Babelfish for PostgreSQL

Jim Nasby
Senior Database Engineer
AWS

Debbie Cerda
Director of Business Development
Command Prompt

Rob Verschoor
Principal Database Engineer
AWS

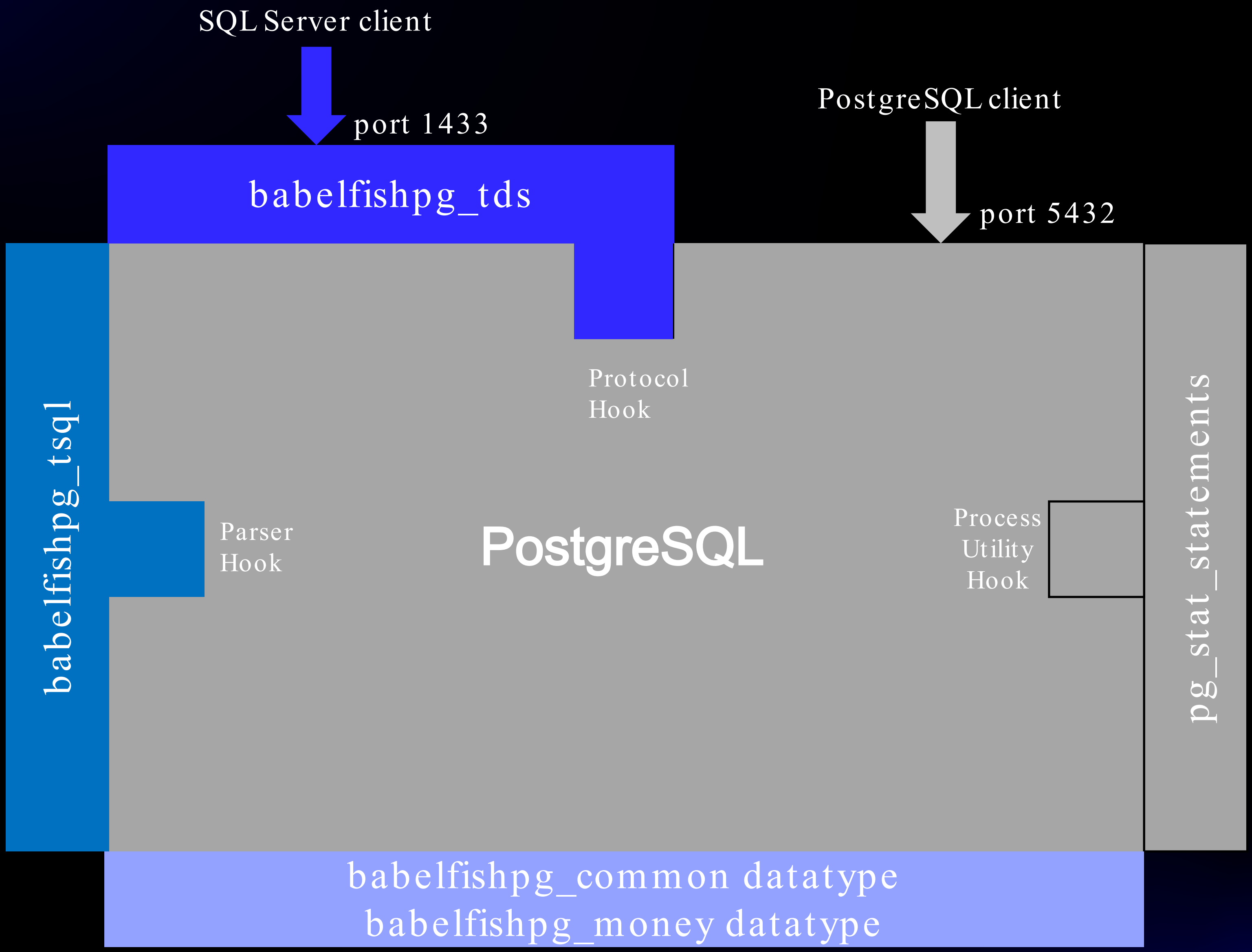
Gopinath Pai
Principal Technical Program Manager
AWS

What is Babelfish for PostgreSQL

Babelfish is a migration accelerator for moving SQL Server Applications to PostgreSQL. It provides the capability for the execution of T-SQL statements over the TDS protocol.

This capability has been natively implemented in PostgreSQL.

Babelfish architecture



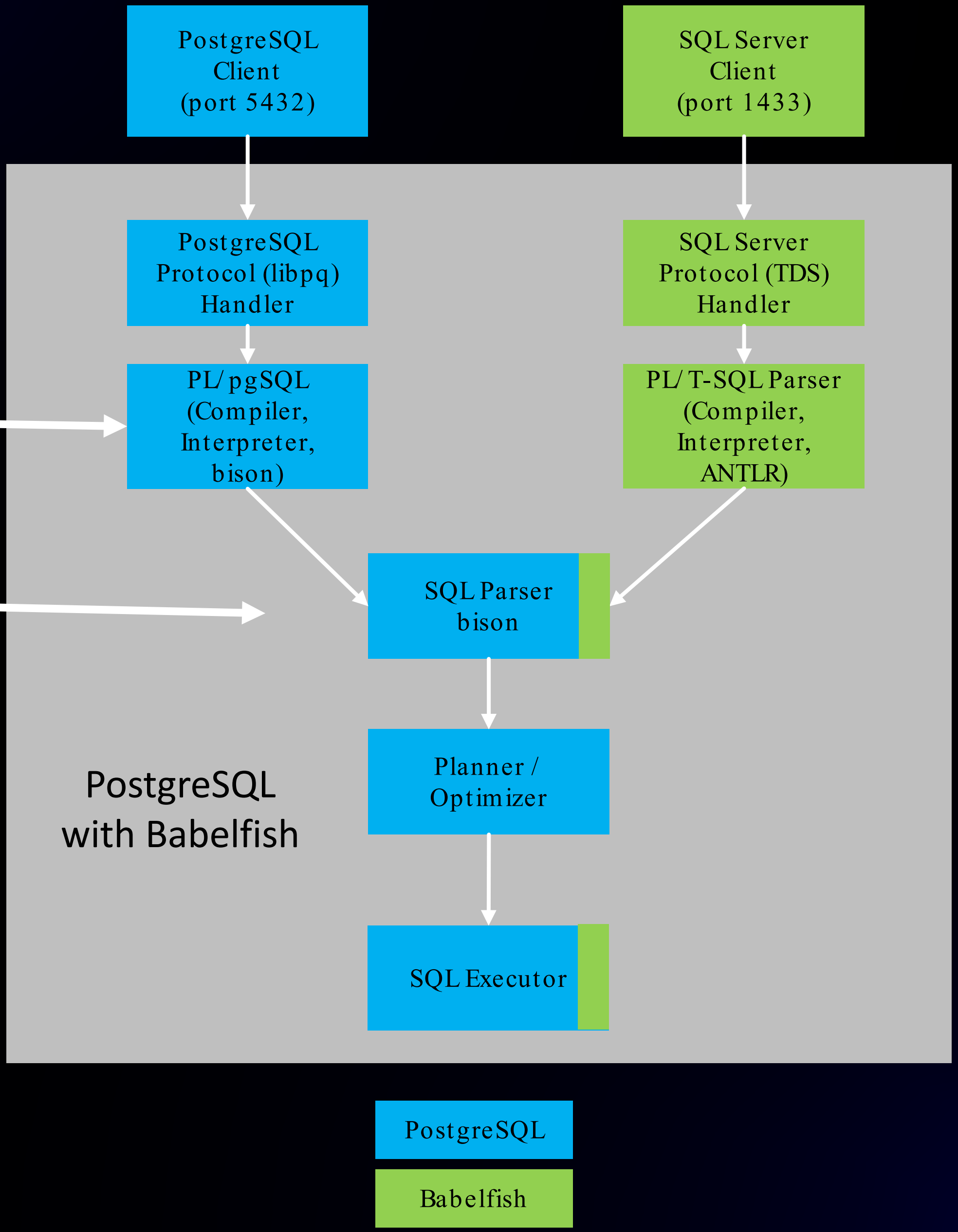
T-SQL Extension

QUERY 1:

SELECT * FROM products WHERE category_id = 1;

Validate syntax
Find identifiers and keywords.
OUTPUT: Raw Parse Tree

Add detailed info
Table OID, column name, column type
OUTPUT: Query Tree



PostgreSQL Client (port 5432)
SQL Server Client (port 1433)
PostgreSQL Protocol (libpq) Handler
SQL Server Protocol (TDS) Handler
PL/pgSQL (Compiler, Interpreter, bison)
PL/T-SQL Parser (Compiler, Interpreter, ANTLR)
SQL Parser bison
Planner / Optimizer
SQL Executor
PostgreSQL with Babelfish
PostgreSQL
Babelfish

```
DECLARE @id int = 100

IF (@a = 42)
BEGIN
  PRINT @a
  DELETE customer WHERE customer_id=@id
END
```

Validate syntax
Find identifiers and keywords.
OUTPUT: Raw Parse Tree

Babelfish for PostgreSQL design tenets

GUIDING PRINCIPLES

No compromises on correctness
Database calls either work the same as in SQL Server or return an error

Wire protocol compatibility
Applications work without changing database drivers

Interoperability
Use PostgreSQL functionality from T-SQL and T-SQL functionality from PostgreSQL code

Deployment model for Babelfish for PostgreSQL

HOW DO I ADD NEW FUNCTIONALITY IN MY MIGRATED APPLICATIONS?

Develop new functionality in T-SQL using SQL Server database drivers

Develop new functionality in PostgreSQL using PostgreSQL database drivers

Develop new functionality in PostgreSQL and call from T-SQL using SQL Server database drivers

Potential migration opportunities

- Home-grown applications
- Database-agnostic applications
- ISV applications
- RDS for SQL Server databases
- On-premises SQL Server databases
- Self-managed SQL Server on Amazon EC2 or Azure VMs
- Azure SQL Distributed Transaction Units

Migration Steps

1. Export DDL (reverse-engineer with SSMS)
   - Make sure to include triggers, logins, owners, and permissions (not included by default)

2. Run Babelfish Compass assessment tool on the DDL to find incompatibilities
   - Rewrite SQL you find to be Babelfish-incompatible. Ex: SELECT..[UN]PIVOT
   - Compass can rewrite selected features with supported T-SQL (MERGE, numeric datetime)

3. Import adjusted DDL script into Babelfish with **sqlcmd**
   - No AWS SCT conversion needed! Babelfish supports T-SQL SQL/ DDL syntax
   - First set Babelfish escape hatches to 'ignore' with sp_babelfish_configure

4. Migrate data using AWS Database Migration Service (DMS)
   - (Or, test with a smaller data set to test getting the app going)

5. Reconfigure the client app to connect to Babelfish instead of SQL Server

Support for SQL Development Tools

- Limited support for SSMS (Query Editor works)

- DBeaver (recommended GUI tool)
  - Free, open source and works on all major OSes (Win/Mac/Linux)
- **sqlcmd** (recommended for script execution)

- With other tools, your mileage will vary
- High priority to support other tools post GA (such as VS Code)

Open Source Project

Project website
https://babelfishpg.org

Freedom from proprietary databases

No vendor lock-in

Apache 2.0 and PostgreSQL licenses

Use it for any purpose, innovate, and distribute your modifications

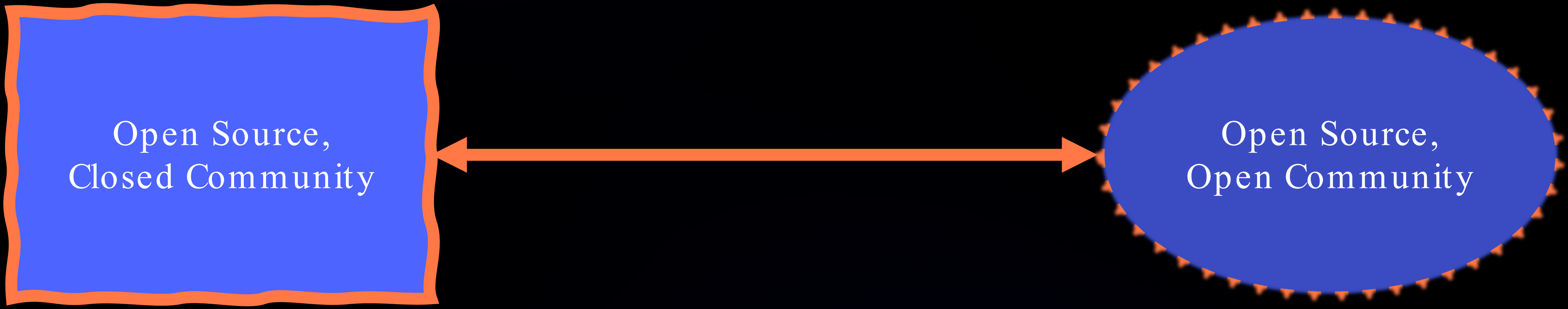
Available on GitHub
https://github.com/babelfish-for-postgresql

Is community driven

Open Source Status

- Initial Babelfish development was closed source (1,155 commits)
- Babelfish launched as open source – as a single code drop
- Babelfish 1.1 (143 commits) and 1.2 (219 commits) open sourced individual commits
- Babelfish 2.0+ will be open development

Open Source vs Open Community



Is Open Source Enough?

Open source Babelfish allows users to control their operations – but only on their own.

Open Community Babelfish allows users to influence the direction of Babelfish development.

Babelfish Open Community

- Babelfish will operate as a stand-alone open community project.
- Amazon developers will work in a Github fork – just like any other contributor.
- Still early on this journey, a lot of open community work pending

Next Steps

- Develop structure and governance
  - Formalize roles
  - Establish commit access policy
- Communicate about the Project
- Build the Community

Open Source Contribution Core Principles

- Projects need contributions with people of all types of skills and all levels of expertise.

- The best project to start working on is one that you use already - or supports another open source tool like PostgreSQL!

Get involved!

- Familiarize yourself with the community project
- Provide input for any feature we build
- Spread the word

Thank you!

Rob Verschoor

rcv@amazon.com