A REAL-TIME BI-DIRECTIONAL DIFFERENTIAL GLOBAL POSITIONING
SYSTEM DATA LINK OVER INTERNET PROTOCOL

A thesis presented to
the faculty of
the Fritz J. and Dolores H. Russ College of Engineering and Technology
of Ohio University

In partial fulfillment
of the requirements for the degree
Master of Science

Sumit Bhattacharya
June 2005
This thesis entitled
A REAL-TIME BI-DIRECTIONAL DIFFERENTIAL GLOBAL POSITIONING
SYSTEM DATA LINK OVER INTERNET PROTOCOL

by
SUMIT BHATTACHARYA

has been approved for
the School of Electrical Engineering and Computer Science
and the Russ College of Engineering and Technology by

Chris G. Bartone
Associate Professor of Electrical Engineering and Computer Science

Dennis Irwin
Dean, Russ College of Engineering and Technology
This research focused on the development of a prototype real-time bi-directional Differential Global Positioning System (DGPS) wherein data was passed over a radio frequency (RF) link using Internet Protocol (IP). This development is unique in that it combines; a typical DGPS architecture, where DGPS correction data are sent from a ground reference station to a mobile user; remote measurements functions typically found in remote positioning/monitoring applications; and using an IP via an ethernet interface to apply at a wide variety of RF data link frequencies.

This project demonstrated a proof of concept for an enhanced data link that would use the ethernet interface to communicate with a wireless transceiver operating at 2.4 GHz, thus providing for high data rates and also supporting additional data e.g., weather imagery, voice over IP, etc. The proof of concept was successfully demonstrated in a laboratory in Stocker Center, Ohio University. The simplified set up of the data link allowed for communication between a reference station and a user over IP via the ethernet interface. The data link supports high accuracy position solution and remote positioning of the user by the ground reference station.

Approved:

Chris G. Bartone
Associate Professor of Electrical Engineering and Computer Science
ACKNOWLEDGEMENTS

As with any serious endeavor, this research involved the effort, guidance and cooperation of many people and to whom I am grateful.

I would like to thank Dr. Chris Bartone for his constant support and guidance throughout the course of this effort that kept me on track in the right direction. Working with him has been a valuable learning experience. I would like to express my gratitude to Dr. Maarten uijt de Haag for his valuable help in fixing some initial problems and for his reviews of this document as a member of my Thesis committee. I am also grateful to Dr. David Matolak and Dr. Eric Steinberg for their reviews this document and helpful comments, serving as member and college representative of my Thesis committee.

I am thankful to the School of Electrical Engineering and Computer Sciences (EECS), for partly supporting this research by way of the Stocker Research Assistantship. I am grateful to Dr. Roger Radcliff, Graduate Chair of the School of EECS, Ohio University, for his extensive co-operation and for some very interesting courses he taught.

I would also like to express my gratitude to Jeff Dickman for his indispensable help in getting familiar with QNX and its intricacies and for some very helpful advice in resolving some hardware and software issues. I am thankful to Sisir Vuyyuru for answering my incessant naïve questions about TCP/IP.

I would like to say thank you to all my friends for their support and comments and for some great memories over the past couple of years.

Last but not the least; I am grateful to my parents, my sister and brother-in-law, for their blessings and unconditional support throughout, and especially through some uncertain times.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................................................................................... 4  
TABLE OF CONTENTS ........................................................................................................... 5  
LIST OF TABLES ..................................................................................................................... 8  
LIST OF FIGURES .................................................................................................................... 9  
LIST OF ACRONYMS .............................................................................................................. 10  
1 INTRODUCTION .................................................................................................................. 14  
   1.1 Thesis outline ................................................................................................................. 15  
2 BACKGROUND .................................................................................................................... 17  
    2.1 Global Positioning System .............................................................................................. 17  
       2.1.1 System Description ............................................................................................... 17  
       2.1.2 Principle of Operation ......................................................................................... 23  
       2.1.3 Error Sources ....................................................................................................... 25  
    2.2 Differential GPS .......................................................................................................... 28  
    2.3 GPS Augmentation Systems .......................................................................................... 33  
       2.3.1 Nationwide Differential Global Positioning System .............................................. 33  
       2.3.2 Local Area Augmentation System ......................................................................... 35  
       2.3.3 Wide Area Augmentation System ......................................................................... 39  
    2.4 Remote Positioning ....................................................................................................... 41  
       2.4.1 Automatic Dependent Surveillance – Broadcast (ADS-B) ................................... 42  
    2.5 Bi-directional DGPS ..................................................................................................... 44  
3 TRANSMISSION CONTROL PROTOCOL/INTERNET PROTOCOL ......................... 47  
   3.1 The Open System Interconnection (OSI) Model and TCP/IP ....................................... 47  
   3.2 The Internet Protocol .................................................................................................... 50
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2.1</td>
<td>IP Packet Format</td>
<td>50</td>
</tr>
<tr>
<td>3.2.2</td>
<td>IP Addressing</td>
<td>53</td>
</tr>
<tr>
<td>3.2.3</td>
<td>IP Address Format</td>
<td>53</td>
</tr>
<tr>
<td>3.2.4</td>
<td>IP Address Classes</td>
<td>54</td>
</tr>
<tr>
<td>3.2.5</td>
<td>IP Subnet Addressing</td>
<td>56</td>
</tr>
<tr>
<td>3.2.6</td>
<td>Address Resolution Protocol (ARP) Overview</td>
<td>56</td>
</tr>
<tr>
<td>3.2.7</td>
<td>Internet Routing</td>
<td>57</td>
</tr>
<tr>
<td>3.2.8</td>
<td>IP Routing</td>
<td>58</td>
</tr>
<tr>
<td>3.3</td>
<td>Transmission Control Protocol</td>
<td>59</td>
</tr>
<tr>
<td>3.3.1</td>
<td>TCP Connection Establishment</td>
<td>60</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Positive Acknowledgment and Retransmission (PAR)</td>
<td>61</td>
</tr>
<tr>
<td>3.3.3</td>
<td>TCP Sliding Window</td>
<td>62</td>
</tr>
<tr>
<td>3.3.4</td>
<td>TCP Packet Format</td>
<td>63</td>
</tr>
<tr>
<td>3.4</td>
<td>User Datagram Protocol</td>
<td>64</td>
</tr>
<tr>
<td>3.5</td>
<td>TCP and UDP: a comparison</td>
<td>67</td>
</tr>
<tr>
<td>3.6</td>
<td>The Socket Interface</td>
<td>68</td>
</tr>
<tr>
<td>4</td>
<td>SYSTEM ARCHITECTURE AND DESCRIPTION</td>
<td>70</td>
</tr>
<tr>
<td>4.1</td>
<td>An Overview</td>
<td>70</td>
</tr>
<tr>
<td>4.2</td>
<td>Ground-Based Station</td>
<td>70</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Data Uplink Message Format</td>
<td>72</td>
</tr>
<tr>
<td>4.3</td>
<td>Mobile User</td>
<td>73</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Data downlink message format</td>
<td>75</td>
</tr>
<tr>
<td>4.4</td>
<td>Hardware Set Up</td>
<td>76</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>The Seven Layers in the OSI Model</td>
<td>48</td>
</tr>
<tr>
<td>3.2</td>
<td>Composition of the IP Packet</td>
<td>51</td>
</tr>
<tr>
<td>3.3</td>
<td>Reference information about the five IP classes</td>
<td>55</td>
</tr>
<tr>
<td>3.4</td>
<td>The Format of a TCP Packet</td>
<td>63</td>
</tr>
<tr>
<td>3.5</td>
<td>The format of a UDP packet</td>
<td>65</td>
</tr>
<tr>
<td>3.6</td>
<td>Comparison of TCP and UDP</td>
<td>67</td>
</tr>
<tr>
<td>4.1</td>
<td>Data Parameters Uplinked</td>
<td>73</td>
</tr>
<tr>
<td>4.2</td>
<td>Data Parameters Downlinked</td>
<td>75</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 2.1  An illustration of the GPS constellation .............................................. 17
Figure 2.2  GPS System Overview........................................................................... 18
Figure 2.3  Principle of navigation using GPS......................................................... 24
Figure 2.4  A conceptual illustration of DGPS .......................................................... 29
Figure 2.5  LAAS System Overview......................................................................... 36
Figure 2.6  WAAS Architecture............................................................................... 40
Figure 2.7  A Bi-directional DGPS illustration......................................................... 45
Figure 3.1  An IP address consists of 32 bits, grouped into four octets.................... 54
Figure 4.1  Bi-directional DGPS Ground Reference Station Block Diagram.......... 71
Figure 4.2  Bi-directional DGPS Mobile User Block Diagram .............................. 74
Figure 4.3  Block Diagram of Hardware Set Up...................................................... 77
Figure 5.1  Remote Positioning data at ground reference ...................................... 89
Figure 5.2  East coordinate of User Position............................................................ 90
Figure 5.3  North coordinate of user position ......................................................... 91
Figure 5.4  Vertical coordinate of user position...................................................... 92
## LIST OF ACRONYMS

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADS-B</td>
<td>Automatic Dependent Surveillance - Broadcast</td>
</tr>
<tr>
<td>As</td>
<td>Anti-Spoofing</td>
</tr>
<tr>
<td>APL</td>
<td>Airport Pseudolite</td>
</tr>
<tr>
<td>APV-II</td>
<td>Approach operations with vertical guidance</td>
</tr>
<tr>
<td>ARP</td>
<td>Address Resolution Protocol</td>
</tr>
<tr>
<td>ATC</td>
<td>Air Traffic Control</td>
</tr>
<tr>
<td>ATM</td>
<td>Air Traffic Management</td>
</tr>
<tr>
<td>ATS</td>
<td>Air Traffic Service</td>
</tr>
<tr>
<td>BGP</td>
<td>Border Gateway Protocol</td>
</tr>
<tr>
<td>bps</td>
<td>bits per second</td>
</tr>
<tr>
<td>BPSK</td>
<td>Bi-Phase Shift Keying</td>
</tr>
<tr>
<td>BSD</td>
<td>Berkeley Software Distribution</td>
</tr>
<tr>
<td>CDTI</td>
<td>Cockpit Display Traffic Information</td>
</tr>
<tr>
<td>CNS</td>
<td>Communication, Navigation and Surveillance</td>
</tr>
<tr>
<td>CONUS</td>
<td>Continental United States</td>
</tr>
<tr>
<td>DGPS</td>
<td>Differential Global Positioning System</td>
</tr>
<tr>
<td>DOD</td>
<td>Department of Defense</td>
</tr>
<tr>
<td>DOT</td>
<td>Department of Transportation</td>
</tr>
<tr>
<td>DSSS</td>
<td>Direct Sequence Spread Spectrum</td>
</tr>
<tr>
<td>FAA</td>
<td>Federal aviation Administration</td>
</tr>
<tr>
<td>FHA</td>
<td>Federal Highway Administration</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>FRA</td>
<td>Federal Railroad administration</td>
</tr>
<tr>
<td>GEO</td>
<td>Geostationary Orbit</td>
</tr>
<tr>
<td>GLS</td>
<td>GNSS Landing System</td>
</tr>
<tr>
<td>GNSS</td>
<td>Global Navigation Satellite System</td>
</tr>
<tr>
<td>GPS</td>
<td>Global Positioning System</td>
</tr>
<tr>
<td>HANDGPS</td>
<td>High Accuracy nationwide Differential Global Positioning System</td>
</tr>
<tr>
<td>I/O</td>
<td>Input/Output</td>
</tr>
<tr>
<td>ICMP</td>
<td>Internet Control Message Protocol</td>
</tr>
<tr>
<td>IESG</td>
<td>Internet Engineering Steering Group</td>
</tr>
<tr>
<td>IETF</td>
<td>Internet Engineering Task Force</td>
</tr>
<tr>
<td>IFR</td>
<td>Instrument Flight Rules</td>
</tr>
<tr>
<td>IGP</td>
<td>Internet Gateway Protocol</td>
</tr>
<tr>
<td>IHL</td>
<td>IP Header Length</td>
</tr>
<tr>
<td>InterNIC</td>
<td>Internet Network Information Center</td>
</tr>
<tr>
<td>IP</td>
<td>Internet Protocol</td>
</tr>
<tr>
<td>IPv4</td>
<td>Internet Protocol version 4</td>
</tr>
<tr>
<td>IPv6</td>
<td>Internet Protocol version 6</td>
</tr>
<tr>
<td>ISM</td>
<td>Industrial Scientific and Medical</td>
</tr>
<tr>
<td>ISP</td>
<td>Internet Service Provider</td>
</tr>
<tr>
<td>km</td>
<td>kilometers</td>
</tr>
<tr>
<td>L1</td>
<td>Link 1</td>
</tr>
<tr>
<td>L2</td>
<td>Link 2</td>
</tr>
<tr>
<td>LAAS</td>
<td>Local Area Augmentation System</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------------------------</td>
</tr>
<tr>
<td>LGF</td>
<td>LAAS Ground Facility</td>
</tr>
<tr>
<td>LLC</td>
<td>Logical Link Control</td>
</tr>
<tr>
<td>LNAV</td>
<td>Lateral Navigation</td>
</tr>
<tr>
<td>m</td>
<td>meters</td>
</tr>
<tr>
<td>MAC</td>
<td>Media Access Control</td>
</tr>
<tr>
<td>Mcps</td>
<td>Mega chips per second</td>
</tr>
<tr>
<td>MCS</td>
<td>Master Control Station</td>
</tr>
<tr>
<td>MF</td>
<td>Medium Frequency</td>
</tr>
<tr>
<td>MTU</td>
<td>Maximum Transmission Unit</td>
</tr>
<tr>
<td>NAS</td>
<td>National Airspace System</td>
</tr>
<tr>
<td>NDGPS</td>
<td>Nationwide Differential GPS</td>
</tr>
<tr>
<td>nmi</td>
<td>nautical miles</td>
</tr>
<tr>
<td>NOAA</td>
<td>National Oceanic and Atmosphere Administration</td>
</tr>
<tr>
<td>OCS</td>
<td>Operational Control Segment</td>
</tr>
<tr>
<td>OSI</td>
<td>Open Systems Interconnection</td>
</tr>
<tr>
<td>PAR</td>
<td>Positive Acknowledgement Re-acknowledgement</td>
</tr>
<tr>
<td>PC</td>
<td>Personal Computer</td>
</tr>
<tr>
<td>PRC</td>
<td>Pseudorange Correction</td>
</tr>
<tr>
<td>PRN</td>
<td>Pseudo Random Noise</td>
</tr>
<tr>
<td>RARP</td>
<td>Reverse Address Resolution Protocol</td>
</tr>
<tr>
<td>RF</td>
<td>Radio Frequency</td>
</tr>
<tr>
<td>RFC</td>
<td>Request for Comment</td>
</tr>
<tr>
<td>RIP</td>
<td>Router Information Protocol</td>
</tr>
</tbody>
</table>
rms  root mean square
RNAV  Terminal Area Navigation
RTCA  Radio Technical Commission of America
SBAS  Satellite Based Augmentation System
SV  Space Vehicle
TCP  Transmission Control Protocol
TCP/IP  Transmission Control Protocol/Internet Protocol
TEC  Total Electron Content
TOA  Time of Arrival
UDP  User Datagram Protocol
US  United States
USACE  United States Army Corps Engineers
USAF  United States Air Force
USCG  United States Coast Guard
UTC  Universal Time Coordinated
VHF  Very High Frequency
VNAV  Vertical Navigation
WAAS  Wide Area Augmentation System
1 INTRODUCTION

Currently various digital data formats are being pursued for the Very High Frequency (VHF) frequency band for civil aviation data links. However, as the civil National Airspace System (NAS) moves from analog to digital exchanges the data links will be limited by the low data rates supported by these VHF systems [19]. A higher frequency for the data link will allow for enhancements in messaging capabilities and the Internet Protocol (IP) allows for better message handling.

The focus of this research was to investigate the potential for a data link at a higher frequency band (e.g., 5 GHz) that will integrate accurate positioning and remote positioning capabilities in a single system. In a typical Differential Global Positioning System (DGPS) architecture, correction data are sent from a reference station to a mobile user for code based systems. For high accuracy applications both the code and carrier measurements are required at the mobile user, and the user combines the code and carrier phase measurements for enhanced performance. Alternatively, remote positioning systems transmit signals for remote monitoring/positioning applications. In such remote positioning systems, the user may send just position data and/or measurements. The development of a DGPS data link over Internet Protocol (IP) would enable the integration of these two functions in a convenient message format thereby providing high performance navigation capabilities for the mobile user and also remote positioning/monitoring capability for the ground reference station. The use of IP over an ethernet interface provides for high data rates, better message handling and encryption capabilities.
This project demonstrated a proof of concept for the above proposed system at the 2.4 GHz Industrial, Scientific and Medical (ISM) frequency band. The system implemented communicates between the ground reference station and the mobile user over IP using two wireless transceivers. The wireless transceivers communicated with the processing computers at each end via the ethernet interface in real-time. The demonstration for this real-time system was set up in a laboratory in Stocker Center at Ohio University. This system lays the foundation for developing the capabilities of the data link including different types of messages including web related data, voice over IP, multiple navigation systems data for a service volume extending to 10 nautical miles (nmi), and possibly beyond.

1.1 Thesis outline

This document discusses the various aspects, implementation and results for the bi-directional DGPS system using IP. The following discussion outlines the organization of this document.

Chapter 2 contains background information relevant to this research project. It includes a discussion of the Global Positioning System (GPS), DGPS and details three GPS based augmentation systems being developed by various government agencies. It also discusses the concept of a bi-directional DGPS architecture and remote positioning applications. Chapter 3 discusses some of the relevant specifications of the Transmission Control Protocol/Internet Protocol (TCP/IP) and the socket interface involved in the implementation for data communication between the end points at the ground reference
and user. It also provides a comparison of TCP and the User Datagram Protocol (UDP) which leads to the justification for choosing UDP over TCP for this proof of concept system. Chapter 4 provides an outline of the system architecture and its implementation. Chapter 5 discusses the results obtained and conclusions. Finally, Chapter 6 provides some recommendation for further enhancement and extension of the system.
2 BACKGROUND

2.1 Global Positioning System

The Global Positioning System (GPS) is a satellite based navigational aid that provides accurate, continuous, worldwide, three-dimensional position, velocity and timing information to users with appropriate receiving equipment. It is a passive ranging system and hence can support an unlimited number of users. It is supported by a constellation of 24 satellites that broadcast ranging signals which can be used by an appropriately equipped user to determine his position [1]. Figure 2.1 provides an illustration of the satellites in their orbits [21].

![An illustration of the GPS constellation](http://gps.faa.gov/GPSbasics/index.htm)

2.1.1 System Description

As illustrated in Figure 2.2, the operation and functionality of GPS is supported by three segments, which are briefly discussed below.
Figure 2.2  GPS System Overview

a) The Space Segment

The space segment consists of a baseline constellation of 24 satellites, otherwise called space vehicles (SVs), in nearly circular orbits with an orbital radius of 26,560 kilometers (km), a period of approximately 12 hours, and repeating ground tracks over the surface of the earth. These satellites are arranged in 6 orbital planes, each of which is inclined at 55° to the equatorial plane and with 4 satellites in unevenly distributed slots [2]. The uneven distribution of the satellites in the orbit is optimized to minimize the effects of a single satellite failure on system degradation [1]. This baseline constellation ensures visibility of at least 4 satellites to any user near the earth’s surface with a clear view of the sky, though the actual visibility may extend to more than the minimum requirement.
Currently, each satellite transmits continuously on two frequencies in the L-Band referred to as Link 1 (L1) and Link 2 (L2). The center frequencies of the L1 and L2 signals are as follows: \( L1 = 1575.42 \text{ MHz} \) and \( L2 = 1227.60 \text{ MHz} \).

Two signals are transmitted on the L1 carrier and one on the L2 carrier. The L1 signal carries both a precision (P) code and a coarse/acquisition (C/A) code to be used by civil users. The L2 signal carries the P-code only. The P-code can be converted to a secure encrypted (Y) code when antispoofer (AS) mode is enabled and is labeled as P(Y)-code. The Y-code is available to Department of Defense (DOD) and authorized users only. The main purpose of the Y-code is to deter potential spoofers and jammers from spoofing the Y-code by generating a replica of the Y-code [1]. Each of the signals essentially consists of three components which are discussed below [3]:

i) **Carrier**: An RF sinusoidal signal at L1 or L2 frequency.

ii) **Ranging code**: The ranging code is a pseudo random noise (PRN) sequence unique to each satellite and generated using Direct Sequence Spread Spectrum (DSSS) techniques such that it allows the satellites to transmit the ranging code at the same frequency with minimal interference from each other. The ranging code allows the receiver to make precise range measurements by determining signal transit time, with the resolution of clock bias errors. Each satellite transmits two different codes: a C/A and the [P(Y)] code. Each C/A-code is a unique sequence of 1023 bits, called chips, with a chipping rate of 1.023 MHz [2]. The P(Y)-code however, is a unique segment of a very long PRN sequence. The chipping rate is 10.23 MHz and the chip width is 29.3
meters (m) as against the C/A-code whose chip width is 293.05 m [2]. This shorter wavelength allows for higher precision on the P(Y)-code which repeats every one week.

iii) **Navigation Message:** The navigation message is a binary coded message comprising of satellite ephemeris data, health status, SV clock error parameters and almanac. The almanac provides reduced precision ephemeris data of all satellites in the constellation. The navigation message is transmitted at 50 bits per second (bps) with bit duration of 20 ms.

Each code is combined with the binary navigation data using modulo-2 addition and then modulated using the binary phase shift keying (BPSK) technique. Each satellite transmits the P(Y)-code in addition to the C/A-code on the L1 frequency. This is accomplished by generating an in-phase and a quadrature signal on the L1 [3]. The C/A-code is used to modulate the quadrature signal and the in-phase signal is modulated by the P(Y)-code. The signal on the L2 frequency, also modulated using the BPSK technique, carries the P(Y)-code only.

b) **The Control Segment**

The functions of the control segment are [3]:

- To monitor satellite orbits,
- To maintain GPS time,
- To monitor and maintain satellite health,
- To predict satellite ephemerides and clock parameters,
• To update satellite navigation messages,
• To command maneuvers of satellites to maintain orbit and relocations to compensate for failure, if required.

The command and control functions are performed by the Master Control Station (MCS), located at Schriever Air Force Base near Colorado Springs, Colorado. In addition to the MCS, the satellite signals are tracked continuously from unmanned United States Air Force (USAF) monitor stations spread around the globe: Ascension Island, Diego Garcia, Kwajalein, and Hawaii [3]. Communications with the satellites is established via the S-Band radio links through dedicated ground antennas co-located at the monitor stations at Ascension Island, Cape Canaveral, Diego Garcia and Kwajalein [3]. These ground antennas are remotely operated by the MCS to receive telemetry from satellites on the status of the subsystems and uplink data to update the navigation messages broadcast by the satellites.

The Control Segment is also referred to as the Operational Control Segment or Operational Control System, both abbreviated as OCS [1].

c) The User Segment

Due to the passive nature of GPS it can support an unlimited number of users equipped with appropriate equipment. Due to the revolution in integrated circuit technology, receivers have continued to become more compact, lighter and cheaper. This has led to a large scale application of GPS in the civil arena. The ranging and timing information obtained from GPS has driven extensive applications in high
precision positioning, aviation, space navigation, land transportation, etc. Some of these applications are discussed below:

- GPS carrier phase measurements are being used by geodesists and geophysicists to achieve millimeter level positioning. This has led to GPS being used to study plate motion and crustal deformation, earthquakes, volcanic processes, ice sheet processes and variations in earth’s rotation. The high precision capability of GPS has become a useful tool for surveyors and mapmakers thus leading to a market for high end GPS receivers.

- GPS has also found application in civil aviation with a potential to enhance the economy and safety of air operations. The avionics involved in civil aviation have stringent performance requirements related to accuracy, continuity, availability and integrity. GPS has been certified for use in the US airspace as a supplemental system for en route, terminal and non-precision phases of flight and as a primary navigation system for oceanic and remote area operations [6].

- Moreover, the development of augmentation aids is underway to enhance the accuracy of position estimates using GPS and to meet other required criteria (as will be discussed later).

- GPS receivers are being used aboard low earth satellite orbits for position, velocity, timing, and attitude rate information, leading to reduction in cost and complexity of the required equipment.
• GPS is used to remotely track aircraft, missiles and vehicles by remote command and control assets.

• The largest current application of GPS is in land transportation, especially vehicle navigation and tracking. GPS has generated a great demand for services to motorists, public transit and emergency operations.

2.1.2 Principle of Operation

GPS works on the principle of Trilateration and Time of Arrival (TOA). Trilateration requires the knowledge of ranges from three known positions to determine the position of the unknown point. In GPS the satellites provide ranging signals to the users and hence, in principle, a user can determine his position if he has three satellites in view with a knowledge of the satellites’ positions.

The distance between the user and the satellite is determined by measuring the transit time of the signal from the satellite to the user. This is accomplished by transmitting accurate timing marks in accordance with highly accurate, stable and synchronized atomic clocks aboard the satellites. For an accurate measure of the transit time of the signal the satellite and receiver clock must be synchronized perfectly. However, this requirement can be obviated by the fact that a bias in the receiver clock at an instant of measurement will affect the transit times from all satellites in view equally. As for the SV clock error, it is estimated by the OCS at the ground and encoded into the GPS navigation message data to be incorporated by the user. Hence, if the receiver clock
bias was to be assumed to be the fourth variable, measurements from four satellites at any instant would enable a position solution for the user.

Mathematically, the equation for the pseudorange from the user to the satellite can be written as [3]:

\[
\rho_k = \sqrt{(x_k - x_u)^2 + (y_k - y_u)^2 + (z_k - z_u)^2} - cb
\]  

(2.1)

where, \((x_k, y_k, z_k)\) = Position coordinates of \(k^{th}\) SV  
\((x_u, y_u, z_u)\) = Position coordinates of User  
\(\rho_k\) = Range from the User to the \(k^{th}\) SV \([m]\)  
\(cb\) = Receiver clock bias \([m]\)  
\(k\) = SV number, where \(k = 1, 2, 3, \ldots, 32\)

Hence, if \(k \geq 4\), the set of equations can be solved to obtain an estimate of the receiver clock bias and the position coordinates of the user. A simple illustration of the principle is provided in Figure 2.3.

Figure 2.3  Principle of navigation using GPS
2.1.3 Error Sources

There are numerous error sources which affect the performance of navigation signals obtained from GPS. The main error sources can broadly be classified as follows:

- **Satellite Clock and Ephemeris Error**: The ephemeris and clock parameters broadcast by the satellites are computed by the Control Segment based on the measurements from the monitor stations. A prediction model is used to generate the ephemeris and clock parameters to be uploaded to the satellites which in turn are broadcast in the 50 bps navigation message. Evidently, the extent of the errors depends on the accuracy of the prediction model and update rate of the parameters, as well as SV geometry. The range error due to the errors in the clock and ephemeris parameters is defined as the root-sum-square value of the clock error and line of sight component of the ephemeris error. With once a day uploads (presently updates are done more than once a day [1]), the current estimates of the rms range errors due to ephemeris and clock errors is 1.5 m each [3].

- **Signal Propagation and Modeling Errors**: RF signals used in GPS are affected by the medium of propagation. GPS signals in particular travel about 20,000 km for a satellite at zenith to 26,000 km for a satellite that is rising or setting, before it reaches the receiver [1]. All of the distance traveled by the signal cannot be regarded as free space wherein electromagnetic waves travel at the speed of light (i.e., 299,792,458 m/sec [2]). Propagation through the atmosphere changes the velocity of propagation of the signals due to the
phenomenon of refraction and dispersion. This change in the speed of travel affects the transit time of the signal, which affects the fundamental measurement of GPS. Errors associated with signal propagation can be discussed primarily as errors due to propagation through the two layers of the atmosphere, namely the Ionosphere and Troposphere.

The ionosphere is a region of ionized gases extending from a height of about 50 km to 1000 km above the earth’s surface [3]. This ionization of the gases is caused by solar activity and hence the physical characteristics of the ionosphere change widely between day and night. The ionosphere is a dispersive media for GPS signals and affects the code and carrier differently. Ionosphere refraction causes a code phase delay and a carrier phase advance by the same amount [1]. This delay and advance is directly proportional to the amount of free electrons in the path of the signal, characterized by Total Electron Content (TEC). Since the path length through the ionosphere will change with the elevation of the satellite, the associated errors change with the movement of the satellites. A dual-frequency GPS receiver can estimate the errors due to the ionosphere refraction and can essentially eliminate the error. Single-frequency receivers can estimate the errors using the Klobuchar model, the parameters of which are broadcast by the satellites [1]. The delay introduced due to the ionosphere can vary from 1 m to 100 m depending on such parameters as height of user, time of day, elevation and azimuth of satellite, solar activity, user latitude, and time of year [3].

The second component of the error associated with atmosphere delay is
due to the troposphere. The troposphere extends above the earth’s surface to about 9 km above the poles and 16 km above the equator and the temperature generally decreases with height [3]. Unlike the ionosphere, the troposphere is a non-dispersive medium at GPS frequencies and delays the code and carrier phase measurements equally. This delay cannot be estimated directly from GPS measurements hence stand-alone users must rely on models to mitigate the effects of this error source. The dry and wet components of the troposphere affect the radio waves differently and hence are typically modeled separately. Various models have been developed to estimate the troposphere delay by estimating the dry and wet refractivities along the signal path. These models differ in their assumptions regarding the variations in water vapor, temperature, and pressure. The troposphere does not significantly attenuate the signal but does introduce a delay from about 2 m for a satellite at zenith to 25 m for a satellite at low elevation angles [1].

- **Receiver Measurement Error and Multipath:** The receiver measurement errors are specific to the receiver and cannot be eliminated using the DGPS technique. This error involves noise due to the antenna, amplifiers, cables, interference from other GPS signals or GPS like broadcasts, and signal quantization noise. Also a receiver cannot follow the changes in the signal waveform perfectly, resulting in delays and distortions. Another error source specific to the receiver or receiver site is signal multipath. Multipath is the phenomenon of a signal reaching the receiver via two or more paths. A
receiver receives the direct signals from the satellites and also reflections from natural and man made structures in the vicinity and ground reflections. These reflected signals are delayed and weaker versions of the signals from the satellites and the receiver sums the direct and reflected signals leading to a range measurement error at the correlation/detector level. Such errors due to receiver noise and multipath may be reduced by signal processing techniques though they cannot be entirely eliminated. Multipath mitigation also involves appropriate siting of the antenna, antenna design to mitigate the effects from these reflections, and a wide variety of receiver hardware and software techniques.

2.2 Differential GPS

Differential GPS (DGPS) enhances the performance of stand-alone GPS by removing common (i.e. correlated) errors from two or more receivers viewing the same satellites. The fundamental idea behind the technique is that errors associated with satellite clock, ephemeris, and atmospheric propagation are similar for users within a geographical area and that these errors exhibit temporal and spatial correlation. This means that the errors get decorrelated with increasing distance between the users and increasing difference in measurement epochs (i.e. time).

For a DGPS reference station that operates on pseudorange measurements, the reference station is placed at a known surveyed location so it can calculate the errors in the pseudorange measurements by comparing the measured and true range to the satellites. The reference station then transmits these correction terms to the users who use
them to correct their pseudorange measurement and obtain better performance (e.g., better accuracy and integrity). It is important to note that due to spatial decorrelation a user closer to the reference station will benefit more from the correction terms than a user farther away. Also, there will be some latency in the reference station calculating and transmitting the corrections terms and the user applying them to its measurements. Since latency cannot be completely eliminated in practical applications, minimizing it will enhance the performance of the DGPS technique. Figure 2.4 shows a simple DGPS set up illustrating the concept.

Figure 2.4  A conceptual illustration of DGPS

Using the DGPS technique, the satellite clock error can be eliminated completely. This is because these errors change very slowly over hours and is very small in
magnitude (2 m rms) [3]. It does not decorrelate with distance and hence is common to
two users viewing the same satellite. Also, similar to the satellite clock error, the
ephemeris error changes slowly over time but decorrelates spatially. But for a satellite
that is 20,000 km from the earth, the angular separation of the satellite of the lines of
sight of two users 100 km apart will be 0.3°. This error can be reduced from 2 m in stand
alone positioning to less than 5 cm of uncompensated error in DGPS [3].

This principle of eliminating the common errors experienced by two or more
users is expressed using the following equations [15].

The geometric range from the ground reference located at \((x_r, y_r, z_r)\) to an SV
located at \((x_k, y_k, z_k)\) can be calculated as:

\[
r_r = \sqrt{(x_k - x_r)^2 + (y_k - y_r)^2 + (z_k - z_r)^2}
\]  \hspace{1cm} (2.2)

where, \((x_k, y_k, z_k)\) = Position coordinates of \(k^{th}\) SV
\((x_r, y_r, z_r)\) = Known position coordinates of reference station

\(r_r\) = Geometric range from the reference station to the \(k^{th}\) SV \([\text{m}]\)
\(k = \text{SV number, where } k = 1, 2, 3, \ldots, 32\)

Also, the reference station GPS receiver makes pseudorange measurements to the
SVs. These measurements consist of the true range and associated errors which can be
represented as follows:

\[
\rho_r = r_r + \varepsilon_{m, \text{space}} + \varepsilon_{m, \text{control}} + \varepsilon_{m, user} + c\delta \hat{\gamma}_m
\]  \hspace{1cm} (2.3)
where, $\rho_r$ = Measured pseudorange by reference station receiver [m]

$\varepsilon_{m,\text{space}}$ = Space segment error for the reference station [m]

$\varepsilon_{m,\text{control}}$ = Control segment error for the reference station [m]

$\varepsilon_{m,\text{user}}$ = User segment error for the reference station [m]

$c\delta_{\text{m}}$ = Reference station receiver clock offset [m]

The difference of the geometric range and measured pseudorange from the reference receiver to the satellite can be differenced to form an error term.

$$PRC = \rho_r - r_r = +\varepsilon_{m,\text{space}} + \varepsilon_{m,\text{control}} + \varepsilon_{m,\text{user}} + c\delta_{\text{m}}$$  \hspace{1cm} (2.4)

where, $PRC$ = Pseudorange correction [m]

Similar to the reference station, a user at an unknown location makes pseudorange measurements using the GPS navigation signals to the satellites which consist of space, control, and user segment errors and a clock offset. This user can apply the differential corrections transmitted by the user to correct its measurements.

$$\rho_u^c = \rho_u - PRC$$

$$= r_u + (\varepsilon_{u,\text{space}} + \varepsilon_{u,\text{control}} + \varepsilon_{u,\text{user}} + c\delta_{u}) - (\varepsilon_{m,\text{space}} + \varepsilon_{m,\text{control}} + \varepsilon_{m,\text{user}} + c\delta_{m})$$  \hspace{1cm} (2.5)

where, $\rho_u^c$ = Corrected Pseudorange user by using differential corrections [m]

$\rho_u$ = Pseudorange measured by the user [m]

$r_u$ = True Range from user to the SV [m]

$\varepsilon_{u,\text{space}}$ = Space segment error for user [m]

$\varepsilon_{u,\text{control}}$ = Control segment error for user [m]

$\varepsilon_{u,\text{user}}$ = User segment error for user [m]

$c\delta_{u}$ = User receiver clock offset [m]
The space and control segment errors as seen by the reference station and user will be the same and hence cancel out (to the extent that they do not decorrelate). The user segment will be similar subject to the distance between the user and reference station and other parameters, and will cancel out leaving some residuals. Hence, the corrected pseudorange for the user, though incorrect, will be more accurate than a measurement without applying the differential corrections.

\[ \rho_u^c = r_u + c\delta t_{combined} + \epsilon_{u,residual} \]  

where, \( c\delta t_{combined} \) = Combined clock offset error [m] \[ \epsilon_{u,residual} \] = Residual user segment error [m]

The ionosphere propagation delay depends on the TEC in the path of the signal. This shows strong variations over time and space depending on solar activity and other factors. For users that are close enough, this error can be significantly reduced using DGPS. In the absence of any significant variations this error may be reduced to almost 0.1 - 0.2 m.

The troposphere propagation delay depends on the weather profile and can vary significantly for users separated by a few kilometers. The residual error after applying differential corrections is usually higher for low elevation satellites. Also for airborne users, the differential height affects the error corrections. For significant altitude difference and over long distances, the troposphere corrections need to be applied at both the user and reference processing ends, depending on user performance requirements.

As already discussed before, the receiver measurement error and multipath are specific to the receiver or receiver site and hence these errors cannot be mitigated using the DGPS technique. It is important to minimize these errors at the ground reference
station to avoid the user inheriting these errors in the transmitted differential correction terms.

### 2.3 GPS Augmentation Systems

The requirement of many civil and military applications for better performance has led to the development of GPS augmentation systems by several agencies to provide differential corrections to users within a geographical area. Three such augmentation systems, namely the Nationwide Differential GPS (NDGPS), Local Area Augmentation System (LAAS) and Wide area Augmentation System (WAAS), are discussed in the following section.

#### 2.3.1 Nationwide Differential Global Positioning System

The Nationwide Differential GPS (NDGPS) is a land-based GPS augmentation initiative supported by a seven agency partnership. The team comprises of the US Air Force (USAF), Federal Railroad Administration (FRA), US Army Corps of Engineers (USACE), Federal Highway Administration (FHA), National Oceanic and Atmospheric Administration (NOAA), Office of the Secretary of the Department of Transportation (DOT), and US Coast Guard (USCG). The USCG in association with USACE has developed and maintained a radio beacon network that broadcasts GPS differential corrections [4]. This system is a collection of broadcast sites that are networked with two control station components. The two Control Stations provide status and remote
operational control over each broadcast site. Each site computes GPS data corrections, combines integrity information, and transmits the resulting corrections at Medium Frequency (MF). The USCG DGPS concept was initially aimed primarily at coastal navigation.

The NDGPS is building on the USCG radio beacon network design and infrastructure to provide the Continental US (CONUS) with dual redundant DGPS radio beacon coverage. The NDGPS expansion effort, scheduled to have over 126 broadcast sites, is designed to provide double terrestrial DGPS coverage across the CONUS to meet most surface transportation navigation requirements [4].

Presently, there are 80 remote transmitting sites broadcasting C/A-code corrections to users in the US [20]. The Nationwide Control Station, fielded in 2002, provides a real-time monitoring and control of the network. The USCG joined in an effort with other government agencies to engineer and provide centimeter-enabling GPS base station data to users by means of a low cost enhancement to existing and planned NDGPS beacons. This prototype system transmits compacted data at 455 kHz, so users can produce centimeter level solutions at repetition rates as low as one second [20]. The message format includes a number of enhancements including the ability to compact dual code and carrier measurements into less than 1000 data bits to be transmitted to users [20]. Presently, the high accuracy NDGPS (HANDGPS) system is at the prototype stage achieving decimeter level accuracy in several minutes.
2.3.2 Local Area Augmentation System

The LAAS is a system being developed to support landing, precision approach and other navigation and surveillance related applications within the local area of an airport, for operational use within the US National Airspace System (NAS) [5]. The “operational goals” of LAAS as defined by RTCA [5] are to provide:

a) Performance equivalent to other Instrument Flight Rules (IFR) radio navigation precision approach and landing systems, so that existing systems can be decommissioned as appropriate;

b) A means to support future terminal area navigation enhancements using augmented GPS;

c) Support for airport surface navigation;

d) High accuracy position, velocity and time information to support future Automatic Dependent Surveillance-Broadcast (ADS-B) applications; and

e) High accuracy, high integrity positioning to support improved obstacle and terrain clearance.

Figure 2.5 provides an overview of the LAAS including its three subsystems.
LAAS is intended to be used to provide radio navigation vertical and lateral guidance for aviation IFR precision approach and landings from about 20 nmi from the runway threshold through touchdown and rollout. The system is being developed to be capable of providing this service to all aircraft in the service volume and to be to be an all-weather navigation service. Additionally, it is intended to be suitable for “precision terminal area navigation (RNAV)” in the terminal area including curved approaches and departures and for surface navigation on the airport [5].

The LAAS is intended to become the primary radio navigation system to support Category II and III precision approach and landing operations in the NAS will also
support Category I precision approach operations, particularly where those operations cannot be supported by the precision approach capability of the WAAS [5].

The performance of the LAAS is classified in Performance Types which define the levels of service. Each Performance type defines a level of integrity, availability, accuracy and continuity. Three levels of are defined namely, Performance Type 1, 2, and 3. Performance Type 1 is the performance adequate to support Category I operations. Performance Type 2 is the performance adequate to support Category II and IIIa operations. Performance Type 3 is the performance adequate to support Category IIIb operations [5].

The LAAS comprises of essentially three subsystems:

- The satellite subsystem provides the ranging signals. These signals may be provided by GPS other Satellite Based Augmentation System (SBAS) or other satellite systems such as GLONASS. These ranging signals are used both by the mobile user and the ground reference station to determine range to the satellites for navigational purposes.

- The ground subsystem (also called the LAAS Ground Facility or LGF) produces ground monitored differential corrections and integrity-related information as well as data including the definition of the final approach path, a geometric path-in-space to which the aircraft on approach will navigate.

The LGF is required to use carrier phase measurements to smooth the pseudorange measurements in a manner such that the steady state noise on the
pseudorange measurements has a spectral density with a 3 dB bandwidth of 0.01 rad/sec [5]. The LGF computes differential corrections for each satellite in view by taking pseudorange measurements from each of the multiple reference receiver sites. The pseudorange measurements from multiple receivers are compared and sufficiently large discrepancies are used to identify receiver failure or excessive measurement errors. Any faulty measurements detected in this way are excluded. The pseudorange corrections are formed by averaging the corrections for a satellite from all the receivers that pass the fault detection and exclusion process. These composite pseudorange corrections are then uplinked to the airborne user to improve positioning accuracy. Also the error statistics (e.g., B-values) and standard deviations of the fault-free measurement error for each pseudorange correction are calculated by the LGF and uplinked to the airborne user. These quantities relate to the integrity and fault-free accuracy of the pseudorange measurement. The methodology for generating these integrity parameters and the specifications for fault detection and exclusion are as mandated by RTCA [5].

Additionally, Airport Pseudolites (APL’s) may be used to increase availability. Pseudolites are ground-based transmitters that simulate the signals transmitted by GPS satellites to provide navigation and ranging information. The data portion of the Pseudolites may also contain differential corrections allowing users to correct their pseudo range measurements.

- The airborne subsystem uses the data transmitted by the LGF to correct the airborne measurements and accurately compute its position with the required
levels of integrity, continuity and availability. Both the ground and air subsystem make measurements in the range domain so the subsystem accuracy is specified in the range domain.

### 2.3.3 Wide Area Augmentation System

The Wide Area Augmentation System (WAAS) is a differential GPS augmentation system developed to extend the capabilities of the GPS [6]. This system comprises of ground stations that calculate GPS integrity and correction data and use geostationary satellites (GEOs) to transmit this data to the WAAS users.

The operational goal of the WAAS is to augment the GPS so that WAAS is the only radio navigation equipment required onboard the aircraft to meet aviation radio navigation performance requirements for oceanic, remote area and domestic en route, terminal, non-precision approach, Lateral Navigation (LNAV)/ Vertical Navigation (VNAV), Approach operations with vertical guidance (APV-II) and Global Navigation Satellite System (GNSS) Landing System (GLS) phases of flight [6]. The WAAS signal provides the augmentation to GPS to obtain the required accuracy improvement for precision approaches, as well as integrity, continuity, and availability of navigation for all phases of flight. Additional goals for WAAS as defined in RTCA guidelines [6] are to provide:

a. flexibility for future enhancements;

b. worldwide primary (sole) means radio navigation;

c. positioning and time for automatic dependent surveillance;

d. ground movement monitoring (with augmentation);
e. growth to GPS/Local Area Augmentation System for Category III precision

f. approach; and

g. replacement of other radio navigation systems.

Figure 2.6 provides an overview of the WAAS architecture as suggested by the RTCA [6].

WAAS reference stations and integrity monitors are data collection sites that contain WAAS ranging receivers which monitor all signals from the GPS, as well as the WAAS geostationary satellites. The reference stations collect measurements from the GPS and WAAS satellites so that differential corrections, ionospheric delay information,
WAAS accuracy, network time, GPS time, and Universal Time Coordinated (UTC) can be determined.

The data from these dispersed reference stations and integrity monitors are forwarded to the central data processing sites. These sites process the data in order to determine differential corrections, ionospheric delay information, and WAAS accuracy, as well as verify residual error bounds for each monitored satellite. The central data processing sites also generate navigation messages for the geostationary satellites and WAAS messages. This information is modulated on the GPS-like signal and broadcast to the users from geostationary satellites.

The WAAS signal is transmitted from a geostationary satellite on the GPS L1 frequency (1575.42 MHz). The WAAS data is a 500 symbols/second stream modulo-2 added to a 1023-bit PRN code. The result is then bi-phase shift-key (BPSK) modulated onto the L1 carrier frequency at a rate of 1.023 Mega-chips/second (Mcps).

WAAS is being developed to provide service for all classes of aircraft in all flight operations - including en route navigation, airport departures, and airport arrivals. This includes precision landing approaches in all weather conditions at all locations throughout the NAS.

2.4 Remote Positioning

Remote Positioning applications require the real-time positioning of a remote user by a reference station for surveillance or purposes. For a GPS based system, the remote user computes its position and transmits the position coordinates to the reference over a data link. The data downlinked to the reference station may include just position coordinates
and/or measurements made by the user. This information is used by the reference station for monitor/control/track operations specific to the application. Some of the applications of remote positioning are as follows:

- Remote positioning systems can be used for applications such as ADS-B and Airport Surface Movement Guidance (ASMG)
- Remote positioning systems can be used to track and eventually maneuver or control an Unmanned Airborne Vehicle (UAV)
- The high accuracy position information obtained from a remote positioning system, such as that employing DGPS, could be used for launch vehicle (LV) or missile tracking and maneuvering.
- Remote positioning can be used for supplementing and/or improving performance of airborne collision avoidance systems (ACAS).
- Apart from tracking airborne users, remote positioning systems such as those employing GPS maybe used for monitoring/tracking surface vehicles movement within a required area. Such an application may lead to better situational awareness within an area, as specified by the application.

2.4.1 Automatic Dependent Surveillance – Broadcast (ADS-B)

ADS-B is a technology associated with Communication, Navigation and Surveillance (CNS) in Air Traffic Management (ATM) that enables an air traffic controller and/or other pilots to have precise information about the location of an aircraft
with additional information such as heading, attitude or velocity [22]. Such a system provides an important application for a remote positioning system.

In a typical application, the ADS-B capable aircraft uses a GNSS (GPS, Galileo) receiver to derive its precise position from satellite constellation. This position information is then combined with other parameters such as heading, altitude or flight number and then simultaneously broadcast to other ADS-B capable aircraft and to ADS-B ground stations, or satellite communications transceivers which then relay the aircraft’s position and additional information to Air Traffic Control (ATC) centers in real-time [22].

Some of the operational applications of an ADS-B system as defined by RTCA are as follows [22]:

- Cockpit Display of Traffic Information (CDTI): ADS-B data can be used to provide data for CDTI displays which provide the flight crew with surveillance information about other aircrafts including their position
- Airborne Collision Avoidance: ADS-B data may be used to improve surveillance performance e.g. by reducing the potential for airborne interference with ground radar interrogations.
- Conflict Management and Airspace Deconfliction: Since both the pilot and the ground-based controller have the same view of the surveillance picture, which includes the position of own and other aircrafts, resolution maneuvers can be better coordinated.
• Air Traffic Service (ATS) Conformance Monitoring: ADS-B data can be used for ATS monitoring which involves the process of ensuring that an aircraft conforms to its agreed-to trajectory.

Apart from above applications, ADS-B which is essentially a one way broadcast format, can be used for applications at low altitudes, ground surveillance or other remote areas where radar coverage is limited by elevation problems or line of sight.

2.5 Bi-directional DGPS

As already discussed, the transmission of DGPS corrections via a data link enables differentially corrected high performance at the mobile user. This provides user benefits for precise positioning, auto pilot, and precision approach landings. Also, a downlink capability via the data link from the mobile user to a reference station would enable remote monitoring and positioning, whereby the mobile user transmitting its measurements and position solution to a ground-based station that could enable remote positioning, an integrity check and control/track operations by the ground-based station. Such a bi-directional DGPS was developed at the AEC at Ohio University which employed a data link at 900 MHz for communication between the ground reference and mobile user using RS-232 based transceiver radios [7]. Figure 2.7 provides an illustration of the bi-directional DGPS architecture.
The bi-directional data link discussed in this research incorporates both of these features over the same functional data link using IP. This section explains the basic functionality of the ground-based station and the mobile user to enable such a capability.

The ground-based reference station is based on a simplified version of the prototype LAAS ground subsystem. The associated antenna is located at a surveyed location and equipped to receive and process GPS signals. The ground station processes the GPS signals to extract carrier phase observables and measured pseudoranges to the GPS satellites in view. The carrier phase measurements are used to smooth the pseudorange measurements. Using the satellites clock and ephemeris data it calculates the geometric range to the satellites based on the “true” known location of the ground-based reference station (2.2). This calculated geometric range is then compared with the smoothed pseudorange to generate the error in measurement. This error is a “lumped”
correction term which includes all the associated errors (2.4). The Tropospheric delay is eventually removed by the mobile user who incorporates a differential tropospheric error correction model. The ground reference station then broadcasts these correction terms and associated timing and integrity information (i.e., B-values) to the users in the local area over the data link.

The mobile user is equipped to receive and process GPS signals from the satellites, to receive data (e.g. PRCs, B-values, etc.) over the data link from the ground reference station, and to transmit its position with pseudorange and carrier phase measurements to the ground station. The mobile user uses its carrier phase measurements to smooth the pseudorange measurements. These pseudorange measurements are corrected by applying the differential corrections from the ground reference station, to eliminate the common errors from its measurement (2.5). This enables the user to generate a high performance solution with added integrity. The mobile user then transmits its measurements, calculated location coordinates and timing information to the ground reference station over the same data link.

The measurement and position data from the airborne mobile user provides the ground subsystem with the remote positioning data to track/control the user. The ground reference uses the measurements from the mobile user to generate an accurate estimate of the user. This remote user solution by the ground station can be compared to the position coordinates calculated and transmitted by the user, to check for any discrepancies. If a discrepancy exists, the ground station can generate an integrity warning message to the user. This process provides for a closed loop integrity check for the system.
3 TRANSMISSION CONTROL PROTOCOL/INTERNET PROTOCOL

Transmission Control Protocol/Internet Protocol or TCP/IP as it is better known defines the protocol suite including all the protocols used by the Internet. It can communicate across any set of interconnected networks and the Internet is the best example of its large scale implementation.

A protocol, like TCP or IP, is a standard that defines a set of rules and specifications. The TCP/IP protocols provides for such standards for communication between networks and computers by specification of message length, error handling etc. One of the most viable features of using the TCP/IP protocol suite is that it allows for exchange of information between computers with no dependence on the hardware or network interface involved.

As a result, programmers or users don’t need to develop different versions of the software as they move it across different machines. In this section, the TCP and IP will be detailed. Also, UDP is discussed here along with the factors that influenced choosing UDP over TCP for communication across the data link in this project.

3.1 The Open System Interconnection (OSI) Model and TCP/IP

The Open System Interconnection (OSI) Model defines the networking framework for implementing protocols in seven layers [8]. Control starts from the top at the application layer, and goes down to the physical layer where it is transmitted across to the user and then goes up the hierarchy to end at the application layer at the user end. Table 3.1 summarizes the seven layers of the OSI model.
<table>
<thead>
<tr>
<th>Layer</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application (Layer 7)</td>
<td>This layer supports application and end-user processes. Communication partners are identified, quality of service is identified, user authentication and privacy are considered, and any constraints on data syntax are identified. Everything at this layer is application-specific. This layer provides application services for file transfers, e-mail, and other network software services.</td>
</tr>
<tr>
<td>Presentation (Layer 6)</td>
<td>This layer provides independence from differences in data representation (e.g., encryption) by translating from application to network format, and vice versa. The presentation layer works to transform data into the form that the application layer can accept.</td>
</tr>
<tr>
<td>Session (Layer 5)</td>
<td>This layer establishes, manages and terminates connections between applications. The session layer sets up, coordinates, and terminates conversations, exchanges, and dialogues between the applications at each end.</td>
</tr>
<tr>
<td>Transport (Layer 4)</td>
<td>This layer provides transparent transfer of data between end systems, or hosts, and is responsible for end-to-end error recovery and flow control.</td>
</tr>
<tr>
<td>Network (Layer 3)</td>
<td>This layer provides switching and routing technologies, creating logical paths, known as virtual circuits, for transmitting data from node to node. Routing and forwarding are functions of this layer, as well as addressing, internetworking, error handling, congestion control and packet sequencing.</td>
</tr>
<tr>
<td>Data Link (Layer 2)</td>
<td>At this layer, data packets are encoded and decoded into bits. It furnishes transmission protocol knowledge and management and handles errors in the physical layer, flow control and frame synchronization. The data link layer is divided into two sub-layers: The Media Access Control (MAC) layer and the Logical Link Control (LLC) layer. The MAC sub-layer controls how a computer on the network gains access to the data and permission to transmit it. The LLC layer controls frame synchronization, flow control and error checking.</td>
</tr>
<tr>
<td>Physical (Layer 1)</td>
<td>This layer conveys the bit stream - electrical impulse, light or radio signal -- through the network at the electrical and mechanical level. It provides the hardware means of sending and receiving data on a carrier, including defining cables, cards and physical aspects. Fast Ethernet and RS232are protocols with physical layer components.</td>
</tr>
</tbody>
</table>
The TCP/IP model, which is often called the TCP/IP stack, is conceptually very similar to the OSI model, different only in some layer functionalities. The TCP/IP model essentially has four layers [9]. The four layers are discussed below with the layers numbered according to the corresponding layers in the OSI model:

- **Layer 5 – Application Layer:** This layer in the TCP/IP model handles the responsibilities of the layers 5, 6 and 7 in the OSI model. It is responsible for authentication, compression and end user services. Some of the available utilities are File Transfer Protocol, electronic mail, telnet etc.

- **Layer 4 – Transport Layer:** The function of the transport layer is similar to that in the OSI model, except for one important difference. In contrast to the OSI model, the transport layer in the TCP/IP model does provide for an option where end-to-end delivery of packets is not guaranteed. The UDP protocol used at this layer does not assure reliable delivery of packets. The TCP also resides at this layer, but provides for connection oriented, reliable delivery of packets.

- **Layer 3 – Network Layer:** This layer is associated with routing of the data packets and uses protocols as the IP, Internet Control Message Protocol (ICMP) etc.

- **Layer 2 – Link Layer:** The link layer is associated with the operating system device driver to the network interface on the computer and also the actual transmission of packets across the network. Also it is concerned with the IP to ethernet address translation and implements such protocols as the Address Resolution Protocol (ARP), Reverse Address Resolution Protocol (RARP).
3.2 The Internet Protocol

The Internet Protocol (IP) is a network-layer (Layer 3) protocol that contains addressing information and some control information that enables packets to be routed across a network [9]. IP is documented in Request for Comment (RFC) 791 and is the primary network-layer protocol in the Internet protocol suite. An RFC is a draft of a standard or protocol for the Internet, which is published by the Internet Engineering Task Force (IETF) on approval by the Internet Engineering Steering Group (IESG) [8].

IP has two primary responsibilities: providing connectionless, best-effort delivery of datagrams through an internetwork; and providing fragmentation and reassembly of datagrams to support data links with different maximum-transmission unit (MTU) sizes [11]. This section details some of the features of IP that support these functions.

3.2.1 IP Packet Format

An IP packet contains several types of information e.g. version number, the header length, type of protocol, source and destination address etc. Table 3.2 provides a representation of the composition of the packet.
The fourteen fields associated with an IP packet, as shown in Table 3.2 are briefly discussed as follows:

- **Version**: Indicates the version of IP currently used.

- **IP Header Length (IHL)**: Indicates the datagram header length in 32-bit words.

- **Type-of-Service**: Specifies how an upper-layer protocol would like a current datagram to be handled, and assigns datagrams various levels of importance.

- **Total Length**: Specifies the length, in bytes, of the entire IP packet, including the data and header.

- **Identification**: Contains an integer that identifies the current datagram. This field is used to help piece together datagram fragments.
• **Flags:** Consists of a 3-bit field of which the two low-order (least-significant) bits control fragmentation. The low-order bit specifies whether the packet can be fragmented. The middle bit specifies whether the packet is the last fragment in a series of fragmented packets. The third or high-order bit is not used.

• **Fragment Offset:** Indicates the position of the fragment's data relative to the beginning of the data in the original datagram, which allows the destination IP process to properly reconstruct the original datagram.

• **Time-to-Live:** Maintains a counter that gradually decrements down to zero, at which point the datagram is discarded. This keeps packets from looping endlessly.

• **Protocol:** Indicates which upper-layer protocol receives incoming packets after IP processing is complete.

• **Header Checksum:** Helps ensure IP header integrity.

• **Source Address:** Specifies the sending node.

• **Destination Address:** Specifies the receiving node.

• **Options:** Allows IP to support various options, such as security.

• **Data:** Contains upper-layer information.
3.2.2 IP Addressing

As with any other network-layer protocol, the IP addressing scheme is integral to the process of routing IP datagrams through an internetwork. Each IP address has specific components and follows a basic format. These IP addresses can be subdivided and used to create addresses for sub networks, as discussed in more detail later in this chapter.

Each host on a TCP/IP network is assigned a unique 32-bit logical address that is divided into two main parts: the network number and the host number [11]. The network number identifies a network and must be assigned by the Internet Network Information Center (InterNIC) if the network is to be part of the Internet [12]. An Internet Service Provider (ISP) can obtain blocks of network addresses from the InterNIC and can itself assign address space as necessary. The host number identifies a host on a network and is assigned by the local network administrator.

3.2.3 IP Address Format

The 32-bit IP address is grouped eight bits at a time, separated by dots, and represented in decimal format (known as dotted decimal notation). Each bit in the octet has a binary weight (128, 64, 32, 16, 8, 4, 2, 1). The minimum value for an octet is 0, and the maximum value for an octet is 255. Figure 3.1 illustrates the basic format of an IP address.
3.2.4 IP Address Classes

IP addressing supports five different address classes: A, B, C, D, and E. Only classes A, B, and C are available for commercial use. The left-most (high-order) bits indicate the network class. Table 3.3 provides reference information about the five IP address classes [12].
Table 3.3  Reference information about the five IP classes

<table>
<thead>
<tr>
<th>IP Address Class</th>
<th>Format</th>
<th>Purpose</th>
<th>High-Order Bit(s)</th>
<th>Address Range</th>
<th>No. Bits Network/Host</th>
<th>Max. Hosts</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>N.H.H.H.H</td>
<td>Few large organizations</td>
<td>0</td>
<td>1.0.0.0 to 126.0.0.0</td>
<td>7/24</td>
<td>16,777, 214 (2^24 - 2)</td>
</tr>
<tr>
<td>B</td>
<td>N.N.H.H</td>
<td>Medium-size organizations</td>
<td>1, 0</td>
<td>128.1.0.0 to 191.254.0.0</td>
<td>14/16</td>
<td>65, 543 (2^16 - 2)</td>
</tr>
<tr>
<td>C</td>
<td>N.N.N.H</td>
<td>Relatively small organizations</td>
<td>1, 1, 0</td>
<td>192.0.1.0 to 223.255.254.0</td>
<td>22/8</td>
<td>245 (2^8 - 2)</td>
</tr>
<tr>
<td>D</td>
<td>N/A</td>
<td>Multicast groups (RFC 1112)</td>
<td>1, 1, 1, 0</td>
<td>224.0.0.0 to 239.255.255.255</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>E</td>
<td>N/A</td>
<td>Experimental</td>
<td>1, 1, 1, 1</td>
<td>240.0.0.0 to 254.255.255.255</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Where, N = Network number, H = Host number.

One address is reserved for the broadcast address, and one address is reserved for the network. The class of address can be determined easily by examining the first octet of the address and mapping that value to a class range in the following table. In an IP address of 132.234.14.15, for example, the first octet is 132 which lies between 128 and 191. Hence, 132.234.14.15 is a Class B address.
3.2.5 IP Subnet Addressing

IP networks can be divided into smaller networks called sub networks (or subnets). Subnetting provides the network administrator with several benefits, including extra flexibility, more efficient use of network addresses, and the capability to contain broadcast traffic (a broadcast will not cross a router).

Subnets are under local administration. As such, the outside world sees an organization as a single network and has no detailed knowledge of the organization's internal structure.

A given network address can be broken up into many sub-networks. For example, 172.16.1.0, 172.16.2.0, 172.16.3.0, and 172.16.4.0 are all subnets within network 171.16.0.0. (All 0s in the host portion of an address specifies the entire network.)

3.2.6 Address Resolution Protocol (ARP) Overview

For two machines on a given network to communicate, they must know the other machine's physical (or MAC) addresses. By broadcasting Address Resolution Protocols (ARPs), a host can dynamically discover the MAC-layer address corresponding to a particular IP network-layer address.
After receiving a MAC-layer address, IP devices create an ARP cache to store the recently acquired IP-to-MAC address mapping, thus avoiding having to broadcast ARPs when they want to re-contact a device. If the device does not respond within a specified time frame, the cache entry is flushed.

In addition the Reverse Address Resolution Protocol (RARP) is used to map MAC-layer addresses to IP addresses. RARP, which is the logical inverse of ARP, might be used by diskless workstations that do not know their IP addresses when they boot. RARP relies on the presence of a RARP server with table entries of MAC-layer-to-IP address mappings.

### 3.2.7 Internet Routing

Internet routing devices traditionally have been called gateways. In today's terminology, however, the term gateway refers specifically to a device that performs application-layer protocol translation between devices. Interior gateways refer to devices that perform these protocol functions between machines or networks under the same administrative control or authority, such as a corporation's internal network. These are known as autonomous systems. Exterior gateways perform protocol functions between independent networks [12].

Routers within the Internet are organized hierarchically. Routers used for information exchange within autonomous systems are called interior routers, which use a
variety of Interior Gateway Protocols (IGPs) to accomplish this purpose. The Routing Information Protocol (RIP) is an example of an IGP.

Routers that move information between autonomous systems are called exterior routers. These routers use an exterior gateway protocol to exchange information between autonomous systems. The Border Gateway Protocol (BGP) is an example of an exterior gateway protocol.

3.2.8 IP Routing

IP routing protocols are dynamic. Dynamic routing calls for routes to be calculated automatically at regular intervals by software in routing devices. This contrasts with static routing, where routers are established by the network administrator and do not change until the network administrator changes them.

An IP routing table, which consists of destination address/next hop pairs, is used to enable dynamic routing. An entry in this table, for example, would be interpreted as follows: to get to network 172.31.0.0, send the packet out of the specific Ethernet interface.

IP routing specifies that IP datagrams travel through internetworks one hop at a time. The entire route is not known at the onset of the journey, however. Instead, at each stop, the next destination is calculated by matching the destination address within the datagram with an entry in the current node's routing table.
Each node's involvement in the routing process is limited to forwarding packets based on internal information. The nodes do not monitor whether the packets get to their final destination, nor does IP provide for error reporting back to the source when routing anomalies occur.

3.3 Transmission Control Protocol

The TCP provides reliable transmission of data in an IP environment and corresponds to the transport layer (Layer 4) of the OSI reference model. Among the services TCP provides are stream data transfer, reliability, efficient flow control, full-duplex operation, and multiplexing [13].

With stream data transfer, TCP delivers an unstructured stream of bytes identified by sequence numbers. This service benefits applications because they do not have to chop data into blocks before handing it off to TCP. Instead, TCP groups bytes into segments and passes them to IP for delivery [13].

TCP offers reliability by providing connection-oriented, end-to-end reliable packet delivery through an internetwork. It does this by sequencing bytes with a forwarding acknowledgment number that indicates to the destination the next byte the source expects to receive. Bytes not acknowledged within a specified time period are retransmitted. The reliability mechanism of TCP allows devices to deal with lost, delayed, duplicate, or misread packets. A time-out mechanism allows devices to detect lost packets and request retransmission.
TCP offers efficient flow control, which means that, when sending acknowledgments back to the source, the receiving TCP process indicates the highest sequence number it can receive without overflowing its internal buffers.

Full-duplex operation means that TCP processes can both send and receive at the same time.

Finally, TCP’s multiplexing means that numerous simultaneous upper-layer conversations can be multiplexed over a single connection.

### 3.3.1 TCP Connection Establishment

To use reliable transport services, TCP hosts must establish a connection-oriented session with one another. Connection establishment is performed by using a "three-way handshake" mechanism.

A three-way handshake synchronizes both ends of a connection by allowing both sides to agree upon initial sequence numbers. This mechanism also guarantees that both sides are ready to transmit data and know that the other side is ready to transmit as well. This is necessary so that packets are not transmitted or retransmitted during session establishment or after session termination.

Each host randomly chooses a sequence number used to track bytes within the stream it is sending and receiving. Then, the three-way handshake proceeds in the following manner [12]:

...
The first host (Host A) initiates a connection by sending a packet with the initial sequence number (X) and SYN bit set to indicate a connection request. The second host (Host B) receives the SYN, records the sequence number X, and replies by acknowledging the SYN (with an ACK = X + 1). Host B includes its own initial sequence number (SEQ = Y). An ACK = 20 means the host has received bytes 0 through 19 and expects byte 20 next. This technique is called forward acknowledgment. Host A then acknowledges all bytes Host B sent with a forward acknowledgment indicating the next byte Host A expects to receive (ACK = Y + 1). Data transfer then can begin.

3.3.2 Positive Acknowledgment and Retransmission (PAR)

A simple transport protocol might implement a reliability-and-flow-control technique where the source sends one packet, starts a timer, and waits for an acknowledgment before sending a new packet. If the acknowledgment is not received before the timer expires, the source retransmits the packet. Such a technique is called positive acknowledgment and retransmission (PAR).

By assigning each packet a sequence number, PAR enables hosts to track lost or duplicate packets caused by network delays that result in premature retransmission. The sequence numbers are sent back in the acknowledgments so that the acknowledgments can be tracked.

PAR is an inefficient use of bandwidth, however, because a host must wait for an acknowledgment before sending a new packet, and only one packet can be sent at a time.
3.3.3 TCP Sliding Window

A **TCP sliding window** provides more efficient use of network bandwidth than PAR because it enables hosts to send multiple bytes or packets before waiting for an acknowledgment.

In TCP, the receiver specifies the current window size in every packet. Because TCP provides a byte-stream connection, window sizes are expressed in bytes. This means that a window is the number of data bytes that the sender is allowed to send before waiting for an acknowledgment. Initial window sizes are indicated at connection setup, but might vary throughout the data transfer to provide flow control. A window size of zero, for instance, means "Send no data."

In a TCP sliding-window operation, for example, the sender might have a sequence of bytes to send (i.e. numbered 1 to 10) to a receiver who has a window size of five. The sender then would place a window around the first five bytes and transmit them together. It would then wait for an acknowledgment.

The receiver would respond with an ACK = 6, indicating that it has received bytes 1 to 5 and is expecting byte 6 next. In the same packet, the receiver would indicate that its window size is 5. The sender then would move the sliding window five bytes to the right and transmit bytes 6 to 10. Once received, the receiver would respond with an ACK = 11, indicating that it is expecting sequenced byte 11 next. In this packet, the receiver might indicate that its window size is 0 (because, for example, its internal buffers are full). At
this point, the sender cannot send any more bytes until the receiver sends another packet with a window size greater than 0.

3.3.4 TCP Packet Format

Table 3.4 illustrates the twelve fields that comprise a TCP packet [13].

<table>
<thead>
<tr>
<th>Table 3.4</th>
<th>The Format of a TCP Packet</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Source Port</td>
</tr>
<tr>
<td></td>
<td>Sequence Number</td>
</tr>
<tr>
<td>Data Offset</td>
<td>Reserved</td>
</tr>
<tr>
<td>Checksum</td>
<td>Urgent Pointer</td>
</tr>
<tr>
<td>Options ( + padding)</td>
<td></td>
</tr>
<tr>
<td>Data (variable)</td>
<td></td>
</tr>
</tbody>
</table>

The following descriptions summarize the TCP packet fields illustrated in Table 3.2:

- *Source Port* and *Destination Port*: Identifies points at which upper-layer source and destination processes receive TCP services.
• *Sequence Number:* Usually specifies the number assigned to the first byte of data in the current message. In the connection-establishment phase, this field also can be used to identify an initial sequence number to be used in an upcoming transmission.

• *Acknowledgment Number:* Contains the sequence number of the next byte of data the sender of the packet expects to receive.

• *Data Offset:* Indicates the number of 32-bit words in the TCP header.

• *Reserved:* Remains reserved for future use.

• *Flags:* Carries a variety of control information, including the SYN and ACK bits used for connection establishment, and the FIN bit used for connection termination.

• *Window:* Specifies the size of the sender's receive window (that is, the buffer space available for incoming data).

• *Checksum:* Indicates whether the header was damaged in transit.

• *Urgent Pointer:* Points to the first urgent data byte in the packet.

• *Options:* Specifies various TCP options.

• *Data:* Contains upper layer information.

### 3.4 User Datagram Protocol

The User Datagram Protocol (UDP) is a connectionless transport-layer protocol (Layer 4) that belongs to the Internet Protocol family. UDP is basically an interface
between IP and upper-layer processes. Unlike TCP, UDP adds no reliability, flow-control, or error-recovery functions to IP [14]. Because of its simplicity, UDP headers contain fewer bytes and consume less network overhead than TCP.

An IP address is used to direct the user datagram to a particular machine, and the destination port number in the UDP header is used to direct the UDP datagram to a specific application process located at the IP address. The UDP header also contains a source port number that allows the receiving process to know how to respond to the user datagram.

UDP has a unique protocol number in the IP header (number 17), which enables the IP software to pass the data portion of the IP datagram to the UDP software. UDP then uses the destination port number to direct the data to the appropriate process.

Table 3.5 illustrates the format of a UDP packet [14].

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>UDP source port</td>
<td>Source port number</td>
</tr>
<tr>
<td>UDP destination port</td>
<td>Destination port number</td>
</tr>
<tr>
<td>UDP message length</td>
<td>Message length</td>
</tr>
<tr>
<td>UDP Checksum</td>
<td>Checksum</td>
</tr>
</tbody>
</table>

The following description summarizes the fields in the UDP packet:

- **Source/Destination Port Numbers**: As discussed before the source and destination port numbers, in conjunction with the IP addresses define the end points of the communication between two (or more) applications. The
destination port number is meaningful only within the context of a particular IP address. The source port, when meaningful, specifies the port to which a reply should be sent in the absence of any other information. If not used, the value is initialized to zero.

- **Length Field:** This 16-bit field contains a count of the total number of octets used in the user datagram, including the header. The minimum value of the length field is 8, which is the number of octets in a UDP header.

- **Checksum:** Checksum is the 16-bit one's complement of the one's complement sum of a pseudo header of information from the IP header, the UDP header, and the data, padded with zero octets at the end (if necessary) to make a multiple of two octets. To include fields from the IP header, a pseudo UDP header that contains the source IP address, destination IP address, protocol number and length of UDP datagram is constructed. This is followed by the UDP header, UDP data, and a pad octet equal to zero, if necessary to force the 16-bit alignment. If the computed checksum is zero, it is transmitted as all ones (the equivalent in one's complement arithmetic). An all zero transmitted checksum value means that the transmitter generated no checksum (for debugging or for higher level protocols).
### 3.5 TCP and UDP: a comparison

Table 3.6 summarizes the discussion on TCP and UDP by comparing the two protocols based on some of their characteristics.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>TCP</th>
<th>UDP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General Description</strong></td>
<td>Full-featured protocol that allows applications to send data reliably without worrying about network layer issues.</td>
<td>Simple, high-speed, low-functionality “wrapper” that interfaces applications to the network layer</td>
</tr>
<tr>
<td><strong>Protocol Connection Setup</strong></td>
<td>Connection-oriented; connection must be established prior to transmission.</td>
<td>Connectionless; data is sent without setup.</td>
</tr>
<tr>
<td><strong>Data Interface To Application</strong></td>
<td>Stream-based; data is sent by the application with no particular structure.</td>
<td>Message-based; data is sent in discrete packages by the application.</td>
</tr>
<tr>
<td><strong>Reliability and Acknowledgments</strong></td>
<td>Reliable delivery of messages; all data is acknowledged.</td>
<td>Unreliable, best-effort delivery without acknowledgments.</td>
</tr>
<tr>
<td><strong>Retransmissions</strong></td>
<td>Delivery of all data is managed, and lost data is retransmitted automatically.</td>
<td>Not performed. Application must detect lost data and retransmit if needed.</td>
</tr>
<tr>
<td><strong>Features Provided to Manage Flow of Data</strong></td>
<td>Flow control using sliding windows; window size adjustment heuristics; congestion avoidance algorithms.</td>
<td>None</td>
</tr>
<tr>
<td><strong>Overhead</strong></td>
<td>Low, but higher than UDP</td>
<td>Very low</td>
</tr>
<tr>
<td><strong>Transmission Speed</strong></td>
<td>High, but not as high as UDP</td>
<td>Very high</td>
</tr>
<tr>
<td><strong>Data Quantity</strong></td>
<td>Small to very large amounts of data (up to gigabytes)</td>
<td>Small to moderate amounts of data (up to a few hundred bytes)</td>
</tr>
</tbody>
</table>
3.6 The Socket Interface

The socket is essentially an end point for inter-process communications. It was first introduced in the Berkeley UNIX as a new operating system abstraction. As a generalization of the UNIX system input/output (I/O), it provides options for several network protocols apart from TCP and IP. The key factor in its introduction and development was the ease of use with which applications could be programmed with TCP/IP [15].

Berkeley UNIX or BSD UNIX refers to the version of UNIX developed by University of California’s Berkeley Software Distribution (BSD) [15]. The BSD software for TCP/IP became popular because of additional protocols, utilities and network services incorporated in the TCP/IP protocol suite.

The computers used for processing the data for the research described in this thesis use the QNX 4.25 operating system. This operating system incorporates the BSD 4.3 TCP/IP stack including its functions, utilities and services. As a result the socket interface was used for data communications over IP via the ethernet.
The socket forms the basis for I/O in BSD UNIX and can be considered as a generalization of the UNIX file access mechanism. As in the case with file access, a request for creation of a socket returns an integer value from the operating system. This integer value serves as the socket descriptor similar to a file descriptor when a file is opened. The difference between file and socket descriptor is that when the application calls an “open” function, the operating system binds the file descriptor to a specific file or device, but it can create a socket without binding them to a specific destination address [15].

The creation of a socket allows the programmer to specify the protocol family and the specific protocol to be used from the protocol family. For this research the UDP was used from the TCP/IP (specified as AF_INET) family [15]. On the creation of the socket the operating system returns the socket descriptor which can be used to perform operations on the socket.

After creation of the socket, the *bind ()* system call may be used to establish a local address for it. This address structure specifies the protocol port and the IP address of the machine.

After creation and binding of a socket, it may be used for read and write operations depending upon the protocol it uses. If it uses the TCP, then a connection has to be established with the destination before the data can be read or written. However, because of a preference for broadcasting the data, the connectionless method involving UDP was chosen for transmission and reception of data in this research. The data can be sent by specifying the destination address, the address of the data to be sent and the maximum size of the data.
4 SYSTEM ARCHITECTURE AND DESCRIPTION

4.1 An Overview

The proposed system demonstrates a real-time bi-directional DGPS architecture using a data link over IP. The set up comprises of a base station and a mobile user, both of which are equipped with a personal computer (PC) for processing measurements, a GPS Receiver and a wireless transceiver. This section provides a description of the system set up including functionality of the base station and the mobile user and also some real-time and practical considerations essential to the performance of the system.

4.2 Ground-Based Station

The base station is located at a surveyed known location so it can compute the errors in pseudo range measurements and transmit them to a mobile user for accurate position solution. The functional block diagram of the ground reference system is shown in Figure 4.1 [7]. The GPS receiver decodes the GPS navigation signals coming in from the GPS antenna. The communication port of the receiver is connected to the serial port of the computer. The ephemeris and time data are used to calculate the satellite positions. The pseudorange measurements are smoothed using the integrated Doppler measurements which are compared to the true geometric range to the satellite. The pseudorange corrections (PRCs) are then generated by adding in the satellite clock correction and clock bias estimations. These PRCs are transmitted using the ethernet to wireless bridge.
Figure 4.1  Bi-directional DGPS Ground Reference Station Block Diagram

After the mobile user computes its accurate position using the differential corrections, it transmits back its measurements and calculated DGPS position coordinates. The ground station, on receiving this data, smoothes the pseudorange measurements using the carrier phase measurements (i.e. integrated Doppler measurements). It also calculates the troposphere correction term using the same method as calculated at the mobile user. It then applies the troposphere correction, PRCs, satellite clock corrections to the smoothed pseudoranges and uses them to obtain a solution (i.e., position and receiver clock bias) for the mobile user. This calculated position is then compared with the transmitted position of mobile user for integrity check.
Additionally, the user position solution type (e.g., stand-alone, BDGPS, LAAS, WAAS, GPS, Inertial Navigation System (INS), etc.) can be embedded into the broadcast via a defined solution type. This feature was not implemented but could be implemented in future research.

4.2.1 Data Uplink Message Format

As illustrated, the data uplinked from the ground reference to the mobile user constitutes 222 bytes and the message format is very similar to the LAAS message type [5]. The header sums up to another 6 bytes. It comprises of the differential correction terms for each SV in view and other timing and integrity related information. Additionally, the ground reference also transmits its pseudorange and carrier phase measurements to the mobile user.

Table 4.1 shows the parameters, name, format, and size, transmitted from the ground station to the mobile user in the prototype software setup [7].
Table 4.1 Data Parameters Uplinked

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Format</th>
<th>Number of Bytes</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPS Time</td>
<td>Double</td>
<td>8</td>
</tr>
<tr>
<td>$\sigma_{pr_{gnd}}$</td>
<td>Integer</td>
<td>1</td>
</tr>
<tr>
<td>Data Checksum</td>
<td>Integer</td>
<td>4</td>
</tr>
<tr>
<td>SV Number (for up to 13 satellites)</td>
<td>Integer</td>
<td>$1 \times 13$</td>
</tr>
<tr>
<td>L1 Pseudorange Corrections (for up to 13 satellites)</td>
<td>Float</td>
<td>$4 \times 13$</td>
</tr>
<tr>
<td>L1 Integrated Doppler (for up to 13 satellites)</td>
<td>Float</td>
<td>$4 \times 13$</td>
</tr>
<tr>
<td>B-Values from first RR (for up to 13 satellites)</td>
<td>Integer</td>
<td>$1 \times 13$</td>
</tr>
<tr>
<td>B-Values from second RR (for up to 13 satellites)</td>
<td>Integer</td>
<td>$1 \times 13$</td>
</tr>
<tr>
<td>B-Values from third RR (for up to 13 satellites)</td>
<td>Integer</td>
<td>$1 \times 13$</td>
</tr>
<tr>
<td>B-Values from fourth RR (for up to 13 satellites)</td>
<td>Integer</td>
<td>$1 \times 13$</td>
</tr>
<tr>
<td>Issue of Data or IOD (for up to 13 satellites)</td>
<td>Integer</td>
<td>$1 \times 13$</td>
</tr>
<tr>
<td>Padding</td>
<td>Integer</td>
<td>27</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>222</td>
</tr>
</tbody>
</table>

where RR = Reference Receiver

4.3 Mobile User

The mobile user end of the data link needs to be equipped so as to receive the differential corrections broadcast by the base station. Figure 4.2 shows a functional block diagram of the processing at the mobile user [16].
The mobile user, similar to the ground station, uses the ephemeris and timing data from the GPS receivers to calculate the satellite positions. The pseudorange measurements from the mobile user receiver are smoothed using the carrier phase measurements (i.e. integrated Doppler measurements) of the user. A differential troposphere model as specified in [5] is implemented to correct the troposphere delay at the mobile user. The PRCs received from the ground reference are used to correct the smoothed measurements of the mobile user. These measurements are then corrected for satellite clock error. This corrected pseudorange is used to compute an accurate position solution for the mobile user.

This position solution, the carrier phase measurements and the measured pseudoranges of the mobile user are transmitted back to the ground reference using the

Figure 4.2  Bi-directional DGPS Mobile User Block Diagram
ethernet to wireless bridge, as remote positioning data for integrity check and remote/control operations by the ground reference.

The data communication between the computer at the ground reference computer and the wireless bridge uses the ethernet interface over UDP/IP. Similarly, the data communication between the computer and the wireless bridge at the mobile user uses the ethernet interface over UDP/IP.

4.3.1 Data downlink message format

Table 4.2 shows the parameters transmitted from the mobile user to the ground station and the format and size of each parameter [7].

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Format</th>
<th>Number of Bytes</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPS Time</td>
<td>Double</td>
<td>8</td>
</tr>
<tr>
<td>Differentially corrected East Coordinate</td>
<td>Float</td>
<td>4</td>
</tr>
<tr>
<td>Differentially corrected North Coordinate</td>
<td>Float</td>
<td>4</td>
</tr>
<tr>
<td>Differentially corrected Up Coordinate</td>
<td>Float</td>
<td>4</td>
</tr>
<tr>
<td>Data Checksum</td>
<td>Integer</td>
<td>4</td>
</tr>
<tr>
<td>SV Number (for up to 13 satellites)</td>
<td>Integer</td>
<td>1×13</td>
</tr>
<tr>
<td>Pseudorange Measurements (for up to 13 satellites)</td>
<td>Float</td>
<td>4×13</td>
</tr>
<tr>
<td>Integrated Doppler (for up to 13 satellites)</td>
<td>Float</td>
<td>4×13</td>
</tr>
<tr>
<td>Padding</td>
<td>Integer</td>
<td>27</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>228</td>
</tr>
</tbody>
</table>
The mobile user transmits its position, pseudorange measurements and carrier phase observables to the ground reference. Although the data is 228 bytes, it could be reduced to 222 bytes and fit into one, yet unidentified Message Type.

4.4 Hardware Set Up

For the demonstration set up in the laboratory, the hardware involved for both the mobile user and the reference station is very similar. The GPS signals for both the users were derived from a single GPS antenna located on the roof of Stocker Centre at Ohio University. The antenna used is a Sensor Systems S67-575-14 dual-frequency L1/L2 passive antenna followed by a low noise JCA amplifier which was fed via a separate +12V power supply. A power splitter was used to feed GPS signals to the two receivers from the same antenna. Since the same antenna was used to for both the reference and mobile user ends, it is expected that the position of the mobile user with respect to the reference station will be the same, which in itself will be a measure of the accuracy of the solution. Figure 4.4 illustrates the set up of the system as it was used for the purpose of the demonstration.
As illustrated in Figure 4.3, the hardware at both the ends comprises of the following essential elements:

- **GPS Receiver**: The GPS receiver used was a NovAtel Beeline receiver at the ground reference and the mobile user. The Beeline receiver has two RF input ports that act essentially like two separate receivers. It should be noted that the calculation of B-values requires at least two reference receivers or ground-based stations [5]. The NovAtel Beeline receiver was used for simplicity at the ground
reference to simulate a two reference receiver environment. Ideally, a high performance ground reference station would have two spatially separated sites to decorrelate common multipath errors for added integrity. Here, a single NovAtel Beeline was used at the reference station to reduce complexity of the system. The receiver treats both inputs as from two GPS antennas but the ephemeris is decoded only from the first RF input of the NovAtel Beeline GPS receivers.

- **Personal Computer:** The computer used to process the measurements and perform the computations at the ground reference is a Gateway E-4200. It uses a Intel Pentium II processor clocking a frequency of 400 MHz. The QNX 4.25 real-time operating system was used to provide a reliable platform for the computations. The serial port of the computer was used to interface to the GPS receiver and the ethernet card was used to transmit and receive data over the data link.

  At the mobile user a Gateway 2000 computer with an Intel Pentium II processor operating at 400 MHz was used. Similar to the reference station the GPS receiver was interfaced to the computer through the serial port and the ethernet port was used to transmit/receive data over the data link.

- **Wireless Ethernet Bridge:** The wireless transceiver used for the research was a D-Link DWL 810+ ethernet to wireless bridge. This bridge essentially takes the data from the ethernet port of the computer and transmits it wirelessly conforming to the 802.11b standards, achieving maximum data rate of 22 Mbps. It operates in the frequency range of 2.4 GHz to 2.4385 GHz and
incorporates the DSSS modulation technology. It uses an input power of DC 5V, 2.5A. The antenna is a detachable 1.0 dB gain antenna with a reverse SMA connector.

The same brand of the wireless ethernet bridge was used at the mobile user and the reference station to transmit and receive data. The reason to go for the D-Link 810+ was driven by the need for an operating system independent ethernet bridge to provide for wireless transmission and reception of data at each end of the data link operating in the 2.4 GHz industrial, scientific, and medical (ISM) band. This bridge was used due to non-availability ethernet radios with a 10 nautical miles range and operating at the 5 GHz frequency band.

4.5 Software Requirements

The software involved for processing the GPS receiver data, perform computations and to use IP for transmission/reception of the data was implemented by incorporating changes in the prototype LAAS software developed and tested at the AEC at Ohio University. The versions of the software for the ground and reference station are similar in some functions but with differences in computations of the corrections and position solution [7].

The version of the software developed for this research did use bi-directional DGPS functions which use serial communications to transmit differential correction data and receive remote positioning data. The major changes made to the existing software
were to route the data through the ethernet port over IP. This section describes the basic functionality of the software used for the bi-directional operation of the data link over IP.

4.5.1 Ground reference station

General Overview: In the previous version of the prototype software for the ground reference station implemented in this research, ground.c is the main program [7]. The novatel.c routine decodes the ephemeris data acquired by the GPS receiver. The coordinates for the satellite positions are calculated in svcalc.c. The g_init.c routine initializes the ground reference parameters including the reference position coordinate. The differential corrections are calculated in intmon.c. The serial.c routine handles the communication of the data read in from the serial ports. The data processing involved in handling the bi-directional link is performed in the program serrem.c. This routine incorporates functions for reading in data from the serial port, verifying the data transfer and setting up the data structures for further processing. Since this routine involves calculating the position for the remote user, it implements a differential troposphere correction model identical to that implemented by the software by the mobile user. The computations for the calculation of the position of the remote user are done in calrem.c. The laasterm.c routine controls all the functions involved in displaying the data on the screen and their updates to reflect real-time calculations. The Freewave_transmit() function in the software handles the transmission of differential correction data to the Freewave RF radios operating at 900 MHz [7].
IP Implementation: To implement the data communication over IP, it was necessary to first initialize the socket manager in QNX 4.25. The operating system uses the BSD4.3 TCP/IP stack which has become an industry standard. The initialization of the socket manager is independent of the ground software and essentially allows for setting up and communicating through sockets using UDP and IP. With the socket manager initialized, the network interface parameters were initialized for communication via ethernet, specifying the interface address and specifying the netmask. Specifying the netmask allows for reserving part of the address space to divide sub-networks in the network. With the network interface parameters initialized the software can be executed for communication over IP.

In the software, the first step involved was the opening and initialization of the sockets. Opening a socket returns a file descriptor which can be used for later sending and receiving functions. The socket was initialized for using the UDP over IP and was set to non-blocking. The reason to set the socket to non-blocking was to avoid the receiving function from blocking further operation if no data was available at the ethernet interface. After opening the socket, the bind () function was used to bind it to a local port. This port number defines the interface for the application program, so that when a datagram arrives at the ethernet interface it gets directed to the correct process requesting the use of the datagram. The Freewave_transmit () function was rewritten to transmit the differential correction via the sockets using UDP over IP.

Also the section of the code in ground.c was rewritten so the data could be read through the socket. The function that handles the receiving of data through the socket by default blocks further operations until it receives data at the interface. This is an
undesirable situation because it hangs the software until further data is received. Ideally, the software should continue processing the receiver data even if it hasn’t received remote positioning data. Hence the socket was set to non-blocking state on initialization.

4.5.2 Mobile User

General Overview: In the air software of the version of the prototype software used in this research, ac.c is the main program [7]. The routines for reading in data from the GPS receiver, processing it and calculating the satellite positions are similar to those in the ground software. The calpos.c routine implements functions to calculate the stand alone position for the mobile user. The functions that implement the troposphere correction are included in atmcor.c. The differentially corrected position is computed in Calculalte_diff_position (). The rem_transmit () function in the software transmits the remote positioning data to the Freewave radios at the mobile user over the serial port, which in turn sends the data to ground reference station [7].

Similar to the ground software, all the display routines to display the data on the screen are controlled by laasterm.c [7].

IP Implementation: Since the mobile user also uses QNX 4.25, the steps involved in the initialization of the socket manager and configuring the ethernet interface are identical to those in the ground reference station. In the software, the difference is that the mobile user receives data from the ground reference first before transmitting the remote data. This is due to the fact that it needs to receive the differential corrections from the ground
reference before it can compute its accurate position and then generate and transmit the remote positioning data.

Originally, the data was read through the serial port which was controlled by the serial.c routine and its functions. The receiving functions were changed to read in the data from the ground reference through over ethernet via the socket. The rem_transmit() function was re-written to transmit data through the wireless transceiver via the ethernet interface.

4.6 Considerations for Choice of Transport Layer Protocol

The choice of the transport layer protocol is critical to the functioning of the system to meet the desired levels of performance requirements. Hence, it is expected that for different applications, this choice will be driven by various considerations specific to the priorities of the specific application. For the purpose of demonstration for this research, UDP was chosen as the transport layer protocol to be implemented for data communications over the data link. This section discusses the rationale behind this choice and considers the trade-offs in performance within the scope of implementing this proof of concept system.

- **Broadcast/Multicast ability:** One of the primary factors for choosing UDP over TCP was that UDP provides the ability to broadcast data within the network without queuing up connection requests from users. This enables the ground reference station to service a large number of users within the service volume of the application.
• **Flexibility:** TCP is a full featured protocol providing a range of services as discussed before. UDP does not provide such a features and this allows the programmer or designer of the system a flexibility to choose and implement the critical services of choice at the application layer. Reliability of data communications is one of the important considerations. TCP provides the reliability but at the cost of bringing in other features which may not be required by the system, thereby adding what may be unnecessary overheads. Data reliability can be assured through other techniques either at the application layer level or at the hardware level. For the purpose of this research, UDP was chosen for the implementation in a laboratory environment where network congestion was minimal. For a real world scenario, this choice will be driven by the network latency involved due to network traffic and the levels of performance required in terms of availability and integrity.

• **Latency:** UDP is a connectionless protocol hence the latency involved in establishment of a connection, as with TCP, is eliminated. Moreover, because of the ability to broadcast data, UDP does not require a user to “wait” for a connection from the ground station before it can receive data. TCP can be implemented for a multi-user environment but limits the number of users that the application at the ground reference station can support. Also, it would require “queuing” up of users at the ground station for connections and the possibility of time-outs in obtaining correction data for navigation is obviated by implementing UDP.
• **Simplicity:** This system was developed as a proof of concept to be developed on, for an enhanced data link. UDP was chosen to provide a simple interface for the application to IP at the network layer.

Hence the choice of UDP was driven primarily by the ability to broadcast data and to minimize risk due to latency. Also, the UDP provides the necessary interface for the application layer to communicate over the data link using IP. Since this system provides a basic foundation to build further enhancements, TCP may be implemented for applications with more specific and stringent requirements.

### 4.7 Real-Time and Practical Considerations

The bi-directional data link over IP has been demonstrated in a very simplified setup to establish proof of concept for the proposed architecture. However, for practical applications in real-time additional factors come into effect that need to be addressed. Some of the factors associated with such an implementation are discussed below:

• **Software Platform:** One of the essential components of the system is the processing unit that takes in the receiver measurements and processes it to output the required data. For this purpose a computer with a real-time operating system was used. The key difference between general-computing operating systems and real-time operating systems is the need for “deterministic” timing behavior in the real-time operating systems. Formally, "deterministic" timing means that operating system services consume only known and expected amounts of time. Random elements in service times
could cause random delays in application software and could then make the application randomly miss real-time deadlines – a scenario clearly unacceptable for a real-time system as in the case of the proposed system.

- **Latency:** Latency is one of the most important issues that need to be accounted for. The differential corrections broadcast to the users are subject to temporal decorrelation and hence need to be applied by the user with minimum latency. For the demonstration, one user was used within 100 m of the base station in a peer to peer fashion. The average round trip transit time of a packet from the base station to the mobile user was reported to be 0 ms by the operating system. This is expected due to the simplified nature of the set up. However, for practical applications with increased number of users and a range of about 10 nmi there will be some latency involved. For this purpose, the correction terms from the base station and the location coordinates from the mobile user are time tagged and when possible the solution is propagated in time by the reference station. For a practical scenario it will be necessary to ensure the update rates are sufficient to ensure availability, integrity, and continuity of the data broadcast to the user. System latency will affect these performance parameters and would eventually impact the “time-to-alarm” for the warning message that is generated by the ground reference station for the mobile user, if there is a discrepancy in the solution provided by the user and that calculated by the ground station.

- **Packet Loss:** As discussed earlier, one of major differences in implementing UDP and TCP is that TCP ensures transmission and reception of the data
while UDP does not. UDP, the connectionless protocol does not implement some of the sophisticated features for data reliability of TCP, which however does allow some of the ease of using UDP over TCP. As a result, it is necessary to handle the issue of data integrity at the application level if using UDP for data transfer. The checksum error check is implemented in the code for the bi-directional DGPS, to ensure integrity of the data received at both the ends. The data sent over the data link conforms to a specific format to enable the mobile user to retrieve the parameters appropriately. As a result, loss of re-ordering of packets will cause a loss of the required data. For a practical implementation it will be necessary to implement checking for and resending lost packets at the software application of hardware level in case of packet loss, to ensure availability of the system. This was not implemented in the demonstration for the initial set up.

- **Physical Range**: The operating range for the data link is expected to be for an area sufficient to provide service for an airport. The physical range will be determined by range capability of the data link ethernet radio. The range of the system will impact the accuracy of the differential corrections, latency and hence the overall system performance for the defined service volume.

- **Data Rate**: The data rate requirement is determined by the data to be transmitted and more importantly the update rate. In this case the ground reference processor receives updates once every second. The data rate requirements will change as additional features are implemented, specific to the application.
5 RESULTS AND CONCLUSION

The demonstration for the proof of concept system was performed in the laboratory in Stocker Centre at Ohio University. The following discussion provides some figures of the screen of the real-time system and also some plots for the data collected on March 10, 2005 for GPS time 414040 to 417202 seconds. Since both the user and reference receiver were located at the same site (i.e., used the same GPS antenna), the errors are expected to be sub-meter level. This is because the use of the same antenna ensures that all the system errors (i.e., satellite clock, ephemeris errors, atmospheric delays, and multipath) are removed and only the receiver specific and errors remain. As a result, the East, North and Up coordinate of the user with respect to the reference may be considered to be the errors in the solution.

5.1 Screen Shot for Remote Positioning Data Received from the User.

Figure 5.1 shows a screen shot that displays in real-time the data received from the mobile user and the position of the remote user as calculated by the ground reference.
Figure 5.1 Remote Positioning data at ground reference

As the screen illustrates, the position of the user as calculated by the ground is displayed along with the position as calculated and transmitted by the user. The update from the user is one second late as can be displayed in the difference of the ground reference time and the time transmitted by the user.

5.2 Errors in position coordinates if user

Figure 5.2 shows a plot for the East coordinate of user with respect to the ground reference for the time interval specified.
The error in the East coordinate varies between -0.65 m to 0.3 m. The mean of the error is -0.0045 m and the standard deviation is 0.0277 m. For the purpose of the demonstration the GPS signals to both the reference and user receivers were from the same antenna. As a result the bulk of the errors may be attributed to be receiver specific errors.

Figure 5.3 provides the error in the North coordinate of user with respect to the reference receiver.
The error in the North coordinate is also within the sub-meter level. The mean of the error is -0.1246 m and the standard deviation is 0.0414 m. This can be considered to be accurate.

Figure 5.4 provides a plot for the error in vertical positioning of the user with respect to the reference.
The error in the vertical positioning is the main component of error, as can be seen from the plot where in the maximum error is almost 1 m. The mean of the error is 0.1130 m and the standard deviation id 0.1381 m.

The overall accuracy of the system is acceptable within the limits of the technology used for implementation and are considered reasonable for a pseudorange based positioning system [2].
5.3 Conclusion

The set up does successfully demonstrate a full-duplex bi-directional DGPS architecture using communication between the user and reference over a high frequency data link over IP, in real-time. The ground reference station uplinked the differential correction data to the mobile user. The mobile user applied those corrections to its measurements to obtain a high performance user solution. This was discussed by way of the error in the position of the mobile user as obtained by the ground reference station. The mobile user also communicated to the ground its measurements to the ground reference station to allow the ground station to compute the position of the remote user to enable an integrity check. The remote positioning ability of the ground reference station was demonstrated by way of the user position data received by the ground reference station. The data communication was established via the ethernet interface over IP using a UDP broadcast format.

This lays the ground work to further develop the data link for additional enhanced features in a move towards a more “robust fully functional real-time bi-directional DGPS architecture using IP” in the 5 GHz frequency band.
6 RECOMMENDATIONS FOR FUTURE WORK

The bi-directional DGPS architecture was successfully demonstrated for data communication via the ethernet interface over IP using a wireless network, though in a laboratory environment. This can be further expanded to include additional features and tested for a more operational environment. The motivation was to exploit the potential of the data link at higher frequency and data rate to integrate communication, navigation and surveillance data into a single high capability data link using IP. Apart from the integration of the data, the challenges would involve maintaining high accuracy, integrity, continuity, and availability of the system. Also, this could be expanded from being a GPS-only link to include other navigation sensors. The following discussion outlines some of the key features that may be implemented towards the objective.

- The software can be migrated from QNX 4.25 to the QNX 6.3 platform or higher. The motivation to do this is that the later version of QNX incorporates an enhanced TCP/IP stack. Also, it is designed to operate on multi-processor platforms and offers better portability. This involve completely over hauling the display routines in the software and some of the serial and TCP/IP routines to be compatible with the new generic compiler in the operating system.

- For applications in civil aviation, the software must be compliant to software certification standards as outline by the FAA. The software could be migrated to a different operating system e.g. LynxOS and VxWorks which are certifiable to level-A certification requirements as outlined by RTCA in [20], to ease the
process of developing the software for the data link, certifiable for use in civil aviation.

- In the present demonstration, wireless transceivers at 2.4 GHz were used for a short range in the laboratory. The earlier version, using serial communications operated at 900 MHz over a range of approximately 10 nmi. Ethernet radios can be used to perform the same function, but with an extended range. This could be demonstrated at 2.4 GHz with the objective for a potential deployment at 5 GHz. The high data rates supported by the interface would reduce the latency to within milliseconds.

- Presently, the only navigation related data being transmitted is from GPS. To enhance the availability, robustness and functionality of the system, navigation data from other navigation systems e.g., an inertial navigation system unit may be added. This would involve modifying the structure of data being transmitted to include and identify the type of navigation data being used.

- The present software uses only position information for remote positioning applications. Velocity information may be implemented in the software which could then be down linked to the ground reference for advanced control/track ability.

- For the purpose of this research a UDP broadcast format was used for data communication. However, TCP may be used for specific applications. The same system can be developed using TCP to perform a comparative analysis between TCP and UDP based on performance parameters of the system e.g., number of users, reliability, continuity, integrity, network congestion, etc.
• To utilize the message handling capabilities of IP, other web related applications may be integrated into the message format uplinked to the mobile user e.g., weather imagery, email, voice over IP, etc.

• Also, the message format can be modified to partition mission critical data and other functions based on the specific application. This would require defining the criticality of the data being transmitted/received and other non-critical data involved in the system.
7 REFERENCES


