Achieving High Goodput Performance in Mars Missions through Application Layer Coding and Transmission Power Trading

Tomaso de Cola
CNIT - Italian National Consortium
for Telecommunications
University of Genoa, Italy
Email: tomaso.decola@cnit.it

Harald Ernst
DLR- Deutsches Zentrum für Luft- und Raumfahrt
German Aerospace Center
Email: harald.ernst@dlr.de

Mario Marchese
DIST
University of Genoa, Italy
Email: mario.marchese@unige.it

Abstract — Transferring data reliably from Mars to Earth stations is becoming an appealing challenge in the design of interplanetary networks. In this view, the use of CCSDS-based protocols along with erasure codes is a promising candidate for mitigating the performance degradation due to large propagation delays and high packet loss rates. Besides, the constraints imposed on the transmission power and on the available link bandwidth introduce further challenges to the design of the telecommunication system. In this paper, the benefits brought are investigated by tuning properly erasure code rate at the upper layers, pointing out the performance gain in terms of goodput performance and analysing power consumption and bandwidth usage issues, as well.

Index Terms – CFDP, Erasure Coding, Transport layer, Satellite Communications*

I. INTRODUCTION

Over last years, National Aeronautics and Space Administration (NASA) agency together with Consultative Committee for Space Data Systems (CCSDS) [1] have made strong efforts to design and deploy a telecommunication network, enabling data communication in Mars missions. The wide gamma of applications supposed to be supported by such a network infrastructure ranges from image retrieval and the download of telemetry data to audio and video streaming. In this perspective, it is of paramount importance to design effective transmission schemes to guarantee, whenever necessary, a reliable transfer of data. Moreover, the large latencies and high information loss ratios, exhibited by interplanetary environments, affect severely the effectiveness of TCP/IP protocols [2], making necessary the design of more appropriate transmission paradigms. In fact, the large free-space loss along with unpredictable shadowing and fading events cause strong signal degradation, that can be properly mitigated by means of efficient satellite link budgets and by tuning Forward Error Correction codes at the physical and datalink layer. These countermeasures, however, are not sufficient since the large costs of deep space satellite and spacecraft equipments imply constraints on the design of the telecommunication system, in terms of the antenna size the and transmission power, which consequently impact on the link budget. Thus, novel techniques able to cope with space impairments and to adapt to the highly dynamic nature of this environments are required, taking into account the transmission power constraints.

It is immediate to see that, in presence of frame loss, guaranteeing reliable data transfer can be achieved only through specific solutions applied at higher layers. From this point of view, the scientific literature had came up in the years with several proposals aimed at guaranteeing high performance data communications in space networks. Ref. [3] proposes a transport protocol suitable to transferring data in lossy space networks, able to recognise outage periods and hence tune transmission algorithms accordingly. Ref. [4] shows as multimedia flows can be handled in deep space networks by means of multi-rate transmission schemes. Protocol architectures able to deal with long outage periods and large latencies are proposed by the Delay Tolerant Network (DTN) working group [5] within IRTF (Internet Research Task Force) and by CCSDS. In particular, CCSDS standardised a protocol stack architecture specifically designed for space missions. Particular attention has to be reserved to the CCSDS File Delivery Protocol (CFDP) spanning over transport and application layer and responsible for guaranteeing reliable data transfer, if necessary, through ARQ mechanisms. Implementing erasure coding techniques may be a viable solution too: [6] shows the advantages of such techniques for computer networks, while [7] presents results of LDPC-based codes applied at the application layer to guarantee efficient multicast data communication over wireless environments. The idea of taking advantages from erasure codes in deep space communications is also addressed in the Long Erasure Codes working group within CCSDS. Finally, [8] contains a proposal for combining advantages of erasure codes and CCSDS protocol architectures, showing the benefits arising in cislunar missions, in terms of effective throughput. Finally, [9] explores the possibility of integrating the ARQ capabilities of CFDP protocol with the powerful LDGM codes [7], showing the performance gain with respect to the standard CFDP implementation.

Despite the large number of protocol proposals, most of them are not addressing the transmission power constraints characterizing the design of telecommunication systems for interplanetary networks. Starting from this observation, this work aims at investigating performance provided by improved CFDP-based solutions in terms of goodput, in dependence of the transmission power and the link budget arranged on the
interplanetary transmission channels. Optimal solutions are identified, by tuning the protocol configuration appropriately, in terms of frame size and code rate of the erasure codes, implemented at the higher layers.

The remainder of this paper is structured as follows. Section II surveys protocol specifications of the CCSDS protocol architecture, while Section III introduces the reference scenario, showing the specifics of the telecommunication system and analysing the peculiarities of deep space environments in terms of a Discrete Time Markov Chain. Section IV presents the protocol solutions investigated in this work and Section V shows the performance analysis. Finally, Section VI draws the conclusions.

II. THE CCSDS PROTOCOL ARCHITECTURE

Since the early eighties, the Consultative Committee for Space Data Systems (CCSDS) has been strongly involved in providing guidelines for the design of space telecommunication systems. Basically, a set of protocol specifications spanning from the physical up to the application layer has been devised. For the sake of the simplicity, only protocols directly connected to this work are presented in the following.

Looking at the whole protocol stack, layer by layer [10]:

- **Application/Transport.** It implements the CCSDS File Delivery Protocol (CFDP), actually spanning application and transport layer, that is responsible for guaranteeing reliable data communication thanks to ARQ schemes, based on NAK issuance.

- **Network.** It implements the CCSDS Space Packet Protocol, performing typical network layer operations, such as addressing and routing.

- **Datalink and Physical.** They implement different protocol solutions in dependence of the data service handled and peculiarities of the transmission link. In more detail, it is possible to classify deep space transmission channels into proximity and long-haul links. First are established between stations and spacecrafts moving in proximity of each other (e.g. the related time distance is in the order of hundreds of milliseconds). They transport data through the Proximity-1 Datalink and Physical layer protocols. On the other hand, long-haul link connect spacecrafts and stations very far from each other (e.g. time distance is in the order of tens of seconds). In this case the reference protocol is represented by the Advanced Orbiting System recommendation, which works together with the TM Channel Coding and Synchronization protocols.

III. THE REFERENCE SCENARIO

A. The Telecommunication System

The investigated scenario is derived from the ongoing space missions aimed at exploring the Mars surface, collecting images and measurements and transmitting them to the Earth gathering centres via a suitable communication infrastructure. In more detail, the reference environment considered in this paper consists of sensors, rovers and landers placed on the Mars’ surface, responsible for taking measures and pictures, which will be sent towards a relay node on the Earth, by means of a two-hop satellite link. As far as the relay nodes are concerned, a spacecraft located nearby Mars surface receives data arriving from the sensors, landers and rovers, while a satellite orbiting around the Earth, collects data arriving from the Mars’ subnet and forward them to the Earth gathering centre.

The photos in the lower part of Fig. 1 show the data path established between the different nodes introduced before.

As far as the protocol architecture is concerned, a full CCSDS-based protocol stack is implemented in each node, as shown in Fig. 1. In particular, the application layer implements the CFDP protocol, while the Space Packet protocol is assumed at the network layer. The protocol implementation at lower layers strictly depends on the characteristics of the medium through which data are transmitted. In practice, relay nodes implement a dual stack that present, on the one side, the Proximity-1 Datalink and Physical protocols and, on the other side, the Advanced Orbiting System protocol. It is straightforward to see that the end nodes, directly connected to the relays, implement Proximity-1 protocols.

![Fig. 1. Reference Scenario and CCSDS Protocol Stack](image)

The focus of this paper is on the deep space link, connecting the relay nodes.

B. The Deep Space Environment: Modelling Issues

Deep space links introduce a number of hazards that affect the transmission channel reliability in terms of shadowing and fading attenuation. In more detail, events such as solar wind and flares, partial unavailability of line of sight between antennas because of asteroids, sun conjunction as well as relative motion between satellite and spacecrafts can be experienced during data communication. As a major result, low signal to noise ratios are registered and, consequently, raw bit error rates are about $10^{-3}$. Since transmission channel exhibits slow fading and shadowing, model commonly used to reproduce the behaviour of wireless links are suited even in this case. For the sake of simplicity this work assumes that fading and shadowing attenuation can be represented as stationary stochastic processes, according to a Rayleigh distribution function. Besides, a four state Discrete Time Markov Chain...
The implementation of erasure coding schemes at the application within the CFDP protocol itself and benefits arising from the protocol, taking advantage of specific features implemented CCSDS File Delivery Protocol (CFDP) is the reference model has been derived. In practice, the bursty nature of bit losses, as reproduced by the 4-state markov chain, gives rise to model has been derived. In practice, the bursty nature of bit losses, as reproduced by the 4-state markov chain, gives rise to the following as error and error-free gaps, indicated in the following as error and error-free gaps, respectively.

The other variables characterising the Markov Chain can be derived by the transmission parameters, such as modulation scheme, channel encoding algorithm and code rate.

Finally, in order to evaluate the impact of bit errors on the datalink frame and hence on the PDUs of upper layers, a GAP model has been derived. In practice, the bursty nature of bit losses, as reproduced by the 4-state markov chain, gives rise to consecutive occurrences of corrupted and correct frames, indicated in the following as error and error-free gaps, respectively.

IV. THE PROTOCOL SOLUTIONS
Proposed solutions are derived from [8-9], where the CCSDS File Delivery Protocol (CFDP) is the reference protocol, taking advantage of specific features implemented within the CFDP protocol itself and benefits arising from the implementation of erasure coding schemes at the application layer. Three protocol solutions have been considered, as follows:

CFDP-deferred. CFDP operates in acknowledged mode, performing the recovery procedures, upon missing CFDP blocks are detected through deferred NAK.

CLDGM. CFDP operates in unacknowledged mode, meaning that recovery mechanism are not available. Nevertheless, in order to make the data communication more robust against link errors, erasure codes based on LDGM (a variant of LDPC) are implemented. Actually, the encoding procedure takes a number of CFDP blocks as input, split them into k packets, whose dimension is ruled by the frame size at the data link layer, and then encodes into n packets. Since LDGM codes perform efficiently in presence of a large number of k, several CFDP blocks are merged together such to fill one only bit vector carrying 1 Mbytes.

CLDGM-deferred. It implements both deferred issuance of NAK, as available from CFDP-deferred, and CLDGM encoding techniques as specified in CLDGM. The idea is to reduce the number of retransmission loops.

V. PERFORMANCE ANALYSIS
A. The Testbed Configuration
In order to evaluate the effectiveness of the presented protocol solutions, a typical case of space mission has been considered: sensor nodes and rovers moving on Mars’ surface have to upload measures and images to Earth gathering centres, through relay nodes and a deep space link. In practice, a big quantity of data, equal to 100 Mbytes, is transferred through the network infrastructure. The goal of this study was to identify the solutions offering the best performance results in terms of the observed goodput. Besides, considerations on the deep space link budget have been provided on the basis of the C/N₀ (Carrier power-to-Noise power spectral density ratio) figures, required to attain the desired optimal performance. The tests have been performed by means of simulation tools, properly adapted to the environment investigated and to the aim of this study. A number of runs sufficient to obtain a width of the confidence interval less than 1% of the measured values for 95% of the cases was imposed.

For the sake of the simplicity, but without loss of generality, the analysis has been mainly focused on the deep space link performance, actually introducing the most relevant challenges from the networking viewpoint. The test campaign has been conducted considering the transfer of 100 Mbytes over the deep space link, exhibiting a forward and reverse bandwidth of 1Mbit/s and 1 Kbit/s, respectively. As far as the propagation delay is concerned, it has been set to 200s, corresponding to the minimum Earth-Mars distance (about 60 millions of Km).

Different configurations of the deep space link have been analysed, by varying parameters of the 4-state markov chain. Actually, configurations in terms of the couple (τ₀, τ₃) have been tested, while π₁ and π₂ were set to constant values. In practice permanence times in states 0 and 3 have been varied between 5s and 60s. Afterwards, these configurations have been revisited in terms of C/N₀ values, taking into account the specifics of modulation and channel coding schemes, implemented at the physical layer.

As far as the protocol solutions are concerned, performance offered by CFDP-deferred, CLDGM and CLDGM-deferred has been thoroughly evaluated, by varying, when applicable, CFDP block size, code rate and frame size, as follows:

- CFDP_deferred: CFDP block ranging from 1024 bytes up to 65536 bytes.
- CLDGM and CLDGM_deferred: CFDP blocks are merged together to fill an unique bit vector carrying 1 Mbytes of data; code rate of erasure codes applied at the application layer has varied from 0.125 to 0.875.
- In all the protocol configurations tested, frame size ranged from 128 bytes to 1024 bytes.

The effectiveness of the proposed solutions has been analysed in terms of the goodput [bit/s], that corresponds to the number of information bits transferred successfully per time unit. Basically, investigated protocol solutions were assessed in terms of goodput, since it conveys information on how reliable the communication was (it accounts only for the received information bits) and numerically shows the transfer rate actually perceived by users.

B. The Results
CLDGM, CLDGM performance is essentially ruled by the channel condition (in terms of τ₀ and τ₃ values), the frame size as well as the applied code rate. In practice, the role played by frame size is really important since link errors cause error gaps,
whole length is strictly dependent on the frame size. Besides, the impact of gap duration on the whole performance is immediate to see: larger the gaps are, more reduced is the number of bytes reliably transferred. In fact, this characteristic can be accounted for in terms of Ploss, evaluated as the ratio between received and transmitted information blocks.

In the case of $\tau_0=5s$ and $\tau_3=60s$, long error gap periods are experienced and, as a result, registered Ploss gap values are moderately high, ranging from 0.1 to 1, as frame size configurations change. In particular, the adoption of code rate 0.875 does not offer satisfactory performance because the reduced number of redundancy packets is not able to counteract erasures caused by link errors. Moreover, from Fig. 2, it is possible to see, that in the case of code rate equal to 0.750, registered performance starts with Ploss equal to 1 when the sent frames carry 128 bytes, then reducing below 0.5 as frame size increases from 256 bytes up to 1024 bytes. As far as code rate settings are considered, all the Ploss registered values just overlap for any frame size. In particular, it is possible to observe that when the frame size is set to 512 bytes, the most satisfactory result is achieved, Ploss of 0.07. This behaviour is due to the fact that, in this case, an intermediate value of frame size (512 bytes) gives rise to error gaps moderately long, whose impact on the correct information delivery is effectively contrasted by the erasure codes implemented, regardless of the code rate.

As far as the goodput analysis is concerned, code rate plays a fundamental role because on the one hand a high number of redundancy packets (low code rate) allows recovering efficiently from losses, while on the other hand higher code rates permit filling almost completely the bandwidth pipe. As a result, an intermediate code rate, equal to 0.625, gives the most satisfactory results in terms of goodput, equal to 600 kbit/s, when a frame size of 512 bytes is considered. When code rates other than 0.625 are analysed, goodput values range from a minimum of 30 kbit/s (for code rate of 0.875) up to a maximum of 580 kbit/s (for code rate 0.750). In this view, a particular note has to be reserved to the case of code rate 0.750. In this configuration the maximum is observed for a frame size of 256 bytes; on the contrary when frame size increases from 384 bytes up to 1024 bytes, performance reduces down to 420 kbit/s because of longer error gap periods.

The case of $\tau_0=5s$ and $\tau_3=20s$ gives place to interesting considerations mainly about Ploss behaviour. The trend shown in Fig. 3 highlights how Ploss increases from 0.1 up to 0.2 as frame size ranges from 128 bytes to 1024 bytes, respectively, when code rate is lower than 0.750. On the contrary, when a code rate of 0.875 is set, Ploss is very high (0.9 – 1) when frame cannot carry more than 512 bytes. Afterwards, as frame size increases up to 1024 bytes, better performance is achieved. As for code rate 0.750, a similar behaviour is observed: Ploss starts from 1, then drops to approximately 0.1 (frame size of 256 bytes) and eventually gets 0.2 as the frame size increases up to 1024 bytes.

The cases of $\tau_0=20s$ and $\tau_3=5s$, and $\tau_0=60s$, $\tau_3=5s$ produce similar results because of long error-free gap periods, that help improve the overall performance. From this viewpoint, it is more interesting to analyse the second case, which gives rise to the best goodput values. In this channel configuration, the impact of link errors is reduced if compared to the previous cases; consequently a lower number of redundancy packets is sufficient to achieve high performance. Actually, goodput values ranging from 110 kbit/s up to 880 kbit/s are registered for code rate varying from 0.125 up to 0.875 respectively. In this case, the role played by the frame size is crucial to the performance: it starts from 110 kbit/s (frame size of 128 bytes), then increases to 880 kbit/s (frame size of 512 bytes). As frame size further increases (up to 1024 bytes) no significant variations are observed.

As for code rate 0.750, a similar behaviour is observed: Ploss starts from 1, then drops to approximately 0.1 (frame size of 256 bytes) and eventually gets 0.2 as the frame size increases up to 1024 bytes.

Fig. 2. Ploss of CLDGM when $\tau_0=5s$ and $\tau_3=60s$

CFDP-deferred

CFDP-deferred performance is strictly dependent on the frame and the block size. In particular, in this study CFDP blocks can carry data from 1024 up to 65536 bytes, differently from CLDGM case where fixed blocks of 65536 bytes were set.

In general, the conducted tests have highlighted goodput values ranging from a minimum of 110 kbit/s (case $\tau_0=5s$ and $\tau_3=60s$) to a maximum of 430 kbit/s (case $\tau_0=60s$ and $\tau_3=5s$). Actually, the double dependence on frame and block size does not show a precise performance trend, because differently from CLDGM, CFDP-deferred does not implement any erasure code mechanisms and communication reliability is ensured by means of specific recovery procedure based on ARQ schemes.
As a result, the large propagation delay (200s) causes long recovery periods when erasures are exhibited by the channel, thus impairing the global performance. The dependence of performance from block and frame size is not immediate to see, whereas observed goodput values are more sensitive to the channel configuration.

**CLDGM-deferred**

CLDGM-deferred takes advantage of both erasure codes (as in CLDGM) and ARQ mechanisms (as in CFDP-deferred). As a consequence, goodput values are strictly dependent on the code rate and frame size settings. In particular, using erasure codes should help reduce the number of lost blocks, hence preventing from long retransmission phases as in the case of CFDP-deferred. Finally, residual information loss will be recovered by the implemented ARQ scheme. From this point of view, it is expected to achieve better performance than in CFDP-deferred.

In the case of $\tau_0=5s$ and $\tau_3=60s$, the maximum goodput is achieved for a code rate of 0.750 and a frame size of 512 bytes, as evident from Fig. 4. Actually, the value of 512 bytes corresponds to a maximum for the goodput independently of the code rates. In fact, in this configuration, performance varies from 100 kbit/s (code rate of 0.875) up to 450 kbit/s (0.750 kbit/s). For lower and bigger values of frame sizes, it is possible to observe smaller values for the goodput, that drops to 20 kbit/s for frame size of 1024 bytes and code rate of 0.875.

As error gap periods get shorter (case $\tau_0=5s$ and $\tau_3=60s$), the observed performance is slightly better than in the previous case. In fact, goodput ranges from a minimum of 30 kbit/s (frame size of 640 bytes and code rate of 0.875) to a maximum of 450 kbit/s (frame size of 768 bytes and code rate of 0.875). It is worth remarkable that in this case both minimum and maximum values for goodput are achieved by the same code rate configuration (i.e. 0.875), for different frame sizes. This behaviour is due to the impact of frame size and link errors on the whole performance.

As far as cases $\tau_0=20s$ and $\tau_3=5s$, and $\tau_0=60s$ and $\tau_3=5s$ are concerned, similar behaviours are experienced by CLDGM-deferred. A particular attention can be reserved to the second case ($\tau_0=60s$ and $\tau_3=5s$), where longer error-free gaps assure the most satisfactory performance results, depicted in Fig. 5. Goodput varies from a minimum of 110 kbit/s (for code rate of 0.125 in the best frame size configuration) up to a maximum of 650 kbit/s (for code rate of 0.875 for the best frame size configuration). In this case, it is immediate to see that for code rates ranging from 0.125 to 0.5, achieved performance slightly varies with the frame size. On the other hand, as code rate gets bigger than 0.625, sketched curves show a “knee”. The reason for this behaviour is due to the trade-off between frame size and code rate. In this specific case, it is immediate to recognize that small code rates are able to recover all the missed information blocks without invoking retransmission loops, independently of the frame size. On the contrary, high code rates are not able to allow the full delivery data, which can be ensured by means of ARQ mechanisms, whose efficiency is thus ruled by the frame size, as highlighted by Fig. 5.

**C. Comparison of Results**

In order to complete the performance analysis and hence identify the most promising solution to be applied for data communication in Mars-Earth operations, comparisons between sketched solutions have been considered. First of all, an other solution, based on the CFDP paradigm, has been considered: CFDP-extended. Basically it implements the “extended operation” feature, consisting in suspending and resuming operations. In this investigation, CFDP-extended implementation is supposed to exploit an “a priori” knowledge of the channel state, forwarded by lower layer protocols and notifying when error gap periods occur. Finally, also power consumption and bandwidth usage issues are addressed.

Goodput values achieved in the most efficient configurations (i.e. code rate, block and frame size, when applicable) for CLDGM, CFDP-deferred and CLDGM-deferred have been compared with CFDP-extended, in dependence of the Carrier power-to-Noise power spectral density ratio, $C/N_0$, exhibited by the deep space link.

From Fig. 6, it is possible to note that CLDGM is able to guarantee the most satisfactory results ranging from 410 kbit/s to 860 kbit/s, regardless of the $C/N_0$ values. CFDP-deferred gives poor performance results, varying from 180 kbit/s to 420 kbit/s, as $C/N_0$ varies from 63 dBHz up to 67.5 dBHz. CLDGM-deferred achieves satisfactory performance results as $C/N_0$ gets bigger than 64 dBHz. Finally, CFDP-extended
presents satisfactory goodput results for low C/N0 values. As C/N0 increases, and accordingly link errors reduce, observed performance is close to CLDGM-deferred. This behaviour is due to the fact that in presence of long error gap periods (low C/N0), suspend and resume option help improve performance.

Finally, in order to complete the evaluation it is necessary to analyse how much bandwidth (at the physical layer) is used by the proposed protocols to meet the desired performance. In this perspective, the bandwidth efficiency factor, defined as ratio between goodput and real bandwidth used has been introduced and taken as reference parameter. Real bandwidth has been computed by taking into account the code rate, applied at both CFDP layer (if any) and datalink layer (turbo-codes), as well as the overhead introduced by each protocol layer.

As sketched in Fig. 7, CFDP-deferred presents poor performance results, with bandwidth efficiency ranging from 0.025 up to 0.18, as C/N0 range from 63 dBHz to 67.5 dBHz. The best results (0.18 – 0.28) are offered by CFDP-extended for C/N0 lower than 66 dBHz. As C/N0 increases, CLDGM is more effective, achieving a performance of 0.33. As far as CLDGM-deferred behaviour is concerned, it performs fairly for C/N0 greater than 64 dBHz; in particular as C/N0 increases, CLDGM-deferred achieves a maximum bandwidth efficiency of 0.25 with respect to 0.29 and 0.33 offered by CFDP-extended and CLDGM, respectively.

It is worth remarking that in this case, differently from the previous comparison, CFDP-extended is able to offer very satisfactory results, thus allowing for a more efficient usage of the channel bandwidth and hence of the network resources.

VI. CONCLUSIONS AND FUTURE WORK

This work studied protocol proposals derived from CCSDS File Delivery Protocol and able to achieve high performance in terms of goodput. As first step, CLDGM, CLDGM-deferred and CFDP-deferred were analysed by identifying the protocol configuration (code rate, block and frame size) more suitable to different channel conditions, expressed in terms of $\tau_0$ and $\tau_1$ values. Finally, the assessment phase was completed by taking into account power consumption and bandwidth usage in terms of C/N0 and bandwidth efficiency values.

From this point of view, tests highlighted that in presence of very strict link budget constraints (dependence of goodput from C/N0) CLDGM offers the most satisfactory results, thanks to the powerful application layer coding. On the other hand, when bandwidth usage was critical for the system design, CFDP-extended proves to be more effective.

Extensions of this work may address the possibility of integrating the “extended operations” option (as in CFDP-extended proposal) into the CLDGM solution, in order to provide a more efficient protocol implementation. Besides, in this case, tuning properly the size of buffers, where data of suspended transmissions have to be stored, is of primary interest. This leads to the design of congestion control policies and hence to the optimisation of transmission mechanisms implemented within the CFDP protocol entity.

REFERENCES