Cloud Computing Topics Preview

Kshiteej Mahajan

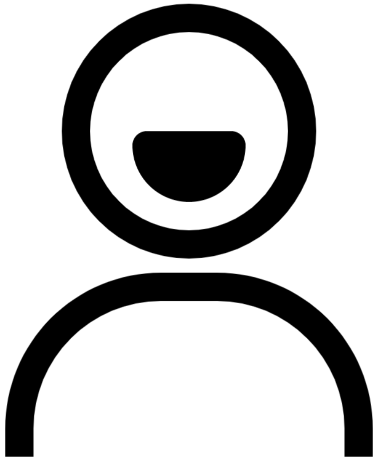
WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

Agenda

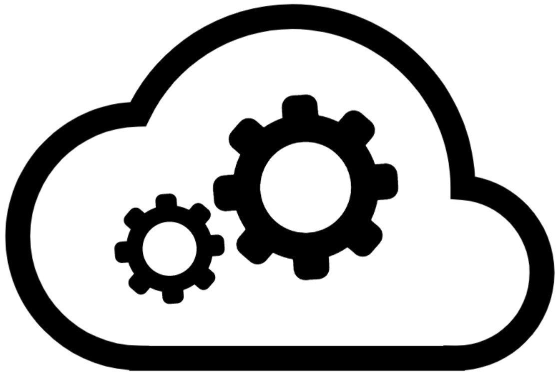
- Cloud Computing for Data Analytics 101
  - Cloud Computing Stack
  - Explaining the Stack
  - Challenges and Evolution
- Cloud Computing Papers at ATC
  - Problems and Key Ideas

Cloud Computing for Data Analytics

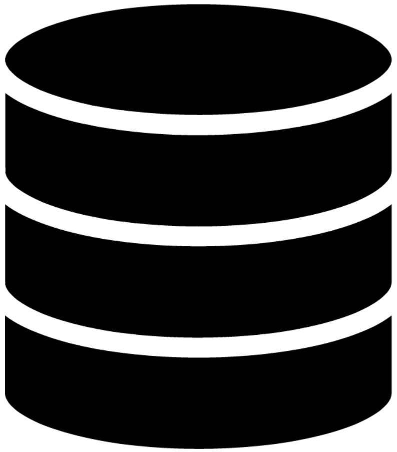
Applications



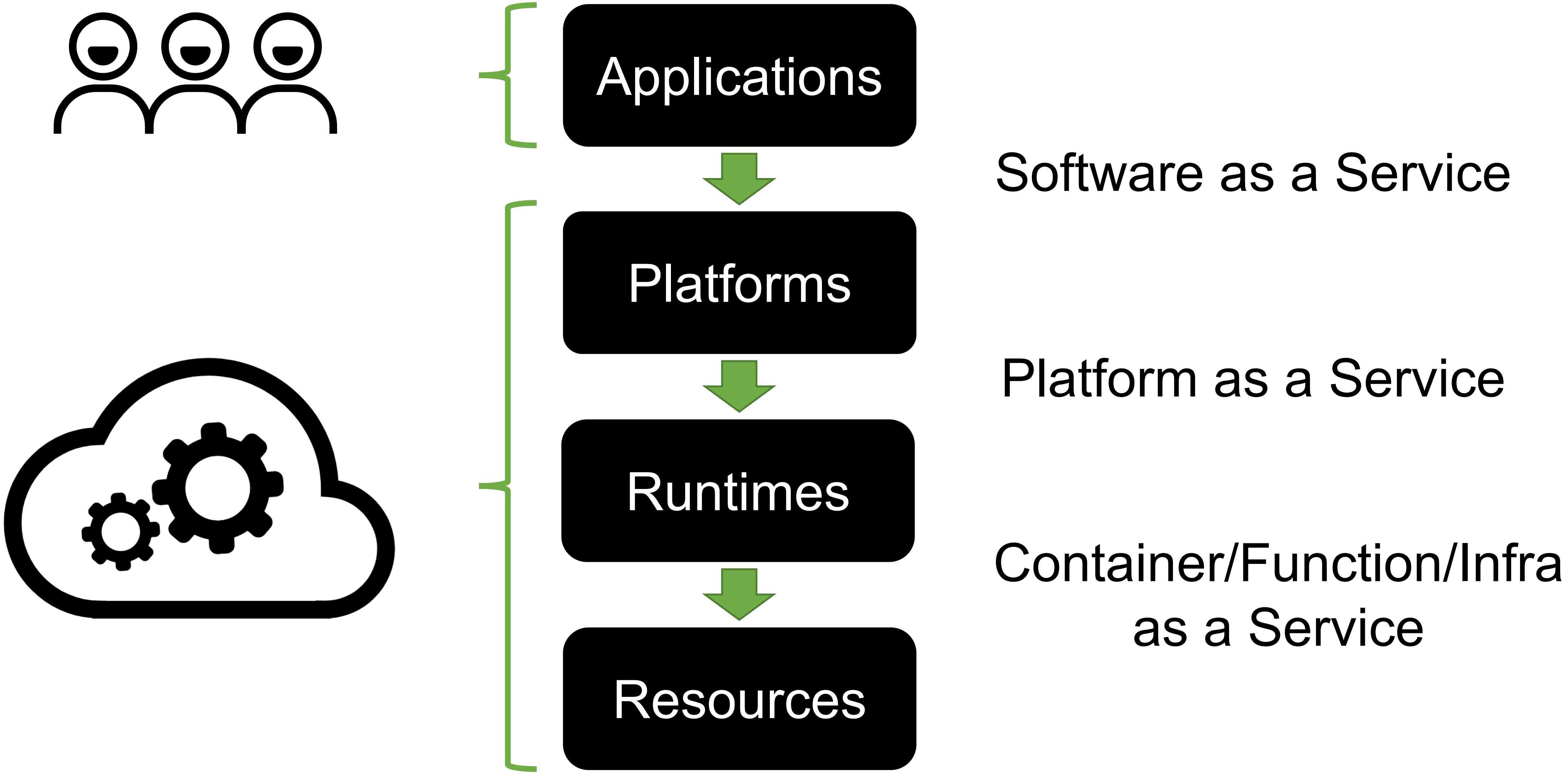
Store
Serve
Analyze



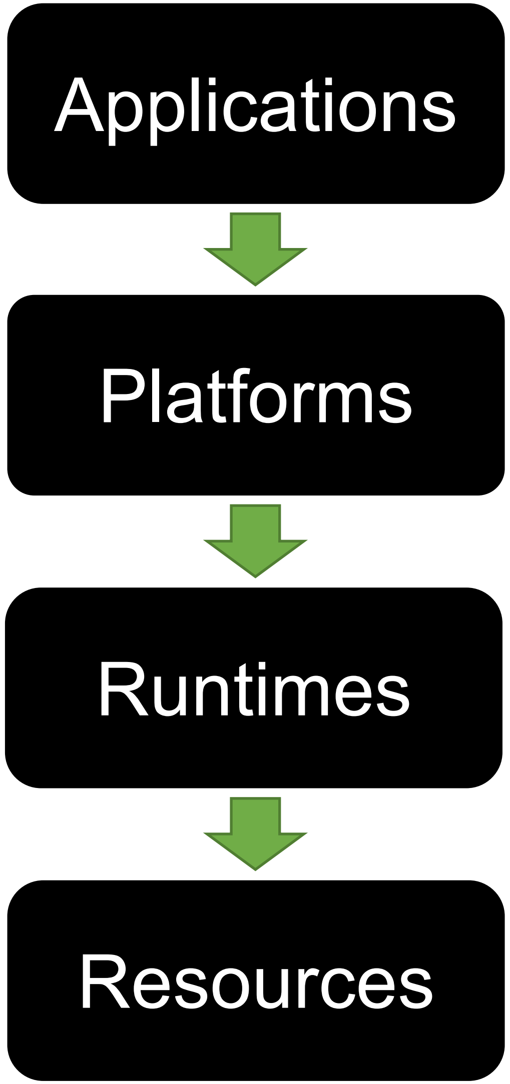
Datasets



Cloud Computing Stack



Explaining the Stack



Explaining the Stack

Applications

Platforms

Runtimes

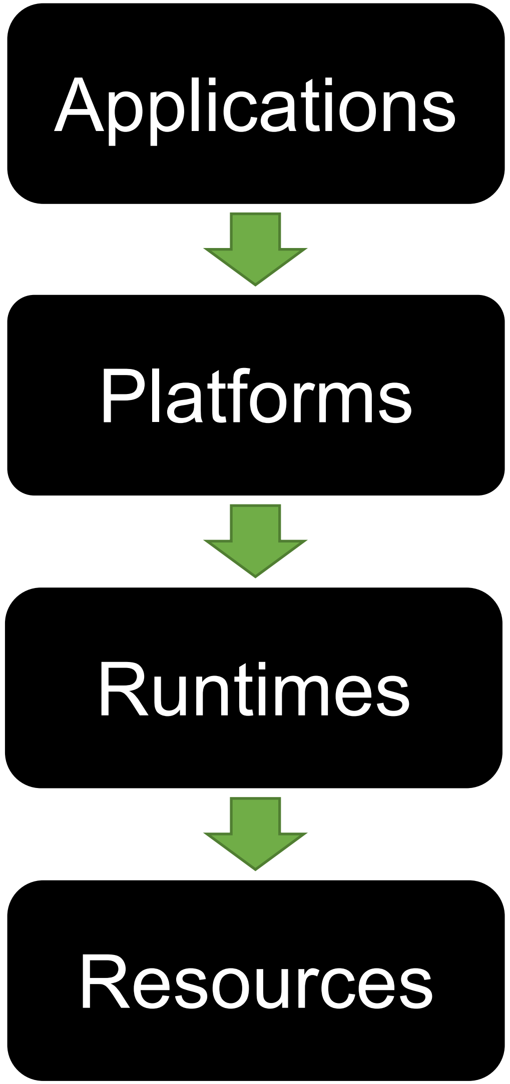
Resources

SQL Query

Apache Spark SQL

Write Query to Analyze Weather Data:
Coldest City in Wisconsin?

Explaining the Stack

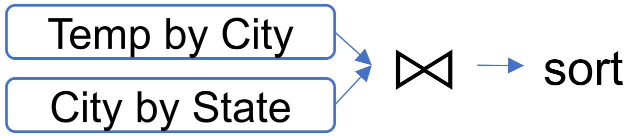


SQL Query

SQL

Write Query to Analyze Weather Data:
Coldest City in Wisconsin?

Compile to a DAG



Explaining the Stack

Applications

Platforms

Runtimes

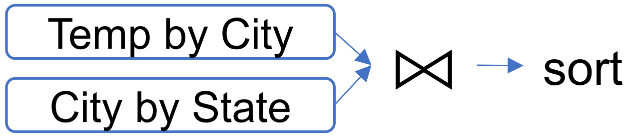
Resources

Write Query to Analyze Weather Data:
Coldest City in Wisconsin?

SQL Query

Apache Spark SQL

Compile to a DAG



Manage Execution and Intermediate State

Submit Ready Tasks

Explaining the Stack

Applications

Platforms

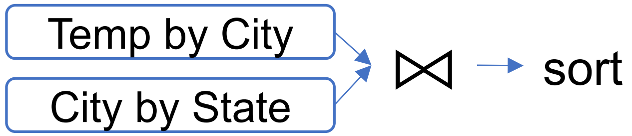
Runtimes

Resources

Write Query to Analyze Weather Data:
Most Cold City in Wisconsin?

SQL Query

Compile to a DAG



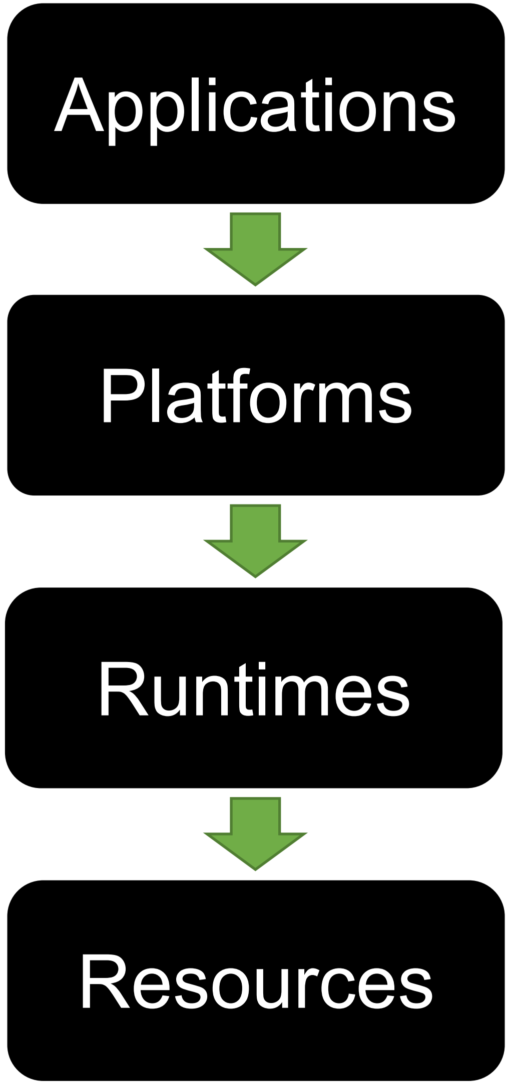
Manage Execution and Intermediate State

Submit Ready Tasks

Linux OS Runtime

VMs in the Cloud

Cloud Computing Stack



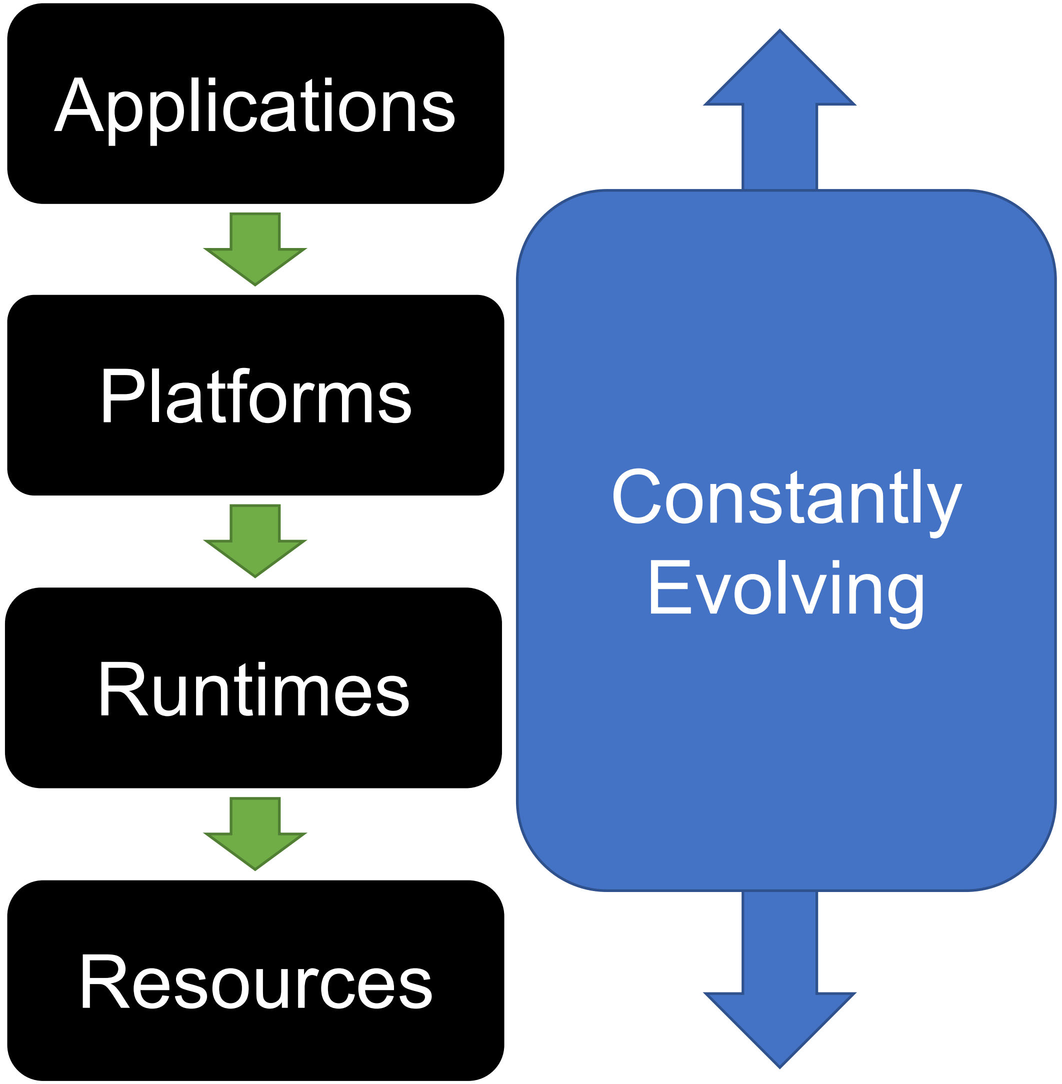
Programming Abstractions / API exposed by Platform to express computation

Compilation as a dataflow of tasks
Manage Execution and Intermediate State

Execute tasks submitted by Platform and give status

Runtime gives access to resources in the Cloud

Cloud Computing Stack is Evolving



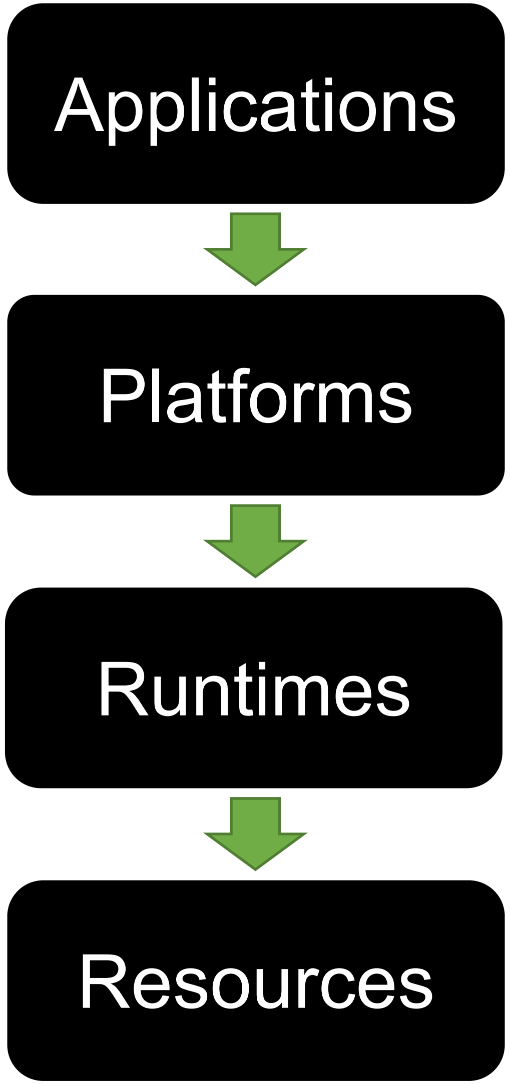
Programming Abstractions / API exposed by Platform to express computation

Compilation as a dataflow of tasks
Manage Execution and Intermediate State

Execute tasks submitted by Platform and give status

Runtime gives access to resources in the Cloud

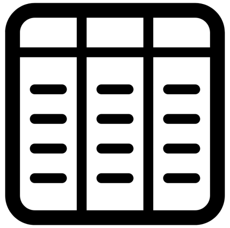
Cloud Computing Stack is Evolving



How are datasets and use-cases evolving?

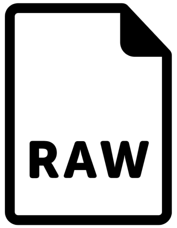
Applications

Structured Data



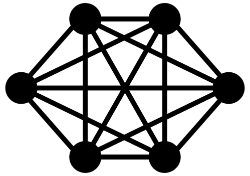
Coldest City from Weather Data

Unstructured Data



Keyword Filter on Tweets

Graphs



Page Rank on Hyperlinks

Dataset Type ***Use Case***

Applications

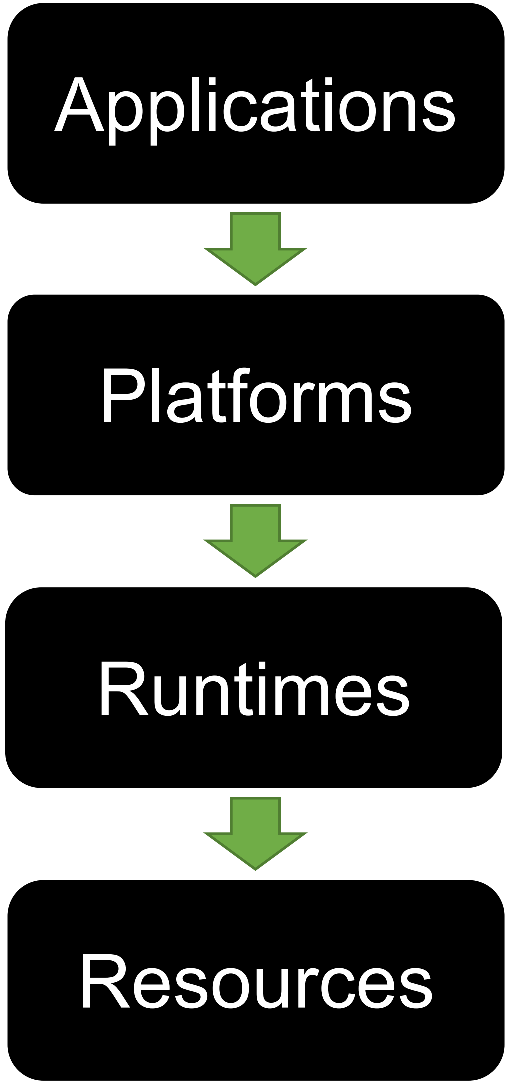
Structured Data — Coldest City from Weather Data

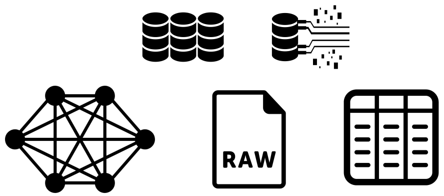
Unstructured Data — Word Search in Tweets

Graphs — Page Rank on Hyperlinks

Dataset Modality	Use Case
Batched Structured Data	Lowest Temperature from last years Weather Data
Streaming Structured Data	Lowest Temperature in last 5 minutes from Real-Time Weather Data

Cloud Computing Stack is Evolving





Tables → Unstructured → Graphs
Batched Data → Streaming Data

How to design faster and resource-efficient platforms for new use-cases?

Platforms

Batched Graph Analytics

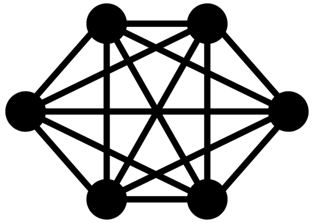
Streaming Graph Analytics

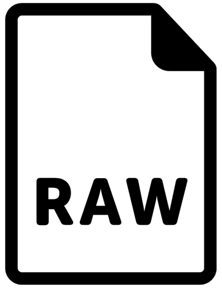
Batched Unstructured Data Analytics

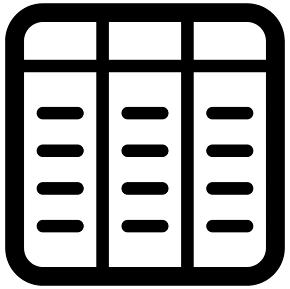
Streaming Unstructured Data Analytics

Batched Structured Data Analytics

Streaming Structured Data Analytics







Cloud Computing Stack is Evolving

Applications

Platforms

Runtimes

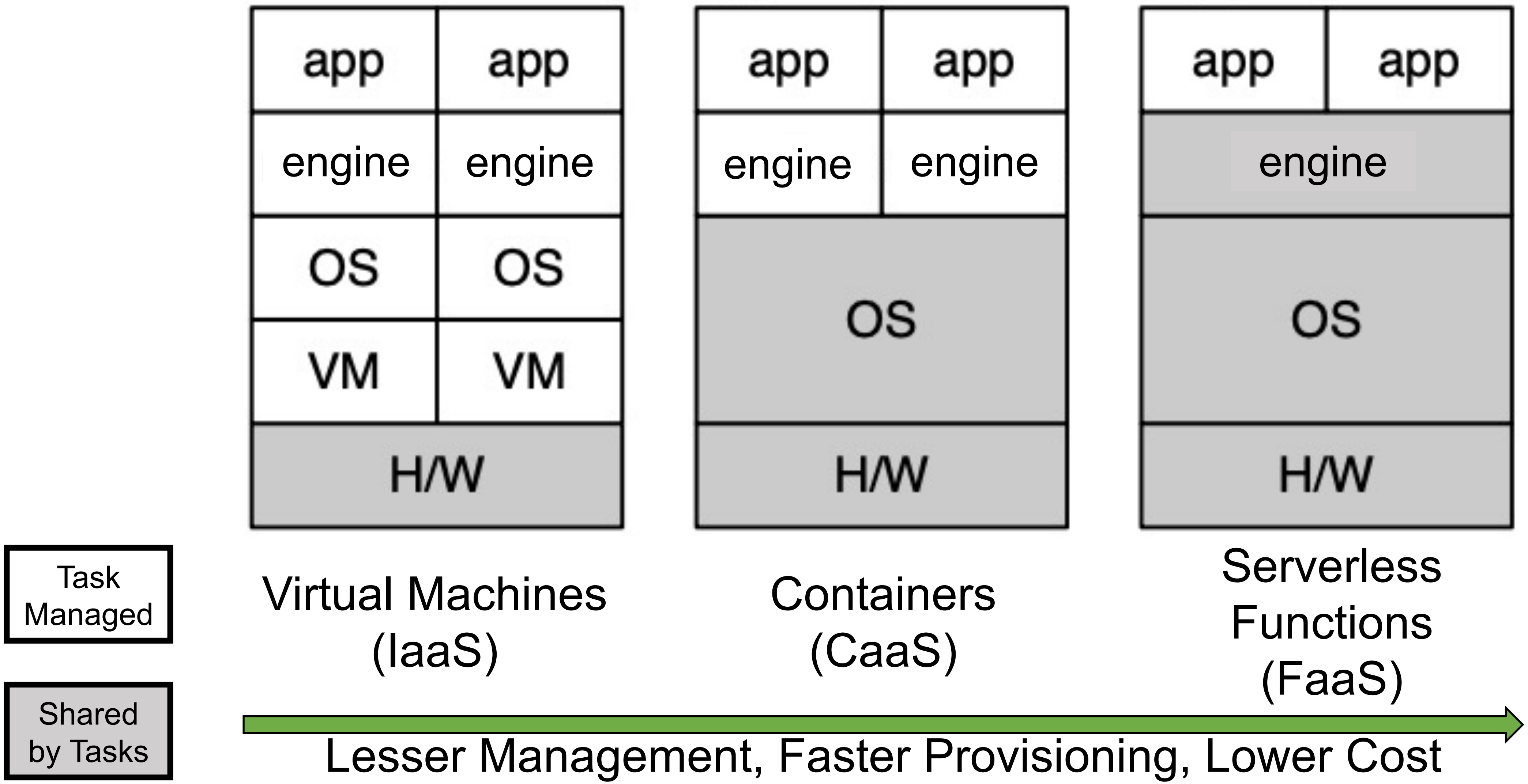
Resources

Tables → Unstructured → Graphs
Batched Data → Streaming Data

Use-case specific Platforms

How to make runtimes fast, efficient and secure?

Runtimes



* Image from OpenLambda, HotCloud'16

Cloud Computing Stack is Evolving

Applications — Tables → Unstructured → Graphs; Batched Data → Streaming Data

↓

Platforms — Use-case specific Platforms

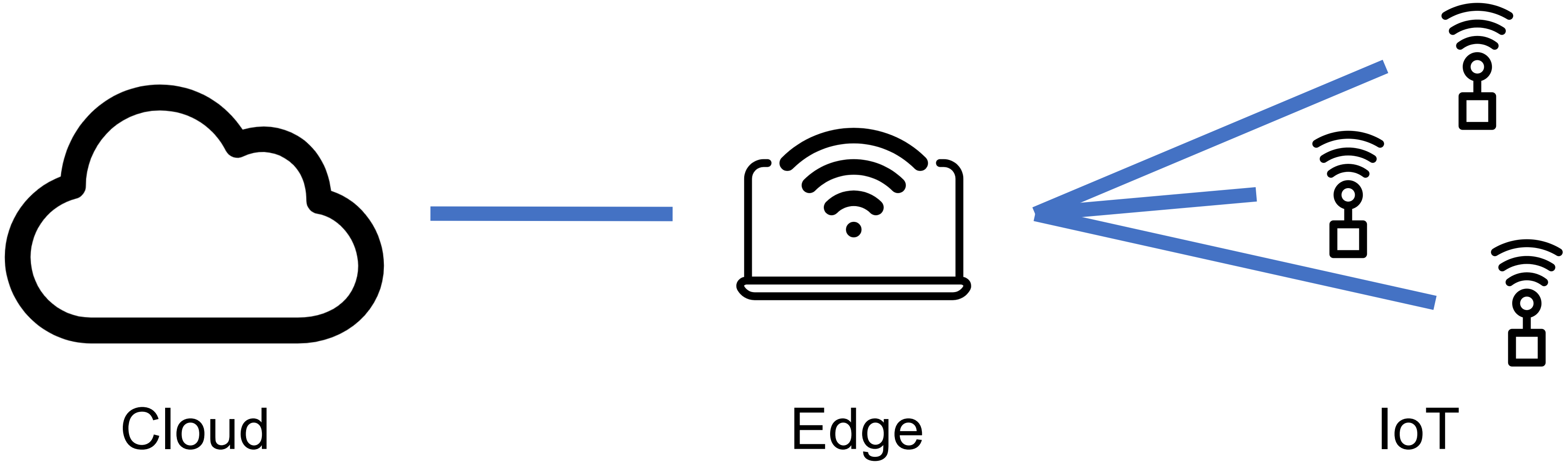
↓

Runtimes — VMs → Containers → Serverless

↓

Resources — What about new types of resources?

Resources



Machines with Accelerators,
More Memory,
Faster Interconnects

Partition Compute Operations across
Cloud, Edge and IoT Resources

Cloud Computing Stack is Evolving

Applications → **Platforms** → **Runtimes** → **Resources**

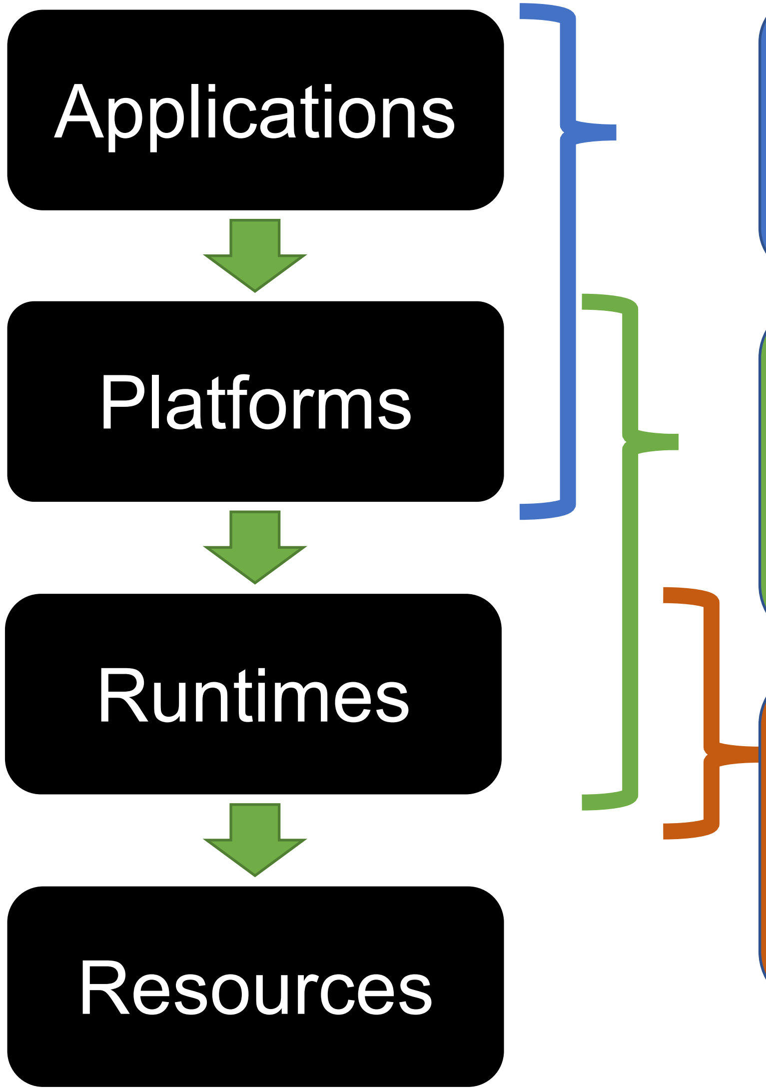
Tables → Unstructured → Graphs
Batched Data → Streaming Data

Use-case specific Platforms

VMs → Containers → Serverless

Cloud → Edge → IoT

Cloud Computing Papers at ATC



- DART
- aDFS
- Controlling Memory Footprint of Stateful Streaming Graph Processing

- CrystalPerf
- Experiences in Managing the Performance and Reliability of a Large-Scale Genomics Cloud Platform
- Fighting the Fog of War: Automated Incident Detection for Cloud Systems (Warden)

- Scaling Large Production Clusters with Partitioned Synchronization (ParSync)
- FaaSNet
- Faastlane
- SONIC

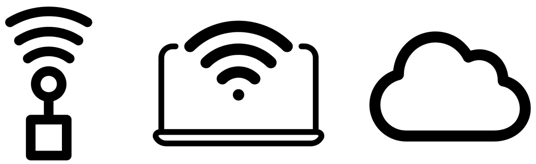
Cloud Computing Papers at ATC

Applications

Platforms

Runtimes

Resources



- DART: A Scalable and Adaptive Edge Stream Processing Engine

Session on July 14 → Searching for Tracks: Graph

Problem: Edge Stream Processing

- Issues
  - Exponential Increase in IoT apps →
    Centralized Monolithic Master → Long Tail-Latency
  - Unexpected and Frequent Failures →
    Provision more Resources / Buffer Data → Poor Fit for Edge

- DART Key Ideas
  - Decentralized architecture
  - DHT based P2P overlay of Streaming Operations for
    Scalability, Low-Latency and Failure Recovery

Cloud Computing Papers at ATC

- Applications
- Platforms
- Runtimes
- Resources

❖ aDFS: An Almost-Depth-First-Search Distributed Graph-Querying System
❖ Controlling Memory Footprint of Stateful Streaming Graph Processing

Session on July 14 → Searching for Tracks: Graph

Problem: Graph Analytics

- Graph Analytics 101
  - vertices partitioned across machines
  - *Compute Phase:* function($vertex_j$) → $intermediate\_state_j$
  - *Aggregation Phase:* $vertex_i$ = aggregate($vertex_{x…z}$)
    - $vertex_{x…z}$ → neighbors($vertex_i$)
- Issues
  - Dynamic Intermediate State
  - Uncontrolled Memory Consumption

Problem: Graph Analytics

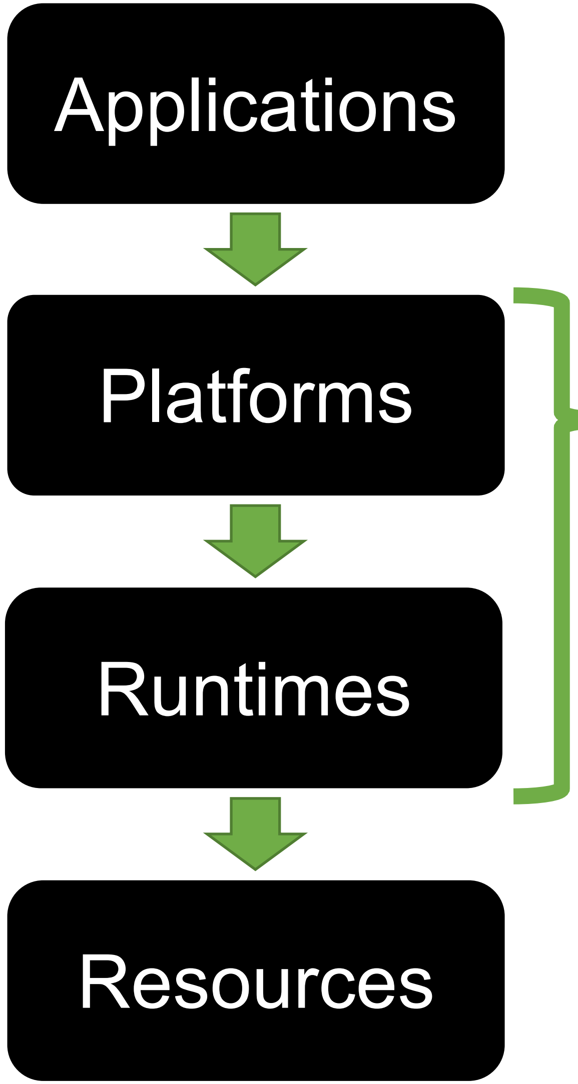
- aDFS Key Ideas
  - Operates under strict Memory Consumption Limits
  - Graph Processing is a combination of BFS- and DFS-traversal
  - Judicious runtime execution decisions to switch across
    - BFS-traversal for more parallel work, more memory and better performance
    - DFS-traversal for eager completion of intermediate work and reduce memory
    - Flow Control to minimize cross-machine chatter and work under target machine memory constraints

Problem: Graph Analytics

- Issues in Streaming Graph Analytics
  - Addition or Deletion of Edges or Vertices → Recompute from Intermediate State for changed portions of Graph
  - Maintain all Intermediate State → Memory Intensive

- Key Ideas in
  Controlling Memory Footprint of Stateful Streaming Graph Processing
  - Selective Stateful Iterative Model
    - How many and which vertex states to track?
  - Minimal Iterative State Model
    - Eliminate the need for intermediate state tracking for several classes of graph algorithms

Cloud Computing Papers at ATC



- CrystalPerf: Learning to Characterize the Performance of Dataflow Computation through Code Analysis

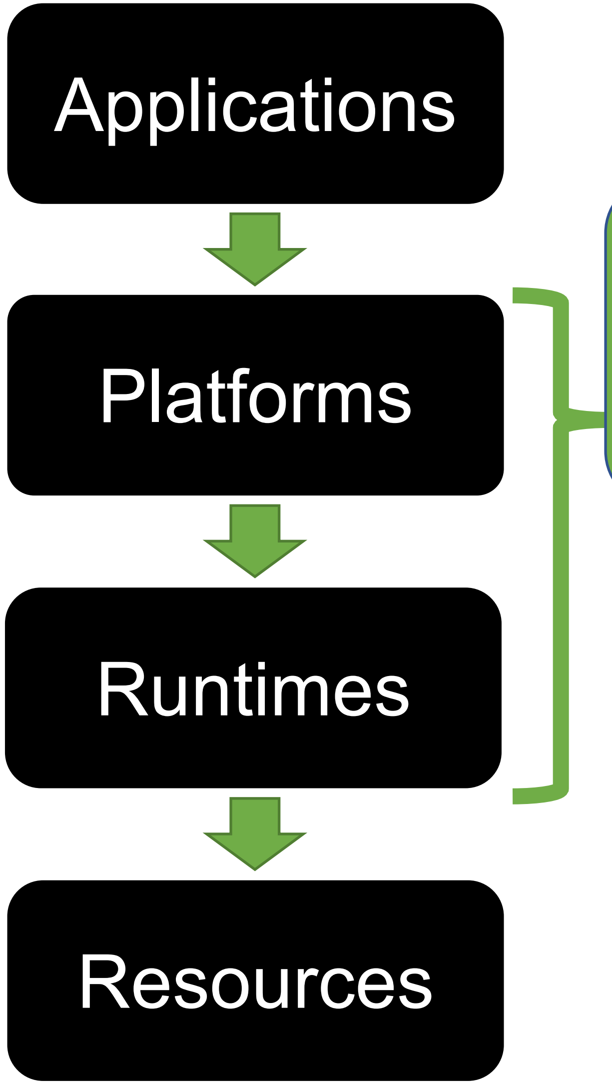
Session on July 14 → Searching for Tracks: Graph

Problem: Perf. Estimation

- Issues
  - Intrusive Platform-specific Code Instrumentation → Doesn’t Generalize
  - Generate Low-level Traces → Requires Manual Analysis

- Key Ideas in CrystalPerf
  - Platform-agnostic and Automated
  - Learning Performance Models for different DAG ops using ML
    - Training from Execution Profiles consisting of
      Code Documentation, Call Traces, Resource Config, Operation Execution Time
  - Prediction using Performance Models

Cloud Computing Papers at ATC



- ❖ Experiences in Managing the Performance and Reliability of a Large-Scale Genomics Cloud Platform

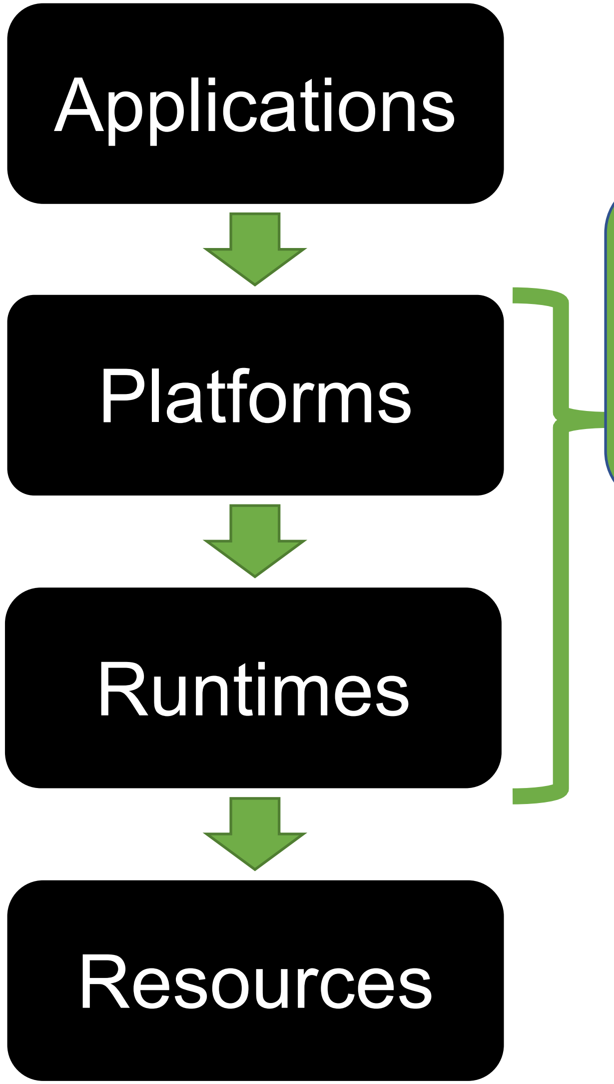
Session on July 15 → Can I Come In? It's Raining!: Cloud Computing

Problem: Perf. Debug

- Issues
  - No trivial performance debugging methodology works

- Key Ideas
  - Found the VM aging problem in OpenStack/KVM stack:
    VM aging → TLB misses → EPT violations → Hypervisor Interrupts → Slowdown
  - Investigated several VM aging mitigation strategies

Cloud Computing Papers at ATC



- Fighting the Fog of War: Automated Incident Detection for Cloud Systems (Warden)

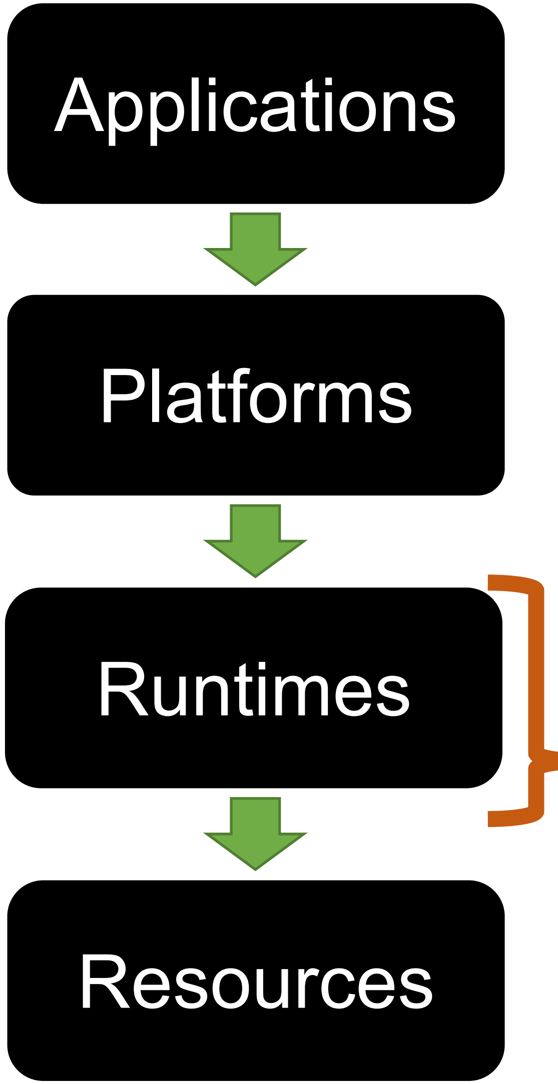
Session on July 15 → Can I Come In? It's Raining!: Cloud Computing

Problem: Perf. Incident Detection

- Issues
  - Handcrafted Rule-based Alert Signals for Incident Detection → Inaccurate and Slow

- Key Ideas in Warden
  - Alert Signal Pruning
  - Learned Model for Incident Detection
    - Input: Features extracted from Pruned Alert Signal and Engineering Activities
  - Auto-Identification of Incident-indicating Alerts via Group-based Model Interpretation
  - Automatic Notification to only the necessary on-call engineers

Cloud Computing Papers at ATC



- Scaling Large Production Clusters with Partitioned Synchronization (ParSync)

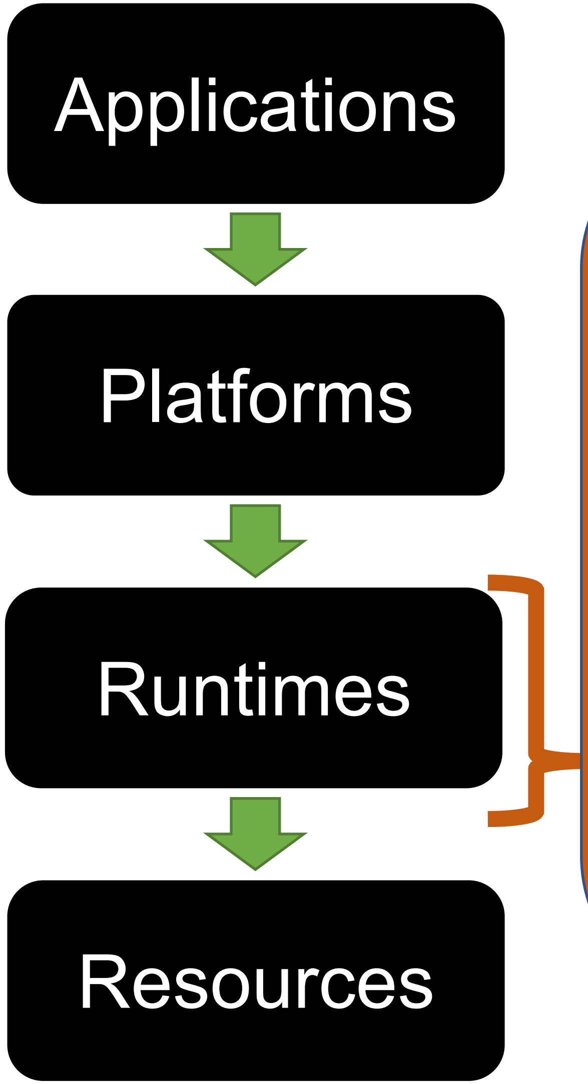
Session on July 15 → Can I Come In? It's Raining!: Cloud Computing

Problem: Scaling Scheduler

- Issues
  - Shared State Scheduler → Not Robust in High Contention Scenarios
  - Two-level Scheduler → Poor Scheduling Efficiency and Quality

- Key Ideas
  - Partitioned Synchronization Architecture
  - Fine-grained Staleness Aware Shared State Scheduler
  - Robustness with High Scheduling Efficiency and Quality

Cloud Computing Papers at ATC



- **FaaSNet:** Scalable and Fast Provisioning of Custom Serverless Container Runtimes at Alibaba Cloud Function Compute

Session on July 15 → Can I Come In? It's Raining!: Cloud Computing

- **Faastlane:** Accelerating FaaS Workflows
- **SONIC:** Application-Aware Data Passing for Chained Serverless Applications

Session on July 16 → But You Played with Me Yesterday: Serverless Computing and Consistency

Problem: Scaling Serverless

- Issues
  1. High Function Provisioning Latency
     - Demand Bursts → Images served from Central Registry → IO and network bottlenecks
  2. High Function Provisioning Costs
     - Higher Request Load → Images served from P2P Registry with more high-performance dedicated root nodes → Increasing Costs
  3. High Function Communication Latency
     - Communicate via object storage or distributed shared memory or global message queues → High interaction latency
     - Function placement is not communication-aware → High Interaction Latency

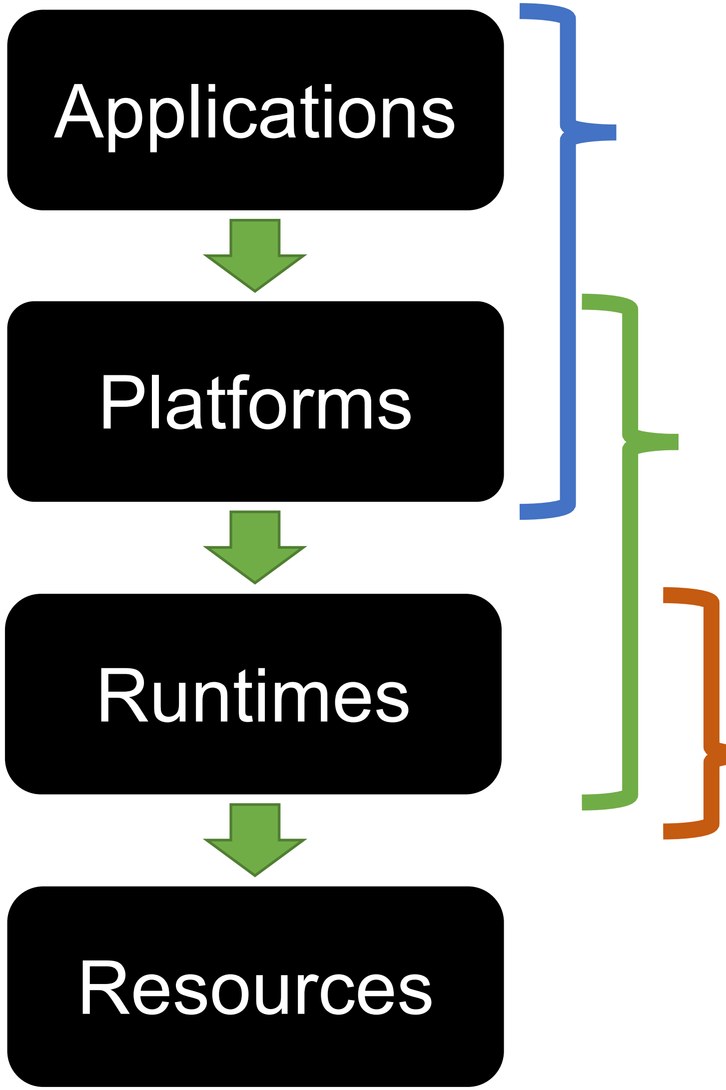
Problem: Scaling Serverless

- Key Ideas in FaaSNet *(Overcoming High Provisioning Latency and Costs)*
  - Per-Function Non-Overlapping Balanced Overlay Trees
    - Use existing VMs for distributing container images

- Key Ideas in Faastlane *(Overcoming High Interaction Latency)*
  - Interacting Functions as separate threads within same process or separate processes in same container

- Key Ideas in SONIC *(Overcoming High Interaction Latency)*
  - Scheduler with communication-aware Function Placement
  - Dynamic choice of data-passing technique for each function interaction

Cloud Computing Papers at ATC



- DART
- aDFS
- Controlling Memory Footprint of Stateful Streaming Graph Processing

- CrystalPerf
- Experiences in Managing the Performance and Reliability of a Large-Scale Genomics Cloud Platform
- Fighting the Fog of War: Automated Incident Detection for Cloud Systems (Warden)

- Scaling Large Production Clusters with Partitioned Synchronization (ParSync)
- FaaSNet
- Faastlane
- SONIC