

Polystore Systeme

Alina Böttcher
Grundseminar 2018

Inhalt

- Was ist das?
- Motivation
- Beispiel mit Aufbau und Funktion
- Aktuelle Probleme
- Konferenzen
- Quellen

Was ist das

Applikationen



RDMS

Graph Datenbank

Datenbank n



Motivation

- Jede Datenbank ist in einer bestimmten Sache gut
- Manche Anwendungen:
 - benötigen unterschiedliche Daten
 - benötigen eine hohe Performance

Leitfäden

- Eine Datenbank passt nicht zu allen Daten
- Echtzeitanalyse ist entscheidend
- Datenbankschnittstellen sind vielfältig
- Komplexe Analysen müssen möglich sein

Beispiel

- BigDAWG Polystore System
 - Big Data Analytics Working Group
- Daten: Intensivstation
 - Name: MIMIC II (Multiparameter Intelligent Monitoring in Intensive Care II)

Daten

- Persönliche Patientendaten
- Wellenformen : Herzfrequenz, Sauerstoffsättigung ...
- Arzt/Krankenschwester Notizen
- Laborergebnisse
- Rezepte

Beispiel Analyseanforderungen

- Standard SQL Anfragen
- Berechnung der schnelle Fourier-Transformation von Wellenformen und anschließende Vergleiche
- Volltextsuche : Finde Notizen wo der Arzt „Medikament hat gut gewirkt“ reingeschrieben hat
- Echtzeitüberwachung der Herzfrequenz

Datenbanken

- **Postgres:** persönliche Patientendaten
- **SciDB:** historische Wellen als Time-Series
- **S-Store:** Ströme von Geräteinformationen
- **Apache Accumulo:** Texte als Key-Value Format

Applications

Examples, Tests, Sample Applications

Middleware

BigDAWG Common Interface/API

Islands

Relational Island

Array Island

Text Island

Shims

Engines



PostgreSQL



SciDB



Accumulo

Cast

Cast

MIMIC II
Data

- Patient History
- Patient Medication

- Patient Timeseries
Waveform

- Doctor/Nurse notes
- Test notes

Applications

Middleware

Islands

Shims

Engines

**MIMIC II
Data**

Examples, Tests, Sample Applications

BigDAWG Common Interface/API

Relational Island

Array Island

Text Island



PostgreSQL

- Patient History
- Patient Medication



SciDB

- Patient Timeseries
Waveform



Accumulo

- Doctor/Nurse notes
- Test notes

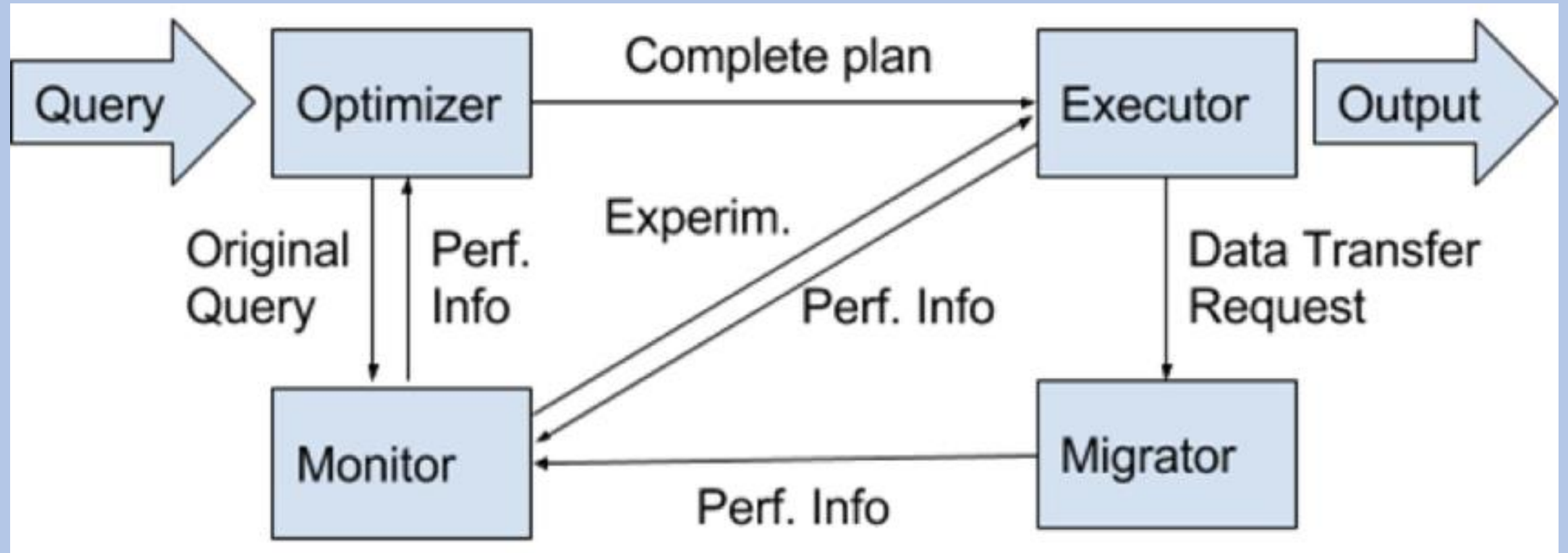
Cast

Cast

Middleware

- Anfragen annehmen
- Anfragen planen
- Bestimmen von effizienten Ausführstrategien
- Pflege der Historie von vorherigen Abfragen mit deren Performance

-> lernt immer dazu



Applications

Examples, Tests, Sample Applications

Middleware

BigDAWG Common Interface/API

Islands

Relational Island

Array Island

Text Island

Shims

Engines

PostgreSQL

SciDB

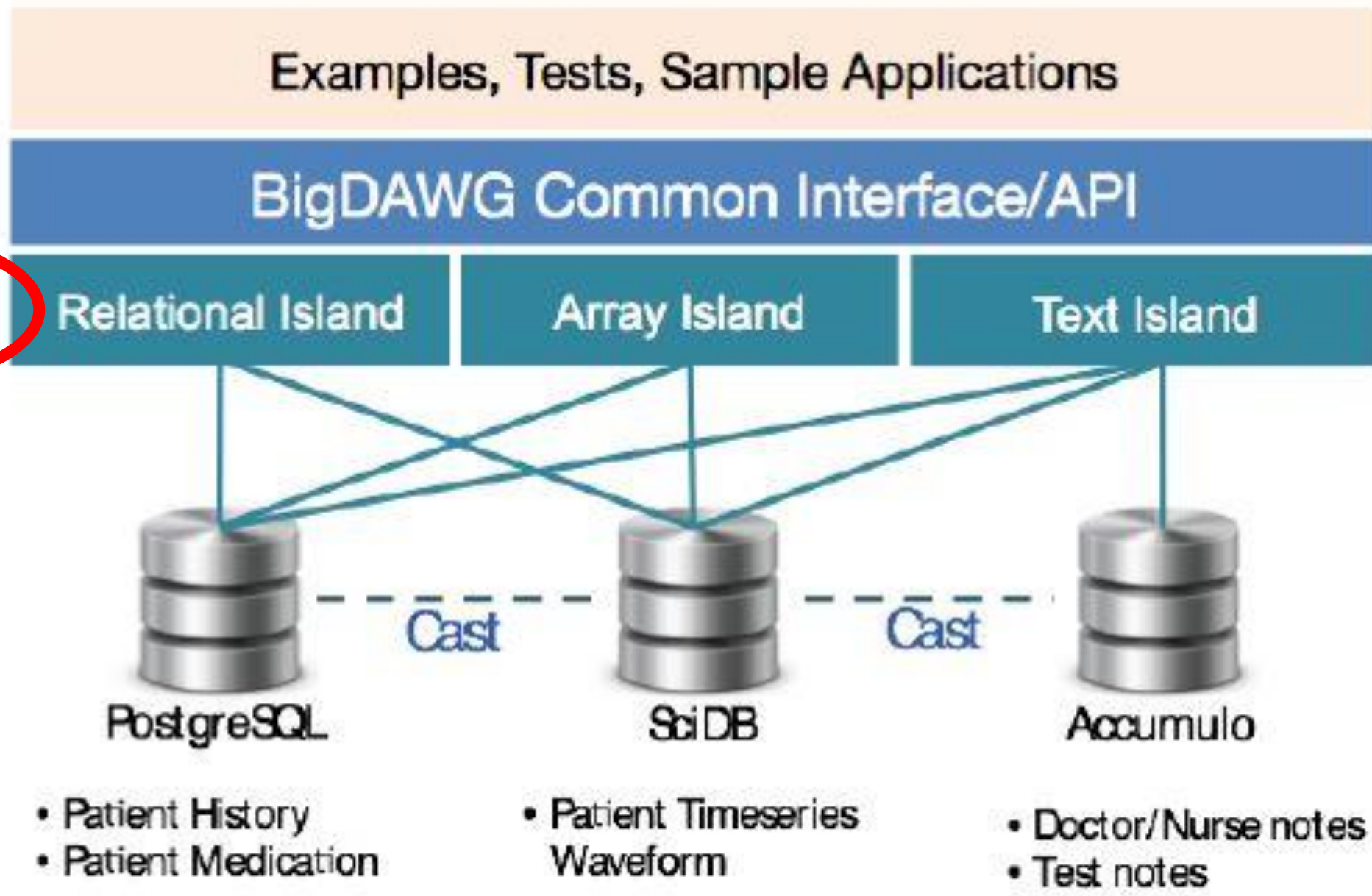
Accumulo

MIMIC II
Data

- Patient History
- Patient Medication

- Patient Timeseries
Waveform

- Doctor/Nurse notes
- Test notes



Islands

- Degenerate islands
- Multi-system islands
 - D4M (Apache Accumulo, SciDB und Postgres)
 - Myria (MyriaX, SciDB und Postgres)

Islands

- Beherrschen die entsprechende/n Datenbanksprache/n
- Kennen das entsprechende Datenmodell
- Haben eine Verbindung zu einer/mehrerer Datenbank/en
- Die Verbindung zur Datenbank erfolgt über Shims

Applications

Examples, Tests, Sample Applications

Middleware

BigDAWG Common Interface/API

Islands

Relational Island

Array Island

Text Island

Shims

Engines

PostgreSQL

SciDB

Accumulo

MIMIC II
Data

- Patient History
- Patient Medication

- Patient Timeseries
Waveform

- Doctor/Nurse notes
- Test notes

Cast

Cast



Applications

Examples, Tests, Sample Applications

Middleware

BigDAWG Common Interface/API

Islands

Relational Island

Array Island

Text Island

Shims

Engines

PostgreSQL

SciDB

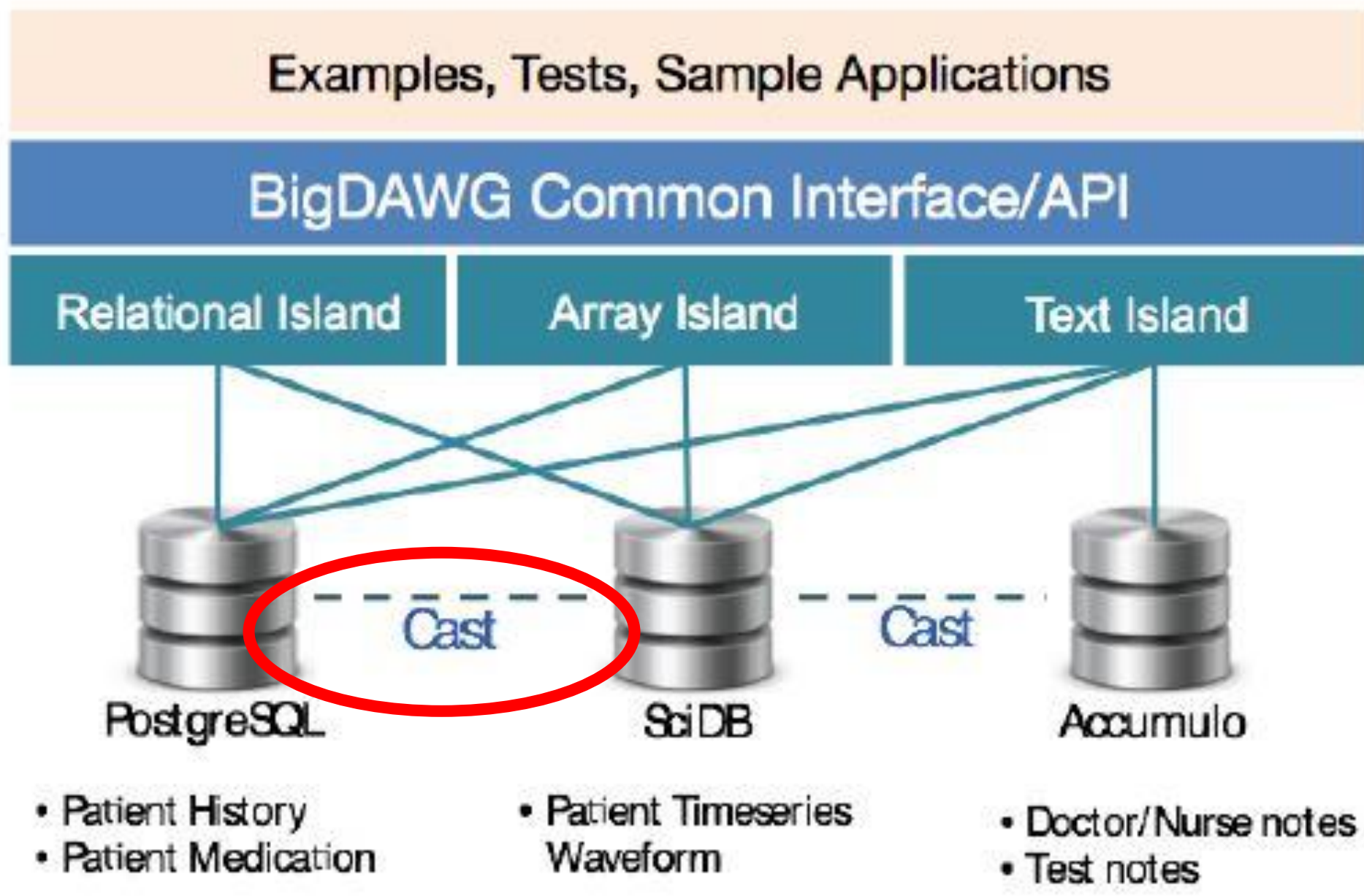
Accumulo

**MIMIC II
Data**

- Patient History
- Patient Medication

- Patient Timeseries
Waveform

- Doctor/Nurse notes
- Test notes



Aktuelle Probleme

- Wenn Anfragen formuliert werden, müssen noch Scopes eingebaut werden -> Ansage welche Info aus welcher Datenbank kommt
-> nicht bei multi-system islands
- Anfragenoptimierung hat noch Potenzial
- Mehr mögliche Datenbanken und entsprechend mehr benötigte Inseln

Konferenzen

- IEEE High Performance Extreme Computing Conference (HPEC)
- IEEE International Conference on Big Data (Big Data)
- IEEE International Conference on Data Engineering (ICDE)

Quellen

- <https://ieeexplore.ieee.org/document/7761636>
(The BigDAWG Polystore System and Architecture)
- <http://www.vldb.org/pvldb/vol8/p1908-Elmore.pdf>
(A Demonstration of the BigDAWG Polystore System)