Towards Transient Resource Usage on Real-Time Stream Processing Systems

Pedro Joaquim
pedro.joaquim@tecnico.ulisboa.pt
INESC-ID, Instituto Superior Técnico,
Universidade de Lisboa
Planner (Group A)

Advisor: Professor Luís Rodrigues
1. The Problem

• How to select cost efficient deployments for stream-processing systems in public cloud environments.
1. The Problem

- How to select cost efficient deployments for stream-processing systems in public cloud environments.

2. Why is it a problem
1. The Problem

- How to select cost efficient deployments for stream-processing systems in public cloud environments.

2. Why is it a problem

- Stream-processing systems are long lived ($$$).
1. The Problem

   • How to select cost efficient deployments for stream-processing systems in public cloud environments.

2. Why is it a problem

   • Stream-processing systems are long lived ($$$).

   • There are thousands of different machine configurations possible.
1. The Problem

- How to select cost efficient deployments for stream-processing systems in public cloud environments.

2. Why is it a problem

- Stream-processing systems are long lived ($$$).
- There are thousands of different machine configurations possible.
- Bad allocation of resources translates into wasted money.
1. The Problem

- How to select cost efficient deployments for stream-processing systems in public cloud environments.

2. Why is it a problem

- Stream-processing systems are long lived ( $$$ ).
- There are thousands of different machine configurations possible
- Bad allocation of resources translates into wasted money

3. ‘Proposed Solution’
1. The Problem

• How to select cost efficient deployments for stream-processing systems in public cloud environments.

2. Why is it a problem

• Stream-processing systems are long lived ($$$).

• There are thousands of different machine configurations possible

• Bad allocation of resources translates into wasted money

3. ‘Proposed Solution’

i. Create a process to derive a performance model for the target system considering machines with different characteristics.
1. The Problem

• How to select cost efficient deployments for stream-processing systems in public cloud environments.

2. Why is it a problem

• Stream-processing systems are long lived ($$$).

• There are thousands of different machine configurations possible.

• Bad allocation of resources translates into wasted money.

3. ‘Proposed Solution’

i. Create a process to derive a performance model for the target system considering machines with different characteristics.
1. The Problem

- How to select cost efficient deployments for stream-processing systems in public cloud environments.

3. ‘Proposed Solution’

i. Create a process to derive a performance model for the target system considering machines with different characteristics.

\[
\hat{I} = x_1 \cdot \left( \frac{\text{packet}_{\text{size}}}{\text{network}_a} + \frac{\text{packet}_{\text{size}}}{\text{network}_b} \right)
\]

2. Why is it a problem

- Stream-processing systems are long lived ( $$$ ).
- There are thousands of different machine configurations possible
- Bad allocation of resources translates into wasted money
1. The Problem

- How to select cost efficient deployments for stream-processing systems in public cloud environments.

2. Why is it a problem

- Stream-processing systems are long lived ($$$).
- There are thousands of different machine configurations possible
- Bad allocation of resources translates into wasted money

3. ‘Proposed Solution’

i. Create a process to derive a performance model for the target system considering machines with different characteristics.

\[
\hat{l} = x_1 \times \left( \frac{\text{packet}_{size_a}}{\text{network}_{a}} + \frac{\text{packet}_{size_b}}{\text{network}_{b}} \right)
\]

No longer a black box function!
1. The Problem

• How to select cost efficient deployments for stream-processing systems in public cloud environments.

3. ‘Proposed Solution’

i. Create a process to derive a performance model for the target system considering machines with different characteristics.

2. Why is it a problem

• Stream-processing systems are long lived ($$$).

• There are thousands of different machine configurations possible

• Bad allocation of resources translates into wasted money
1. The Problem

- How to select cost efficient deployments for stream-processing systems in public cloud environments.

2. Why is it a problem

- Stream-processing systems are long lived ($$$).
- There are thousands of different machine configurations possible
- Bad allocation of resources translates into wasted money

3. ‘Proposed Solution’

i. Create a process to derive a performance model for the target system considering machines with different characteristics.

ii. Leverage the dynamic availability and hardware heterogeneity of transient resources to reduce the operational costs.
1. The Problem

- How to select cost efficient deployments for stream-processing systems in public cloud environments.

2. Why is it a problem

- Stream-processing systems are long lived ($$$).
- There are thousands of different machine configurations possible.
- Bad allocation of resources translates into wasted money.

3. ‘Proposed Solution’

i. Create a process to derive a performance model for the target system considering machines with different characteristics.

ii. Leverage the dynamic availability and hardware heterogeneity of transient resources to reduce the operational costs.

4. Expected outcomes

- A real-time stream processing system that is able to self manage its physical resources on public cloud environments with minimal user interaction.
1. The Problem

- How to select cost efficient deployments for stream-processing systems in public cloud environments.

2. Why is it a problem

- Stream-processing systems are long lived ($$$).
- There are thousands of different machine configurations possible
- Bad allocation of resources translates into wasted money

3. ‘Proposed Solution’

   i. Create a process to derive a performance model for the target system considering machines with different characteristics.

   ii. Leverage the dynamic availability and hardware heterogeneity of transient resources to reduce the operational costs.

4. Expected outcomes

- A real-time stream processing system that is able to self manage its physical resources on public cloud environments with minimal user interaction.
  - Cost
  - Performance
Thank You

Questions?