

SAS-Programme für Anwender verfügbar machen mit Stored Processes

Andreas Bachert
HMS Analytical Software
Rohrbacher Str. 26
69115 Heidelberg
andreas.bachert@analytical-software.de

Zusammenfassung

Stored Processes, die Teil der Funktionalität des SAS 9 BI-Servers sind, machen es einfach, SAS-Programme für Anwender verfügbar zu machen, die keinerlei SAS-Kenntnisse haben müssen. Dadurch kann der Programmierer das Aufrufen und das Parametrieren von SAS-Programmen sowie die Weiterverarbeitung der Ergebnisse in die Verantwortung des Anwenders übergeben.

Zunächst wird eine Einführung zu Stored Processes und der notwendigen Architektur-Umgebung gegeben.

Danach werden die wichtigsten Techniken für das Erstellen von Stored Processes erklärt und demonstriert. Der Programmierer kann Stored Processes komfortabel erstellen mit SAS Enterprise Guide oder alternativ mit Hilfe der SAS Management Console.

Stored Processes können durch die Endanwender über eine Web-Oberfläche, aus Microsoft Office heraus oder wieder aus SAS Enterprise Guide heraus aufgerufen werden. Thematisiert werden dabei auch die Steuerung von Stored Processes über Parameter (Eingabefelder, Auswahllisten etc.) und die verschiedenen Ergebnisformate (unter anderem HTML, PDF für Adobe Acrobat, RTF für Microsoft Word).

Der Umstieg von bestehenden SAS-Programmen und –Anwendungen auf Stored Processes ist das nächste Thema. Was muss man beachten und anpassen, wenn ein bestehendes SAS-Programm als Stored Process eingesetzt werden soll?

Im letzten Teil wird das Session Konzept erläutert, das es ermöglicht, Informationen über Anfragen hinweg im Zugriff halten zu können.

Schlüsselwörter: BI-Plattform, Intelligence Plattform, Stored Process, Add-In for Microsoft Office, Enterprise Guide, Management Console, Session

1 Einführung in Stored Processes

Stored Processes sind einfach nur SAS Programme, die bestimmte Eigenschaften haben bzw. für die bestimmte Aussagen zutreffen.

Sie ...

- ... sind an zentraler Stelle registriert
- ... werden von einem zu bestimmenden Server ausgeführt
- ... können über unterschiedliche Clients aufgerufen werden
- ... können von Benutzern aufgerufen werden, die keine Programmierkenntnisse benötigen

- ... können mit Ausführungsberechtigungen versehen werden
- ... können beim Aufruf mit bestimmten Parametern versorgt werden, wobei die Clients dabei dynamisch eine Eingabeoberfläche generieren
- ... können Ergebnisse direkt an den Client zurück senden

1.1 Die SAS Intelligence Plattform

Das Stored Process Konzept ist Teil der SAS Intelligence Plattform, einer Mehrschichten-Architektur, die von SAS mit der Version 9 eingeführt worden ist.

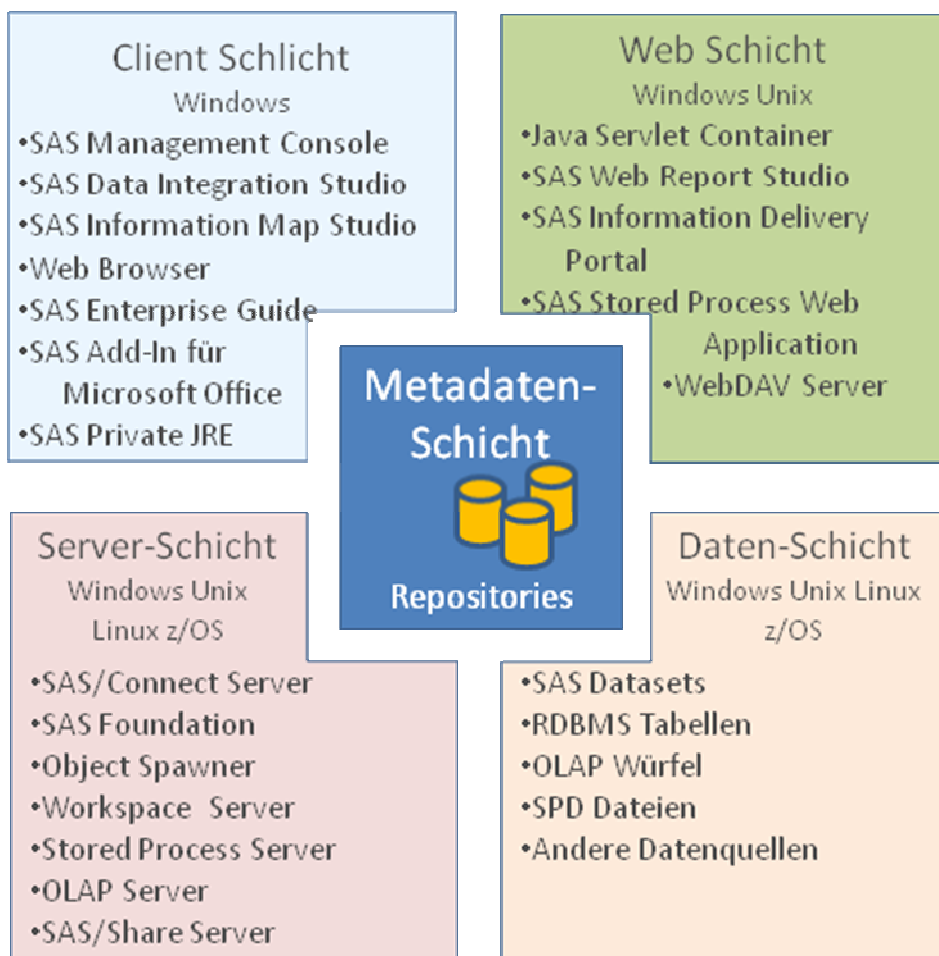


Abbildung 1: Die Schichten der SAS Intelligence Plattform

Es lassen sich 5 unterschiedliche Schichten identifizieren:

- Metadaten-Schicht
- Client Schicht
- Web Schicht
- Server Schicht
- Daten-Schicht

Das zentrale Herzstück der SAS Intelligence Plattform ist der Metadaten-server, der die Metadaten-Schicht repräsentiert. In sogenannten Repositories werden beschreibende Informationen zu folgenden Objekttypen zentral gespeichert:

- Benutzer und Gruppen
- Server
- Daten-Bibliotheken
- Datasets bzw. Tabellen in diesen Bibliotheken
- Ordnerstrukturen für Programm- und Ergebnisablage
- Stored Processes

Administratoren dürfen diese Einträge verwalten, wobei insbesondere auch Zugriffsberechtigungen geregelt werden können. Als Werkzeug wird dabei die SAS Management Console eingesetzt.

Zu beachten ist, dass es sich nur um beschreibende Informationen handelt. So wird für Datasets nur registriert, über welche Bibliothek zugegriffen wird und welche Spalten mit welchen Eigenschaften darin enthalten sind. Die Daten selbst sind allein im Dataset selbst gespeichert.

Metadaten, die für einen Stored Process gespeichert werden, sind im Wesentlichen

- Beschreibende Daten
- Ablageort des Stored Process Programms
- Ausführender SP-Server
- Eingabeparameter
- Ausgabeoptionen (z.B. Kein Output, Streaming Output oder Package)
- Berechtigungen

Als ausführender Server kommen zwei unterschiedliche Serverarten in Frage, nämlich ein SAS Workspace Server oder ein SAS Stored Process Server.

1.2 Der SAS Workspace Server

Der SAS Workspace Server ist ein Einbenutzer-Server. D.h., dass immer, wenn ein Benutzer eine Verbindung zu einem Workspace Server öffnet, ein neuer SAS Prozess auf dem entsprechenden physischen Computer gestartet wird.

Dieser Prozess läuft so lange, bis der Benutzer die Verbindung wieder schließt. So können Folgeanfragen jederzeit auf WORK Datasets oder Makrovariablen zugreifen, die bei einer früheren Anfrage generiert bzw. initialisiert wurden.

Der Prozess wird unter dem Account des Benutzers gestartet, der die Verbindung öffnet. Das bedeutet, dass der Prozess auch die Rechte hat, über die der betreffende Benutzer im Server-Betriebssystem verfügt. Dies gilt insbesondere für Zugriffsrechte auf Datasets im Dateisystem.

Der Enterprise Guide arbeitet ausschließlich mit Workspace Servern zusammen.

1.3 Der SAS Stored Process Server

Beim SAS Stored Process Server handelt es sich um einen Mehrbenutzer-Server, also um einen Prozess, den sich mehrere Clients teilen. Dadurch wird Zeit gespart, die sonst bei jeder Anfrage für das Starten einer SAS Sitzung anfallen würde. Da aus Betriebs-

systemsicht ein Prozess nur unter dem Account genau eines Benutzers laufen kann, folgt daraus, dass der Stored Process Server üblicherweise unter dem Account eines technischen Benutzers läuft. Diesem technischen Benutzer werden dann alle Rechte im Betriebssystem eingeräumt, die für die geplanten Stored Processes benötigt werden.

In der Praxis wird das Stored Process Szenario häufig so aufgebaut, dass in Abhängigkeit der zu erwartenden Benutzerzahl mehrere SAS Prozesse als Stored Processes Server in den Metadaten registriert werden. Dabei wird zusätzlich ein Load Balancer eingesetzt, der die Anfragen entgegennimmt und an den nächsten freien Server Prozess weiterleitet.

Stored Process Server sind also skalierbar und dadurch hoch performant.

Abschließend folgt eine Aufzählung weiterer spezieller Merkmale eines Stored Process Servers.

Im Gegensatz zum Workspace Server unterstützt der SAS Stored Process Server ...

- ... Streaming Output
- ... Sessions
- ... *multi-value* Parameter, d.h. Mehrfachselektion aus Auswahllisten

1.4 Das Ablaufprinzip bei Stored Processes

Stored Processes laufen in folgenden Schritten ab:

- Der Client fragt den Metadatenserver nach einem bestimmten Stored Process (über Metadatenpfad und Name).
- Der Client erhält Metadaten, insbesondere auf welchem Server der Stored Process laufen soll und welche Parameter abgefragt werden müssen.
- Sofern Parameter erforderlich sind generiert der Client eine Eingabeoberfläche, in der der Benutzer die notwendigen Einstellungen vornehmen kann.
- Beim Start des Stored Process übergibt der Client die Informationen über Parameterwerte und das auszuführende SAS Programm über einen Spawner an einen freien Stored Process Server. Dieser initialisiert globale Makrovariablen entsprechend der übergebenen Parameterwerte und startet dann das gewünschte Programm.
- Der Stored Process Server gibt die Ergebnisse an den Client zurück. Dies kann als Stream oder als Package erfolgen.
- Der Client zeigt die Ergebnisse in geeigneter Form an.

2 Erstellen von Stored Processes

Wie bereits erwähnt sind Stored Processes zunächst nichts anderes als SAS Programme. Man muss also ein SAS schreiben oder sich dieses ggf. vom Enterprise Guide erzeugen lassen. Üblicherweise handelt es sich bei SPs um Makroprogramme mit Parametern.

Danach muss der Stored Process im Metadatenserver registriert werden. Dies kann über die Management Console oder mit Hilfe des Enterprise Guide geschehen.

2.1 Erstellen von Stored Processes mit dem Enterprise Guide

Der Enterprise Guide ist das zu bevorzugende Werkzeug für diesen Arbeitsschritt. Der Benutzer markiert einfach den Programmknoten bzw. die Anwendungsroutine, der/die zu einem Stored Process konvertiert werden soll und wählt dann über das Kontextmenü die Funktion *Stored Process erstellen...*. Danach wird er von einem Assistenten durch die einzelnen Schritte geleitet.

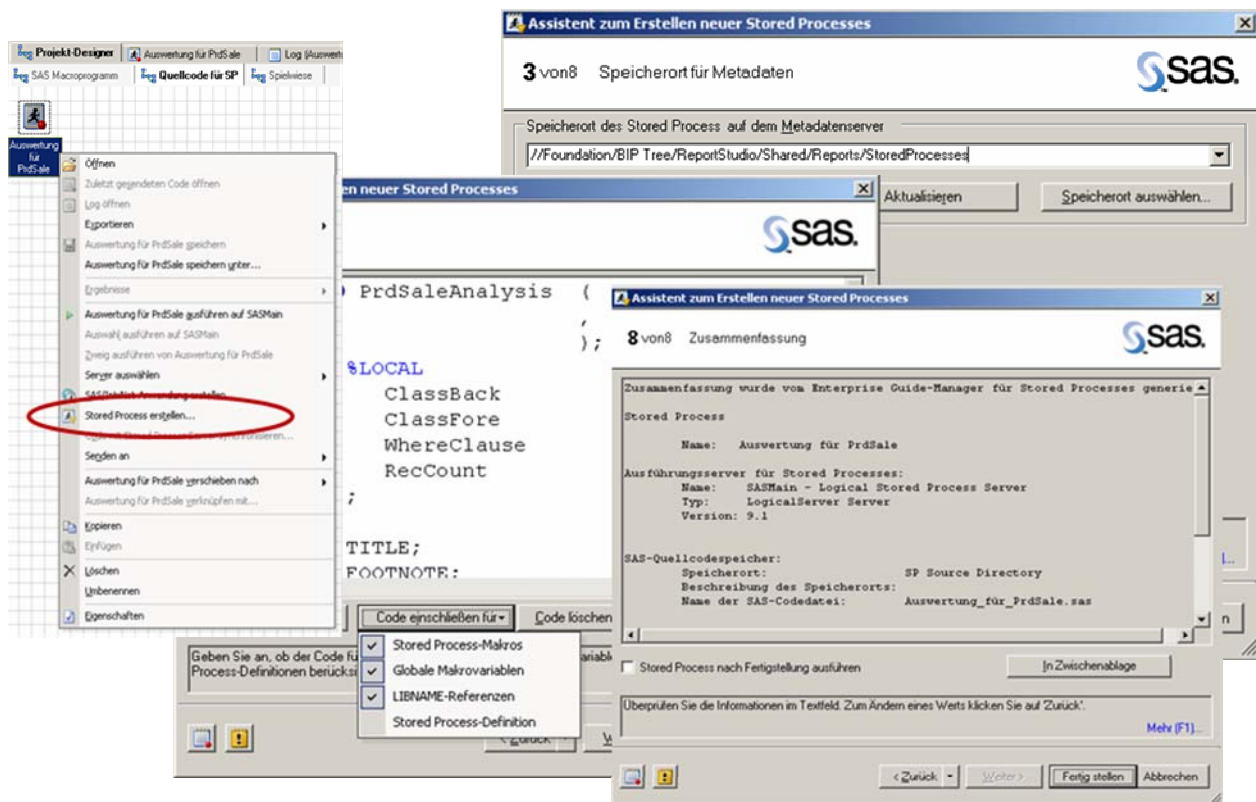


Abbildung 2: Erstellung eines Stored Processes mit dem Enterprise Guide

2.2 Testen von Stored Processes mit dem Enterprise Guide

Nach erfolgreicher Registrierung erscheint im Enterprise Guide Projekt ein Stored Process Symbol, das unlösbar mit dem Ausgangsknoten verknüpft ist. Im Kontextmenü dieses Symbols kann der Stored Process dann ganz einfach durch Aufruf des entsprechenden Menüpunktes *<sp-name> ausführen* getestet werden.

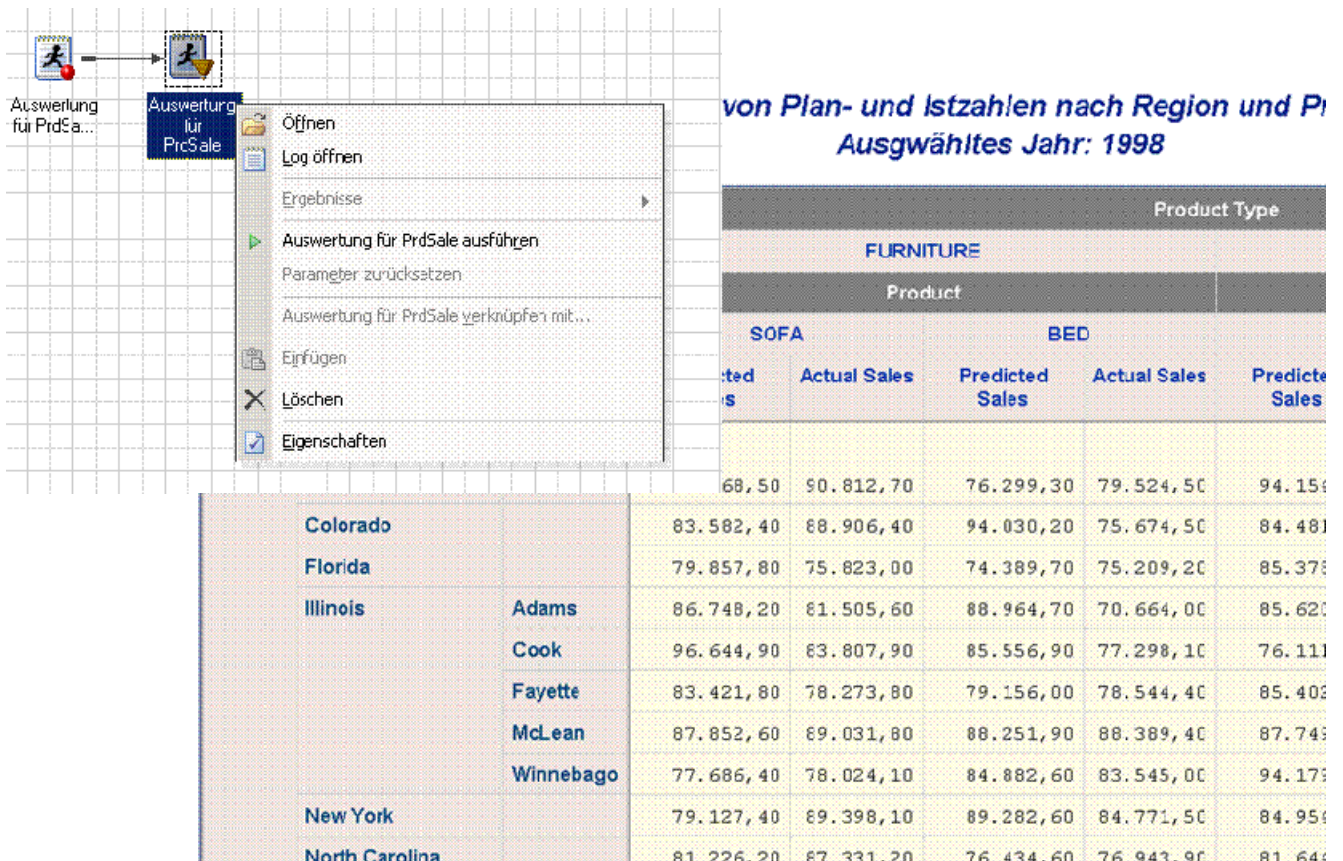


Abbildung 3: Testen eines Stored Processes mit dem Enterprise Guide

2.3 Registrieren von Parametern für Stored Processes

Als mögliche Parameter für einen Stored Process stehen dem Benutzer eine Reihe unterschiedlicher Typen zur Verfügung:

- Integer
- Numerisch
- Alphanumerisch
- Datum
- Dateibezeichner

Für jeden Parametertyp kann man spezifische Einschränkungen, sogenannte Constraints definieren. Bei Integer-Parametern kann z.B. ein erlaubter Bereich festgelegt werden mit *Von*, *Bis* und einem *Inkrement*.

Für alphanumerische Parameter lassen sich Auswahllisten definieren. Die Daten können aus Datasets geladen werden oder direkt am Bildschirm eingegeben werden. Bei Auswahllisten lässt sich festlegen, ob der Benutzer auch einen Wert eingeben kann, der nicht in der Liste enthalten ist oder ob mehr als eine Auswahl erlaubt ist.

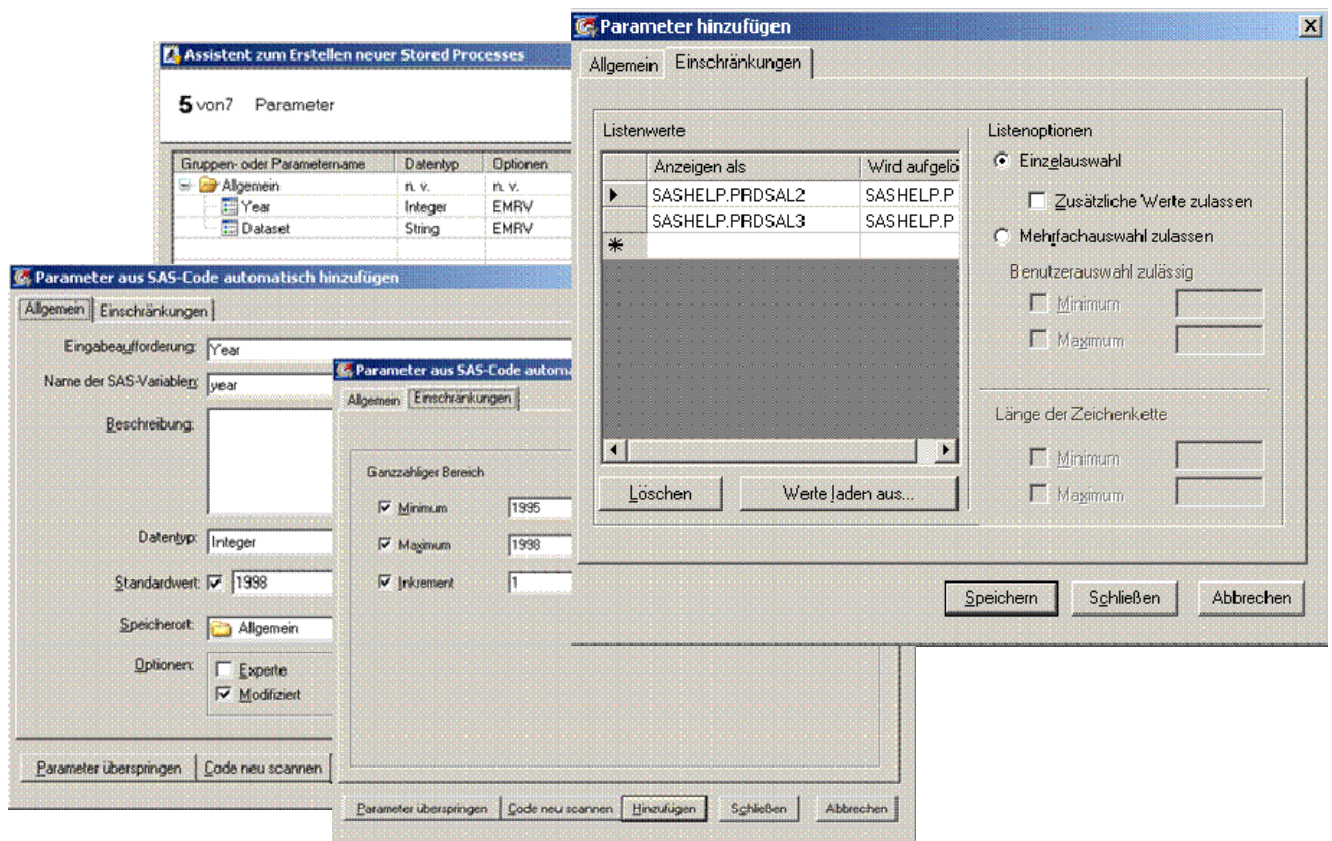


Abbildung 4: Definition von Parametern für Stored Processes

In jedem Fall muss man beachten, dass die Einschränkungen für die Parameter und somit auch die einzelnen Optionen in Auswahllisten als fixer Teil der Metadaten eines Stored Processes im Metadatenserver gespeichert werden. D.h. dass sich spätere Änderungen an den Ausgangsdaten, z.B. wenn ein neues Jahr hinzugekommen ist, nicht in den Metadaten widerspiegeln. Vielmehr muss man dann die Definition des Stored Process entsprechend überarbeiten und der Auswahlliste für das Berichtsjahr die neu zur Verfügung stehende Ausprägung hinzufügen.

Parameter lassen sich auch zu Parametergruppen zusammenfassen. Auch ist es möglich, Parameter als unsichtbar oder als erforderlich zu definieren. Bei erforderlichen Parametern ist es später nicht möglich, den Stored Process zu starten, ohne zuvor explizit einen Wert festgelegt zu haben. Jeder Parameter kann auch mit einem Standardwert versehen werden.

2.4 Code-Modifikation an bestehenden Stored Processes

Mit dem Enterprise Guide ist es sehr einfach möglich, Änderungen am Programm vorzunehmen, das als Basis für die Stored Process Erstellung diente.

Im Kontextmenü des Programmknotens wird dazu die Funktion *Code mit Stored Process Server synchronisieren...* aufgerufen.

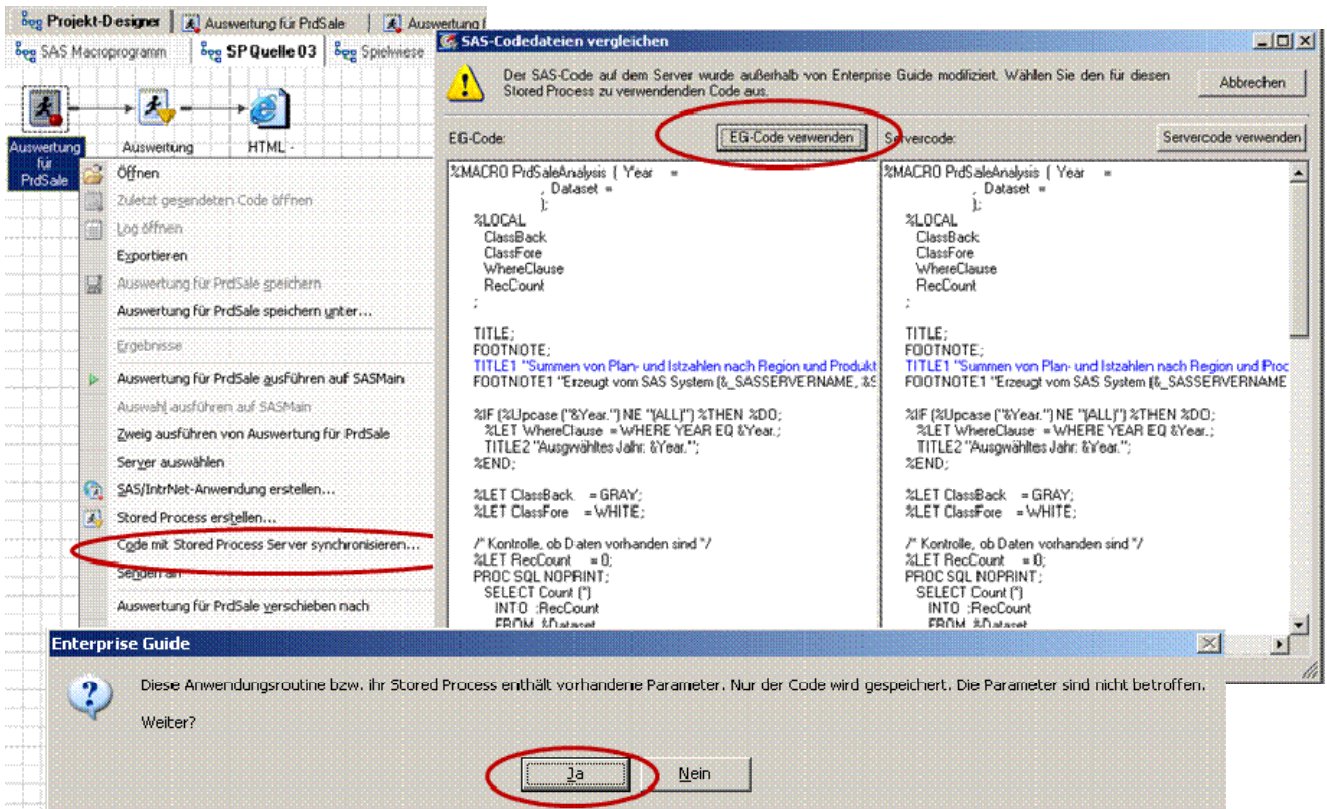


Abbildung 5: Synchronisation nach Änderung im Programm-Code eines Stored Process

Sofern Parameter definiert sind, wird der Benutzer darauf hingewiesen, dass nur der Code synchronisiert wird. Sollte man z.B. neue Parameter für das Programm hinzugefügt haben, können diese nicht automatisch zu Parametern des Stored Process werden. Im letzten Fenster dieser Funktion sieht man eine Gegenüberstellung des Programmcodes, wie er aktuell im Enterprise Guide Projekt vorhanden ist und dem aktuellen Programmstand auf dem Stored Process Server. Üblicherweise entscheidet man sich dann für *EG Code verwenden*. Dadurch wird der Enterprise Guide-Programminhalt an den Speicherort im Stored Process Server übertragen.

2.5 Gegenüberstellung zwischen Enterprise Guide und Management Console

Wie bereits erwähnt ist der Enterprise Guide das Mittel der Wahl, wenn es um die Registrierung von Stored Processes geht.

Es folgt eine Bewertung der Möglichkeiten, die der Enterprise Guide in diesem Kontext bietet:

- + Sehr Komfortabel, ausgehend von Code im SEG Projekt
- + Weitergehende Möglichkeiten bei Parameterdefinition, z.B. Trennung von Anzeige- und Schlüsselwert bei Auswahllisten
- + Stored Process-spezifische Codefragmente werden automatisch eingetragen
- + Code wird automatisch an den Speicherort im SP Server transferiert

- + Einfache Synchronisation nach Programm-Änderungen
- Speicherorte für Metadaten und Programmdateien können nur gewählt, aber nicht neu definiert werden

Demgegenüber bietet die Management Console nur wenig Vorteile, wie die folgende Bewertung zeigt:

- + Neue Speicherorte für Metadaten und Programmdateien können definiert werden
- + Möglichkeit der Vergabe von Zugriffsberechtigungen
- Weniger Möglichkeiten bei Parameterdefinition
- Parameter bzw. –Gruppen können nicht mehr nachträglich verschoben werden
- Programmdatei muss sich bereits am Speicherort im Zugriff des SP Servers befinden
- Für Änderungen muss die Programmdatei überschrieben werden

3 Ausführen von Stored Processes

Nachdem das Umfeld und der Registrierungsprozess für Stored Processes beschrieben wurden, werden in diesem Abschnitt die unterschiedlichen Clients gegenübergestellt, mit denen man Zugriff auf Stored Processes hat und diese ausführen lassen kann.

In der SAS Intelligence Plattform sind die folgenden Programme und Werkzeuge enthalten, mit denen Stored Processes ausgeführt werden können:

- SAS Enterprise Guide (Windows-Client)
- SAS Add-In für Microsoft Office (Windows-Client)
- SAS Web Report Studio (Web Applikation via Web Browser)
- SAS Information Delivery Portal (Web Applikation via Web Browser)
- SAS Stored Process Web Application (Web Applikation via Web Browser)

Tabelle 1: Gegenüberstellung der SAS Clients zur Ausführung von Stored Processes

	Ausgabe-Format	Ausgabe-Style	Parameter-Seite	Speichertort für SP Metadaten
Enterprise Guide	<ul style="list-style-type: none"> • HTML • RTF • PDF • SAS Report 	<ul style="list-style-type: none"> • Frei Wählbar aus SAS Styles • Erweiterbar um eigene Styles 	<ul style="list-style-type: none"> • .NET Controls • Parametergruppen als Reiter 	<ul style="list-style-type: none"> • Kann beliebig sein • Keine Suchmöglichkeit • Manuelle Navigation im Baum des BI Managers

	Ausgabe-Format	Ausgabe-Style	Parameter-Seite	Speichertort für SP Metadaten
Add-In für MS Office	<ul style="list-style-type: none"> • HTML • SAS Report • CSV 	<ul style="list-style-type: none"> • Frei Wählbar aus SAS Styles • Erweiterbar um eigene Styles 	<ul style="list-style-type: none"> • .NET Controls • Parametergruppen als Reiter 	<ul style="list-style-type: none"> • Kann beliebig sein • Über Suche ist jeder Stored Process auffindbar
Web Report Studio	<ul style="list-style-type: none"> • SAS Report 	<ul style="list-style-type: none"> • Wählbar • Nur wenige Styles stehen zur Auswahl 	<ul style="list-style-type: none"> • HTML-Seite • Teilw. Listboxen statt Comboboxen • Keine Parametergruppen 	<ul style="list-style-type: none"> • Unterhalb von BIP-Tree/Reports bzw. BIP-Tree/Users/<identity>Reports
Information Delivery Portal	<ul style="list-style-type: none"> • HTML • RTF • PDF 	<ul style="list-style-type: none"> • Frei Wählbar aus SAS Styles 	<ul style="list-style-type: none"> • HTML-Seite • Parametergruppen via Navigation 	<ul style="list-style-type: none"> • Kann beliebig sein • Über Suche ist jeder Stored Process auffindbar
Stored Process Web Application	<ul style="list-style-type: none"> • HTML • RTF • PDF 	<ul style="list-style-type: none"> • Frei Wählbar aus SAS Styles 	<ul style="list-style-type: none"> • HTML-Seite • Parametergruppen via Navigation 	<ul style="list-style-type: none"> • Kann beliebig sein • Keine Suchmöglichkeit • Manuelle Navigation im Baum des BI Managers

In allen genannten Tools kann der Benutzer das Ausgabeformat und den Style anpassen. Teilweise geht das nur global für alle bevorstehenden Ausgaben, teilweise kann man auch nur die Ausgabe für genau einen aktuellen Stored Process variieren. Weitere Unterschiede finden sich bei der Abfrage von Parametern. Die Eingabeseiten im Web-Umfeld sehen naturgemäß anders aus als die Oberflächen der Windows-Clients.

4 Umwandeln von bestehenden SAS Programmen zu Stored Processes

Wie eingangs erwähnt, sind Stored Processes nichts weiter als SAS Programme. Es bedarf nur einiger Modifikationen um aus einem bestehenden SAS Programm einen Stored Process zu erstellen.

Die Struktur eines Stored Process sieht wie folgt aus:

- ***ProcessBody;**
Startanweisung; Trigger für Stored Process
- **%GLOBAL**
Alle Parameter zu globalen Macrovariablen machen
- **%STPBEGIN**
Initialisierungsmakro, insbesondere mit ODS-Anweisungen
- Danach folgt der eigentliche, vom Entwickler bereitgestellte SAS Programmcode
- **%STPEND**
Abschlussmakro, insbesondere mit ODS-Anweisungen

Beim Programmcode selbst sollten wenigstens die folgenden beiden Punkte beachtet werden:

- Keine eigenen ODS Statements im Code, damit der Benutzer das Ausgabeformat im Client wählen kann. Würde im Programm z.B. ein Statement für ODS HTML mit Style=DEFAULT enthalten sein, würde man damit die Einstellung eines Benutzers, wonach z.B. PDF mit Style=BRICK erzeugt werden soll, überschreiben.
- Fehlende Libname Statements ggf. eintragen, wenn nicht z.B. durch autoexec auf Server bereits gesetzt. Es gilt grundsätzlich zu beachten, dass auf dem Stored Process Server das gleiche Umfeld wie beim bisherigen Programmablauf existieren muss. D.h., dass alle Libnames, die im Programm verwendet werden, gesetzt sein müssen und dass der technische Benutzer auch die Zugriffsberechtigung auf die Dateien darin haben muss.

Am einfachsten führt man die Umwandlung mit dem SAS Enterprise Guide durch, da dieser die wesentlichen Punkte (Eintragen von %STPBEGIN und %STPEND, %GLOBAL für Parameter) bereits umsetzt.

5 Stored Processes und Sessions

Normalerweise ist das Web (Internet, Intranet) "stateless". D.h. dass in einer 2. Anfrage keinerlei Information mehr aus einer 1. Anfrage zur Verfügung steht.

Durch das Session-Konzept wird ein mehrstufiges Verfahren ermöglicht. Zu beachten ist, dass Sessions nur vom Stored Process Server unterstützt werden.

Es gibt spezielle SAS-Funktionen, um mit Sessions zu arbeiten:

- `stpsrv_session` (create | delete)
- `stpsrvset` ('session timeout' | 'request timeout', <seconds>)
- ...

Beim Anlegen einer Session werden automatisch die folgenden globalen Makrovariablen initialisiert:

- `_SESSIONID`
Enthält die ID einer zuvor mit `stpsrv_session` (create) erzeugten Session
- `_THISSESSION`
Enthält die URL zum Aufruf der Stored Process Web Application samt SessionID, um eine Anfrage an eine zuvor erzeugte Session zu leiten.

Sofern man nicht über ein eigenes Framework zum Aufrufen von Stored Process–Aufrufen verfügt, dient als Ausführungsumgebung für Stored Processes die SAS Stored Process Web Application.

Neben den genannten Makrovariablen wird beim Anlegen einer Session zusätzlich automatisch eine SAS Library namens `SAVE` allokiert.

Alle Datasets und Catalogs, die man in der Library `SAVE` gespeichert hat, bleiben für spätere Zugriffe erhalten. Darüber hinaus bleiben alle Makrovariablen erhalten, deren Namen mit dem Präfix `SAVE_` beginnen.

Die Voraussetzung dafür, dass die permanent gespeicherten Informationen aus Library `SAVE` tatsächlich bei Folgeanfragen derselben Session zur Verfügung stehen, ist, dass alle Anfragen einer Session an denselben Stored Process Server weitergeleitet werden. Dadurch wird zwangsläufig das eingangs angesprochene Load Balancing *ausgehebelt*, was evtl. zu Wartezeiten führen kann.

Aus diesem Grunde sollten Sessions nur dann verwendet werden, wenn sehr viele Informationen bzw. große Dateien über mehrere Anfragen hinweg im Zugriff sein sollen, wenn also die Informationsübergabe in Form von URL-Parametern nicht mehr möglich ist.

6 Ausblick

Stored Processes sind aus Sicht des Autors eine herausragende Möglichkeit, die Vorteile der SAS Intelligence Plattform auszunutzen. Die Programme werden von Programmierern entwickelt und können den Anwendern ohne weiteren nennenswerten Aufwand zur Verfügung gestellt werden.

Es entsteht insbesondere kein Aufwand mehr für Installation und Verteilung geänderter Programm-Versionen, da immer genau die zuletzt freigegebene Version über den Metadaten-server zur Verfügung gestellt wird.

Ein weiterer Aspekt, der für den Einsatz von Stored Processes spricht, ist dass durch die Client Server Architektur keine lokalen Lizenzkosten mehr bei den Endanwendern anfallen.

Besonderen Charme hat sicherlich auch die Tatsache, dass mit unterschiedlichen Clients zugegriffen werden kann, die letztlich auch für unterschiedliche Benutzergruppen konzipiert sind.

In der bald zur Verfügung stehenden SAS Version 9.2 werden Neuerungen enthalten sein, die die Nachteile des aktuell mit SAS 9.1.3 zur Verfügung stehenden Frameworks aufheben werden. Parameter werden dann selbst Metadatenobjekte sein und stehen somit zur Wiederverwendung in unterschiedlichen Stored Processes zur Verfügung.

Auch wird es ebenso dynamische wie kaskadierende Eingabeaufforderungen (Prompts) geben. Man spricht von kaskadierenden Prompts, wenn die zur Verfügung stehenden Optionen eines Parameter B von der zuvor getroffenen Auswahl für einen Parameter A abhängen.

Literatur

- [1] http://support.sas.com/rnd/itech/doc9/dev_guide/stprocess/index.html
SAS Stored Process Developer's Guide im Internet, zuletzt besucht am 26. Mai 2009