Architecture for Cognitive Networking within NASA's Future Space Communications Infrastructure

Gilbert J. Clark III[*] and Wesley M. Eddy[†]
MTI Systems, Cleveland, OH, 44135, USA

Sandra K. Johnson[‡]
NASA / Glenn Research Center, Cleveland, OH, 44135, USA

James Barnes[§]
MTI Systems, Cleveland, OH, 44135, USA

and

David Brooks[**]
Vantage Partners LLC, Cleveland, OH, 44135, USA

Future space mission concepts and designs pose many networking challenges for command, telemetry, and science data applications with diverse end-to-end data delivery needs. For future end-to-end architecture designs, a key challenge is meeting expected application quality of service requirements for multiple simultaneous mission data flows with options to use diverse onboard local data buses, commercial ground networks, and multiple satellite relay constellations in LEO, MEO, GEO, or even deep space relay links. Effectively utilizing a complex network topology requires orchestration and direction that spans the many discrete, individually addressable computer systems, which cause them to act in concert to achieve the overall network goals. The system must be intelligent enough to not only function under nominal conditions, but also adapt to unexpected situations, and reorganize or adapt to perform roles not originally intended for the system or explicitly programmed. This paper describes architecture features of cognitive networking within the future NASA space communications infrastructure, and interacting with the legacy systems and infrastructure in the meantime. The paper begins by discussing the need for increased automation, including inter-system collaboration. This discussion motivates the features of an architecture including cognitive networking for future missions and relays, interoperating with both existing endpoint-based networking models and emerging information-centric models. From this basis, we discuss progress on a proof-of-concept implementation of this architecture as a cognitive networking on-orbit application on the SCaN Testbed attached to the International Space Station.

I. Introduction

SPACECRAFT and mission architectures are growing increasingly complex and networked. Networks onboard spacecraft are connecting ever-growing numbers of instruments, flight computers, and other components. For future human spaceflight, whether on the International Space Station (ISS), or for exploration missions to the moon, Mars, or other destinations, networking technology is critical for not only on-board systems, but also tying together mission systems between the ground and space segments[1,2]. Networking technology also has

---

[*]   Software Engineer, gclark@mti-systems.com
[†]   Chief Technologist, wes@mti-systems.com
[‡]   Electronics Engineer, Advanced Communications Division, sandra.k.johnson@nasa.gov
[§]   Network Engineer, jbarnes@mti-systems.com
[**]   Electrical Engineer V, GESS-3 Communications Group, David.E.Brooks@nasa.gov

a role in formation-flying missions, fractionated spacecraft architectures, hosted payload systems, and other mission concepts. NASA's space communications infrastructure is also undergoing changes in order to keep up with the needs of its mission customers[3,4].

This paper begins with a description of relevant features of the NASA future space communications architecture, identifies the need for cognitive communications, and highlights the research challenges that we are pursuing at the NASA Glenn Research Center (GRC) in this area. Section II of this document describes architecture concepts required to meet these challenges. Section III discusses an initial intelligent routing software effort being worked at GRC, and Section IV concludes with a selection of technology gaps and future work that we believe is necessary going forward towards developing cognitive networking systems that can be integrated into SCaN's infrastructure.

A. Future NASA Space Communications Infrastructure

The NASA Space Communications and Navigation (SCaN) program operates the Space Network (SN), Near-Earth Network (NEN), and Deep Space Network (DSN) that provide critical services for many space missions. The three separate SCaN networks are currently transitioning into elements of a single integrated network, with a primary goal of reducing operations costs. Figure 1, from the SCaN architecture definition[5] provides a high-level overview of the future architecture. This new architecture is expected to be completely seamless, offering standard services, service interfaces, and scheduling across all network elements. Additionally, this architecture is expected to have the capability to support controlling any network element from any SCaN operations center, with the possibility to rotate control of specific elements from center to center[4].

The future service offerings will range across the communications stack, from raw signal services, to bitstream, link-layer frame, network / packet, and delay-tolerant networking. The network will also provide spacecraft timing or clock calibration services, tracking services, and radiometric measurements. A wide range of missions need to be supported, including human spaceflight (e.g. ISS, EM-2), near-Earth flagship missions (e.g. JWST), cubesats, and deep-space robotic exploration missions (e.g. Mars rovers). Commercial and international missions and service providers are expected to be part of the future Solar System Internet (SSI).

Major drivers for cognitive networking in this architecture include:

- Reducing the user burden will make it easier for missions to perform service management and obtain access to diverse types of services across the different network elements.
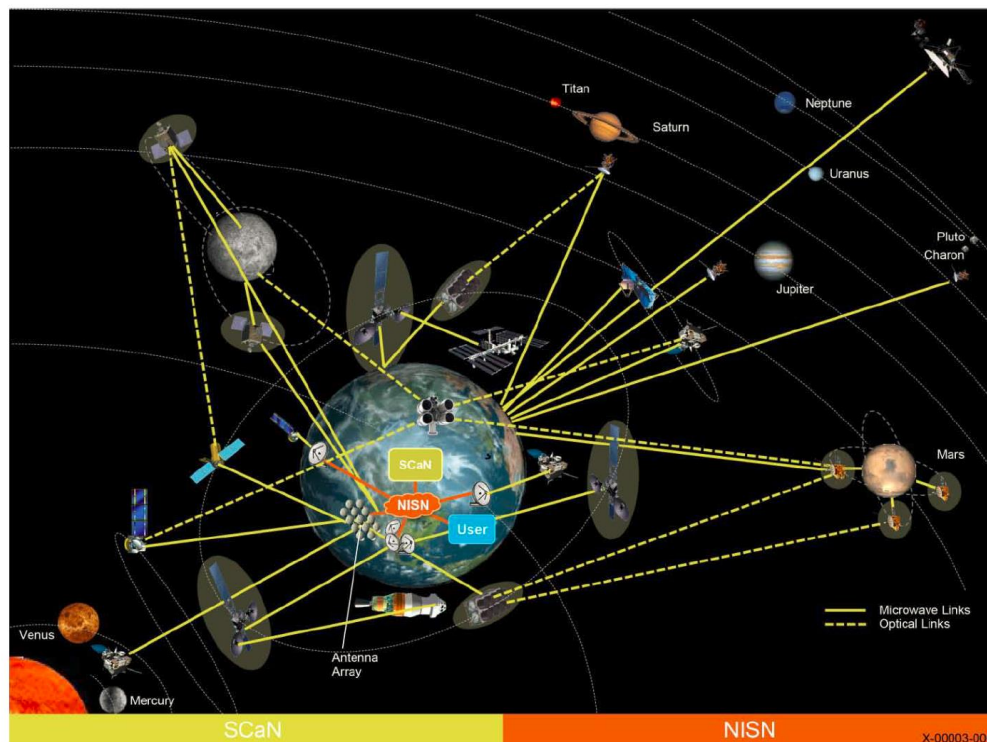

Figure 1. Future SCaN Network Architecture

American Institute of Aeronautics and Astronautics

- Mitigating operational risks due to growing complexity will be key in supporting the many different types of services, configurations, and scenarios that can take place, especially as the network grows in size. If this complexity can be managed by machines, and offload human operators in SCaN, it will allow more missions and more complex missions to be supported in a cost-effective way.
- Providing communications and navigation services in locations where humans will not frequently be in real-time control (e.g. on and around the moon and other planets) will be important. The network will need to operate semi-autonomously or fully autonomously in order to provide critical services in absence of human control.

The legacy infrastructure supporting lower-layer services will coexist with new networking services. Many resources such as spectrum, access times, etc, will continue to be managed and scheduled to accommodate legacy users that do not have cognitive radios, media access control technologies, or operations across large delays or unidirectional links. This will be the case even after cognitive networking and other technologies supporting on-demand scheduling capabilities become available for use.

In addition to ground-based antennas, space-based relay services are a part of both the current SN and the future infrastructure[6], and new relay spacecraft capabilities are currently being investigated for support of internetworking services. In order to support missions on the surfaces and in-orbit around Mars, the moon, and other bodies, there may be relays around other bodies and frequent disconnections from the terrestrial Earth-based infrastructure and operations center(s) with human operators.

Cognitive networking's flexibility to support multiple types of mission architecture with differing networking needs will be a key capability. Flagship missions are very complex, and have multiple internal fault-tolerance mechanisms, helping to increase their reliability, mission lifetime, and ability to meet mission success criteria. Alternative mission concepts have emerged, such as fractionated satellites, clusters, swarms, and other architectures involving a network of cooperating flight systems. These collections of systems meet their mission needs by working together over the network, and the availability, reliability, and performance of the network can be a major factor in mission success. Different types of cognitive technology are needed to meet the needs of specific missions, ranging across simple automation tasks, orchestration of large systems, and up to fully autonomous cooperation of distributed agents. For either flagship or small spacecraft missions, autonomous network operations will be critical, and the technologies employed for cognitive networking will be flexible for a wide range of mission needs.

B. Intelligent Routing Efforts at NASA Glenn Research Center

In the past, we have done a number of loosely-related networking experiments onboard the SCaN Testbed, in orbit attached to the International Space Station. For instance, we worked on experimentation with running IP over protocols published by the Consultative Committee for Space Data Systems (CCSDS), supported testing and profiling of Delay/Distruption-Tolerant Networking (DTN) applications, and supported multi-layer security experiments[7]. Going forward, the SCaN Testbed project has changed scope and focus to become the Cognitive Communications Project at NASA's Glenn Research Center (GRC). We have unified and correspondingly transformed our networking efforts towards intelligent routing. The NASA InTelligent ROuting (NITRO) effort is our new direction. NITRO is developing and maturing technology that will improve future space networking operations to meet the top-level challenges of being adaptive, autonomous, cross-layer, secure, and scalable.

Within the Cognitive Communications Project, there are other researchers working on waveforms and systems for adaptive coding and modulation, user-initiated services, and other technologies that will create more dynamic space links and dynamic sets of multiple space links at any given time. NITRO technology will need to adapt network routing in accordance with the variable data rates, and dynamic links, including support for opportunistic (unscheduled) contacts. Additionally, NITRO will accommodate multiple upper-layer protocols and applications, attempting to select proper routing based on the specific traffic or data flow requirements.

NITRO technology will operate autonomously, to the greatest extent possible. It will include features for self-configuration and self-management, as basic cognitive technology goals.

Because intelligent routing requires knowledge and even control of lower layer protocols and configurations, NITRO will require advanced cross-layer interfaces. Because of the mix of legacy systems and future technologies that will coexist for some time, we expect NITRO to need to interact with multiple lower-layer protocols. This includes both COTS industry and CCSDS protocols.

Security is fundamental to NITRO, and performing security functions autonomously is a key challenge. We are pursuing security techniques at multiple layers due to the architecture and need to support many possible mission protocol stacks.

A final top-level goal for NITRO is scalable operation across both high and low-rate data links, including optical, Ka-band, or other high-bandwidth trunk links that may be needed for relay spacecraft. Additionally, the processing

capabilities of network nodes is wide-ranging from wearable spacesuit radios or other embedded computers to datacenter systems on the ground. NITRO will need to be able to stitch together networks of diverse systems and meet high data-rate demands.

II. Cognitive Space Network Architecture Concepts

This section describes several important architecture features and concepts for introducing cognitive technology into space networks, especially with regard to the future SCaN infrastructure. There is a large body of work on routing technology for satellite constellations and space networks. Figure 2 attempts to put intelligent routing and cognitive networking into scope with prior work in the field[††].
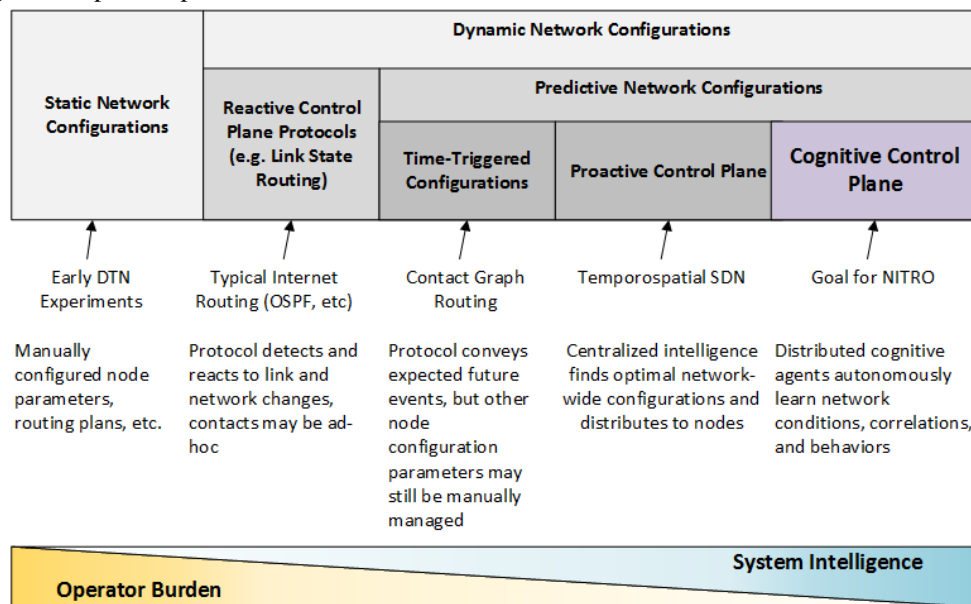

Figure 2. Progression of routing technology towards cognitive control plane capabilities

At the highest level, there is a distinction between static and dynamic network configurations. In many of our early DTN experiments, we used static network configurations. All parameters in these networks were manually configured by humans, at "design time", prior to operating. This has high operator burden, potential for error, and issues in scaling for large networks, which drives the need to move to dynamic configurations. These can be divided between reactive and predictive techniques. Typical terrestrial networks heavily use reactive protocols in the network "control plane", where real-time decision making happens regarding service performance. Reactive control plane protocols probe the network, sense changes, and alter configurations accordingly. However, they have no concept of what may happen in the future, which is where predictive techniques have an advantage. Time-triggered configuration changes can be pre-computed based on knowledge of orbits, schedules, and other physical factors or planned events. As an example, classical DTN Contact Graph Routing (CGR) distributes the set of expected future inter-node contacts throughout the network, which is used by each node to make data forwarding decisions. A proactive control plane extends time-triggering capabilities further into the control plane, and is able to use Software Defined Networking (SDN) technology to actually alter protocol behaviors across layers, and tune other parameters under guidance of a centralized intelligence. SDN technology supports automation and adaptation capabilities, but by itself does not include machine learning features. SDN can be part of an intelligent routing implementation, when used in combination with other cognitive technology.

A major goal of NITRO is to minimize operator burden for routine network management tasks. This will be achieved using learning within the network, adapting both configuration parameters and behaviors (e.g. algorithms, logic, code). The network intelligence will be able to discover novel solutions to provide the needed services, outside of the constraints of pre-coded logic present in other non-intelligent routing technologies. The rest of this section describes architectural features of a network with cognitive control plane capabilities.

[††] This is a highly simplified perspective, to fit in the available space for this paper, and a more detailed technical comparison and contrast between intelligent routing techniques and other routing techniques is future work that we are preparing.

American Institute of Aeronautics and Astronautics

A. Cognitive radio and cognitive networking

Cross-layer coordination and information-sharing beyond today's capabilities is essential. Cognitive radio[8] is an active research area, but much work to-date has been limited to dynamic spectrum access techniques. Cognitive networking[9] research builds on cognitive radio, and additionally involves higher-layers of the protocol stack. With cognitive technology, traditional static design parameters or configuration settings can now be dynamic throughout the entire communications stack.

Similar to how cognitive radio systems are built on software-defined radios (SDRs), cognitive networking systems are assumed by most recent papers to be built on SDN technology[10]. Space SDR technology is continuing to advance, and SDRs are beginning to have capabilities for supporting networking protocols in addition to traditional waveforms[11]. Some extensions and enhancements to terrestrial SDN may be needed for space[12], but SDR and SDN platforms will be crucial to enabling cognitive networking.

Multiple cognitive engines may be operating at different layers, with differing scope, and at different points in the network. Cognitive processes with wider scope (e.g. system-wide across several radio links on a relay spacecraft) may overlap and even encompass other processes with smaller scope (e.g. running on individual radio devices or links). Local and global optimization attempts by different cognitive engines will need to be coordinated efficiently. This leads to additional architecture features described in the next several subsections on multiple scopes of cognitive operations, different roles, and computational offload capabilities.

B. Multiple Administrative Domains

Because the future SCaN infrastructure is intended to include or interact with commercial and international partner systems, there can be multiple administrative domains involved in end-to-end data flows. Cognitive systems are able to communicate across different administrative domains, but may not always be able to alter or have detailed insight into systems belonging to other domains. For instance, SCaN is a service provider network, and SCaN relays and ground stations are operated independent of science mission spacecraft. While a SCaN relay may contain detailed knowledge of its radio capabilities and control over all link and routing parameters within the relay, it may not always have the ability to alter the user mission spacecraft accordingly. Missions may or may not have their own cognitive systems, and the NASA enterprise encompassing both SCaN and the missions may one day even have cognitive systems to coordinate operations.
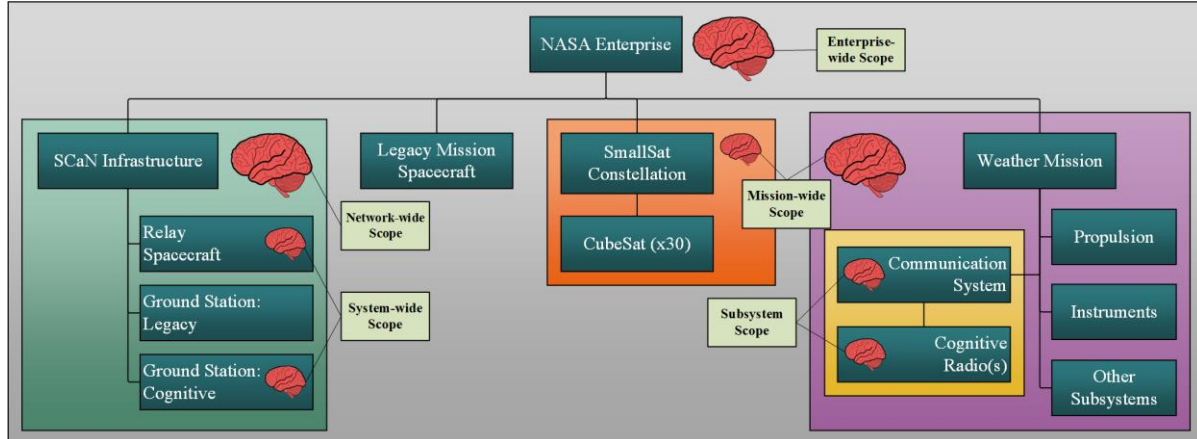

Figure 3. Domains and Hierarchy of Cognitive Systems Example

Figure 3 illustrates examples of domains with differing administrative domains, from NASA enterprise level, down through SCaN and some example mission systems, into individual relay, ground station, and spacecraft platforms. Note that this is simply an example. Not all systems need to be cognitive, incremental deployment is expected, and some legacy mission and ground station examples are included. When cognitive systems are operating in different administrative domains, and attempting to learn and optimize parameters and behaviors, it may be important for them to coordinate in order to avoid detrimental interactions. As one example, if routing decisions made by a SCaN relay are not in concert with routing decisions made by a partner relay, then temporary routing loops could result, causing high link utilization, loss of data, or other impacts. Existing control plane protocols attempt to avoid these problems, but do not learn or adjust in the same way that a cognitive system might. Protocols for cognitive systems to express intent and coordinate learning and adaptation actions are an area of future research.

C. Global and Local Cognitive Processes

Figure 3 shows brain icons indicating cognitive agents operating with different scopes within the example systems. There are two sizes of icon illustrating differing roles of "big brain" and "little brain" cognitive systems in the architecture. This sizing is related both to available hardware resources, as well as the role or function of the cognitive system. "Little brain" cognitive systems have lower computing power (e.g. because they are onboard spacecraft platforms), tighter scope, and are operating with view and control of a smaller subset of resources. In comparison, the "big brain" systems may be operated terrestrially, for instance, with wider-scope and capabilities. For instance, a "big brain" could be optimizing resource allocations across the entire SCaN network, in order to maximize some utility function (e.g. based on meeting mission support requests) while maintaining high availability, whereas a "little brain" within a cognitive radio modem could be optimizing its parameters real-time during an individual contact event. The global and local cognitive processes can work in parallel towards their different goals.

Cognitive radio systems are mostly limited to local scope, and tied to the radio (or optical) front-ends and apertures that are available to them for use. Other cognitive systems operating globally, may perform functions totally outside the radio. As an example, intelligent routing can use network optimization techniques to plan routes across multiple systems, balance load, and perform other global traffic engineering functions. Introducing cognitive technology into the service management system may allow a reduction of the role that scheduling plays in today's SCaN operations, but for existing missions and services, legacy scheduling will be necessary to maintain for at least some of the resources. As an early demonstration of cognitive technology, it could be applied to service management tasks while running in a "shadow mode" (operating in parallel to the existing systems, and without impacting them). This could identify potential scheduling efficiencies, note anomalous conditions, audit system configurations, predict future loading, and be used to evaluate benefits without the risk of letting it immediately make operations decisions while running "live".

D. Computation Offload

As noted previously, "little brains" can operate on limited computing systems (like typical flight hardware), whereas "big brain" systems may have more computing resources (e.g. more CPU cores, GPUs, or other specialized machine learning hardware). Typically, the more significant computing resources will be terrestrial, but the architecture may be able to include some shared infrastructure nodes in-space, with greater computing powers. In either case, it could be possible to offload some heavier computations, enabling the "little brains" to run expensive training and analysis computations offboard of their own limited hardware.

As an example, a small in-space cognitive system can collect measurements, logs, and other historical behavior data relevant to itself. This can be compressed and sent to a more capable computing system where training algorithms, genetic / evolutionary techniques, or other computationally expensive algorithms can be performed. The resulting outputs (e.g. neural network weights, "fittest" genetic programming results, etc.) can then be uploaded back to the smaller in-space system for execution.

E. Autonomic computing and autonomic networking

Simple automation systems (e.g. scripting) as commonly used in spacecraft operations will not be sufficient to replace the role of humans, for which more fully *autonomic* systems are required. Terminology and concepts for autonomic computing were established in seminal IBM work on the topic[13], and many autonomic computing efforts exist, including for space systems, and in terrestrial networking. IBM described a set of "self-management" concepts, including self-configuration, self-optimization, self-healing, and self-protection, and identified several engineering challenges for autonomic systems, all of which are applicable to space communications services. Of particular interest, this includes:

- Design, test, and verification – Current design techniques, heavily based on link-budget analysis, will not be sufficient when a cognitive system can alter all properties of the link(s), and can completely alter protocol configurations. Current mission compatibility testing is primarily physical layer, and may also include rudimentary link layer properties, but covers a fixed set of intended communications modes. With a cognitive radio, and additional networking layers, the scope and test methodology will need to be reconsidered.
- Monitoring and problem determination – Autonomic systems monitor their performance, detect, and solve problems on their own. However, like humans, cognitive communications systems can make mistakes. Communications, tracking, and navigation services are critical to mission success, and the SCaN elements today operate with high reliability requirements. In order to prevent risks to the mission or crew (for manned missions), one challenge is to apply the technology in ways that avoid the potential for terminal states and/or limit the impact of erroneous decisions. This has been less of an issue for many terrestrial uses of machine learning and artificial intelligence, but similar concerns apply in self-driving cars, aeronautics/UAVs, and

other emerging fields. Cognitive space systems will need to support reporting and forensic capabilities in case of issues, in addition to their internal autonomic problem-solving processes.

- Goal specification – For many simple optimization tasks, goals can easily be specified. As an example, an adaptive rate waveform might maximize the modulation and coding bitrate for science data downlink. However, more holistic network-wide goals can be harder to specify to the machines. While legacy non-cognitive elements are controlled via specific commands, cognitive systems will be directed with descriptions of operator intent, and specific actions will then be chosen by the cognitive system.

F. Flow-centric versus Information-centric Communications

Most current networking technology is "flow-centric", and makes routing decisions based on destination addresses, rather than the information content of packets. Many mission systems have used message buses and similar data distribution systems, but it has not been part of a SCaN service offering. With more complex missions, and intelligent routing able to determine packet contents and predict shared interests in data, there may be value in providing information-centric networking services in addition to purely flow-based services. As one hypothetical case, space weather information collected by some spacecraft could be shared directly with other relevant space systems, without needing to pass through many intermediate systems first. For instance, detected solar events that could be harmful to astronauts could be conveyed more quickly, particularly during operations away from Earth and omni-present communications with mission control. Information-centric networking (ICN) technology being developed for terrestrial use[14] may be adapted for space use.

G. Cognitive Systems Engineering

There are many types of artificial intelligence (AI) and machine learning (ML) technologies that can be used for different problems in cognitive networking. For instance, selecting, configuring, and optimizing protocols for particular services is similar to classification tasks typically implemented by conventional neural networks. Other problems may be more appropriate for expert systems, support vector machines, self-organizing maps, recurrent neural networks, learning automata, and/or agent-based approaches to intelligence.

Some AI techniques are computationally expensive, beyond levels that may be appropriate to implement in space. For instance, genetic programming and evolutionary algorithms are often (but not always) done with networks of dozens or hundreds of servers, with computing power far beyond what even a large relay spacecraft will be capable of. The "big brain" architecture concept, however, allows computationally intensive activities to happen offline on capable nodes, and the results to be pushed back to lesser capable nodes. This is similar, for instance, to how SDN already works in terrestrial networks, where routing solutions are computed centrally and distributed to all of the individual networking devices. Neural network training, genetic algorithms, and other operations can happen centrally with results uplinked to relays and other network nodes, including potentially to user spacecraft. This would allow cognitive networking technology to be supported on missions where the resources need to be devoted to primary science goals, and the size, weight, and power budgets can go towards science instruments or other mission systems, rather than being monopolized for cognitive communications.

In terms of learning, the use of online versus offline learning approaches for different problems is a consideration for systems engineering of cognitive communications systems. There will be large amounts of training data, including historical time-series, sets of specific "edge cases" or other pathological scenarios, regression testing data, and other sets of records important for learning and testing. The provisioning of proper data sets to support the specific learning algorithms operating in different parts of the network is an important engineering task, and may be an ongoing part of operating a cognitive network.

III. Initial Work Engineering a Cognitive Agent for SCaN Testbed

The ScaN Testbed onboard the International Space Station has three SDRs with five CCSDS AOS/ENCAP network interfaces[7]. Two of the interfaces go through S-band radios that connect directly to ground stations. Two other interfaces go through the same radios but are relayed via the SN, and one interface is routed through a Ka-band radio that relays through the SN. There are multiple service "on-ramps" on the flight system, supporting experiments directly over the CCSDS data links, experiments using Internet protocols (e.g. via IP-over-CCSDS standard), experiments using DTN, and other protocol stack configurations. Gateway systems on the ground provide similar service on-ramps, at multiple layers of the communications stack.
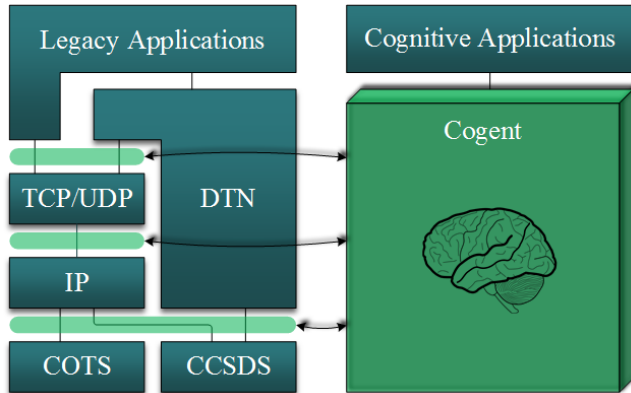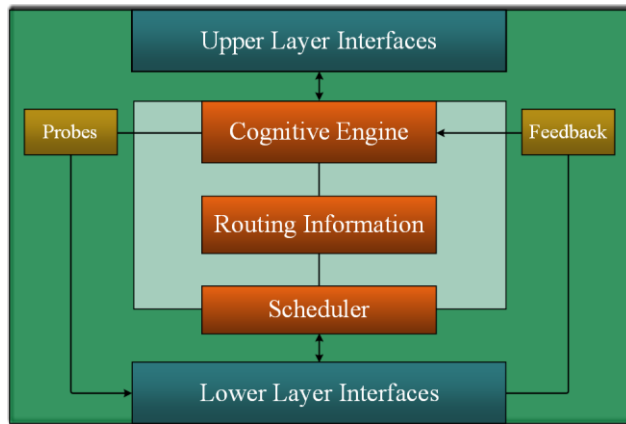
Figure 4. Cogent software stack design

Figure 5. High-Level Cogent Architecture

For NITRO, using the SCaN Testbed allows us to tap into the stack in multiple places, and make configuration, routing, and data handling decisions across the whole stack. The initial software package, called "*Cogent*", discussed in this section supports adaptive routing, multilink operations, and handles both DTN and IP-based protocols. It does not currently control the radio systems, where there are either fixed or adaptive waveforms running. It treats anything between the CCSDS ENCAP interfaces as black boxes, and probes for connectivity, estimated data rate, and other parameters per ENCAP interface. Figure 4 shows how the *Cogent* design supports multiple link types, layers, and protocol stacks, and can work for either legacy applications or new cognitive applications (e.g. that may use more of an information-centric networking paradigm).

At the boundaries between layers, *Cogent* creates virtual interfaces that represent ensembles of the lower-layer options. This allows it to operate without impacting data flows that are routed directly to the actual lower-layer interfaces. For instance, between IP and the link layers, the IP routing table is used to select an appropriate outgoing interface. Particular experiment subnets can be routed to a *Cogent* virtual interface, while other traffic continues to be routed normally. This demonstrates one way to operate cognitive networking in incremental deployment with legacy services, and with "failsafe" options to bypass the cognitive decision making.

The current revision of *Cogent* software is sensing and adapting to link properties, but is not yet examining its own decisions and implementing learning algorithms. These are planned later capabilities, discussed further with other future work in section IV of this document.

A high-level view of the system design is shown in Figure 5. Key elements of the *Cogent* design include:

1. A cognitive engine that accepts link messages as data input, and other informational input to help with decision making. There are not standards or commonality across the SCaN Testbed for communicating status of the data links. Each SDR has its own distinct telemetry format, and includes different information. The *Cogent* software uses small discovery messages to establish the presence or absence of connectivity across a pair of ENCAP interfaces. These messages are also used to probe and measure link characteristics, specifically including latency and throughput, that are measured and provided as input to the cognitive engine.

2. An API on top of the cognitive engine that accepts data messages, and an optional set of constraints (or tags) attached to them. For this prototype, only a few constraints are supported: "low-latency", "low-priority", "high-priority", and "high-throughput".

3. A scheduler algorithm sends messages on the available links. Based on user-provided tags or inference of desired service parameters, the elements in the queue, and the discovered properties of each link, messages are selected for output to specific ENCAP interfaces. A minimal wrapper protocol is used to identify and time-tag their transmission and receipt, offering a way to feedback information about success and updated link status in-band with the data stream. This information is used to update link properties and guide selections in the future.

4. A content cache to store data messages and metadata about them is available in order to support future information-centric networking, but is not currently used by the system.

American Institute of Aeronautics and Astronautics

Initial implementation for this effort has focused on enabling DTN data flows to adapt to data rate changes on the underlying links, and also providing DTN with the capability to utilize an ensemble of multiple links. Additionally, for applications with different quality of service concerns, the current iteration of *Cogent* is able to optimize radio link selection to minimize delay or maximize bandwidth. For instance, *Cogent* can select a direct-to-Earth rather than SN relay interface for low-latency, or a Ka-band rather than S-band link for high throughput.

The prototype construction of *Cogent* is primarily intended to act as a tool to help us explore this research space. The goal is to identify the needed areas for future efforts by implementing an early cognitive agent on the SCaN Testbed. *Cogent* has largely been a success in this regard. Through its design and implementation, we have been able to identify a number of gaps that we believe should be further explored to facilitate future efforts in the area of cognitive networking.

In addition to using *Cogent* as an exploration tool, we have also designed *Cogent* to collect data throughout routine operations in order for us to build a corpus for future learning. Data is collected for each of the onboard interfaces, on the flight system, and at each of the ground gateway systems. Records are offloaded (in the form of custom telemetry packets) to various destinations, including local gateways and a central database for network-wide operations. We expect to use this information to support gateway-level "little brain" operations and network-wide "big brain" operations in the future, but we are early in the process of implementing appropriate learning algorithms. We currently expect to be able to extend the *Cogent* framework in this regard for NITRO, but there are likely to be other cognitive agents developed as well: for instance, we believe network-wide scheduling of data links could benefit from cognitive support.

IV. Technology Gaps and Future Work

To conclude this paper, we have a list of technology gaps and topics for future research work, based on our efforts so far. This is not an exhaustive list, but identifies a subset of interesting and promising directions.

- Cross-layer signaling – Probing as done in the early *Cogent* prototype is feasible for many cases in the near-earth regime, but not feasible for space networks in-general. There are difficulties with large (e.g. interplanetary) delays, unidirectional link connectivity, asymmetric link properties, asymmetric routing, and other peculiarities of space communications topologies. A better cross-layer signaling mechanism is desirable so that cognitive networking processes can communicate directly with local (or remote) link layer systems. New standards or APIs may be needed in this area. Functions including discovery, monitoring, and control are relevant. Without standards, each mission or system composition will have difficulty achieving code-reuse, and will need to devote resources to interface code development rather than core cognitive engineering efforts.

- Algorithm development – We are still gaining experience with particular learning algorithms and AI techniques for specific problems relevant to space communications. While there is a large body of research for communications in general, most work applies to lower layers, and the cognitive networking field is less well-explored, even for terrestrial communications. In some cases, techniques being applied for 5G communications systems could be considered for applicability to SCaN, but there are large differences in the service offerings of 5G providers versus SCaN's service catalog, so additional algorithms are likely to be needed.

- Computational Offload – Due to computational expense of training and learning algorithms, it may be useful to develop protocols and practices that enable computational offload of some parts of the cognitive functionality. This would be similar to how some AI functions on mobile devices are offloaded to the cloud today, terrestrially.

- Debugging and management of intelligent systems – Cognitive systems may reduce operator burden significantly in nominal scenarios and conditions, but during off-nominal or contingency situations, it may be more troublesome for human operators to debug and apply remedies alongside cognitive algorithms. Failsafe techniques and practices will need to be developed for operating in conjunction with human administrators.

- Self-knowledge - One of the fundamental aspects of cognitive radio, defined by Joe Mitola, is the capability to have self-knowledge and a comprehensive internal model of the system available for adaptation. This is needed for more advanced capabilities like allowing the system to more extensively adapt and develop new protocol mechanisms, compositions, and other communications methods of its own devising. Since space systems are now often being developed with Model-Based Systems Engineering (MBSE) techniques, it may be possible to leverage MBSE artifacts as one way to describe the system to itself. Some MBSE models contain behavioral descriptions and parametric models that

could be used in simulations by a learning algorithm, as one example. Machine-readable MBSE artifacts can be modified to reflect and capture adaptations made by the cognitive system as well.

- Self-design– Learning of new protocols and behaviors is more complex than simply optimizing parameter sets, but is possible with techniques such as genetic programming, among others, that output software code or other design descriptions. One practical example is the well-known NASA Ames work using genetic programming to develop novel antenna designs[15]. Similar techniques could be extended into and across other parts of the space communications system, but has not yet been done.

In all areas of future work, solutions can be implemented piecemeal for individual systems, and multiple approaches may be valid, however, in some cases it will be beneficial to establish standards. Early-on, the most important cases for standards are in cross-layer interfaces, since there are many different device vendors, and multiple devices may be integrated into a communications system.

In summary, SCaN's future space communications infrastructure has aspects that drive the need to develop cognitive communications technologies, including cognitive networking. Cognitive networking technology is still maturing terrestrially, and applying it for intelligent routing in space will be a significant challenge. However, the benefits of cognitive networking can be worth the effort: cognitive engines can be used for everything from assisting mission operations personnel (allowing for more missions to be supported per staff member) to extending the life span and range of missions beyond that which has been historically possible (by allowing both objects and networks to adapt to unexpected circumstances without human intervention).

In this paper, we described the differences between a cognitive control plane for intelligent routing, in comparison to well-known network routing and control-plane techniques. We identified several architectural features and concepts for incorporation of cognitive networking into SCaN's architecture. We described early prototype code, currently addressing a small subset of the goals for cognitive networking, but with the ability to grow more features over time as we continue research in this area. Finally, we have identified several fruitful research areas to fill the current technology gaps in future work.

References

[1]Bhasin, K., Putt, C., Hayden, J., Tseng, S., Miller, R., Hudiburg, J., Miller, D., Jeffries, A., Sartwell, T., Biswas, A., Kennedy, B., and Jennings, E., "Architecting the Communication and Navigation Networks for NASA's Space Exploration Systems", *Proceedings of the 2007 IEEE System of Systems Engineering (SoSE) Conference*, 2007.

[2]Schier, J., "An Update on NASA's Space Communication Architecture from 2010 to 2030", *INCOSE INSIGHT*, Vol. 11, No. 5, pp. 26–31, December, 2008.

[3]Liebrect, P., Schier, J., Bhasin, K., Bibyk, I., Butler, M., Hudiburg, J., Tai, W., Shames, P., "NASA's Space Communications Integrated Architecture", *Proceedings of the AIAA SpaceOps 2010 Conference*, Huntsville, Alabama, 2010.

[4]Tai, W., Wright, N., Prior, M., and Bhasin, K., "NASA Integrated Space Communications Network", *Proceedings of AIAA SpaceOps 2012 Conference*, Stockholm, Sweden, June, 2012.

[5]NASA, "Space Communications and Navigation (SCaN) Network Architecture Definition Document (ADD) Volume 1: Executive Summary", Revision 2, SCaN-ADD-V1, October, 2011.

[6]NASA, "NASA's Space-Based Relay Study: Overview and Direction", white paper posted to Internet, March, 2013, https://www.nasa.gov/sites/default/files/files/SBRSWhitePaper3-7-2013.pdf

[7]Brooks, D., Clark, G., Eddy, W., and Johnson, S., "In-Space Networking on NASA's SCaN Testbed", *Proceedings of the 34th AIAA International Communications Satellite Systems Conference (ICSSC)*, October 2016.

[8]Mitola, J., and Maguire, G., "Cognitive radio: making software radios more personal", *IEEE Personal Communications Magazine*, vol 6, number 4, pp. 13-18, August, 1999.

[9]Thomas, R., DaSilva, L, and MacKenzie, A., "Cognitive networks", *Proceedings of the First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, 2005.

[10]Zorzi, M., Zanella, A., Testolin, A., De Grazia, M., and Zorzi, M., "Cognition-Based Networks: A New Perspective on Network Optimization Using Learning and Distributed Intelligence", *IEEE Access*, volume 3, pp. 1512 – 1530, 2015.

[11]Kacpura, T., Eddy, W., Smith, C., and Liebetreu, J., "Software Defined Radio Architecture Contributions to Next Generation Space Communications", *2015 IEEE Aerospace Conference*, Big Sky, Montana, 2015.

[12]Barritt, B., and Eddy, W., "Temporospatial SDN for Aerospace Communications", *Proceedings of AIAA SPACE 2015 Conference and Exposition*, Pasadena, California, 2015.

[13]Kephart, J., and Chess, D., "The Vision of Autonomic Computing", *IEEE Computer*, January, 2003.

[14]Ahlgren, B., Dannewitz, C., Imbrenda, C., Kutscher, D., and Ohlman, B., "A Survey of Information-Centric Networking", *IEEE Communications Magazine*, vol. 50, no. 7, pp. 26-36, 2014.

[15]Hornby, G., Globus, A., Linden, D., and Lohn, J., "Automated Antenna Design with Evolutionary Algorithms", AIAA SPACE Conference and Exposition 2006 (SPACE 2006), San Jose, California, 2006.