

Seminar: Next-Generation Data Systems

Session 2: Q&A and Doing a Presentation

Michalis Georgoulakis

Garching, October 21st, 2026



Today

Recap and questions on the papers

- Preferences are due Sunday, 25.10

Doing a presentation

- What makes a good talk
- Content and slides, with examples
- Delivery and preparation

Next week (28.10): topics assigned, instructions, academic writing

The Papers at a Glance

Modern Hardware (4)

GPUs

CoddSpeed · Sirius

Memory and CPUs

Pond · CHARM

Transactional Systems (6)

Lake storage

Delta Lake · Interoperable ACID

Disaggregation

Amazon Aurora · Socrates

Geo-distribution

Aurora DSQL · SunStorm

Analytical Systems (7)

Object storage

AnyBlob · Snowflake pruning · NVMe spilling

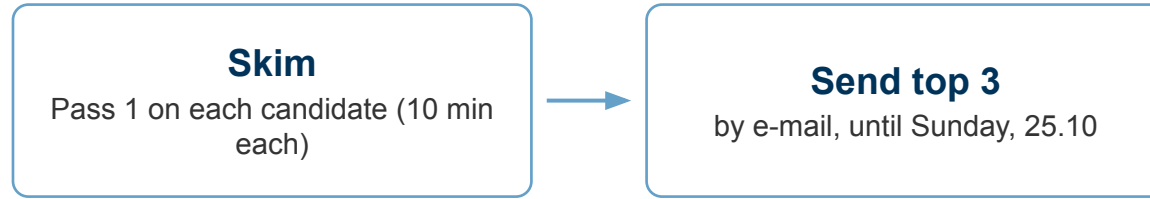
Compute pushdown

FlexPushdownDB · LiquidCache

Elastic and serverless

POLARIS · Skyrise

Sending Your Preferences



Choosing:

- Pick a problem that interests you
- You will spend weeks with it: talk, discussion, report

Questions on the Topics?



Why Give Talks in Science?

Paper

- structured, asynchronous
- understanding it takes effort (> 1 h)
- read alone, questions stay open

Talk

- structured, synchronous
- replaces the first end-to-end read
- everyone is in the room, including the expert: you

After your talk, everyone knows the paper well enough to discuss it.

“A research talk gives you access to the world’s most priceless commodity: the time and attention of other people. Don’t waste it!”

All You Need for a Good Talk

You

your voice carries the talk

Slides

enrich the talk with visuals and a simple structure

Board and demos

do not be afraid to draw or show something live

The slides support you, but do not replace you.

Don't Lose the Audience

Once you have lost the audience, it is hard or impossible to win them back.

You lose them when they:

1. **can't follow:** a step was skipped, or it went too fast
2. **get bored:** no clear problem, no story
3. **stop listening:** everything you say is already on the slide

Check every slide: would someone who blinked still follow?

Know Your Audience

Adapt your talk to who is in the room:

- What background knowledge do they have?
- What are they interested in?
- What do you want from them (questions, a discussion, feedback)?

In this seminar

- Fellow students: know databases, have read the paper once
- Supervisors: know the paper in depth
- Goal: a good discussion after your talk

Pitch the talk at your fellow students, not at the supervisors.

Presenting Someone Else's Paper

Explain it

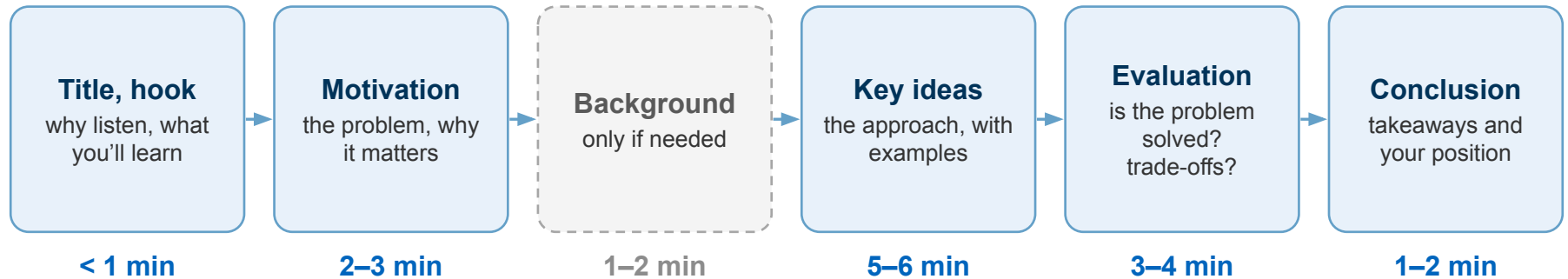
- Not section by section: the audience can read the paper
- The key idea, with your own examples and figures
- Only the experiments that support or break the claims

Assess it

- Take a position: what convinces you, what you doubt
- Criticize politely, with arguments: the authors are experts
- You don't have to defend the paper

Your position is what starts the discussion.

Structure of a Talk



Suggested split of 15 minutes. Optional before the conclusion: open questions, future work.

Most of your time goes to the key ideas.

The First Two Minutes

Hook: why should I listen?

*“Mirrored MySQL needs 7.4 I/Os per transaction.
Aurora needs less than one.”*

Promise: what will I learn?

*“In 15 minutes, you will know how Aurora gets there,
and what it pays for it.”*

You have about two minutes before attention drifts.

The talk is an advertisement for the paper, not the paper itself.

Simon Peyton Jones, How to Give a Great Research Talk · Patrick Winston, How to Speak (MIT)

Do You Need an Outline Slide?

Outline

- Motivation
- Background
- Design
- Evaluation
- Conclusion

Fits every talk, so it tells the audience nothing

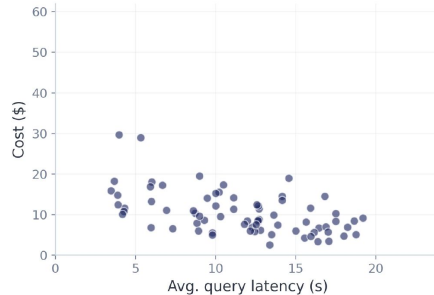
Instead

- Use the first minute for the hook and the promise
- Say the plan in one sentence: “first the problem, then Aurora’s one idea, then whether it pays off”
- Signpost while you go: “now the second part”

In 15 minutes, an agenda slide costs time and buys nothing.

Example: Motivation and Problem

One workload, many outcomes



6× latency
4× cost

from configuration alone

2

Problem statement



How can we enable **secure** and **verifiable** GDPR compliance for **key-value stores** on **untrusted** cloud infrastructure, while preserving **performance** and **compatibility**?

6

One plot, one number to remember: 6× latency, 4× cost

The research question in one sentence

From the CloudGlide talk, VLDB 2026 (left) and the GDPRuler talk, CCS 2026 (right)

Example: Show Where You Are

Challenges for GDPR-compliant KVS 

 #1 Security Protect data, metadata and audit logs	 #2 Policy enforcement Enforcing legal rules requires invasive engine code changes	 #3 Performance GDPR queries, metadata explosion and audit logging degrade performance
--	--	--

Credits: Flatiron.com 9

Challenges for GDPR-compliant KVS 

 #1 Security Protect data, metadata and audit logs	 #2 Policy enforcement Enforcing legal rules requires invasive engine code changes	 #3 Performance GDPR queries, metadata explosion and audit logging degrade performance
--	--	--

Credits: Flatiron.com 12

Challenges for GDPR-compliant KVS 

 #1 Security Protect data, metadata and audit logs	 #2 Policy enforcement Enforcing legal rules requires invasive engine code changes	 #3 Performance GDPR queries, metadata explosion and audit logging degrade performance
--	--	--

Credits: Flatiron.com 15

The overview of the three challenges comes back before each part of the design, with the current one highlighted.

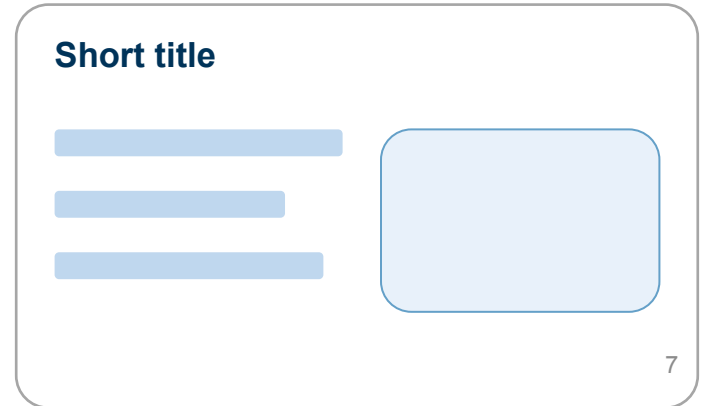
Come back to your overview, and highlight where you are.

From the GDPRuler talk, CCS 2026

Style Guide for Slides

Based on what your audience and examiners expect

- **Simple look** is a good look
- **Classic colors:** dark text on a light background
- **No flashy animations** or slide transitions
- **Readable** from the last row: large fonts, few words
- **Slide numbers**, so questions can say “back to slide 7”



How to Avoid Death by PowerPoint

- 1 **One message per slide**: two messages? Make two slides
- 2 **Don't write what you say**: the audience cannot read and listen at the same time
- 3 **Size**: the most important thing is the biggest
- 4 **Contrast**: highlight what you talk about, dim the rest
- 5 **At most six objects per slide**: beyond that, the slide competes with you

David JP Phillips: How to Avoid Death by PowerPoint, TEDxStockholmSalon

Our options

Measure



Live trials

How does this exact configuration actually perform?

[Cloud Analytics Benchmark, PVLDB 2023]

	Measure
Works before you build it	✗
Explains why	—
Matches Reality	Ground truth

Our options

Measure



Live trials

How does this exact configuration actually perform?

[Cloud Analytics Benchmark, PVLDB 2023]

Predict



ML performance predictors / autotuners

How do I optimize the system I am already running?

[Automatic Database Management System Tuning Through Large-scale Machine Learning, SIGMOD 2017]

	Measure	Predict
Works before you build it	✗	~
Explains why	—	✗
Matches Reality	Ground truth	In-distribution

Our options

Measure



Live trials

How does this exact configuration actually perform?
[Cloud Analytics Benchmark, PVLDB 2023]

Predict



ML performance predictors / autotuners

How do I optimize the system I am already running?
[Automatic Database Management System Tuning Through Large-scale Machine Learning, SIGMOD 2017]

Model



White-box cloud models

What if I change architecture, caching, or policy?
[Towards Cost-Optimal Query Processing in the Cloud, PVLDB 2021]

	Measure	Predict
Works before you build it	✗	~
Explains why	—	✗
Matches Reality	Ground truth	In-distribution

Model

✓

✓

?

A Talk Is Not a Paper in Slide Format

Relaxed format requirements:

	In a paper	In a talk
Citations	on every claim	only for borrowed figures
Figures and tables	numbered, with captions	you explain them
Text	full sentences	keywords

Reduce Text

- **Don't write full sentences!**
- Avoid bullet points that read like sentences
- Avoid bullet points altogether when a figure can show it

Next: the same slide, twice.

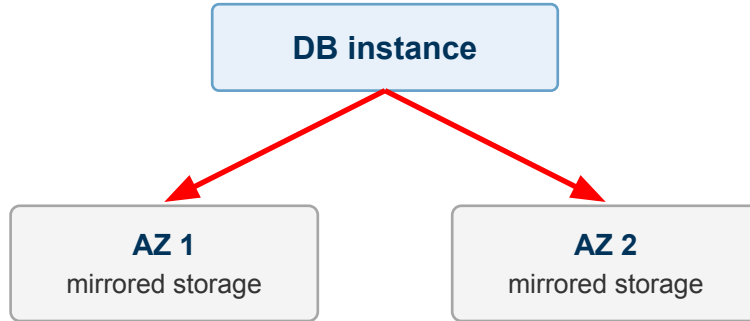
Example: Too Much Text

- Traditional databases write data pages, the redo log, the binlog and a double-write buffer to storage; with storage mirrored across availability zones (AZs), each write turns into many synchronous network I/Os.
- Aurora's key idea is to offload redo processing to a multi-tenant, scale-out storage service, so that the database instance only sends redo log records over the network ("the log is the database").
- Storage nodes apply the log records in the background to materialize pages and keep 6 copies of the data across 3 AZs, with a write quorum of 4/6 and a read quorum of 3/6.
- In the SysBench write-only evaluation, Aurora processes 35× more transactions than mirrored MySQL, while issuing 7.7× fewer I/Os per transaction.

✗ The audience reads instead of listening.

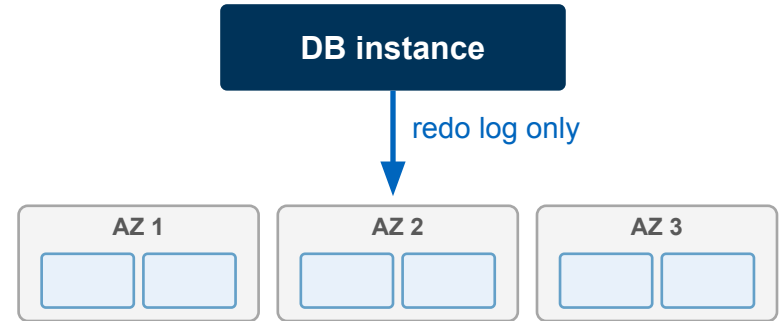
Example: Text Reduced

Mirrored MySQL



pages, redo log, binlog, ...

Aurora



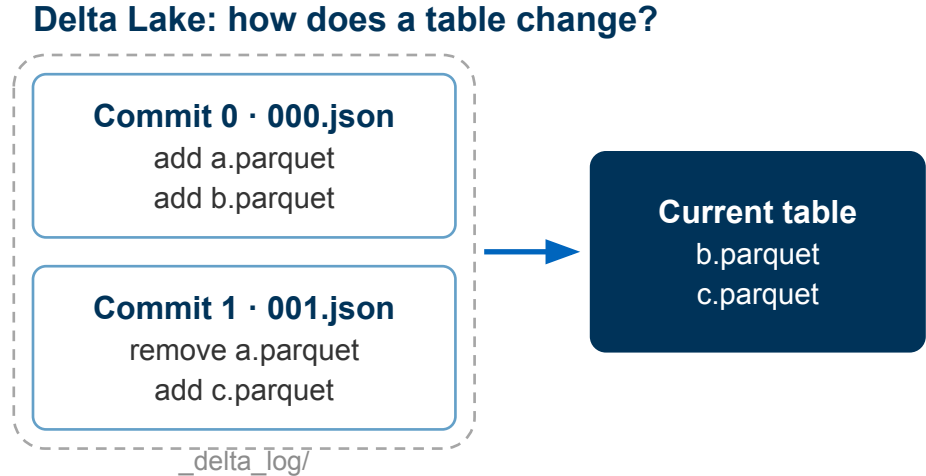
6 copies, write quorum 4/6

Gain: 35× transactions **7.7×** fewer I/Os per transaction

Examples Are Gold

- (Visual) examples are gold
- They build intuition and show where the problems are
- Keep them small and simple

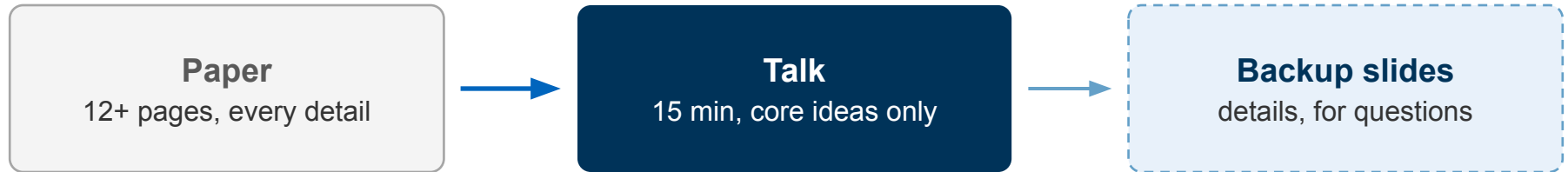
A table is a set of data files plus a log of commits.
Replaying the log gives the current files: a and b are added, then a is swapped for c.



Two log entries explain more than a paragraph about “ACID over object storage”.

Reduce Content and Slides

- Reduce the talk to the core ideas and insights
- You will need to cut content: that is normal
- Keep cut slides as **backup slides** after your final slide; they help in the discussion



Backup Slides

- Put them after your final slide
- One question per slide, and name it in the title
- Good candidates: details you cut, extra experiments, the setup, limitations

Prepare backups for the questions you fear most.

Backup — Experimental setup



Backup — Why not approach Y?



Backup — Effect of parameter X



Backup — Where does it break?



Omit Technical Details

- Avoid formulas
- Avoid detailed algorithms and pseudocode
- Show the intuition instead: an example, a figure

If you need a formula, explain what each part means.

Schöning's Probabilistic Algorithm
Mathematical Analysis of Correctness Probability

Generalized Binomial series

$$B_2(z) = \sum_{i=0}^{\infty} \binom{2i+1}{i} \frac{z^i}{2i+1} = \frac{1 - \sqrt{1-4z}}{2z}$$

$$B_2(z)^r = \sum_{i=0}^{\infty} \binom{2i+r}{i} \frac{r}{2i+r} z^i$$

$$\Rightarrow \Pr(N_j < \infty) = q^j \sum_{i=0}^{\infty} \binom{2i+j}{i} \frac{j}{2i+j} (q(1-q))^i = q^j (B_2(q(1-q)))^j$$

$$= q^j \left(\frac{1 - \sqrt{1-4(q(1-q))}}{2(q(1-q))} \right)^j = q^j \left(\frac{1 - \sqrt{1-4q+4q^2}}{2q(1-q)} \right)^j$$

$$= q^j \left(\frac{1 - (1-2q)}{2q(1-q)} \right)^j = q^j \left(\frac{1}{1-q} \right)^j$$

A real slide: what does the audience take from it?

Figures and Plots from the Paper

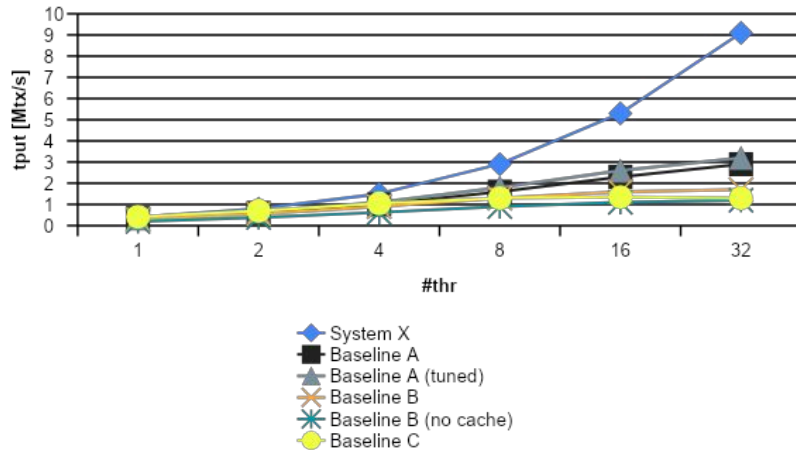
The paper's plots were made for readers, not for a room:

- **Redraw or simplify:** keep only the lines that matter
- **Enlarge** axes, labels and legends for the last row
- **Explain the axes first**, then the result
- **Highlight the takeaway:** color, an arrow, or one short note
- **Say how to read it:** “higher is better”
- **Color-blind safe:** no red–green pairs or rainbow maps; add labels or markers

One plot, one message.

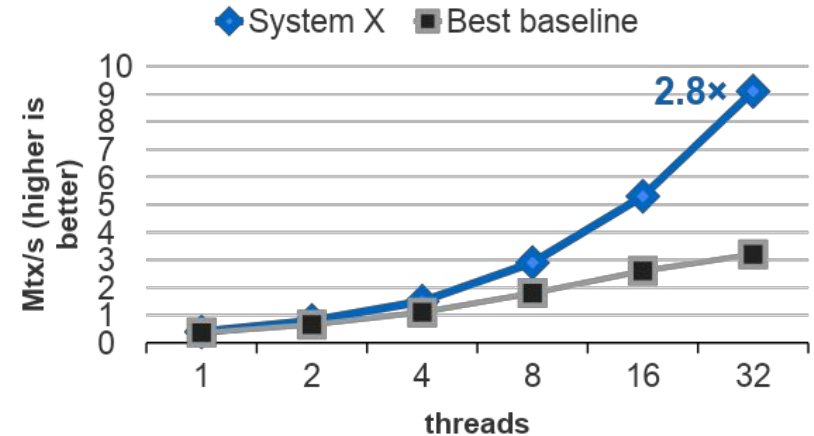
Example: One Plot, One Message

As in the paper



Made-up data, for illustration only

In your talk



Example: Evaluation, Questions First

Evaluation

Q1 Is it faster than the baseline?

Q2 Does it scale with more nodes?

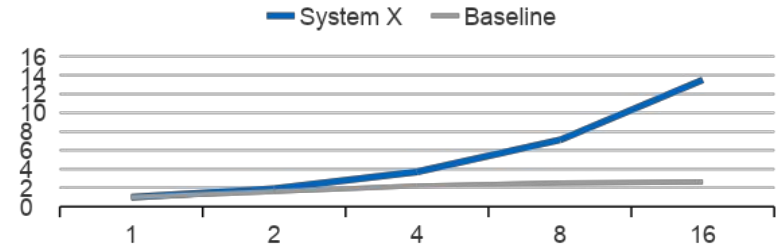
Q3 What does it cost?

Q4 How does it handle skewed data? → in the paper

Open with the questions; gray out the ones you skip

Made-up system and data, for illustration only

Q2: Does it scale?



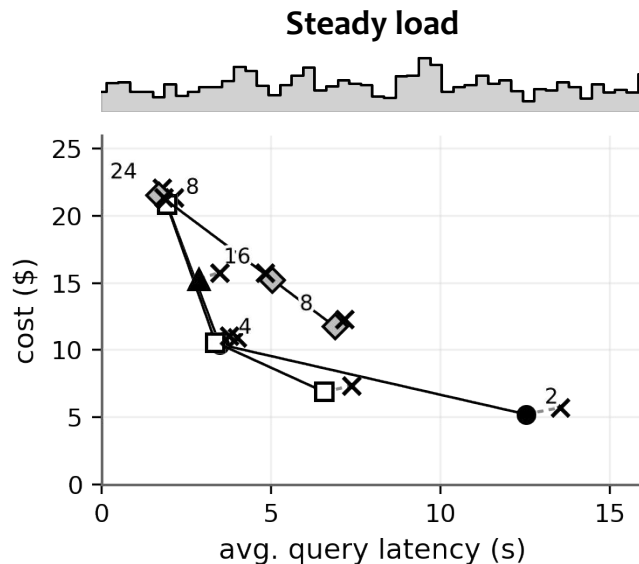
Near-linear to 16 nodes; the baseline flattens after 4

Then one plot per question, with the answer below it

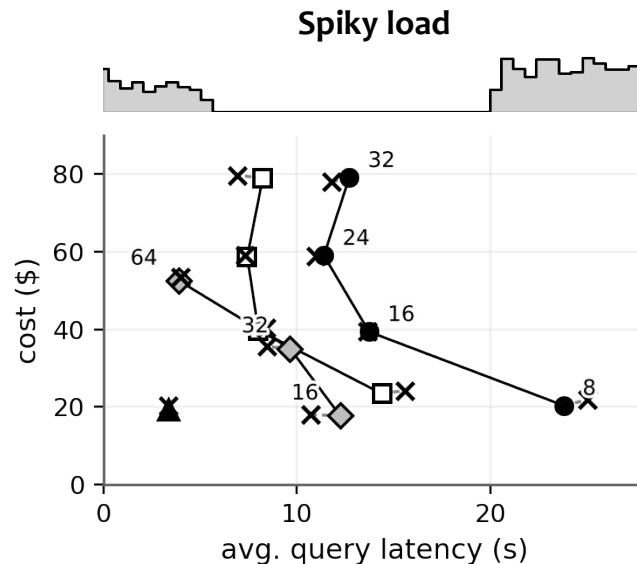
Which architecture, for which workload

- Same workload, two arrival patterns; lower-left is better

● DWaaS □ Autoscaling DWaaS ◇ Elastic pool ▲ QaaS × measured



Fixed cluster: 1.7–2× cheaper



Serverless: 1.5–6× lower latency

No architecture dominates — the arrival pattern decides, not the engine.

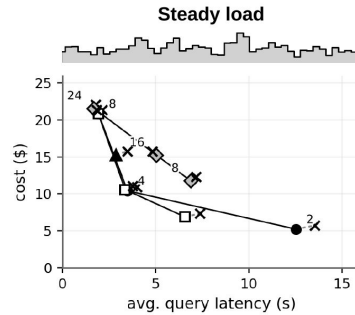
Example: One Takeaway per Plot

Which architecture, for which workload

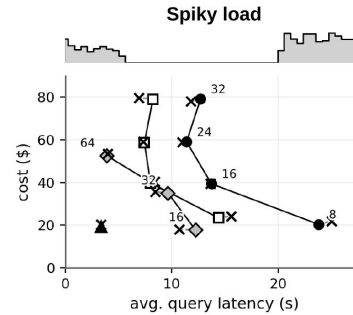


- Same workload, two arrival patterns; lower-left is better

● DWaaS □ Autoscaling DWaaS ◇ Elastic pool ▲ QaaS × measured



Fixed cluster: 1.7–2× cheaper



Serverless: 1.5–6× lower latency

No architecture dominates — the arrival pattern decides, not the engine.

Michail Georgoulakis (TUM) | CloudGlide: Deconstructing the Landscape of Cloud-Based Analytics | VLDB 2026

14

1

How to read it

“lower-left is better”, said before the result

2

Numbers on the plot

1.7–2× cheaper, 1.5–6× lower latency, right where they apply

3

A one-sentence takeaway

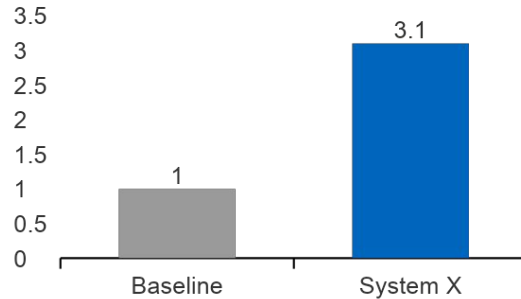
“No architecture dominates — the arrival pattern decides”

From the CloudGlide talk, VLDB 2026

Example: The Same Pattern on Every Plot

Q1: Is it faster?

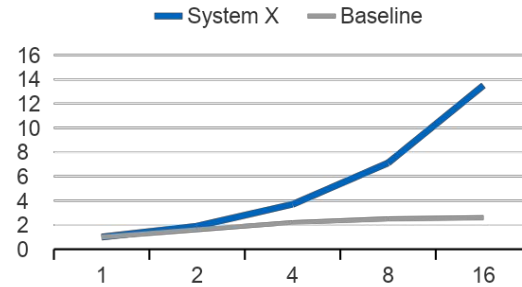
higher is better



3.1× the baseline's throughput

Q2: Does it scale?

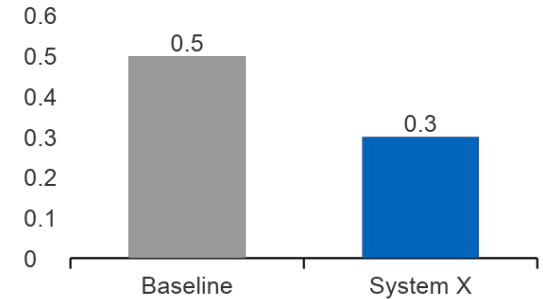
higher is better



Near-linear to 16 nodes

Q3: What does it cost?

lower is better



40% cheaper per query

Same layout every time: the question, the plot, the answer.

Made-up data, for illustration only

Animations

- Keep the audience focused on you
- Reveal parts only once they become relevant
- Also works for figures, as in the Aurora example
- Keep effects simple: “appear” is enough

Check your PDF backup: animated slides show only their final state.

Presentation Style

- **Speak freely**
 - Practice and prepare the talk enough
 - A transcript can help you prepare; do not read it out
- **Look at the audience**, not at the screen
- **Show that you find it exciting**
- **Avoid distracting habits**: pacing, fidgeting, filler words

Rehearse aloud with a timer: 15 minutes is a hard limit.

Repeat and Signpost

- 1 **Say the key idea three times**: in the hook, when you explain it, in the conclusion
- 2 **Signpost out loud**: “Now the second part”, “This is the key slide”
- 3 **Point with the slide**: arrows, highlights, animations, rather than a laser pointer

Repetition is not wasted time: some listeners missed it the first time.

Patrick Winston, How to Speak (MIT)

Dealing with Nerves

Memorize the first minute

not the whole talk: that sounds like reading

Arrive early

stand where you will present, test the clicker

Slow down

breathe between sentences; pauses feel longer to you than to them

Nerves peak in the minute before you start, and then go down.

The audience notices them much less than you do.

Interaction with the Audience

- Decide whether you take questions during the talk
- Is everyone familiar with what you assume? Ask
- Humor is welcome
- Watch faces during the talk:
 - Who is still following?
 - Who is getting lost?

In this seminar, most questions come in the moderated discussion after your talk.

Handling Questions

- 1 **Repeat the question:** so the whole room heard it, and you gain a few seconds
- 2 **Clarify if unclear:** “Do you mean ... ?” before answering the wrong question
- 3 **Answer briefly:** then check: “Does that answer your question?”
- 4 **Don't know? Say so:** “I don't know, but I would check ...”, never guess
- 5 **Use your backup slides:** this is what they are for

A question is interest, not an attack.

End on Your Takeaways, Not on “Questions?”

Questions?



Stays up for the whole discussion, and shows nothing

Aurora: Takeaways

- Only the redo log crosses the network
- 35× the throughput of mirrored MySQL

My position

- Convincing: far fewer I/Os per transaction
- Unclear: what do six copies cost?

Keeps your message and position in view

Your last slide is on screen longer than any other.

Preparation

PDF backup

works on any machine

Room check

02.09.014: laptop, adapter, projector

Clicker?

bring one or ask us

Plan B for time

know which slides to skip if you run late

Live demo?

bring a recording as a backup

Live demos work in every rehearsal, and fail exactly once.

Further References

- **Markus Püschel**: How to Give Good Technical Presentations
- **Simon Peyton Jones**: How to Give a Great Research Talk
- **David JP Phillips**: How to Avoid Death by PowerPoint (TEDxStockholmSalon talk)
- **Patrick Winston**: How to Speak (MIT lecture)
- **Ashwin Ram**: How to Present a Paper

All are short and available online.

Takeaways

- 1 **Tell one story:** hook, key idea, evidence, your position
- 2 **Less on the slide:** one message per slide; figures and examples over text
- 3 **One plot, one answer:** say how to read it, then what it shows
- 4 **Rehearse with a timer:** 15 minutes is a hard limit
- 5 **End on your takeaways:** they stay up during the whole discussion