ALTO
Internet-Draft
Intended status: Standards Track
Expires: August 17, 2014

S. Kiesel
University of Stuttgart
M. Scharf
Alcatel-Lucent Bell Labs

February 13, 2014

ALTO metrices for expressing availability information
draft-kiesel-alto-availability-metrics-00

Abstract

This document specifies new metrices to be used with the ALTO protocol. The goal is to provide information about the availability of physical network, host, and storage infrastructures to management systems that orchestrate virtual infrastructures on top of them.

Terminology and Requirements Language

This document makes use of the ALTO terminology defined in [RFC5693] and [RFC6708].

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on August 17, 2014.

Copyright Notice

Table of Contents

1. Introduction

Various virtualization technologies allow to instantiate virtual hosts, virtual storage, and virtual networks on top of their physical counterparts. They can be combined to build complex virtual infrastructures. Management systems automate the task of mapping virtual to physical resources, considering various optimization goals. Mechanisms like live migration of virtual machines or re-shaping the topology of overlay networks allow to dynamically react on changing conditions, both in the virtual infrastructure (e.g., change in demand) and in the underlying physical infrastructures (e.g., change in available resources).

A typical example is, that in a cluster of several physical servers, in times of low demand, all virtual machines could be migrated away from one node. This node would be powered down, in order to save precious energy. If resource utilization is the only optimization goal, the input for the placement/scheduling manager can be gathered by measurements.

If, however, other optimization goals have to be considered, the management system needs external information. For example, if all but two nodes of the cluster are to be shut down, the remaining two nodes should be selected in a way that minimizes the risk of both nodes failing at the same time due to a single root cause. This optimization problem becomes more difficult if not only hosts and storage but also network resources are considered.

This document shows that the ALTO protocol [I-D.ietf-alto-protocol] offers the required base mechanisms for providing a standardized interface to virtual infrastructure orchestration managers, for conveying information about the availability / reliability of the underlying physical infrastructure, This document further defines appropriate metrices for this use case.

2. Classification of availability-related parameters

Important concepts of the ALTO protocol are Endpoints, which are identified by Endpoint Addresses. Endpoints can be grouped in PIDs. Endpoints (and by means of protocol extensions [I-D.roome-alto-pid-properties] also PIDs) may have properties that can be queried using the ALTO protocol. Paths between PIDs may have one or more path costs according to some cost metric. These path costs can be queried for individual pairs of PIDs, or a whole cost map (i.e., a "PID x PID -> cost" matrix) can be downloaded. The path cost concept can easily be generalized to a path property concept.

This section discusses how these base mechanisms can be used to convey information related to availability of physical infrastructures to systems that manage virtual infrastructures on top of them.

2.1. Identification of physical resources

In order to identify physical resources within the ALTO protocol, an appropriate endpoint address type has to be used. The ALTO base protocol specification [I-D.ietf-alto-protocol] only defines IPv4 and IPv6 addresses, and establishes a process to register further types. In fact, IP addresses may be used to identify physical resources in many cases, e.g., the loopback address of a router, or the management address of a physical server, etc. For a discussion of VPNs and ALTO see [I-D.scharf-alto-vpn-service].

TBD: discussion of further options for endpoint addesses.

2.2. Classification of cost types and properties

Information related to availability of physical resources may be of different fundamental natures, requiring different encodings and different update intervals. This section itemizes several criteria.

2.2.1. Static vs. dynamic facts vs. probabilites

Information may be static facts that change never or very infequently. For example: "Electrical power outlets A and B are both connected to circuit breaker F1".

Information may also be more frequently changing, e.g., "Uninterruptible power supply UPS1 is now running on battery power, 82% capacity left". TBD: further investigation and guidance is needed on the maximum update frequency that can reasonably be done using ALTO.

Another type of information are statistical measures such as the average relative availability of a subsystem, e.g., the famous "five nines".

2.2.2. Causality and Correlation

Many initial incidents can cause a series of events, according to some kind of "failure propagation topology", which is independent of the IP network topology. There may be even hierarchies.

For example, "Servers S1 and S2 are connected via circuit breaker F1 to uninterruptible power supply UPS1 while S3 and S4 are connected via F2 to UPS1" implies that a failure in S1 triggering F1 will also interrupt operation of S2. Furhermore, shutting down S2, S3, and S4 in case of a power grid failure could strech UPS1's battery lifetime and thereby prolong S1's survivability time. Similar considerations can be made for different kinds of problems, e.g., the impact of a fire.

Modeling diffent risk types (e.g., power outage, fire, flooding, physical intruders, etc.) in their respective terminology would require the definition of many new data types.

A more generic approach is to use an ALTO cost map as a matrix, which indicates the level of isolation against "fate sharing" of any two PIDs with respect to a given (physical) risk. In other words, for every specific risk R the coefficients of that matrix could be calculated as

C_R(x,y) = 1 - P(y fails due to R | x fails due to R)

For example, if the risk type is "fire", then a coefficient of 0 could mean "these two physical resources are in the same rack. If one is on fire for any reason, the other one will almost inevitably fail within seconds, too.", a value of 0.3 could mean "the resources are in adjacent buildings" and 0.99999 could mean "these two resources are on different continents and only a natural disaster causing global destruction could disable both of them in one single event".

Note that these conditional properties only indicate how likely it is that the second resource will become unavailable due to the same event that disabled the first resource. They do not indicate how likely it is that the event will actually occur.

TBD: discuss to which extent a single "endpoint address to PID" network map is useful when considering different risk types. The idea behind PIDs is to reduce map size by grouping topologically

close endpoints, but the "failure propagation topologies" may be very unalingned for different risk types. We will probably end up with many very small PIDs.

3. Specification of new Endpoint Address types

TBD.

4. Specification of new Cost and Property types

TBD.

We need: the "isolation level agains fate sharing" matrix, and a list of risk types, in order to give the absolute probability of that risk for a given resource.

5. Obtaining Availability Information

For any ALTO information, it is important to consider whether the ALTO service realistically can discover that information, if the distribution of that information is allowed, if the data is useful, if a client can get that information without excessive privacy concerns, and if the information cannot be gathered easily be found in some other way.

Availability-related parameters can both refer to properties of the network infrastructure (e.g., network resiliency mechanisms) as well as non-networking effects (e.g., redundancy of power supply). In both cases, an application typically cannot measure that information, neither by passive monitoring nor by active probing. Yet, availability information and insight into impact of incidents matters to many applications and can be an important criteria for resource selection decisions. Since typical use cases would be limited to one administrative domain, privacy is not a major concern; in addition, the suggested correlation metrics provide an abstraction over the actual physical infrastructure.

Gathering availability information may be more challenging than, for instance, IP routing topologies. For instance, it may require access to inventory databases. Yet, within one domain, the organization that is responsible for the physical network topology may also take care of other parts of the physical infrastructure, such as the power supply or hardware installation. An organization that operates an ALTO server for exposing network topology information could therefore also have access to other inventory data. Therefore, providing availability information to an ALTO server as described in this document is realistic.

6. IANA Considerations

TBD.

7. Security Considerations

TBD.

8. References

8.1. Normative References

[I-D.ietf-alto-protocol]
Alimi, R., Penno, R., and Y. Yang, "ALTO Protocol", draft-ietf-alto-protocol-25 (work in progress), January 2014.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

8.2. Informative References

[I-D.roome-alto-pid-properties]
Roome, B. and Y. Yang, "PID Property Extension for ALTO Protocol", draft-roome-alto-pid-properties-00 (work in progress), October 2013.

[I-D.scharf-alto-vpn-service]
Scharf, M., Gurbani, V., Soprovich, G., and V. Hilt, "The Virtual Private Network (VPN) Service in ALTO: Use Cases, Requirements and Extensions", draft-scharf-alto-vpn-service-01 (work in progress), July 2013.

[RFC5693] Seedorf, J. and E. Burger, "Application-Layer Traffic Optimization (ALTO) Problem Statement", RFC 5693, October 2009.

[RFC6708] Kiesel, S., Previdi, S., Stiemerling, M., Woundy, R., and Y. Yang, "Application-Layer Traffic Optimization (ALTO) Requirements", RFC 6708, September 2012.

Authors' Addresses

Sebastian Kiesel
University of Stuttgart Information Center
Networks and Communication Systems Department
Allmandring 30
Stuttgart 70550
Germany

Email: ietf-alto@skiesel.de
URI: http://www.rus.uni-stuttgart.de/nks/

Michael Scharf
Alcatel-Lucent Bell Labs
Lorenzstrasse 10
Stuttgart 70435
Germany

Email: michael.scharf@alcatel-lucent.com
URI: www.alcatel-lucent.com/bell-labs