Survivability and Scalability of Space Networks

Md. Shohrab Hossain, Mohammed Atiquzzaman
Telecommunication and Network Research Lab
University of Oklahoma, OK 73019
Email: {shohrab, atiq}@ou.edu

William Ivancic
NASA Glenn Research Center
Cleveland, OH 44135
Email: wivancic@grc.nasa.gov

***Abstract*—Network survivability, which reflects the ability of a network to continue to function during and after natural or man-made disturbances, is a crucial aspect for any kind communication. It goes beyond security and fault tolerance to focus on delivery of essential services and rapid recovery of full services when situation improves. Wireless and space networks have the challenge of survivability, since users and network entities are mobile and the communication channels are accessible to anyone. Network scalability reflects the ability of a network to support increased amount of load. Future Low Earth Orbiting spacecrafts will contain several IP-enabled devices, such as Earth observing equipment, that will be accessible by terrestrial users through the Internet. Satellites are being used to capture real-time images, video for various purposes, such as, observing the Earth, weather data, live images for tornado, cyclones, tsunami, etc. In future, these data will be accessed by terrestrial users through the Internet. As number of Internet users are growing very rapidly, such satellite services will have a significantly large number of client requests which will overload the bandwidth-limited satellite links. Therefore, the survivability and scalability of space networks may become a major issue and this should be addressed in order to provide seamless services to ongoing communication to all users irrespective of mobile or stationary. However, most of network survivability and scalability analysis are based on terrestrial networks. We have listed various approaches used in those analysis. This paper clearly indicates the issues regarding survivability and scalability of space networks that need to be addressed.**

I. Introduction

Network survivability is a very important and crucial aspect for any kind communication. It reflects the ability of a network to continue to function during and after natural or man-made disturbances. It goes beyond security and fault tolerance to focus on delivery of essential services and rapid recovery of full services when situation improves. Mobile and Wireless networks have the challenge of survivability, since users are mobile and the communication channels are accessible to anyone.

Satellite communication has its vital application in telephony, weather forecasting, satellite television, in-flight Internet, navigation (GPS) and military communications. Satellite Internet can serve as an alternate means to connect aid workers and troops to coordinate rescue and recovery missions in case of catastrophic events, such as, massive earth quakes, tornados. ISPs can use satellite technology such as, very small aperture terminal (VSAT) that connects remote earth stations to satellites in geosynchronous orbit. Spacecrafts with sensing elements, such as, microwave imager, Earth radiation sensor, lightning imaging sensor, etc. are used for observing the Earth, surveillance, and monitoring. Data are periodically downloaded from the spacecrafts using dedicated links with ground stations. The spacecrafts work autonomously, downloading data to Earth whenever they come in contact with a ground station.

The research work reported in this paper was supported by NASA Grant NNX06AE44G.

Modern communications satellites use a variety of orbits including Geostationary Orbits (GEO), Medium Earth orbit (MEO) and Low Earth Orbits (MEO). A constellation of spacecrafts (such as Iridium, Globalstar, Disaster Monitoring Constellation (DMC), GPS, etc.) form space networks where the spacecrafts can communicate among themselves using inter-satellite links, and also switch data between other spacecrafts and ground stations. The continuous movement of the spacecrafts relative to Earth (such as, LEO satellites) requires IP-mobility management protocols for managing the handoff of connections between ground stations on Earth as spacecrafts have IP-enabled devices that are stand-alone devices or a collection of devices connecting to an onboard LAN to form a mobile network [1].

Increasing demand for mobility in wireless data networks has given rise to various mobility management schemes. IETF proposed Mobile IPv6 [2], Hierarchical MIPv6 [3] to support host-mobility, NEMO Basic support protocol [1] to support network mobility, allowing a TCP connection to remain alive while mobile nodes are on the move. NASA has been investigating the use of Internet protocols for space communications [4]–[6] and handover management [7], [8] for quite some time. A number of projects studied the possible use of Internet technologies and protocols to support all aspects of data communication with spacecrafts [9], [10].

Different mobility agents, such as Home Agent (HA), Foreign Agent (FA) are responsible for mobility management. With the growing importance, necessity and popularity of mobile computing, these agents are now required to serve increasing number of mobile nodes. Having a single HA or FA to serve all the mobile nodes of a network can affect the reliability and performance of the mobility protocols. Moreover, too much load on a HA or FA may cause several problems to the agents. Thus the mobility agents could be the performance bottlenecks for IP-mobility protocols in the near future. So the ability of wireless network infrastructures to handle the growing demand of data communication is now questionable.

Communication network must have a high survivability and scalability in order to provide reliable services to increasing number of end users, requiring satellite access for various military and commercial applications. There have been research works on the network survivability for terrestrial networks and the papers studying network survivability for satellite network is few. The approaches used for improved survivability and scalability in terrestrial networks can be used in space networks though there exists issues and challenges in space networks. In this paper, we have presented a comprehensive survey of the approaches used for improved survivability and scalability in terrestrial networks after classifying them into various groups. The issues and challenges to apply those approaches to space networks are also illustrated.

The rest of the paper is organized as follows. Section II explains briefly host and network-mobility protocols. Section III explains the use of IP-mobility protocols in space network. In Section IV, existing works on survivability and scalability are classified according to the approaches used. Section V explains issues and challenges relating to survivability and scalability analysis of space network. Finally, Section VI has the concluding remarks.

II. IP-mobility Protocols

Increasing demand for mobility in wireless data networks has given rise to various mobility management protocols. Mobility management protocol aims at providing essential technology to allow mobile users change their point of attachment without affecting an ongoing communication. Mobility management thus require signaling messages to be exchanged among various mobility agents to keep track of mobile nodes current locations. There exists many mobility management schemes [11] for cellular networks. However, Next Generation Wireless Network (NGWN) is supposed to be all IP-based unified network which works differently than cellular network. Therefore, it is essential to analyze the survivability and scalability of IP-mobility protocols. There exists two types of mobility protocols: host-mobility and network mobility protocols.

A. Host-mobility

There are several host-mobility protocols. IETF proposed Mobile IPv6 [2], Hierarchical Mobile IPv6 [3]. Mobile IP protocol provides simple solution for IP-mobility support by forwarding packets through Home Agent (HA). However, Mobile IP has several limitations: inefficient routing, high packet loss, handover latency, changes in Internet infrastructure. On the other hand, the IP-diversity based host-mobility protocol, SIGMA [12] ensures uninterrupted connectivity using make-before-break strategy through its IP-diversity feature.

The host-mobility protocols work as follows. As shown in Fig. 1, the satellite can be considered as an Mobile Host (MH). The satellite's footprint is moving from ground station A to B, while the satellite is bound with an IP address from ground station A. During movement, the satellite should maintain continuous connectivity with ground stations on earth. Thus, the IP address of the satellite has to be changed when it is handed over to ground station B. Whenever the Satellite acquires new care-of-address from ground station B, it informs the HA about the new care-of-address. So whenever any Correspondent Node (CN) wants to communicate with the satellite, it sends query message to the HA to find out the current location of the satellite. The HA replies the query message with the current CoA of the Satellite. The CN then can send setup and data packets to the Satellite for communication.

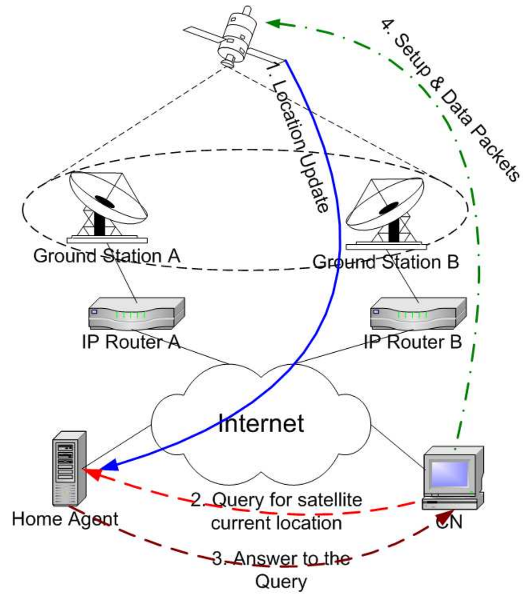

Fig. 1. Satellite as a Mobile Host.

B. Network-mobility

Mobility of a network in motion can be managed in an aggregated way using NEMO protocols [1] where Mobile Router (MR) act as gateways for the nodes inside the mobile network and ensures their Internet connectivity when the MR changes its point of attachment while moving from a home network to a foreign network. In Fig. 2, a mobile network can be formed with the on-board IP-enabled devices, laptops of an aeroplane and in-flight Internet connectivity can be provided. Here MR communicates with HA via the satellite link and data is transmitted to the CN through the ground stations as shown in the figure.

III. IP-mobility in Space Networks

Space networks include satellite communication networks composed of satellite constellations, in-flight mobile network (inside aeroplanes, helicopters) can take advantage of IP-mobility protocols to maintain Internet connectivity in next generation network.

IP-mobility protocols can manage host-mobility and network-mobility in space network. Satellite with IP-enable device transmitting or receiving data is an example of host-mobility in space network. In-flight Internet connectivity is an

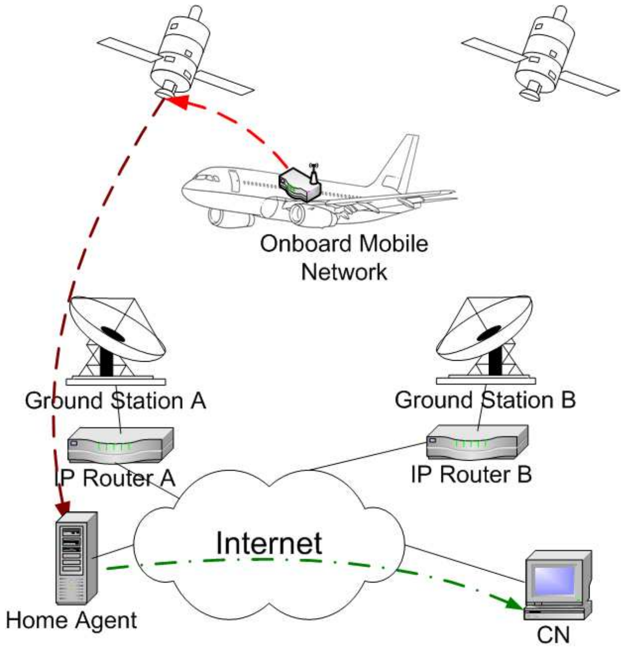

Fig. 2. In-flight Internet connectivity through satellites.

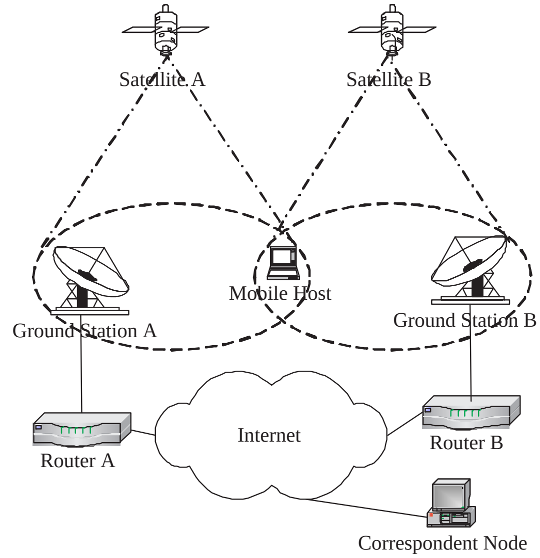

Fig. 3. User handover between Satellites where satellite act as routers.

example of network-mobility where a high capacity mobile router communicate with satellite transponders and ground station while providing Wi-fi in the aircraft.

A. *Satellite as a Mobile Host / Network*

Satellites can act as communication endpoints with onboard IP-enable device which exchange data with ground stations on earth (see Fig. 1)). In fact, multiple onboard IP-enabled devices on the Satellites can form mobile network and a mobile router (with high transmission capacity) can manage the mobility of all the hosts in an aggregated way using NEMO protocol.

B. *Satellite as a Router*

As shown in Fig. 3, satellites do not have any onboard equipment to produce or receive data; rather they merely act as routers in the Internet. Each satellite can be assigned an IP address prefixes, and they can provide IP-connectivity to Mobile hosts in other spacecrafts (such as, laptops in aeroplane, helicopters, etc.) or in remote location on earth. Hosts are handed over between satellites as they come under the footprint of a new satellite.

IV. Research on Survivability and Scalability

With the proliferation of wireless and mobile networks, the ability of network infrastructure to manage such growing demand is now questionable which raise the survivability issue of wireless and mobile network.

To improve survivability of wireless networks, there have been several research works. These works can be grouped into different categories: redundancy based schemes, load Balancing schemes, removal of FA, using Location Register, hierarchical approach, refresh-based schemes, DNS as mobility agent and quantitative approach.

A. *Redundancy-based schemes*

Redundancy based schemes [13]–[19] focus on the vulnerability issues of mobility agents that may be the single point of failure. As shown in Fig. 4, multiple redundant mobility agents are used in these schemes as backup for primary mobility agents (such as home agent) and they are all usually synchronized. In case of failure of the primary mobility agent, another redundant (backup) mobility agent takes charge of the mobility management. In IP-mobility protocols, mobility agents (such as, home agent, mobility anchor points, etc.) are potential points of failure that may lead to loss of network connectivity to all the mobile nodes in the wireless network. The redundancy based schemes generally work in two ways: Static and dynamic approaches.

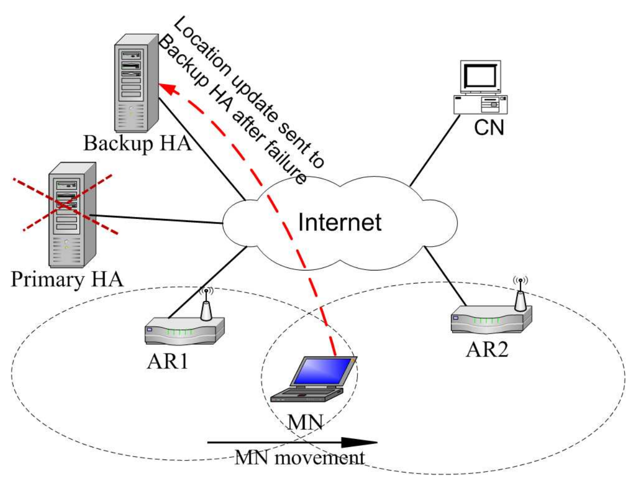

Fig. 4. Redundancy based schemes.

1) Static redundancy: In this approach, vulnerable mobility agent is statically equipped with one or more redundant agents as its backup set [13], [15], [18], [19]. The mobility agent can cooperate with its backup set to work in the standby or load-sharing model. If a mobility agent fails, one member in its backup set will be selected to act as the primary mobility agent.

In [13], multiple home agents are deployed to enhance survivability of the mobile network and they are synchronized through exchanging periodic messages. This can reduce the recovery time when primary HA is under attack or not available. Hot Standby Router Protocol (HSRP) [18], developed by Cisco provides enhanced robustness and transparent recovery mechanism from the first hop failure in network devices. Virtual Router Redundancy Protocol (VRRP) [19] is an IETF standardized mechanism for redundancy among home agents.

2) Dynamic redundancy: Second, on detecting a failure in a MA, a failure-free mobility agent is dynamically selected to act as the new mobility agent on its behalf [14], [16], [17].

You et al. [14] proposes Robust Hierarchical Mobile IPv6 (RH-MIPv6) that provides fault tolerance and robustness in HMIPv6 networks. In RH-MIPv6, each mobile node registers primary and secondary regional care-of-address to two different MAPs (primary and secondary) simultaneously. In case of failure of primary MAP, packets are sent through the secondary MAP, thus reducing packet loss rate significantly.

B. Load balancing schemes

Theoretically, a home network can host unlimited number of mobile nodes. While these nodes away in the foreign network, one single home agent has to keep the location database updated after processing registration messages, forward IP packets to the nodes. As the number of mobile nodes (MNs) in a network increases, the load on the mobility agent (HA) also increases, raising several scalability issues, such as degrading the performance the MAs with increased delay and packet loss. In order to scale large number of mobile nodes in a network, the load of the mobility agents need to be distributed among multiple agents. In Fig. 5, the load of mobile nodes are distributed between two HAs; mobile nodes A_1 and A_2 sends binding updates to HA_1 whereas mobile nodes B_1 and B_2 sends binding updates to HA_2. Thus the signaling load of all the mobile nodes can be distributed among multiple mobility agents.

There are a number of load balancing schemes of mobility protocols in the literature [13], [15], [17], [20]–[24] where the load on mobility agents are distributed among two or more redundant agents resulting in performance improvement of the whole network. This can be done in two ways: Statically and dynamically.

1) Static: In the static load balancing approach, N Home Agents are kept and load balancing is distributed among these N agents statically. So when an agent fails, the MNs attached to that HA cannot communicate with others.

Huang et al. [15] proposed a distributed protocol that maintains double mobility bindings in the whole system and provides load balancing between HAs. MIPv6-based reliability and load balancing schemes have been proposed for HAs in [23], where failure detection and recovery mechanisms are transparent to the mobile nodes.

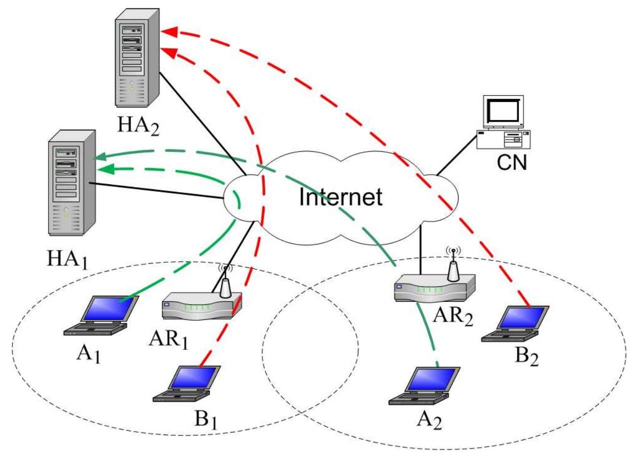

Fig. 5. Load balancing schemes.

In [24], another HA reliability and load balancing solution for Mobile IPv6 is proposed where multiple HAs are provided over different Home Links. This assumption is a modification to Mobile IPv6 protocol where all the HAs should exist on the same link. Mobile node registers with in the a number of HAs and chooses one of them as the primary HA. Primary HA or any other HA can tunnel packets from CN to MN. If the primary HA fails, MN can switch to any other HA on any home link.

2) Dynamic: It is possible that static assignment of mobile nodes may overload the queue of some mobility agents while others are almost idle. This will result in poor performance of the overall system. On the other hand, in dynamic load balancing approach [17], [20]–[22], mobile nodes are dynamically assigned among several HAs, thus making the scheme more robust. In order to distribute the load dynamically, some control packets (with state information) will have to be exchanged among the mobility agents.

Jue et al. [17] have designed a replicated server architecture where mobile nodes are divided among several Home agents facilitating dynamic load balancing. When any HA gets overloaded, it de-registers the binding of mobile nodes and sends registration messages to the least loaded home agent on the home link. Thus it performs load balancing by load shedding from overloaded HA, and adding that node to under-loaded HA.

Faizan et al. [21] introduces Virtual HA Reliability Protocol (VHARP) an extension to Mobile IPv6 that introduces reliability and load balancing. Deng et al. [22] proposed a hybrid load balance mechanism by deploying multiple HAs to share the traffic in the home network, taking into account the traffic information and the registration information at each HA to distribute traffic load among the HAs. HAs are dynamically assigned to the MNs; a HA is reallocated when there is a considerable traffic load difference between HAs.

C. Removal of FA

To comply with the requirement of IETF that the mobility solution would work without any changes to the majority of routers in the Internet, some researchers try to remove the Foreign Agent and proposed MN to perform the foreign agent function. To make this work, however, a Mobile Agent must be able to dynamically acquire an IP address that is located in the address space of the foreign network. This address will then be used as the care-of address. One way of acquiring such address may be using DHCP. Thus mobility can be achieved with only one agent and some new software on the mobile node, assuming DHCP is used on the foreign network.

Several research works [2], [12], [25] eliminates FA to improve survivability. In [2], [12], DHCP server is assumed to provide IP address to the MH and the decapsulation task of FA is done by the MH and Jain et al. [25] introduce an Advertisement Agent (AA) in each subnet. The function of AA is to issue broadcasts specifying the its subnet and to respond to agent solicitation messages, making it is very simple and any surviving host can take over as an AA. The conflict among two or more AAs while broadcasting can be resolved using exponential backoff.

D. Using Location Register

In Mobile IP, the Home Agent maintains the mapping of MN's IP address to Care-of-Address and Foreign Agent keeps the reverse mapping of that. So these data should not be kept in a vulnerable environment, rather they should be replicated and distributed.

Ravi et al. [25] propose MIP-LR (Mobile IP with Location Register) that places the mobility database outside a vulnerable Home network, replicating and distributing the database to improve survivability and robustness in case of attack. Before sending a packet to the mobile host, the sender first queries a database, called the Home Location Register (HLR), to obtain the recipient's current location. MIP-LR also provides improved performance over Mobile IP by avoiding triangular routing and encapsulation of data packets.

Multilayered mobility management involving MIP-LR are presented in [26], [27], [28] for improved survivability. In [29], Dutta et al. designed and prototyped an application layer solution based on Mobile-IP with Location Registers (MIP-LR) in a laboratory environment. Results demonstrate that one can attain up to 50% reduction in management overhead and 40% improvement on latency compared to standard Mobile IP in co-located mode, also providing survivability features in an ad hoc environment.

E. Hierarchical approach

The amount of signaling on the mobility agents can be reduced by introducing hierarchy in mobility management, thus facilitating scalable services as number of mobile nodes increases. Examples of such hierarchical approaches include Hierarchical Mobile IPv6 [3], Robust Hierarchical Mobile IPv6 (RH-MIPv6), [14], Fast Hierarchical Mobile IPv6 (FH-MIPv6), Optimal Multi-level HMIPv6 (OM-HMIPv6) [30].

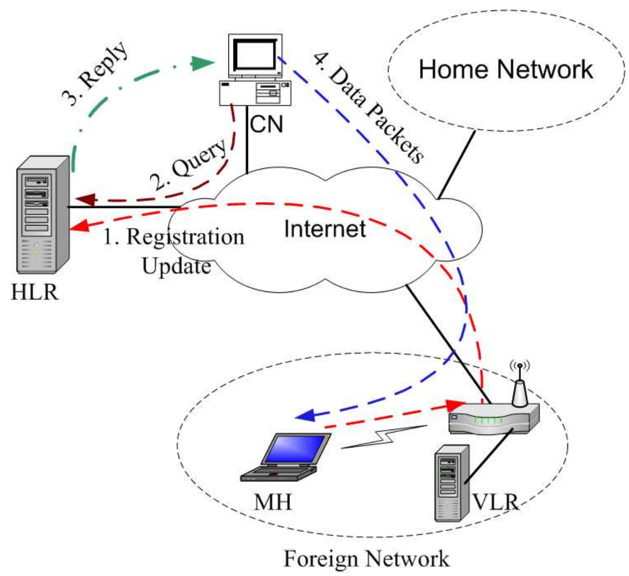

Fig. 6. Mobile IP with Location Register.

To reduce the amount of signaling to the correspondent nodes and to the HA, Hierarchical Mobile IPv6 (HMIPv6) [3] is designed that allows mobile nodes to locally register in a domain using Mobility Anchor Point (MAP). In a distributed MAP environment, multiple MAPs can exist on any level in a hierarchy including the access router. The distributed MAP environment has several advantages, like load balancing and scalable service. In HMIPv6 architecture, the mobile nodes do not send updates to the HA, if it is within the domain of the MAP, reducing the load on the HA. In addition, the failure in one MAP does not affect all the mobile nodes, having improved survivability.

The Robust Hierarchical Mobile IPv6 (RH-MIPv6) [14], an enhancement of HMIPv6, provides survivability and fault tolerance to the existing architecture. Unlike other fault-tolerant schemes, it does not require any synchronization between mobility agents (e.g., HA and MAP). When Router Advertisement messages are received from multiple MAPs, an MN configures two regional care-of addresses (RCoAs): one is primary RCoA (P-RCoA) and the other is secondary RCoA (S-RCoA). Then the MN registers two RCoAs to two MAPs (primary and secondary MAPs). In addition, the MN registers the HA and CNs. These multiple RCoAs configured in advance improves the robustness of the network and are dynamically changed in case of MAP failure. Thus, it is possible to reduce the failure recovery time compared with the HMIPv6.

Pack et al. [30] proposed optimal multi-level Hierarchical Mobile IPv6 that produces least cost. Thus the network survivability and scalability is improved. Instead of using single-level MAP-hierarchy, multi-level hierarchical structure of HMIPv6 [30], shown in Fig. 7, can be used to accommodate larger number of mobile nodes. However, as multi-level architecture has more packet delivery cost due to encapsulation /decapsulation of packets. Hence it is crucial to determine the

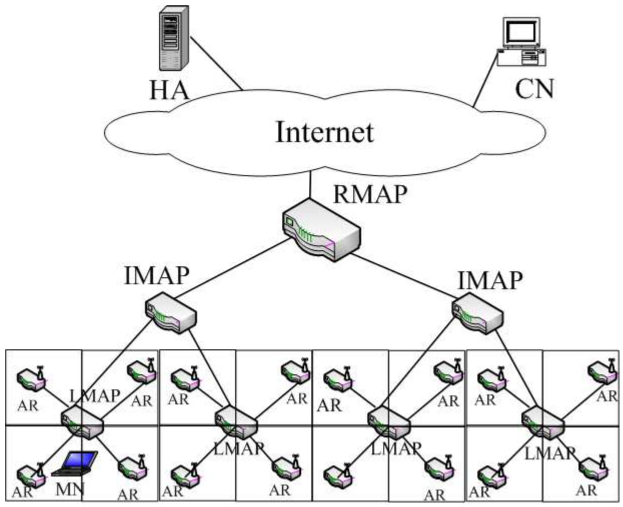

Fig. 7. A three-level HMIPv6 architecture of a foreign network.

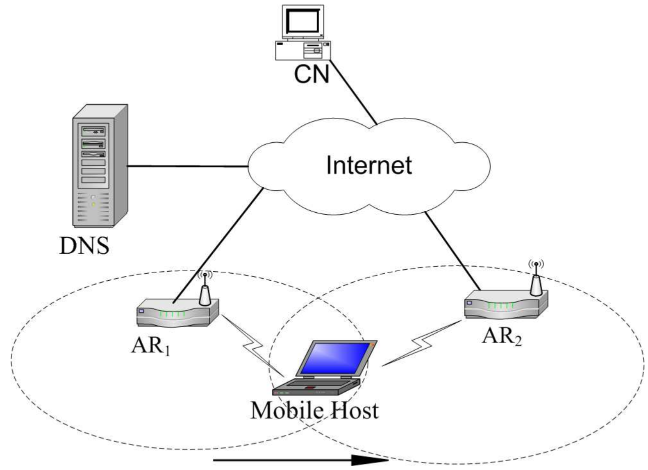

Fig. 8. DNS as the Location Manager.

optimal hierarchy that minimizes the total cost. OM-HMIPv6 architecture is dynamically configured after the computation of optimal hierarchy level.

F. Refresh-based Schemes

There may be cases when the mobility database in mobility agents (HA) are damaged, corrupted or lost, rather than the HA itself. Such problem can be fixed by sending periodical (refreshing) binding updates. Yuguang et al. [31] discuss the active recovery procedure of Home Location Regisgter (HLR) failure in the cellular network. They present an analytical model to compute the recovery time and the average number of lost calls due to a failure of mobility database. Jain et al. [25] also used refreshed based schemes to recover HLR from failure. Similar works in [32] and [33] worked on HLR recovery in PCS networks using checkpointing and failure restoration

G. DNS as mobility agent

The idea of having home network (where one or more home agents are situated) in some location on Earth may lead to survivability issues. If the complete home network is located in a hostile environment, such as a battlefield, the possibility of all HAs being destroyed is relatively high, leaving all the MHs belonging to the home network would be inaccessible. The distributed mobility agents can improve survivability of mobility protocols, as they can be secured from attack.

The mobility agents can be arranged on a DNS-like structure, or can be combined with a DNS server [34] for IP-diversity based mobility management techniques, such as SIGMA [12]. In fact, the location management scheme proposed in SIGMA has no concept of HA or FA. The Location Manager (LM) does not have to be located in a specific network to intercept data packets destined to a particular MH. It is thus possible to avoid physically locating the LM in a hostile environment; it can be located in a secure environment, making it highly available and survivable. Also the location management and data traffic forwarding functions are decoupled, allowing it to overcome many of the drawbacks of MIP in terms of scalability.

H. Quantitative approaches

There are several research works on survivability and scalability analysis that focus on defining certain metrics or functions for quantitative study of terrestrial and space networks. In [35], a methodology is proposed to measure the electronic and physical survivability of satellite communication networks as a function of the architecture, cost of implementation, cost of enemy to destroy the network, traffic rate or throughput, and mean delay of the message. This methodology can compare different types of Satellite networks on the basis of cost and technical capability.

He et al. [36] discuss the network survivability for satellite network based on Walker Constellation. A new survivability metric and survivability function for satellite network is introduced based on topology structure and traffic capacity, using which some quantities of interest are derived. Survivability performance evaluation on Iridium satellite network is presented as an example.

Heegaard et al. [37] have developed an analytical model to assess the survivability of a network with virtual connections exposed to link or node failures, also validated by simulations. The modeling approaches are applied to both small and real-sized network examples with three different scenarios: single link failure, hurricane disaster, and instabilities in a large block of the system.

V. Future Research: Issues and Challenges

Most of the works on survivability and scalability of mobility protocols are based on terrestrial networks. Commercial Industry and Government are increasingly using satellite networks to exchange data, voice and video for critical services, such disaster monitoring and preparedness, weather information, military communications, etc. Though there exists some similarity between space and terrestrial networks,

TABLE I
COMPARISON AMONG THE SCHEMES.

Schemes	Basic Principle	Advantages	Limitations
Redundancy-based	Multiple redundant MAs are used to avoid single point of failure	Less failure time	Synchronization required among the agents
Load-balancing	Loads on MAs are distributed to allow scalable services	Support large number of nodes with less delay and loss	Some control message with state information (such as queue size) need to be exchanged among MAs
FA removal	MN performs the function of FA	reduces changes in Internet infrastructure	More work in MN
Using LR	Location database are kept in distributed LR instead of HA / FA.	Eliminate tunneling and triangular routing, location database are distributed and secured in case of attack	Every CN may be configured with the address of HLR
Hierarchical	Amount of signaling to MAs are reduced by introducing multi-level hierarchy	Less location update cost	More processing due encapsulation and decapsulation in multiple agents
Refresh-based	Damaged mobility database may be restored by periodic refreshing updates	The MAs are not required to be replicated	Additional signaling required. Location database may be invalid for longer period if refreshing period is high
DNS-based	DNS can be used as location database instead of HA in hostile environment	Location database can be kept secured	DNS was not designed to modify too frequently
Quantitative approach	Use of several metrics to quantify survivability	In-depth understanding and comparison possible	Depends on assumptions, sometimes not realistic

the approaches discussed for improving the survivability and scalability of IP-based mobility protocols may not be directly applicable to space networks. In this section, we discuss some of the issues and challenges that need to be addressed in order to use the techniques to ensure survivability and scalability in space networks.

A. Removal of home network

The idea of having home network (where one or more home agents of the satellite constellations are situated) in some location on Earth may lead to survivability issues in case of military communication satellites, such as MILSTAR, Defense Satellite Communications System (DSCS), Ultra High Frequency Follow-On (UFO), etc. Troops, military vans can use IP-enabled devices in the battlefield which act as mobile nodes. They can communicate with the commanders in military bases through the use of military satellites. Even if the HA is replicated (statically or dynamically) to multiple agents for improved survivability, all these HAs have to be located in the home network for intercepting packets sent by the Mobile nodes (troops). If the complete home network is located in a hostile environment, the possibility of all HAs being destroyed is relatively high. In that case, all the MHs belonging to the home network would be inaccessible. In this respect, the IP-mobility protocols (such as, LR-based or DNS-based schemes) that are designed without any requirement of having the home network will survive such hostile environment.

B. Bandwidth limitation

The bandwidth in satellite networks is limited and is an important constraints unlike in terrestrial networks. Therefore, IP-mobility protocols should be designed in such a way that the satellite bandwidth is utilized efficiently, so that more nodes can be served by the space network. The design should focus on reducing signaling overhead in the inter-satellite links as well as in the link between satellite and ground stations.

C. Propagation delay

Another important factor is the long propagation delay in space network. Compared to terrestrial communications, all geostationary satellite communication experience high end-to-end delay due to the reason that the signal has to travel to an altitude of 35,768 km above the sea level to a GEO satellite and back to Earth. This kind of delay is unacceptable in real-time communication, such as VoIP, videoconferencing etc. These time critical sessions should be given higher priority over non-real time services, such as ftp to improve the quality of service of space networks. This can become an important issue when the number of people using satellite networks become large.

D. Remote diagnosis and repair

Communications can often be disrupted due to terrestrial interferences or interference by other spacecrafts. It is essential to remotely access, and re-configure the hybrid networks of satellite and terrestrial network equipments, such as routers, switches, transponders, VSAT modems, antenna controllers, GPS devices, to reduce the outage time minimal. This will improve the survivability of space networks and also reduce the cost of repair. This will also facilitate other routine maintenance tasks such as, upgrading operating system in various entities of space network with the latest security patches, a lot more easier.

E. Constellation design issues and use of IP-diversity

The constellation of satellites and installation of ground stations should be such that the number of handoffs between the ground stations is reduced, i.e., satellite remains connected to the same earth stations for longer period of time. In addition, use of IP-diversity mobility protocols (such as, SIGMA [12]) can facilitate seamless handover between ground stations, ensuring least packet loss and handoff latency.

VI. Conclusion

With the proliferation of wireless and mobile computing, increasingly larger number of mobile nodes will require satellite communications, specially in remote areas where there is no terrestrial coverage, and in critical situations, such as, in disaster areas, battlefields, etc. In this paper, we have presented a comprehensive survey of the approaches used for improved survivability and scalability in terrestrial networks after classifying them into various groups. The issues and challenges to apply those approaches to space networks are also illustrated.

References

[1] V. Devarapalli, R. Wakikawa, A. Petrescu, and P. Thubert, "NEtwork MObility (NEMO) basic support protocol," RFC 3963, Jan 2005.

[2] D. Johnson, C. E. Perkins, and J. Arkko, "Mobility support in IPv6," IETF RFC 3775, June 2004.

[3] H. Soliman, C. Castelluccia, K. E. Malki, and L. Bellier, "Hierarchical Mobile IPv6 mobility management (HMIPv6)," IETF RFC 5380, Oct 2008.

[4] M. Atiquzzaman and W. Ivancic, "SIGMA for Space Sensor Web Networks," in *ESTO AIST Sensor Web Technology Meeting*, Orlando, FL, April 2-3 2008.

[5] K. Bhasin and J. L. Hayden, "Space internet architectures and technologies for NASA enterprises," *Journal of Satellite Communications*, vol. 20, no. 5, Sept 2002.

[6] W. Ivancic, P. Paulsen, D. Stewart, D. Shell, and L. W. et al., "Secure, network-centric operations of a spacebased asset - an abridged report," in *Earth-Sun System Technology Conference*, College Park, MD, June 2005.

[7] J. Noles, K. Scott, M. Zukoski, and H. Weiss, "Next generation space internet: Prototype implementation," in *NASA Earth Science Technology Conference*, Pasadena, CA, June 2002.

[8] W. Ivancic, D. Stewart, T. Bell, P. Paulsen, and D. Shell, "Use of Mobile-IP priority home agents for aeronautics, space operations and military applications," in *IEEE Aerospace Conference*, Big Sky, MT, March 2004.

[9] "Omni: Operating missions as nodes on the internet."

[10] S. B. S. R. G. Minden, J. Evans and L. Searl, "Routing in space based internets," in *NASA Earth Science Technology Conference*, Pasadena, CA, June 2002.

[11] e. a. I.F. Akyildiz, "Mobility management in next-generation wireless systems," in *Proceedings of the IEEE*, vol. 87, no. 8, 1999, pp. 1347–1384.

[12] S. Fu and M. Atiquzzaman, "SIGMA: A Transport Layer Handover Protocol for Mobile Terrestrial and Space Networks," *e-Business and Telecommunication Networks, Springer*, pp. 41–52, 2006.

[13] G. V. Rajib Ghosh, "Fault-tolerant Mobile IP," Washington University, Technical Report WUCS-98-11, 1998.

[14] T. You, S. Pack, and Y. Choi, "Robust Hierarchical Mobile IPv6 (RH-MIPv6): an enhancement for survivability and fault-tolerance in Mobile IP systems," in *IEEE Vehicular Technology Conference*, vol. 3, 2003, pp. 2014–2018.

[15] Y. F. Huang and M. H. Chuang, "Fault tolerance for home agents in Mobile IP," *Computer Networks*, vol. 50, no. 18, pp. 3686–3700, Dec 2006.

[16] J.-W. Lin and J. Arul, "An efficient fault-tolerant approach for Mobile IP in wireless systems," *IEEE Transaction of Mobile Computing*, vol. 2, no. 3, 2003.

[17] J. Jue and D. Ghosal, "Design and analysis of replicated server architecture for supporting IP-host mobility," *ACM Mobile Computing and Comm Revue*, vol. 2, no. 3, pp. 16–23, 1998.

[18] C. C. Ltd, "Mobile IP Home Agent Redundancy," http://www.cisco.com/en/US/docs/ios/12_0t/12_0t2/feature/guide/HAredun.html.

[19] S. Knight and D. W. et al., "Virtual router redundancy protocol," IETF RFC 2338, April 1998.

[20] A. Vasilache, J. Li, and H. Kameda, "Load balancing policies for multiple home agents Mobile IP networks," in *Web Information Systems Engineering*, 2001.

[21] J. F. H. EL-Rewini and M. Khalil, "Introducing reliability and load balancing in Mobile IPv6-based networks," *Wireless Communication and Mobile Computing*, vol. 8, no. 4, 2008.

[22] H. Deng, X. Huang, K. Zhang, Z. Niu, and M. Ojima, "A hybrid load balance mechanism for distributed home agents in Mobile IPv6," *Personal, Indoor and Mobile Radio Communications*, pp. 2842–2846, 2003.

[23] R. Wakikawa, V. Devarapalli, and P. Thubert, "Inter home agnets protocol (HAHA)," IETF Draft, draft-wakikawa-mip6-nemohaha-00.txt (work in progress), Oct 2003.

[24] H. Deng, R. Zhang, X. Huang, and K. Zhang, "Load balance for distributed home agents in Mobile IPv6," draftdeng-mip6-ha-loadbalance-00.txt (work in progress), Nov 2003.

[25] R. Jain, T. Raleigh, D. Yang, L. Fung, C. C. Graff, M. Bereschinsky, and M. Patel, "Enhancing survivability of mobile internet access using Mobile IP with location registers," in *IEEE INFOCOM*, 1999.

[26] F.-M. Chen, S. Su, and J.-L. Chen, "Multilayered mobility management integrating SIP with MIP-LR," *Journal of Beijing University of Posts and Telecommunications*, vol. 27, no. 6, pp. 97–100, Dec 2004.

[27] A. Dutta, J. Burns, K. D. Wong, R. Jain, and K. Young, "Multilayered mobility management for survivable network," in *IEEE MILCOM*, Oct 2001.

[28] C. Yap, M. Kraner, N. Fikouras, and S. Cvetkovic, "Novel and enhanced mobile internet protocol for third generation cellular environments compared to MIP and MIP-LR," in *Int. Conference on 3G Mobile Communication Technologies*, London, UK, March 2000, pp. 143–147.

[29] A. Dutta, J. Burns, R. Jain, K. D. Wong1, and K. Young, "Implementation and performance evaluation of application layer MIP-LR," in *International Conference on Wireless Networks, Communications and Mobile Computing*, vol. 2, Maui, HI, June 13-16 2005, pp. 999–1004.

[30] S. Pack, Y. Choi, and M. Nam, "Design and analysis of optimal multi-level Hierarchical Mobile IPv6 networks," *Wireless Personal Communications*, vol. 36, no. 2, January 2006.

[31] Y. Fang, I. Chlamtac, and H.-B. Fei, "Analytical results for optimal choice of location update interval for mobility database failure restoration in PCS networks," *IEEE Transactions on Parallel and Distributed Systems*, 2000.

[32] Y.-B. Lin, "Per-user checkpointing for mobility database failure restoration," *IEEE Transactions on Mobile Computing*, vol. 4, no. 2, 2005.

[33] A. O. Abdelgwad, D. Ahmed, and A. Saleh, "A reliable - double edged strategy for HLR mobility database failure detection and recovery in PCS networks," in *International Conference on Networking and Media Convergence*, Cairo, Egypt, March 24-25 2009, pp. 9–15.

[34] A. S. Reaz, M. Atiquzzaman, and S. Fu., "Performance of DNS as location manager for wireless systems in IP networks," in *IEEE Global Telecommunications Conference (GLOBECOM)*, St. Louis, MI, 28 November - 2 December 2005, pp. 359–363.

[35] A. Glenn, "Methodology for quantitatively evaluating satellite communication network survivability," *IEEE Communication Magazine*, vol. 23, no. 6, June 1985.

[36] Y. He and H. Zhao, "Survivability performance evaluation for satellite communication network based on walker constellation," in *International Conference on Space Information Technology*, Wuhan, China, Nov 2007.

[37] P. E. Heegaard and K. S. Trivedi, "Network survivability modeling," *Computer Networks*, vol. 53, no. 8, June 2009.