

# An Efficient Data collection method for Digital Twin Network

draft-zhu-nmrg-digitaltwin-data-collection-00

<https://datatracker.ietf.org/doc/draft-zhu-nmrg-digitaltwin-data-collection/>

# Contents

- Table of Contents
- 1. Introduction . . . . . 2
- 2. Definitions and Acroyms . . . . . 3
- 3. Overview . . . . . 3
- 4. Conclusion . . . . . 6
- 5. Security Considerations . . . . . 6
- 6. IANA Considerations . . . . . 6
- 7. References . . . . . 6
- 7.1. Normative References . . . . . 6
- 7.2. Informative References . . . . . 7
- Authors' Addresses . . . . . 7

# Background & Introduction

## The definition of Digital Twin Network

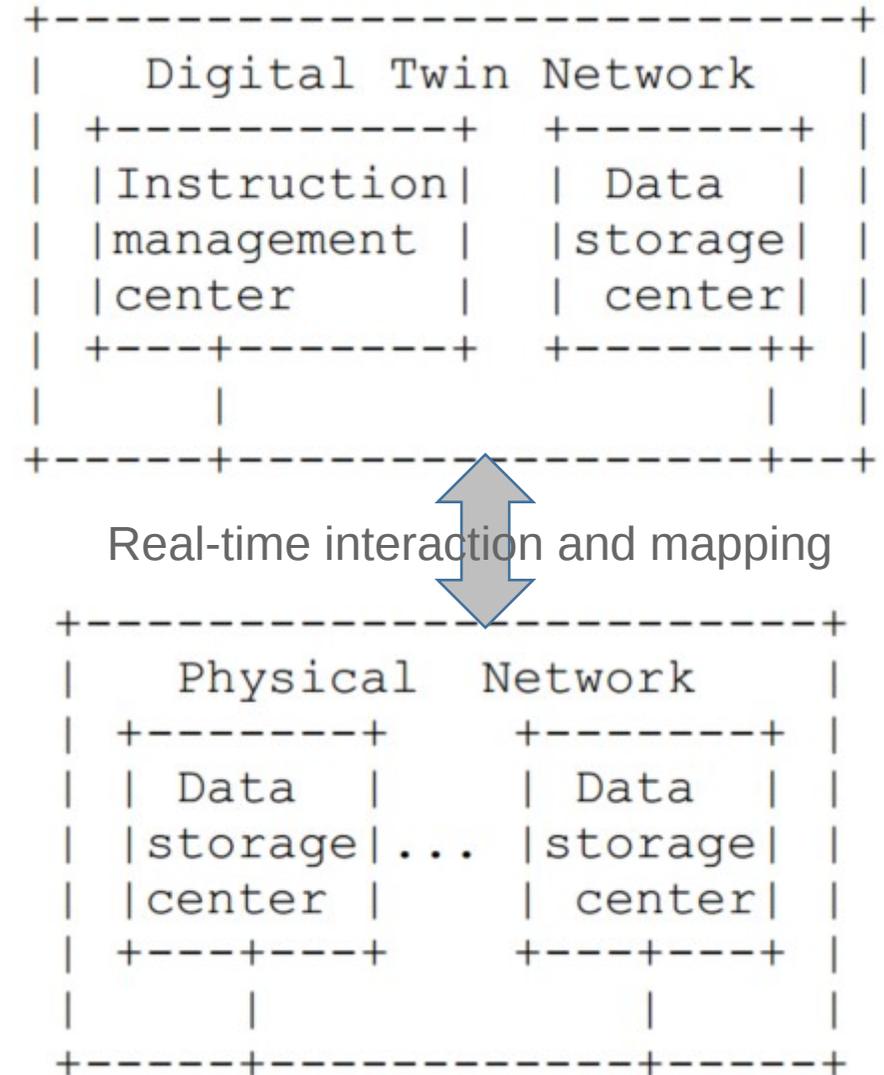
Digital Twin Network (DTN) is a network system with Physical Network and Twin Network, which can be mapped interactively in real time. The construction of DTN requires real-time data of Physical Network to update the state of Twin Network.

Reference: draft-zhou-nmrg-digitaltwin-network-concepts-04

## The disadvantages of the existing method

The existing method collects the full amount of data from the Physical Network for modeling, and does not consider the problems,

- insufficient storage resources
- low computational efficiency
- waste of bandwidth resources caused by data transmission.

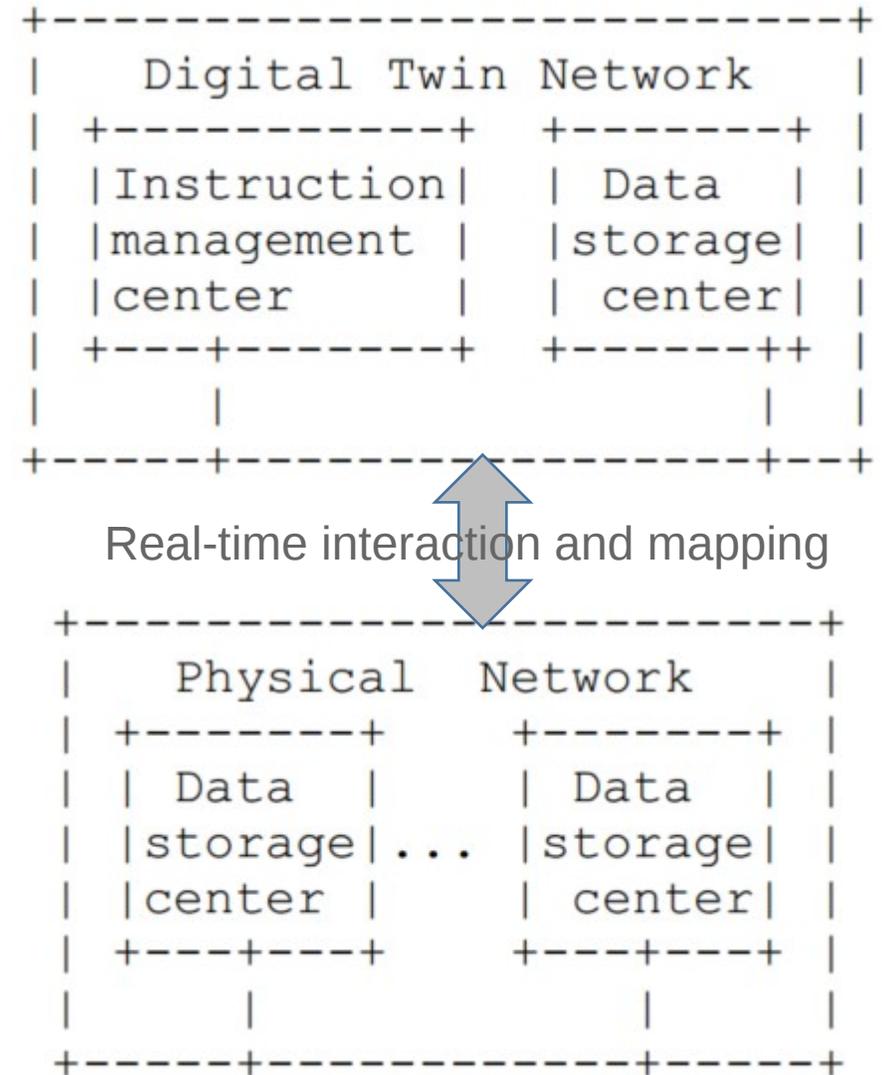


# Overview □

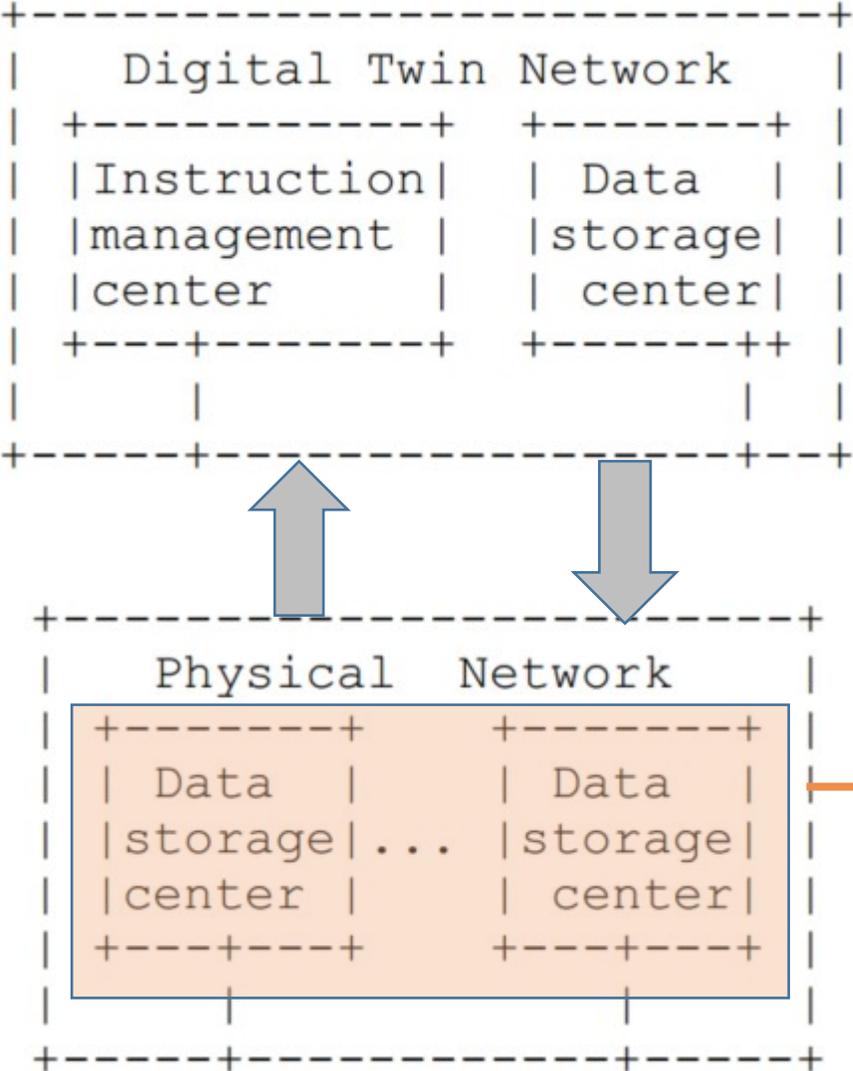
In order to solve these problems, this memo introduces an efficient data collection method for DTN. This data collection method is to send instructions in the Twin Network to the Physical Network to collect data on demand, and then the Physical Network parses and executes instructions such as data cleaning or knowledge representation, and then sends the processed or represented data to the DTN.

# The composition of DTN

- DTN consists of Physical Network and Twin Network.
- The Physical Network includes multiple Data Storage Centers
- The twin Network includes the Instruction Management Center and Data Storage Center.



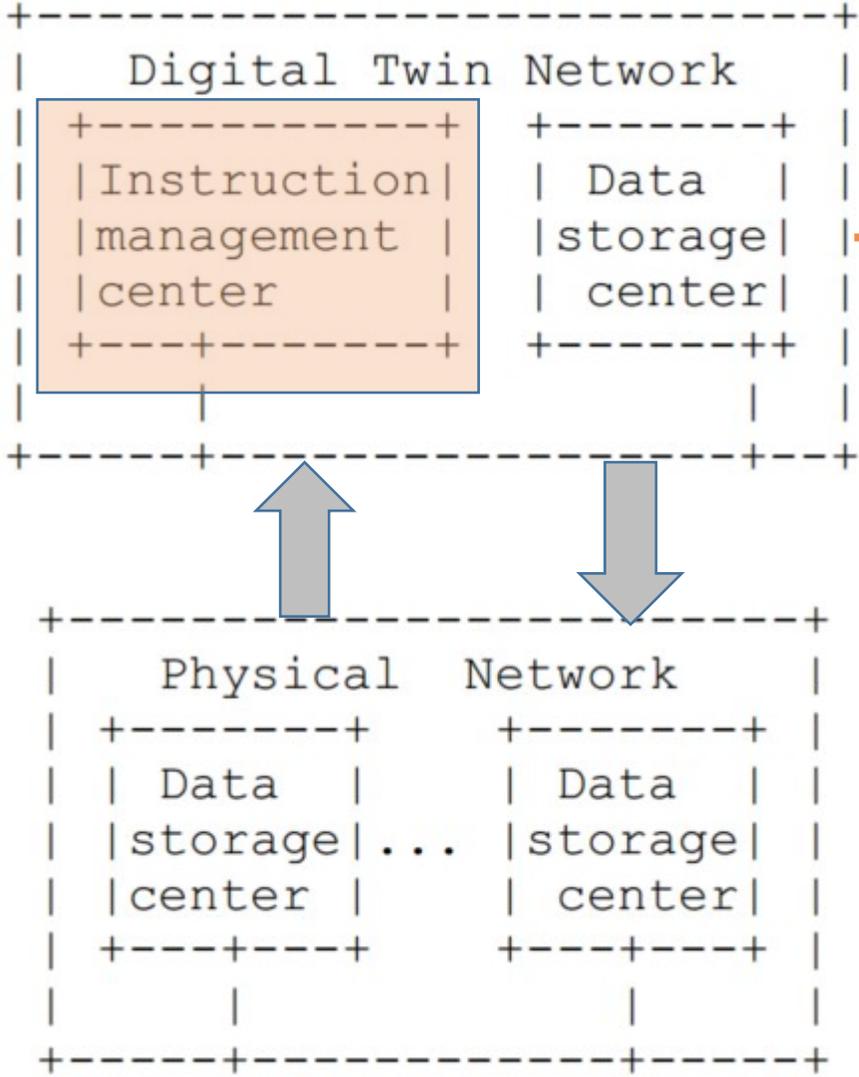
# The introduction of Data Storage Center in the PN



Data Storage Center in the Physical Network has two functions

- It can store data, such as performance indicators, operational status, logs, traffic scheduling, business requirements, etc.
- It has the function of automatically parsing the instructions sent by the Instruction Management Center in the Twin Network.

# The introduction of Instruction Management Center



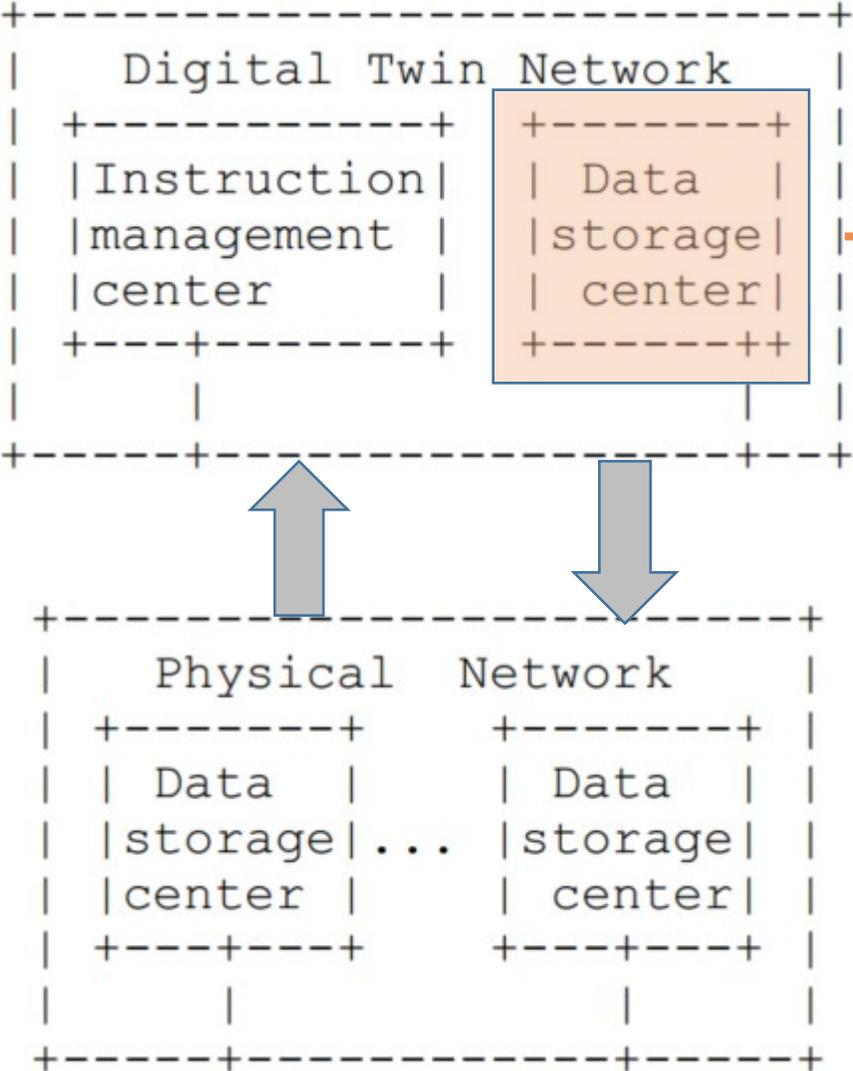
The Twin Network includes the Instruction Management Center and Data Storage Center.

The Instruction Management has two functions.

- It is mainly used to manage the registration of the Data Storage Center in the Physical Network
- 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.

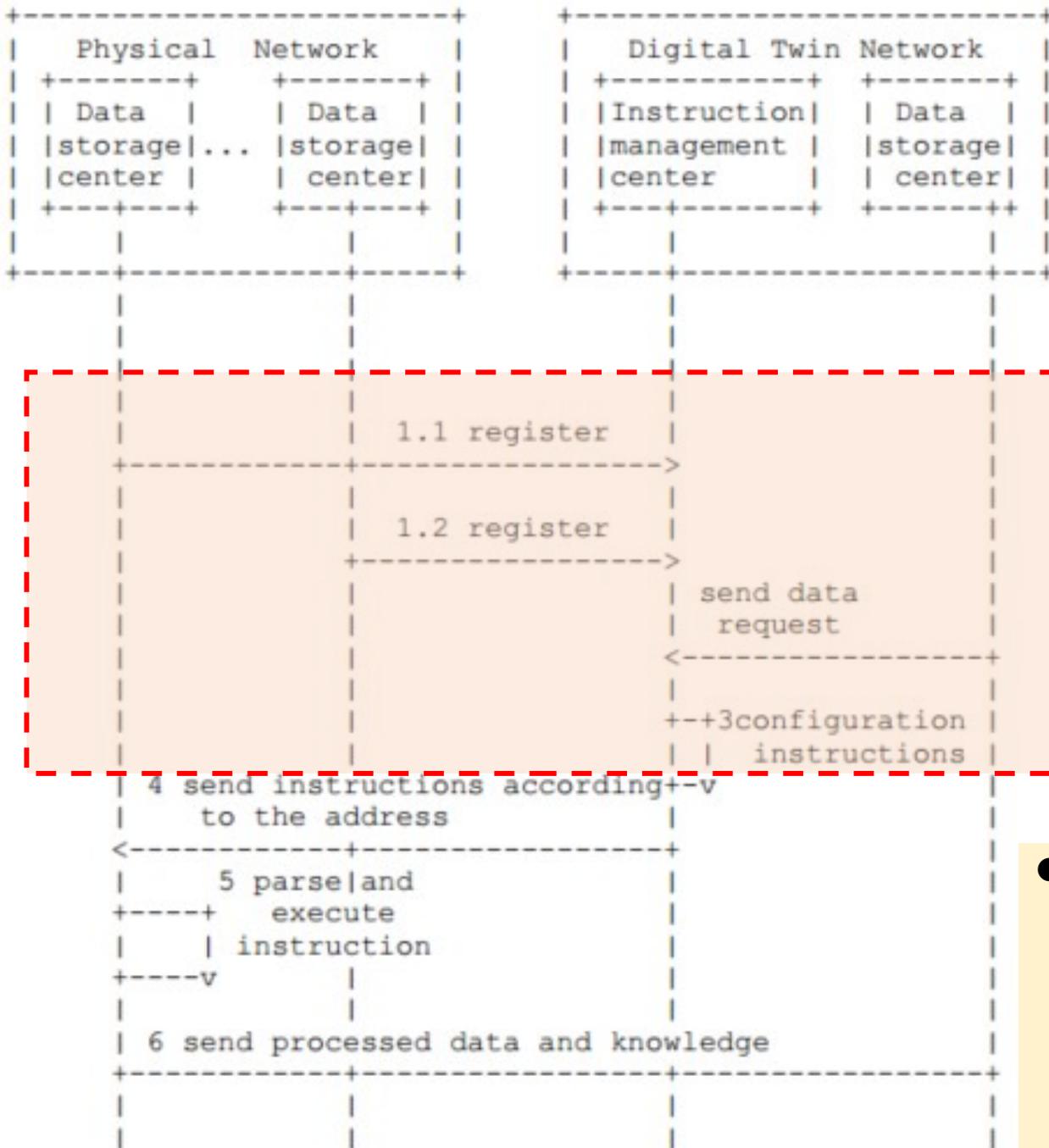
- Registration information include various key information, such as the IP address of the Data Storage Center in the Physical Network, data type, and various index names of the data ,data source name and data size, etc;

# The introduction of Data Storage Center in the TN



The function of The Data Storage Center:

- 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.



The specific process is as follows:

1. 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.
2. The Data Storage Center in the Twin Network sends the data collection request to the Instruction Management Center.
3. 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 specific process is as follows:

4. 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.
5. 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.
6. The Data Storage Center of the Physical Network will send the processed and represented data to the Data Storage Center in the Twin Network.

# Advantages

1. The need for storage resources in the twin network is reduced.
2. The need for computing resources and improve the efficiency of data processing and representation is reduced.
3. The bandwidth resource consumed by data transmission is greatly reduced.

# Next Steps

- Detailed analysis on the instruction-carrying information
- To discover more concrete examples of Data collection
- Consider about the extension of the Data Collection method
  
- **Looking forward to the comments, suggestions and questions.**

Thank you!