

Tweaking APEX: A New Approach for Dynamic Interactive Reports

Eckhart Guthöhrlein¹, Bernhard Schirm²
F. Hoffmann-La Roche AG¹, quattro research GmbH²
Basel¹, Martinsried²

Schlüsselworte

APEX API, Interaktive Reports, Dynamisches SQL, MS Excel

Abstract

„Das Data Warehouse ist fertig, nun fehlt nur noch eine gute Reporting-Plattform. Schön, dass es APEX gibt ...Upps, Interactive Reports funktionieren nicht mit dynamischem SQL... Gibt es da keine Lösung?“

Dies ist die wohl schwerwiegendste Einschränkung interaktiver Reports, über die wahrscheinlich jeder APEX-Entwickler irgendwann stolpert. Wie kann man also diese Einschränkung interaktiver Reports umgehen?

Unsere Lösung entstand aus den Anforderungen der Datenanalyse biologischer Fermentationsprozesse bei der Roche Diagnostics GmbH. Bei diesen Fermentationen erzeugen biologische Zellen oder Bakterien potentielle Wirkstoffe, die später als Medikamente z.B. gegen Krebs eingesetzt werden können.

Um den Verlauf und Erfolg einer Fermentation beurteilen zu können, benötigt der Wissenschaftler eine Vielzahl verschiedener Parameter und Messwerte. Diese werden normalerweise als tabellarischer Report zur Verfügung gestellt. Einige Routineaufgaben können dabei durch vordefinierte Reports abgedeckt werden, meistens möchten die Wissenschaftler jedoch ihre Reports je nach Fragestellung dynamisch erstellen.

Um dies zu ermöglichen, wurde ein flexibles PL/SQL Framework für dynamische User-definierte interaktive Reports entwickelt. Die wichtigsten Teile des Frameworks werden innerhalb des Vortrags behandelt, insbesondere die automatische Erzeugung von interaktiven Reports über die APEX API, und über die API hinausgehendes „Hacking“. Ein zusätzliches Augenmerk wird auf die Interaktion mit MS Excel und Performanceoptimierungen gelegt.

Einleitung

Roche ist ein global operierendes Pharmaunternehmen mit mehr als 80000 Mitarbeitern und Weltmarktführer für verschreibungspflichtige Krebsmedikamente. Der bayrische Standort Penzberg beherbergt sowohl Forschungs- und Entwicklungsabteilungen als auch große Produktionsanlagen für therapeutische Proteine. Während es sich bei klassischen Arzneimitteln um kleine Moleküle handelt, die meist mittels chemischer Synthese hergestellt werden können, sind Proteintherapeutika sehr große Makromoleküle, die nur über biotechnologische Fermentationsprozesse zugänglich sind. Hierbei werden zunächst spezielle Zellen gezüchtet, die das gewünschte Protein produzieren. Diese Zellen werden anschließend in große Bioreaktoren transferiert, dort vermehrt und schließlich geerntet.

Die Effizienz der Fermentationsprozesse hängt von einer großen Zahl von Bedingungen ab, beispielsweise Temperatur, Verfügbarkeit diverser Nährstoffe, pH-Wert und vielen anderen mehr.

Eine Erfassung und effiziente Speicherung all dieser Parameter ist nötig, um zum einen laufende Prozesse zu überwachen, zum anderen um retrospektiv mögliche Fehlerquellen zu erkennen und ggf. Optimierungspotential für die Verfahren zu erschließen.

Im Rahmen eines großen Projektes zur Optimierung der Produktionskapazität (siehe <http://www.roche.de/presse/biotechnologie/tp-expand>) sollten Reporting-Werkzeuge zur Überwachung und Auswertung der bei den Fermentationsprozessen anfallenden Daten entwickelt werden.

APEX ist eine bei Roche seit längerem etablierte Plattform für Web- und Reporting-Anwendungen. Daher wurde es als Basis für diese Aufgabe ausgewählt. Im Folgenden werden die Erweiterungen von APEX beschrieben, die notwendig waren, um den besonderen Anforderungen dieses Projekts gerecht zu werden.

Dynamische interaktive Reports

Für die Analyse des Fermentationsverlaufs werden tabellarische Reports herangezogen, die aus hunderten verschiedener Parameter und Messwerte bestehen. Hierfür sind Interaktive Reports prädestiniert. Neben ihren hinreichend bekannten Vorzügen spricht für sie, dass sie sich unter den Anwendern bei Roche einer außerordentlichen Beliebtheit erfreuen.

Allerdings können im Vorfeld nur wenige Standardreports vollständig spezifiziert werden. Im Regelfall muss ein Techniker oder Wissenschaftler eine eigene Auswahl aus allen verfügbaren Werten treffen, die er in einer Übersicht benötigt. Dies bedeutet, dass die dem Report zugrundeliegende Abfrage dynamisch erzeugt werden muss. Mit APEX-Bordmitteln ist dies für interaktive Reports jedoch unmöglich.

Eine recht häufig eingesetzte Möglichkeit, diese Problematik anzugehen, liegt in der Verwendung von APEX Collections. Man baut die Abfrage des Reports auf einer Collection auf, befüllt diese dynamisch und passt die Eigenschaften des Reports wie Anzahl der Spalten und Spaltenüberschriften zur Laufzeit an. Allerdings ist bei Verwendung von Collections die Anzahl der Spalten auf 50 limitiert, eine für die beschriebene Anwendung nicht hinnehmbare Beschränkung.

Die hier vorgestellte Lösung besteht in einem in PL/SQL geschriebenen Reportgenerator. Anstatt einen existierenden interaktiven Report zu verändern, wird für jede neue Fragestellung programmatisch eine neue APEX-Seite mit einem eigenen interaktiven Report erzeugt. Dem Anwender steht ein grafischer Reportdesigner zur Verfügung, mit dem er nach Belieben aus allen verfügbaren Daten Spalten auswählen kann. Anhand dieser Auswahl wird die SQL-Abfrage zusammengesetzt, die dem neuen Report zugrunde liegt.

Architektur

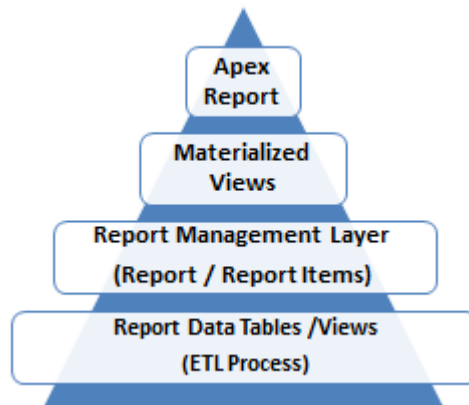


Abb. 1: Architekturübersicht

Abb. 1 zeigt schematisch die Architektur des Reportgenerators.

Als Datenbasis liegt der Anwendung ein Data Warehouse zugrunde, in dem die relevanten Daten aus verschiedenen Quellsystemen integriert werden, vor allem von elektronischen Laborbüchern, Laborinformationssystemen und Messgeräten zur Erfassung von Online- und Offlinedaten. Insbesondere aus der Online-Erfassung von Messwerten laufender Fermentationsprozesse fallen dabei sehr große Datenmengen an (in der Größenordnung 10^9 Datensätze).

Der programmatische Kern der Anwendung liegt in einer Report Management Schicht. Deren Hauptkomponenten sind:

- Zwei Tabellen zum Speichern der Reports (REPORT) und der zugehörigen Spaltenauswahl (REPORT_ITEM).
- Mehrere Tabellen mit den Definitionen für verschiedene Parametergruppen, aus denen der Anwender wählen kann.
- PL/SQL-Packages für das Anlegen der Datenbankobjekte (Views, Materialized Views) basierend auf den Reportdefinitionen, das Erstellen und Aktualisieren der zugehörigen APEX-Seiten sowie die Interaktion mit dem graphischen Report-Designer.

Aus Gründen der Performance wird für jeden Report ein eigener Materialized View angelegt. Auf diesem baut schließlich der eigentliche Interactive Report auf, der in der für den Anwender sichtbaren APEX-Schicht angelegt wird.

Dem Anwender steht für die Definition neuer und Verwaltung existierender Reports ein graphisches Frontend zur Verfügung. Hier können Spalten ausgewählt, benannt, angeordnet und formatiert werden. Jedes Speichern eines neuen oder veränderten Reports löst im Hintergrund einen dreistufigen Prozess aus.

Zunächst wird anhand der Spaltendefinitionen eine SQL-Abfrage generiert, die alle gewünschten Informationen zusammenführt. Diese wird als View in der Datenbank angelegt. Darauf aufbauend wird ein Materialized View erstellt, der anschließend asynchron aktualisiert wird.

Zuletzt wird in der APEX-Zielapplikation eine neue Seite angelegt (bzw. ein existierender Report synchronisiert). Dies geschieht hauptsächlich unter Verwendung der APEX WWV_FLOW_API.

Leider stellt die APEX API (noch) nicht für alle benötigten Anpassungen Funktionen bereit. Um alle Anwenderwünsche abbilden zu können, ist es daher teilweise notwendig, direkt auf APEX-Tabellen zuzugreifen.

Weitere Besonderheiten

Auf eine Reihe weiterer Feinheiten und APEX Kniffe wird im Vortrag näher eingegangen. Beispielsweise ist ein intensiv genutztes Feature der Reports der direkte Download des aktuellen Reports als MS Excel Datei mit Hilfe eines speziell entwickelten APEX Plugins. Anwender-definierte Einstellungen wie Spaltenanordnung und Filter bleiben dabei erhalten.

Kontaktadresse

Dr. Eckhart Guthöhrlein
Roche Diagnostics GmbH
Nonnenwald 2
D-82377 Penzberg
+49 (0) 8856-60 10220
eckhart.guthoehrlein@roche.com

Dr. Bernhard Schirm
quattro research GmbH
Am Klopferspitz 19
D-82152 Martinsried
+49 (0) 89-9901629 0
schirm@quattro-research.com