Infrastructure to Application Information Exposure and Communications (i2aex) -- DC Case

IETF i2aex BoF
Draft version: v0.6

Mar. 26, 2012
Based on Input from Many People

- Kamil Bajda-Pawlikowski
- Florin Balus
- Nabil Bitar
- Harry Liu
- Hui-Lan Lu
- Ramki Gummadi
- Vijay Gurbani
- Enrico Marocco
- David McDysan
- Tom Nadeau
- Ping Pan
- Mircea Pisica
- Sabine Randriamasy
- Alexander Tian
- Andreas Voellmy
- Ye Wang
- Henderickx Wim
- Richard Yang (coordinator)
Scope

- Limited to applications with significant components that are (or could be) deployed in data centers (DC)

- Limited to infrastructure -> application info flow
  - could be query/response, but info bits are from infra -> app
  - focus on information that
    - applications require (or benefit in a significant way)
    - cannot be made available easily or through existing mechanisms in a practical way

- Not limited to information that infrastructure already has
  - assume that if there is a strong need, infrastructure can collect

- Use Case/actual projects driven
Basic Entities in an App

• Node entities
  – Compute element
  – Storage element
  – Middlebox element
  – External client
  – ...

• Inter-entity relation
  – On same/not_the_same (node, subnet, VLAN, IP, VPN, availability zone, update domain)
  – Latency/bw/loss
  – ...

Node entities:
- Compute element
- Storage element
- Middlebox element
- External client
- ...

Inter-entity relation:
- On same/not_the_same (node, subnet, VLAN, IP, VPN, availability zone, update domain)
- Latency/bw/loss
- ...
Entities Example

Internet

Consumers

L1/2/3 VPNS

Data Center 1

Data Center 2

Infrastructure A

Infrastructure B

App

MBox

VM

LUN
Why Infrastructure Info Exposure

- **Discovery**: App/other infrastructure could monitor its current inventory, but does not know the invisible (resources/policies)/could-be-available

- **Aggregation/service**: The infrastructure is already monitoring, reduce App complexity and provide (monitoring) information as a service

- **Coordination/Joint Optimization (JO)**: Observe across Apps, signaling for joint optimization
Challenges of Infrastructure Info Expo

• Consistency
  – The infrastructure info could be highly dynamic.

• Security and privacy
  – The infrastructure may not want to reveal some info, in particular, if across different administrative domains.

• Interdomain
  – Information may come from multiple domains.

• Transparency
  – Exposed info may remove infrastructure flexibility (e.g., VM migration); note that invisible actions from infrastructure may violate app constraints/expectation or lead to the need of notification.

• Heterogeneity
  – Diverse infrastructure technologies and construction.

In addition to other considerations such as scalability
Use Case: Network Rack/Location Awareness

- Example project: Hadoop/MapReduce

- Setting and goal: app uses topology awareness for
  - block placement: multiple copies of same block at different racks for (1) reliability, (2) flexibility in task scheduling
  - task placement: place a task close to block, and/or close to communicating tasks

- Current I2A API: A RackID resolver API to map from node IP/DNS name to a rack ID
  - e.g., 192.168.10.20 -> /dc1/rack2

- Info type: App entity DC location discovery

- Relationship w/ ALTO:
  - ALTO can implement the API using network map, and cost map can be more general than the tree distance assumption
Use Case: Hybrid Cloud Bandwidth On Demand

• Example: Hybrid cloud

• Setting and goal:
  – Discover topology/bandwidth/latency between two infrastructures (e.g., a private cloud and a public (virtual private) cloud)

• Potential I2A: (WAN) topology/bandwidth/latency between/among infrastructures’ boundaries

• Info type: Infrastructure interconnect capacity discovery/service

• Relationship w/ ALTO: potential extension to handle changes in interconnection state/capacity/performance.
Use Case: DC Hosted Virtual Desktop

- Example project: ATIS Cloud Service Forum (CSF) for hosted virtual desktop services for enterprises

- Setting and goal:
  - A virtual desktop (VD) is mapped to a VM in a DC
  - The VM should be close to the end user
  - Federation of VD providers to choose close-by VD

- Potential I2A: QoS between end user and candidate VD

- Info type: Cross-domain resource/location discovery

- Relationship w/ ALTO: ALTO appears to provide the basic abstractions; Inter-server communication (Cross-Domain Coordination) can make the topology and cost map available across domains.
Use Case: Network QoS Awareness

- **Example project:** QoSaaS in the context of Microsoft Lync

- **Setting and goal:** provide QoS metrics (e.g., delay, loss) between end hosts and media servers deployed at data centers, for
  - diagnosis,
  - user QoS expectation (indication of QoS bars), and
  - app adaptation (e.g., choosing the right media gateway)

- **Current I2A info:** QoS prediction between entities

- **Info type:** Aggregation/service

- **Relationship w/ ALTO:** ALTO appears to provide the basic abstractions; can it handle the dynamic info required? Will a sub/pub framework better for such a service?
Use Case: Inter-DC Bulk Transfer

- Example project: NetStitcher

- Setting and goal:
  - many large organizations run backup/replication among multiple sites (DCs), e.g., Google inter-DC copy service
  - app: leveraging delay elasticity of such apps to rescue non-peak bw

- Potential I2A: leftover bw prediction at different locations, time
- Info type: Coordination/Joint Optimization
- Relationship w/ ALTO: ALTO cost map may carry left over bw, but it does not have the time dimension
Use Case: Cloud Resource Monitoring

- Example project: Amazon CloudWatch

- Setting and goal: monitoring predefined/user defined metrics on infrastructure resources, allows alert, connection to infrastructure-provided auto-scaling action

- Current I2A: retrieve/report metrics/simple statistics; specify some actions on metrics
- Info type: Aggregation/service
- Relationship w/ ALTO: Do we want to substantially expand the current schema? Add sub/pub/triggering?
Thank You
Example project: Microsoft D3