

University of Massachusetts Deploys Advans' BI Conversion Software for Business Intelligence Migration



*Advans IT Services, Inc. White Paper
September 2014*



UMass Data Warehouse Case Study / OBIEE / Cognos / netPulz

University of Massachusetts Deploys Advans' BI Conversion Software for Business Intelligence Migration

Abstract	3
UMass	3
Architecture	4
Business Problem	5
Solution	5
Migration	6
Conclusion	7
References	7

Abstract

The University of Massachusetts' ("UMass) enterprise-wide data warehouse application ("Summit") had multiple ETL and BI tools as part of its environment. UMass proposed to re-architect their Summit application by first upgrading two of the tools, DataStage and Oracle Business Intelligence Enterprise Edition ("OBIEE") and then retire its Informatica and Cognos applications. To do this, the Informatica mappings were to be migrated to the upgraded DataStage application while the Cognos dashboards and reports were to be migrated to the upgraded OBIEE application.

To facilitate these efforts, UMass requested that the vendor chosen for this work also provide an offsite lab to host the appropriate software to test the upgrade and migration procedures.

Advans IT Services, Inc. ("Advans"), the vendor chosen to perform this project, deployed servers, with the applicable software installed on them, in its datacenter, to test the installation, configuration, upgrade and retirement procedures. All procedures were performed several times, optimized and documented in the lab before they were carried out on the applications hosted by UMass.

This white paper will address the retirement of Informatica and Cognos, the migration of their objects, and the development of a custom Java application to programmatically recreate the Cognos reports and dashboards in OBIEE.

A previous white paper addresses the use of Advans' Managed Virtual WAN service to rapidly deploy secure point-to-point network connections between sites, (see, Figure 1, below). In this case, both of UMass' datacenters were connected to Advans' datacenter and the resulting virtual WAN facilitated the testing of the procedures used in this project in Advans' lab using live data, (see, the Advans white paper, "University of Massachusetts Deploys the Advans netPulz Transparent Network Appliance for Data Warehouse Upgrades").

UMass

The University of Massachusetts has been providing high quality educational opportunities for Massachusetts residents and students from all over the world for more than 140 years. There are five campuses in the UMass system:

- UMass Amherst
- UMass Boston
- UMass Dartmouth
- UMass Lowell
- UMass Medical School

The university's mission is to provide affordable and accessible education of high quality and to conduct programs of research and public service that advance knowledge and improve the lives of the people of the Commonwealth, the nation, and the world.

Architecture

Prior to initiating this project, the UMass Summit application consisted of the following software applications:

Current	Proposed
DataStage 7.5	To be upgraded to 8.5
Informatica	To be retired (migrate to DataStage)
OBIEE 10g	To be upgraded to 11g & Install BI Publisher
Cognos	To be retired (migrate to OBIEE)
Oracle 10g Database (back-end)	No upgrade

UMass maintains Development, Test, QA and Production environments in two geographically separated datacenters. Source databases, consisting of the back-ends of various applications used by the individual campuses, are also hosted in these datacenters. Data is extracted from the source databases, transformed and inserted into the data warehouses, daily.

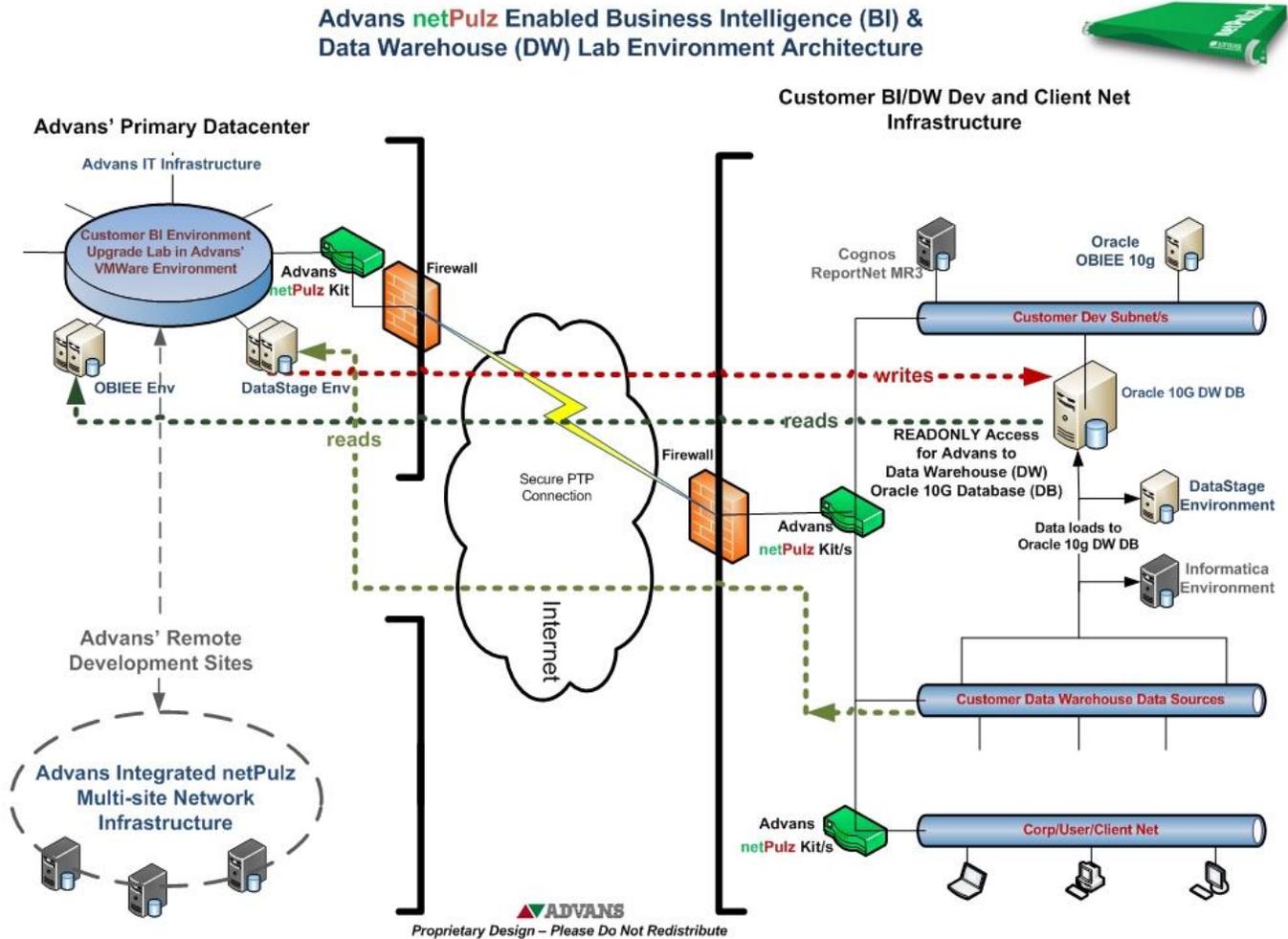


Figure 1, netPulz Virtual WAN

Business Problem

Ordinarily, when migrating from one ETL or BI application to another, the objects in the software application to be retired are manually recreated in another application. Doing so requires identifying each object, determining its intended function, recreating the object and testing that it works properly.

Given the scope of the UMass system, the various campuses combined had nearly 1000 ETL objects and over 20 thousand reports and dashboards included in the Summit application. While it was possible to manually recreate the Informatica mappings as DataStage jobs, as approximately 500 to 750 mappings and SQL statements were in scope, a manual effort would be impossible for reports and dashboards. It would simply take too long and cost too much to manually recreate the 2500 to 3000 reports that required migration from Cognos to OBIEE. The only solution would be to create a software tool to automate the report and dashboard migration.

Moreover, UMass had calculated the amount of developer time required to recreate all the reports and dashboards in scope. From this calculation, UMass determined that in order to make a sound business case justifying the cost and time expenditures for the Cognos to OBIEE migration, the conversion tool had to eliminate at least 80% of the manual effort associated with recreating the reports and dashboards. If such a time savings could not be achieved, then the migration effort would not be undertaken.

Solution: Advans Software Development

Both Cognos and OBIEE have application programming interfaces, and UMass determined that a Java application could be created to programmatically migrate Cognos reports to OBIEE. While it was reasonably believed that such a program could be created, at issue, however, was whether the Java application would automate enough of the report and dashboard recreation process to eliminate at least 80% of the manual work.

Advans provided a senior Java developer to create such an application. Within two months, the developer created an application that would connect to Cognos, identify all the objects in Cognos BI Content Store, and extract their properties, including the name, format and the SQL statement(s) upon which they were based. Once that was accomplished, the Java application connected to the OBIEE API and recreated the objects within the Oracle BI Presentation Catalog.

To use this application, all that was required was to point the application at the appropriate Cognos BI Content Store, identify which objects were to be extracted and then point the application at the appropriate Oracle BI Presentation Catalog. The user then initiates the conversion process. After it completes, and thousands of objects could be recreated in a matter of minutes, the application provides a list of all the source and target object names which were used to validate the new OBIEE object's functionality against the corresponding Cognos object.

Migration

For both the ETL and BI migrations, the application objects were grouped by subject area. Two teams were formed consisting of DataStage and OBIEE developers. One of the DataStage developers in Chennai was also familiar with Informatica, and that developer served to identify the mappings in scope and determine their functionality, so the other OBIEE developers could recreate those objects as DataStage jobs.

Project Teams

Advans used blended onsite/offshore development teams for both the ETL and the BI migrations. The offshore developers were located in Advans' Chennai, India office. The Advans virtual WAN allowed ETL processes to act upon production data rather than a stale reference database. The output of DataStage processes as well as reports and dashboards generated by OBIEE in the offsite lab, were compared to live UMass applications.

In other words, the daily production data load executed on not only the UMass environments, but also on the upgraded Advans environments in the lab. Daily side-by-side comparisons between the production UMass and Advans' environments were made, accelerating the development and QA cycles in the migration of ETL and BI objects because issues could be resolved in real-time.

ETL

As mentioned previously, the source mappings and SQL statements, were identified in Informatica and grouped by subject area. Groups of mappings were manually recreated as DataStage jobs. The developers creating the DataStage jobs conducted Unit Testing to verify their functionality. This testing consisted of AB Testing, row counts of tables affected by jobs and regression testing. Again, since a netPulz virtual WAN was in place, both source and target objects and tables could be viewed simultaneously.

Every object was catalogued and the process from source identification, target object creation and testing was tracked. Once Unit Testing was completed and approved by the development team leader, the catalogue listing was shared with UMass employees for User Acceptance Testing.

Once all the ETL jobs were created in DataStage and they had passed UAT, a code promotion path was identified for promoting the code through the live UMass DataStage environments.

BI

As was the case with the ETL objects, the reports and dashboards in the Cognos application were grouped according to subject matter and catalogued. Once the BI conversation tool was developed, groups of reports and dashboards could be migrated from a Cognos test application to an OBIEE test environment to determine the accuracy with which the reports and dashboards were recreated. Several migration runs were performed to identify issues with the software.

After approximately two months of development, refinement and testing, UMass concluded that the BI conversion tool easily surpassed the automation threshold required to move forward, making the conversion effort feasible.

The same process that was used for the ETL objects was followed by the OBIEE developers. Each report and dashboard was catalogued, compared to its source object, including the various elements that made up the report, such as SQL statements, formatting and code execution. Specifically, report elements such as totals, subtotals and other calculations were checked, by the BI developers, directly against the source object to ensure that each report and dashboard functioned as designed. Once the developers conducted Unit Testing the reports were given to UMass for UAT.

Lastly, once all the reports and dashboards successfully completed UAT, the objects were promoted through the UMass live environments.

Conclusion

While having a netPulz WAN and onsite and offshore developers following the best practices for project management and development were all vital for the successful outcome of a project of this scope, such an undertaking would only have been possible with the use of an automated tool.

References

1. "5 Minute WAN: Rapid WAN Deployment with netPulz", Advans IT Services, Inc., white paper, <http://www.netPulz.com>
2. "University of Massachusetts Deploys the Advans netPulz Transparent Network Appliance for Data Warehouse Upgrades", Advans IT Services, Inc., white paper, http://www.advansit.com/pdfs/UMass_Summit_Upgrade-Advans_netPulz_Virtual_WAN.pdf



University of Massachusetts Deploys the Advans netPulz Transparent Network Appliance for Data Warehouse Upgrades
September 2014
Published in Collaboration with Jeffrey Glatstein, Shahriar Panahi and Sean Blood of the University of Massachusetts
Paul Angelo, Peo Nathan, and Vik Solem of Advans IT Services, Inc.

Authors: Paul Angelo, Peo Nathan, Advans IT Services, Inc.



Advans IT Services, Inc.
65 Boston Post Road W
Marlborough, MA 01752
U.S.A.
www.AdvansIT.com

Inquiries:
Phone: +1.508.624.9900
Fax: +1.508.624.9905

Copyright © 2014, Advans. All rights reserved.
This document may not be reproduced or transmitted in
any form or by any means, electronic or mechanical, for
any purpose without our prior written permission.