## **Online Performance Summaries**

A platform for deep analytics of event streams in real-time

Zaheer Chothia < zchothia@inf.ethz.ch >

Advised by: Timothy Roscoe

Department of Computer Science, ETH Zürich

EuroSys Doctoral Workshop – 23.04.2017



Big picture: understanding the datacenter

Applications Middleware

Network

Infrastructure





#### NETFLIX The Netflix Tech Blog Monday, December 31, 2012 Links Netflix US & Canada Blog A Closer Look At The Christmas Eve Outage Netflix America Latina Blog by Adrian Cockcroft etflix Brasil Blor **NETWORKWORLD** Most read: Netflix streaming was impacted Services (AWS) Elastic Load E services supporting streaming. We apologize for the inconven Home > Cloud Computing and how we continue to invest **UPDATE:** Amazon.com suffers outage: Nearly **Partial Outage** \$5M down the drain? The problems at AWS caused PM Pacific Time on December **BUSINESS DAY** Amazon.com, the major onl affected playback on TV conne in the UK, Ireland and Nordic of minutes

Netflix uses hundreds of ELBs service and provides a network Netflix streaming has been imp the last few years, and groups from devices are passed by the Netflix application. Out of hund ability to pass requests to the

A Computer Problem Delays American Airlines Flights MORE (in) By Brandon Butler Follow By JAD MOUAWAD SEPT. 17, 2015

🖌 Email

American Airlines flights were delayed for several hours on Thursday lentified computer problem, just weeks before the

#### The Mystery of Link Imbalance A metastable failure state at scale

Network World | Jan 31, 2013 4:55

P

 $\mathbf{ }$ 

failure state



after a burst, traffic is on overloaded link

es its move to a single reservation system after its merger

at its latest problems were caused by a network affecting various computer systems, including its website in counters at three of its biggest hubs: Chicago, Dallas

d the Federal Aviation Administration to ground and American Eagle flights to and from these airports op program that lasted for about an hour and 45

ad fixed the problem, which it did not disclose, but was he cause. It ruled out an outside breach of its systems.

ith delays that averaged a little over an hour. Six regional ed, according to a spokesman for the airline.

## Big picture: understanding the datacenter

Observation: faced by an abundance of deployment choices



Emblematic operational questions:

- What are the most critical pieces of infrastructure?
- Which applications are responsible for global network traffic?
- What will happen to link utilization if load doubles?
- ... or if I change load balancing strategy?

Approach: deep analytics on shadow infrastructure

# Enterprise DCStrymonImage: Strymon strymon

- The volume of datacenter logs is huge (multi-GB per second)
- Keeping archives is not a viable solution (evolution, privacy)
- Insight: can process logs online with modest cluster resources

Approach: deep analytics on shadow infrastructure



## **CROSS-LAYER PERF. ANALYSIS** Globality, traffic matrices, what-ifs

# 2

## **EXPLAINING OUTPUTS**

"Why is this record in the output of my dataflow?"

## **GENERALIZING CRITICAL PATH**

"Why is my distributed dataflow computation slow?"

## Grok the data center

## **Goal:**

A flexible, reusable infrastructure for online, cross-layer modeling and analysis of the state of an enterprise data center

online: continuous process, results emitted in seconds (vs. minutes/hours) cross-layer: reusable abstractions spanning a wide portion of the stack

## Service and application view





## Network View



## A twist: Cross-layer view



## A twist: Cross-layer view



## Result: explanation of traffic globality

Detailed reasons for hops traversing the distribution layer



## Fork(): real-time what-if scenarios



## Fork(): real-time what-if scenarios



Approach	Drawback
Simulators and emulators	Typically don't scale to data center size, or accurately reflect performance
Live testbeds	Prohibitively expensive, and fundamentally a <i>different</i> data center
Formal verification	Largely focused on properties at L2/L3

## **CROSS-LAYER PERF. ANALYSIS**

Globality, traffic matrices, what-ifs

## **EXPLAINING OUTPUTS**

"Why is this record in the output of my dataflow?"

## **GENERALIZING CRITICAL PATH**

"Why is my distributed dataflow computation slow?"

















## Approach: iterative backward tracing



## Approach: iterative backward tracing



#### Augment the original dataflow with a shadow dataflow

## **Explaining Outputs** concisely and correctly



Perhaps surprisingly:

1. It works for non-monotonic computations!

- 2. It's about 150 lines.
- 3. Updates in real-time
- 4. Limited overhead

Experiments:

- 1. Graph connectivity (monotonic) over Twitter graph (1.5B edges; 42M nodes)
- 2. Stable matching (non-monotonic) over LiveJournal (68M edges; 5M nodes)

## CROSS-LAYER PERF. ANALYSIS

Globality, traffic matrices, what-ifs

## **EXPLAINING OUTPUTS**

"Why is this record in the output of my dataflow?"

## **GENERALIZING CRITICAL PATH**

"Why is my distributed dataflow computation slow?"

## Bottleneck detection is hard



- Many segregated processes and program activities
- The cause is usually not isolated but spans multiple distributed workers





- Models happens-before relationships
- Vertices: events with timestamps
- Edges: duration of activities
- Wait edges: time spent in waiting for a message



- Models happens-before relationships
- Vertices: events with timestamps
- Edges: duration of activities
- Wait edges: time spent in waiting for a message



- Models happens-before relationships
- Vertices: events with timestamps
- Edges: duration of activities
- Wait edges: time spent in waiting for a message



- Models happens-before relationships
- Vertices: events with timestamps
- Edges: duration of activities
- Wait edges: time spent in waiting for a message



## Critical path analysis

The critical path is the path of non-waiting activities in the execution history of the program with the longest duration



## How can we compute the critical path in long-running, dynamic distributed applications, with possibly unbounded input?

- There may be no "job end"
- The PAG is evolving while the job is running
- Stale profiling information is not useful

## Transient critical paths

An adaptation of standard critical path on trace snapshots

- tumbling, sliding, or custom windows
- more than one transient critical path per snapshot



# Transient critical paths are widely applicable



- data transformation
- data exchange
- control messages
- data (de)-serialization
- buffer management

common set of low-level primitives!

## SUMMARY

#### PART II: Iterative Backward Tracing



#### Part III: Online Critical Path Analysis



#### concise explanations output reproduction guarantees interactive response times

#### transient critical paths real-time performance summaries continuous computations

#### Online Reconstruction of Structural Information from Datacenter Logs

