

IETF®

# Network and Computing Exposure

# draft-contreras-alto-service-edge-05 draft-lcsr-alto-service-functions-01

Luis M. Contreras (Telefonica), Jordi Ros Giralt (Qualcomm) Yokohama, ALTO WG, March 2023

draft-contreras-alto-service-edge-07

Luis M. Contreras (Telefonica) Sabine Randriamasy (Nokia) Jordi Ros-Giralt (Qualcomm) Danny Lachos (Benocs) Christian E. Rothenberg (Univ. Of Campinas) draft-lcsr-alto-service-functions-02

Luis M. Contreras (Telefonica) Sabine Randriamasy (Nokia) Xufeng Liu (IBM Corporation)

# Relationship among drafts



### draft-contreras-alto-service-edge

Document describing ALTO as the element to combine compute and network information to determine the more convenient Edge or Compute facility to deploy an application

### draft-lcsr-alto-service-functions

Document describing ALTO as the element to combine service function(s) and network information to retrieve path characteristics to reach a specific SF or for the interconnection paths among a sequence of SFs Overarching document describing the role of ALTO as IETF Network Exposure Function, including existing (i.e., as WG documents or RFCs) and proposed/future capabilities

Others ...

# draft-contreras-alto-serviceedge

## Problem statement

- Multiple (heterogenous) DC Data Centers across the network featuring resources (CPUs, memory, storage, bandwidth, etc)
- Identify the suitable DC to deploy a given application considering both compute and transport information

## <u>Solution</u>

- Leverage the ALTO protocol (+ext) to assist on the selection of the "best" edge, combining both network & compute info.
  - Optionally complemented with other inputs such as performance metrics, etc

#### • Current version -07

 Related with Compute Aware Networking discusión, being ALTO an off-path solution



## Updates in version -06

- Two use cases describes to support the discussion based on real operational situations: (1) Open Abstraction for Edge Computing, (2) Optimized placement of microservice components
- Added Jordi as co-author
  - Further use cases to be added <sup>3</sup>

#### **Split Computation Use Cases**



# draft-lcsr-alto-service-

- Network services formed by means of the concatenation of several atomic service functions (SF), resulting in a connected graph of functions
- Identify suitable path characteristics in the chain

## <u>Solution</u>

- Leverage on ALTO protocol (+ext) to assist on the identification of "best" chains, combining both network & SF info.
  - Optionally complemented with other inputs such as performance metrics, etc
- Extensions foreseen for [I-D.ietf-alto-pathvector], [I-D.ietf-alto-unified-props-new] and mechanisms for collecting SF information

- Current version -02
- Related with Compute Aware Networking discusión, being ALTO an off-path solution



# Updates in version -02

- Reference to [I-D.ldbc-cats-framework] as additional case to motivate this work
- Added reference to [RFC9015], [I-D.xu-lsrisis-service-function-adv] and [I-D.xu-lsr-ospfservice-function-adv] as a mean for rerieving SF information

# *draft-lcsr-alto-service-functions* Some ALTO information of interest (examples)

Assuming that application endpoints are located in PIDs

- Path characteristics, from a PID, to any instance of a service function type.
- Path characteristics, from a PID, to a specific instance of a service function type.
- Path characteristics among any instance of a service function type X to any other instance of a service function type Y.
- Path characteristics among a specific instance of a service function type X to any other instance of a service function type Y.
- Path characteristics, from a PID, to a chain of service functions.
- Path characteristics, from a PID, to a chain of specific instances of service functions.

• etc

# Network and Computing Exposure Summary

- ALTO provides value to this problem space as an off-path solution
- Exposure towards applications is crucial
  - For applications to get by themselves combined information of networking and compute (then deciding instead of delegating decisions)
  - For orchestrators or managers to instantiate functions or applications in the network with information about expected behavior
- Not all the needs of exposing combined compute and network information are the ones related to traffic steering ...

... but in that case ALTO also could feed network elements/controllers to take such steering decisions