Overview

West Virginia Department of Education (WVDE) is required by law to collect and store student and educator records, and takes seriously its obligations to secure information systems and protect the privacy of data collected, used, accessed, and stored by the Department. Educational data is an asset that is essential to WVDE's business and must be diligently protected. WVDE has adopted their Data Access & Management Guidance policies to ensure that it is fully compliant with the requirements of the Family Educational Rights and Privacy Act (FERPA) and its other legal and regulatory obligations.

The ZoomWV Education Data project leverages the Edvantage Data Warehouse (Edvantage) technology platform provided by Versifit Technologies. The Edvantage solution employs best security practices to ensure that both data at rest and data in transit are protected. The subsequent pages below describe in detail how the Edvantage technology secures the ZoomWV data, the ZoomWV Secure/Internal environment infrastructure, and the ZoomWV Public environment infrastructure.

Edvantage Security

The elevated interest in data governance has refocused the Business Intelligence (BI) community on the criticality of security and data quality to the BI solution. In truth, these have always been important elements and particularly so in education where privacy has been mandated not only by local district policies but also by state and federal legislation. The Edvantage Solution conforms to the BI best practices for security and offers the foremost data quality technology available.

Security, like the other architectural elements of a BI solution, is ideally layered to allow for all types of users to access the appropriate data. Layered data access allows one report to service many users, thereby easing the change management burden and rendering a more flexible and lean system. Users may have their data access restricted because the data is inappropriate for them or because they have not been fully trained on use of that data.

The security platform to administer BI application access should ideally be able to support the following:

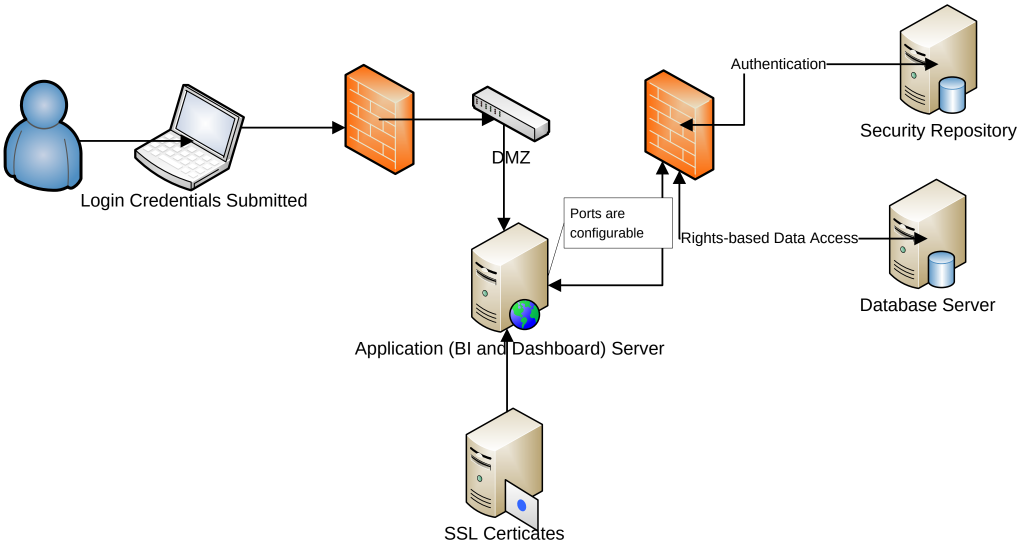
- Mixed Mode Authentication (LDAP, Active Directory, Custom, etc.)
- Domain-Level Access (subject area or table)
- Attribute Level Access (columns, dimensions, measures, calculations)
- Row Level Access (list of vendor number, territories, school codes, etc.)
- Document Level Access (reports, spreadsheets, URL, etc.)
- Application Functionality Access (view, refresh, upload)

The Edvantage Solution is secured at two levels: system and application.

System Level—The database and application servers are secured at a system level in the operating system (OS) and in the database engine (DBMS). Following BI security best practices, only network and system administrators are granted system level access. The BI applications themselves have system level logins and via this proxy account, all other users of the applications have access to the resources and data. There may be a desire or requirement for a small group of developers/designers to have direct access to the database. This is possible in the Edvantage Solution as long as it is managed properly. As a rule, very little security is applied at the database level in BI solutions: those who have direct data access have full access to the relevant schemas.

Application Level—Most security definitions and administration will be performed at the BI application level. The Edvantage Analytical Dashboard has comprehensive security models which allow for user- (or optimally role-) level definition of security at the object level. Security administration can be delegated in the Dashboard so that a district can administer the security for his/her users in a regional or state solution.

Basic Security Architecture



User Authentication

All users have unique ID's and required passwords as their login credentials, regardless of the security repository used. Users cannot access the system without authentication and membership in a defined group. These credentials are required for login with every session a user opens. Anonymous, generic, or default logins are not allowed. Privilege escalation protection is inherent in the security challenges for every session. Users never have access to any underlying security model to be able to make alterations. Users of the Edvantage Solution do not have access to the underlying databases. All data requests for Dashboards and reports are executed by the application server (proxy accounts). Every session has a designated inactivity timeout period that is configurable, but should be set within an upper range of no more than 10 minutes due to the sensitivity of the content.

This security level identifies a user and verifies correct user name and password. The Edvantage Dashboard supports Active Directory (AD), Lightweight Directory Access Protocol (LDAP), and custom security repositories, and is capable of heterogeneous authentication (can use multiple directories in the same instance to authenticate a user). We recommend where feasible that a local AD or LDAP repository be utilized, but this can be difficult in some regional or state deployments. In these cases, we have advocated either the creation of a separate AD or LDAP repository for the region/state, or the use of our application's security repository for authentication. Single-Sign-On (SSO) is possible with the proposed solution so long as the conditions required by the authentication engine are met.

Content Access

Reports and data are protected through the underlying security model. Administrators create groups with defined roles and then assign users to these groups according to their data needs. Users are thus granted permission to access reports, databases, tables, fields, or even individual records, according to their defined roles and groups.

Dashboard users are classified into groups and each group has a defined level of access to content (i.e. Reports/Analyses, Dashboard Pages, Metrics, etc.). These groups are defined in the Dashboard. If LDAP or AD is used for authentication, the groups are typically defined there as well and then mapped into the corresponding Dashboard groups.

Row Level Data Access

Based on data that associates Dashboard users to a subset of data (e.g. schools or students), users are presented with data sets specific to their scope of responsibility. The Edvantage Solution provides a comprehensive model for the definition and maintenance of row-level security which will leverage the row-level definitions in the district's student management system.

The Edvantage Solution has a role-based security model which can restrict access at the object level. The groups (roles) of security repositories such as LDAP or AD are utilized to provide access to the Edvantage Solution and to the content it contains. All of the required layers of security (Folder, Content [Dashboard or Report], Table, Column, or Row) can be managed by group or role. The roles or groups defined in the Student Management System (SMS/SIS) provide access to the data presented in that content (reports or dashboards); consequently, a teacher only sees the reports and dashboards designated for her role, and then she only sees the data on those reports or dashboards for the students she has in her classes.

The Edvantage Solution leverages the data pertaining to a staff member's data in the underlying SMS. We link authenticated users to their corresponding logins or teacher/staff records in the SMS and load our own row-level security tables. Our security tables have user interfaces to augment the security definitions of the SMS so that individuals requiring row-level access to the analytical data, but who are not defined in any way in the SMS, can have restricted access. These security tables are joined to the analytical data tables in the semantic metadata of the reporting and analysis software.

Heterogeneous Security Model

The heterogeneous security model offers a unique layering of security policies. It is flexible enough to allow for global and regional security policies to be applied to any user because it uses multiple directories in the same instance to authenticate a user. Combined with the role-based or row-based security policies, hierarchical delegation of control can be restricted to specific classes of users, groups, sites, and modules.

Hardened Security

The Edvantage Solution is hardened and resistant to internal and external compromise. The security model and application are kept up-to-date with the latest security patches. Any and all vulnerabilities are addressed in an immediate fashion. Because the solution is hardened, we opt not to encrypt data at rest as it adds significant overhead that negatively impacts performance.

Encryption

To protect sensitive data (content, passwords, etc.) encryption is employed in three ways within the Edvantage Solution.

- Passwords for local users and database connections are encrypted with a 128-bit key and the private key itself is encrypted with another internal key.
- We recommend that SSL certificates be used to secure communications among the remote connections to the application servers. The applications in the Edvantage Solution support SSL.
- Finally, transfers of data between the Districts and the State should be encrypted. We recommend transferring this data via a password-protected, compressed data file through an encrypted SFTP tunnel (we support either FTPS or SFTP).

Safe SQL

The applications in the Edvantage Solution are written to protect against SQL insertion attacks and test any values which will be added to the dynamically generated SQL. The Edvantage Dashboard goes further, only allowing for entry of string values in the login screen and in the restricted access administrative/developer interfaces.

Validation & Auditing

Trust in the quality of the data in a business intelligence solution is critical. If the users of the system do not believe that it is accurate then they will not use the system and it requires much effort to restore confidence in the solution’s data quality. One of our initial project tasks is to define data quality metrics, as the validity, consistency, and integrity of the data can be measured in different ways. For example, we could compute a student’s GPA in the ETL and be absolutely correct in our calculation, but if the computed value varies from the value in the student management system, it will be considered invalid.

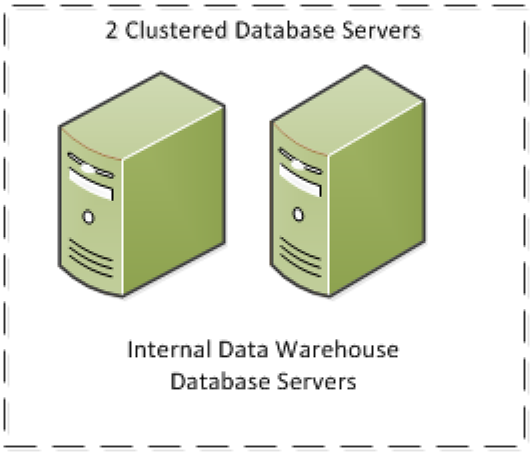
Logging

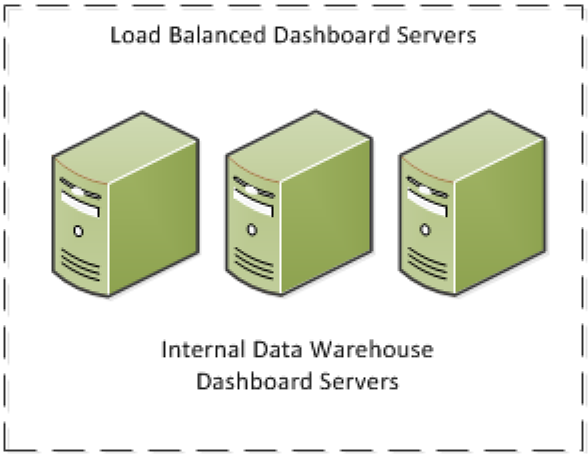
There is extensive logging which is performed by the different applications in the recommended technology stack. Logging plays a vital role in the solution. Logs are generated by the ETL process so that administrators can verify the relative success of each loading procedure and identify any issues which will require attention. Logs are read by the ETL processes to determine how a dependent procedure should behave. Logs of Dashboard processing and user activity are used by administrators to optimize the system performance and debug reported issues.

Educational clients utilize logs to demonstrate and/or verify compliance of the solution with FERPA and other privacy guidelines. A collection of reports are provided with the solution which allow for the analysis of system utilization by user and by report. The logs can retain the actual SQL generated by the solution for each data access request by user and date. This level of transactional detail allows for more detailed analysis in the future.

ZoomWV Internal/Secure Infrastructure

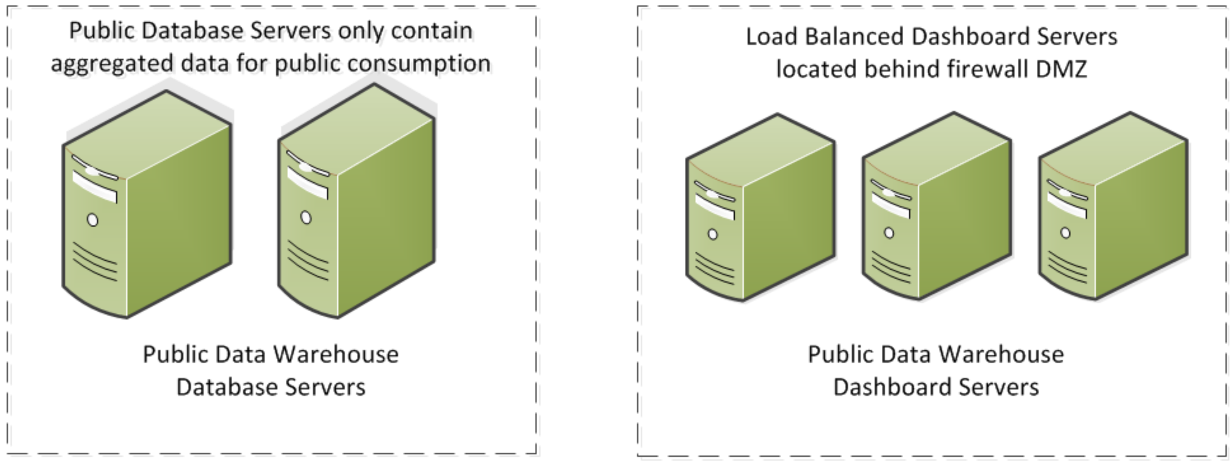
The internal/secure ZoomWV environment is only accessible on the local network or external users through WVEIS on the Web (WOW). Only users with valid credentials will be able to access the Dashboard Application and only a very limited number of typical IT Admin personnel will have access to the physical servers.





ZoomWV Public Infrastructure

The public ZoomWV environment will be available to all Internet users; however, the data contained in that environment will only be aggregated student data that is WVDE approved for public consumption. In following best security practices and consistent with how other states have deployed a public version of the Edvantage Dashboard, the servers are located in a “demilitarized zone” (DMZ) behind a firewall. The DMZ allows the servers to be isolated from the internal network hosting the database servers so that in the event they are compromised, the attacker can only access those 3 systems.


Public Database Servers only contain
aggregated data for public consumption
Public Data Warehouse
Database Servers
Load Balanced Dashboard Servers
located behind firewall DMZ
Public Data Warehouse
Dashboard Servers