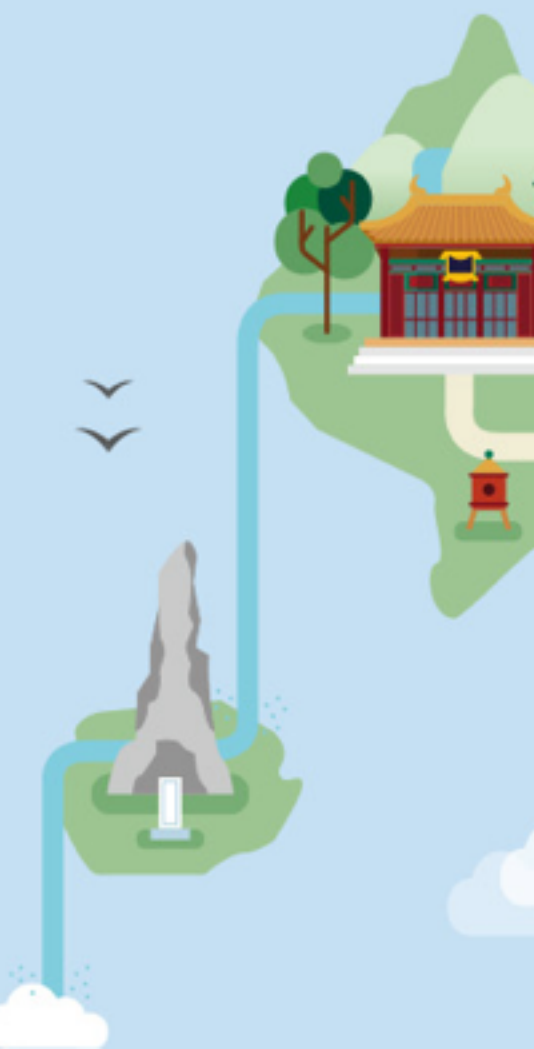




Go in TiDB

申砾@PingCAP



Agenda

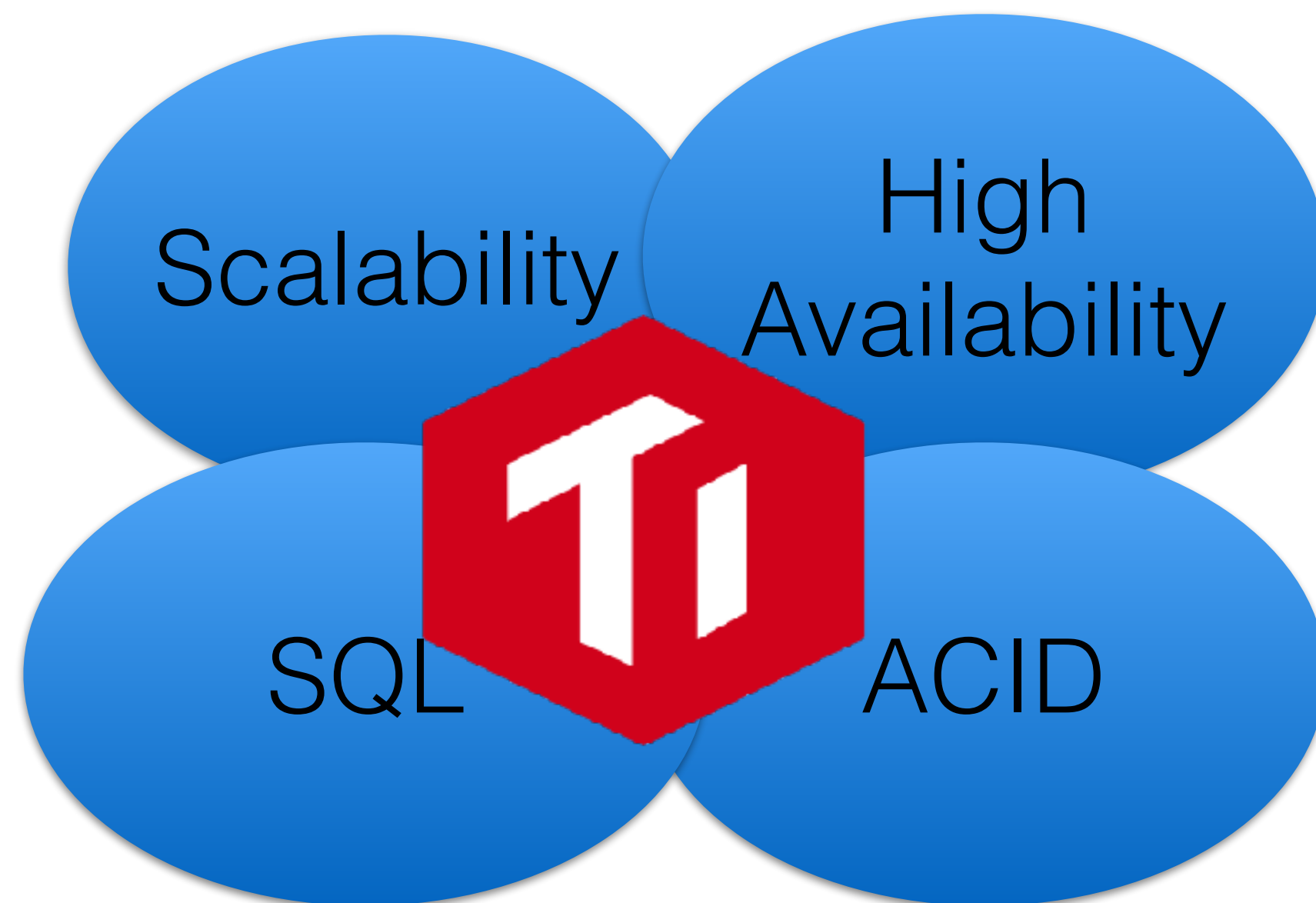
- About Me
- What is TiDB
- TiDB Architecture
- SQL Layer Internal
- Golang in TiDB
- Next Step of TiDB

About Me

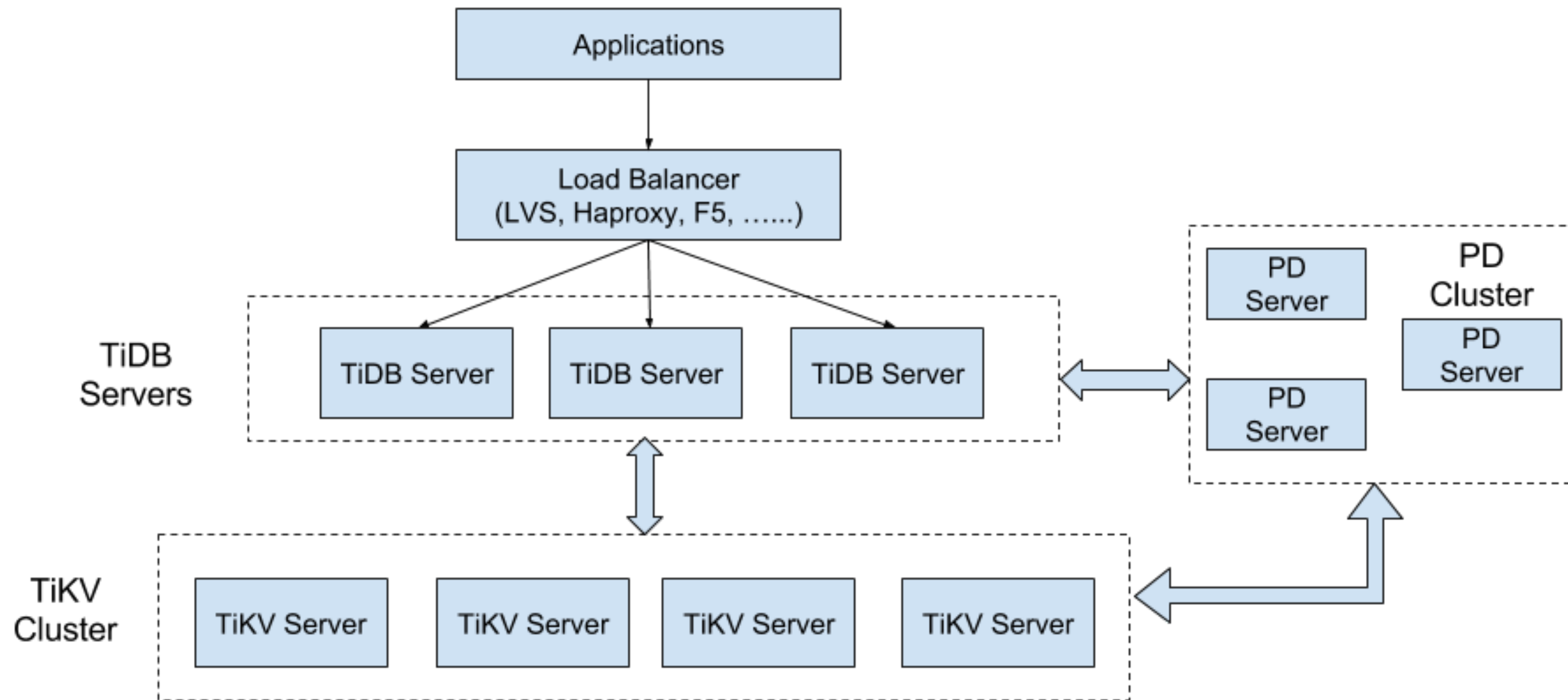
- Shen Li (申砾)
- Tech Lead of TiDB, VP of Engineering
- Netease / 360 / PingCAP
- Infrastructure software engineer

What is TiDB

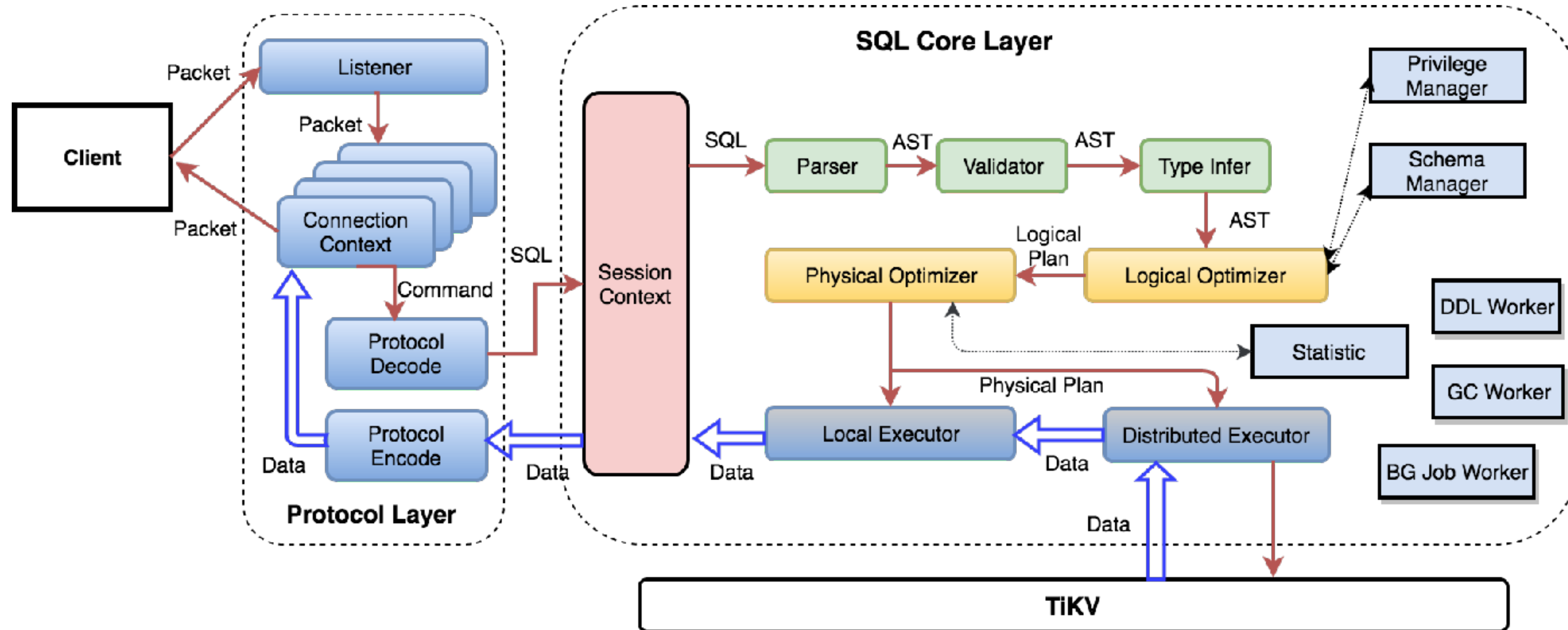
A **Distributed, Consistent, Scalable, SQL** Database that supports the best features of both traditional RDBMS and NoSQL.



Architecture - the Whole Picture



Architecture - SQL Layer



Example - SQL

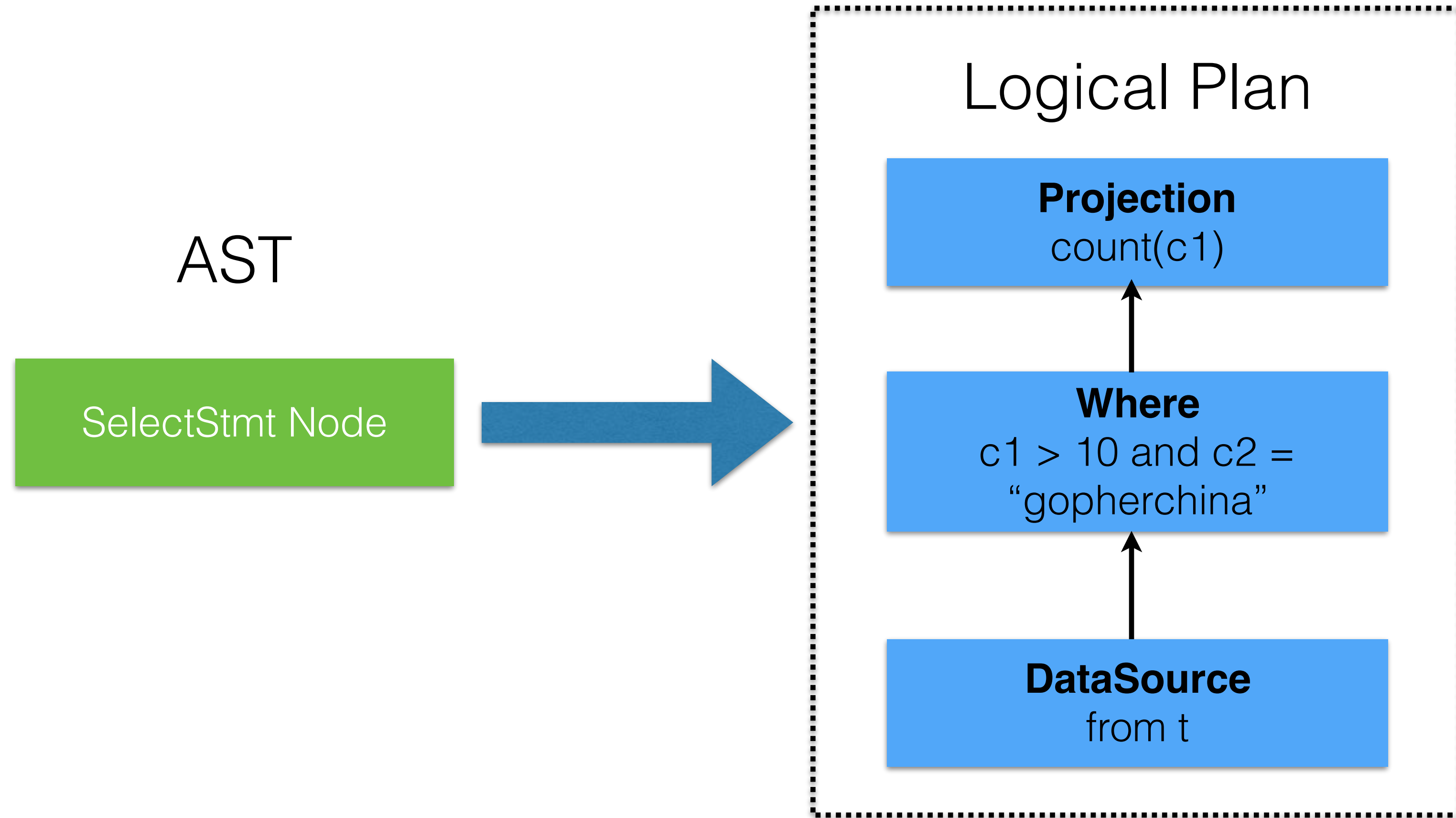
Schema:

```
CREATE TABLE t (c1 INT, c2 VARCHAR(32), INDEX  
idx1 (c1));
```

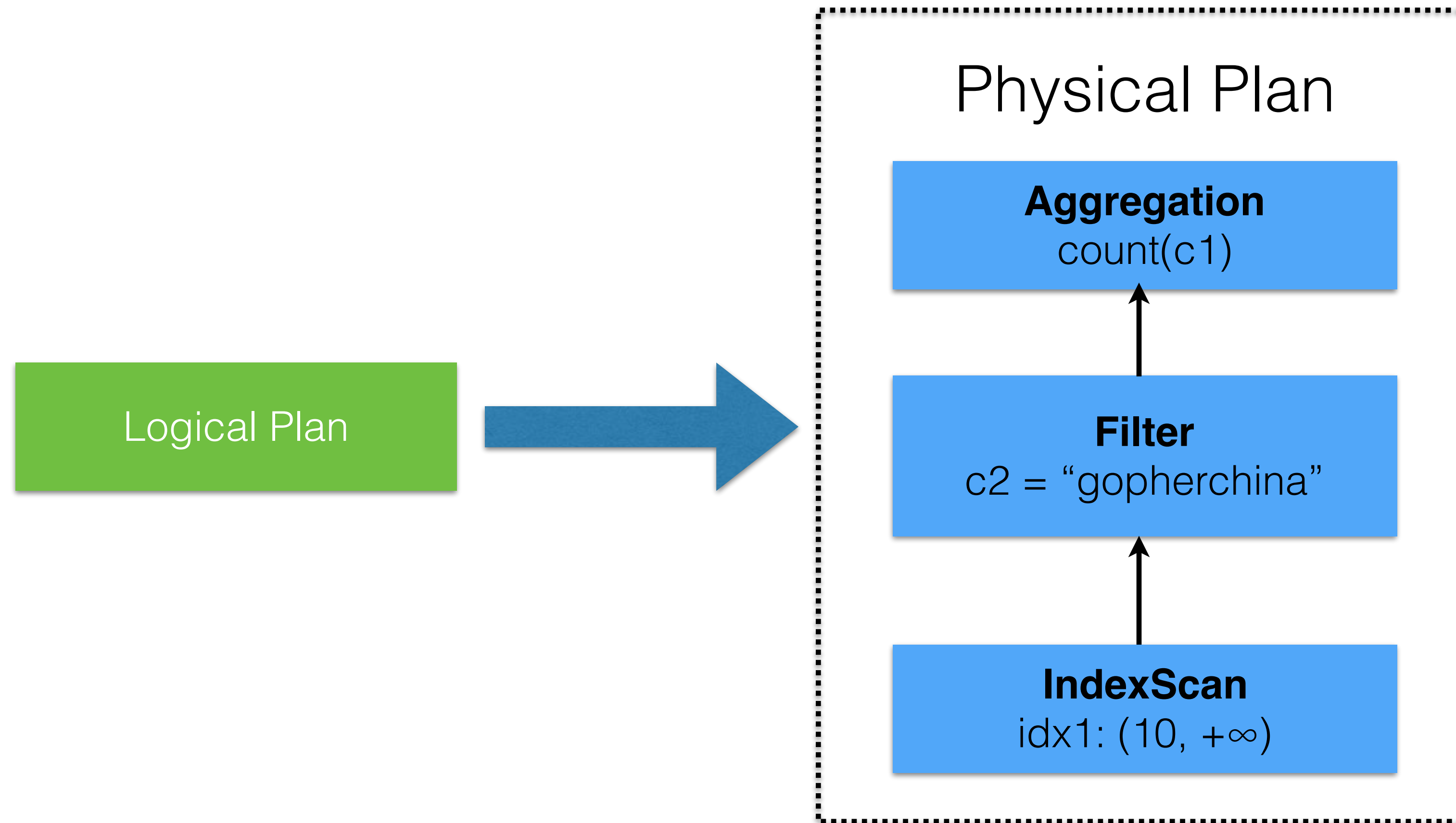
Query:

```
SELECT COUNT(c1) FROM t WHERE c1 > 10 AND c2  
= "gopherchina";
```

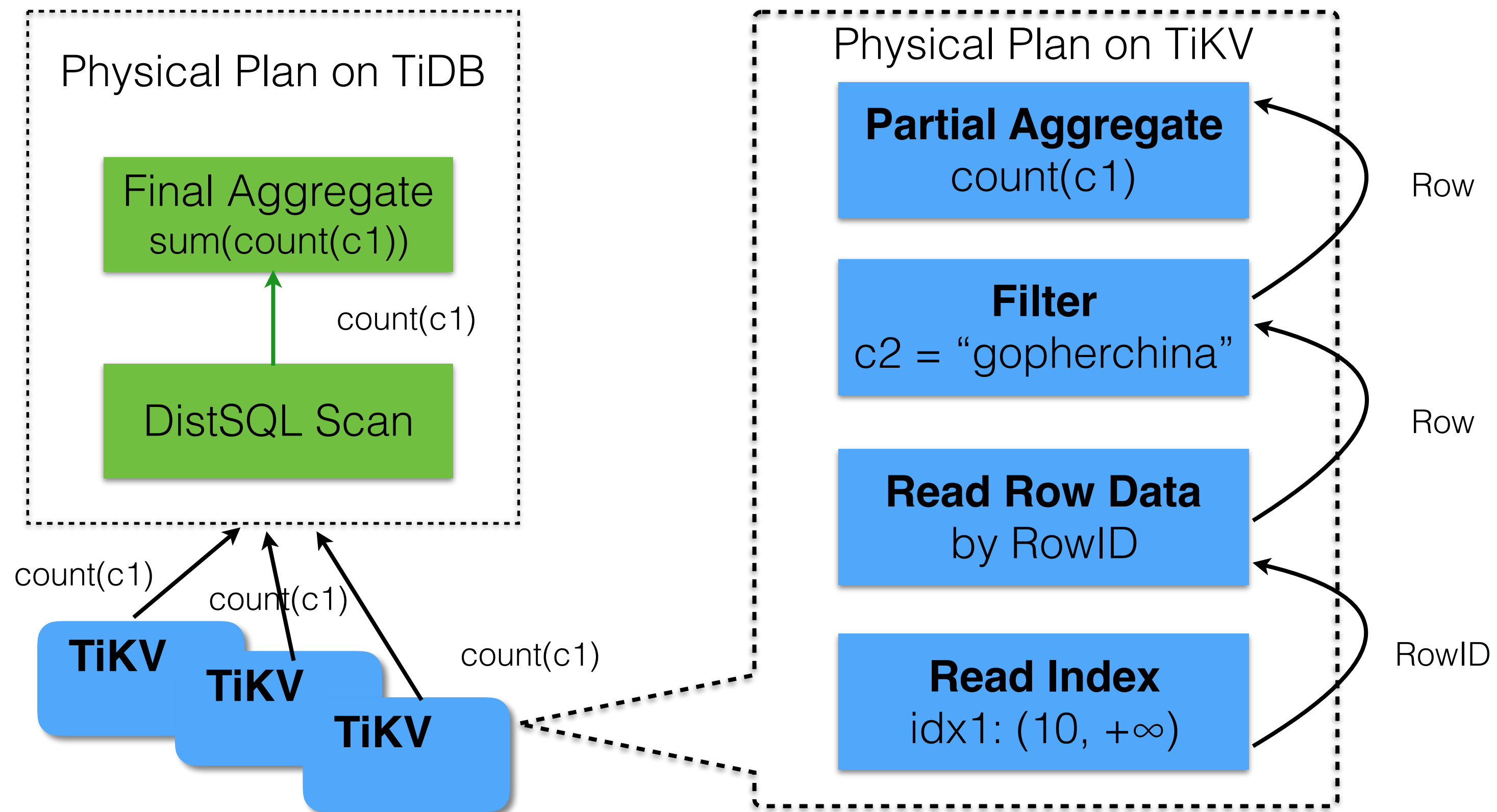
Example - Logical Plan



Example - Physical Plan



Example - Distributed Physical Plan



Challenges in Building a Distributed Database

- A very complex distributed system
- A lot of RPC work
- High performance
- Tons of data
- Huge amount of OLTP queries
- Very complex OLAP queries



Why Go?

- Productivity
- Concurrency
- Great for network programming
- GC
- Standard libraries && Tools
- Good performance
- Quick improvement

Go in TiDB

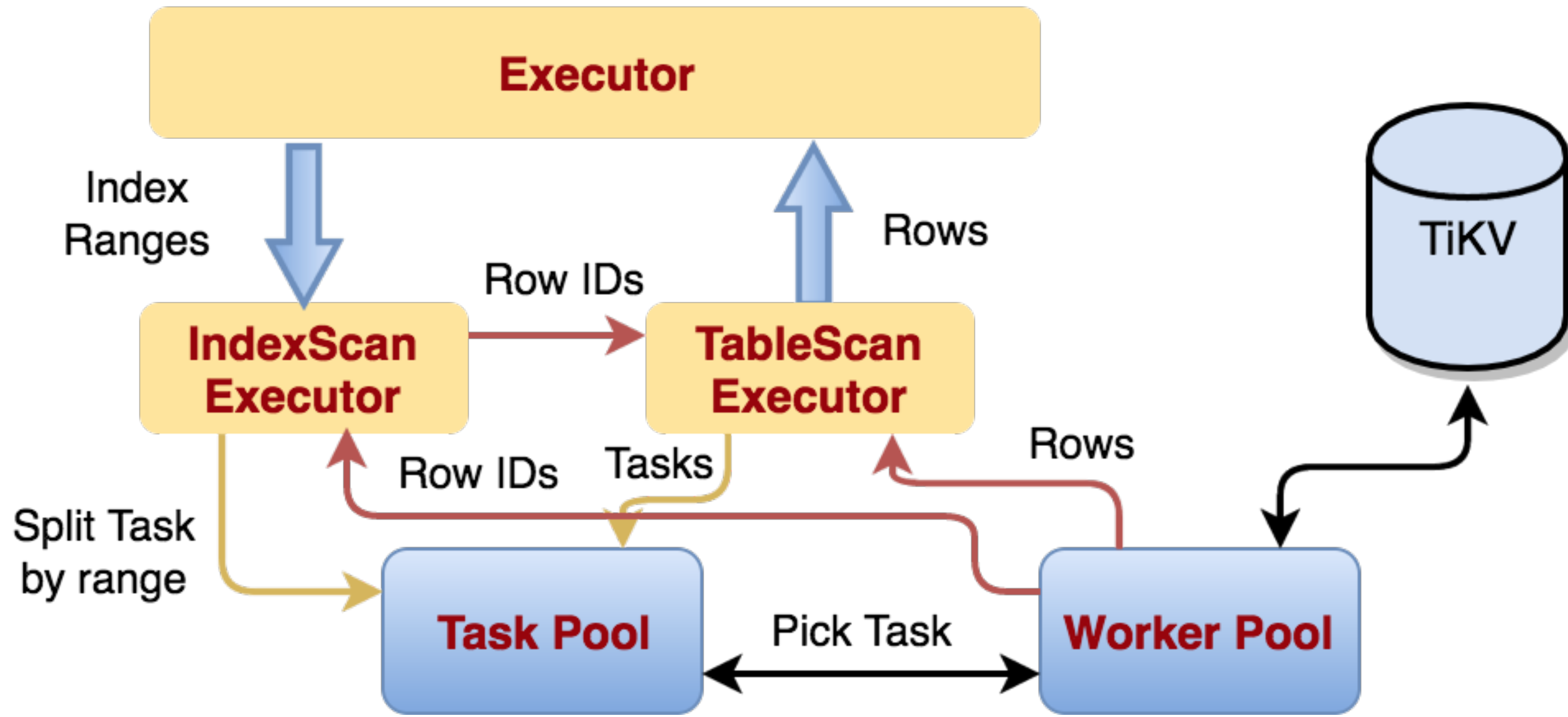
- More than 100k lines of Go code and 94 contributors.

Language	files	blank	comment	code
Go	457	13924	14782	112255
yacc	1	363	227	6021
YAML	3	4	0	191
make	1	33	2	102
Bourne Shell	4	12	5	53
SUM:	466	14336	15016	118622

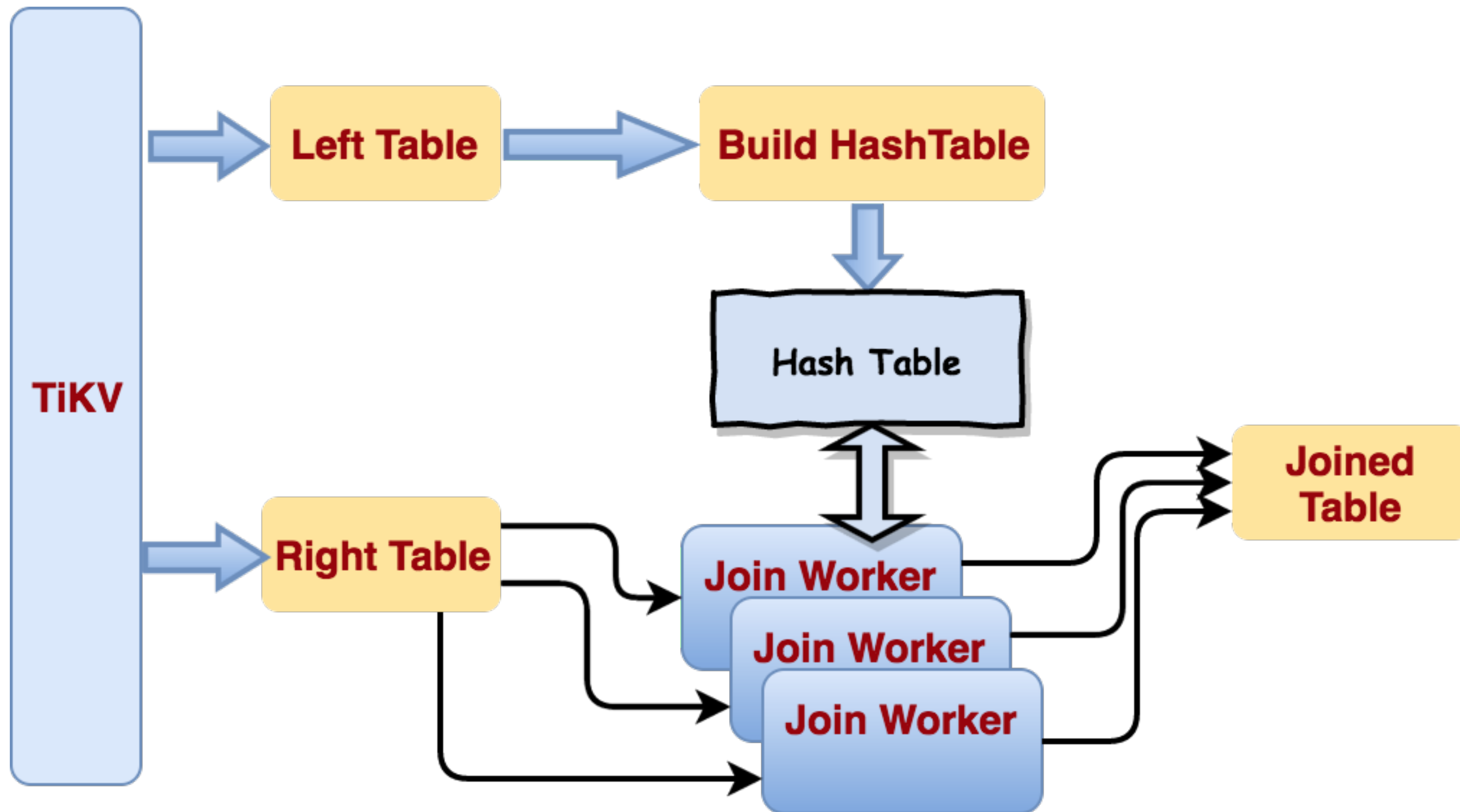
Goroutine

- Starting a goroutine is easy and cheap.
- Goroutines come with built-in primitives to communicate safely between themselves (channels).
- Very easy to write concurrent programs.
- Make full use of multi-core processors.

Parallel Data Scan Operator



Parallel HashJoin Operator



Goroutine Leak

- Write to a chan with no reader.
- Read from a chan with no writer.
- How to resolve?
 - Block profile
 - Timeout
 - Context

Goroutine Leak Test

- A useful tool from net/http

github.com/pingcap/tidb/util/testleak

Usage:

```
func (s testMiscSuite) TestRunWithRetry(c *C) {  
    defer testleak.AfterTest(c)()  
    .....  
}
```

Memory && GC

- Query may touch a huge number of data.

-

So we should take care of the memory.

-

- Degrade the performance of SQL.

- OOM sucks!

Reduce the Number of Allocation

- Get enough memory in one allocation operation

```
a := []int{1, 2, 3, 4, 5}
b := []int{}
// a much better way:
// b := make([]int, 0, len(a))
for _, i := range a {
    b = append(b, i)
}
```

Reuse Object

- Introduce a cache in goyacc
- Share a stack for all queries in one session

Before:

insert test data finished ... elapse : 34.563343s tps:28932.386438

After:

insert test data finished ... elapse : 28.859153s tps:34651.051825

sync.Pool

- Thread safe
- Reuse objects to relieve pressure on the GC.
- Case study: bytes pool

Test on a query reads 1 million rows

Before

runtime.MemStats

Alloc = 1985803616

TotalAlloc = 4365147096

After

runtime.MemStats

Alloc = 1303384504

TotalAlloc = 1676585832

gogo/protobuf

- Fast marshalling and unmarshalling
- Fields without pointers cause less time in the garbage collector.
- Used by etcd/docker/k8s/tidb
- [Benchmark](#)

Monitor the Memory Usage

- Monitor the memory usage of the server is easy
 - `runtime.MemProfile()`
- Monitor the memory usage of a session is hard
 - Account for large memory allocation
 - Account for memory consuming operators

New in Go1.8

- Better GC
 - Pauses 50μs
- `sort.Slice`

```
Benchmark100Slice-4      2000000      630 ns/op
Benchmark100Sort-4      300000      4784 ns/op
Benchmark1000Slice-4    300000      3988 ns/op
Benchmark1000Sort-4     10000     108745 ns/op
Benchmark10000Slice-4   1000     1228942 ns/op
Benchmark10000Sort-4    1000     1440223 ns/op
Benchmark100000Slice-4  100     14483929 ns/op
Benchmark100000Sort-4  100     17661242 ns/op
```

Next Step of TiDB

- Switch to gRPC
- Document support
- Statistic & CBO (Cost-based Optimizer)
- Better optimizer
- Faster and more general execution engine
- TiDB + Spark = TiSpark

Thanks!

We are hiring!
shenli@pingcap.com

May all the data find their own place,
May all the SQL queries speed up!
Come, and join us,
To make the best NewSQL database!



TiDB 交流群 - Gopher
China 2017

