

## *NatCDC versus NatCDCSP*



NatWorks Incorporated  
454 South Main Street  
Northfield, VT 05663  
(802) 485-6112

Website: [www.natworks-inc.com](http://www.natworks-inc.com)  
E-Mail: [info@natworks-inc.com](mailto:info@natworks-inc.com)

**Table of Contents**

Overview ..... 3  
Statement on NATCDC ..... 4  
Introduction of NATCDCSP..... 5  
Summary of Differences & Changes ..... 7  
    Generate NATCDC Objects Function Modified ..... 7  
    Generate NATCDCSP Objects for Multi File Function Added ..... 7  
    Different JCL Templates..... 8  
        Review of NATCDC Processing ..... 8  
        Overview of NATCDCSP Processing ..... 8  
    Different Processors..... 9  
    Different Parameter Handling..... 9  
    Different Output..... 10  
Continuing Use of NATCDC..... 11  
Migrating from NATCDC to NATCDCSP ..... 12  
Creating New Processes Using NATCDCSP ..... 13

## Overview

With the release of NatQuery version 5, new support was added for NATCDCSP, with NATCDCSP representing the next generation of the current product NATCDC. The intent of this document is to clarify the use of NATCDC versus NATCDCSP, and to outline the enhancements found in NATCDCSP which differentiates it from its predecessor NATCDC.

Historically, NatWorks only offered one solution to ADABAS PLOG processing, and that solution was and is called NATCDC.

If there was any one major deficiency with NATCDC – it would be that NATCDC was designed to only be “single file” processor. This meant that if there were 100 files against which ADABAS PLOG processing was needed: There would by necessity have to be 100 individual executions of NATCDC, each with its own processing Parameters, to provide this data.

While all of these NATCDC processes could be “fed” from a single execution of ADASEL (mainframes) or ADAPLP (UNIX and Windows), and they could be job-streamed together or run concurrently; the fact remained that as PLOG processing grew, the more individual processes there would be, and the sheer number of individual processes could eventually prove cumbersome to manage – despite our ability to generate everything needed for that processing.

The solution to that problem was NATCDCSP, which NatWorks released late in 2006 in conjunction with version 5 of NatQuery.

All existing NATCDC customers active with maintenance will receive NATCDCSP free-of-charge, these customers can use one, the other or both, and NatWorks will actively support the use of both solutions under the NATCDC maintenance umbrella.

In summary: The NatWorks approach to the release of NATCDCSP allows an existing customer to continue to use NATCDC with their existing processing, it allows existing NATCDC processes to be re-generated into NATCDCSP processes as the customer sees fit, and it allows the customer to have the flexibility of creating new PLOG processing using either NATCDC or NATCDCSP.

## Statement on NATCDC

The first thing that an existing customer of NATCDC must be made aware of is that NatWorks has no plans for dropping support for NATCDC. The primary reasons for this are as follows:

1. NATCDC is a stable and proven solution, and
2. NatWorks has no intention of forcing a customer to change simply because there is now a new way of doing things.

As noted in the previous section, NATCDCSP is the intended replacement for NATCDC. Since NATCDCSP provides all the functionality found in NATCDC, and given that NATCDCSP has been specifically designed to support the easy migration from NATCDC, it is generally believed that the benefit of adopting a single process over multiple processes will entice customers into adopting NATCDCSP as a matter of course.

However: NatWorks has gone to significant lengths to insure that there is no requirement whatsoever that the installation of a newer version of NatQuery will force a customer to adopt NATCDCSP. NatQuery version 5 and above introduced support for NATCDCSP, while at the same time retaining all support for the use of NATCDC.

Unless a new version of NATCDC is required to correct a reported bug, or a version incompatibility is created with a new release of NATURAL or ADABAS that demands attention, NatWorks is not planning any further development of the product known as NATCDC beyond the currently released version, which is 3.2.1.

Again, existing customers using NATCDC may continue to do so without any worry that support for NatCDC processing will be dropped or any existing functionality of NATCDC will be dropped, and customers may move to NATCDCSP or not at their own discretion independently of the installation of newer versions of NatQuery.

## Introduction of NATCDCSP

To address the one perceived weakness of NATCDC (a single file processed for a single execution of NATCDC), NatWorks modified NATCDC to create a new program called NATCDCSP.

While the introduction of a new program might warrant both a new Product Name and a Release Number starting at 1, due to the fact that NATCDCSP is essentially a major enhancement to NATCDC, NatWorks opted to continue the version numbering that was based on NATCDC releases, but give the new program a different name (NATCDCSP). The first release of NATCDCSP was therefore version 4.0, with the current release of NATCDCSP as of this writing being 4.2.0.

While the basic nature of NATCDC and NATCDCSP are the same (I.E. PLOG Processing), the adoption of a new name primarily addressed how these two processors could co-exist in the same NATURAL library, since both processors are ultimately NATURAL programs.

In terms of changes to processing: Both NATCDC and NATCDCSP read in a parameter file that informs the respective processors of the expected layout of a file. However: NATCDC only dealt with a single File Specification from the Parameter File. This meant that all NATCDC had to do was read the File Specification for a single file into memory and then start reading ADASEL output records looking for a match between a transaction record's FNR and the Parameter File's FNR. Any and all other transactions were then ignored by that execution of NATCDC.

In contrast, NATCDCSP performs processing that is generally referred to as "two-file-match". This means that NATCDCSP will begin its processing by reading the first File Specification into memory from the Parameter File, and it will then read the first record from the ADASEL output. If the FNR's match, then NATCDCSP will continue reading and processing the ADASEL output until a "new" FNR is found on the ADASEL output. When this occurs, NATCDCSP will then read in the File Specification for that file from the Parameter File, and then process the ADASEL output transactions for that specific file. This process continues until all transactions have been read, with all transactions that have FNRs which match the File Specifications in the Parameter File being processed.

It should be noted that NATCDCSP has the intelligent to handle and then report on any disparities it finds between the input Parameter File and the input ADASEL output. This means that NATCDCSP will continue to process even when it finds:

- No PLOG transactions for an given File Specification, or
- No File Specification for an input PLOG transaction

Any and all such situations are reported on by NATCDCSP, along with all reporting that previously occurred against a given file by NATCDC.

## Summary of Differences & Changes

The primary difference between NATCDC and NATCDCSP is that NATCDC was designed to only handle the PLOG transactions against a single ADABAS file in one execution against PLOG data, whereas NATCDCSP is designed to handle the PLOG transactions against as many files as is required in one execution against PLOG data. The main benefit then of using NATCDCSP over NATCDC is that with NATCDCSP there only needs to be one process executed and managed to handle the PLOG data from  $x$  number of files - versus the management of  $x$  number of NATCDC processes that deliver the exact same data.

The following sections detail the difference between NATCDC and NATCDCSP. For more in-depth processing descriptions of both NATCDC and NATCDCSP, please refer to the NATCDC / NATCDCSP Installation and Operation Manual.

### ***Generate NATCDC Objects Function Modified***

To support NATCDCSP, the function previously called **Generate NATCDC Objects** was renamed to be **Generate NATCDC Objects for Single File**.

Additionally, a new option was introduced in this function as seen on the **General Options** tab in the frame entitled **Single Pass Processing**. In this frame there are now two radio buttons: **Stand Alone Process** and **Component of Multi File Process**.

When a NATCDC process is generated, it is generated from the **Generate NATCDC Objects for Single File** function which has the **Single Pass Option** set to **Stand Alone Process**. When this generation occurs: A JCL / Script file is generated, a Parameter File is generated, and a post-NATCDC processing program is generated at a minimum (other files may be generated depending upon the setting of Data Warehouse Integration option as set on the **Output Options** tab of the **Generate NatCDC Objects for Single File** function). At the user's option – NatQuery can then automatically introduce these generated objects into the server environment.

If the **Generate NATCDC Objects for Single File** is used but has the **Single Pass Option** set to **Component of Multi File Process**, then a Parameter file for the specific file is internally generated along with a file-specific NATURAL sub-program, with this generation being linked to a specific **Process Name**. In this case – no movement of generated objects is performed as this action is deferred until the **Generate NATCDCSP Object for Multi File** is used.

### ***Generate NATCDCSP Objects for Multi File Function Added***

To support NATCDCSP, a new function called **Generate NATCDCSP Objects for Multi File** has been added.

This new function allows for the generation of the objects needed for an execution of

NATCDCSP. Specifically, a Parameter File is generated (which contains the file specifications for all files linked to a given **Process Name**), a “Traffic” Program is generated, and the JCL to execute the NATCDCSP process is executed. At the user’s option – NatQuery can then automatically introduce these generated objects, along with the previously generated NATURAL sub-programs, into the server environment.

### ***Different JCL Templates***

The JCL / Script processing is physically different between NATCDC and NATCDCSP.

### **Review of NATCDC Processing**

The following processing steps occur with a NATCDC process:

1. NATCDC is run against the output of an ADABAS PLOG utility such as ADASEL (mainframes) or ADAPLP (UNIX and Windows).

Beyond inputting the pre-processed PLOG data, NATCDC also inputs a Parameter File generated by NatQuery, with this Parameter File being specific to a single ADABAS file.

NATCDC outputs several reports, but its primary output is the PLOG transactions that pertain to the file specified in the Parameter File, which are written out as fixed-length records (but written as VB).

2. The output of NATCDC is then sorted using a System Sort, such that the records are put into ascending sequence based on ISN and transaction Date Time.

These records are read as VB, and are written as VB.

3. The output of the System Sort is then processed by a NatQuery-generated NATURAL program that is specific to the file being processed. This generated program then acts on the incoming records, usually by distilling these records down into “Delta” images, as well as performing any needed data conversions.

This data is then ready to be used as needed.

### **Overview of NATCDCSP Processing**

The following processing steps occur in NATCDCSP processing:

1. A System Sort is run against the output of an ADABAS PLOG utility such as ADASEL (mainframes) or ADAPLP (UNIX and Windows).

This Sort outputs a file that is sorted into ascending sequence by DBID, FNR, ISN, and relative transaction number.

The records are read as VB, and are output as VB.

2. NATCDCSP is then run against the output of the System Sort. Beyond inputting the sorted data, NATCDCSP also inputs a Parameter File generated by NatQuery, with this Parameter File containing the file specifications for all files that are requested to be processed.

NATCDC outputs several reports, but its primary output is the PLOG transactions that pertain to the file specified in the Parameter File, which are written out as fixed-length records (but written as VB).

When writing its data out, NATCDCSP writes “File Header” and “File Trailer” records to the output file in addition to data.

3. The output of NATCDCSP is then processed by a collection of generate NATURAL programs: One NATURAL “Traffic” program and n number of File-Specific NATURAL sub-programs.

Processing begin with an execution of the “Traffic” program, which reads the first input record – which will be a Header Record for a specific file. The Traffic Program then CALLNATs a generated sub-program that is specific to processing the given file. This sub-program reads and processes records for the given file until it reaches the file’s Trailer record, at which point control is returned to the Traffic Program. The Traffic program then reads the next input record, which will be a Header record for the next file (at which point it will CALLNAT the appropriate sub-program) or an EOF condition (at which point processing ends).

ALL DATA FROM A NATCDCSP PPROCESS IS WRITTEN TO THE SAME PHYSICAL FILE; however each record is preceded by a standard header which identifies the File involved as well as the given ISN.

### ***Different Processors***

In a NATCDC process, a NATURAL module provided by NatQuery and named NATCDC is executed.

With a NATCDCSP process, a NATURAL module provided by NatQuery and named NATCDCSP is executed.

### ***Different Parameter Handling***

With a NATCDC Process, a single Parameter file is generated from the use of the **Generate NATCDC Objects for Single File**, and this Parameter file only contains information on a single file of interest.

With a NATCDCSP Process, a single Parameter file is generated from the use of the

**Generate NATCDCSP Objects for Multi File;** however this Parameter file contains information on one or more files (depending upon how many files were linked to a given **Process Name**).

### ***Different Output***

With a NATCDC process, the output is a fixed length file that only contains the transactions against the single file of interest. This file is created with Fixed-Length records and is written in FB format.

With a NATCDCSP process, the output is a Variable Length file that contains all transactions against the requested files. This output is presented in ascending sequence by ISN within File Number.

To facilitate the “splitting out” of this single file on likely target platforms where the data is likely to be used, a NatQuery installation provides two splitter programs, one specific to UNIX and the other specific to Windows that will take the single output file and create x number of individual output files on the target platform.

## Continuing Use of NATCDC

If a customer intends to continue to use NATCDC, then the generation of a NATCDC process remains essentially the same as with previous versions.

Specifically, the function **Generate NatCDC Objects for Single File** is used (prior to NatQuery version 5 this function was known as **Generate NATCDC Objects**) to generate the required processing. When this function is used, and to insure that a NATCDC process is generated as opposed to a NATCDCSP process, the user will select the **Stand Alone Process** option as opposed to the **Component of a Multi-File Option**; these options are found in a frame called **Single Pass Processing** located on the **General Options** tab of the **Generate NATCDC Objects for Single File** function.

As described above, existing NATCDC processes can be re-generated as NATCDC processes, or new files can have NATCDC PLOG processing generated for them.

## Migrating from NATCDC to NATCDCSP

Migrating from NATCDC to NATCDCSP is a simple task. The pre-requisites to performing this migration would be:

1. NatQuery 5.x or high must be installed,
2. The NATCDCSP module must be installed into NATURAL, and
3. The JCL template name **Production NATCDC SP Process** must be created.

With the above pre-requisites met (please refer to the NatCDC / NATCDCSP Installation Manual for detailed instructions), the process of taking an existing NATCDC process and converting it into a NATCDCSP process is as follows:

### 1. **Generate NATCDC Objects for Single File function**

The **Generate NATCDC Objects for Single File** function is used for each file which needs to be included in a NATCDCSP process.

This function is found in an Administrator version of NatQuery by clicking on the **Administer** drop-down menu and then moving the mouse to point at the **NATCDC / NATCDCSP** menu item.

In versions of NatQuery prior to version 5, this function was called **Generate NatCDC Objects**, and it was found under the menu item name of **NATCDC**.

When using the **Generate NATCDC Objects for Single File** function, there has been a new frame added with the release of NatQuery 5 to the tab named **General Options**, with this frame labeled **Single Pass Processing**.

The **Single Pass Processing** frame contains two radio buttons, one called **Stand Alone Process**, and the other called **Component of Multi-File Process**.

If the **Stand Alone Process** is selected and generation is then invoked, then the output of the Generate NATCDC Objects for single file will be a NATCDC process and all objects necessary to run this process will be generated (JCL, NATCDC Parameter File and the post-NATCDC Natural processing program). This is essentially the same generational processing as before for NATCDC.

If on the other hand the **Component of Multi-File Process** is selected and generation is invoked, then the output will be used for a NATCDCSP process and only the Parameter file and a file-specific NATURAL sub-program will be generated. Further: When the user selects the **Component of Multi-File Process** option, the user will be prompted to provide (or select) a **Process Name** that will be associated with the new NATCDCSP process. The generation of remaining objects necessary for a full NATCDCSP execution is then deferred

until the new function called **Generate NATCDCSP Objects for Multi-File** is used.

## 2. **Generate NATCDCSP Objects for Multi-File function**

When all currently-defined NATCDC process specifications are designated as being a **Component of a Multi-File Process**, and all have been linked to a specific **Process Name**, then the user can use the new function called **Generate NATCDCSP Objects for Multi-File**. This function is found in an Administrator version of NatQuery by clicking on the **Administer** drop-down menu, moving the mouse over the **NATCDC / NATCDCSP** menu item, and then clicking on **Generate NATCDC Objects for Multi File**.

The **Generate NATCDCSP Objects for Multi-File** then displays to the user the names of all Process Names that were named during the processing of the **Generate NATCDC Objects for Single File** function (usually just one).

This allows a user to indicate which of the individual file specifications they wish to include in the NATCDCSP process, and when generation is invoked the result is the JCL necessary to run a NATCDCSP process, a single Parameter file (containing all File Specifications linked to the Process Name), and a generated NATURAL “traffic program” that will serve to call the file-specific NATURAL sub-programs generated through the use of **Generate NATCDC Objects for Single File**.

When generation is completed, the NATCDCSP process will request permission to move the JCL, Parameter File, “Traffic Program” and all file-specific sub-programs onto the server via FTP for execution.

3. After the NATCDCSP process is placed on the server, and for the initial execution, it is expected that JCL problems may exist with the JCL template such that manual “tweaks” will need to be made to the **Production NATCDC SP Process** template so that it will properly execute.

## Creating New Processes Using NATCDCSP

To create a new NATCDCSP process, the procedure is essentially the same as described in the previous section.

Essentially: The **Generate NATCDC Objects for Single File** function is used to provide the basic file processing specifications, with the **Single Pass Option** being set to **Component of Multi File Process** and then linking that process to a given **Process Name**.

Once all file specifications for the Multi File process have been defined through the **Generate NATCDC Objects for Single File**, the **Generate NatCDCSP Objects for Multi File** is used.