Transport Layer Mobility Protocols: Issues, Challenges and a Comparative Study

Chowdhury Nawrin Ferdous, Md. Ahsan-Ullah Asif-Ur-Rahaman
Department of Computer Science and Engineering
Military Institute of Science and Technology
Dhaka, Bangladesh
Email: bipashachowdhury03@gmail.com

Md. Shohrab Hossain
Department of Computer Science and Engineering
Bangladesh University of Engineering and Technology
Dhaka, Bangladesh
Email: shohrab29@gmail.com

***Abstract*—Mobility of Internet hosts allows computing nodes to move between subnets while being connected to the remote nodes. To provide seamless connectivity to the roaming users, several mobility protocols have been developed at different layers. Some schemes have been proposed in the transport layer. However, no research has been performed to analyze different transport layer mobility schemes. In this paper, we have discussed a number of transport layer mobility protocols. We have identified different mobility related issues and have chosen a number of evaluation criteria to critically analyze existing transport layer mobility schemes. Our critical analysis can help reader understand the stengths and weaknesses of these mobility schemes.**

***Index Terms*—Mobility protocols, Transport layer protocols, handover, wireless networks, Mobile IP.**

I. Introduction

The Internet is based on five layer architecture: physical, data link, network, transport and application layers, with each layer having specific responsibilities. Since mobility [1] can be managed at different layers, so at which layer the mobility should be managed is a natural question. Several works have investigated the weakness and strengths of mobility management at the different layers.

Mobility can be handled at different layers of the protocol stack, but network and transport layer mobility being the most widely studied. Transport Layer Mobility allows mobile nodes to not only change their point of attachment to the Internet, but also to control which network interfaces are used for the different kinds of data leaving from and arriving at the mobile nodes. Transport layer mobility can overcome many of the limitations of network layer schemes.

The Internet was originally designed for static hosts connected through wired networks. Proliferation of wireless networks has given rise to an increasing demand for mobility of hosts, resulting in various mobility management schemes. Mobility management consists of two fundamental operations: Handoff and Location Management. Handoff occurs when a mobile device changes its point of attachment while still communicating with its peer. Handoff can be implemented by Location management that refers to the task of locating (finding the IP address) a Mobile Host (MH) in order to initiate and establish a connection by a node. A good location management scheme should provide a valid address to the MH, and be transparent to its peers.

There are many transport layer mobility protocols [2] with different criteria. While Mobile IP is a network layer scheme which makes mobility transparent to upper layers by increasing the burden and responsibility of the Internet infrastructure, transport layer schemes are based on an end-to-end approach to mobility that attempt to keep the Internet infrastructure unchanged by allowing the end hosts to take care of mobility. MSOCKS [3], SIGMA [4], RCP [5], Freeze-TCP [6], R^2CP [7], I-TCP [8], M-TCP [9], M-UDP [10], BARWAN, TCP-R, mSCTP [11] etc. are the different mobility protocols working in the transport layer. However, there is lack of works that tries to analyze the performance of different existing mobility shcemes proposed to work in transport layer. Such an analysis is important to judge the strength and weakness of the existing schemes, thereby choose a scheme based on its merits.

Handoff, connection migration, and location management are the main fundamentals of a complete mobility management scheme. To determine and compare the effectiveness of mobility schemes, many evaluation criteria can be used. In this paper, we have used handoff type, packet loss and delay, fault tolerance, requirement to change in the network infrastructure, mobility type, support for IP diversity, security, scalability, etc. to classify and evaluate different existing transport layer mobility schemes.

The *objective* of this work is to analyze different existing transport layer mobility protocols based on different evaluation criteria to find out which protocols perform better in specific scenario.

Our *contribution* in this paper is to critically analyze different transport layer mobilty protocols: MSOCKS, SIGMA, RCP, R^2CP, Freeze TCP and Migrate TCP and to critically analyze their performance based on some basic evaluation criteria.

The rest of the paper is organized as follows. In Section II, the existing transport layer mobility protocols are explained along with their protocol operation. Section III presents the issues of mobility management. Evaluation criteria are listed in Section IV and comparison among the schemes are presented in Section V. Finally, we conclude the paper in Section VI.

II. Mobility Protocols

In this section, we explain six major mobility protocols working in the transport layer, namely, MOCKS, SIGMA, Freeze TCP, Migrate TCP, RCP and R^{2}CP as follows:

A. MSOCKS

MSOCKS [3] is built around a proxy that is inserted into the communication path between a mobile node and its correspondent hosts. For each data stream, from a mobile node to a correspondent host, the proxy is able to maintain one stable data stream to the correspondent host, thereby isolating the correspondent host from any mobility issues. Meanwhile the proxy can simultaneously make and break connections to the mobile node as needed to migrate data streams between network interfaces or subnets.

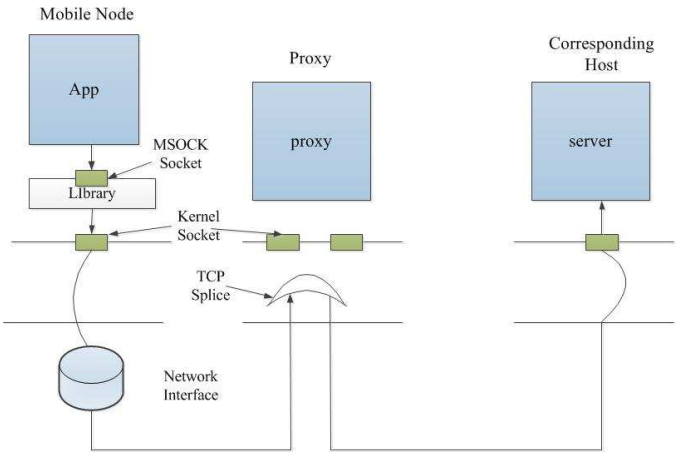

Fig. 1. Architecture of MSOCKS

The proxy can then mediate the communication between server and client, and provide services on behalf of either. MSOCKS, is built around a technique we call TCP Splice. TCP Splice allows the machine where two independent TCP connections terminate to splice the two connections together, effectively forming a single end-to-end TCP connection between the endpoints of the two original connections. MSOCKS architecture consists of three pieces: a user level MSOCKS proxy process running on a proxy machine; an in-kernel modication on the proxy machine to provide the TCP Splice service; and a MSOCKS library that runs under the application of the mobile node.

B. SIGMA

SIGMA [4] stands for Seamless IP-diversity based Generalized Mobility Architecture. The handover preparation procedure begins when MH moves into the overlapping radio coverage area of two adjacent subnets. Once the MH receives the router advertisement from the new access router (AR2), it begins to obtain a new IP address. After the MH obtained the IP address, MH notifies CN about the availability of the new IP address. When MH moves further into the coverage area of wireless access network2, CN can redirect data traffic to

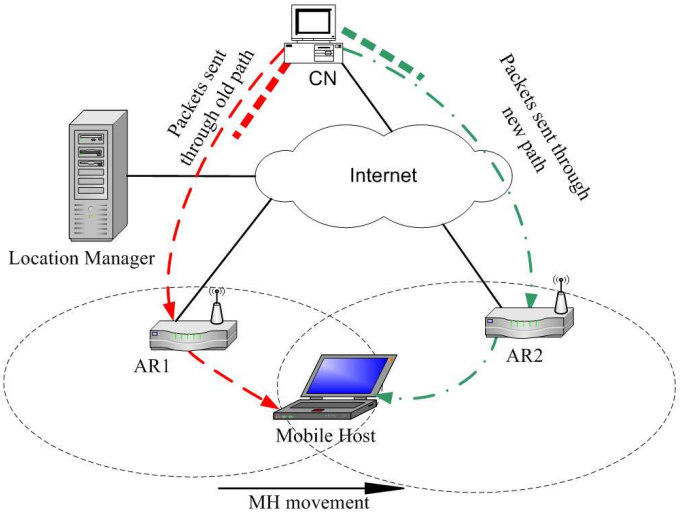

Fig. 2. Architecture of SIGMA

new IP address IP2 to increase the possibility that data can be delivered successfully to the MH. SIGMA supports location management by employing a location manager which maintains a database recording the correspondence between MHs identity and MHs current primary IP address. MH can use any unique information as its identity such as home address like MIP, or domain name, or a public key defined in Public Key Infrastructure(PKI). We can observe an important difference between SIGMA and MIP: the location management and data traffic forwarding functions are coupled together in MIP, while in SIGMA they are decoupled to speed up the the handover proces. When MH moves out of the coverage of wireless access network1, no new or retransmitted data should be directed to address IP1.

In SIGMA, MH notifies CN that IP1 is out of service for data transmission by sending an ASCONF chunk to CN to delete IP1 from CNs available destination IP list. A less aggressive way to prevent CN from sending data to IP1 is MH advertising a zero receiver window (corresponding to IP1) to CN. By deactivating, instead of deleting, the IP address,SIGMA can adapt more gracefully to MHs zigzag movement patterns and reuse the previously obtained IP address (IP1) as long as the IP1s lifetime is not expired. This will reduce the latency and signalling trafc caused by obtaining a new IP address.

C. Freeze TCP

Freeze-TCP [6] mechanism which is a true end-to-end scheme and does not require the involvement of any intermediaries (such as base stations) for flow control. When a mobile node certainly monitored the signal strength is fading, In such case, it can advertise a zero window size, to force the sender into the ZWP mode and prevent it from dropping its congestion window.

If the receiver can sense an impending disconnection, it should try to send out a few (at least one) acknowledgements,

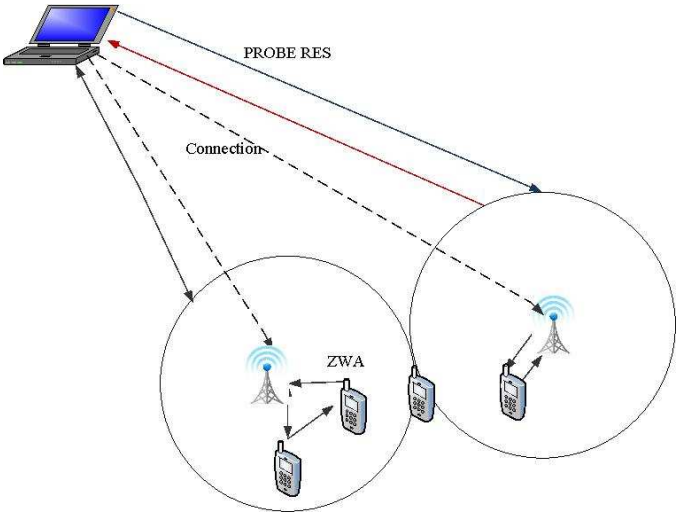

Fig. 3. Architecture of FREEZE TCP

wherein its window size is advertised as zero. Freeze-TCP lets the MH 'freeze'or stops an existing TCP connection during handoff by advertising a zero window size to the CN, and unfreezes the connection after handoff. This scheme reduces packet losses during handoff at the cost of higher delay.

D. Migrate TCP

Migrate TCP (M-TCP) [9] is a transparent mobility management scheme which is based on connection migration and uses DNS for location management. In Migrate TCP, when an MH initiates a connection with a CN, the end nodes exchange a token to identify the particular connection. A hard handoff takes place when the MH reestablishes a previously established connection using the token, followed by migration of the connection. Similar to SIGMA, this scheme proposes

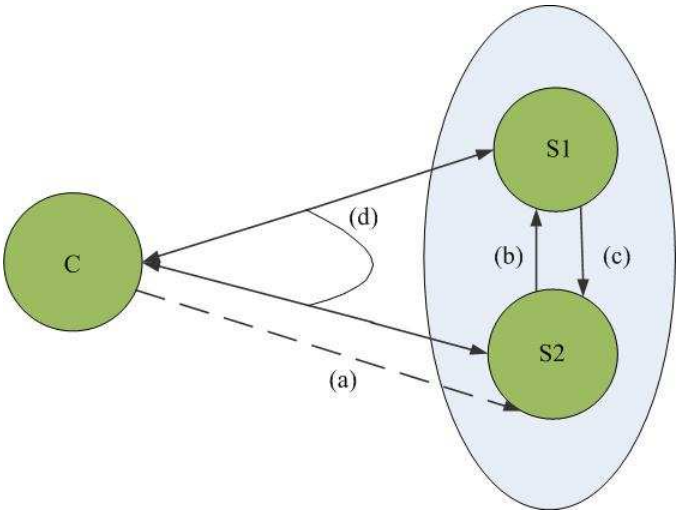

Fig. 4. Architecture of MIGRATE TCP

to use DNS for location management. The main achievement of M-TCP is service continuity. The basic application of M-TCP is for long lived connections like multimedia streaming services or videos, end users expect both correctness and good response time like, Internet banking, e-commerce etc. S2 sends a request (b) to S1 and receives the state (c). If the migrating endpoint is reinstated successfully at S2, then C and S2 complete the handshake, which ends the migration (d). Upon accepting the migrated connection, the server application at S2 imports the state snapshot. It then resumes service using the snapshot as a restart point, and performs execution replay for a logbased recovery supported by the protocol. The execution replay restores the state of the service at the new server and synchronizes it with the protocol state. To support the replay, M-TCP logs and transfers from S1 data received and acknowledged since the last snapshot. It also transfers unacknowledged data sent before the last snapshot, for retransmission from S2.

E. RCP

RCP (Reception Control Protocol) [5] is a receiver centric protocol that moves the responsibility from the sender to the receiver for performing reliability and congestion control. RCP is a TCP clone in its general behavior, but allows for better congestion control, loss recovery, and power management mechanisms compared to sender-centric approaches. In RCP, since the control of data transfer is shifted from the sender to the receiver, the data acknowledgement style of handshaking in TCP is no longer applicable. Instead RCP uses the request data handshake for data transfer, It has advantages over a sender-

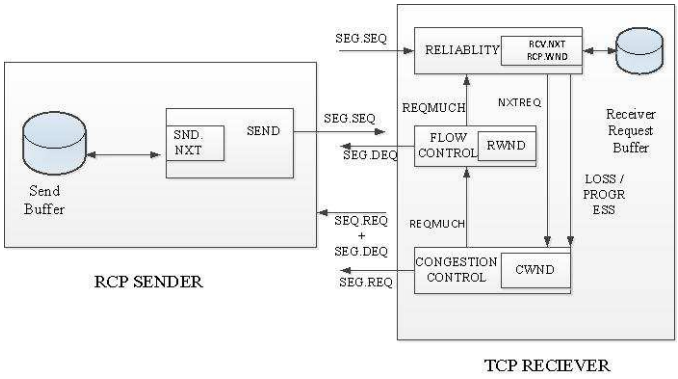

Fig. 5. Architecture of RCP

centric one since the receiver can accurately control which and how much data to send through each pipe based on the status (say, signal strength) of each interface. Moreover, when the receiver decides to switch to another interface specific congestion control mechanism after handoffs, such decision does not need to involve the sender. While the receiver in TCP merely sends back ACKs with no control over which and in what sequence data is transmitted by the sender, in RCP the receiver explicitly controls these factors and the reliable delivery of data. Moreover, the RCP receiver also assumes total control over the bandwidth the connection can consume, using the same window based algorithm employed by the TCP sender. Finally, although ow control in TCP involves

the sender, it is performed solely by the receiver in RCP. Therefore, the receiver in RCP determines how much data the sender can send (via congestion control and ow control), and which data the sender should send (via reliability).

F. R^2CP

R^2CP (Radial Reception Control Protocol) [7] is based on Reception Control Protocol, a TCP clone. At the time of the event of handoff from one access network to another the mobile host is initially connected to Server-I through network A and hence one RCP pipe (RCP-1) is created in the R^2CP connection. A little later the mobile host decides to handoff to network B, so a second RCP pipe (RCP-2) is created (using the new network address). However, RCP-1 is not closed until some more time, and hence during this time two pipes co-exist in the connection to collaboratively deliver data for the application. When the mobile host moves to network B, it has access to a replicated server (Server-II).

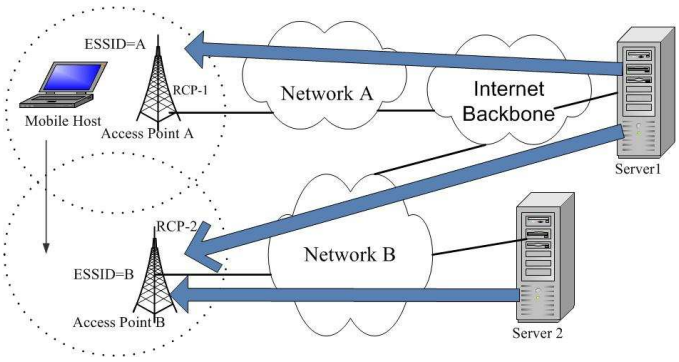

Fig. 6. Architecture of R^2CP

Because of shorter round-trip time and a larger bandwidth the mobile host decides to perform server migration from Server-I to Server-II. Initially, the R^2CP connection creates an RCP pipe (RCP-1) using network address A and the address of Server-I. When the mobile host moves to network B, R^2CP creates a new RCP pipe (RCP-3) using network address B and the address of Server-II. At first mobile host does not perform server migration, and hence the second RCP pipe created (RCP-2) is between network address B and the address of Server-I.

III. Issues of Mobility Management

In mobility management, change in the IP address of a MH gives rise to the challenges of maintaining an uninterrupted data flow. Morevover, minimizing loss of packets, maintaining security, identication of the newer location, etc. are also crucial issues. Some of these issues are discussed in the following subsections:

A. Connection Migration

An MH acquires a new IP address when it changes its subnet. Since the old IP address is retained, then the problem is how the CN will continue communicating with the MH which now has multiple IP addresses. Connection Migration notifies the CN about this change and migrate the connection from the old to the new address. To avoid data ow through the old address of MH, connection migration may result in a temporary stop in the data ow during the migration process. A gateway in the middle of the connection or proxy may be used to handle the connection switching.

B. Packet Loss and Latency

When an MH acquires a new IP address, unless theMH and the underlying protocols support multiple addresses, the MH can only be contacted via the new address. Packets destined to the MH via the old address cannot reach the destination, resulting in packet loss, latency and wastage of internet bandwidth. Mobility schemes must come up with techniques to mitigate packet losses and latency during handoffs.

C. Infrastructure Requirement

The Internet was no initially designed with mobility in mind. Consequently many of the proposed schemes require changes in the existing Internet infrastructure, such as gateway or proxy in the middle of the connection, to support mobility.

D. Location Management

Following the change of IP address of an MH, a CN should be able to locate the MH. A location manager keeps track of the current IP address of an MH, and provides the current address to any entity trying to initiate communication with the MH.

IV. Evaluation Criteria

In this section, we discuss a set of evaluation criteria [12] which we use to compare various mobility schemes.

A. Handoff Process

The performance of a mobility management scheme depends on the type of handoff which can be either soft or hard. Soft handoff (also called seamless handoff) permits a smooth handoff by allowing a mobile to communicate and exchange data with multiple interfaces simultaneously during handoff. Communication through the old interface is dropped when the signal strength from the corresponding access point drops below a certain threshold. On the contrary, hard handoff results in disconnecting from the old access point when the signal strength is below a threshold before connecting to the new access point.

B. Scalability and Fault Tolerance

Scalability refers to the ability of a mobility management scheme to handle a large number of MHs and CNs. A scheme is scalable when its performance does not drop with an increase in the size of the network size or the number of MHs and CNs. A system is said to be fault tolerant when it can function in the presence of system failures. For example, a scheme with a single point of failure in said to be faulting intolerant.

TABLE I
COMPARISON AMONG THE PROTOCOLS BASED ON DIFFERENT CRITERIA

Criteria	MSOCK	SIGMA	FREEZE TCP	MIGRATE TCP	R^2CP	RCP
Handoff	Hard	Soft	N/A	Hard	Soft	Soft
Loss/Delay	Only the fly packets are lost	No	Avoids data transfer during hand-off to prevent loss	No, but stops transmission If MH is the server	No	No
Fault tolerance	Single point of failure: proxy	New connections would fail if location manager fails	Yes	would fail if location manager fails	Yes	Yes
Change in infrastructure	Yes	No	No	No	No	Yes
Transparency	Yes	Yes	Yes	Yes	Yes	Yes
Conflicts with security solution	Yes	No	No	No	No	No
IP Diversity	No	Yes	No	No	Yes	No
Change in protocol stack	Yes	Yes	No in CN, Yes in MH	Yes	Yes	Yes

C. Application Transparency

A mobility scheme is transparent to an application when the application does not need to know about handoff taking place in the lower layers, and hence does not require any modication to the application.

D. Loss/Delay

Packets in in flight may not be delivered to the MH during the handoff period. This may result in packet losses, packet delay, and a false indication of congestion in the network.

E. Security Solution

Internet is vulnerable to many security threats. Many of the solutions, such as ingress ltering and rewalls, to the threats do not allow network entities to process packet headers as may be required by some of the mobility schemes.

F. Path Diversity/IP Diversity

Increasing number of mobile devices now a days comes with multiple communication interfaces. During handoff, an MH may be able to take advantage of multiple IP addresses (called IP diversity), obtained from separate subnets, associated with the multiple interfaces.

G. Change in Infrastructure

A mobility management scheme may require additional software agents (such as Home / Foreign agents in the case of MIP) or hardware to be deployed in the existing network infrastructure. Such additional agents / hardware may result in scalability and deployment issues for the scheme to be implemented in the real world.

H. Change in Protocol

A transport layer mobility management scheme may require change in the transport protocol, or may require applications to use a new transport protocol or API.

V. COMPARATIVE STUDY

We have discussed different evaluation criteria in the previous Section IV. Now we use those evaluation criteria to classify the six transport layer mobility protocols. Table I presents such a comparison.

In MSOCK, TCP Splice splits a TCP connection at a proxy by dividing the host-to-host communication into host-proxy and proxy-host communications. MSOCKS uses TCP Splice for connection migration. During handoff, it obtains a new IP address from the new subnet, and establishes a new connection with the proxy using its second interface. The handoff proceess is hard. The communication between proxy and CN, however, remains unchanged. The data flow between MH and CN thus continues, with the CN being unaware of the mobility. Location management is done through the proxy who is always aware of the location of the MH; this limits the mobility within the coverage of the proxy. Only the flying packets are lost here. But a single point of failure, if the proxy fails then the whole system breaks. The distadvantage of this protocol is that it needs to change the infrastructure of the existing network and as well as the protocol stack.

SIGMA is a complete mobility management scheme implemented at the transport layer, and can be used with any transport protocol that supports IP diversity. SIGMA supports IP diversity-based soft handoff. As an MH moves into the overlapping region of two neighboring subnets, it obtains a new IP address from the new subnet while still having the old one as its primary address. When the received signal at the MH from the old subnet goes below a certain threshold, the MH changes its primary address to the new one. When it leaves the overlapping area, it releases the old address and continues communicating with the new address thus achieving a smooth handoff across subnets. Location management in SIGMA is done using DNS as almost every Internet connection starts with a name lookup. Whenever an MH changes its address, the DNS entry is updated so that subsequent requests can be served with the new IP address. The handoff it supports is soft.

There is less delay/loss of packets than the other protocols. New connection will fail if the location manager fails. Here, it is not needed to change the infrastucture but need to change in the protocol stack.

Migrate TCP is a transparent mobility management scheme which is based on connection migration using Migrate TCP, and uses DNS for location management. In Migrate TCP, when an MH initiates a connection with a CN, the end nodes exchange a token to identify the particular connection. A hard handoff takes place when the MH reestablishes a previously established connection using the token, followed by migration of the connection. Similar to SIGMA , this scheme proposes to use DNS for location management. The handoff it supports is soft. It avoids data transfer during handoff so that no packet is loss. Here, it is not needed to change the infrastructure. It needs to change the protocol stack in CN but not in MH.

Freeze-TCP is a connection migration scheme that lets the MH 'freeze'or stop an existing TCP connection during handoff by advertising a zero window size to the CN, and unfreezes the connection after handoff. This scheme reduces packet losses during handoff at the cost of higher delay. Although it provides transparency to applications, Freeze TCP requires changes to the transport layer at the end nodes. Freeze-TCP only deals with connection migration, but does not consider handoff or location management. It can be employed with some other schemes like Migrate to implement a complete mobility management scheme. It supports hard hand-off. New connections would fail if location manager fails. Here, it is not needed to change the infrastucture but need to change in the protocol stack.

RCP moves the responsibility for performing reliability and congestion control from the sender to the receiver. It allows for better congestion control, loss recovery, and power management mechanisms compared to sender-centric approaches. The handoff is soft. It does not conflict with the security solution. It supports IP diversity. Here, it needs to change in the infrastructure. With further improvement, R^2CP is proposed.

R^2CP is based on Reception Control Protocol (RCP), a TCP clone in its general behavior but moves the congestion control and reliability issues from sender to receiver on the assumption that the MH is the receiver and should be responsible for the network parameters. R^2CP has some added features over RCP like the support of accessing heterogeneous wireless connections and IP diversity that enables a soft handoff and bandwidth aggregation using multiple interfaces. A location management scheme might be integrated with R^2CP to deploy a complete scheme. The handoff is soft. It does not conflict with the security solution. It supports IP diversity. Here, it does not need to change in the infrastucture.

Table I presents a comparison among the schemes, showing the strengths and weaknesses of different existing transport layer mobility schemes.

VI. Conclusion

In this paper, we have discussed different transport layer mobility protocols and compared their performances based on different evaluation criteria, such as, Handoff type, loss / delay, fault tolerance, change in infrstructure, conflict with security solutions, IP diversity, change in protocol stack, etc. Our work can help in understanding the difference among different transport layer mobility schemes. In addition, our critical analysis can help reader understand the stengths and weaknesses of these mobility schemes.

References

[1] C. Perkins, D. Johnson, and J. Arkko, "Mobility support in IPv6," IETF RFC 6275, Jul 2011.

[2] G. Bolch, S. Greiner, H. de Meer, and K. S. Trivedi, *Computer networking*. Pearson Education, 2012.

[3] Maltz, David, and P. Bhagwat, "MSOCKS: An architecture for transport layer mobility," *Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings*, vol. 3, 1998.

[4] S. Fu, M. Atiquzzaman, and Y. J. Lee, "Architecture and performance of SIGMA: A seamless mobility architecture for data networks," *Journal of High Speed Networks*, vol. 5, 2005.

[5] Kevin and S. Singh, "A receiver-centric transport protocol for mobile hosts with heterogeneous wireless interfaces," *Wireless Networks*, vol. 11, no. 4, pp. 363–382, 2005.

[6] Goff, Tom, J. Moronski, D. S. Phatak, and V. Gupta, "Freeze-TCP: A true end-to-end tcp enhancement mechanism for mobile environments," *Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 3, pp. 1537–1545, 2000.

[7] Kim, Kyu-Han, Y. Zhu, R. Sivakumar, and H.-Y. Hsieh, "A receiver-centric transport protocol for mobile hosts with heterogeneous wireless interfaces," *Wireless Networks*, vol. 11, no. 4, pp. 363–382, 2005.

[8] Goff, J. Moronski, D. S. Phatak, and V. Gupta, "I-TCP: Indirect TCP for mobile hosts," *Distributed Computing Systems, Proceedings of the 15th International Conference*, 1995.

[9] Brown, Kevin, and S. Singh, "M-TCP: TCP for mobile cellular networks," *ACM SIGCOMM Computer Communication Review*, vol. 27, no. 5, pp. 19–43, 1997.

[10] Brown and S. Singh, "M-UDP: UDP for mobile cellular networks," *ACM SIGCOMM Computer Communication Review*, vol. 11, no. 5, pp. 60–78, 1996.

[11] Phan, Thomas, K. Xu, R. Guy, and R. Bagrodia, "Handoff of application sessions across time and space," *IEEE International Conference*, vol. 5, 2001.

[12] M. Atiquzzaman and A. S. Reaz, "Survey and classification of transport layer mobility management schemes," *Personal, Indoor and Mobile Radio Communications*, vol. 4, 2005.