SATELLITES AS A SERVICE: LEVERAGING RELIABLE ON-BOARD HIGH-PERFORMANCE COMPUTING TO DEMOCRATIZE ACCESS TO SPACE

David Selčan [1], Iztok Kramberger [2]

[1] SkyLabs d.o,o., Zagrebška cesta 104, 2000 Maribor, Slovenia, david.selcan@skylabs.si
[2] University of Maribor, Faculty of Electrical Engineering and Computer Science, Koroška cesta 46, 2000 Maribor, Slovenia

ABSTRACT

This paper presents a baseline for the creation of a Space Cloud – a constellation of satellites designed in a way to give external users direct control of the high-performance computing resources, thus enabling them to run custom applications and store data on-board a satellite, while also scheduling communication operations with external Ground Stations. To achieve sufficient reliability and availability of the platform, a three-level FDIR policy is presented and further supported with a method of utilizing a "building blocks" approach to satellite electronic system design to improve reliability of COTS components-based systems. Further, a model of a modular Space Cloud satellite is presented, where the modularity is achieved through the use of scalable satellite subsystems. Finally, the user's perspective of using such a constellation to solve real satellite-supported problems in the form of applications is presented.

1 INTRODUCTION

The last couple of decades have resulted in a tremendous increase in the capabilities of small satellites. By leveraging the tremendous performance and miniaturization gains of the semiconductor and related industries, it has become possible to fit the functions and performance of a large traditional satellite into the form factor of a much smaller satellite. However, as the drivers of this "new space" approach were in majority Universities and SMEs, much of the improvements were realized by directly taking the COTS equipment and using it in a satellite, without taking into account the specifics of operating that equipment in a space environment or trying to fully leverage the new functions now available. As this approach finally gained acceptance in the space community and is beginning to be used to provide serious added value to more challenging missions, it is time to evaluate how to fully and completely utilize the benefits offered by the "real space" approach it has transformed into.

Most small satellites following this "new space" approach have taken advantage of COTS-based high-performance processors and components to provide processing power and sensor performance that can rival or even surpass traditional larger satellites. Yet the way that the mission data is processed has not changed much, data is still gathered by the payload subsystem, transferred and potentially stored by the OBDH system and finally transmitted in a mostly raw format to ground for further processing. However, with the trend of increasing satellite computing performance not showing any signs of slowing down, the discrepancy between the amount of data generated on-board and the amount of data that can be successfully transferred to ground will keep growing. This is currently being addressed by trying to use AI algorithms on board satellites (AI@Edge), nevertheless, due to the necessity of using AI-capable hardware that is not suited for use in a radiation environment and the rigidity of reducing all satellite-based computing problems to a single approach, we propose an alternative.

The proposal presented in this paper is based on the creation of a Space Cloud, allowing its users access to the high-performance computing resources of satellites which are part of the Cloud. This enables users to optimize and run their application on-board a satellite, storing the processed information on-board the satellite and finally transferring the data to the user on demand, without the previously mandatory step of investing in the production and operation of space infrastructure.

The key enabler of this approach is the adaptation of the use of COTS-based components in systems, in a way that the platform is sufficiently reliable both from the point of view of assuring the user sufficient computing power without interruption and from the point of view of preventing the user from affecting the safe operation of the satellite. To achieve this, a hierarchical FDIR policy based on the use of LCLs, supervisors and watchdogs is proposed. By basing the design of the satellites of the Space Cloud on these aspects, high reliability and availability to all users can be assured.

2 RELIABILITY AND AVAILABILITY OF COTS COMPONENT-BASED SATELLITE SYSTEMS

Two critical aspects of satellite reliability and availability must be addressed. The first is the aspect of reliability – as each satellite that is part of the proposed Space Cloud must still be assembled, tested, launched and commissioned, it represents a significant financial investment. As such, it is critical that it remains operational for as much time as possible. The other aspect is that of availability – as each satellite is effectively made available for use to external users, it is imperative that it is fully operational for as much of the time the user requires from it as possible.

Both of the aspects address the need for the platform to be reliable against the harsh environment that is space. The design of the satellite must take into account ionizing radiation aspects, the limitations of remote operation of electronic systems and thermal constraints. An additional important focus of the Space Cloud is the capability of users to upload their own code and execute it on the satellite itself – safeguards must be in place to prevent a user from, either accidentally or even maliciously, causing any permanent damage to the satellite.

To tackle these issues, an approach based on three separate aspects is proposed:

- A hierarchical, three level FDIR policy is extensively used. This policy has two primary purposes. The first is that it allows for the protection of COTS components and subsystems against SEE (Single Event Effects), which are still present even if using the previously mentioned “building blocks” approach. The other important aspects are that it allows for the isolation of the external-user-accessible systems from the systems critical for satellite operation.
- The system design is based on a “building blocks” based approach, where careful component selection is exclusively employed, and the number of different components used is minimized. The common aspects of each subsystem (e.g. power regulation, local command and control, interfaces) are solved using identical electrical circuits, which are carefully analyzed for proper operation from a radiation tolerance point of view.
- Both, the on-board interfaces and the external RF interfaces, are implemented in a tightly-coupled manner. Extensive use of existing standards (e.g. CCSDS) is made for interoperability purposes, but on the subsystem level itself, subsystems are interconnected with as little overhead as possible. This allows for removal of as much overhead as possible, increasing the level of availability of the platform.

2.1 Advanced three-level FDIR policy

The primary idea behind the proposed FDIR approach is that no fault should cause the satellite to become inoperable. The approach is based on the attempt to try to isolate each potential fault as much as possible, but have many levels of backup in case this cannot be achieved. Three primary mechanisms form the backbone of this three-level FDIR policy [1].

The first mechanism is to employ redundancy at all levels where space, mass and energy constraints allow for this action. The most relevant part of the application of this aspect is the use of redundant interfaces.

Next is the use of three different types of overcurrent protection, which function both as Single Event Latchup protection as well as allow restarting of components and subsystems in case of other faults. The three levels are the Component level Latching Current Limiter (C-LCL), the Subsystem level Latching Current Limiter (SS-LCL) and the System level Latching Current Limiter (S-LCL), which form the base of the hierarchical three-level FDIR policy [2]. In addition, two levels of watchdog protection are used – a digital watchdog which monitors the individual subsystems and an analog watchdog, which is used to protect the central FPGA responsible for the operation of the subsystem.

The second level is to define an FDIR policy executor subsystem, which is tasked with assuring the minimum functionality of the satellite in case of a fault propagating past the subsystem level. The FDIR policy executor is a subsystem, which has a way to access the power distribution of the satellite, and, based on keep-alive messages from all subsystems on the satellite, can start or stop the power distribution to each individual subsystem.

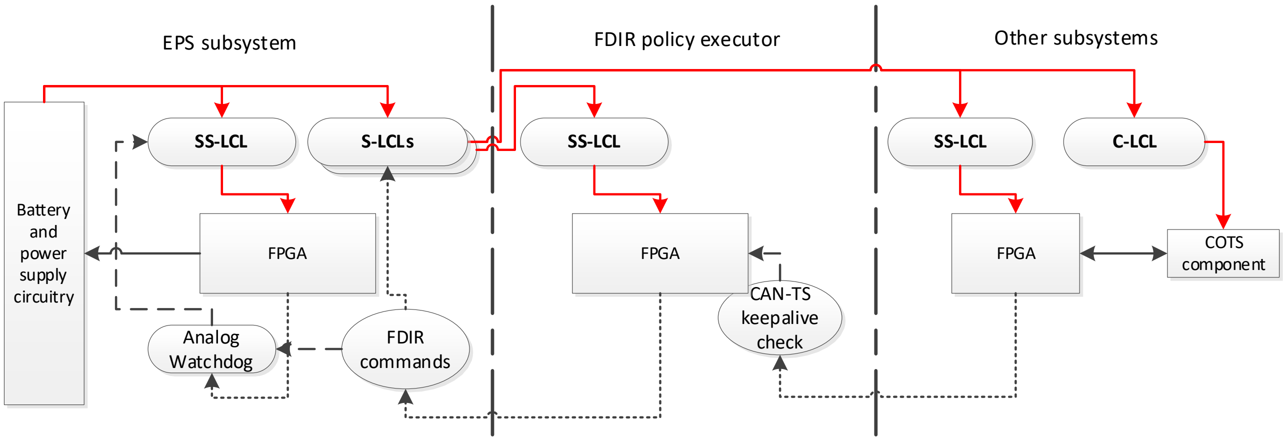

Figure 1: Presented three-level FDIR policy diagram

This approach results in a satellite that is very robust to errors. Most faults are isolated to the subsystems themselves. In case there is a fault on any subsystem, it is possible to try to mitigate it on the subsystem itself, if this is not possible, the whole subsystem can be power cycled by the FDIR policy executor. In the worst case, the whole satellite can be reset based on a watchdog loop between the FDIR policy executor and the EPS subsystem. The redundancy on the interfaces means that even if a whole interface become inoperable, it is possible to trigger the same action using a different interface.

2.2 Component selection based on a "building blocks" approach

By using a careful process to guide the component selection, it is possible to design circuits that are suitable for use in a high-radiation environment as encountered in space. The following process is proposed:

- Passive components (resistors, inductors, capacitors, diodes) and discrete bipolar transistors can be presumed immune to a TID of up to 30 krad [3].
- P-type MOSFET transistors and NPN-type bipolar transistors should be preferred where possible, as these types are more latch-up tolerant than their complements [4]. N-type MOSFETs should be avoided, as due to the fact that they are produced on a P-type substrate, a possible thyristor configuration can become active in the presence of radiation.
- More complex COTS components can be used, provided that an analysis is done into their radiation-hardness. Specifically, the most important characteristics are either the existence of radiation testing results and the fabrication process used in their production.
- The Silicon-on-Insulator – SOI fabrication process should be preferred, as it provides an inherent protection against SEL effects. Examples of this include the Texas Instruments' BiCOM or Analog Devices' XFCB process. Some components produced with a classical bipolar process can also be suitable, provided enhanced low-dose radiation sensitivity – ELDRS hardness can be assured.
- Finally, if a specific functionality cannot be effectively implemented otherwise, COTS parts of at least an automotive grade, which can also be obtained as a radiation hardened version, should be used. It can be presumed that the underlying silicon structure of the COTS part is the same to the radiation-hardened part meaning that they are more resistant to radiation than other parts, for which such assurances cannot be made.
- If the functionality cannot be achieved in any other way, use a COTS part and protect it with the three-level FDIR policy (using an LCL and a watchdog-controlled control procedure – e. g. periodic reconfiguration or periodic monitoring of functionality).

This process is especially important for the design of critical parts of the system which affect the power distribution, including the Electrical Power System, the Latching Current Limiters. These components form the primary backbone of the system, and their careful design is critical for the suitability of the platform for use in a high radiation environment.

Based on this component selection process, the common "building blocks" of each subsystem are implemented in a reliable way. For example, the subsystem-level power regulators are all based on a single block, where the only difference is the resistors used to set the voltage levels. The primary command and control blocks are based on non-volatile Flash-based FPGAs running a Fault-tolerant PicoSkyFT processor, which has been shown to be resistant to radiation effects. The local telemetry blocks are built from an identical delta-sigma ADC circuit.

2.3 Tight coupling of inter-subsystem interfaces

The external interfaces to the Space Cloud Satellite (namely, the RF interfaces) have to be standard compliant to enable interoperability with the existing Earth-based Ground Station infrastructure. For this purpose, the CCSDS standards, based on the use of the TC [5] and TM [6] protocol, in combination with Space Packet Protocol [7] is used. An important part of allowing external user access to the satellite is a well defined and implemented security policy. For this reason, the Space Data Link Security [8] and its associated Extended Procedures [9] are used.

In contrast, the on-board interfaces do not have to be standards-compliant to such a high level, as the user is never directly interfacing with them, only using them to support their target application. As such, the on-board interfaces consist of two types of interfaces. The primary bus that connects all the systems to one another is the CAN bus. The protocol that is used is the custom CAN-TS protocol. Additionally, for direct, high throughput applications, a LVDS interface is used for point-to-point connections. The LVDS protocol is based on a a custom 8b10b protocol named LVDS-TS.

The low-level nature of the CAN-TS protocol means that all the necessary functions can be implemented directly (e.g. one command one action). This allows for tight integration between the CAN-TS protocol and the CCSDS stack, which means that operating the satellite is equivalent to issuing commands as any other system on the satellite. This allows for efficient commanding of the satellite, as the procedures are equivalent regardless of the source of the commands. Additionally, both CAN-TS and LVDS-TS define efficient data burst transfers of up to 512-byte chunks. This means that an efficient protocol to transfer data is also present. A fixed addressing scheme is used – each subsystem has a custom address region of multiples of 512-byte chunks, which is used to transfer and map data. This allows a simple on-board addressing scheme, where the complexity of the location of data storage is managed on ground and enforced on a subsystem level.

Another factor for the efficient use of the on-board interfaces is the fact that they are tightly coupled to the communication stack of the communication systems. The CCSDS protocol was supports the COP-1 protocol [10]. In this way, the uplink of the satellite can be operated in a fully managed mode – when a message is acknowledged to the GS, it means that it was successfully processed on board [11]. This allows for efficient communication with the satellite, where the low-level protocol guarantees message order and execution. In this way, it is not necessary to explicitly check every action in a command script, but only the general result after each action.

3 MODULARITY OF SATELLITE HARDWARE

The proposed approach is based on a modular satellite platform, which consists of:

- An Electrical Power System (which includes the solar panels, the batteries and all related electronics for power management and distribution.
- An Attitude and Orbit Control System, which includes all the required sensors and actuators to perform attitude determination and control.
- A high-throughput RF transceiver to facilitate the communication of the satellite with the Earth-based Ground Stations, over a secure communication link.
- A Mass storage system to store raw and processed data.
- A High-Performance On-Board Computer to perform the data processing and schedule payload activities.
- Payloads, which gather the user requested data and perform a limited amount of data pre-processing.

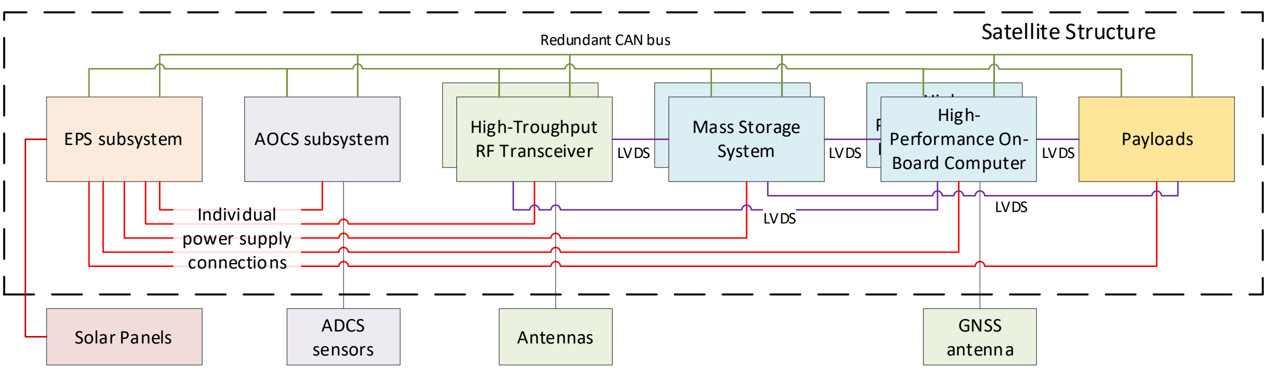

Figure 2: Modular satellite platform for Space Cloud Satellites

The modularity of the satellite platform comes from the fact that it is possible to achieve scalability of the approach trough the utilization of multiple individual units on one satellite. The necessary subsystems, which can be used in multiple instances, include the High-Performance On-Board Computer, the Mass Storage System, and the High-Throughput RF Transceiver supporting a secure communication link. The rationale for the scalability includes:

- For the High-Performance On-Board Computer, multiple units may process different data simultaneously, improving availability.
- For the Mass Storage System, multiple units improve the data storage capacity of the satellite.
- For the High-Throughput RF Transceiver, it is possible to utilize multiple transceivers on different frequencies to improve communication bandwidth.

Additionally, all multiple units can also be thought of as redundant units and used to improve reliability of the platform.

3.1 High-Performance On-Board Computer

The NANOhpc-obc is a high-performance microcontroller in a single-board computer, designed for LEO applications. NANOhpc-obc provides a versatile design in terms of variety of resources, extension possibilities and available interfaces.

It is based on a RISC-V 64-bit processor cluster (RV64IMAFDC – four processor cores) in a PolarFire SoC FPGA, which includes 32 kB L1 instruction cache with SECDED and 32 kB L1 data cache with SECDED. Additionally, a RISC-V 64-bit monitor processor (RV64IMAC – single processor core) with 16 kB L1 cache with SECDED is also present. The processor is capable of achieving a CoreMark score of 1875 or 1.7DMIPS/MHz while running at 600 MHz. The commercial grade PolarFire SoC FPGA it is based on features a SEU immune FPGA fabric, with a TID tolerance on the fabric of over 300kRad [12].

The NANOhpc-obc itself can be used as a single board computer, or in dual or even multiple redundant configurations. It integrates a Supervisor Module, which is tasked with supervising the RISC-V operation, gathering critical housekeeping data and performing a reconfiguration in case a serious anomaly is detected. The NANOhpc-obc features 2GB of LPDDR4 memory (ECC protected) and 2 GB NVM Flash storage, in 1GB redundancy configuration (EDAC protected).

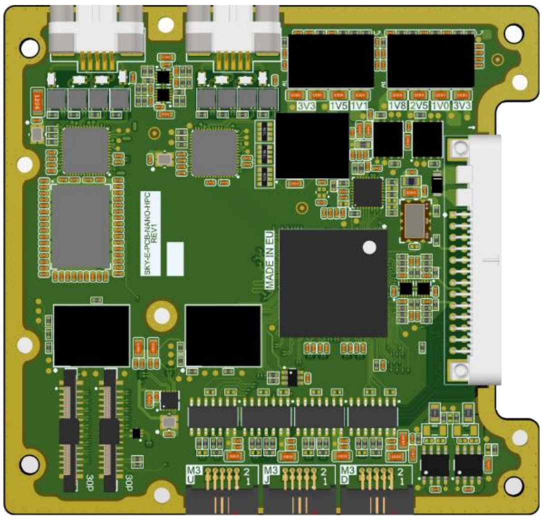

Figure 3: NANOhpc-obc High Performance On-board Computer

3.2 Mass Storage System

The Mass Storage System is a radiation hardened by design highly miniaturised, high density NAND flash based solid state mass memory system. The system is designed for autonomous memory management operation and exposes a redundant CAN bus, which is used as the primary TMTC interface. this primary interface is intended to gather general board telemetry, storage telemetry and issue access restrictions, memory management commands and unit management commands. The data interface is implemented over a redundant LVDS with bitrates of up to 100 Mbits and supports single page access and burst access.

Mass memory units can be clustered in a multiple unit storage solution enabling device redundancy and storage space expansion. The storage cluster can include a data link hub, expanding the primary bus interface to be able to serve multiple subsystems.

The unit has a built-in hardware accelerated memory management featuring address translation, garbage collection, bad block management and wear levelling. The memory controller incorporates a matrix based ECC scheme with Hamming codes and Read Solomon codes. The data is periodically scrubbed and automatically rewritten if user defined symbol error threshold is achieved. The maximum achievable raw access speed is 250 MB per second and the maximal raw memory size of the module is 1TB.

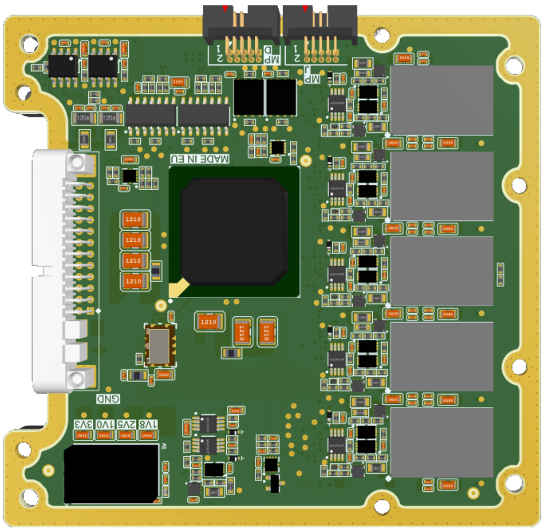

Figure 4: Mass memory subsystem

3.3 High-Throughput RF Transceiver

The NANOlink subsystem is a high-throughput CCSDS compliant communication module. It is built in a nanosatellite compatible PC-104 form factor, which consists of a primary board and an add-on RF amplifier module for the higher output power. In addition, NANOlink provides a communication channel for TM/TC via redundant CAN-TS bus or high-speed LVDS interface.

The NANOlink features an adjustable SDR-based transceiver, with a raw bitrate capability of up to 4 Mbps using O-QPSK modulation. It features full support for the CCSDS TM and TC packet protocols and the Space Packet Protocol. Additionally, it offers an out-of-the-box security solution in the form of support for the CCSDS SDLS and CCSDS SDLS-EP standards.

The NANOlink integrates an LNA on the RX path and a PA on the TX path with a power output of up to 30 dBm in the base configuration, with an optional boost addon to boost the power output to

37 dBm. An additional optional diplexer addon is supporter, featuring a splitter/combiner on two antenna ports, a diplexer connected to the splitter/combiner, an additional TX filter and an additional RX filter coupled with an LNA.

Figure 5: NANOlink-boost-dp High-throughput RF Transceiver Subsystem

4 USER PERSPECTIVE OF THE SPACE CLOUD

From the user perspective, the Space Cloud enables users to perform activities that typically require investments into satellite infrastructure. Today's application that are based on the acquisition of Earth Observation data from space are limited primarily by the available satellite communication bandwidth, which, due to the lack of on-board data processing, imposes a hard cap on the amount of data that can be generated. An additional limit, is, due to the limited forms of use of data acquisition (e.g. services that are geared towards specific use cases that focus only on image acquisition), that data is available with a certain delay, that usually cannot be affected by the user – this imposes a severe limit on any applications that require real-time satellite data.

Solutions to the previously mentioned bandwidth limit have begun to emerge in the form of the use of AI algorithms on spacecraft (AI@Edge concept). However, a major limitation of this approach is the premise that all space processing activities that are required for the varied space-data-based applications can be reduced to a machine learning algorithm, which is what we normally understand under the modern term of AI algorithms. Due to this premise, users only have access to a very small part of a satellite's resources and as such are quite limited in the types of processing that can be performed on-board. As such, our assertion is that the AI@Edge approach is inherently flawed as it is too limiting to the user. Instead, we propose an approach to give the user broader access to the satellite's resources, which is backed by the FDIR policy and satellite platform previously presented.

From the user's perspective, the advantages of this approach are numerous – they do not require to invest into satellite infrastructure, which means that the user is free to focus on the tasks that are actually relevant to their application. The satellite maintenance and critical operations is handled by the cloud operator, meaning that user's resources do not need to be spent also on satellite command and control. The direct access to the satellite's high-performance computing resources allows the user direct control over the most key areas of data processing, allowing fine-grained control over the amount of data processing (directly correlated to required downlink data to be transferred) and the amount of time that passes between data acquisition and data downlink.

An envisioned Space Cloud usage scenario consists of the five discrete steps:

- First, the user identifies an application that must be executed on a satellite. He prepares all the necessary development effort to support the execution of this application on the satellite. Finally, he identifies the regions where he would like to make use of the satellite's resources.
- Afterwards, the user locates a ground station or a ground station provider through which the algorithm will be uploaded to the satellite and through which the requested activities will be scheduled. He then performs the satellite configuration and scheduling process.
- A configured satellite performs all the payload-related activities that were scheduled and stores the raw data in the mass storage memory.
- Afterwards, the satellite runs the uploaded algorithms, processing the raw data based on the user's algorithm and stores the processed data in the mass storage memory.
- Finally, the user locates a ground station or a ground station provider through which the processed data will be downlinked to Earth. The user downlinks this data and uses it according to his requirements.

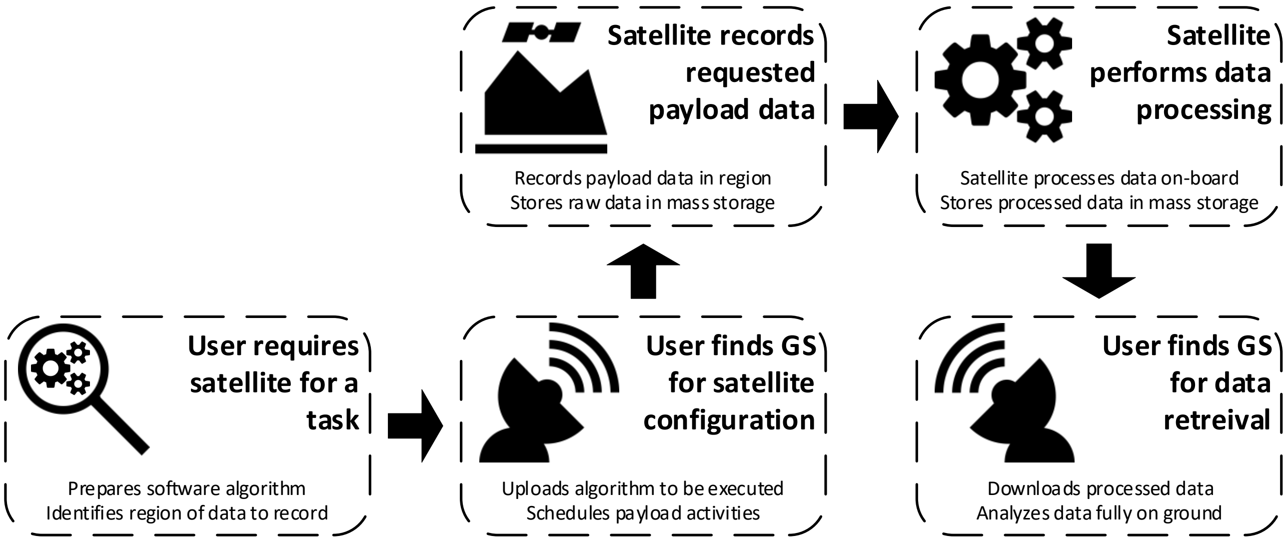

Figure 6: Space Cloud usage scenario

A final important advantage of the presented Space Cloud is the fact that, in the same way that the user does not need to invest in space infrastructure, there is also no need to invest in ground station infrastructure. As the Space Cloud is provisioned on demand, it makes sense to also provision the Ground Station infrastructure on-demand, using one of the many Ground Station providers offering an on-demand service.

5 CONCLUSION

The presented proposal for a Space Cloud constellation is a potential first step in the democratization of access to space. By giving external users access to computing resources of a satellite, it is possible to foster a rapid development of new applications based on on-board payload data processing and real-time data dissemination, bringing a paradigm shift in the way we think of satellite-supported applications.

The Space Cloud constellation is primarily based on the work on the presented “building blocks” based approach, by which it is possible to improve the reliability of COTS-based electronic systems for space to high enough level, where it is possible to use such systems in demanding high-availability satellite applications. The proposed three-level FDIR process is required to give user the access they require – it allows external code to run on the system, while keeping all the critical functions effectively isolated and transitioning the whole satellite to a safe state in case of any anomalies. The tightly coupled interfaces additionally improve the availability of the platform, while also allowing interoperability with existing ground station providers, while the use of CCSDS SDLS and CCSDS SDLS-EP protocols enables secure communication with the satellite even over ground station provider networks.

Finally, the satellite model presented forms a modular base as a satellite to be part of the proposed Space Cloud. The inclusion of a modular High-throughput RF transceiver, High-performance On-board Computer and Mass Storage System forms the base of the high-performance computing resources required to power the most demanding on-board data processing tasks.

6 REFERENCES

[1] D. Selčan, G. Kirbiš, and I. Kramberger, “*IMPROVING RELIABILITY OF NANOSATELLITES USING A LOW LEVEL SEE PROTECTION APPROACH*,” presented at the 4S Symposium, 2018.

[2] D. Selčan, G. Kirbiš, and I. Kramberger, “*Nanosatellites in LEO and beyond: Advanced Radiation protection techniques for COTS-based spacecraft*,” *Acta Astronaut.*, vol. 131, pp. 131–144, 2017.

[3] D. Sinclair and J. Dyer, “*Radiation Effects and COTS Parts in SmallSats*,” presented at the 27th Annual AIAA/USU Conference on Small Satellites, 2013.

[4] M. Maher, *Radiation Owners Manual - Issues, Environments, Effects.* National Semiconductor Corporation, 1999.

[5] “*CCSDS 232.0-B-4 - TC Space Data Link Protocol.*” CCSDS Secretariat, Sep. 2010.

[6] “*CCSDS 132.0-B-2 - TM Space Data Link Protocol.*” CCSDS Secretariat, Sep. 2003.

[7] “*CCSDS 133.0-B-1.2 - Space Packet Protocol.*” CCSDS Secretariat, Sep. 2012.

[8] “*CCSDS 355.0-B-1 - Space Data Link Security Protocol.*” CCSDS Secretariat, Sep. 2015.

[9] “*CCSDS 355.1-B-1 - Space Data Link Security Protocol—Extended Procedures.*” CCSDS Secretariat, Feb. 2020.

[10] “*CCSDS 232.1-B-2.1 - Communication Operation Procedure-1.*” CCSDS Secretariat, Sep. 2015.

[11] A. Di Stefano, G. Fiscelli, and C. G. Giaconia, “*An FPGA-based software defined radio platform for the 2.4 GHz ISM band*,” *IEEE PRIME 2006 IEEE Ph Res. Microelectron. Electron.*, vol. 12, pp. 73–76, 2006.

[12] J. J. Wang, N. Rezzak, F. Hawley, G. Bakker, J. McCollum, and E. Hamdy, “*RADIATION CHARACTERISTICS OF FIELD PROGRAMMABLE GATE ARRAY USING COMPLEMENTARY-SONOS CONFIGURATION CELL.*” Microchip.