SCIENCE CHINA
Information Sciences

• REVIEW •

April 2018, Vol. 61 040303:1–040303:16
https://doi.org/10.1007/s11432-018-9386-5

Special Focus on Deep Space Communications

Network protocol architectures for future deep-space internetworking

Kanglian ZHAO[1] & Qinyu ZHANG[2*]

[1]*School of Electronic Science and Engineering, Nanjing University, Nanjing* 210023*, China;*
[2]*Harbin Institute of Technology Shenzhen Graduate School, Shenzhen* 518055*, China*

Received 27 February 2018/Revised 8 March 2018/Accepted 15 March 2018/Published online 19 March 2018

Abstract In the next two decades, humans are going to experience a grand age of deep-space exploration, especially in Mars and Lunar spaces. These relatively frequent and long-term activities provide the opportunity, and at the same time, demands the necessity for a true interplanetary network as an essential infrastructure for future deep-space exploration. In this study, we try to provide a picture and a perspective in the current network protocol architectures for future deep-space internetworking. We first investigate the recent technical advances for deep-space internetworking and the challenges to their network protocol architecture. Detailed technical characteristics of three effective network protocol architectures are presented. A special focus is casted on delay tolerant networking (DTN), which is a dedicated network protocol architecture for deep-space internetworking. Finally, several open questions in DTN for future deep-space internetworking are proposed for further study.

Keywords network protocol architecture, deep-space internetworking, deep-space communications, space internetworking, delay-tolerant networking

Citation Zhao K L, Zhang Q Y. Network protocol architectures for future deep-space internetworking. Sci China Inf Sci, 2018, 61(4): 040303, https://doi.org/10.1007/s11432-018-9386-5

1 Introduction

Space science and exploration have been one of the most important targets of space agencies since the very beginning of their activities. Although no strict definition of "deep-space" has been provided, the International Telecommunication Union responsible for radio communication (ITU-R) defines the beginning of deep space at 2×10^6 km [1]. Another well-known definition is that any region beyond the cislunar space, including the Moon, is called deep space. Since the first space probe flight to the Moon in 1959, over 200 deep-space exploration missions have already been made to all eight planets, various asteroids, comets in the solar System and beyond. The relationship of Earth with the other planets is better understood with deep-space exploration, and exploration activities will continue to augment our knowledge of the solar system and our universe.

On January 11, 2016, China approved its first Mars program, which is a robotic probe mission to Mars. The mission will launch a Mars probe consisting of an orbiter, a lander, and a rover, in the third quarter of 2020 [2]. Meanwhile, impressive Chinese Lunar missions are also continuing. Aside from China, the National Aeronautics and Space Administration (NASA) of the United States, the European Space Agency (ESA), the Russian Space Agency (Roscosmos), the Japanese Aerospace Exploration Agency

* Corresponding author (email: zqy@hit.edu.cn)

(JAXA), India, and United Arab Emirates have also announced Mars missions around the 2020 time window. On October 8, 2015, NASA published its official plan for human exploration of Mars aiming for a manned surface landing in 2030. SpaceX, a private company that just recently launched a sports car to the Mars orbit in a test flight of the Falcon Heavy rocket, also set up ambitious goals to send its first cargo mission to Mars in 2022 and both cargo and crew in a second mission in 2024. A flourishing age of deep space/Mars exploration is taking off.

Currently, deep-space missions are always supported by dedicated ground infrastructures to guarantee telemetry, tracking and control (TT&C) and data transfer. Deep-space communications are characterized as impaired dynamic links with very long delay and disruption, highly asymmetric channel rates and high error rates. Deep-space nodes can cooperate through space internetworking for better connections, more communication opportunities and broader bandwidth. Moreover, base stations on other planets are planned in different missions which can provide higher communication capabilities. As more and more probes and crews are to be sent into deep space in the next two decades, the current ground-based deep-space communications are limited to support such a large number of nodes in deep-space, providing opportunity and demanding the necessity for space internetworking [3]. With the rapid increase in the capability of on-board computation and inter-satellite link, space internetworking is getting increasingly feasible and flexible for practical implementation. Consequently, future space mission scenarios are envisioned to involve deep-space internetworking with data flowing across multiple hops and over multiple paths to achieve end-to-end data transfers.

This evolution of space exploration scenarios toward more complex communication networks is worth noting to ensure reliable communications for common TT&C messages along with data transfers (e.g., images and files) with reliable and efficient network protocols [4]. However, commercial terrestrial network protocols, especially the end-to-end transmission control protocol (TCP), are not suitable for direct application in space networks [5, 6]. No matter how fast the broadband links provided by Ka band or optical links are, impairments in the network protocols can result in a deteriorated goodput performance [7] or even in a network failure [8]. Network protocol architectures for deep-space internetworking will be a key technology to support future deep-space explorations.

This study tries to provide a picture and a perspective on various network protocol architectures for future deep-space internetworking, also known as interplanetary internet (IPN). The remainder of this paper is organized as follows. Section 2 presents the recent technical advances and challenges in deep-space internetworking. Section 3 provides information about the standardization in network protocol architectures in space internetworking. This section also analyzes the detailed technical characteristics of the three current effective standardized network protocol architectures (i.e., TCP/IP architecture, space packet protocol (SPP)-based architecture, and delay-tolerant networking (DTN) architecture). Section 4 proposes several open questions in network protocol architectures for future deep-space internetworking. Finally, Section 5 concludes the whole paper.

2 Technical advances and challenges in deep-space internetworking

The idea of deep-space internetworking, or IPN, was proposed in the end of the last century by researchers from the terrestrial Internet community to adapt commercial Internet technologies to support the communication needs of space exploration [9], which inspired the research on the network protocol architectures for space internetworking [10–13]. This section presents the technical advances and the challenges in deep-space internetworking. Section 3 will describe the detailed network protocol architectures for space internetworking.

Since 1963, NASA has built the most complete and advanced communication infrastructure supporting its own and its partner's deep-space exploration missions. This infrastructure is known as NASA's Deep Space Network (DSN). The DSN supports both Earth orbiting and deep-space science missions with three ground stations located approximately 120° apart on Earth. The ESA also has similar deep-space ground station networks supporting European deep space missions [14]. The DSN is part of NASA's

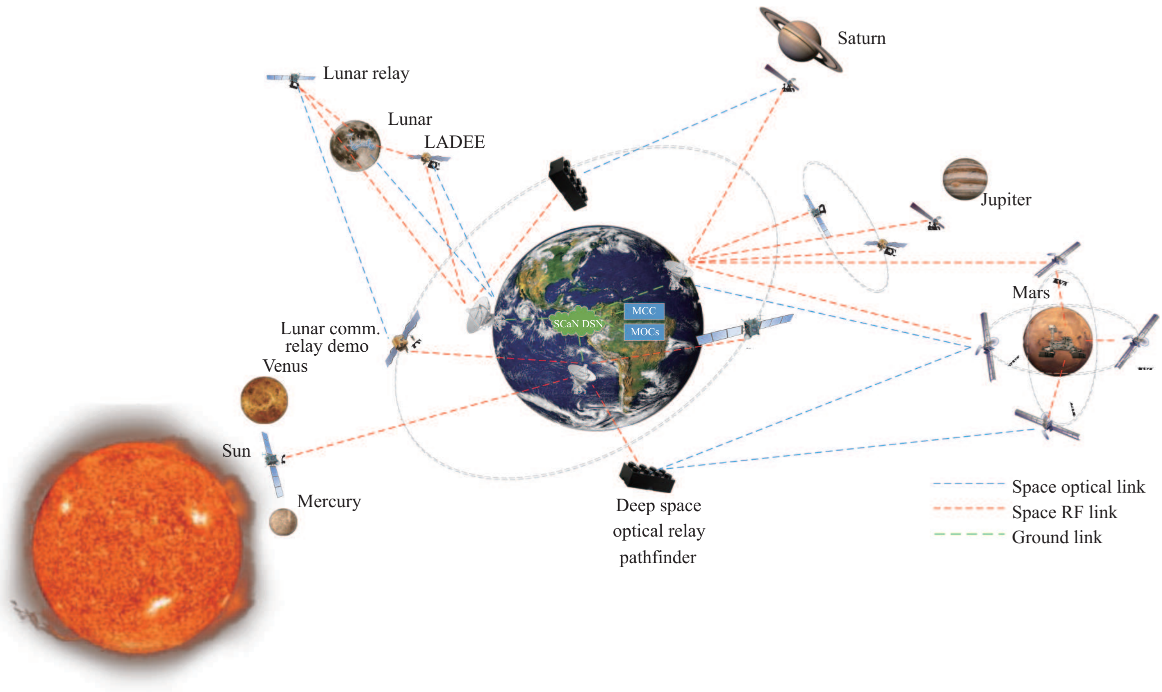

Figure 1 (Color online) SCaN integrated communication architecture for solar system exploration.

entire communication and navigation systems, which consists of three separate networks. Aside from the DSN, the other two networks are the Near Earth Network (NEN) and the Space Network (SN). The NEN supports non-deep-space missions with ground stations all over the world. Meanwhile, the SN, which is also known as the Tracking and Data Relay Satellite System (TDRSS), currently consists of ten geosynchronous satellites and two ground facilities to provide 24×7 coverage for Earth-orbiting vehicles. Among the three tracking networks, the DSN provides the command, telemetric, and tracking services to many deep-space missions. In 2007, the NASA Space Communications and Navigation (SCaN) office was set up to unify the management of NASA's disparate networks. Figure 1 depicts the SCaN integrated communications network's architecture [15]. With the development of high-speed inter-satellite link technology and the increase of various space nodes, deep-space probes, satellite relays, and ground stations are being connected to support space science and exploration missions. With the limited capability of on-board processing, current space relays are all in "bent-pipe" mode and do not provide on-board switching or routing services.

2.1 Advances in the deep-space communication technology

Some major advances in the deep-space communication technology have recently been made: antenna array-based ground station architecture, deep-space optical communications and the capability of on-board processing which is especially important for a network protocol architecture. The antenna array-based ground station employs an array of smaller antennas instead of a single large antenna, which can achieve an equivalent or a larger antenna gain with a lower cost [14, 16].

The fast development of deep-space optical communications has two main motivations.

The first one is that the radio frequency used for space communications diverges or spreads largely over the deep-space distances [17], which considerably weakens the signal for reception. Optical waves in much higher frequencies have less beam spread for better reception and detection, which also means higher data/power efficiency (higher data rates for the same transmission power).

The second reason is that a higher directivity of the optical beam also allows a theoretically infinite spectrum, which is precious for the radio frequency (RF) band. We did not provide details on the development of deep-space optical communication technologies herein. Interested readers can refer

Table 1 Mission concepts driving the next generation SCaN architecture [22]

Key changes in mission concepts	Impact on next-generation architecture
Human explorers return to cis-lunar space and eventually reach Mars with steadily increasing surface capabilities	• Lunar and Mars networks and services resemble Earth network and services • Provide sufficient deep-space communications and tracking capacity for human and robotic missions
Planetary missions with robotic sample return to Earth — followed by human exploration and return	• Added complexity in mission definition and navigation; planetary orbit and position determination needs accuracy beyond Global Positioning System (GPS)
Increasingly capable compact spacecraft that lead to missions consisting of larger clusters and fleets with distributed capabilities	• Increasing quantity of spacecraft simultaneously needing service and more autonomous operations • Services to disadvantaged and compact spacecraft impose high burden on networks
Missions continuing to increase their need for temporal, spatial resolution in science measurements and reduce data delivery latency	• Increasing near Earth capacity (10-100x) and deep-space capacity (100-1000x) • Balanced capacity between lower cost ground stations and higher cost space relays driven by mission latency
Continued focus on mission affordability through collaboration with external partners	• Increasing need for secure cross-support with domestic and international partners that is interoperable and easy to arrange
New space entrepreneurs establishing new markets in space increasing demand for non-government communication and navigation services	• Industry drives needs for increasing capacity and interoperability with reduced cost of services
Increasing sophistication and diversity of mission design and operations concepts as mission complexity and goals increase	• Need for faster, less labor-intensive network support for service negotiation and mission design • Increasing need for service flexibility and rapid response to mission requests during operations • More agile operations and development process

to [17,18]. Free space optical communications are highly sensitive to weather and atmospheric turbulence, especially cloud blockage, which causes bad link conditions or even link disruptions [19].

In January 2016, the NASA Glenn Research Center, on behalf of NASA SCaN office, started the "NASA Next-Generation Space Relay Architecture Concept Study" project to define a top-level end-to-end next generation architecture concept that evolves NASAs space communication and navigation networks, which is broadly defined over the next 25 years to the 2040 timeframe and more specifically defined over the next 10 years to the 2025 timeframe when its initial capabilities will start to become operational [20–22]. The future mission concepts range from planetary science missions to human exploration to fleets of mission spacecraft performing coordinated science investigations. To enable these missions, the SCaN network must transition and become part of an end-to-end "system of systems" which includes many elements from research spacecraft, mission control and science processing centers, network infrastructure, partner agencies both domestic and foreign, and commercially provided services and partnerships. One of the key technical advancements enabling such a transition is the advanced capability of on-board processing, which makes it possible for on-board switching and routing in deep space. The mission concepts driving the requirements of the next-generation architecture are summarized in Table 1 by [22].

Dedicated planetary relays, such as Mars relays, will provide a near-continuous trunk link availability to Earth. The trunk lines to Earth will operate using RF (Ka- and X-band) up to 125 Mbps and with optical links up to 300 Mbps with forward links operating at 50 Mbps. Each relay orbiter communicates with the mission orbiters (science or exploration spacecraft) and surface vehicles (e.g., habitats, communications stations, landers, and rovers) via proximity links, whose maximum data rate for each mission is nominally 50 Mbps (considering either Ka-band or optical) [22].

New space relay satellites with on-board processing/routing capabilities display traditional bent-pipe relays; hence, the data transfer between any two nodes on Mars or on the Moon can only require proximity links as shown in Figure 2, which can meet the increasing needs for various science missions and future human missions. The legacy infrastructure will continue during the transition toward future architectures (Figure 3).

By comparing Figures 1 and 2, we can better understand the main difference between the current

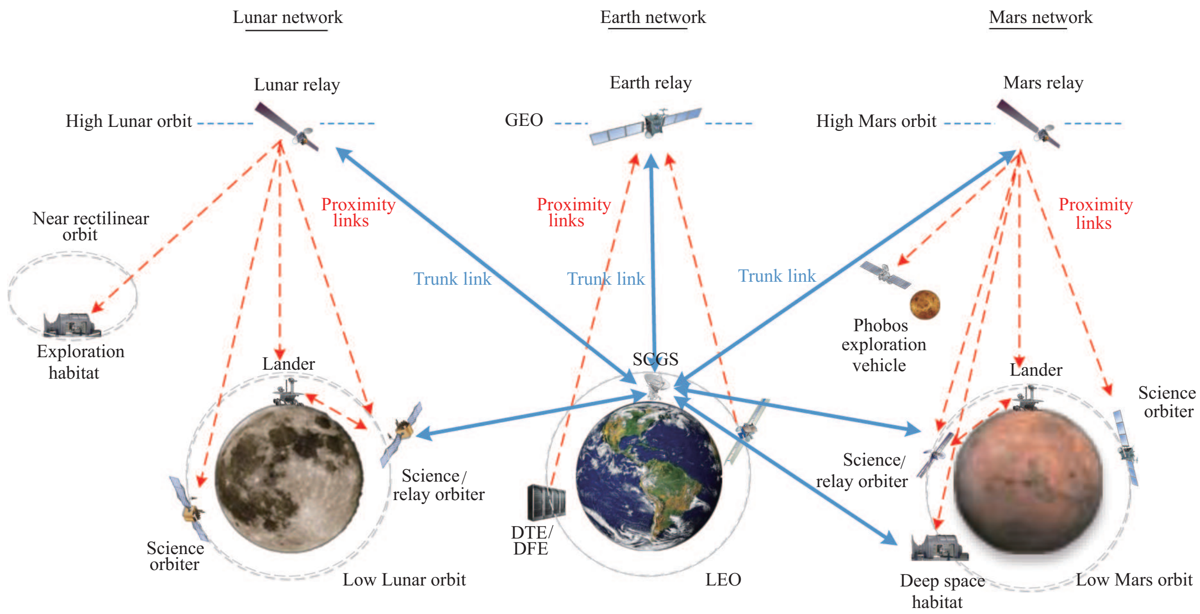

Figure 2 (Color online) Near Earth, Mars, and Lunar network concept with next-generation space relays.

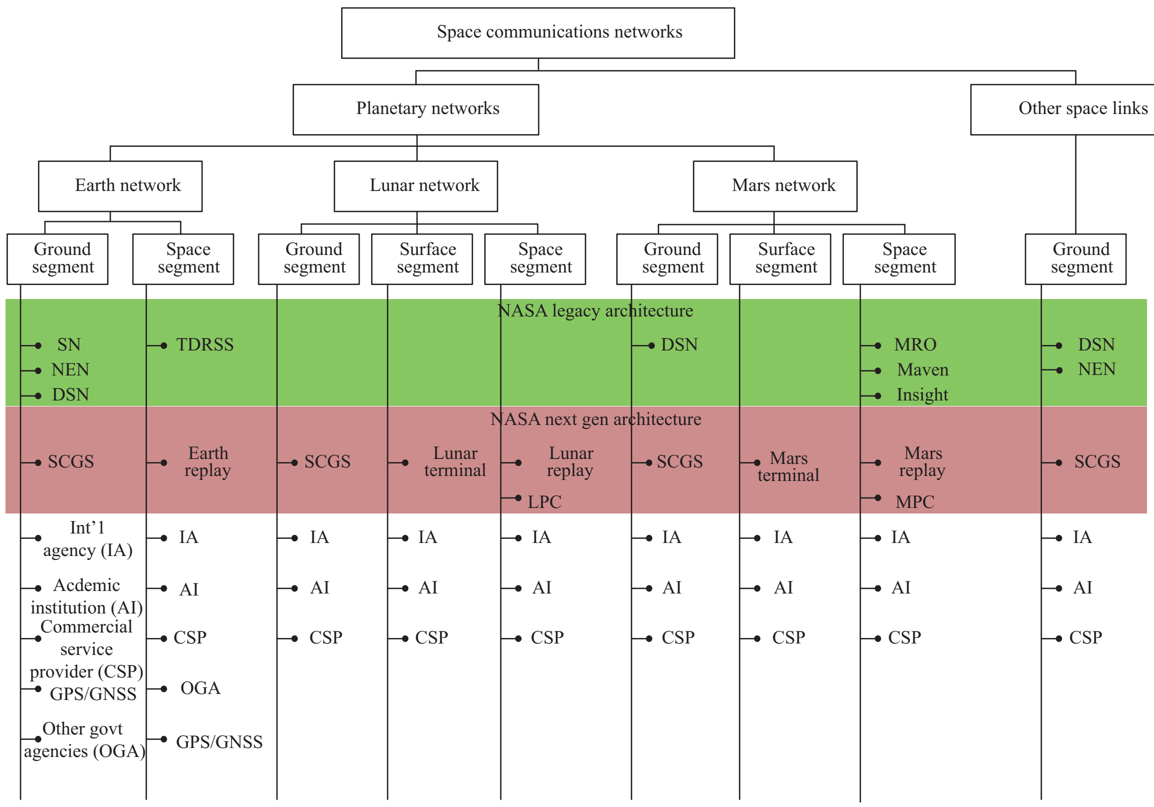

Figure 3 (Color online) Evolution of NASA's SCaN network infrastructure from 2015 (green) to 2040 (red).

network infrastructure and the next generation infrastructure in 2040. The infrastructure today is a ground-based DSN-centric network with deep space nodes attached through direct links or bent-pipe relay links, which can be named as "Earth network with deep-space access". In this case, the Earth-based DSN is the backbone network, and the deep-space trunk links blong to the access networks. While the next-generation SCaN network is a true "network of planetary networks" (Figure 2), in which the planetary relay satellites (Mars relays or Lunar relays), the Earth-based DSN, and the deep-space trunk

Table 2 Distance and relative difficulty (1/distance^2) for deep-space communications

Place	Distance (km)	Difficulty
GEO	4×10^4	Baseline
Moon	4×10^5	100
Mars	3×10^8	5.6×10^7
Jupiter	8×10^8	4.0×10^8
Pluto	5×10^9	1.6×10^{10}

links form the backbone network. All the planetary networks (i.e., Earth network, Mars network and Lunar network) are homogeneous in the next-generation infrastructure.

2.2 Constraints and challenges in deep-space internetworking

In implementing the "network of planetary networks", we are facing various specific constraints and challenges [9–13, 15] that will finally affect the efficiency, reliability and flexibility in different layers of the network. We first categorized these constraints into four different levels, namely node, link, network, and operation levels.

(1) Node level. Space nodes (networking payloads) are accommodated in satellites or other dedicated space platforms; hence, both hardware and software constraints exist in implementing the network functions.

• Constraints on hardware. The size weight and power (SWaP) constraints, system complexity and reliability, and space qualification costs [23] are always in first place for the design of space systems, which have typically limited the storage capacity and the capability of processing and routing on-board. Moreover, it is still not currently possible for hardware upgrade during the missions' lifecycle, indicating that the hardware should keep working in space for years to decades.

• Constraints on software. Dedicated Real-time operating systems (OS) [24, 25], commercial or customized, are adopted for space missions. The development of on-board software, from drivers to various applications, is usually hardware platform-dependent and must conform to the flight software qualification. High development costs are expected, and the open source strategy is still not as popular as in the terrestrial software community. All these conditions result in limited resources and a slow progress for the flight software development.

(2) Link level. The extremely long distance between two nodes (e.g., from Earth to Mars) in deep-space causes most of the problems in the link level. Table 2 shows some examples of the distance in deep space.

• Extremely long propagation delays. Figure 4 presents some examples to illustrate the large propagation delay in deep-space communications.

• High error rate. Table 2 shows the relative difficulty of deep-space communications. A lower receiving power results in a higher error rate.

• Limited data rate. A low receiving power also limits the transmission data rate.

• Asymmetric channel rates. The rate between the forward and backward links in deep-space communications might be as high as 1 : 1000 [26].

• Intermittent connectivity. Predictable long disruptions happen because of orbital mechanics. Unpredictable disruptions may be caused by node failure, channel blockage, and bad link state.

(3) Network level. When multiple deep-space nodes and ground stations are connected with links, the network we are interested is formed and brings new challenges in the network level.

• Dynamic topology. Aside from user mobility and unpredictable node failure, the topology of the whole network is continuously evolving with time because of orbital mechanics.

• Heterogeneous links and protocols. The proximity links in planetary networks and the chunk links in deep-space backbone networks have totally different link level characteristics and employ different physical/link layer protocols, thereby challenging the network protocol architecture.

(4) Operation level. We should notice that all the networking nodes are also science equipment for deep-space exploration, which are constrained by operation rules and strategies.

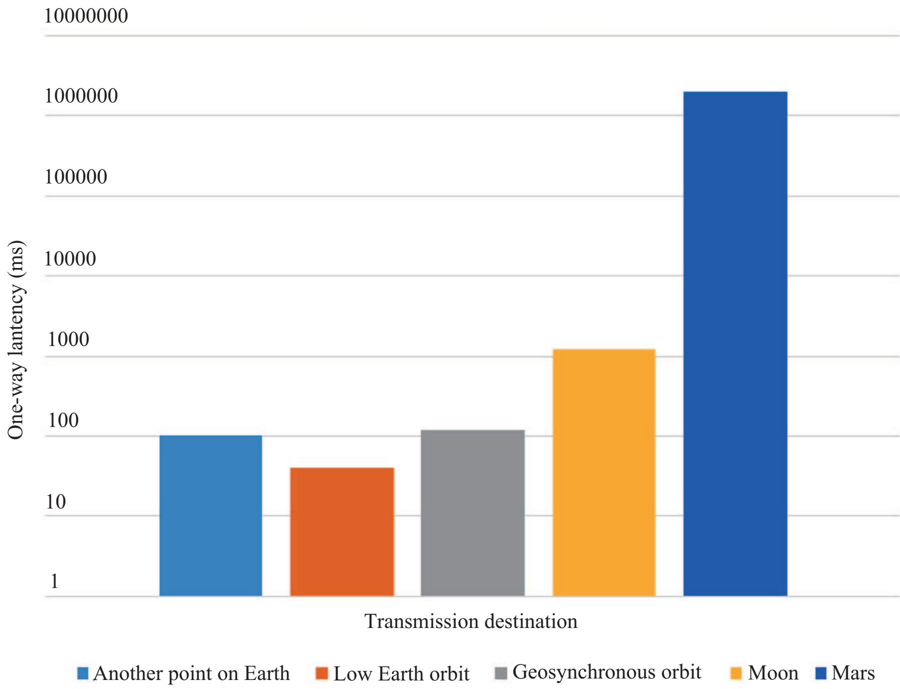

Figure 4 (Color online) Maximum latency (transmit from Earth).

- Backward compatibility. Compatibility with functional legacy systems means more heterogeneous systems.
- Scheduled connection. With limited access and directional connectivity, the connections in deep-space communications are scheduled in advance. The connections are sometimes not established because of operational assignment.
- Centralized management. In the current "Earth network with deep-space access" network infrastructure, the resource and network management is conducted in a centralized manner. As explained in Section 2, future space relays will provide more powerful on-board processing and routing capabilities, which might introduce distributed network functions instead of the current centralized controlled static mechanism.
- Multi-party cooperation. The network security should be considered with international or domestic cooperation to provide the network infrastructure [27].

Figure 5 shows the mapping between the mentioned constraints and the related network protocols or functions in different layers.

3 Network protocol architectures for deep-space internetworking

A network protocol architecture is the design of layered protocols in a communication network, which specifies the physical components of the network and the functional organization and configuration, as well as the data formats used in common. The Open Systems Interconnection model (OSI model) might be the most famous network protocol architecture standardized by both the International Organisation for Standardization (ISO) and the Telecommunications Standardization Sector of the International Telecommunication Union (ITU-T). In the field of Internet, the TCP/IP architecture is the de facto Internet architecture with protocols standardized by Internet Engineering Task Force (IETF) [28]. The Consultative Committee for Space Data Systems (CCSDS) is the most important standard-developing organization (SDO) in the field of space internetworking.

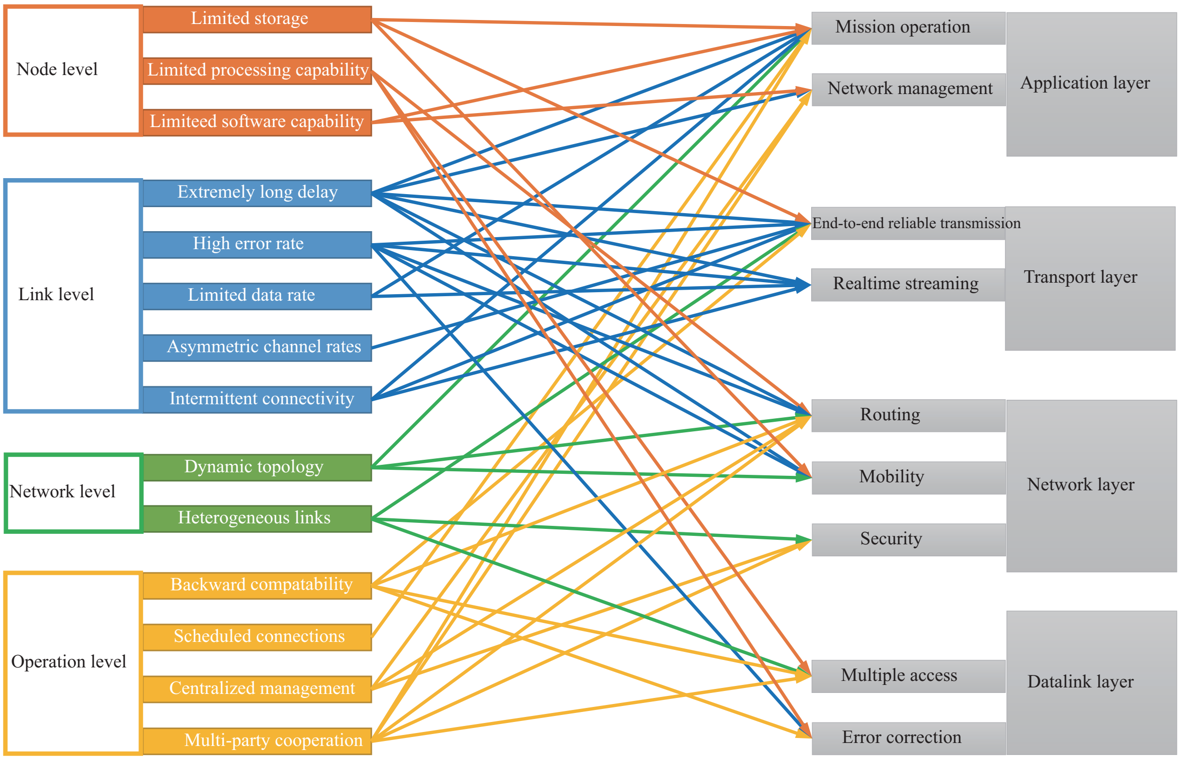

Figure 5 (Color online) Mapping between the mentioned constraints and the related network protocols or functions in different layers.

3.1 Standardization for space internetworking

The CCSDS was formed in 1982 by the major space agencies of the world. It provides a forum for discussion of common problems in the developing and operating space data systems. Since its establishment, the CCSDS has been continuously developing recommendations for data- and information system standards to promote interoperability and cross support among cooperating space agencies and enable a multi-agency spaceflight collaboration and new capabilities for future missions. In 1990, ISO Technical Committee 20 Subcommittee 13 (ISO TC 20/SC 13) was formed and designated the "Space Data and Information Transfer Systems". ISO TC 20/SC 13 is the ISO administrative subcommittee of the CCSDS. Through a special arrangement with the ISO, the CCSDS documents were processed as ISO TC 20/SC 13 projects at the Draft International Standard (DIS) stage. A total of 84 CCSDS recommendations are currently published as ISO standards, while 11 recommendations are under development within ISO TC 20/SC 13.

The technical organization of the CCSDS are divided into six areas, in which the Space Internetworking Service Area (SIS) is fully dedicated on SN protocol architectures. The rest of the technical areas include Space Link Service Area (SLS), Cross-Support Service Area (CSS), Space On-board Interface Service Area (SOIS), Mission Operations and Information Management Services Area (MOIMS), and Systems Engineering Area (SEA), which are all network related. The CCSDS provides data formats and the characteristics and specifications of the physical and transport layer [29]. More than 800 space missions are already complying with the CCSDS recommendations.

The CCSDS also recommends network protocol stack for space internetworking which is shown in Figure 6. Four different network protocol architectures can be found in the CCSDS stack, (i.e., the CCSDS Space Communications Protocol Specifications (SCPS) architecture [30], the Space Packet Protocol (SPP) [31]-based architecture, the IP architecture, and the Delay/Disruption Tolerant Networking (DTN) architecture [12]). The SCPS architecture is a space-dedicate TCP/IP-alike network protocol architecture [30], which consists of the SCPS File Protocol, Transport Protocol (TP), Network Protocol,

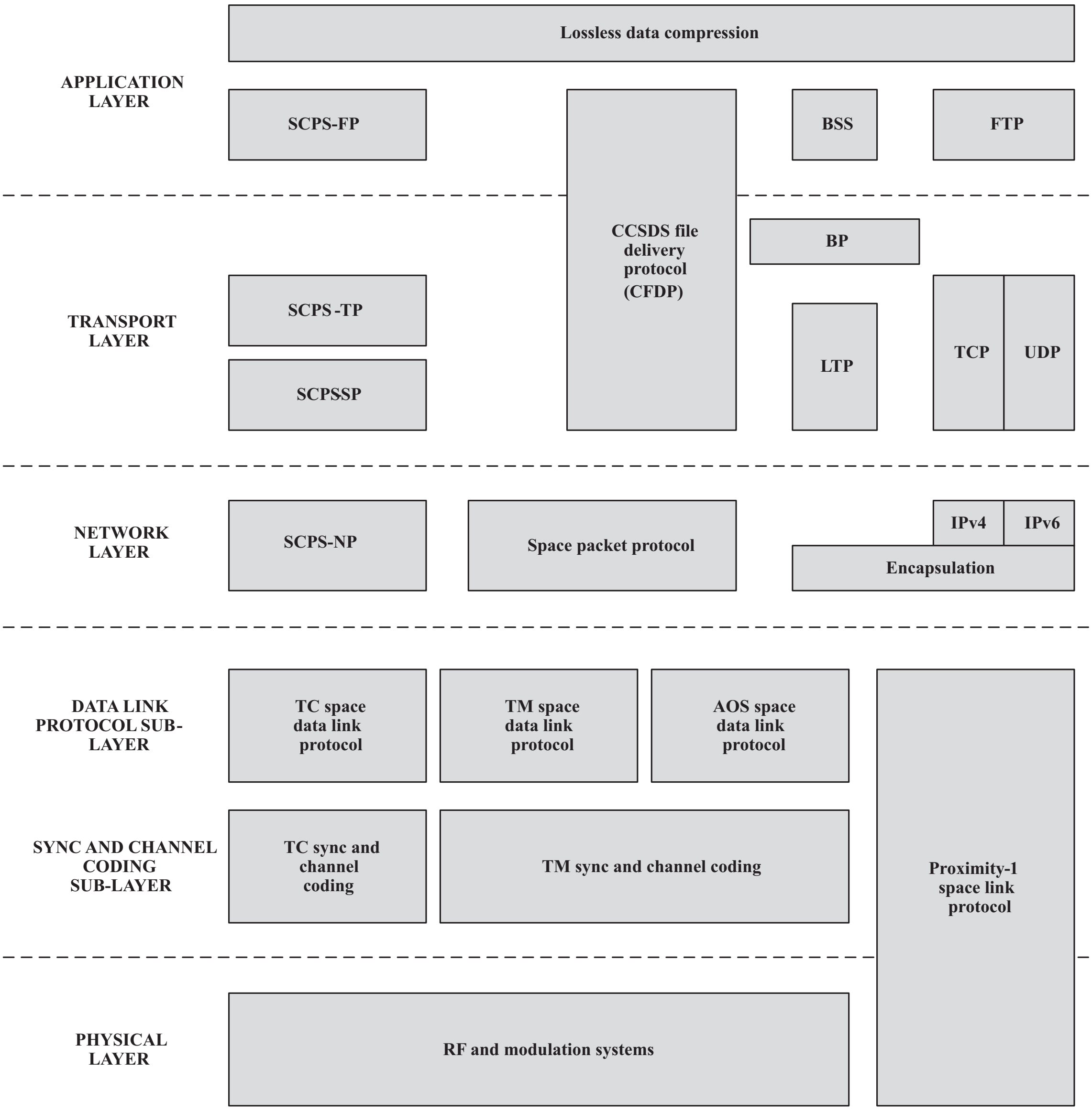

Figure 6 CCSDS network stack.

and Security Protocol. Almost no adoption can be found in real missions, hence, the SCPS suites have been downgraded to historical (silver book) status since 2016. Only the SCPS-TP is being maintained by the CCSDS as an active recommendation (Blue Book) [32] because it was taken up by the commercial industry as a TCP enhancement, and is still in use. Therefore, we will only discuss herein the details of the other three effective network protocol architectures in the next sections.

Other space internetworking relevant SDOs include ITU-R, European Telecommunications Standards Institute (ETSI) and IETF. ITU-R not only standardizes the rules for RF spectrum allocation, but also develops a standard for satellite communications and networking. ETSI also works in the field of satellite communications and networking. It developed the well-known digital video broadcasting series standards including DVB-S, DVB-S2, DVB-RCS, and DVB-RCS2 standards.

Although IETF focuses on Internet protocols, the DTN Working Group (DTNWG) has started the standardization process of the DTN in IETF since 2016. The first research efforts for the DTN were also initiated in the Internet Research Task Force (IRTF) Delay-Tolerant Networking Research Group in 2002.

Figure 7 presents the space internetworking-related SDOs and some research and development activities

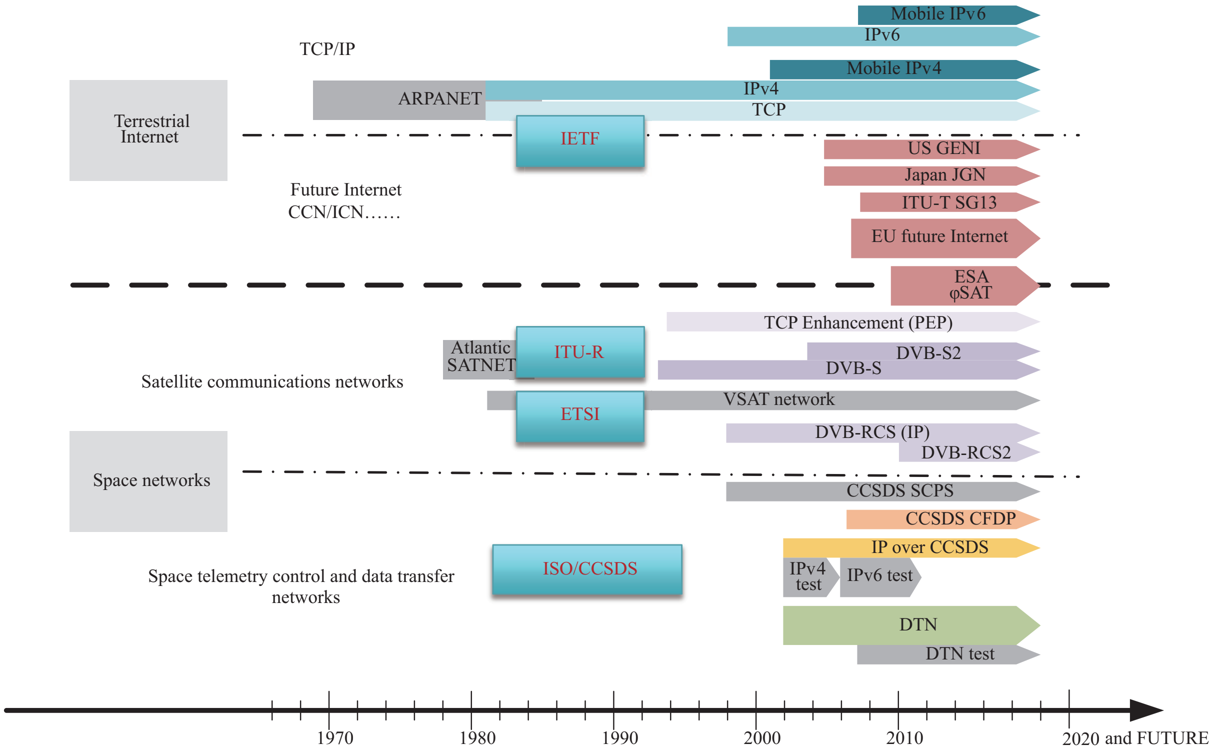

Figure 7 (Color online) Different SDOs and their working areas relevant to space internetworking.

in their relevant areas.

Moreover, deep-space networking needs to comply with other standards like the National Telecommunication and Information Administration standard, which describes the rules governing the RF spectrum use in the USA by government agencies [29].

3.2 TCP/IP architecture for proximity networking: IP over CCSDS

The CCSDS recommendation for IP over CCSDS Space Links [33] (IPoC) specifies the implementation of the IP Protocol over CCSDS Space Data Link Protocols (SDLPs). IP Protocol Data Units are transferred by encapsulating them, one-for-one, within CCSDS Encapsulation Packets. The Encapsulation Packets are directly transferred within one or more CCSDS SDLP Transfer Frames. This method uses the CCSDS Internet Protocol Extension convention [33] in conjunction with the CCSDS Encapsulation Service [34] over CCSDS SDLPs: Telecommand (TC) [35], Telemetry (TM) [36], Advanced Orbiting Systems (AOS) [37], and Proximity-1 (Prox-1) [38].

With IPoC, commercial-off-the-shelf (COTS) IPv4 or IPv6 technology can be directly implemented for space applications. IP is still the "narrow waist" in IPoC, which extends terrestrial TCP/IP architecture into space. Hence, the same upper layer protocols and applications can be adopted in space missions. This has already been proven to be cost-effective in near-Earth scenarios [39]. With the space-optimized variants of the TCP, a good data transfer performance can also be achieved. However, considering deep-space scenarios with extremely long delays not only fails the TCP, it also fails the routing protocols relying on the TCP (e.g., Boarder Gateway Protocol [12]). This is the reason why researchers started to seek for new network protocol architectures for deep-space internetworking instead of TCP/IP.

3.3 Space packet protocol based architecture for deep-space internetworking

The SPP is an IP-alike packet-based network layer protocol. As shown in Figure 6, the SPP is the new "narrow waist" in this architecture instead of the IP. Various CCSDS-recommended datalink layer

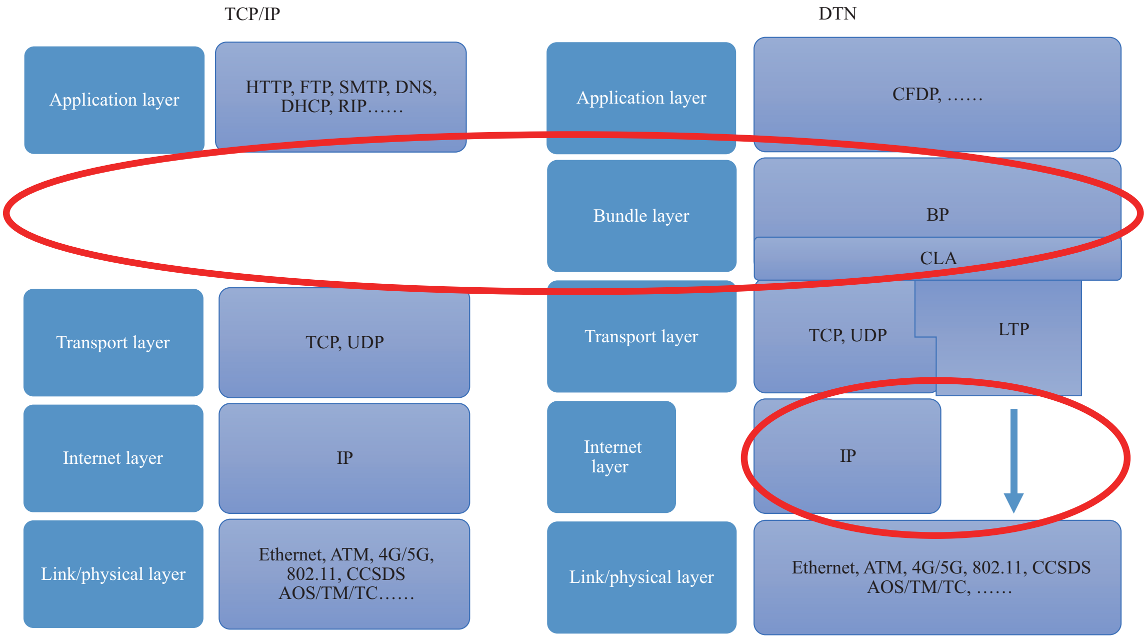

Figure 8 (Color online) Comparison between the TCP/IP network protocol architecture and the DTN architecture.

protocols are underneath and converged by the SPP. Above the SPP, the CCSDS File Delivery Protocol (CFDP) [40] currently exists for deep-space internetworking [41–46]. With the SPP in network layer handling naming and routing, the CFDP accomplishes reliable file transmission over deep-space channels [41]. Various experiments were conducted to evaluate the performance of the CFDP [42–44], among which a true DSN communication experiment was performed on July 4, 2005 to verify the performance of the CFDP in the Deep Impact mission [44]. Further studies were continued to improve the performance of the CFDP. Erasure coding [45] and packet interleaving [46] were proposed to be combined with the CFDP for a higher data transfer efficiency.

As a space-customized equivalent of the IP with very limited users, the SPP does not receive much attention. Moreover, as a new network protocol for heterogeneous SNs, it definitely brings complicated gateways for translation which is not welcome. Based on these reasons, this new architecture also failed to obtain popularity in deep-space internetworking. The CFDP has been transformed into an application layer protocol in the DTN architecture with evolution.

3.4 Delay/disruption tolerant networking for deep-space internetworking

In comparison to the conventional TCP/IP architecture, the DTN architecture is designed to provide communications in highly stressed environments characterized by long or variable delays, intermittent loss of link connectivity, high error rates, and asymmetric data rates [6]. Originally developed for the IPN, the DTN was extended to terrestrial rural networks, wireless sensor networks (WSNs), vehicular networks, and underwater communications networks because these networks also suffer from link disruption and delay [47].

3.4.1 *DTN architecture*

Figure 8 shows a comparison of the DTN architecture and the TCP/IP architecture. A new Bundle Protocol (BP) [48] was introduced in the DTN architecture, which was the new "narrow waist" in the DTN. The communication assets on which BP engines run (analogous to the hosts and routers in an IP-based network) are termed as DTN nodes. A DTN region is defined as a set of DTN nodes that can communicate among themselves using a single common protocol family that is suitable for the networking environment, in which all of the nodes must operate. The DTN network protocol architecture follows

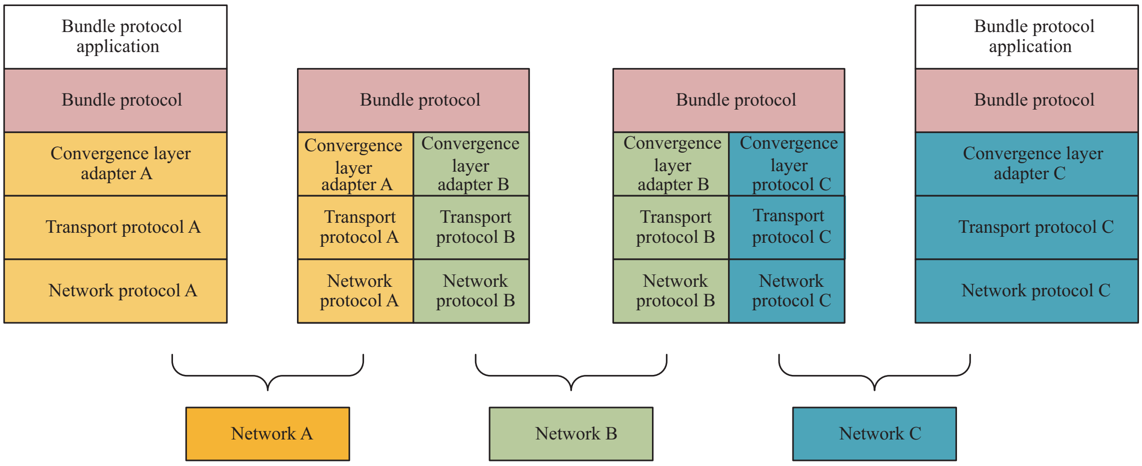

Figure 9 (Color online) DTN architecture, the bundle protocol, and protocol stack in a network.

two basic design principles: (1) hop-by-hop transport [49], and (2) overlay network. With the hop-by-hop transport, each DTN node, end or intermediate, takes a store-carry-forward strategy to efficiently cope with the disruptions and the dynamics in each hop. This strategy is entirely different from the end-to-end design principle of the TCP/IP architecture. The fundamental notion behind the end-to-end principle is that for two processes communicating with each other via some communication means, the reliability beyond a certain margin should only be achieved by mechanisms such as positive end-to-end acknowledgments and retransmissions in the end hosts, rather than in the intermediary nodes, especially when the latter are beyond the control of, and not accountable to the former. However in deep-space communication links with a long transmission delay and a high error rate, the end-to-end principle can significantly decrease the transport efficiency [12, 15]. Nonetheless, the end-to-end route might not even exist sometimes because of orbital mechanism. Note that a hop between two DTN nodes can be a direct link or a region in which there might be intermediate nodes might be operating based on other protocols without BP support. These protocols designed for use within various local regions already exploit whatever favorable circumstances the region offer while operating within their constraints. With the second design principle, BP, which is one layer higher in the stack, is enabled to rely on the capabilities of these protocols in different regions and perform any required additional functions that the local protocols typically cannot. For example, the Licklider Transmission Protocol (LTP) is designed as a reliable local transport (optional) dedicated for long-delay deep-space links, while TCP is a reliable local transport for short proximity links. BP can exploit each transport's capability in its dedicated environment. Moreover, as an overlay, BP provides a unified layer through the whole heterogeneous complex network. Figure 9 shows an example of the DTN architecture across a network. The convergence layer serves as a software adapter to different underlay networks.

3.4.2 *DTN functionality*

According to the design principles, the main structural elements of the DTN are as follows [12].

Tiered forwarding. As mentioned before, the DTN architecture generally relies on regional protocols. The forwarding of bundles among DTN nodes that are in different regions is performed by BP, while the forwarding of bundles in the regional network relies on the regional network layer protocols, such as IP in Internet-like regions.

Tiered naming and addressing. For a bundle to reach its destination within a given region, the source and destination expressions of bundles are defined as concatenated identifiers termed as tuples comprising both region identifiers and regional destination identifiers that can be mapped to regional addresses (or equivalent): {region ID, regional destination identifier}. Regional destination identifiers are late bound; that is, they are mapped to regional addresses (or equivalent) only upon arrival at the destination region, rather than at the time of original transmission.

Tiered routing. The forwarding performed by BP supports new routing protocols different from the routing protocols operating in the network protocols within regional networks. Contact graph routing (CGR) is currently employed by BP for routing calculation, which is sensitive to future contacts (link establishment opportunities) and recalculates the routing choice based on the current contacts' status.

Tiered ARQ. The DTN architecture depends on regional TPs, such as TCP and LTP, for an assured transmission of bundles among DTN nodes within each regional network. BP implements an additional ARQ mechanism, called "custody transfer" (optional). A node that explicitly "takes custody" of a bundle guarantees that it can and will devote sufficient resources to retain a copy of the bundle until some downstream node subsequently takes custody of it. This enables custodial retransmission in case no such notice of custody transfer arrives.

Tiered congestion control. Access to links in the DTN network was assumed scheduled and controlled, hence, the competition for the link access was resolved by reservation, rather than contention. The bandwidth reservation in BP is currently being partially implemented by the CGR.

3.4.3 *Recent advances in the DTN architecture for deep-space internetworking*

DTN technology keeps evolving with time. Survey and review papers in different specific areas must be explored by researchers. We list herein some recent important advances in the DTN architectures.

(1) Reliability: end-to-end transport revisited. As we have discussed, the hop-by-hop transport is one of the most important design principles in the DTN architecture. However, in [50], the hop-by-hop principle was questioned for not 100% reliable in data communications. Imagine if any intermediate DTN node took the custody and accidentally failed. The bundles are also gone with the failure and can never be recovered, which will never happen with the TPs designed with the end-to-end principle. In case of this scenario, several new TPs were developed (e.g., Deep-Space Transport (DS-TP) Protocol [51], Delay Tolerant Transport Protocol (DTTP) [52], and Delay Tolerant Payload Conditioning (DTPC) protocol [53]). The DTPC protocol was integrated into the JPL's Interplanetary Overlay Network (ION) DTN reference implementation [54]. The bundle layer end-to-end retransmission mechanism was also proposed in [55].

(2) Efficiency: erasure coding vs Automatic Retransmission reQuest (ARQ) in a higher layer. Erasure coding and ARQ schemes both support reliable transmissions. Although usually implemented in the datalink layer, erasure coding is feasible for implementation in higher layers [6]. Different erasure coding schemes in different layers have been proposed for a higher transmission efficiency. Some examples are given as follows: application layer erasure coding with CFDP [56], erasure coding with an end-to-end ARQ scheme in bundle layer [57], erasure coding with custody transfer in the bundle layer [58], a rateless coding-based LTP variant [59], etc. The research results showed that erasure coding in higher layers benefits the data transfer efficiency in deep-space communications. Seeing further research on the optimum layer to perform coding in different scenarios would be interesting.

(3) More services: bundle streaming service for multimedia communications. The BP was designed for bundle delivery; hence, no function block can support data streaming in the original DTN architecture. With the increasing number of robotic and manned missions in future space exploration, a streaming service was expected for the DTN architecture [60]. In [61], the authors proposed the bundle streaming service (BSS), a communication framework that allows reliable data streaming over delay/disruptive tolerant networks, which has also been implemented in ION DTN.

4 Several open questions in the network protocol architectures for future deep-space internetworking

A grand age of deep-space exploration is taking place. Designing the network to support these great missions is inspiring while challenging. Many open questions for future deep-space internetworking, especially in network protocol architectures, remain.

(1) Is the DTN scalable for future deep-space internetworking?

Network scalability is the capability of a network to expand and work efficiently when the work load is increased. The scalability of the CGR has raised many questions because the original CGR [62] conducted a route computation according to its local contact plan for every forwarding bundle, which seemed computationally intensive, especially when the network topology was big. The results in [63] proved the exponential time complexity of the CGR with growth in the size of the topology. The CGR is certainly evolving with time [64] to solve these problems. Path encoding has been proposed in [65] as a solution. The calculated path is attached to the bundle to avoid recalculation on each DTN node. Further study is still needed to determine the scalability problem in the DTN.

(2) What is missed in the DTN to support future deep-space missions?

Future deep-space missions will involve more robotic and human activities [66]. Some functions missing in the current DTN network architecture can be used to support various new services in future deep-space missions.

- Mutlicast/Broadcast for group exploration activities.
- Dynamic routing in addition to the quasi-static CGR.
- Improved real time streaming based on the BSS.
- More DTN based applications for better user experience.

5 Conclusion

This study investigated the current network protocol architectures for future deep-space internetworking. With research on the current technical advances in the communication technology and the transition in the physical infrastructure, we categorized most of the constraints and challenges in deep-space internetworking into four different levels and listed their effects on the network functions in different layers. The detailed technical characteristics of three current effective architectures in the CCSDS (i.e., TCP/IP architecture, SPP based architecture and DTN architecture) were also analyzed. The DTN currently seems to be the most persuasive architecture for future deep-space internetworking. Finally, two important problems of the DTN (i.e., the scalability problem and the missing function problems) are proposed for a further study to better support future deep-space internetworking.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant No. 61032003), Training Program of the Major Research Plan of the National Natural Science Foundation of China (Grant No. 91338112), and National Science Fund for Distinguished Young Scholars (Grant No. 1525103).

References

1 International Telecommunication Union. ITU-R Radio Regulations, Article 1, Terms and definitions, Section VIII, Technical terms relating to space, paragraph 1.177. http://life.itu.int/radioclub/rr/art1.pdf
2 Ning X, Li Z, Wu W, et al. Recursive adaptive filter using current innovation for celestial navigation during the Mars approach phase. Sci China Inf Sci, 2017, 60: 032205
3 Burleigh S, Cerf V G, Crowcroft J, et al. Space for Internet and Internet for space. Ad Hoc Networks, 2014, 23: 80–86
4 de Cola T, Paolini E, Liva G, et al. Reliability options for data communications in the future deep-space missions. Proc IEEE, 2011, 99: 2056–2074
5 Wang R, Taleb T, Jamalipour A, et al. Protocols for reliable data transport in space Internet. IEEE Commun Surv Tutorials, 2009, 11: 21–32
6 Vasilakos A, Zhang Y, Spyropoulos T V. Delay Tolerant Networks: Protocols and Applications. Boca Raton: CRC Press, 2016. 288–313
7 Zhang W R, Yang G N, Jiang F, et al. Licklider transmission protocol for GEO-relayed space internetworking. Wirel Netw, 2018, 1–11
8 Farrell S, Cahill V, Geraghty D, et al. When TCP breaks: delay- and disruption- tolerant networking. IEEE Internet Comput, 2006, 10: 72–78
9 Hooke A J. The interplanetary Internet. In: Proceedings of the 3rd Annual Symposium on Advanced Radio Technologies, Colorado, 2000. 1
10 Akyildiz I F, Akan B, Chen C, et al. InterPlaNetary Internet: state-of-the-art and research challenges. Comput Netw, 2003, 43: 75–112

11 Akan O B, Fang J, Akyildiz I F. TP-Planet: a reliable transport protocol for interplanetary Internet. IEEE J Sel Areas Commun, 2004, 22: 348–361
12 Burleigh S, Hooke A, Torgerson L, et al. Delay-tolerant networking: an approach to interplanetary internet. IEEE Commun Mag, 2003, 41: 128–136
13 Burleigh S, Cerf V, Durst R, et al. The interplanetary Internet: a communications infrastructure for Mars exploration. Acta Astronaut, 2003, 53: 365–373
14 Bozzi M, Cametti M, Fornaroli M, et al. Future architectures for European space agency deep-space ground stations [antenna applications corner]. IEEE Antennas Propag Mag, 2012, 54: 254–263
15 National Aeronautics and Space Administration. SCaN 101. https://www.nasa.gov/sites/default/files/files/SCaN_101_2013.pdf
16 Bagri D S, Statman J I, Gatti M S. Proposed array-based deep space network for NASA. Proc IEEE, 2007, 95: 1916–1922
17 Hemmati H. Deep Space Optical Communications. Hoboken: John Wiley & Sons, Inc., 2006. 1–82
18 Wu W, Chen M, Zhang Z, et al. Overview of deep space laser communication. Sci China Inf Sci, 2018, 61: 040301
19 Consultative Committee for Space Data Systems. Real-time weather and atmospheric characterization data. CCSDS 140.1-G-1. 2017. https://public.ccsds.org/Pubs/140x1g1.pdf
20 NASA space communication and navigation program next generation space relay architecture concept study. Space relay architecture background information. NNC16ZLC002L. 2016. https://elibrary.gsfc.nasa.gov/_assets/doclibBidder/tech_docs/NASA
21 Munger J, Ladrach W, Hetrick J. The next generation space relay architectures. In: Proceedings of the 34th AIAA International Communications Satellite Systems Conference, Cleveland, 2016. 1–18
22 Reinhart R C, Schier J S, Israel D J, et al. Enabling future science and human exploration with NASA's next generation near earth and deep space communications and navigation architecture. In: Proceedings of International Astronautical Congress, Adelaid, 2017. 1–10
23 Rosborough V, Gambini F, Snyder J, et al. Integrated transmitter for deep space optical communications. In: Proceedings of 2016 IEEE Avionics and Vehicle Fiber-Optics and Photonics Conference, Long Beach, 2016. 207–208
24 Brooks D E, Eddy W, Johnson S K, et al. In-space networking on NASA's SCaN testbed. In: Proceedings of the 34th AIAA International Communications Satellite Systems Conference, Cleveland, 2016. 1–9
25 Brandon C, Chapin P. The use of SPARK in a complex spacecraft. ACM SIGAda Ada Lett, 2016, 36: 18–21
26 Sabbagh A, Wang R, Zhao K, et al. Bundle protocol over highly asymmetric deep-space channels. IEEE Trans Wirel Commun, 2016, 16: 2478–2489
27 Farrell S, Cahill V. Security considerations in space and delay tolerant networks. In: Proceedings of the 2nd IEEE International Conference on Space Mission Challenges for Information Technology, Pasadena, 2016. 1–8
28 Cath C, Floridi L. The design of the Internet's architecture by the Internet engineering task force (IETF) and human rights. Sci Eng Ethics, 2017, 23: 449–468
29 Mukherjee J, Ramamurthy B. Communication technologies and architectures for space network and interplanetary Internet. IEEE Commun Surv Tutorials, 2013, 15: 881–897
30 Akyildiz I F, Akan O B, Chen C, et al. The state of the art in interplanetary Internet. IEEE Commun Mag, 2004, 42: 108–118
31 Consultative Committee for Space Data Systems. Space packet protocol. CCSDS 133.0-B-1. 2003. https://public.ccsds.org/Pubs/133x0b1c2.pdf
32 Consultative Committee for Space Data Systems. Space Communications Protocol Specification (SCPS)-Transport Protocol. CCSDS 714.0-B-2. 2006. https://public.ccsds.org/pubs/714x0b2.pdf
33 Consultative Committee for Space Data Systems. IP over CCSDS Space Links. CCSDS 702.1-B-1. 2012. https://public.ccsds.org/Pubs/702x1b1c1_tc1413.pdf
34 Consultative Committee for Space Data Systems. Encapsulation Service. CCSDS 133.1-B-2. 2009. https://public.ccsds.org/Pubs/133x1b2c2.pdf
35 Consultative Committee for Space Data Systems. TC Space Data Link Protocol. CCSDS 232.0-B-2. 2010. https://public.ccsds.org/Pubs/232x0b3.pdf
36 Consultative Committee for Space Data Systems. TM Space Data Link Protocol. CCSDS 132.0-B-1. 2003. https://public.ccsds.org/Pubs/132x0b2.pdf
37 Consultative Committee for Space Data Systems. AOS Space Data Link Protocol. CCSDS 732.0-B-2. 2006. https://public.ccsds.org/Pubs/732x0b2c1s_tc1230.pdf
38 Consultative Committee for Space Data Systems. Proximity-1 Space Link Protocol-Data Link Layer. CCSDS 211.0-B-5. 2013. https://public.ccsds.org/Pubs/211x0b5.pdf
39 Cheng M P, Clare L P. Prototyping IP over CCSDS for manned space applications. In: Proceedings of 2010 SpaceOps Conference, Huntsville, 2010. 1–14
40 Consultative Committee for Space Data Systems. CCSDS File Delivery Protocol (CFDP). CCSDS 727.0-B-4. 2007. https://public.ccsds.org/pubs/727x0b4.pdf
41 Burleigh S. Operating CFDP in the interplanetary Internet. In: Proceedings of SpaceOps 2002, Houston, 2002. 1–6
42 Wang R, Shrestha B L, Wu X, et al. Unreliable CCSDS file delivery protocol (CFDP) over cislunar communication links. IEEE Trans Aerosp Electron Syst, 2010, 46: 147–169
43 Yang Z, Li H, Jiao J, et al. CFDP-based two-hop relaying protocol over weather-dependent Ka-band space channel. IEEE Trans Aerosp Electron Syst, 2015, 51: 1357–1374

44 Sanders F A, Jones G, Levesque M. Transfer of files between the deep impact spacecrafts and the ground data system using CFDP: a case study. In: Proceedings of 2007 IEEE Aerospace Conference, Big Sky, 2007. 1–5

45 de Cola T, Ernst H, Marchese M. Performance analysis of CCSDS file delivery protocol and erasure coding techniques in deep space environments. Comput Netw, 2007, 51: 4032–4049

46 Jiao J, Guol Q, Zhang Q Y. Packets interleaving CCSDS file delivery protocol in deep space communication. IEEE Aerosp Electron Syst Mag, 2011, 26: 5–11

47 Fall K, Farrell S. DTN: an architectural retrospective. IEEE J Sel Areas Commun, 2008, 26: 828–836

48 Consultative Committee for Space Data Systems. CCSDS Bundle Protocol Specification. CCSDS 734.2-B-1. 2015. https://public.ccsds.org/Pubs/734x2b1.pdf

49 Heimlicher S, Baumann R, May M, et al. The transport layer revisited. In: Proceedings of the 2nd International Conference on Communications Systems Software and Middleware, Bangalore, 2007. 1–8

50 Papastergiou G, Samaras C V, Tsaoussidis V. Where does transport layer fit into space dtn architecture? In: Proceedings of the 5th Advanced Satellite Multimedia Systems Conference and the 11th Signal Processing for Space Communications Workshop, Cagliari, 2010. 81–88

51 Papastergiou G, Psaras I, Tsaoussidis V. Deep-space transport protocol: a novel transport scheme for space DTNs. Comput Commun, 2009, 32: 1757–1767

52 Samaras C V, Tsaoussidis V. Design of delay-tolerant transport protocol (DTTP) and its evaluation for Mars. Acta Astronaut, 2010, 67: 863–880

53 Papastergiou G, Alexiadis I, Burleigh S, et al. Delay tolerant payload conditioning protocol. Comput Netw, 2014, 59: 244–263

54 ION-DTN: Delay-Tolerant Networking suitable for use in spacecraft. Version 3.6.1. NASA JPL. 2018

55 Koutsogiannis E, Tsapeli F, Tsaoussidis V. Bundle layer end-to-end retransmission mechanism. In: Proceedings of 2011 Baltic Congress on Future Internet Communications, Riga, 2011. 109–115

56 de Cola T, Ernst H, Marchese M. Performance analysis of CCSDS file delivery protocol and erasure coding techniques in deep space environments. Comput Netw, 2007, 51: 4032–4049

57 Hou D, Zhao K. Application layer channel coding for space DTN. In: Proceedings of 2017 International Conference on Machine Learning and Intelligent Communications, Weihai, 2017. 347–354

58 de Cola T, Marchese M. Joint use of custody transfer and erasure codes in DTN space networks: benefits and shortcomings. In: Proceedings of 2010 IEEE Global Telecommunications Conference, Miami, 2010. 1–5

59 Gu S, Jiao J, Yang Z, et al. RCLTP: a rateless coding-based Licklider transmission protocol in space delay/disrupt tolerant network. In: Proceedings of 2013 International Conference on Wireless Communications & Signal Processing (WCSP), Hangzhou, 2013. 1–6

60 Lenas S A, Burleigh S C, Tsaoussidis V. Reliable data streaming over delay tolerant networks. In: Proceedings of 2012 International Conference on Wired/Wireless Internet Communications, Santorini, 2012. 358–365

61 Lenas S A, Burleigh S C, Tsaoussidis V. Bundle streaming service: design, implementation and performance evaluation. Trans Emerging Tel Tech, 2015, 26: 905–917

62 Burleigh S. Contact graph routing IETF Internet draft. Internet-Draft. 2009. https://tools.ietf.org/html/draft-burleigh-dtnrg-cgr-00

63 Wang G, Burleigh S, Wang R, et al. Scoping contact graph-routing scalability: investigating the system's usability in space-vehicle communication networks. IEEE Veh Technol Mag, 2016, 11: 1556–6072

64 Araniti G, Bezirgiannidis N, Birrane E, et al. Contact graph routing in DTN space networks: overview, enhancements and performance. IEEE Commun Mag, 2015, 53: 38–46

65 Birrane E, Burleigh S, Kasch N. Analysis of the contact graph routing algorithm: bounding interplanetary paths. Acta Astronaut, 2012, 75: 108–119

66 Kazz G J, Burleigh S C, Cheung K M, et al. Evolution of the Mars relay network end-to-end information system in the Mars human era (2030-2040). In: Proceedings of 2016 SpaceOps Conference, Daejeon, 2016. 1–10

67 Jiao J, Hu Y, Zhang Q, et al. Performance modeling of LTP-HARQ schemes over OSTBC-MIMO channels for hybrid satellite terrestrial networks. IEEE Access, 2018, 6: 5256–5268