CS 839: Topics in Database Management Systems

Lecture 17: Auto-Scaling

Xiangyao Yu

11/1/2023

Logistics

Guest lecture next Monday (no paper review)

Title: Build an open source, high performance, cloud native time series database

by **Jeff Tao** (founder and CEO of TDengine)

Group Discussion from Lecture #16

Imagine you are building a cloud OS following the architecture of a cloud database + some cloud runtime (e.g., Lambda function or Kubernetes). Please consider the following questions:

- What can be a good initial killer app for such a system?
- Are there key missing technologies before we can build such a system? If so, what are they?
- Which existing systems/services would you use as building blocks for your initial prototype?

Group Discussion from Lecture #16

What can be a good initial killer app for such a system?

- Social media texting application, like Facebook messenger. Database is geo-distributed based on location of users. Treat each connected user as running process.
- Unified backup platform for multiple businesses. Handles recovery and provides a rigorous platform for configuration.
- Stock trading platform on a worldwide scale
- OLAP applications implemented using MapReduce or Spark. Job scheduling and IPC can all be done through the OS.
- Special hardware devices, shared programming environment like Google docs, Spark/Hadoop

Group Discussion from Lecture #16

Key missing technologies?

- IPC should be faster and locality-aware
- Kernel level implementation and concrete OS and DB APIs.
- Multi-partition transactions

Existing systems/services to build the prototype?

- Kubernetes as the runtime
- DynamoDB + Lambda + S3 + API Gateway

Discussion Questions

Auto-scaling has two aspects: (1) mechanism (e.g., how to add a new process without affecting correctness) and (2) policy (e.g., when should a new process be added or removed). Pick one paper you have read; how does it contribute to both aspects?

What do you see as the biggest challenge to enable an auto-scaling database? You can consider either mechanism or policy in either OLTP or OLAP setting.

Submit by **11:59pm CT on Thursday** (11/2). Title starts with "[Discussion L17]"