CONCEPT OF OPERATIONS

For the

CASCADE GATEWAY BORDER DATA WAREHOUSE 3.0 UPGRADE

And the

WHATCOM COUNCIL OF GOVERNMENTS

Date	May 5, 2020
Version	1.0
Prepared by	Whatcom Council of Governments 314 E. Champion Street Bellingham, WA 98225 (360) 685-8385 melissa@wcog.org

1.0 PURPOSE OF THE DOCUMENT

The purpose of the Concept of Operations is to describe the expected operations of the Cascade Gateway Border Data Warehouse Version 3.0 (BDW3.0) and the rationale behind the development. It will help guide design and implementation of the final system and is intended to be used by project stakeholders including the project management team, software developers, and advisory team.

2.0 SCOPE OF THE PROJECT

The BDW3.0 will build on investments made in 2016 to integrate U.S. Customs & Border Protection's (CBP) booth status data system into southbound wait time calculations by upgrading the existing border wait time data warehouse located at www.borderdata.org. The database stores passenger and commercial vehicle wait times, volume, and other data relating to vehicle traffic at the four land border crossings between the Lower Mainland of British Columbia and Whatcom County, Washington State. The objective of the database is to provide timely and accurate access to this data for a broad range of stakeholders.

This project will:

- Add booth status data to currently maintained datasets
- Improve warehouse performance and output
- Upgrade commercial vehicle datasets

3.0 REFERENCED DOCUMENTS

Documents relating to this project include:

- Project Traceability Matrix
- Cascade Gateway Border Data Warehouse 3.0 Design Document
- System Validation Plan

4.0 BACKGROUND

The existing Border Data Warehouse (BDW2.0) is a proven data solution that provides critical metrics to multiple border stakeholder agencies including federal and regional transportation agencies, inspection agencies, planning organizations, and municipal governments.

Applications using these data are numerous and include traffic models, border throughput simulation models, staffing allocations, reporting by agencies on delays, transportation agency performance metrics, academic research, and more.

However, since its inception in 2007 several challenges have developed that need to be addressed:

- **Data accuracy**: because the current version doesn't use the newly available booth status data from CBP, it is erroneously archiving wait time for NEXUS and non-NEXUS passenger cars.
- **Increasing hosting expense**: because the booth status data is archived in raw format (a data field for every car) and not binned, it is growing much faster than before and costing more to host.
- **Loss of U.S. Bureau of Transportation Statistics (BTS) compatibility**: since the BTS converted its transborder freight data query tool to Tableau it is no longer possible for the warehouse to "screen scrape" relevant regional data for local queries.
- **End of maintenance funding**: this project is the only ongoing archive of five-minute increment border wait times going back to 2007. Keeping this resource maintained and hosted is very important to regional stakeholders.
- **No way to access to new datasets**: the booth status data and archived commercial wait times are not currently accessible on the existing website, even though the data are being saved.

5.0 CONCEPT FOR THE PROPOSED SYSTEMS

This section describes system improvements that address the challenges listed above. Each improvement relates to a system requirement identified at the outset of the project when stakeholders provided feedback on the current BDW.

Requirement 1: The warehouse needs to categorize data for each booth and lane dynamically, using the booth status data.

Solution 1.1: The warehouse shall be restructured so that data are categorized by processing type rather than by lane.

Solution 1.2: The warehouse shall recategorize historic data to fit the new dynamic architecture.

Requirement 2: Southbound booth status data needs to be accessible for queries, reports, and visualizations.

Solution 2.1: Booth status data shall be queriable in the custom query tool.

Solution 2.2: Booth status data shall be displayed in the primary crossing visualizations.

Solution 2.3: Booth status data shall be available in the reports section.

Requirement 3: Southbound booth status data needs to be binned appropriately to reduce the size of the storage requirements and improve functionality.

Solution 3.1: Booth status data shall be binned in five-minute increments.

Requirement 4: The warehouse needs to update its connection with the U.S. Bureau of Transportation Statistics freight data to show filtered Cascade Gateway data results.

Solution 4.1: The warehouse shall provide a feed of regionally relevant freight data from the BTS transborder surface freight database (if an API is made available by the BTS site).

Requirement 5: The warehouse needs an updated subscription service for notifications of border delays.

Solution 5.1: The warehouse shall have a subscription tool that allows users to define email/text preferences for border alerts.

Solution 5.2: The warehouse shall update all security requirements to allow for safe subscription practices.

Requirement 6: The warehouse needs an API registration function to provide contact information for agencies/organizations using the API resource.

Solution 6.1: The warehouse shall require those using the API to register for a key so that contact information can be collected on developers using the tool.

Requirement 7: The warehouse needs an updated visualization interface.

Solution 7.1: The warehouse shall maintain existing features of mapping, displays, reporting, and custom query tools, but updated to currently available software capabilities and visualizations.

Solution 7.2: The warehouse shall use Tableau to create its data visualizations and data sharing tools.

Requirement 8: The warehouse needs a dynamic web interface that detects device type for optimal layout.

Solution 8.1: The warehouse shall be designed on a web platform with a dynamic display based on device type (computer, tablet, phone).

Requirement 9: The warehouse needs to update the loop detector data in the back end of the archive.

Solution 9.1: The loop detector data shall be reviewed and updated as necessary.

Requirement 10: The warehouse needs an improved loop detector interface system for maintaining loop detector records.

Solution 10.1: The loop detector data input tool (back end) shall be upgraded to make it easier to locate and update information.

Solution 10.2: The loop detector input tool shall be upgraded to make it easier to select one or a series of loops to be used for volume counts.

Solution 10.3: The loop detector data input tool shall use improved mapping/graphics to place each loop on location.

Requirement 11: The warehouse needs to send a simplified daily email to system administrators reporting if all data feeds and backups are complete.

Solution 11.1: The warehouse shall send a simplified daily email with symbols showing whether all data feeds were received.

Solution 11.2: The warehouse shall send a simplified email stating if all backups were completed effectively.

Requirement 12: The warehouse needs an improved backend reporting system to provide monthly status reports and logs to system administrators.

Solution 12.1: The warehouse shall have a backend reporting tool to make it easier to compile monthly maintenance logs.

Requirement 13: The warehouse needs an improved method for backing up existing and backfilling missing data.

Solution 13.1: The warehouse shall have an updated backfill process for when data needs to be added in bulk.

Solution 13.2: The warehouse shall evaluate and implement an improved backup routine.

Requirement 14: The warehouse needs an improved help section function in the backend for system administrators to easily update each page/section.

Solution 14.1: The warehouse shall have an improved help page administrative feature that shows completed pages, and which are missing.

Solution 14.2: The warehouse shall have an improved interface for adding text and images to the help sections.

All requirements listed above will be completed by IBI Group, overseen by the Project Manager at the Whatcom Council of Governments, and reviewed by the Project Advisory Team.

6.0 USER-ORIENTED OPERATIONAL DESCRIPTION

Currently, the system functions as described in the Design Document. A collection service is responsible for polling feeds from partner agencies and capturing detector data. That data is identified as crossings, lanes, and detectors by the XML schema. These relationships are maintained within the database. The collected data is associated with the corresponding lane configuration allowing for users of the website and APIs to query data and generate reports for end users.

Stakeholders

Primary system stakeholders include:

- **B.C. Ministry of Transportation & Infrastructure**: Data providers and system users
- **Border Policy Research Institute, Western Washington University**: Power system users
- **Canada Border Services Agenc**y: System users
- **IBI Group**: System developers and maintenance team
- **Transport Canada**: Project funding partners and system users
- **U.S. Bureau of Transportation Statistics**: Data partners and system users
- **U.S. Customs & Border Protection**: Data providers and system users
- **U.S. Federal Highway Administration**: Project funding partners and system users
- **WA State Department of Transportation**: Data providers and system users
- **Whatcom Council of Governments**: Project manager and system users

7.0 OPERATIONAL NEEDS

The requirements of the system are outlined in **Section 5.0** of this document.

8.0 SYSTEM OVERVEW

The new system will have an updated extract-transform-load engine (ETL) that will improve notifications to system administrators about data source availability and system health. To optimize performance and support improved querying capabilities, there will be separate repositories for reporting data and storing data. New external system status tools will generate notifications if any service becomes unavailable and unable to trigger alerts.

The new database design will incorporate the booth status lane identification components into the back end of the architecture. This update will improve processing of incoming data as well as categorization and correction of historical data.

The system will be upgraded to the most recent version of .NET technology for improved performance.

Lastly, Tableau Server will be used to replace the existing data visualizations on the website for end users to analyze data outputs and download results.

A full list of envisioned system upgrades is discussed in the Design Document.

9.0 OPERATIONAL ENVIRONMENT

Equipment and Hardware

All hardware and equipment for collecting the data that the system uses is owned by partner agencies – B.C. Ministry of Transportation and Infrastructure, and WA State Department of Transportation. Therefore, no changes or additions to hardware or equipment are part of this project.

Software

BDW 3.0 will be hosted at its current location, on an instance at Amazon Web Service (AWS). Configuration of the instance will be determined after an evaluation of costs and benefits. The development team at IBI Group will evaluate and recommend services for backups and maintenance.

Personnel

Whatcom Council of Governments

Melissa Fanucci has managed the Cascade Gateway Border Data Warehouse since its initial development in 2007. She also manages the Whatcom Regional ITS Architecture and is Principal Planner at WCOG, helps coordinate the International Mobility & Trade Corridor (IMTC) Program, and has been involved in other ITS-related border projects for the region.

IBI Group

A description of all personnel working on the project, and their skills and training, is available in the IBI Group RFP Document.

Operational Procedures

The system will automatically update border volume, wait time, and other data in five-minute increments using the new ETL engine. Data will be immediately available for querying through the front-end warehouse at the website borderdata.org. The warehouse will query the day's data as well as use an extract of historic data that has been pulled and stored to streamline the process and reduce demand on the primary database.

If there are any errors in the ETL process or missing data, the system will notify managers at WCOG and IBI Group through an external data notification system. The missing data will also be logged to be part of the maintenance logs for the system.

At the end of the day, a backup process will be completed (to be determined later exactly how this process will be completed most effectively). Currently a full backup of the entire database is completed every night. The future version may make a backup of only new data compiled during the twenty-four-hour period since the prior backup.

Ongoing maintenance will be completed but the design should keep requirements to a minimum. Monitoring of resource requirements, backup health, the size of the databases, and keeping a focus on incoming data feeds are the top priorities for ongoing maintenance.

10.0 SUPPORT ENVIRONMENT

The system is supported by the existing infrastructure already in place and funded by partner agencies WA State Department of Transportation, B.C. Ministry of Transportation and Infrastructure, and U.S. Customs & Border Protection. IBI Group has been selected through a request for proposal process to complete the work on the project by December 2020. Once the project is completed, the Whatcom Council of Governments will continue to administer the database and establish a maintenance agreement with an appropriate entity to keep the database running.

The Whatcom Council of Governments will also be responsible for hosting the archive at AWS. Funding for this project is for the build of BDW 3.0. Funding to maintain the archive will be sought upon completion.

11.0 OPERATIONAL SCENARIOS

Operational scenarios help convey what is expected to be achieved with the implementation of this project. The following scenario has been developed with the members of the project advisory team in mind to illustrate how the warehouse will function. "System Manager" refers to Whatcom Council of Governments. The following operational scenarios are included:

- Routine operation
- Failures and other unusual events
- Upgrades

Routing operational scenarios

Researcher compiling statistics

A researcher at a university is looking at the correlation of border traffic and the price of milk in the United States.

After gathering daily average milk prices for Whatcom County in Washington State for 2018 and 2019, the researcher then visits www.borderdata.org and goes to the custom query section of the website. She selects January 1, 2018 – December 31, 2019 and specifies all days of the week, all months, but selects only the primary hours of travel (0800 – 2000). She chooses all ports-of-entry and both directions, and all passenger vehicles as she is not interested in whether they are NEXUS, standard, or Ready lane.

She then chooses average wait time then hits download. Once the data are downloaded as a .csv file, the researcher can now compare the data with the milk prices and run an analysis. If the researcher wanted to compare these results with other years, she can change the fields in the custom query to change the date range. Or she may choose to filter the southbound data by license plate province and view BC data only, seeing only the wait times experienced by those in that province.

Because the data are downloadable by a variety of measurements, the researcher can get the data directly in a format she needs to run an analysis specific to her requirements.

Transportation agency staff

A staff member at the state or provincial transportation agency is interested in seeing how construction along the highway approaching one of the regional land ports-of-entry is impacting border wait times.

He logs onto www.borderdata.org and on the primary dashboard, is able to quickly view wait times at all four Cascade Gateway ports-of-entry simultaneously. At a glimpse, he can see how the crossing in question compares to the other crossings in the region.

However, this crossing typically has lower traffic, and shorter wait times, than the other crossings in the region, so a historic comparison will be more valuable. Therefore, he chooses to select "compare to last year" on the dashboard. This shows how the wait times at this crossing in the southbound direction are noticeably different than they were one year ago at the same time. To glean more information about the scenario, he can then choose "expand this data" to go to another data visualization that will expand the view and give him a more detailed analysis of the historic trends of this crossing and allow him to pinpoint when and where the delay is taking place, and when it started. If it started in the same time frame as construction on the highway, he may be able to make a correlation between the two.

To do his own analyses, he then chooses to download the data.

Inspection agency staff

An inspection officer in charge of resource allocation signs up for the email subscription service that alerts her any time wait times at a crossing exceed forty-five minutes. As soon as she receives the email, she realizes that there is a substantial delay and can notify her staff before the wait time exceeds sixty minutes.

Failures and other unusual scenarios

System manager gets a data loss alert

A system manager checks her email in the morning and sees an alert from the BDW 3.0 notification system that shows that one of the expected data feeds is no longer providing data.

The report also states when the data stopped coming into the system and confirms the error is in receiving the data, not in processing the data. This allows the system manager to contact her partners at the respective data providing agency and inquire as to the status of the data feed.

System manager gets notified of the backfill of data

When data are missing in the system, the system will attempt to locate the missing data the next time in connects to the data feeds of partner agencies.

The system manager receives an alert on the system health status dashboard that notifies him that a backfill is being performed on a specific date. The backfill start and end time is listed, as well as listing what data date range and ports-of-entry are being backfilled. This allows him to note if it matches any data losses noted with data gaps he was alerted to earlier.

Upgrade scenarios

System managers

After the system has been running, another agency contacts the system managers with a proposal to share border-related data on the website and archive that data in the database.

The system manager will work with the maintenance team to:

- Evaluate if the new data service is an enhancement that improves an existing function, or provides additional functionality
- Determine if the data feed fits the existing XML schema
- Estimate how much customization/changes will be needed to fit the data feed into the existing architecture and outputs
- Determine total costs and timeline to implement the addition

Once a cost estimate is developed, the system manager would work with project stakeholders to evaluate the cost benefits of the proposed inclusion of the new dataset. Should the new dataset be deemed a worthy inclusion, and funding identified, the system manager would work with the maintenance team to develop a plan for integrating the new feed into the existing architecture. Upgrades to the system will be catalogued in the same manner as suggested by the software developers and system managers as per the original project.

12.0 SUMMARY OF IMPACTS

The following lists potential impacts of the proposed system on each project stakeholder:

As the system manager, the **Whatcom Council of Governments** will work with the system developers (IBI Group), Project Advisory Team, and end users to oversee the entire project from start to finish. WCOG also has the responsibility to maintain the archive after completion for at least one year of hosting and maintenance beyond the budget of the original project, and to seek additional funding for ongoing maintenance of the archive. WCOG will continue to work with end users to assist in queries; to develop system engineering documentation; to develop help articles and products for making the solution more accessible; and to provide outreach and education about the process of development and end results to help others developing border warehouse archives.

IBI Group, as the system developers, will be leading the design and software development process for the new archive. In addition to maintaining the original archive, IBI Group will be transferring over the data, testing the new version, and developing a maintenance plan at the completion of the project to successfully move it forward upon completion.

BC Ministry of Transportation and Infrastructure, U.S. Customs & Border Protection, and the **WA State Department of Transportation** will be asked to continue to share their data with the Cascade Gateway Border Data Warehouse system.

U.S. Bureau of Transportation Statistics will be asked to assist in developing an applied programming interface (API) between their database of transborder freight statistics and the BDW 3.0 in order to make sharing between the two archives easier.

Performance metrics

The following are potential measures of the system's performance:

- Completion of all 14 established project requirements
- Availability and access to data
- Presentation of the data in a clear and concise format
- Ease of maintaining the warehouse
- Overall functionality of the website