Prepared for:

NATIONAL ASSOCIATION OF INDEPENDENT SCHOOLS

The National Association of Independent Schools

Technical Recommendations.

10.02.2009 – First Draft delivered to Client.

Prepared by:

10 Notes

Nate Mow	Email	nmow@10notes.com
	Direct	202.553.4000
Paul Bhorjee	Email	pbhorjee@10notes.com
	Direct	917.612.5835
Lori Evans	Email	levans@10notes.com
	Direct	202.577.8873

CLIENT INITIALS _____

10 Notes, LLC

10 Notes, LLC • 1725 I Street, NW, Suite 300 • Washington, DC 20006
tel 202.349.1115 • fax 202.280.1341 • web www.10notes.com

CLIENT INITIALS

10 Notes, LLC • 1725 I Street, NW, Suite 300 • Washington, DC 20006
tel 202.349.1115 • fax 202.280.1341 • web www.10notes.com

1. Objectives.

This document outlines the technical requirements and recommendations for the successful implementation of the Management Reporting System, leveraging all documentation generated to date. The objectives for this document are to:

- Document and visually illustrate the flow and/or synchronization of data between disparate systems (e.g. iMIS, Comp*Assist Online, StatsOnline, Solomon, etc.)

- Name the product that will be implemented (e.g. SQL Server Reporting Services, Crystal Enterprise Server, a custom in-house solution, etc.)

- Document the implementation procedure for the named Reporting System.

- Describe the process for developing reports identified by the Use Case documents (as well as future reports). Group the reports based on high-level departmental relevance, or target audience.

2. Data Sources.

In order to connect each of the externally-hosted data sources, NAIS will need to make sure that each vendor in question has a push/pull mechanism of some sort that can be leveraged to send data to and fro from the Management Reporting System. Roy Kallman specifically cited FolderWave's existing connector, which will of course be critical here. As for the rest of the vendors/systems in question, the following describes what each NAIS vendor will need to put in place:

- **FolderWave** – During the kickoff meeting FolderWave did mention some sort of XML-based data exchange mechanism...we did follow up with them on 7/21 about the format, delivery method, etc. but received the data dictionary in response. So presumably that data connector is still under active development at this point.

  Notes: The data dictionary contains names of ~450 fields including application info, person info, tax info, income info, asset, expenses, and allowances info, contribution calculation info, “unusual conditions”, judgments, etc.

- **Databank** – Data exchange with C*AO is already in place. So we should be able to reliably state that FolderWave is the primary data source for all parent financial data. Its unlikely that NAIS will need to take any further action with Databank, as far as the Reporting System is concerned.

- **Experient** – Currently POCC and Annual Conference data comes back to NAIS once or twice a year; this will not be sufficient for frequent, accurate querying from the Reporting System. We recommend that NAIS engage Experient to provide a more real-time data exchange mechanism, even if its something as simple as a web service that can be hit by the Reporting System.

- **BrightKey** – Based on our findings, BrightKey currently pushes data directly in to iMIS. However, inventory has historically been a problem. We recommend that BrightKey be engaged to provide a broader data set to NAIS. Those inventory records might be dumped to iMIS' Product table, or to an alternate location; regardless, this data must become more visible/available.

CLIENT INITIALS

10 Notes, LLC • 1725 I Street, NW, Suite 300 • Washington, DC 20006
tel 202.349.1115 • fax 202.280.1341 • web www.10notes.com

Notes: BrightKey uses iMIS to manage NAIS orders and inventory. On a daily basis, BrightKey exports data from iMIS to their order fulfillment database, which manages their interaction with the warehouse and shipping. NAIS does not have access to this database, though in the future they plan to upload order tracking numbers from their system into iMIS.

- **Salesforce** – Salesforce does offer web service API connectivity to external applications. The Reporting System will need to be responsible for harvesting this data on a regularly scheduled interval. For now, NAIS should make sure that these web services are available to their account, and ideally, also find out if there is any alternate data dump mechanism available.

  Notes: The Contact and Organizational reports that were received are screenshots of the fields that are presently culled from iMIS and other places to *upload* into Salesforce. Currently NAIS is running about 150 queries in MS Access to consolidate, audit, and produce the data uploaded to Salesforce. The queries obtain the data from ETS (the current SSS provider), StatsOnline, and iMIS.

- **EasiDemographics** – NAIS should engage this vendor to create a web service for use by the reporting system. Based on the interviews that we conducted, it sounds as if EasiDemographics has developed a number of custom utilities for NAIS in the past, so we expect that this particular data exchange should go smoothly.

  Notes: The interface NAIS has with the Demographics Center is purely related to access authentication. A schema for this data is not available, nor is an ODBC connection. An iMIS username and password is required to login. The system is available for review at: http://www.easidemographics.com/cgi-bin/naisdemocenter.asp

- **Customer Service Center** – The data delivery format from the Customer Service Center is not currently known. NAIS should make sure that this vendor can offer metrics in bulk format. Our guess is that there may be some sort of account management website that contains analytics, etc. NAIS will need to make sure that this database is available and exposed in some programmatic manner.

- **StatsOnline** – Through StatsOnline NAIS collects over 1,700 data elements from each school. StatsOnline has a second layer of security that is role-based (e.g., head, business manager, development director, etc.), which will need to be accounted for in the data delivery transfer routine.

  Notes: It used to be that activities were created for schools that had participated in StatsOnline, but NAIS has not done that for several years. Every year, prior to sending out dues statements, NAIS compares school enrollment numbers collected in StatsOnline with those in iMIS, which in recent years have also been self-reported. (The latter exercise serves as a consistency check, since StatsOnline benchmarking numbers are important to members, and also because NAIS assesses dues – through iMIS – based on school enrollment numbers.) Even though NAIS is not storing other StatsOnline values in iMIS, they often merge iMIS records with data exported from StatsOnline to produce various analyses required for business decisions. Some of the fields obtained from StatsOnline for these merges include endowment value, total annual expenses, total

annual revenues, teachers' median salary, annual gifts and grants received, financial aid applications, admissions applications, financial aid awarded, school staff count, etc.

3. Data Connectors.

This section of the document describes known data exchange mechanisms and key data elements that will be required to join and connect various data sources across the enterprise.

3.1. Data Exchange Mechanisms.

- **SSS Data Web Service** – This web service is used for passing SSS School and User data to FolderWave. The following methods are available:
    - GetSchoolRecord – Get one record based on iMIS ID; queries vSSS_School SQL view.
    - GetUserRecordByID – Get one user based on iMIS ID; queries vSSS_User SQL view.
    - GetUserRecordByEmail – Get one user record based on e-mail address; queries vSSS_User SQL view.
    - SendUserRecordByID – Send iMIS ID to FolderWave's web service, which then triggers FW to call GetUserRecordByID.

3.2. Key Data Fields.

NAIS Field	FolderWave Field	Description
vSSS_School.SSSID	SSSSubCode	School's SSS subscriber code.
vSSS_School.SCID	SCID	ID of the individual who is the school's SSS administrative contact.
vSSS_School.Individual_ID	RosterUserIds	A comma-delimited list of individuals rostered to the school, used by FW to determine who is eligible (though not necessarily authorized) to use Comp*Assist.
vSSS_School.Prior_Year_START	PriorYrStartDate	Start date of the SSS subscription for the prior processing year.
vSSS_School.Prior_Year_THRU	PriorYrThruDate	End date of the SSS subscription for the prior processing year.
vSSS_School.Current_Year_START	CurrentYrStartDate	Start date of the SSS subscription for the current processing year.
vSSS_School.Current_Year_Thru	CurrentYrThruDate	End date of the SSS subscription for the current processing year.

4. Data Federation, Data Warehouse, or Data Mart?

We are tasked with recommending a technical solution to a complex problem – namely, what the mechanism is for collecting the data needed in the reports, how it is stored, and how that data is made available for querying, viewing, and distribution.

Through analysis of the data universe, in which the separate functions and departments of NAIS operate, we have determined that data must be centralized in a common and easily accessible repository. Due to the use of many unique vendors which supply NAIS with separate services and products, the operational and historical data used to run these services and products are housed with the individual vendors – separately, physically very far away from each other, and behind proprietary firewalls and security

topologies. This is also the data which will be needed to run the reporting engine NAIS is looking to construct. Data from disparate sources, with disparate purposes, must be aggregated into a common repository for a single purpose – enterprise reporting. We researched three common techniques for accomplishing this task:

4.1. Data Federation.

Data Federation aims to efficiently join data from multiple heterogeneous sources, leaving the data in place -- without creating data redundancy. The data federation pattern supports data operations against an integrated and transient (virtual) view where the real data is stored in multiple diverse sources. The source data remains under the control of the source systems and is pulled on demand for federated access.

With data federation, the consumer will see a single uniform interface. Location transparency means the consuming application of the pattern does not need to be aware of where the data is stored. Nor does it need to know what language or programming interface is supported by the source database, thanks to invocation transparency. For example, if SQL is used, it does not matter to the application what dialect of SQL the source supports. The application also does not need to know how the data is physically stored due to physical data independence, fragmentation and replication transparency – or what networking protocols are used, known as network transparency.

Due to the data retrieval, transport, aggregation, and processing time to construct one requested report, we are recommending against the data federation pattern. All vendors would have to expose their data, in place, probably through a web-service, with read-only access to all data that the reports need. They would have to give us a fast connection to this data and give us the processing power we needed to select the required data very quickly, then move it to our reporting engine very quickly. Each vendor would likely have to provide separate storage and hardware for our reporting needs, since they would not want to expose their proprietary, operational data stores and hardware to NAIS. Each vendor would also need to provide continuing support in the event that reporting needs change or that there was a technical, transport, or programmatic error. It is very unlikely that NAIS could persuade all vendors to agree to these requirements, and so we are recommending against the data federation approach.

4.2. Data Warehouse.

A *Data Warehouse* is a database geared towards the business intelligence requirements of an organization. The data warehouse integrates data from the various operational systems and is typically loaded from these systems at regular intervals. Data warehouses contain historical information that enables analysis of business performance over time. It is a repository of operational data from one or more sources within an organization (e.g. iMIS), together with data derived from a variety of external sources (e.g. EasiDemographics or Salesforce) that have been arranged into meaningful information, and rendered easily accessible, allowing for effective analysis and decision-making.

A data warehouse is built with an eye to serving all reporting needs including data mining, On-Line Analytical Processing (OLAP), ad-hoc reporting, general, and unknown reporting needs. They are developed with no eye for a specific end-use, and thus are built using very normalized data schemata, which is very difficult to implement, very difficult to query against, and very difficult to maintain.

10 Notes, LLC • 1725 I Street, NW, Suite 300 • Washington, DC 20006
tel 202.349.1115 • fax 202.280.1341 • web www.10notes.com

4.3. Data Mart.

A *Data Mart* is a repository of data gathered from operational data and other sources that is designed to serve a particular community of knowledge workers. In scope, the data may derive from outside data sources, proprietary vendor data, an enterprise-wide database or data warehouse, or be more specialized. The emphasis of a data mart is on meeting the specific demands of a particular group of knowledge users in terms of analysis, content, presentation, and ease-of-use. Users of a data mart can expect to have data presented in terms that are already familiar to them.

In practice, the terms data mart and data warehouse each tend to imply the presence of the other in some form. However, most writers using the term seem to agree that the design of a data mart tends to start from an analysis of user needs and that a data warehouse tends to start from an analysis of what data already exists and how it can be collected in such a way that the data can later be used. A data warehouse is a central aggregation of data (which can be distributed physically); a data mart is a data repository that may derive from a data warehouse, or not, and emphasizes ease of access and usability for a particular, designed purpose. In general, a data warehouse tends to be a strategic but somewhat unfinished concept; a data mart tends to be tactical and aimed at meeting an immediate need.

Reasons to build a data mart rather than a data warehouse:

- Easy access to frequently needed data.
- Creates collective view for a group of users.
- Improves end-user response time.
- Ease of creation.
- Lower cost than implementing a full data warehouse.
- Potential users are more clearly defined than in a full data warehouse.
- More flexibility in implementation schemata.
- Focused-need development, as opposed to General-need development.

We recommend the building of a focused-need data mart database, whose primary purpose is to feed the flat and drill-down reports proposed by 10 Notes in the Use Case documents, with a schema that will facilitate easy integration of new data, and the building of new, similar flat and drill-down reports in the future.

5. Implementation.

This section of the document names the reporting system, recommends the initial implementation procedures for the system, and provides an overview of the features available to support maintaining report history, securing access to reports, and general output and delivery formats.

5.1. Database and Reporting Engine.

We recommend SQL Server 2005 for the central data store and SQL Server 2005 Reporting Services for the reporting infrastructure.

SQL Server 2005 Reporting Services (SSRS) is a server-based reporting platform that provides comprehensive reporting functionality for a variety of data sources. Reporting Services includes a complete set of tools for NAIS to create, manage, and deliver reports, and APIs that enable developers to integrate or extend data and report processing in custom applications. Reporting Services tools work

10 Notes, LLC • 1725 I Street, NW, Suite 300 • Washington, DC 20006
tel 202.349.1115 • fax 202.280.1341 • web www.10notes.com

within the Microsoft Visual Studio environment and are fully integrated with SQL Server tools and components.

Although there are several other well known and robust business intelligence products in the marketplace, they would be difficult for NAIS to implement and maintain. These include Cognos by IBM, Crystal Reports by Business Objects, and Actuate; historically, these products are expensive and unwieldy. 10 Notes' experience in the field has taught us that SQL Server's Reporting Services provides a robust, yet easy to use, infrastructure tightly coupled with SQL Server which serves the needs of an organization the size of NAIS perfectly and flexibly.

This solution is also very cost-effective for NAIS, since the organization already owns licenses to SQL Server 2005 Standard. Reporting Services comes bundled with that product and can easily use SQL Server as a data source. NAIS is familiar with SQL Server and this would product leverages one of NAIS' IT department's core competencies.

Per Microsoft's website, SQL Server Reporting Services offer the following features:

- Output to Excel.
- Output to PDF.
- Output to Image (RGDI, Print).
- Interactive, web-based DHTML rendering of reports.
- A wide variety of possible data source(s).
- Report subscription, scheduling, and delivery.
- Data-driven report subscriptions.
- Extensibility (e.g. SharePoint integration).
- Management through SQL Server Management Studio and Report Manager.
- Report Builder desktop report building application.
- Report Builder Data Sources.
- Report caching.
- Report history.
- Custom authentication mechanisms.
- Flexible, role-based security features.
- Model-level security features.
- Scale-out Report Servers.

In this list of features, it is important to highlight Role-based Security and Scheduled Delivery. It is easy to wrap user and role-based security around any single report or group of reports. These reports can also be emailed directly to individual users or groups of users on a schedule. The reports can also be served up through a built-in intranet tool behind the firewall.

Given the delivery requirements for the NAIS Management Reporting System, we recommend using the built-in Intranet report server. If NAIS determines that reports need to be accessible outside of the firewall, or be delivered to clients over the Internet, we recommend building a custom .NET web application and wrapping reports with appropriate Forms-based security. This application could also be used to sell access to data.

5.2. Web Services or ETL?

There is still the question of how to technically get the data out of the many vendors' repositories and in to the SQL Server data mart. As we have recommended against the Data Federation pattern, the procedures to do this will be run on a schedule with each vendor which is report-independent. There are two main ways to accomplish this:

5.2.1. Web Service.

A *Web Service* is defined by the W3C as "a software system designed to support inter-operable machine-to-machine interaction over a network. It has an interface described in a machine-processable format (specifically WSDL). Other systems interact with the Web service in a manner prescribed by its description using SOAP-messages, typically conveyed using HTTP with an XML serialization in conjunction with other Web-related standards".

In developing web services with all of NAIS' vendors, the vendors would also need to provide resources to initially develop their web service and test each connector individually, and also through our data consumer. Web services are less tightly coupled than other forms of machine-to-machine and proprietary data exchange formats, but are coupled nonetheless. Negotiations of data types and structures will need to occur. Security would need to be implemented individually and load testing performed. This is a very labor-intensive task when working concurrently with many vendors.

Web services are a good solution for fetching small amounts of on-demand data in a Data Federated system. As we have recommended against this pattern, and for the reasons listed above, we recommend against using web services to retrieve data from vendors.

5.2.2. ETL.

We recommend that an ETL set of processes be implemented to populate the data mart. With *ETL* (Extract, Transform, Load), there is a 3 stage process which delivers data from source systems into the data mart: (1), the *Extract* functions read data from a specified source database or data source and extract a desired subset of data, (2) the *Transform* function works with the acquired data – using rules or lookup tables, or creating combinations with other data – to convert it to the desired state, and (3) the *Load* function is used to write the resulting data (either all of the subset or just the changes) to a target database.

ETL is a very loosely coupled set of functions. The implementation vendor would be tasked with the Extract and Transform functions. NAIS would not need to be intimately acquainted with the workings of these functions, only with the end product. The end product of the Extract and Transform functions is typically a flat file of delimited text of a pre-specified format or a validated XML which conforms to a pre-defined schema.

The Load function is the responsibility of both the implementation vendor and NAIS. Usually, the vendor will drop the data product in an FTP directory wherein NAIS can download the file at an automated and scheduled interval. Security can easily be implemented using the SFTP protocol using strong-password credentials. The implementation vendor can then write a simple program in .NET to programmatically download the file, and transport it for processing via the Internet. This procedure can be set to run as often as is needed for the specific vendor and dataset provided. For most vendors and data, a nightly, scheduled batch ETL job is recommended. This nightly job will also be responsible for reporting transfer or processing failures to NAIS' IT department, so that the causes can be investigated and problems dispatched.

10 Notes, LLC • 1725 I Street, NW, Suite 300 • Washington, DC 20006
tel 202.349.1115 • fax 202.280.1341 • web www.10notes.com

At this point, with a validated data product, any last transformations can be run before the data is deconstructed and inserted into the data mart for the reporting engine to consume. We recommend writing this portion of the process in a strongly-typed language (such as .NET) to more easily validate each data element's data type.

5.3. Developing Reports.

These are the steps for developing the reports:

1. Develop the SQL Server database schema using Microsoft's SQL Server Management Studio. The purpose of this schema is to store the data from all vendors and to facilitate querying by the reporting engine. Care should be taken in developing the schema it is not required that that data be overwritten, and that dates from the vendor and create dates be properly stored so that historical and comparative data may be retained, and that data integrity is maintained. A table with process information, such as last-import dates per vendor and failure-codes and dates for ETL procedures which do not perform as expected should be included.

2. Send a list of the required data fields to each vendor.

3. Negotiate with each vendor the format of each raw data product, including how that data will be validated for null values, data type and completeness. Depending on the technical expertise of the vendor, an XML schema ma be using for validating XML files, or a well-known delimited file pattern may be used to validate the individual data rows prior to processing.

4. Negotiate with each vendor how often and when the data will be picked up, the FTP endpoint, and security including protocol (FTP or SFTP) and login credentials. Again, as the *extract* and *transform* parts of ETL are the responsibility of the vendor, NAIS need not concern itself with how exactly the data is culled from the vendor's operational or primary data store, or how that data is processed to produce the the vendor's final data product, just that the data product is made available at the specified point and at by the specified time.

5. Develop a .NET executable for each vendor which is tasked with the *load* procedure. This includes retrieving the data product from the vendor, validating the data, performing any internal transformations which are required to cleanly insert the data, and finally, inserting the data into the data mart database. A separate executable is recommended for easy maintenance and for separation of processes to mitigate cascading faults. If the connection to one vendor hangs, or the process of loading their data causes any number of problems such as infinite loops or unreleased resources, these failures will not affect loading data from other vendors. These separate executables may use shared resources such as common code libraries, but should remain separate with the code and credentials specific to individual vendors remaining separate.

6. Write the individual reports in Microsoft Report Builder. A short exploration of this process is discussed here: http://www.developer.com/db/article.php/3520116/Introduction-to-SQL-Server-Report-Builder.htm.

7. Publish these reports importing them into Report Manager and assign User and Role based security to them as discussed here: http://msdn.microsoft.com/en-us/library/ms157147.aspx.

8. Inform the NAIS user community of the URL of the Reports Manager so that they may run them and set up delivery options.

   OPTIONAL steps for publishing reports outside of the NAIS domain:

9. Build a web page to display the list of available reports using the reportingservice web service API discussed here: http://msdn.microsoft.com/en-us/library/aa274396(SQL.80).aspx. Forms based authentication should be used to validate and authenticate the web user so that security may be applied to the individual reports.

10. Build a web page to display individual reports as discussed here: http://msdn.microsoft.com/en-us/library/ms251671%28VS.80%29.aspx.

6. Conclusions.

The technical recommendations described in this document will provide for a robust and extensible reporting infrastructure for NAIS.

- We have outlined the possible data storage patterns (Data Federation, Data Warehouse, and Data Mart) and have determined that the data mart is the best solution for the immediate and long term reporting needs of the organization. In this analysis, we have included in our decision considerations for ease of implementation, specificity of goals, and maintenance of the system. A well-constructed data mart schema will be extensible for future reporting needs, robust for the data loading procedures, and easy to query against.

- We have explored different reporting engines and have determined that Microsoft's Reporting Services are an excellent choice when considering cost, user community, enterprise footprint, NAIS technical core competencies, and extensibility. Total cost of ownership of reporting services is substantially less that any other platform.

- We have explored how to deliver reports with an eye towards future delivery options such as website and mail outside of the NAIS domain.

- We have outlined the possible methods of technically interacting with the vendors (web service vs. ETL) to retrieve data. We have determined that simple ETL processes should be developed with each vendor. In this way, NAIS will dictate what data is required and can ask for more than what will needed for the reports described in the Use Cases. The vendor will be independently responsible for extracting and transforming the data that NAIS requests, and NAIS will not be responsible for understanding how these processes are implemented. ETL uses simple but robust and secure protocols, and well-known methodologies to transport the data to NAIS.

- We have listed a high-level, step-by-step pattern for building and implementing reports in MS Reporting Services. An in-depth examination of the details of this process are beyond the scope of this document.

It should be understood that this technical implementation is robust and scalable. When the NAIS reporting infrastructure grows to encompass more vendors, more reports, more databases, more users, more types of users (school or parent web users, for instance), and reports are run more frequently, the pieces of the infrastructure may be easily decoupled and scaled out to new hardware.

CLIENT INITIALS _____

10 Notes, LLC • 1725 I Street, NW, Suite 300 • Washington, DC 20006
tel 202.349.1115 • fax 202.280.1341 • web www.10notes.com

7. Appendix.

7.1. System Logins.

Access to the SonicWall VPN gateway should be obtained directly from the NAIS IT department.

SSS Comp*Assist Online demo system access:

host:	http://www8.student-1.com/NAIStest/
username:	10notes
password:	fw10notesfw

Remote Desktop access to a workstation with Access and iMIS installed:

host:	Access1 (192.168.100.165)
username:	notes@nais.org
password:	Notes#16$%20

Login to iMIS and to SSS:

username:	10notes
password:	fw10notesfw

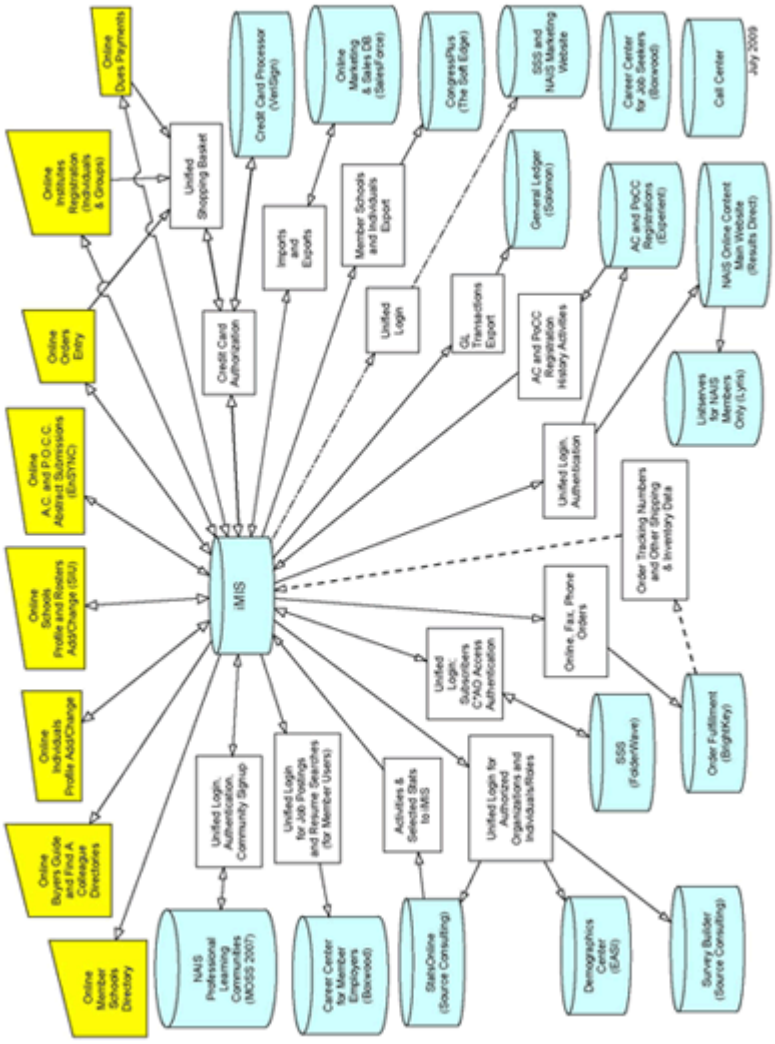
7.2. Data Flow Diagrams.

CLIENT INITIALS _____

10 Notes, LLC • 1725 I Street, NW, Suite 300 • Washington, DC 20006
tel 202.349.1115 • fax 202.280.1341 • web www.10notes.com

7.2.1. Data Universe.

Notes: The diagram below is from July 2009.

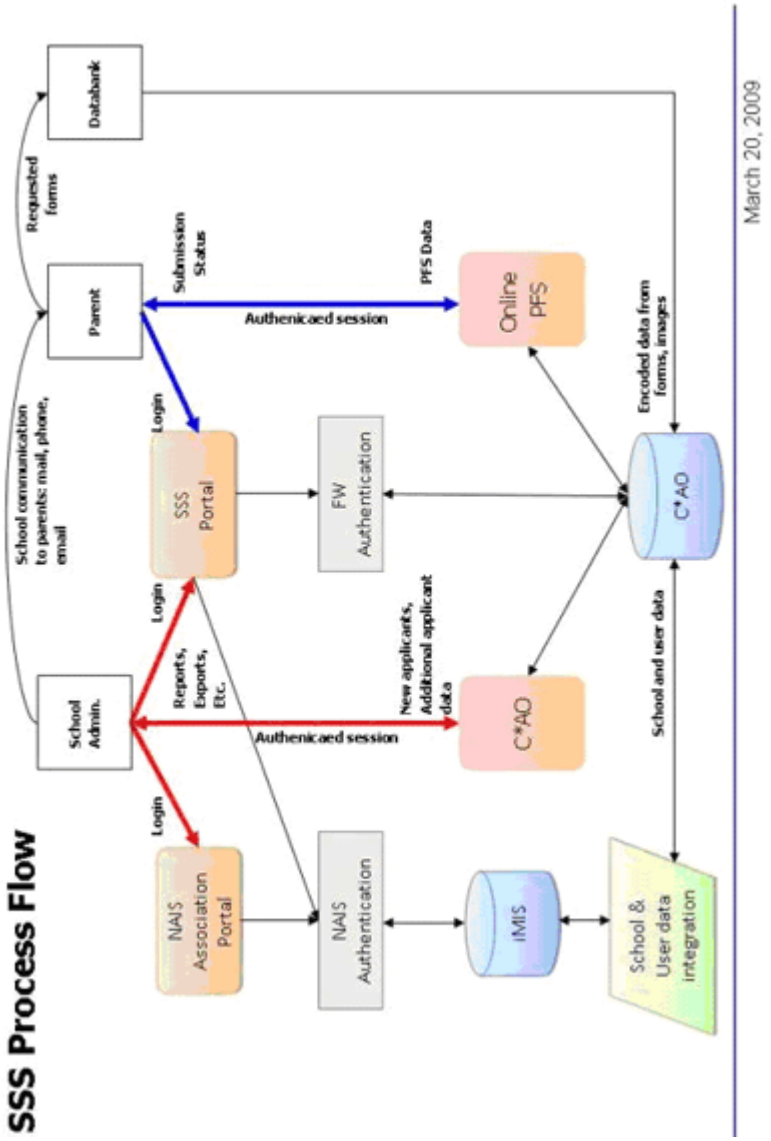


CLIENT INITIALS

10 Notes, LLC • 1725 I Street, NW, Suite 300 • Washington, DC 20006
tel 202.349.1115 • fax 202.280.1341 • web www.10notes.com

7.2.2. SSS – Comp*Assist Online.

Notes: Red lines show school administrator login and work session; blue lines show parent login and work session; black lines are system data flows. Curved lines show data flow outside SSS. Data flow titles are placed near data destinations. Green trapezoid indicates new process outside C*AO and PFS applications, but included in development effort.

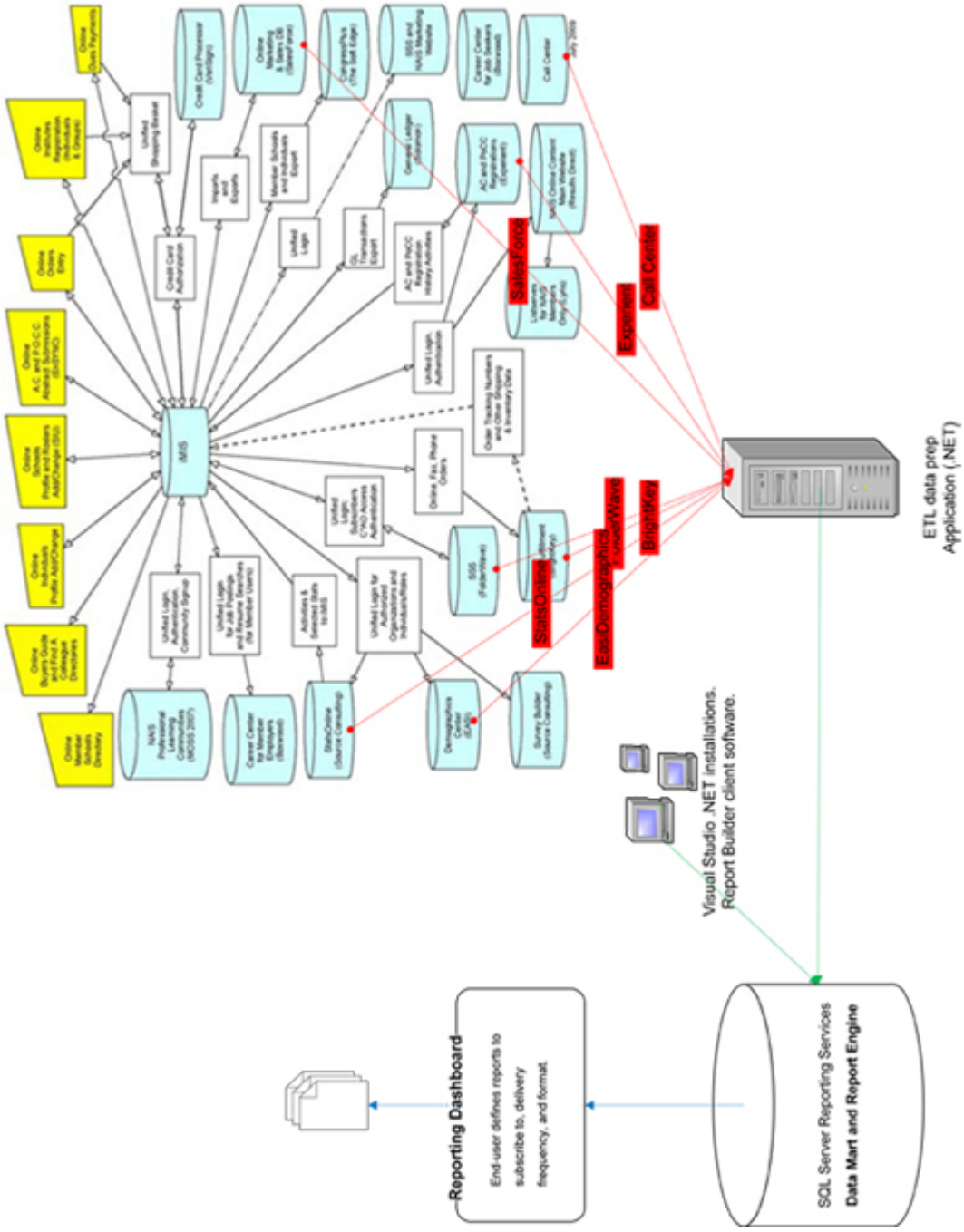


CLIENT INITIALS

10 Notes, LLC • 1725 I Street, NW, Suite 300 • Washington, DC 20006
tel 202.349.1115 • fax 202.280.1341 • web www.10notes.com

7.2.3. Reporting Data Mart.

Notes: The diagram below illustrates the Reporting System as related to its data sources.



CLIENT INITIALS

10 Notes, LLC • 1725 I Street, NW, Suite 300 • Washington, DC 20006
tel 202.349.1115 • fax 202.280.1341 • web www.10notes.com