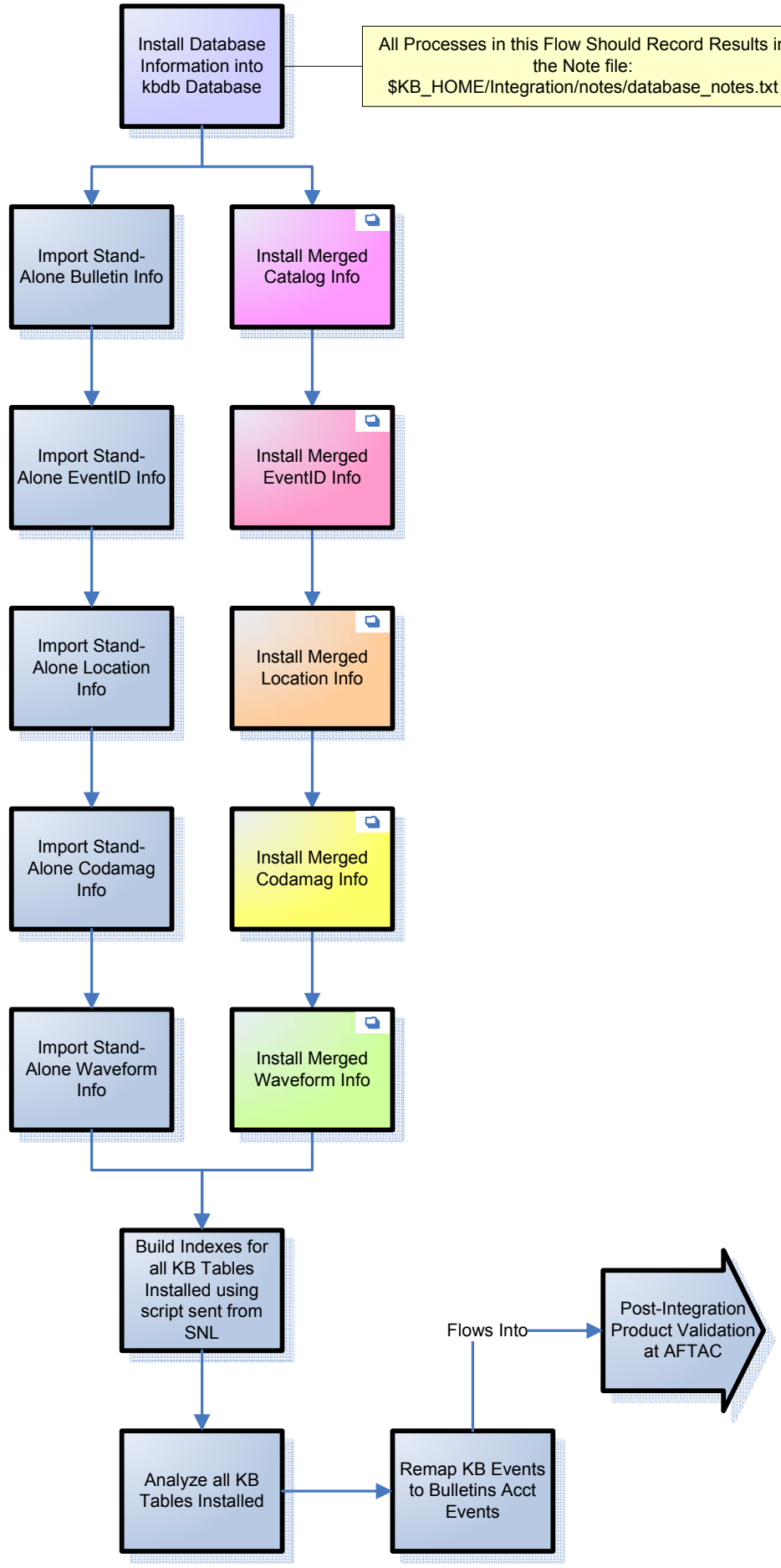


Install Products, Documents and Database Information at AFTAC
Stage Products, Documents, Database Info

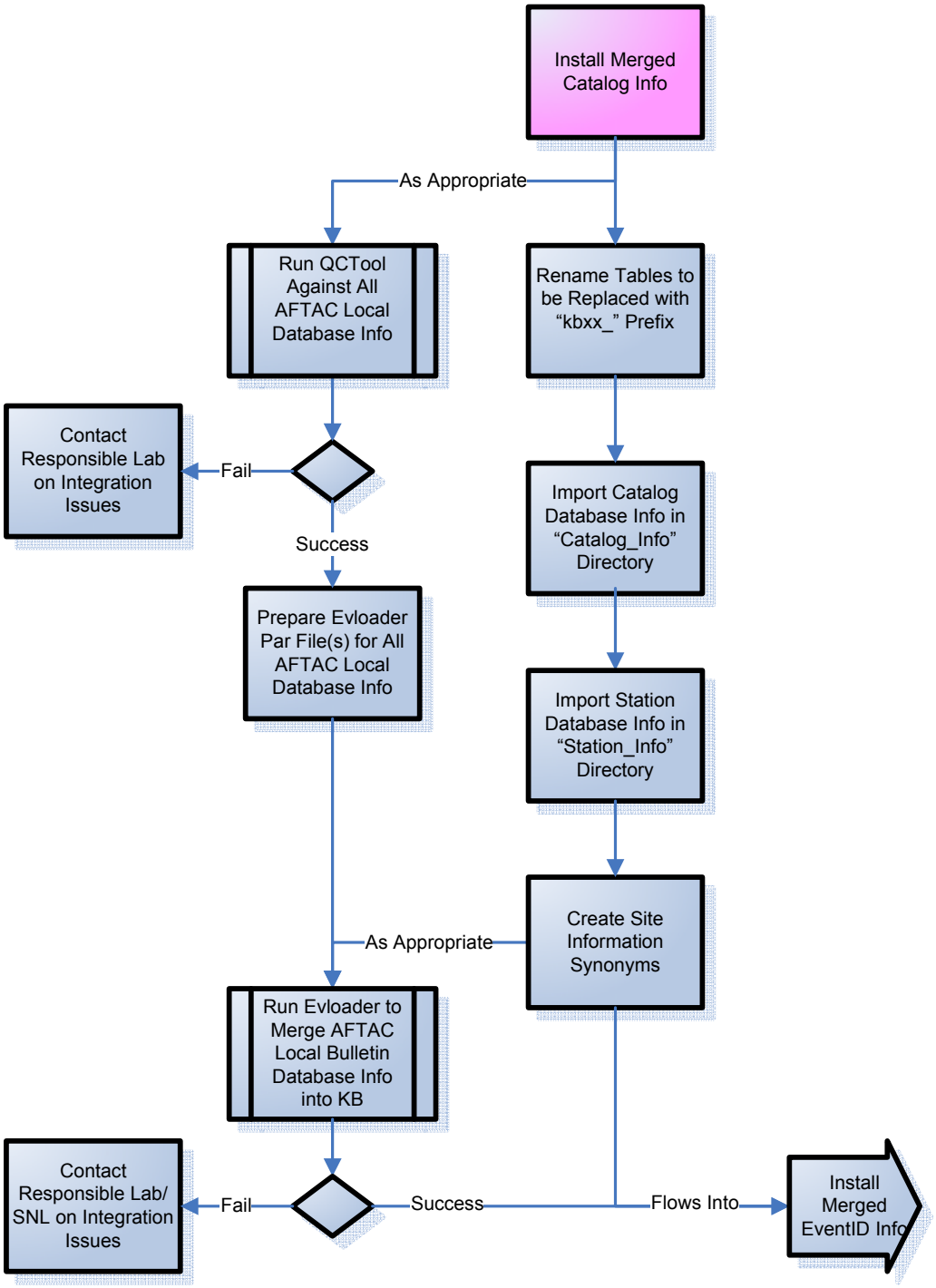


Install Products, Documents and Database Information at AFTAC
Integrate AFTAC Local Products at AFTAC

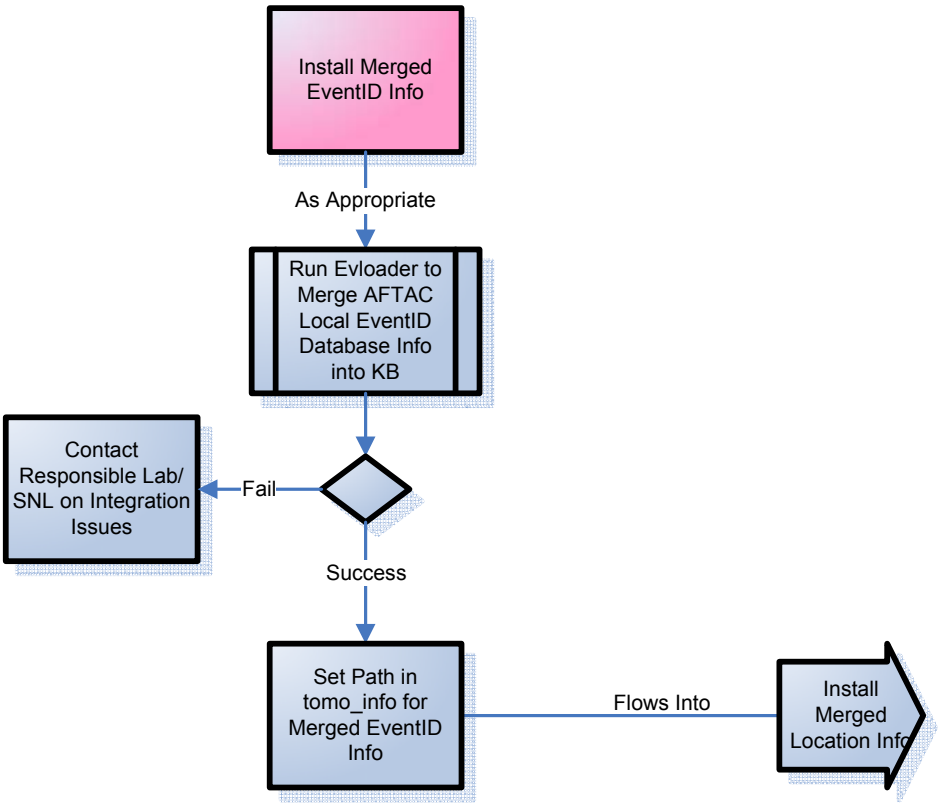

Install Product Tars into Product Directories
Product Category?
New/Updated
Existing Product Change
List Out Product Tar file and Verify Directory Path Used
Maintain a copy of the Existing Product Dir as "OBSO_<product dir>" in its Working Group Directory
Reconcile/Adjust to Directory Path Used and Install
Diffs Found
Any Diffs?
No Differences
Untar Product file into Its Working Group Directory
List Out Product Tar file and Verify Directory Path Used
Any Diffs?
Diffs Found
Reconcile/Adjust to Directory Path Used and Install
No Differences
Untar Product file into its Working Group Directory
Review/Reconcile Differences Between Untarred Dir and "OBSO" dir
Yes
More Prods to Install?
No
Update Q Model files in $KB_HOME/Support/Discrim_PAR/Q0_MODEL as needed
Flows Into
Post-Integration Product Validation at AFTAC



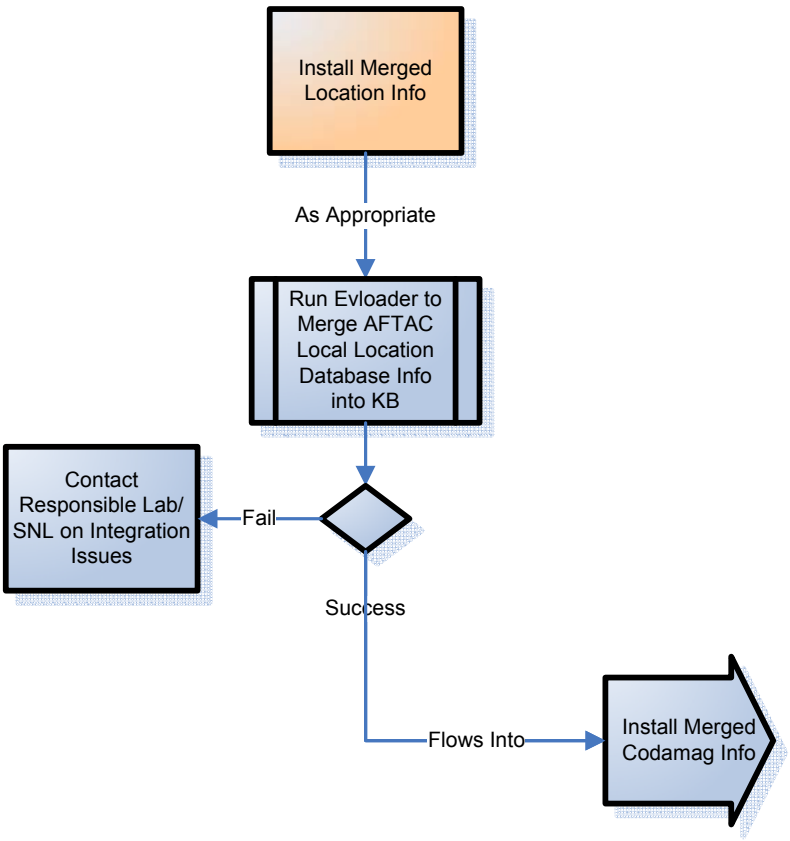
Install Products, Documents and Database Information at AFTAC
Install Database Information into kbdb Database



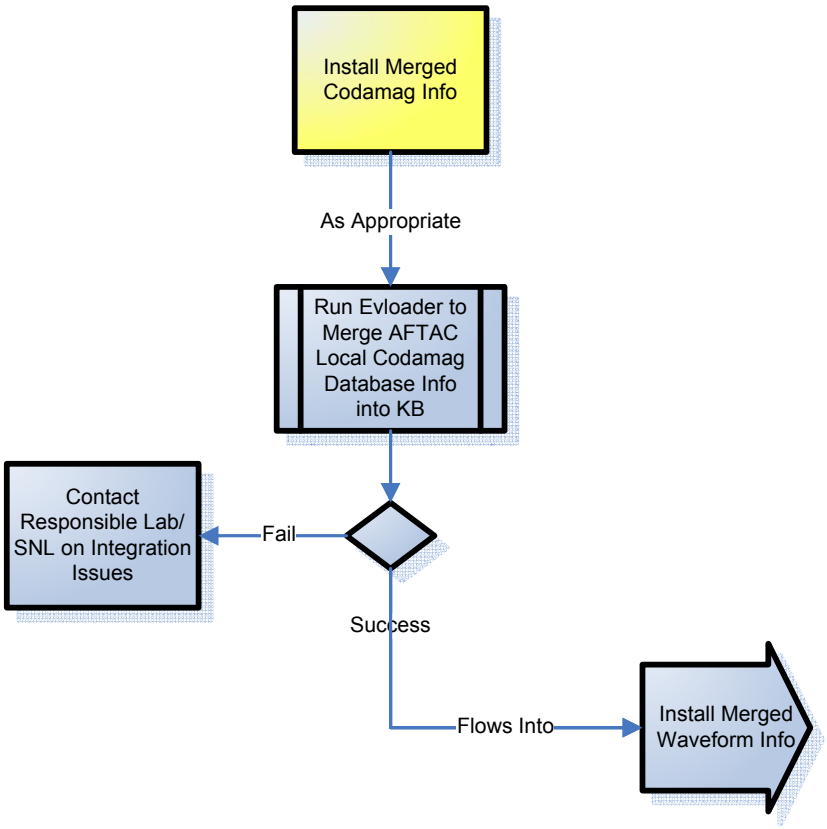
Install Products, Documents and Database Information at AFTAC
Install Database Information into kbdb Database
Install Merged Catalog Info



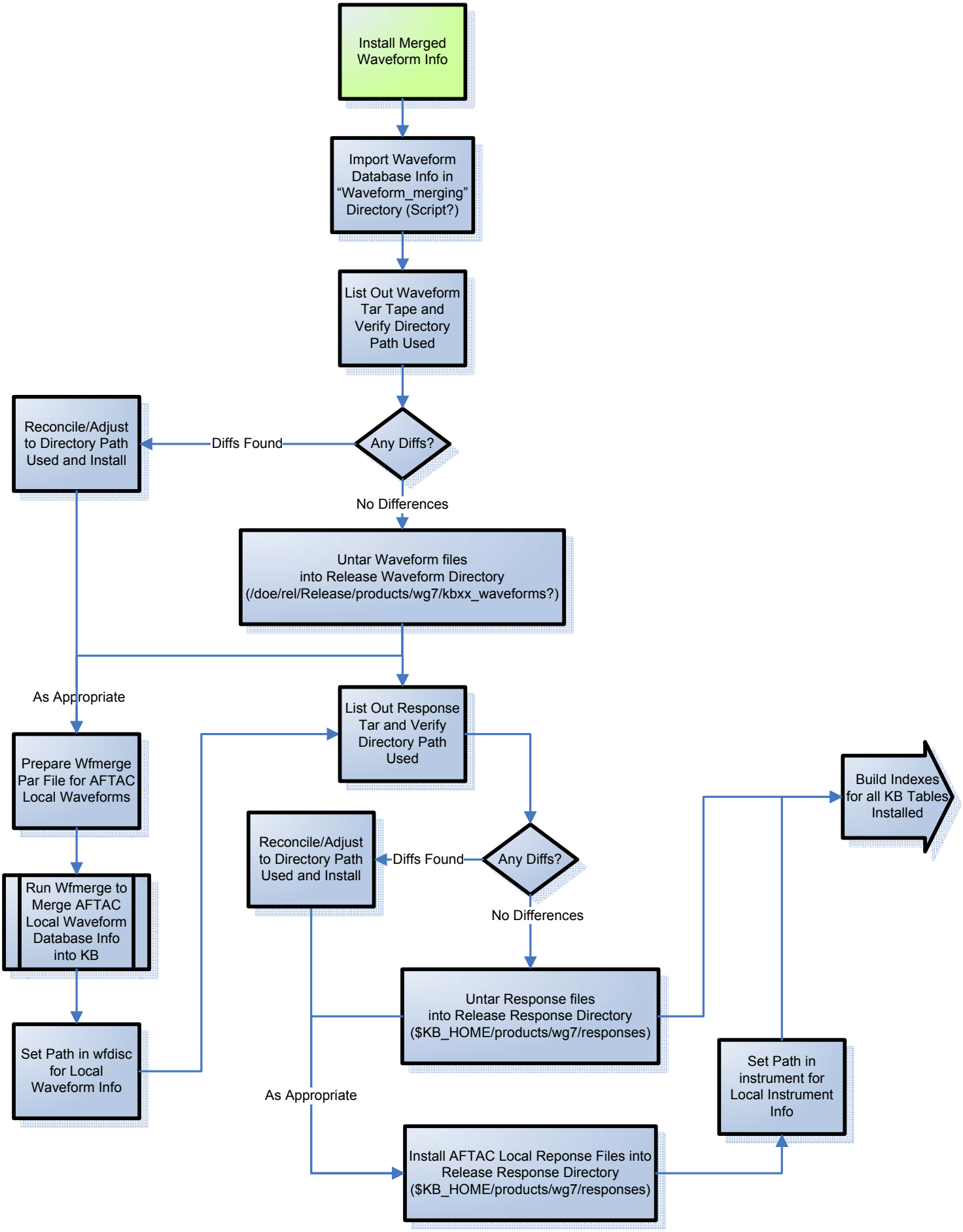
Install Products, Documents and Database Information at AFTAC
Install Database Information into kbdb Database
Install Merged EventID Info



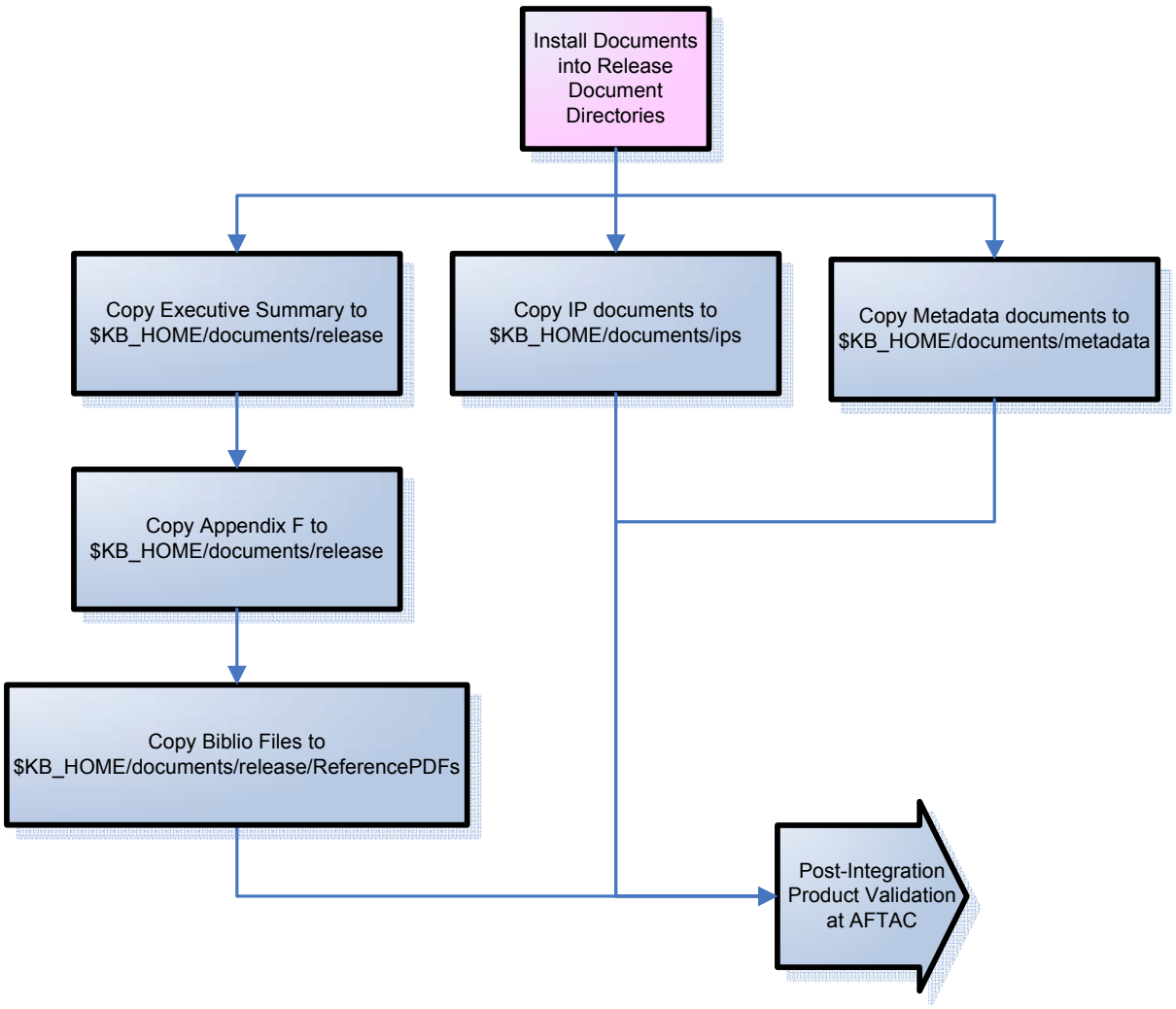
Install Products, Documents and Database Information at AFTAC
Install Database Information into kbdb Database
Install Merged Location Info



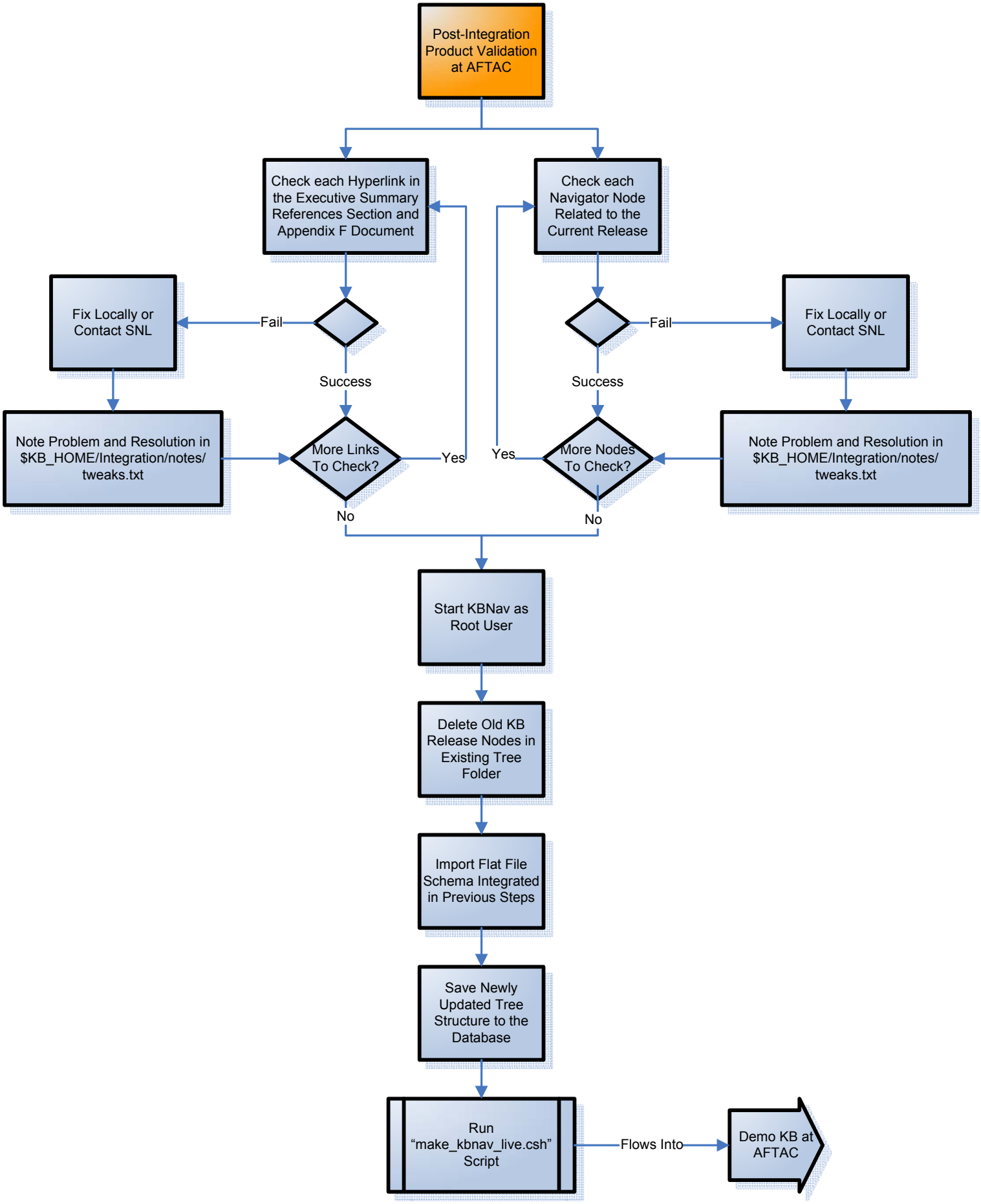
Install Products, Documents and Database Information at AFTAC
Install Database Information into kbdb Database
Install Merged Codamag Info

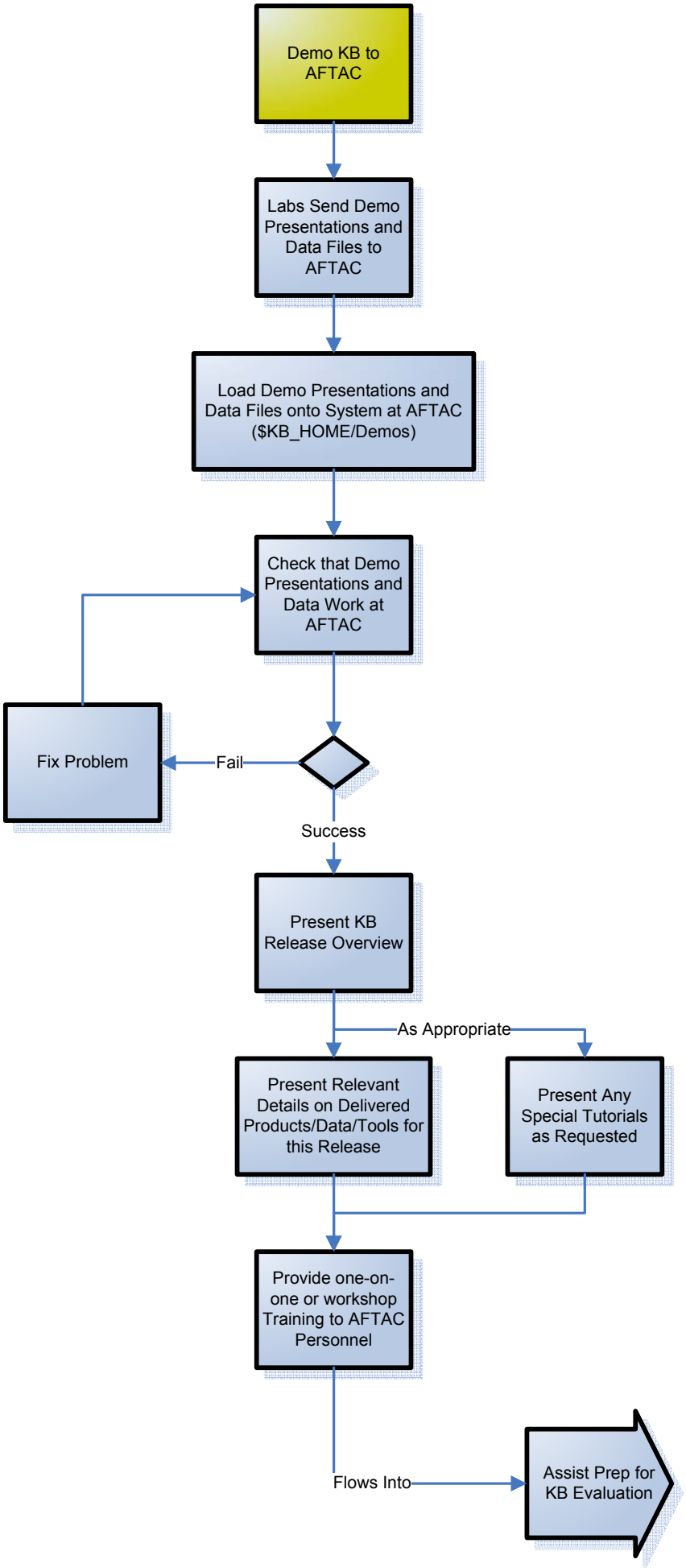


Install Products, Documents and Database Information at AFTAC
Install Database Information into kbdb Database
Install Merged Waveform Info

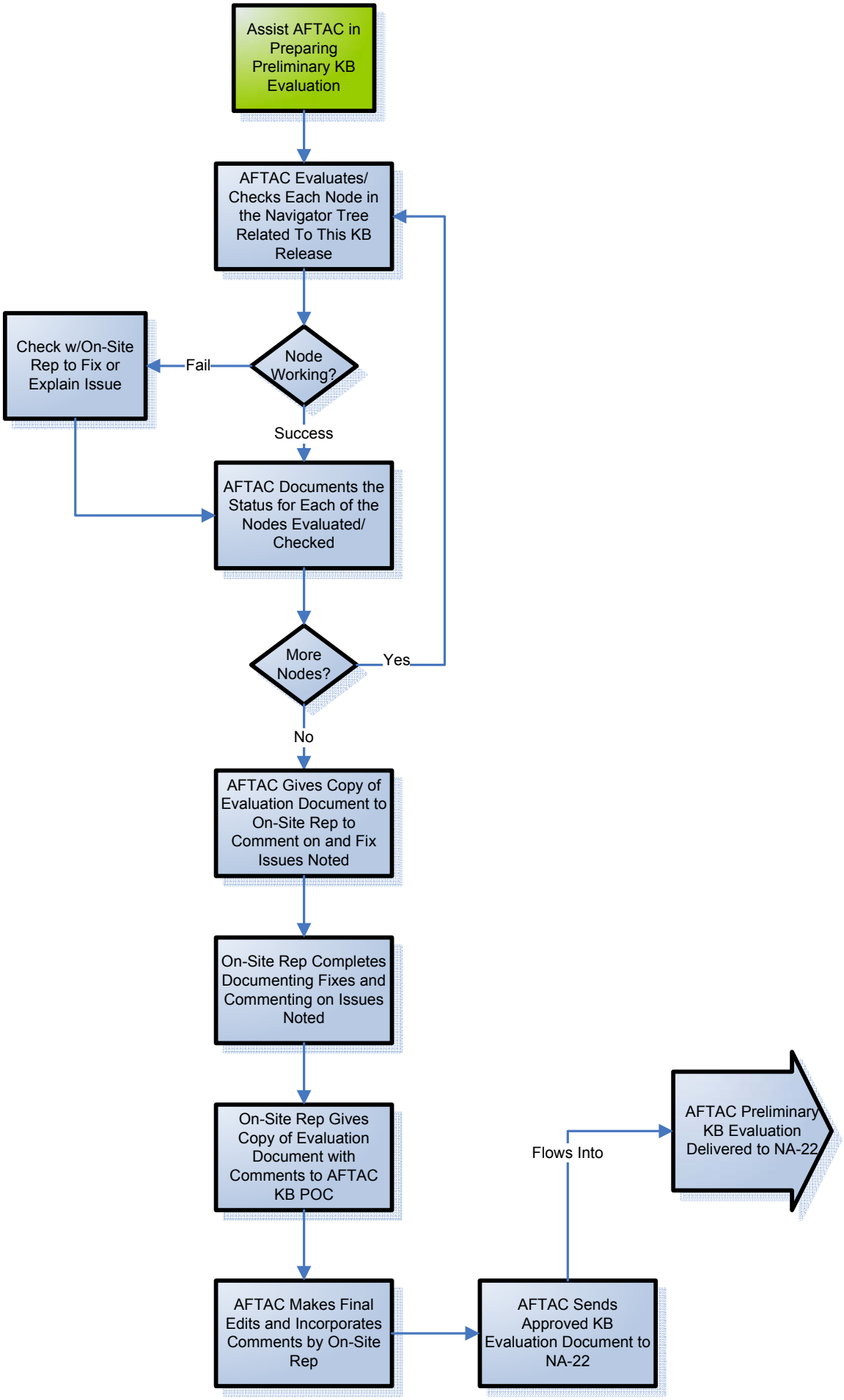


Install Products, Documents and Database Information at AFTAC
Install Documents into Release Document Directories

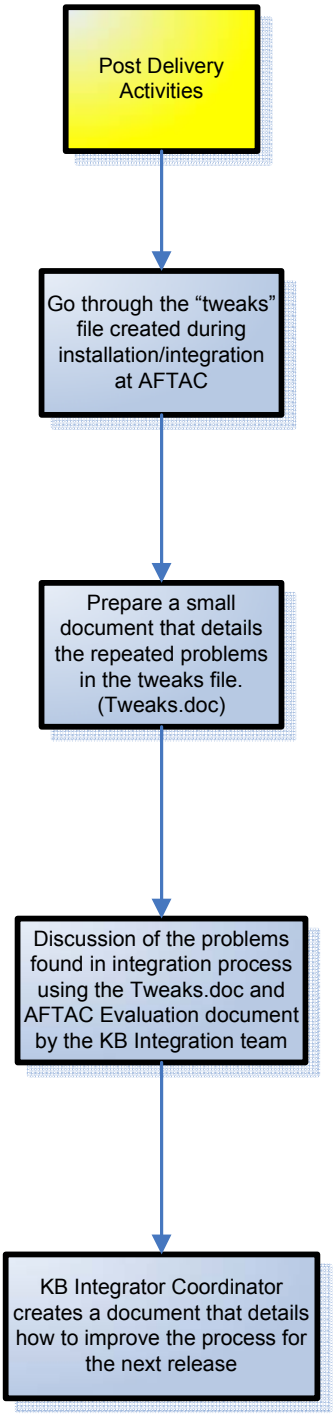

Post-Integration Product Validation at AFTAC
Check each Hyperlink in the Executive Summary References Section and Appendix F Document
Check each Navigator Node Related to the Current Release
Fix Locally or Contact SNL
Fail
Success
Fail
Success
Fix Locally or Contact SNL
Note Problem and Resolution in $KB_HOME/Integration/notes/ tweaks.txt
More Links To Check?
Yes
No
Yes
More Nodes To Check?
No
Note Problem and Resolution in $KB_HOME/Integration/notes/ tweaks.txt
Start KBNav as Root User
Delete Old KB Release Nodes in Existing Tree Folder
Import Flat File Schema Integrated in Previous Steps
Save Newly Updated Tree Structure to the Database
Run "make_kbnav_live.csh" Script
Flows Into
Demo KB at AFTAC



Demo KB to AFTAC



Assist AFTAC in Preparing Preliminary KB Evlauation



Post Delivery Activites

DISTRIBUTION

1 Leslie A. Casey
NNSA Office of Nonproliferation Research and Development/NA-22
1000 Independence Avenue SW
Washington, DC 20585

1 John Dwyer
Air Force Technical Applications Center/TTR
1030 S. Hwy A1A
Patrick AFB, FL 32925-3002

1 Mark Woods
Air Force Technical Applications Center/TTR
1030 S. Hwy A1A
Patrick AFB, FL 32925-3002

1 Michelle Crown
Air Force Technical Applications Center/TTR/BAE Systems
1030 S. Hwy A1A
Patrick AFB, FL 32925-3002

1 Jorge Roman-Nieves
Air Force Technical Applications Center/TTR/BAE Systems
1030 S. Hwy A1A
Patrick AFB, FL 32925-3002

1 Jeff Miller
Air Force Technical Applications Center/TTR/BAE Systems
1030 S. Hwy A1A
Patrick AFB, FL 32925-3002

1 Kia Tavares
Air Force Technical Applications Center/TTR/BAE Systems
1030 S. Hwy A1A
Patrick AFB, FL 32925-3002

1 Ward Hawkins
Los Alamos National Laboratory
MS F665
P.O. Box 1663
Los Alamos, NM 87545

1 Julio Aguilar-Chang
Los Alamos National Laboratory
MS F659
P.O. Box 1663
Los Alamos, NM 87545

1 Bill Walter
Lawrence Livermore National Laboratory
MS L-046
P.O. Box 808
Livermore, CA 94551

1 Stanley Ruppert
Lawrence Livermore National Laboratory
MS L-046
P.O. Box 808
Livermore, CA 94551

1 Jeffery Hampton
4022 Carolina Shores Lane
League City, TX 77573

20	MS0401	Dorthe Carr, 05527
1	MS0401	David Gallegos, 05527
1	MS0401	Christopher Young, 05527
1	MS1243	Elaine Martinez, 05533
1	MS0576	Lisa Wilkening, 05533
1	MS0671	James M. Harris, 05627
1	MS0899	Technical Library, 9536 (electronic copy)


Sandia National Laboratories