

SAS 9.4 bei der Fraport AG

Mannheim, 10. September 2013
Stefan Sabatzki



Seite 2



Inhaltsverzeichnis

- Fraport auf einen Blick
- Historie der BI-Umgebung
- Architektur der BI-Umgebung
- SAS 9.4
- Fazit



Seite 3

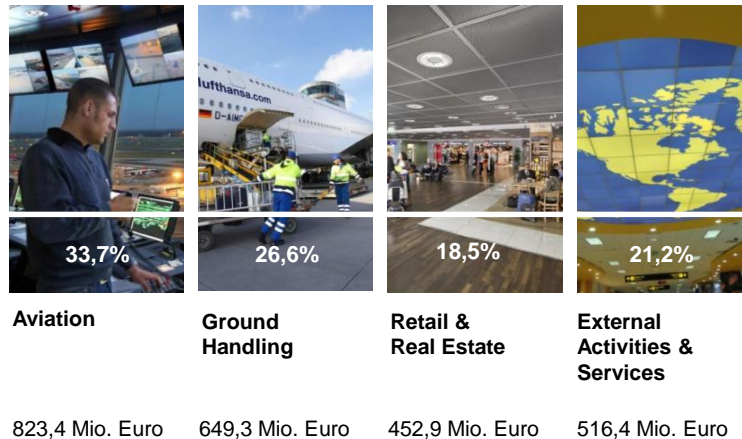


Fraport auf einen Blick

Seite 4



Umsatzträger sind unsere Kerngeschäftsfelder



Umsätze im Gesamtkonzern nach Segmenten 2012

Frankfurt Airport – Zahlen und Fakten

- 21 km² Fläche
- bis zu 94 Flugbewegungen/Std.
- 2 Start- und Landebahnen
- 1 Startbahn, 1 Landebahn
- 2 Fluggast-Terminals und
- 1 AIRail Terminal (ICE-Bahnhof)
- 154 Gates; 204 Positionen
- 295 Ziele in 107 Ländern

An einem Tag*:

- 158.000 Passagiere
- 78.000 Stück Abfluggepäck
- 5.800 Tonnen Cargo
- 397 Züge an den Flughafen-Bahnhöfen (Fahrplan 2010)



*durchschnittliche Angaben aus 2012

Historie der BI-Umgebung

Hintergrund

- Inbetriebnahme des „Business Intelligence Architecture Framework“, der BI-Umgebung für operative Daten im Jahr 2005
- Ziel: Ganzheitliche Sicht auf Betriebszustände für das operative Management mittels
 - historischer Berichte
 - online Reports
- Nutzerkreis erweitert sich um Disponenten, Betriebsleiter, Mitarbeiter der Betriebssteuerung
- BIAF wird nicht mehr nur rein informativ unterstützend, sondern für die betriebliche Steuerung genutzt
- 2013 wird BIAF in die Liste der betriebskritischen, sogenannten „Prio 1 Systeme“ aufgenommen
- ➔ **Die Verfügbarkeit des Systems wirkt sich auf die Qualität der operativen Prozesse aus. Bei Ausfall des Systems stehen wichtige Informationen zum „Systemzustand“ Flughafen Frankfurt nicht mehr zur Verfügung.**

Neue Anforderungen an die Architektur

- Höhere Ausfallsicherheit und Disaster Recovery Fähigkeit
- Kein Datenverlust bei Ausfall des Rechenzentrums → Wahrung der Datenkonsistenz
- Steigerung der Performance im Bereich der Analytics, im speziellen der „freien“ Analysefähigkeit (unberechenbarer Workload)
- Priorisierung von Diensten und Nutzergruppen (Abbildung der Serviceklassen operational BI vs. classic BI)
- Einfache Skalierbarkeit
- ➔ **Die neue Architektur soll in der Lage sein, heutige und anstehende Probleme lösen zu können, ohne dass es notwendig ist, die Herausforderung an die Hardwareleistung anzupassen.**

Zeitverlauf

Phase 1 (BIAF-Cluster & SAS-Grid)

- Anfang 2011 PoC SAS-Grid
- 2011 Aufbau eines BIAF-Clusters mit zwei Knoten aufgespannt über zwei Brandabschnitte im primären Rechenzentrum (PDC)
- Einsatz von SAS-Grid und Workloadmanagement (Serviceklassen)
- Mitte 2012 Inbetriebnahme BIAF GRID und 9.2 auf AIX

Phase 2 (Analytische Datenbank, SAS 9.4 auf x86, HA&DR)

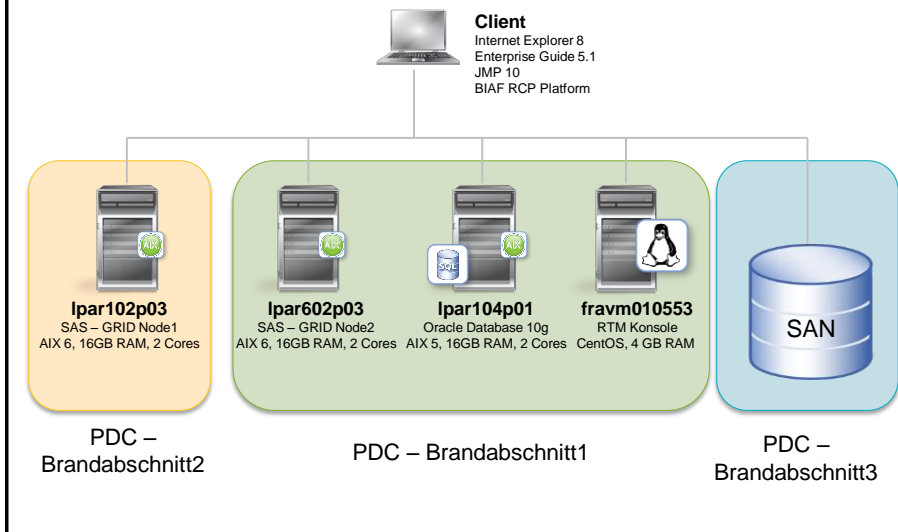
- 2012 PoC – Analytische Datenbanken
- 2013 Aufbau einer SAS/Greenplum Appliance-Umgebung über zwei Rechenzentren
- Migration der SAS-Umgebung auf x86 Architektur unter SUSE Linux
- Fraport als Early Adopter für SAS 9.4
- Geplante Inbetriebnahme Oktober 2013

Architektur der BI-Umgebung

Seite 11



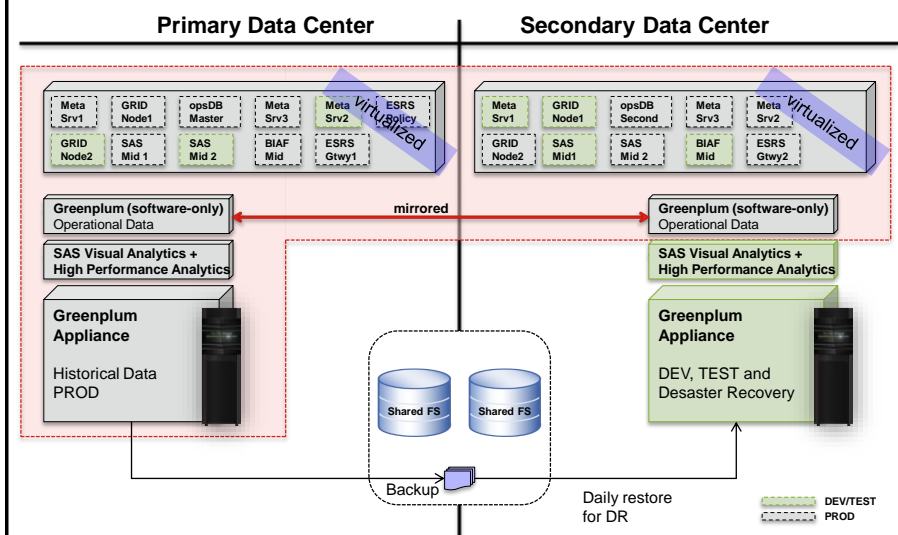
Systemüberblick SAS 9.2 + Oracle



Seite 12



Systemüberblick SAS 9.4 + Greenplum



Hardware in Zahlen

➤ SAS

- 14 virtuelle Maschinen
- 2 physische Maschinen
- 60 Cores
- 188 GB RAM

➤ Greenplum software-only

- 2 virtuelle Maschinen
- 2 physische Maschinen
- 24 Cores
- 195 GB RAM

➤ Greenplum Data Computing Appliance (x2)

- 10 physische Maschinen (2 Master, 8 Segmente)
- 96 Cores
- 384 GB RAM

Spotlights

Storage

- GPFS als shared FS
- Zwei gespiegelte HP-Storageboxen
- Auf zwei Rechenzentren verteilt

Netzwerk

- 10G Leitungen erforderlich („gebondet“ zwischen Rechenzentren)
- Netzwerktechnische Abnahme von GP-Komponenten

High Availability

- Erste Ebene von HA durch Virtualisierung
- „HA Extension“ als Cluster-Manager
- Failover von Prozessen und IP-Adressen
- Lokale Monitor-, Start- und Stopskripte im cron



Spotlights II

Disaster Recovery

- Verteilung von „Hot-Data“ über Rechenzentren („stretched cluster“)
- Täglicher Backup/ Restore von „Cold-Data“
- GPFS ist über Rechenzentren gespiegelt

Greenplum

- Sowohl Appliance, als auch Software-only im Einsatz
- Schnittstelle Access2Greenplum stellte sich als Herausforderung dar
- Bulkloading zwingend erforderlich
- Append-only
- WLM innerhalb von Greenplum
- Datenverteilung wichtig (shared nothing)
- Zeitfaktor: Oracle um den Faktor 60 – 100 langsamer
- Summierung von 1Mrd. Passagieren in 400ms



Key-Features

- SAS High-Performance Analytics Umgebung
- Einstieg in SAS Visual Analytics (In-Memory Processing) & MobileBI
- Einsatz einer analytischen Datenbank (EMC/Greenplum)
 - Reduktion der Datenlayer und Ladestrecken (schnellere Entwicklung neuer Projekte)
 - Höhere Performance
 - Deutlich weniger Aufwand für Optimierung und Tuning notwendig (keine Indizes, etc.)
 - Innerhalb der DB teilweise mehrfach redundant ausgelegt
 - In-Database Analytics
- Online Replikation von virtualisierten Servern möglich
- High-Availability und Disaster Recovery

SAS 9.4

Early Adopter Programm

- Laufzeit von Mai – Juli 2013
- Installation eines EA-Images (Stand April 2013)
- Regelmäßige Telefonkonferenzen mit SAS Deutschland und Cary
- Unterstützung vor Ort durch SAS Deutschland
- Dedizierte Unterstützung seitens SAS Support
- Diverse Bugs und Feature-Requests platziert
- ➔ **Sehr hilfreich, da erste Erfahrungen mit 9.4, jedoch mit zeitlichen Aufwänden verbunden**



General Availability

- Ab 10. Juli 2013
 - Parallele Installation von 9.4 und VA 6.2
 - 9.4 Installation dauerte ca. 4 Tage (3 Umgebungen E, Q und P)
 - VA Installation dauerte ca. 3 Tage (2 Umgebungen Q und P)
- **9.4 scheint die richtige Entscheidung gewesen zu sein.**



Neue Features

- Clustered MetadataServer
- TCServer im Portfolio
- WorkspaceServer unter LSF-Kontrolle
- Grid Option Sets
- (Visual Analytics und High Performance Analytics)
- (Suchfunktion in der SMC)



Migration von 9.2 auf 9.4

- Läuft (bisher) mit weniger Hindernissen, als Migration von 9.1.3. auf 9.2
- Profitieren vom Early Adopter Programm
- Umfang
 - Ca. 400 WRS-Berichte (+ ca. 1000 „private“ Berichte)
 - Ca. 170 STPs
 - Ca. 110 InfoMaps
 - Ca. 520 DI-Jobs
- Portalinhalte stellen Herausforderung dar
- Geschätzte Dauer: ca. drei Monate



Feature Requests

- Shared bin und config Verzeichnisse
- STP unter LSF-Kontrolle
- Clustered MidTier umfänglicher, nicht nur TCServer
- Unterstützung des Clustered MDS durch den EG
- „Kosmetisches“, wie z.B. Servertypen im EG

Fazit

Zu guter Letzt...

- Projektlaufzeit von ca. 6 Monaten
- 9.4 ist (bisher) die richtige Wahl
- Greenplum ist (bisher) die richtige Wahl

On a side note:

- Kein Bulkloading, kein DWH
- Netzwerktechnische Abnahme erforderlich