# University of Massachusetts Deploys the Advans netPulz Transparent Network Appliance for Data Warehouse Upgrades

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


**UMass Data Warehouse Case Study / OBIEE / DataStage / netPulz**

# University of Massachusetts Deploys the Advans netPulz Transparent Network Appliance for Data Warehouse Upgrades

## 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 OBIEE.

To facilitate these efforts, UMass requested that the vendor chosen for this work also provide an offsite lab 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 UMass data warehouse applications.

This white paper focuses on the use of Advans' Managed Virtual WAN service to rapidly deploy secure point-to-point network connections between sites. In this instance, a virtual WAN was created connecting the test lab in Advans' datacenter with the two UMass' datacenters. Using a virtual WAN, dramatically shortened the project timeline resulting in very significant cost savings for UMass. This paper explores how Advans deploys a virtual WAN using its proprietary Transparent Network Appliance, ("netPulz"), how Advans uses netPulz in support of its IT consulting work and how a virtual WAN impacts the project timeline and associated costs.

A second white paper will address the retirement of the legacy ETL and BI tools, including how their objects were programmatically migrated with a proprietary software application developed by Advans.

## 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 |

The Summit data warehouse back-end, an Oracle 10g database, was not to be upgraded to 11g, however. This necessitated then, the creation of a separate 11g database to act as the OBIEE application's metadata repository. Lastly, Oracle BI Publisher was installed as part of the OBIEE 10g to 11g 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 daily.

One other project requirement included deploying both DataStage and OBIEE in high availability clusters. Moving both applications to this architecture was included as part of the upgrade procedures.

To test the upgrade procedures, Advans installed both DataStage 7.5 and OBIEE 10g in its lab, in the same configuration as used by UMass. Upgrades were performed as listed above, with the step-by-step procedures documented. The upgrade procedures were performed several times to optimize them. After the upgrades of DataStage and OBIEE were validated, they were provided to UMass' staff so that they could perform them on the applications hosted in their datacenters.

Work then began to retire the Informatica and Cognos applications, by migrating their objects into DataStage and OBIEE and BI Publisher, respectively. While the retirement/migration phase of this project will be the focus of another white paper, the use of the Advans virtual WAN also significantly reduced the timeline of that phase of the project. For that reason, the retirement/migration phase will be briefly discussed here.

## Business Problem

In order to meet the aggressive project schedule, it was necessary to stand-up an offsite testing lab in Advans' datacenter very quickly. Moreover, Advans proposed connecting it's lab to the two UMass datacenters, using a virtual WAN, so that current data could be used rather than stale data from database copies.

## Solution: Advans netPulz

Advans has a Managed Virtual WAN solution based on its proprietary Transparent Network Appliance, ("netPulz"). The netPulz network appliance, is a 1-U miniserver that is shipped pre-configured to customers (a virtual netPulz appliance is also available). Installed in just minutes, the netPulz appliance creates a secure point-to-point connection across the Internet backbone, although any other telecom circuit may be used. This solution is technology agnostic, works with any existing infrastructure and has no impact on the LAN's where it is installed, i.e., no network devices must be removed or reconfigured for netPulz to work.

Using netPulz obviates the need to create a traditional WAN, as netPulz resolves the packet collisions between the networks, so separate IP address spaces are not required. This is accomplished by combining the functionality of a router, firewall and VPN in a netPulz appliance and all sites are connected by an encrypted tunnel. Lastly, Advans remotely manages the netPulz appliances which may be re-configured on the fly, (see, Figure 1, below).

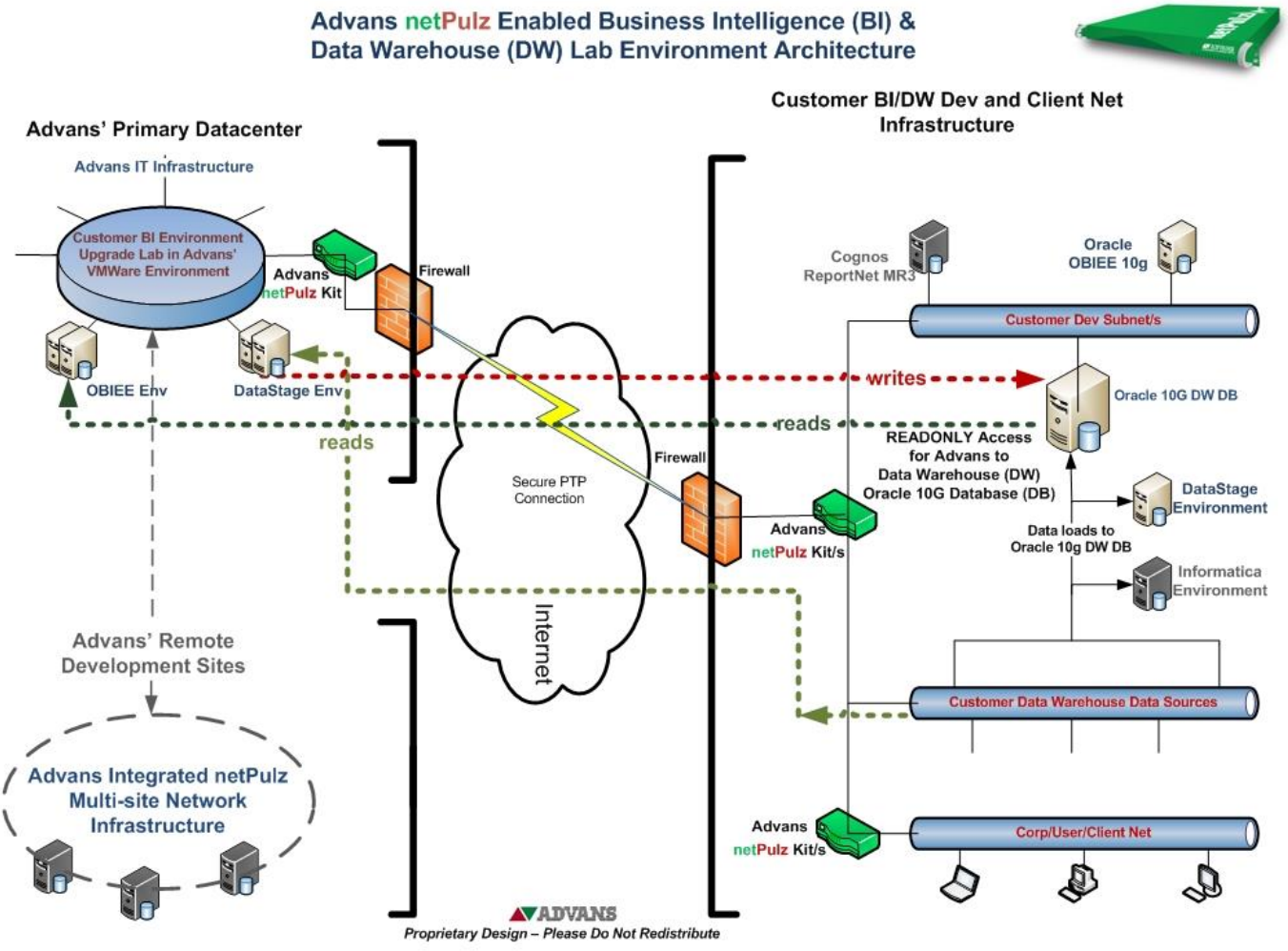


Figure 1, netPulz Virtual WAN

## netPulz Virtual WAN

UMass placed one netPulz appliance in each datacenter. Advans recommends connecting a netPulz appliance to a VLAN to restrict its access to only required resources. In this case, access was made available to only the applicable source and target databases associated with the Summit data warehouse application. A third netPulz appliance was deployed in Advans' datacenter, establishing connectivity between it and the two hosted by UMass, across the Internet backbone.

The virtual WAN, established by the netPulz appliances, has the effect of making Advans DataStage and OBIEE installations in Advans' lab, appear to be part of the two UMass networks. As a result, data in the source databases, was pulled from both of UMass' datacenters across the Internet into the lab. The source data was then transformed by the DataStage installation in Advans' datacenter.

After the ETL processing was completed by the DataStage application in the lab, the transformed data was then pushed back across the Internet where it was loaded into the target data warehouses hosted in UMass' datacenters. Because of the netPulz virtual WAN, the OBIEE installation in Advans' lab, connected to the target data warehouses in UMass' networks, then was used to generate reports and dashboards. In effect, the DataStage and OBIEE applications hosted in Advans' lab function as if they were both hosted in UMass' network.

## Project Team

Advans used a hybrid onsite/offshore development team to complete this project. The offshore developers were located in Advans' Chennai, India office. The staff in Chennai did not have direct access to UMass' networks. Instead, the offshore developers connected directly to the data center in Advans' headquarters in Westborough, Massachusetts. From there, all developer and administrative traffic was routed to UMass' networks. Doing so assures UMass that all network connections originated from within Advans' network, rather VPN connections from an offsite location.

Using netPulz to create a virtual WAN conferred several advantages for UMass. Firstly, no client data was stored outside of UMass' network. While source data was pulled into Advans' datacenter, it was only acted upon by the ETL processes of the DataStage application but not stored there. After ETL processing, the data was loaded into the target data warehouses hosted in UMass' datacenters. Data, therefore, only passed through the Advans datacenter and was never stored locally, thereby reducing privacy and security concerns.

Ordinarily, when performing such a project, the source and target databases must be cloned and replicated in the offsite lab. Database cloning is often repeated several times during the course of a project. Moving data between sites and replicating databases, which may be quite large, is typically a time consuming process. Moreover, to verify the accuracy of ETL processes and BI output, a reference copy of the data must be maintained by the customer to maintain concurrent data.

By eliminating source and target database cloning, a significant time savings was realized, and not just the time saved for replicating databases in the lab. 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 upgrade and retirement/migration phases because issues could be resolved in real-time.

## Upgrades

To perform the upgrades, both the DataStage 7.5 and OBIEE10g applications were installed in Advans' datacenter, once the sites were connected with a netPulz virtual WAN. The applications in the lab were configured exactly as they were in UMass' datacenters. The upgrade procedures were performed several times, optimized and documented, including the creation of the 11g repository for OBIEE.

During the upgrade, bugs in DataStage were encountered necessitating patches to be created by IBM and then applied. Once the upgraded applications were functioning, DataStage began transforming the source data and inserting it into target data warehouses hosted in UMass' datacenters.

Both the DataStage jobs and OBIEE reports and dashboards were compared side-by-side to the actual jobs and reports in UMass' systems to validate the upgrades. A netPulz virtual WAN allowed a developer to log into both a UMass application, such as OBIEE, and the OBIEE application in Advans' datacenter simultaneously. This is how issues could be resolved in real-time, since production data was used by both the UMass application and the one in Advans.

The upgraded or migrated objects were logically divided into groups according to their subject matter. Project sprints were constructed to analyze each object, identify and correct any deficiencies and lastly to test their functionality, first by the Advans developers and then by UMass' end users.

For ETL, DataStage jobs were subjected to Minus Testing utilizing SQL scripts, to compare the tables in the Advans target data warehouse to a reference UMass data warehouse to validate the upgrade. Once the jobs successfully completed all the testing, the jobs were promoted into UMass development, test, QA and production environments.

In parallel to the ETL work, a similar process occurred for all the OBIEE dashboards and reports. In this case each object was logically grouped by subject area and project sprints devised to review the object's appearance and overall functionality. All calculated totals and widget functions were compared to a UMass reference environment. Again, since netPulz was deployed, the developers performing the analysis could simultaneously log into Advans' OBIEE

environment in the lab and the UMass OBIEE reference environment. Each report and dashboard was then simultaneously opened so that the developers could display the reports and dashboards side-by-side for comparison.

## Conclusion

While the techniques used to perform the upgrades, object analysis and code promotion followed the best practices for each application and the entire project was managed with Agile project management concepts, the use of a netPulz virtual WAN to connect the Advans and UMass networks enabled the rapid deployment of an offsite lab environment which yielded several benefits. First, it was not necessary to clone large data warehouses, nor store any client data offsite. Moreover, the most current production data could be used for testing, rather than making adjustments for stale data as part of application validation. Secondly, developers could simultaneously log into an application hosted in Advans' lab and a UMass reference environment allowing for side-by-side object comparisons in real-time with production data. Both of these factors significantly reduced the project timeline and cost.

## References

1. "5 Minute WAN: Rapid WAN Deployment with netPulz", Advans IT Services, Inc., white paper, http://www.advansit.com/?q=node/41

University of Massachusetts Deploys the Advans netPulz Transparent Network Appliance for Data Warehouse Upgrades
May 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