Internet Research Task Force
Internet-Draft
Intended status: Informational
Expires: January 9, 2022

Y. Zhu
D. Chen
C. Zhou
China Mobile
July 8, 2021

An Efficient Data collection method for Digital Twin Network
draft-zhu-nmrg-digitaltwin-data-collection-00

Abstract

Digital Twin Network (DTN) is a network system with Physical Network and Twin Network, which can be mapped interactively in real time. The construction of Digital Twin Network requires real-time data of Physical Network to update the state of Twin Network. However the existing method collects the full amount of data from the Physical Network for modeling, and does not consider the problems such as insufficient storage resources, low computational efficiency and waste of bandwidth resources. This document introduces an efficient data collection method in which the Twin Network sends instructions to the Physical Network to collect data on demand, and then the Physical Network parses and executes instructions such as data cleaning and knowledge representation, and sends the processed or requested data to the Digital Twin Network.

Requirements Language

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 https://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 January 9, 2022.

Copyright Notice

Copyright (c) 2021 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (https://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction

With the deployment of Internet of Things, cloud computing and data center, etc., the scale of the current network is expanded gradually. However, the increase of network scale leads to the increasing complexity of the current network, and that induces plenty of problems. In order to improve the autonomy ability of network and reduce the negative effect on Physical Network, it is considered that an endogenous intelligent and autonomous network architecture which achieves self-optimization and decision is indispensable. Digital twin, as an innovative technology, has the potential to realize this architecture because it can optimize and validate policies through real-time and interactive mapping with physical entities.[I-D.zhou-nmrg-digitaltwin-network-concepts]

Data is the cornerstone of DTN construction. In the face of large network scale, data collection, storage and management are faced with great challenges. If the full-data collection method is adopted, huge storage space and bandwidth resource are needed, especially for complex scenarios that require real-time data and traffic from multi-

source heterogeneous devices. Therefore, it is extremely important to propose a lightweight and efficient data collection method.

2. Definitions and Acroyms

DTN: Digital Twin Network

PN: Physical Network

IMC: Instruction Management Center

DSC: Data Storage Center

TN: Twin Network

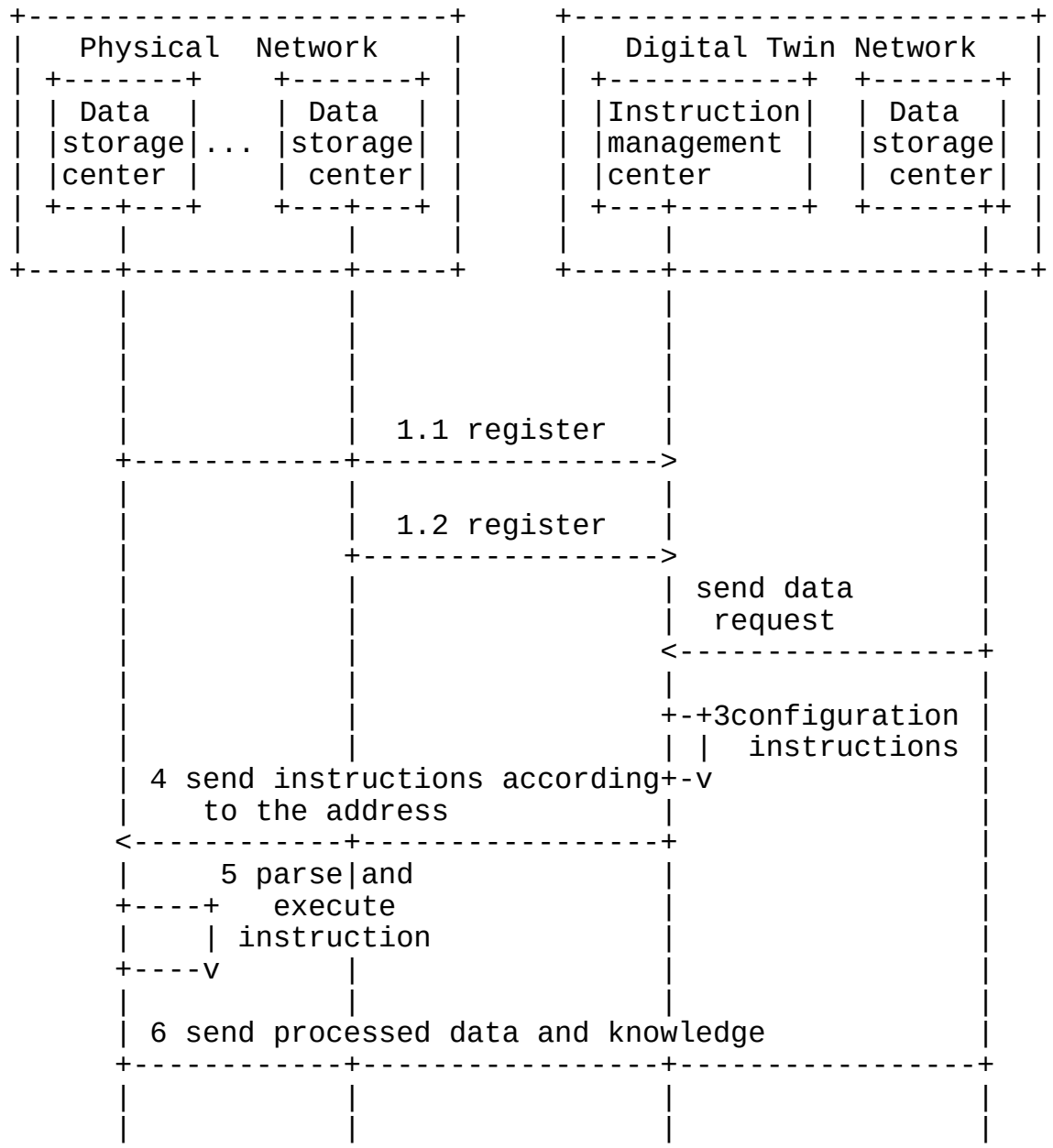
3. Overview

Digital Twin Network (DTN) is a network system with Physical Network and Twin Network, which can be mapped interactively in real time. The construction of DTN requires real-time data of Physical Network to update the state of Twin Network. However the existing method collects the full amount of data from the Physical Network for modeling, and does not consider the problems such as insufficient storage resources, low computational efficiency and waste of bandwidth resources caused by data transmission. In order to solve these problems, this memo introduces an efficient data collection method for DTN. This data collection method is to send instructions in the Twin Network to the Physical Network to collect data on demand, and then the Physical Network parses and executes instructions such as data cleaning or knowledge representation, and then sends the processed or represented data to the DTN.

DTN consists of Physical Network and Twin Network. The Physical Network includes multiple Data Storage Centers, and the Twin Network includes the Instruction Management Center and Data Storage Center. The Instruction Management has two functions. On the one hand, the Instruction Management Center of the Twin Network is mainly used to manage the registration of the Data Storage Center in the Physical Network, and its registration information can include various key information such as the IP address of the Data Storage Center in the Physical Network, data type, and various index names of the data , data source name and data size, etc; on the other hand, it is mainly used to adaptively configure data collection instructions according to the collection requirements of the Data Storage Center in the Twin Network, and search for IP addresses to send instructions. The instruction-carrying information includes rule-based mathematical expressions, executable models in .exe format, dynamic collection frequency, parameter lists, program text files in .m format, text

files with parameter configuration, and other types of files. Instructions are flexible and programmable, and can be created, modified, combined, and deleted at any time according to requirements. When the Data Storage Center of the Twin Network initiates data collection requests to the Instruction Management Center, the Instruction Management Center searches for IP addresses of Data Storage Center from registration information according to critical information such as data type and data name, and functional instructions for data processing or knowledge representation can be implemented depending on the demand configuration. The Data Storage Center of the Twin Network is mainly used to store the effective information after data processing and knowledge representation returned by the Data Storage Center in the Physical Network.

Data Storage Center in the Physical Network has two functions. On the one hand, it can store data, such as performance indicators, operational status, logs, traffic scheduling, business requirements, etc. On the other hand, it has the function of automatically parsing the instructions sent by the Instruction Management Center in the Twin Network. Then the operating environment of the instruction is configured according to the instruction needs, and data processing or knowledge representation is performed based on the instruction. Data processing mainly includes data cleaning, filling missing data, normalization, conflict verification, etc. The role of knowledge representation is to represent the original data as a data structure that can be used to efficiently calculate. Such representation results are closer to the machine language, which is conducive to the rapid and accurate construction of the model.



The specific process is as follows:

- The Data Storage Centers in the Physical Network register to the Instruction Management Center in the Twin Network. The registration information includes the IP address of the Data Storage Center, the data type, the data source, the data size, etc.

- The Data Storage Center in the Twin Network sends the data collection request to the Instruction Management Center.

- According to the data collection request, the Instruction Management Center intelligently searches the registration information for addressing, and configures the data processing instruction. The instruction-carrying information includes rule-based mathematical expressions, executable models in .exe format, dynamic collection frequency, parameter lists, program text files in .m format, text files with parameter configuration, and other types of files. And these are created, modified, combined and deleted flexibly according to requirements

- The Instruction Management Center in the Twin Network sends the corresponding instruction according to the address to the Data Storage Center in the Physical Network.

- After receiving the instructions, the Data Storage Center in the Physical Network will parse and execute them according to the instructions, such as filling missing data, data association, knowledge representation, etc.

- The Data Storage Center of the Physical Network will send the processed and represented data to the Data Storage Center in the Twin Network.

4. Conclusion

This memo introduces an efficient data collection method for DTN. This data collection method is to send instructions model in the Twin Network to the Physical Network to collect data on demand, and then the Physical Network completes instructions such as data cleaning or knowledge representation, and then sends the processed and represented data to the DTN. With this method, DTN can build and maintain it’s data porosity more efficiently and effectively.

5. Security Considerations

TBD.

6. IANA Considerations

This document has no requests to IANA.

7. References

7.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

7.2. Informative References

[I-D.zhou-nmrg-digitaltwin-network-concepts] Zhou, C., Yang, H., Duan, X., Lopez, D., Pastor, A., Wu, Q., Boucadair, M., and C. Jacquenet, "Concepts of Digital Twin Network", draft-zhou-nmrg-digitaltwin-network-concepts-03 (work in progress), February 2021.

Authors' Addresses

Yanhong Zhu
China Mobile
Beijing 100053
China

Email: zhuyanhong@chinamobile.com

Danyang Chen
China Mobile
Beijing 100053
China

Email: chendanyang@chinamobile.com

Cheng Zhou
China Mobile
Beijing 100053
China

Email: zhouchengyjy@chinamobile.com