WCWWA 2007 Conference & Trade Show
October 23 - 26, 2007
Edmonton, Alberta

IMPLEMENTING A REGION-WIDE MOBILE SOLUTION FOR WATER AND WASTEWATER DATA COLLECTION AND QUALIITY ASSURANCE

M. Branch[1] and B. Bradley[2]
[1] Inovex Inc., Software Architect, Oakville, ON
[2] Inovex Inc., Software Architect, Oakville, ON

ABSTRACT

Recent technological advances in hardware and software, such as Tablet PCs and Smart Client architecture, have facilitated the development of Water and Wastewater data acquisition applications for use in disconnected environments. Smart Client packages provide end users with the connectivity of a web application and the rich user experience of a Windows Form application on a mobile device. The days of providing operators with clipboards or maintaining volumes of spreadsheets are behind us with these emerging technologies. Using modern software packages, Water and Wastewater divisions can empower their operators and in turn reduce data entry time, eliminate erroneous data entries, and greatly increase efficiency.

Today's operators are equipped with ruggedized, waterproof Tablet PCs for remote data collection. On a daily basis, the operators collect data on-site using the Tablet PC in a disconnected mode. Custom input controls present in the software recognize the operator's handwriting and convert it to digital information, thus making the use of the Tablet PC completely analogous to writing with pen and paper. Operators gain immediate feedback when entering readings that could lead to erroneous data or when events such as meter rollover and/or replacement occur. Collected data is immediately validated against associated thresholds and rules.

Information is stored locally on the Tablet PC until operators synchronize collected data with a central server, and download any necessary updates. The process of data synchronization is transparent to the end users and provides a central volume of collected data for real-time reporting, querying, and data mining. Complete user, data, and application administration tasks are performed through a web application and updates are sent to individual operators upon data synchronization.

Both the Tablet PC and web components of the applications provide users with the capability to create real-time trends in various formats. Operators can quickly and easily assess performance and safety metrics using the collection of fully customizable trends.

It is becoming increasingly important for Water and Wastewater divisions to be aware of the available technologies that will make their workforce a mobile one. This paper presents York Region's Department of Water and Wastewater as a driving force behind the development of a Water and Wastewater mobile data collection solution.

INTRODUCTION

Offering the highest level of water quality services to the public is undoubtedly one of the most critical outcomes of each municipality's water and wastewater operations. Of equal importance to water quality, however, are data quality, integrity, and dissemination to the appropriate audience in an expedient and salient manner. This paper will take you through a case study at York Region's Department of Water and Wastewater, clearly illustrating how they were able to develop a robust mobile system for their water and wastewater operators to effectively collect, and interpret field data, while sharing that data with billing, compliance, and management via a collection of real-time reports and trends.

RATIONALE FOR CHANGE

Prior to August 2007, York Region operators were collecting field data through a mix of Excel spreadsheets and traditional paper-based methodologies from seven different hubs, spanning nine municipalities, and servicing 900,000 residents, over 1800 square kilometers. Having such a large geographic span, and nearing 200 different facilities, York Region was faced with several challenges to maintain its exceptional record of water quality:

- Collection methodologies varied between paper and pen, and electronic capture on MS Excel (via Tablet PC in the field)
- Facilities in the north are far more rural, and operate differently than southern facilities, where there is close proximity to the city of Toronto
- Excel data collection templates varied from hub to hub, causing difficulties in standardization across the Region
- Synchronization issues (lost data, dual entries, etc.) were prevalent with Microsoft Excel
- Lack of detailed mechanisms for granular user-level data auditing
- Lag times between collection of field data, and alarm indicators for data points outside of compliance limitations

The decision was therefore made to implement a standardized, region-wide methodology for data capture and reporting, that would fulfill the needs of all stakeholders. While a certain percentage of the system requirements could be fulfilled by an automated Supervisory Control and Data Acquisition (SCADA) system, there were several reasons why automated collection could not meet the Region's needs:

- Not every facility in the Region was equipped for SCADA (remote location, equipment updates, etc.)
- Not every data point that is collected through the Tablet PC interface is capable of being collected by a SCADA system
- Values collected from the SCADA system are based on calculations which do not always accurately reflect what the operator is visibly able to see and record at the facility in question

A mobile solution for the Region's operators was therefore required to fill a gap where SCADA systems were unable to fulfill requirements due to accuracy, timeliness, and expense.

DEVELOPMENT PROCESS

Because the much needed mobile system would be used predominantly by operators on a daily basis, it was decided to include them as a key component throughout the development process. The project began in October 2006 by identifying the needs of all key stakeholders in the system, including: Operators, Superintendents, Team Leads, IT, Compliance, Management, Billing, Accounting, Maintenance, and Capital Delivery. Structured interviews and ride-a-longs were conducted with identified proponents in each of the seven hubs in order to clearly ascertain a comprehensive user requirements document (URD).

Once complete, the URD was used as a basis to formulate screen shots of the entire proposed application. In doing so, the operators were intimately involved in helping to create a look-and-feel that would be most salient to them.

After the completion of all software architecture models and design documentation, the development process began in December 2006, and continued to the beginning of April 2007 (when the application was officially launched in pilot). A single hub was selected to pilot the application, where it was periodically used for four months (in conjunction with their existing Excel-based system) to test for usability, data verification, and performance. After several iterations in pilot, the application was fully released to all hubs for the August 2007 billing cycle.

MOBILE ARCHITECTURE

The disconnected nature of the data collection process made a traditional networked application impractical. To meet the Region's requirements, a Smart Client architecture was used to build a flexible, robust water and wastewater management system (see figure 1 below). A Smart Client Architecture consists of a collection of disconnected client machines operating independently on their own local databases, with the capability of synchronizing on demand with a larger central database. The Smart Client architecture provides the Region all of the robustness and speed of a local windows-based application while still benefiting from the connectivity of a web-based solution. This is attained by maintaining smaller, localized versions of the application data on each Tablet PC, and synchronizing this data with the central server.

The Command and Control Centre web portal is the interface to the central data warehouse where all of the synchronized data resides. Data collected by each individual operator is propagated to the site upon synchronization, and made available for analysis and reporting. In addition to gathering information from the Tablet PCs, the web portal provides users with the ability to push new or updated information out to all operators. Region users can create new water treatment facilities, collection points, calculated formulas, data collection schedules, etc. and publish this information to the Command and Control Centre for propagation to field users.

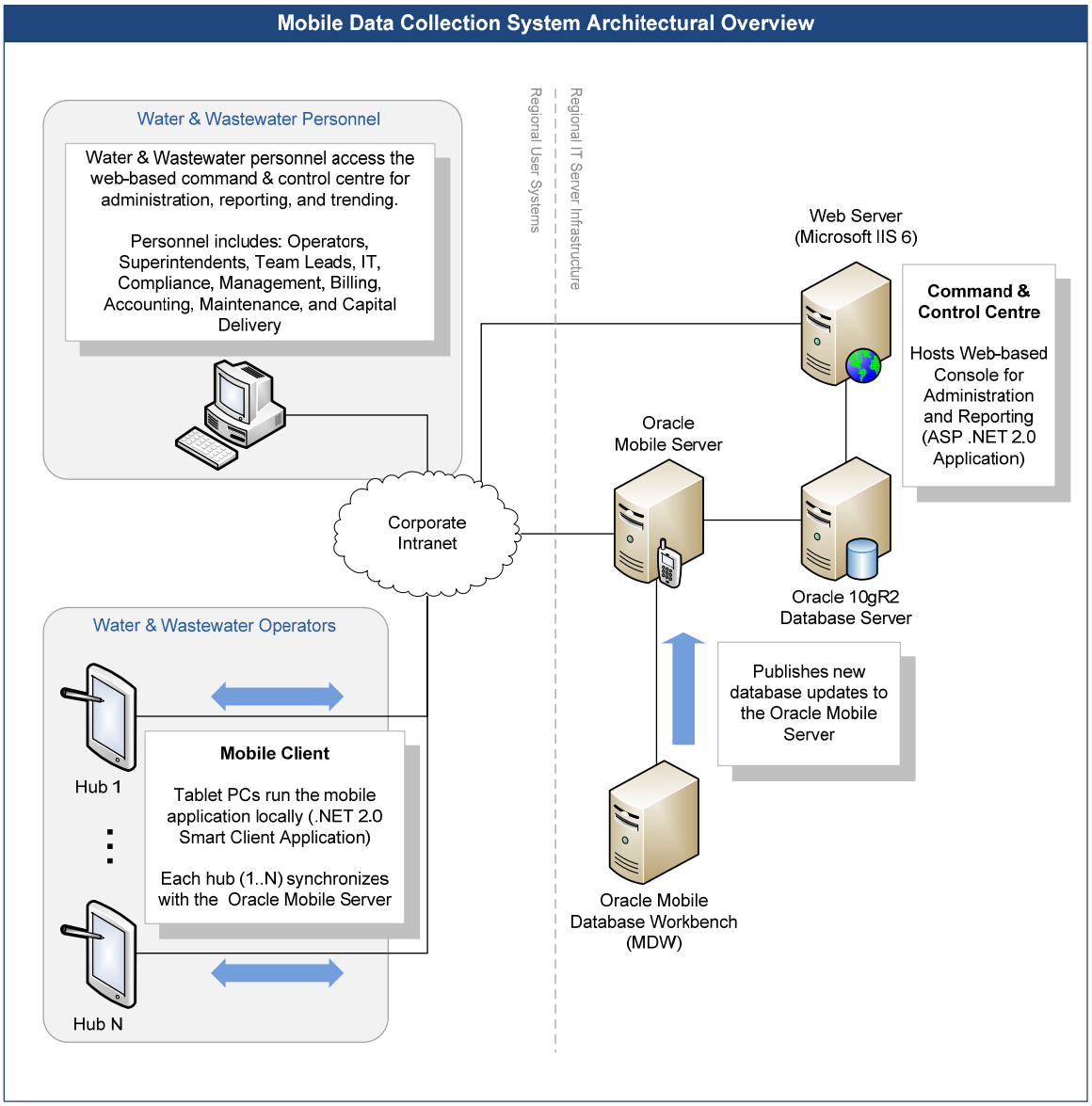

Figure 1: Mobile Data Collection System Architecture

MOBILE CLIENT

The Mobile Client is a light-weight Windows-based application that will operate on any laptop, desktop, or Tablet PC. It synchronizes with the Command and Control Centre automatically and by the user's request. It was designed to include both data collection features and trending

features (as outlined below).

Data Collection

The interface for data collection allows the user to page through data points for a particular facility on a particular day. Depending on the facility, the user may only have to collect a few data points, or may have to collect a number of pages of data points (see figure 2 below). These data points can be grouped and ordered into specific sections (such as "Chemicals", "Flow", "Residuals", etc.) and can consist of multiple different types: numeric, drop down lists, text, or check boxes.

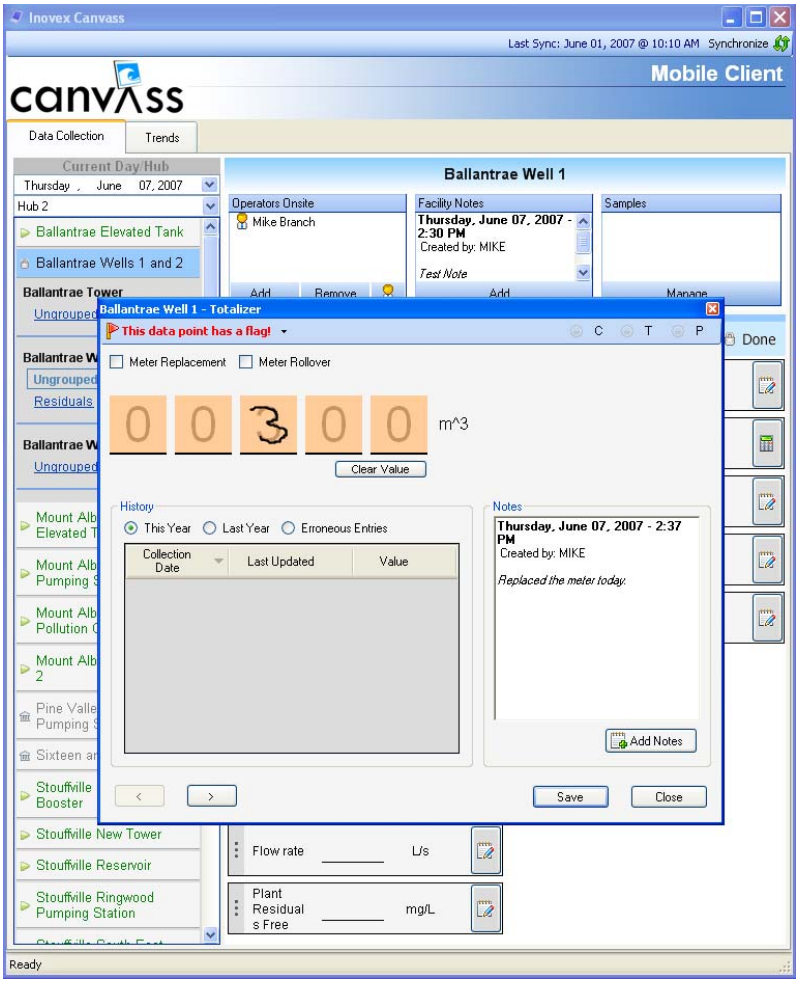

Figure 2: Data Collection Interface

Once the operator chooses to record data for a particular point, the input validation screen is displayed illustrating a plethora of information specific to the data point at hand, including:

- Short-term data history (tabular and trend-based)

- Threshold warning indicators
- History of erroneous data entry
- Notes associated with the current data point
- Flags associated with the current data point
- Automatic detection of: meter-rollovers, meter replacements, and chemical top-ups/cylinder changes

In addition, all screens have been optimized for the Tablet PC, and include custom ink recognition controls to facilitate the ease of data entry. Fields that are numeric in nature will only recognize numeric values, and numeric entries will scale to the number of significant digits and decimal places required for specific data points.

Trending

Operators in the field have an integrated dynamic trending component with the data collection application (see figure 3 below). Field workers have the capability to instantly trend any data point(s) they collect and compare current values to historical values in the system. Operators can save common trend templates and share them with other users in the field. Commonly, an operator would trend the flow at a particular facility, checking for any anomalies or changes that would not be apparent when viewing data in a tabular format. In this manner an operator can quickly find and mitigate any potential risks immediately at the first sign of a poorly shaped trend or scatter plot.

In synchronizing their data with the Command and Control Centre, the field operator passes this information and these trends along to the central office. They can choose to share trends, create templates, and even use trends created by other system users. They can also direct management to view a trend that they feel requires immediate attention.

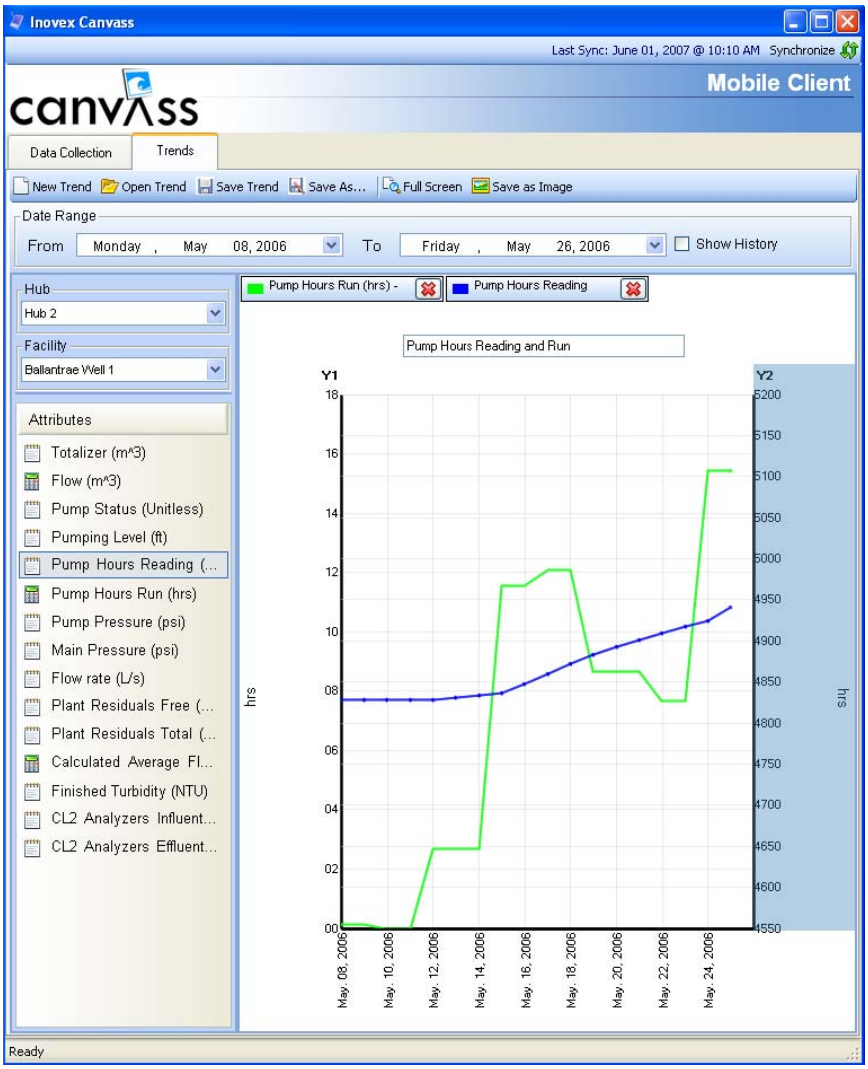

Figure 3: Dynamic Trending Interface

Sample Collection

Samples are collected at various intervals depending on the facility visited and time of year (as dictated by the Certificate of Authority for each facility). For this reason, the application supports the collection of samples for any visit to a particular facility – the precise interval is left to the judgment of the operator.

Each sample collected has an associated field identifier which directly corresponds to a data point at the facility. In addition to containing a field identifier, data points marked for sampling must also have a related sample type (raw water, distributed system, etc.). See figure 4 below for a screenshot of the sample collection screen.

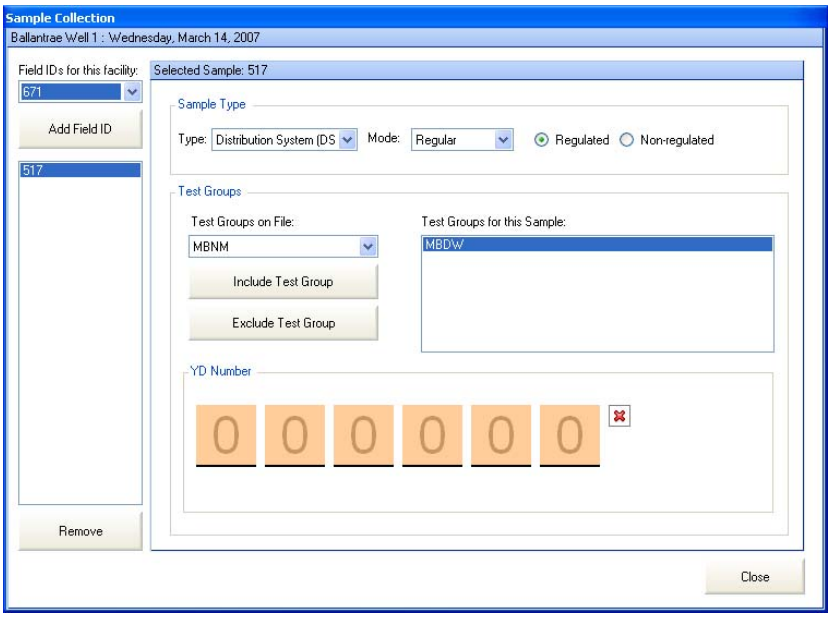

Figure 4: Sample Collection Screen

Once collected, samples can then be associated with one or more test groups and can have multiple regional lab numbers associated with them (depending on whether or not re-sampling was performed, or another test group was required). At the end of the sample collection process, the system will automatically generate the necessary Chain of Custody documents on a per-facility basis which can be emailed to the lab for further processing.

Audit Trails

Every action that is taken within the system is stamped with the user who created or modified the data, as well as the date and time the data was last changed. In the event any erroneous data is propagated to system reports, the system's audit trail gives management the necessary information to understand how data arrived at its current state.

Not only does the system have the ability to track data changes, but it also tracks certain events (such as the commencement and termination of data collection at each facility). Prior to collecting data on a particular facility visit, the operator must indicate that they will be starting data collection (otherwise, controls on the screen are disabled). In a similar fashion, the operator must also indicate when the data collection process has stopped (otherwise, they will be unable to continue collecting data at subsequent facilities). This mechanism allows management to create effective route schedules for their operators by understanding the amount of time taken at specific facilities.

In the event that the operator must make a change to the data collected after it has been "signed

off" for the day, a valid reason for the change must be supplied. The historic value is then expired and recorded with the change reason, while the new value is made current.

COMMAND AND CONTROL CENTRE

Managing Data Points

As facilities change, data collection points will need to be added, altered, and removed. All of these tasks and more are handled centrally by the Command and Control Centre. It is through this application that new locations and collection points are added to the Region. The administrator will publish collection points with the necessary metadata including unit of measurement, significant digits, number of decimal points, physical boundaries, tolerance boundaries, and compliance boundaries, among others. These new changes will take effect when the operator next synchronizes their Tablet PC with the central server. When the operator begins collection, they will see this new collection point as part of their run.

In addition to metadata, data collection points are also associated with a collection of data profile types. Data profile types identify the behavior of the data collection point. Using data profile types, the administrator can specify what type of data is being collected, whether the data point can have meter rollover or replacement associated with it, and whether it participates in chemical top ups or cylinder changes.

Calculations

In addition to the standard data collection points, an administrator may also create calculated collection points. These collection points will appear as part of an operator's daily run, but will automatically be calculated once the operator has completed collecting all dependant data point. The Region uses two types of calculated data points: custom and accumulator.

Custom calculated data points allow administrators to specify a custom mathematical formula in which operations are performed on any number of other data points at a given facility. These formulas are usually used to perform measurement conversions, average a collection of points, or calculate plant dosages on demand.

Accumulator data points were designed specifically by the Region to automatically calculate the daily rate of change of another collected data point. The most common use for accumulator data points is to calculate the total daily flow and chemical consumption at a given facility.

Alarm Thresholds

One of the most valuable components of this system lies in its ability to give the operator the ability to interpret data on the spot and determine whether or not certain thresholds are being upheld. Specifically, each data point in this system is equipped with a set of metadata that

defines a set of alarm thresholds in the form of physical boundaries, tolerance boundaries, and compliance boundaries. These can be set from the administration console on a granular basis at the data point level, or can be set for specific types of data points (i.e. flow within all facilities in hub 3 must not exceed 1,000,000 m^3 / day). In either case, the operator is immediately notified of the alarm threshold(s) being violated by on-screen, colour-coded indicators.

As mentioned above, these thresholds fall into one of the following three categories:

Physical Thresholds

These are numeric thresholds that represent hard upper and lower boundaries for data points. For example, the number of hours in operation for a particular pump must always be a positive number; therefore, the lower physical boundary limitation is zero.

Tolerance Thresholds

Tolerance thresholds help to raise awareness about data points that seem to be out of the normal range of values within a short-term average. For example, let us take the data point "chlorine consumption" and assign it a tolerance of 5% above and below average. If consumption has reached an average of 100 m^3 /day over the last month, this means that the tolerance thresholds will throw a warning if the values being collected are outside of the range 95-105 m^3 /day.

Compliance Thresholds

These are thresholds usually imposed by the Ministry of Environment, related to Certificate of Authority (C of A) limitations. For compliance purposes and reporting to the Ministry, it is important for operators, management, and compliance personnel alike to be warned as soon as C of A limitations have been violated.

Both physical and tolerance thresholds help prevent against errors in number transposition during data entry, such as when an operator enters 1000 for a particular data point instead of 100. Erroneous data may still exist within the system, but having these checks and balances in places helps to reduce the margin of error within the system as a whole. From a reporting standpoint, regular data alarm reports can be run which clearly identify exactly where thresholds are being violated, along with the type of threshold (in a salient, colour-coded fashion).

Data Point Scheduling

Not all data points being collected throughout the Region are collected on every scheduled run day by operators. For this reason, the administration console gives superintendants and team leads the power to schedule data points for collection for virtually any time period desired. For example, if flow metrics are gathered on Monday, Wednesday and Friday during the winter, and on a daily basis throughout the spring, this can be accommodated through the administration

console. Once data points have been scheduled and saved for various facilities, the next time operators synchronize their Tablet PCs with the system, they will have their updated schedules intact.

Command and Control Analytics

The Command and Control Centre has access to data submitted by all field users and, as a result, users can perform analytics on all data Region-wide. The same trending component found on the Tablet PCs is found in the Command and Control Centre. Trends created and saved by operators on the Tablet PCs can be shared with all intranet users.

In addition to trending, the Command and Control Centre houses a robust reporting engine. Intranet users now have access to a large repository of custom reports that pull any number of statistical measures from their centralized database on demand. Custom reports can be turned around quickly and easily, can conform to current standard regional reports, and can be published in a number of industry standard formats. Moreover, since data is synchronized from all operators to the central store on a regular basis, the reports can compare data from across the entire region.

The region leverages these custom reports to perform tasks including: output their daily flow rates, generate billing information for water usage by other regions, show all alarms and issues that have been recorded over a time period, etc. Many custom reports have also been designed to meet with the Ministry of Environment regulatory reporting requirements.

BENEFITS

Since implementing this mobile solution, the Region has noticed improvements in many areas of their business process, some of the key benefits are outlined below:

Improved Data Integrity

- Specified compliance regions notify end users of potentially erroneous data values
- Audit trails makes end users accountable for data they record and/or change
- Data is stored using modern database techniques, eliminating the need for a collection of independent spreadsheets
- Updates to published data force end users to record a reason for change

Powerful Operator Toolset

- Tablet PC interface provides users with pen and paper feel with handwriting recognition
- Calculations allow operators to automate data conversions, perform complex mathematical operations, and generate daily rate of change data for daily flow and chemical consumption

- Meter rollover, meter replacement, and chemical top-up are automatically detected and appropriate calculations are performed
- Trending tools give operators charts on demand to better visualize changes to collected data
- The sample collection interface drives the necessary reporting for lab analysis

Improved Data Visibility

- A highly flexible and scalable reporting solution gives management detailed reports on demand and affords the Region the ability to generate new reports as changes in their business process dictate
- Centralized trending tools generate comparative trends across facilities to identify system-wide problems
- Standardized reports are automated for submission to the Ministry of Environment and other governing bodies
- Data collected in the field is propagated to a central data store ensuring reports contain the most recently collected information

Simple Data Management Interface

- Intuitive Command and Control Centre is used to make all application changes
- Changes and additions to the data are seamlessly pushed out to field users upon synchronization
- Scheduling of data collection is handled centrally for all locations within the region

ACKNOWLDEMENTS

The authors would like to acknowledge the efforts put forth for development of the system by Inovex software developers Bernard Cheng, and Daniel Ng. The authors also extend their thanks to York Region operators, management, and IT staff, specifically Simon Yates, Sue Kelsey, Mark Williamson, and Richard Newton, for their combined effort toward ensuring the success of this project.