bidasco SCALE

DATA SCIENCE SYMPOSIUM 2014

Inhalt

MOTIVATION

IDEE

- BIDASCO SCALE

VORAUSSETZUNGEN

UMSETZUNG

- REPLICATION
- QUERY ROUTING
- SQL REWRITING & DATA CONVERSION

FAZIT

bidasco GmbH

Gründung bidasco im Herbst 2013

Spinoff einer Berliner Firma

Übernahme der IP-Rechte, des Codes und von Schlüsselmitarbeitern

Büro an der Schlösserstraße in Erfurt

Bratwurst

MOTIVATION

DATENBANKEN AUF PHYSISCHEN SERVERN

SIND **KOSTENINTENSIV**

- hohe Vorabinvestition in Hardware und Lizenzen
- Erweiterungen verursachen erhöhte Lizenzkosten

KÖNNEN NICHT OHNE **MASSIVE** EINGRIFFE ERWEITERT WERDEN

- Wartungsfenster für Erweiterungen

MOTIVATION

DATENBANKEN AUF PHYSISCHEN SERVERN

SIND AUF **LASTSPITZEN OPTIMIERT**

- keine dynamische Anpassung an Lastaufkommen
- auf vordefinierte Leistung festgelegt

LEIDEN UNTER **ZUGRIFFSKONFLIKTEN**

- Applikationen (schreibend/lesend)
- Reporting (lesend)
- Data Warehouse (lesend)
- Kundenservice (lesend)

MOTIVATION

DATENBANKEN AUF PHYSISCHEN SERVERN

REPLIKATION MIT STANDARD-TOOLS (STREAMS/DATA GUARD) LIMITIERT

- Abhängigkeiten von Versionen und Editionen
- Standby-DBs nur mit Aufwand oder Kosten (Active DG) für Abfragen verwendbar

MOTIVATION

DATENBANKEN AUF PHYSISCHEN SERVERN

KÖNNEN NICHT BELIEBIG FÜR VERSCHIEDENE ANWENDUNGSFÄLLE OPTIMIERT WERDEN

- Lese-Optimierung/ Indizierung

  ⇕

  Schreib-Optimierung

- Replikate können nur mit viel Aufwand individuell angepasst werden

MOTIVATION

AM BEISPIEL ORACLE

bidasco
big data, scalable operations

GOAL:

To set-up an elastic environment to cope with performance peaks (high user/customer load) and high availability.

Setup

- Oracle Database (stand-alone)

Setup

- Oracle Database (Cluster/RAC)

Setup

- Oracle Database extended by bidasco®scale

IDEE

EINSPARUNGEN

- keine zusätzlichen Lizenzen für Replikate
- Einsparung neuer Lizenzen für die Master-Datenbanken
- Lastreduzierung durch Umleiten der Abfragen auf Replikate
- Replikate als VMs = reduziert zusätzliche Investitionen für Hardware

IDEE

REPLIKATION

- Oracle und non-Oracle Datenbanken
  - *MySQL, PostgreSQL*
  - *Optimierung für lesende Zugriffe z.B. NoSQL Datenbanken*
- keine Eingriffe in die Master-DB
- komplett anpassbar an Kundenbedürfnisse
- Replikation in heterogene Datenbank-Umgebungen
  - *z.B. parallele Replikation in Oracle und MySQL Satelliten*

IDEE

SKALIERUNG & ELASTIZITÄT

- bedarfsorientierte Hinzu- bzw. Wegnahme von Replikaten
- Replikate als VMs in Clouds
- Integration von Cluster-Lösungen
  - *wie z.B. Percona MySQL*

IDEE

LESEOPTIMIERUNG

- umlenken der meisten lesenden Zugriffe auf Replikate
  - T*ransparente Schicht zwischen Applikationen und Datenbanken*
- Zugriffsoptimierte Indizierung
  - *Individuelle Datenmengen in Replikaten*
  - *Reduzierung auf benötigte Daten*
  - *Archivierung (Blockierung von löschenden Aktualisierungen)*

VORAUSSETZUNGEN

MASTER-**DB**

- Stand-alone oder RAC
- Standard, Standard One oder Enterprise Edition
  - *ab Oracle 11R1*

Satelliten-**DBs**

- **Oracle**
  - *Standard, Standard One oder Enterprise Edition*
- **MySQL** *ab 5.1*
  - *empfohlen: Percona MySQL*
- **PostgreSQL** (in Roadmap)
- **NoSQL** (in Roadmap)

VORAUSSETZUNGEN

DATENBANK /**APPLIKATIONEN**

- hoher Anteil lesender Zugriffe auf die Master-DB (~80:20)
- Abfragen mit möglichst allgemeinen Statements
- Zugriff via JDBC

UMSETZUNG

DATABASE-MIDDLEWARE

bidasco
big data, scalable operations

UNABHÄNGIGKEIT VON **ORACLE-**VERSIONEN UND EDITIONEN

EINFACHE INSTALLATION & **ADMINISTRATION**

- eigener Server für Replikationsschicht
- Konsole zur Steuerung der Replikation
- kein Queueing auf Festplatte/Storage

REPLIKATION MIT STANDARD-**TOOLS/APIS**

- neuer DB-User für Replikation in Master-DB
- Standard-APIs
  - LogMiner
  - JDBC

UMSETZUNG

DATABASE-MIDDLEWARE

BEI BEDARF ABSCHALTBAR **OHNE** PRODUKTIONS-UMGEBUNG ZU **BEEINFLUSSEN**

TRANSPARENTE SCHICHT ZWISCHEN DER DATENBANK UND DEN ZUGREIFENDEN **APPLIKATIONEN / ANWENDERN**

- erweiterter JDBC-Treiber (Wrapper)
  - *umleiten von lesenden Zugriffen*
  - *1:1 Durchgriff von schreibenden und Oracle-spezifischen Abfragen auf Master*
  - *integriertes Load Balancing*
- Übersetzung von SQL-Dialekten

UMSETZUNG

DATABASE-MIDDLEWARE

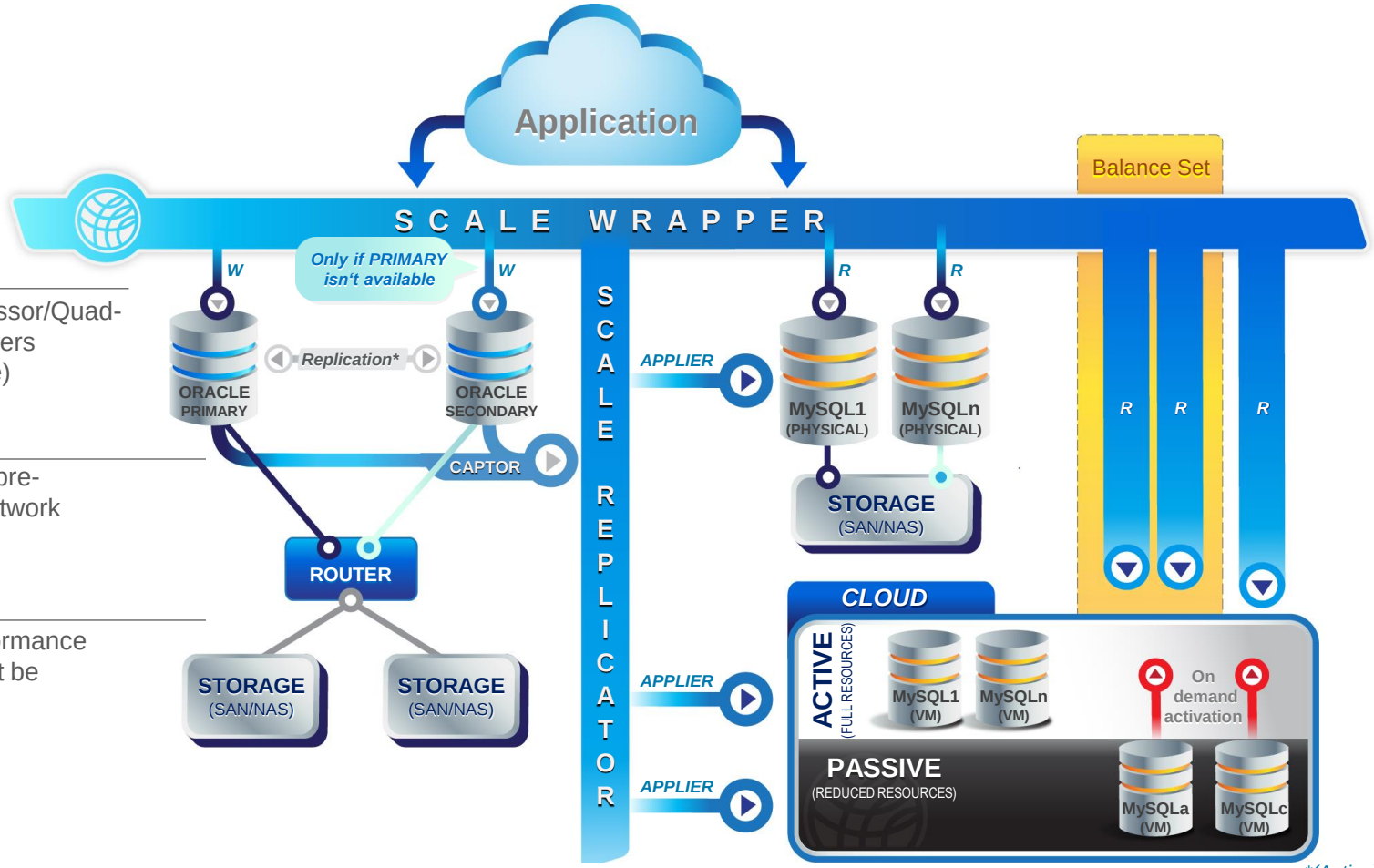
Setup

Oracle extended by bidasco®scale

2 or 4 - Processor/Quad-Core DB-Servers (i.e. 2x 4-Core)

Redundant Fibre-Channel & Network

Mid-size/performance storage (might be redundant)



*(Active) DataGuard

UMSETZUNG

Oracle erweitert um bidasco®scale

PROS

Hochverfügbarkeit
- MASTER: anfällig für Server Fehler (manuelle Ausfallsicherung durch Sekundäre Knoten)
- SATELLITEN: mehrere identische Instanzen

Eingebaute Lastverteilung
- Scale Wrapper verteilt Lesezugriffe
- Verteilungs-Sets (Bindung von Applikationen/Nutzern zu spezifischen Satelliten)

Verringert Kosten
- Kleiner dimensionierte Master Hardware/Speicher
- Keine zusätzlichen Kosten für Satelliten

Hohe Performance erreichbar
- Lesezugriffe werden auf Satelliten verteilt (~80%)
  - Optimierung von Lesezugriffen (zusätzl. Indices, Lese Cache/Puffer etc.)
- Schreibzugriffe verbleiben auf dem Master (~20%)
  - Performance von Schreib- Transaktionen können optimiert werden (weniger Indices, Puffer, etc.)

Satelliten können je nach Andwendungs-/Nutzeranforderung dimensioniert werden
- Verteilte Hochleistungsserver (aber kleiner dimensioniert als Oracle Server)
- Virtualisierte Server
- Individuelle Datensätze
  - Schema für Schema Replikation – Volle Replikation ist optional

Hohe Skalierbarkeit
- Hinzufügen von Satelliten on-demand
  - Aktivierung von inaktiven VMs – Klonen von physischen Servern

Heterogenes Setup möglich
- MASTER & SATELLITEN können unterschiedliche OS nutzen (AIX, Linux, etc)
- MASTER & SATELLITEN können verschiedene Datenbanken nutzen (M: 10g S: 11g)

CONS

Komplexes Set-up

Drittanbieter Produkt (nicht Oracle)

REPLICATION

REPLIKATIONSSCHICHT

bidasco
big data, scalable operations

AUSLESEN DER REDOLOGS UND ARCHIVELOGS KEIN QUEUEING AUF FESTPLATTE/STORAGE

IN MEMORY PROCESSING

PRO SCHEMA

AUF KUNDENBEDÜRFNISSE ANPASSBAR

- limitierte Datenmengen
- Blacklist für Tabellen
- Blacklist für SQL-Anweisungen
  - *z.B. nur INSERT und UPDATE*

REPLICATION

REPLICATION MODES

KÖNNEN PRO SATELLIT ODER GLOBAL **AKTIVIERT WERDEN**

PERSISTENTE **SPEICHERUNG** DER KONFIGURATION

- neue Satelliten übernehmen automatisch globale Replication Modes

SKIP

- gezieltes Ignorieren von SQL-Anweisungen

OVERWRITE

- Überschreiben der Daten in den Satelliten mit den Master-Daten

REPLICATION

REPLICATION MODES

COMPAIR & **REPAIR** (Roadmap)

- interaktive Behebung von Konflikten

COMPRESSION AND **ENCRYPTION** (Roadmap)

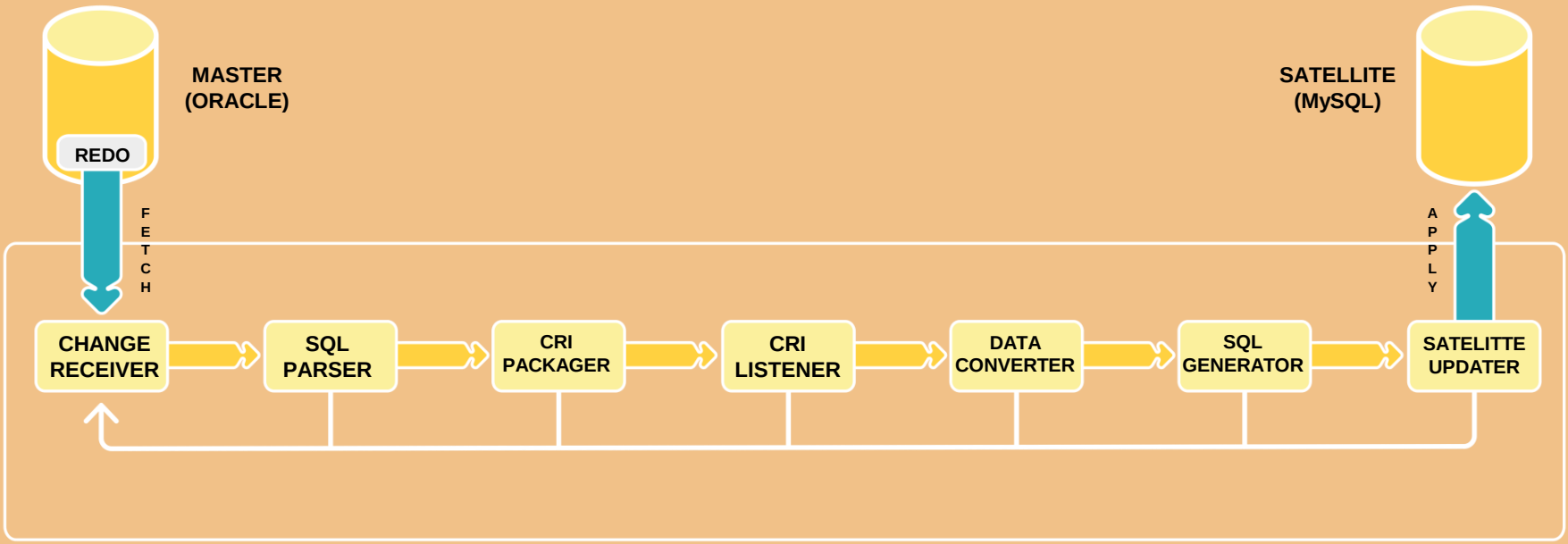
- Optimierung für WAN Replikation

BULK

- Umschalten auf JDBC BATCH UPDATE/INSERT
- Beschleunigung bei vielen, gleichartigen Statements (z.B. Migrationen)

REPLICATION

REPLICATION MODES



CRI – Change **R**ecord **I**tem

QUERY ROUTING LAYER

TRANSPARENTE SCHICHT ZWISCHEN **APPLIKATION UND DATENBANK**

- Übersetzung von SQL-Dialekten
- Result Sets werden als Oracle-Result Sets zurückgeliefert

INTELLIGENTES **QUERY ROUTING**

- lesende, kompatible / übersetzbare Abfragen gegen Satelliten
- schreibende und Oracle-spezifische Abfragen gegen Master-DB

QUERY ROUTING LAYER

KONSISTENZ-**CHECKS**

- Routing nur auf konsistente Satelliten
- Re-Routing auf Master falls keine Satelliten-DB konsistent

INTEGRIERTES **LOAD BALANCING**

- Round Robin
- Satellit mit kürzester Antwortzeit

SQL REWRITING & DATA CONVERSION

SQL **REWRITING**

- DML STATEMENTS
  - *INSERT, UPDATE und DELETE*
- STANDARD-SQL FUNKTIONEN
  - *AVG, SUM, MIN, MAX, etc.*
- JOINs
  - *INNER JOIN, OUTER (left/right) JOIN, FULL OUTER JOIN*

SQL REWRITING & DATA CONVERSION

DATA **CONVERSION**

- Basic data types (NUMBER, CHAR, VARCHAR2, DATE, etc.)

  Oracle: ja

  MySQL: ja

- Binary data types (RAW, CLOB, LONG, etc.)

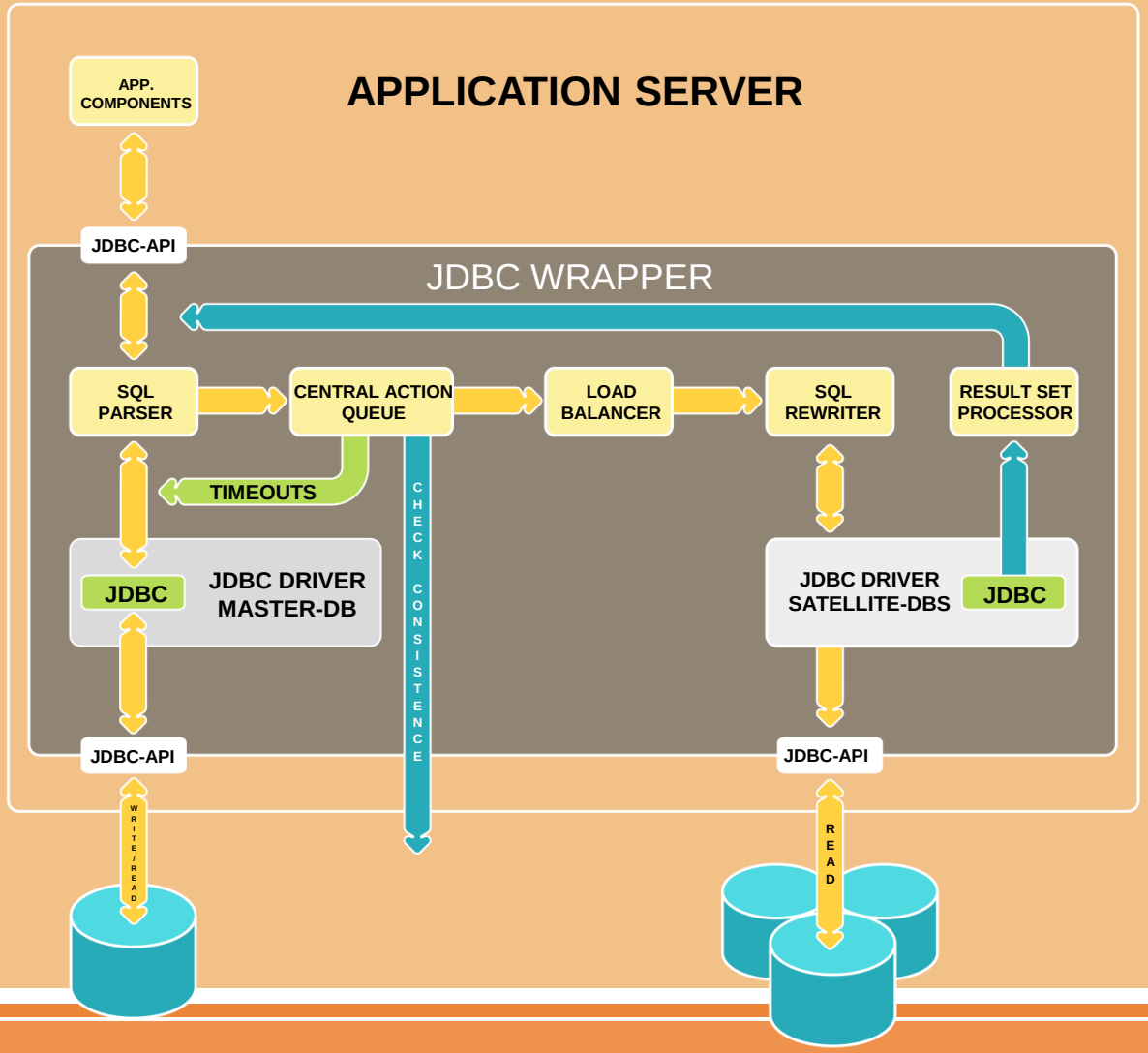
  Oracle: ja

  MySQL: eingeschränkt, wird aktuell entwickelt

TRANSLATION API (in Roadmap)

- Schema/Daten Abstraktion

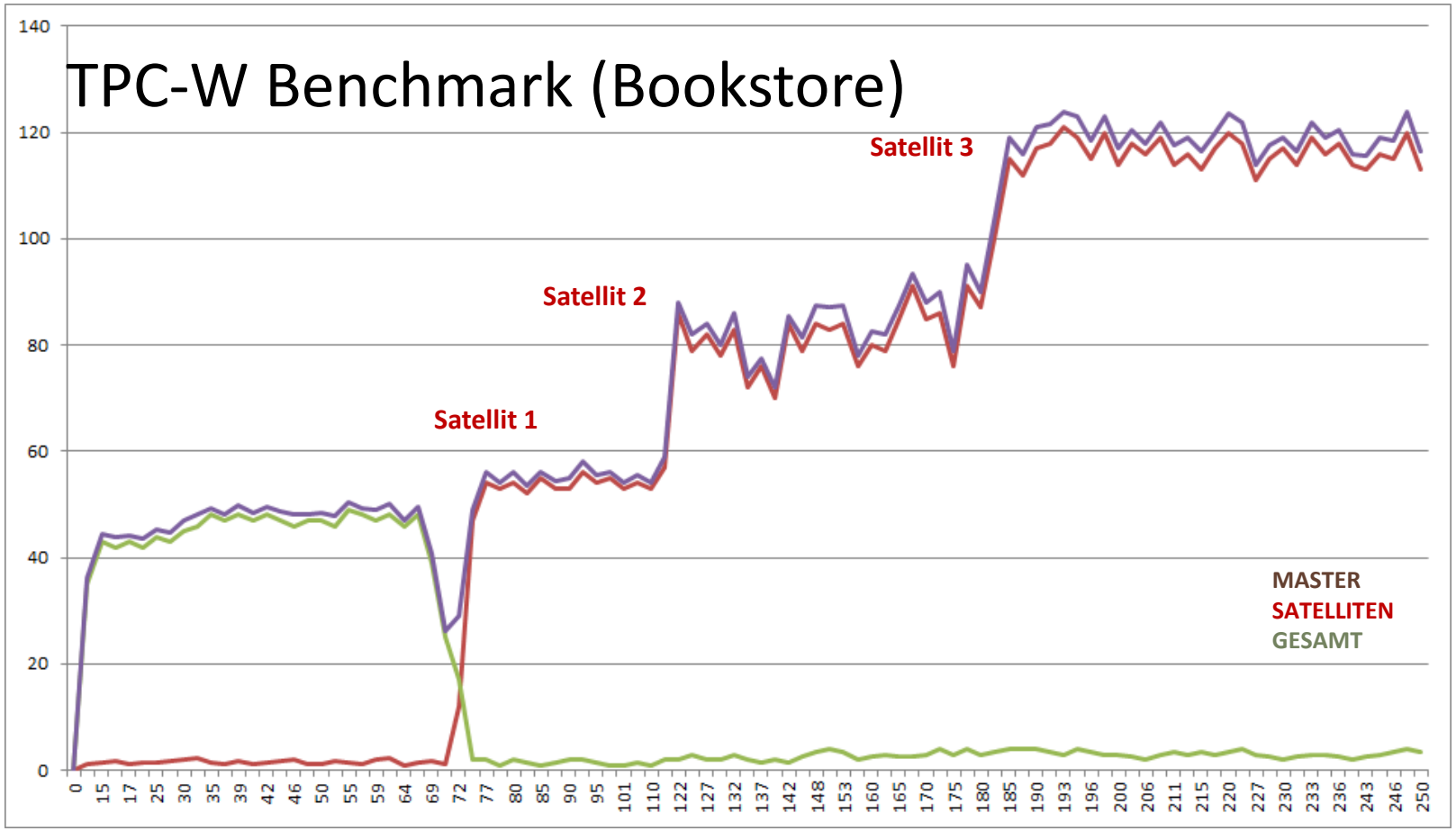
QUERY ROUTING LAYER



KONSOLE

Command	Description
version	Shows the Replicator version
help <command>	Shows help for a command
connect <IP/hostname>	Connects to a Replicator instance
disconnect	Disconnects from the Replicator instance
status	Shows the client status
addsat	Adds a satellite to the Replicator (interactive)
remsat <satellite_ID>	Removes a satellite from the Replicator
showsat <satellite_ID>	Shows properties about a satellite
showsat all	Shows properties about all satellites
start <satellite_ID>	Starts replication to a satellite (interactive)
start all	Starts replication to all satellites (interactive)
restart <satellite_ID>	Restarts replication to a satellite
restart all	Restarts replication to all satellites
stop <satellite_ID>	Stops replication to a satellite
stop all	Stops replication to all satellites
stop abort <satellite_ID>	Hard stops replication to a satellite
stop abort all	Hard stops replication to all satellites
stats <satellite_ID>	Shows statistics about a satellite
clearsat <satellite_ID>	Removes replicator related state from a satellite
bulk <enable/disable> <satellite_ID>	Toggles Batch Mode On or Off
createtables <satellite_ID>	Creates master db tables at the specified satellite.
createindexes <satellite_ID>	Creates indexes on specified satellite.
exit	Disconnects and exits from the console
quit	Disconnects and exits from the console
help	Shows help

MESSUNGEN


TPC-W Benchmark (Bookstore)
Satellit 1
Satellit 2
Satellit 3
MASTER
SATELLITEN
GESAMT
140
120
100
80
60
40
20
0
0 15 17 25 30 35 39 42 46 50 55 59 64 69 72 77 80 85 90 95 101 110 122 127 132 137 142 148 153 160 165 170 175 180 185 190 193 196 200 206 211 215 220 227 230 233 236 243 246 250

CONCLUSION

bidasco®scale erhöht die Datenbank-Performance durch Verteilen der Lesezugriffe auf skalierbare, elastische Datenbank-Umgebungen, in der on-demand Hardware und Software Ressourcen verfügbar sind. (in der Cloud, etc.)

bidasco®scale ermöglicht mittels heterogener Setups signifikante Kostenersparnis durch das Betreiben der Satelliten auf unterschiedlichen Datenbanksystemen.

bidasco®scale integriert heterogene Datenbanksysteme und übernimmt die Verwaltung. Dazu gehören Replikation, Anfrageübersetzungen, Datenkonvertierungen und Selektierung.

FRAGEN ?

Felix Egloff

Geschäftsführer

Michael Woidak

Research & Development

bidasco GmbH

Schlösserstr. 39

99084 Erfurt

www.bidasco.com