Lightning Talks
Platform Lab Students
Stanford University
Lightning Talks

1. TorcDB: Some Updates - Jonathan Ellithorpe
2. NanoLog: A Nanosecond Scale Logging System - Stephen Yang
3. λ-NIC: Interactive Serverless Compute on SmartNICs
4. A Managed and Model-less Inference Serving System - Franky Romero, Qian Li
5. Arachne: Core Aware Thread Management - Henry Qin
7. CPUJailing: Dynamic and Hyperthread-aware Isolation for Shared Clusters – Kostis Kaffes
8. Millisort: a distributed sorting for millisecond-scale time budget – Seo Jin Park, Yilong Li, Collin Lee
9. Smart Harvesting of Spare CPU Cores in Cloud Servers – Yawen Wang
10. DejaVu: Enhancing Videoconferencing with Prior Knowledge – Pan Hu
11. SPN Chatbot: Handling Compound and Domain-Specific Queries
12. Tackling Frontier in Pull-based Graph Processing on CGRA - Timothy Chong
13. Multiagent Adaptive Scheduling, Planning and Control – Kyle Brown, Oriana Peltzer
14. Programmable Packet Scheduling - Stephen Ibanez
15. Task Assignment for Flexible Manufacturing – Patrick Washington
16. Implementing Linearizability at Large-Scale and Low-Latency – Seo Jin Park, Collin Lee
17. ...
TorcDB: Some Updates

Jonathan Ellithorpe

- TorcDB originally designed to provide low latency for simple update and read queries
  - Comment on things
  - Get friend’s posts
- Now looking at “complex” queries
  - Trending tags in your network
  - Ranking friends in terms of similarity
Lightning Talks

1. TorcDB: Some Updates - Jonathan Ellithorpe
2. **NanoLog: A Nanosecond Scale Logging System** - Stephen Yang
3. λ-NIC: Interactive Serverless Compute on SmartNICs
4. A Managed and Model-less Inference Serving System - Franky Romero, Qian Li
5. Arachne: Core Aware Thread Management - Henry Qin
7. CPUJailing: Dynamic and Hyperthread-aware Isolation for Shared Clusters – Kostis Kaffes
8. Millisort: a distributed sorting for millisecond-scale time budget – Seo Jin Park, Yilong Li, Collin Lee
9. Smart Harvesting of Spare CPU Cores in Cloud Servers – Yawen Wang
10. DejaVu: Enhancing Videoconferencing with Prior Knowledge – Pan Hu
11. SPN Chatbot: Handling Compound and Domain-Specific Queries
12. Tackling Frontier in Pull-based Graph Processing on CGRA - Timothy Chong
13. Multiagent Adaptive Scheduling, Planning and Control – Kyle Brown, Oriana Peltzer
14. Programmable Packet Scheduling - Stephen Ibanez
15. Task Assignment for Flexible Manufacturing – Patrick Washington
16. Implementing Linearizability at Large-Scale and Low-Latency – Seo Jin Park, Collin Lee
17. ...
NanoLog: A Nanosecond Scale Logging System
Stephen Yang

- 10-100x faster than existing solutions such as Log4j2 or spdlog
  - Achieves a throughput of up to 80M log messages/second at a 7-18ns median latency
  - Maintains a simple printf-like API

- Key Concept: Shift work out of Runtime
  - Deduplicates static log info at compile time to save I/O
  - Only log dynamic information in binary format at Runtime
  - Defers message formatting until post execution

<table>
<thead>
<tr>
<th>Latency / System</th>
<th>NanoLog</th>
<th>spdlog</th>
<th>Log4j2</th>
<th>Boost</th>
<th>glog</th>
<th>ETW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median(ns)</td>
<td>7</td>
<td>214</td>
<td>174</td>
<td>1764</td>
<td>1198</td>
<td>161</td>
</tr>
<tr>
<td>99.9th (ns)</td>
<td>37</td>
<td>2546</td>
<td>3364</td>
<td>3772</td>
<td>5969</td>
<td>2967</td>
</tr>
</tbody>
</table>

GitHub: https://github.com/PlatformLab/NanoLog
Lightning Talks

1. TorcDB: Some Updates - Jonathan Ellithorpe
2. NanoLog: A Nanosecond Scale Logging System - Stephen Yang
3. $\lambda$-NIC: Interactive Serverless Compute on SmartNICs
4. A Managed and Model-less Inference Serving System - Franky Romero, Qian Li
5. Arachne: Core Aware Thread Management - Henry Qin
7. CPUJailing: Dynamic and Hyperthread-aware Isolation for Shared Clusters – Kostis Kaffes
8. Millisort: a distributed sorting for millisecond-scale time budget – Seo Jin Park, Yilong Li, Collin Lee
9. Smart Harvesting of Spare CPU Cores in Cloud Servers – Yawen Wang
10. DejaVu: Enhancing Videoconferencing with Prior Knowledge – Pan Hu
11. SPN Chatbot: Handling Compound and Domain-Specific Queries
12. Tackling Frontier in Pull-based Graph Processing on CGRA - Timothy Chong
13. Multiagent Adaptive Scheduling, Planning and Control – Kyle Brown, Oriana Peltzer
14. Programmable Packet Scheduling - Stephen Ibanez
15. Task Assignment for Flexible Manufacturing – Patrick Washington
16. Implementing Linearizability at Large-Scale and Low-Latency – Seo Jin Park, Collin Lee
17. ...
**λ-NIC: Interactive Serverless Compute on SmartNICs**

Sean Choi

- Key Concept: an open-source platform for running latency sensitive serverless compute on ASIC-based SmartNICs.

- Up to 880x improvements in latency and 736x improvements in throughput than existing serverless compute platforms
1. TorcDB: Some Updates - Jonathan Ellithorpe
2. NanoLog: A Nanosecond Scale Logging System - Stephen Yang
3. ƛ-NIC: Interactive Serverless Compute on SmartNICs
4. **A Managed and Model-less Inference Serving System** - Franky Romero, Qian Li
5. Arachne: Core Aware Thread Management - Henry Qin
7. CPUJailing: Dynamic and Hyperthread-aware Isolation for Shared Clusters – Kostis Kaffes
8. Millisort: a distributed sorting for millisecond-scale time budget – Seo Jin Park, Yilong Li, Collin Lee
9. Smart Harvesting of Spare CPU Cores in Cloud Servers – Yawen Wang
10. DejaVu: Enhancing Videoconferencing with Prior Knowledge – Pan Hu
11. SPN Chatbot: Handling Compound and Domain-Specific Queries
12. Tackling Frontier in Pull-based Graph Processing on CGRA - Timothy Chong
13. Multiagent Adaptive Scheduling, Planning and Control – Kyle Brown, Oriana Peltzer
14. Programmable Packet Scheduling - Stephen Ibanez
15. Task Assignment for Flexible Manufacturing – Patrick Washington
16. Implementing Linearizability at Large-Scale and Low-Latency – Seo Jin Park, Collin Lee
17. ...
A Managed and Model-less Inference Serving System

Franky Romero, Qian Li

Problem

😊 User: hard to configure and manage
😊 Provider: waste resources
😊 Designing a good inference serving system is non-trivial

What We Need

- **Managed**: Automatic model resource management and scaling
- **Model-less**: Select model variant to meet SLO-requirements
Lightning Talks

1. TorcDB: Some Updates - Jonathan Ellithorpe
2. NanoLog: A Nanosecond Scale Logging System - Stephen Yang
3. λ-NIC: Interactive Serverless Compute on SmartNICs
4. A Managed and Model-less Inference Serving System - Franky Romero, Qian Li
5. Arachne: Core Aware Thread Management - Henry Qin
7. CPUJailing: Dynamic and Hyperthread-aware Isolation for Shared Clusters – Kostis Kaffes
8. Millisort: a distributed sorting for millisecond-scale time budget – Seo Jin Park, Yilong Li, Collin Lee
9. Smart Harvesting of Spare CPU Cores in Cloud Servers – Yawen Wang
10. DejaVu: Enhancing Videoconferencing with Prior Knowledge – Pan Hu
11. SPN Chatbot: Handling Compound and Domain-Specific Queries
12. Tackling Frontier in Pull-based Graph Processing on CGRA - Timothy Chong
13. Multiagent Adaptive Scheduling, Planning and Control – Kyle Brown, Oriana Peltzer
14. Programmable Packet Scheduling - Stephen Ibanez
15. Task Assignment for Flexible Manufacturing – Patrick Washington
16. Implementing Linearizability at Large-Scale and Low-Latency – Seo Jin Park, Collin Lee
17. ...
Arachne: Core Aware Thread Management

Henry Qin, Qian Li, Jacqueline Speiser, Peter Kraft, John Ousterhout

[Diagram of Arachne system architecture and performance graphs showing latency and throughput under different conditions.]
Lightning Talks

1. TorcDB: Some Updates - Jonathan Ellithorpe
2. NanoLog: A Nanosecond Scale Logging System - Stephen Yang
3. Λ-NIC: Interactive Serverless Compute on SmartNICs
4. A Managed and Model-less Inference Serving System - Franky Romero, Qian Li
5. Arachne: Core Aware Thread Management - Henry Qin
7. CPUJailing: Dynamic and Hyperthread-aware Isolation for Shared Clusters – Kostis Kaffes
8. Millisort: a distributed sorting for millisecond-scale time budget – Seo Jin Park, Yilong Li, Collin Lee
9. Smart Harvesting of Spare CPU Cores in Cloud Servers – Yawen Wang
10. DejaVu: Enhancing Videoconferencing with Prior Knowledge – Pan Hu
11. SPN Chatbot: Handling Compound and Domain-Specific Queries
12. Tackling Frontier in Pull-based Graph Processing on CGRA - Timothy Chong
13. Multiagent Adaptive Scheduling, Planning and Control – Kyle Brown, Oriana Peltzer
14. Programmable Packet Scheduling - Stephen Ibanez
15. Task Assignment for Flexible Manufacturing – Patrick Washington
16. Implementing Linearizability at Large-Scale and Low-Latency – Seo Jin Park, Collin Lee
17. ...
DeepCut: Bandwidth and Task Aware Offloading From Edge Devices

Eyal Cidon, Jenya Pergament, Sandeep Chinchali
Lightning Talks

1. TorcDB: Some Updates - Jonathan Ellithorpe
2. NanoLog: A Nanosecond Scale Logging System - Stephen Yang
3. λ-NIC: Interactive Serverless Compute on SmartNICs
4. A Managed and Model-less Inference Serving System - Franky Romero, Qian Li
5. Arachne: Core Aware Thread Management - Henry Qin
7. CPUJailing: Dynamic and Hyperthread-aware Isolation for Shared Clusters – Kostis Kaffes
8. Millisort: a distributed sorting for millisecond-scale time budget – Seo Jin Park, Yilong Li, Collin Lee
9. Smart Harvesting of Spare CPU Cores in Cloud Servers – Yawen Wang
10. DejaVu: Enhancing Videoconferencing with Prior Knowledge – Pan Hu
11. SPN Chatbot: Handling Compound and Domain-Specific Queries
12. Tackling Frontier in Pull-based Graph Processing on CGRA - Timothy Chong
13. Multiagent Adaptive Scheduling, Planning and Control – Kyle Brown, Oriana Peltzer
14. Programmable Packet Scheduling - Stephen Ibanez
15. Task Assignment for Flexible Manufacturing – Patrick Washington
16. Implementing Linearizability at Large-Scale and Low-Latency – Seo Jin Park, Collin Lee
17. ...
CPUJailing: Dynamic and Hyperthread-aware Isolation for Shared Clusters

Kostis Kaffes

**CPUJailing:**
- Reduced CPU interference without knowledge of application specific-characteristics
- Works with any workload
- Integrated with Google’s Borg and showed significant performance gains.
Lightning Talks

1. TorcDB: Some Updates - Jonathan Ellithorpe
2. NanoLog: A Nanosecond Scale Logging System - Stephen Yang
3. ƛ-NIC: Interactive Serverless Compute on SmartNICs
4. A Managed and Model-less Inference Serving System - Franky Romero, Qian Li
5. Arachne: Core Aware Thread Management - Henry Qin
7. CPUJailing: Dynamic and Hyperthread-aware Isolation for Shared Clusters – Kostis Kaffes
8. Millisort: a distributed sorting for millisecond-scale time budget – Seo Jin Park, Yilong Li, Collin Lee
9. Smart Harvesting of Spare CPU Cores in Cloud Servers – Yawen Wang
10. DejaVu: Enhancing Videoconferencing with Prior Knowledge – Pan Hu
11. SPN Chatbot: Handling Compound and Domain-Specific Queries
12. Tackling Frontier in Pull-based Graph Processing on CGRA - Timothy Chong
13. Multiagent Adaptive Scheduling, Planning and Control – Kyle Brown, Oriana Peltzer
14. Programmable Packet Scheduling - Stephen Ibanez
15. Task Assignment for Flexible Manufacturing – Patrick Washington
16. Implementing Linearizability at Large-Scale and Low-Latency – Seo Jin Park, Collin Lee
17. ...

...
Millisort: Distributed Sorting for Millisecond-Scale Time Budget

Yilong Li

- How many records can you sort in 1 ms given unlimited number of nodes?
  - Tight end-to-end latency budget
  - Harness massive parallelism from thousands of machines using tiny tasks
  - Bursty communication with complex patterns

- Build granular computing infrastructure by need
  - MPI-like collective communication operations
  - RPC system atop Homa transport
Lightning Talks

1. TorcDB: Some Updates - Jonathan Ellithorpe
2. NanoLog: A Nanosecond Scale Logging System - Stephen Yang
3. λ-NIC: Interactive Serverless Compute on SmartNICs
4. A Managed and Model-less Inference Serving System - Franky Romero, Qian Li
5. Arachne: Core Aware Thread Management - Henry Qin
7. CPUJailing: Dynamic and Hyperthread-aware Isolation for Shared Clusters – Kostis Kaffes
8. Millisort: a distributed sorting for millisecond-scale time budget – Seo Jin Park, Yilong Li, Collin Lee
9. **Smart Harvesting of Spare CPU Cores in Cloud Servers** – Yawen Wang
10. DejaVu: Enhancing Videoconferencing with Prior Knowledge – Pan Hu
11. SPN Chatbot: Handling Compound and Domain-Specific Queries
12. Tackling Frontier in Pull-based Graph Processing on CGRA - Timothy Chong
13. Multiagent Adaptive Scheduling, Planning and Control – Kyle Brown, Oriana Peltzer
14. Programmable Packet Scheduling - Stephen Ibanez
15. Task Assignment for Flexible Manufacturing – Patrick Washington
16. Implementing Linearizability at Large-Scale and Low-Latency – Seo Jin Park, Collin Lee
17. ...
Smart Harvesting of Spare CPU Cores in Cloud Servers

Yawen Wang

On average, less than 50% of allocated CPU resources are used in a cluster from Google [1]

Harvest unused CPU cores from client VMs to run batch jobs to improve CPU utilization on VM host servers

Use online cost-sensitive multiclass classification algorithm to predict peak CPU utilization of client VM

1. Maximize harvested spare CPU resources while minimizing impact on co-located client VMs
2. Quickly adapts to different CPU usage patterns

[1] Reiss et al. Heterogeneity and dynamicity of clouds at scale: Google trace analysis. SOCC’12
Lightning Talks

1. TorcDB: Some Updates - Jonathan Ellithorpe
2. NanoLog: A Nanosecond Scale Logging System - Stephen Yang
3. λ-NIC: Interactive Serverless Compute on SmartNICs
4. A Managed and Model-less Inference Serving System - Franky Romero, Qian Li
5. Arachne: Core Aware Thread Management - Henry Qin
7. CPUJailing: Dynamic and Hyperthread-aware Isolation for Shared Clusters – Kostis Kaffes
8. Millisort: a distributed sorting for millisecond-scale time budget – Seo Jin Park, Yilong Li, Collin Lee
9. Smart Harvesting of Spare CPU Cores in Cloud Servers – Yawen Wang
10. DejaVu: Enhancing Videoconferencing with Prior Knowledge – Pan Hu
11. SPN Chatbot: Handling Compound and Domain-Specific Queries
12. Tackling Frontier in Pull-based Graph Processing on CGRA - Timothy Chong
13. Multiagent Adaptive Scheduling, Planning and Control – Kyle Brown, Oriana Peltzer
14. Programmable Packet Scheduling - Stephen Ibanez
15. Task Assignment for Flexible Manufacturing – Patrick Washington
16. Implementing Linearizability at Large-Scale and Low-Latency – Seo Jin Park, Collin Lee
17. ...
DejaVu: Enhancing Videoconferencing with Prior Knowledge

Pan Hu
Lightning Talks

1. TorcDB: Some Updates - Jonathan Ellithorpe
2. NanoLog: A Nanosecond Scale Logging System - Stephen Yang
3. A-NIC: Interactive Serverless Compute on SmartNICs
4. A Managed and Model-less Inference Serving System - Franky Romero, Qian Li
5. Arachne: Core Aware Thread Management - Henry Qin
7. CPUJailing: Dynamic and Hyperthread-aware Isolation for Shared Clusters – Kostis Kaffes
8. Millisort: a distributed sorting for millisecond-scale time budget – Seo Jin Park, Yilong Li, Collin Lee
9. Smart Harvesting of Spare CPU Cores in Cloud Servers – Yawen Wang
10. DejaVu: Enhancing Videoconferencing with Prior Knowledge – Pan Hu

11. **SPN Chatbot: Handling Compound and Domain-Specific Queries**
12. Tackling Frontier in Pull-based Graph Processing on CGRA - Timothy Chong
13. Multiagent Adaptive Scheduling, Planning and Control – Kyle Brown, Oriana Peltzer
14. Programmable Packet Scheduling - Stephen Ibanez
15. Task Assignment for Flexible Manufacturing – Patrick Washington
16. Implementing Linearizability at Large-Scale and Low-Latency – Seo Jin Park, Collin Lee
17. ...
SPN Chatbot: Handling Compound and Domain-Specific Queries

- **A dashboard that ...**
  - Shows you what is going on in your data center
  - In as much detail as you want

- **A chatbot that ...**
  - Understands your data center and can answer your questions
  - Talk to it or point at things from the dashboard and ask

![Diagram of User + Dashboard, Conversational Agent, and Database with SQL queries for handling data center utilizations.](image)
Lightning Talks

1. TorcDB: Some Updates - Jonathan Ellithorpe
2. NanoLog: A Nanosecond Scale Logging System - Stephen Yang
3. λ-NIC: Interactive Serverless Compute on SmartNICs
4. A Managed and Model-less Inference Serving System - Franky Romero, Qian Li
5. Arachne: Core Aware Thread Management - Henry Qin
7. CPUJailing: Dynamic and Hyperthread-aware Isolation for Shared Clusters – Kostis Kaffes
8. Millisort: a distributed sorting for millisecond-scale time budget – Seo Jin Park, Yilong Li, Collin Lee
9. Smart Harvesting of Spare CPU Cores in Cloud Servers – Yawen Wang
10. DejaVu: Enhancing Videoconferencing with Prior Knowledge – Pan Hu
11. SPN Chatbot: Handling Compound and Domain-Specific Queries
12. **Tackling Frontier in Pull-based Graph Processing on CGRA** - Timothy Chong
13. Multiagent Adaptive Scheduling, Planning and Control – Kyle Brown, Oriana Peltzer
14. Programmable Packet Scheduling - Stephen Ibanez
15. Task Assignment for Flexible Manufacturing – Patrick Washington
16. Implementing Linearizability at Large-Scale and Low-Latency – Seo Jin Park, Collin Lee
17. ...
Tackling the Frontier in pull-based Graph Processing on CGRA

Timothy Chong

- **Problem**: Pull-based graph processing is more performant than push-based one but cannot easily take advantage of **frontier optimization**.
- **Objective**: Develop a graph processing framework that exploits the **parallelism** of CGRA, Plasticine, and incorporates an **edge frontier prefetching engine** that allows **high-throughput parallel frontier processing**.
Lightning Talks

1. TorcDB: Some Updates - Jonathan Ellithorpe
2. NanoLog: A Nanosecond Scale Logging System - Stephen Yang
3. λ-NIC: Interactive Serverless Compute on SmartNICs
4. A Managed and Model-less Inference Serving System - Franky Romero, Qian Li
5. Arachne: Core Aware Thread Management - Henry Qin
7. CPUJailing: Dynamic and Hyperthread-aware Isolation for Shared Clusters – Kostis Kaffes
8. Millisort: a distributed sorting for millisecond-scale time budget – Seo Jin Park, Yilong Li, Collin Lee
9. Smart Harvesting of Spare CPU Cores in Cloud Servers – Yawen Wang
10. DejaVu: Enhancing Videoconferencing with Prior Knowledge – Pan Hu
11. SPN Chatbot: Handling Compound and Domain-Specific Queries
12. Tackling Frontier in Pull-based Graph Processing on CGRA - Timothy Chong
13. Multiagent Adaptive Scheduling, Planning and Control – Kyle Brown, Oriana Peltzer
14. Programmable Packet Scheduling - Stephen Ibanez
15. Task Assignment for Flexible Manufacturing – Patrick Washington
16. Implementing Linearizability at Large-Scale and Low-Latency – Seo Jin Park, Collin Lee
17. ...
Teams of agents operate in a dynamic and uncertain environment

Objective: optimally serve sequences of tasks by transporting objects

Challenges:
- Partial observability of the environment: balancing exploration vs. exploitation
- Simultaneous task assignment and vehicle routing
Lightning Talks

1. TorcDB: Some Updates - Jonathan Ellithorpe
2. NanoLog: A Nanosecond Scale Logging System - Stephen Yang
3. λ-NIC: Interactive Serverless Compute on SmartNICs
4. A Managed and Model-less Inference Serving System - Franky Romero, Qian Li
5. Arachne: Core Aware Thread Management - Henry Qin
7. CPUJailing: Dynamic and Hyperthread-aware Isolation for Shared Clusters – Kostis Kaffes
8. Millisort: a distributed sorting for millisecond-scale time budget – Seo Jin Park, Yilong Li, Collin Lee
9. Smart Harvesting of Spare CPU Cores in Cloud Servers – Yawen Wang
10. DejaVu: Enhancing Videoconferencing with Prior Knowledge – Pan Hu
11. SPN Chatbot: Handling Compound and Domain-Specific Queries
12. Tackling Frontier in Pull-based Graph Processing on CGRA - Timothy Chong
13. Multiagent Adaptive Scheduling, Planning and Control – Kyle Brown, Oriana Peltzer
14. **Programmable Packet Scheduling - Stephen Ibanez**
15. Task Assignment for Flexible Manufacturing – Patrick Washington
16. Implementing Linearizability at Large-Scale and Low-Latency – Seo Jin Park, Collin Lee
17. ...
Motivation:

- Be able to deploy new scheduling policies in production networks

Questions:

1. Can we find a programmable abstraction for scheduling policies?
2. Is our abstraction practical to implement in hardware?
Lightning Talks

1. TorcDB: Some Updates - Jonathan Ellithorpe
2. NanoLog: A Nanosecond Scale Logging System - Stephen Yang
3. λ-NIC: Interactive Serverless Compute on SmartNICs
4. A Managed and Model-less Inference Serving System - Franky Romero, Qian Li
5. Arachne: Core Aware Thread Management - Henry Qin
7. CPUJailing: Dynamic and Hyperthread-aware Isolation for Shared Clusters – Kostis Kaffes
8. Millisort: a distributed sorting for millisecond-scale time budget – Seo Jin Park, Yilong Li, Collin Lee
9. Smart Harvesting of Spare CPU Cores in Cloud Servers – Yawen Wang
10. DejaVu: Enhancing Videoconferencing with Prior Knowledge – Pan Hu
11. SPN Chatbot: Handling Compound and Domain-Specific Queries
12. Tackling Frontier in Pull-based Graph Processing on CGRA - Timothy Chong
13. Multiagent Adaptive Scheduling, Planning and Control – Kyle Brown, Oriana Peltzer
14. Programmable Packet Scheduling - Stephen Ibanez
15. Task Assignment for Flexible Manufacturing – Patrick Washington
16. Implementing Linearizability at Large-Scale and Low-Latency – Seo Jin Park, Collin Lee
17. ...
Task Assignment for Flexible Manufacturing

Patrick Washington

- Factory with several task stations
- Objects arrive for manufacturing
- Agents carry objects between stations
Lightning Talks

1. TorcDB: Some Updates - Jonathan Ellithorpe
2. NanoLog: A Nanosecond Scale Logging System - Stephen Yang
3. λ-NIC: Interactive Serverless Compute on SmartNICs
4. A Managed and Model-less Inference Serving System - Franky Romero, Qian Li
5. Arachne: Core Aware Thread Management - Henry Qin
7. CPUJailing: Dynamic and Hyperthread-aware Isolation for Shared Clusters – Kostis Kaffes
8. Millisort: a distributed sorting for millisecond-scale time budget – Seo Jin Park, Yilong Li, Collin Lee
9. Smart Harvesting of Spare CPU Cores in Cloud Servers – Yawen Wang
10. DejaVu: Enhancing Videoconferencing with Prior Knowledge – Pan Hu
11. SPN Chatbot: Handling Compound and Domain-Specific Queries
12. Tackling Frontier in Pull-based Graph Processing on CGRA - Timothy Chong
13. Multiagent Adaptive Scheduling, Planning and Control – Kyle Brown, Oriana Peltzer
14. Programmable Packet Scheduling - Stephen Ibanez
15. Task Assignment for Flexible Manufacturing – Patrick Washington
16. Implementing Linearizability at Large-Scale and Low-Latency – Seo Jin Park, Collin Lee
17. ...
Difficult to use RPCs for update operations
- No response for an RPC: updated or not??
  - Possible causes: message dropped, server crashed, etc…
- Most systems: resend the request until succeed.
  - Re-execution endangers consistency!