

Architecture



The Canadian Geospatial Data Infrastructure

Architecture Description Version 2.0

**GeoConnections
2005**



Preface

This document is an update and revision of the original Canadian Geospatial Data Infrastructure (CGDI) Architecture Description that was created in 2001. Since that time, there has been significant progress made in the development of the CGDI. The first five years have resulted in the development of an infrastructure, data, and services, as well as establishing key relationships and partnerships.

The updates to this document are based on extensive stakeholder consultations. Rich contributions were received through a multi-faceted approach that included both traditional and on-line consultation techniques. This has allowed for the preparation and presentation of an Architecture document that encompasses the reflections and considerations of those with a keen interest in the future of the Canadian Geospatial Data Infrastructure.

Intended Audience

This document is intended to introduce the Canadian Geospatial Data Infrastructure (CGDI) architecture to data providers, service providers, and application developers so that they can gain an overall understanding of the technical aspects of the CGDI.

Scope

This document provides a high-level overview of the architecture of the CGDI. It describes the range of services that comprise the CGDI, and provides contextual and reference information for the more technical standards and specifications documents that detail the individual services and other specific components of the architecture. It also describes the underlying architecture that is common to all services.

Table of Contents

PREFACE	I
INTENDED AUDIENCE	I
SCOPE.....	I
TABLE OF CONTENTS	II
EXECUTIVE SUMMARY	IV
1 INTRODUCTION.....	1
1.1 <i>ALIGNMENT WITH A FEDERATED ARCHITECTURE.....</i>	<i>1</i>
1.2 <i>ARCHITECTURE FRAMEWORK.....</i>	<i>2</i>
2 OVERVIEW	4
2.1 <i>A GEOSPATIAL DATA INFRASTRUCTURE.....</i>	<i>4</i>
2.2 <i>THE CGDI ARCHITECTURE</i>	<i>4</i>
2.2.1 <i>The Business Rationale.....</i>	<i>5</i>
2.3 <i>CHARACTERISTICS OF THE CGDI.....</i>	<i>5</i>
2.4 <i>CONCEPTUAL ARCHITECTURE ELEMENTS.....</i>	<i>6</i>
2.4.1 <i>Data.....</i>	<i>7</i>
2.4.2 <i>Services</i>	<i>8</i>
2.4.3 <i>Applications</i>	<i>9</i>
2.4.4 <i>Users</i>	<i>9</i>
3 ARCHITECTURAL COMPONENTS.....	11
3.1 <i>OBJECTS.....</i>	<i>11</i>
3.2 <i>OPEN STANDARDS AND SPECIFICATIONS.....</i>	<i>11</i>
3.2.1 <i>The Open Geospatial Consortium</i>	<i>12</i>
3.2.2 <i>The International Organization for Standardization (ISO)</i>	<i>12</i>
3.3 <i>CGDI-ENDORSED STANDARDS</i>	<i>13</i>
3.4 <i>REGISTRIES.....</i>	<i>14</i>
3.5 <i>METADATA.....</i>	<i>15</i>
3.5.1 <i>The Importance of Metadata</i>	<i>16</i>
3.5.2 <i>Metadata Standards.....</i>	<i>16</i>
3.6 <i>SECURITY AND AUTHENTICATION.....</i>	<i>16</i>
3.7 <i>ROBUSTNESS AND REDUNDANCY.....</i>	<i>17</i>
4 ARCHITECTURAL REQUIREMENTS.....	18
4.1 <i>INTEROPERABILITY.....</i>	<i>19</i>
5 DEVELOPMENT STRATEGY	20
5.1 <i>OVERALL APPROACH.....</i>	<i>20</i>
5.2 <i>ARCHITECTURE PERSPECTIVES</i>	<i>20</i>
5.2.1 <i>Operational Architecture.....</i>	<i>22</i>
5.2.2 <i>Technical Architecture</i>	<i>22</i>

5.2.3	<i>Systems Architecture</i>	23
5.3	<i>USING THE CGI</i>	23
GLOSSARY OF TERMS		24
REFERENCES		28
APPENDIX: AN ARCHITECTURE BASED ON WEB SERVICES		29

Executive Summary

Since March of 2000, important strides have been made towards developing the CGDI (Canadian Geospatial Data Infrastructure) into a distributed network based on open standards. The premise that distributed networked access makes data more readily available is the driving force behind the CGDI. The architecture is designed to enable the implementation of systems to support **service providers, data providers** and **application developers**, using interoperable and re-usable components. The ability to interconnect computers nationally and globally to access geospatial information is a powerful proposition. The commitment for CGDI to provide universal access has placed it at the forefront of the global geospatial community.

The CGDI has the potential to be the starting point for a business revolution. A networked data infrastructure based on open standards creates greater equality amongst users looking to leverage geospatial data to serve their business and operational needs. As geospatial data becomes more integrated in everyday society, the importance of the CGDI and the infrastructure's path to self-sustainability become clearer.

Core elements of the CGDI

The CGDI architecture is made up of four key elements; **data, services, applications** and **users**. From a conceptual perspective, the infrastructure is the highway traveled by all players participating in the CGDI. Data is at the core of the infrastructure, with services and applications developed to access and contribute data and produce information for consumption and decision-making.

The underpinning of a distributed architecture is the means by which independent systems communicate and collaborate with one another. Web services based on open standards provide the basis for interaction across the Internet and allow users to contribute, access, and exchange geospatial data. The future success of the infrastructure requires a continued commitment to the on-going development and adoption of CGDI-endorsed standards.

Client applications provide a doorway to geospatial data using one or more services. Applications address a span of user needs, from generic end-user requirements to addressing the needs of data suppliers who want to contribute to the CGDI. A rich set of client software applications based on core CGDI components, interfaces and services, built by CGDI developers and suppliers, will ultimately deliver most of the anticipated benefits to Canadians.

Overview of the Architecture

The CGDI is based on open and interoperable standards and specifications for operational transactions and information exchange. The CGDI focuses on identifying and promoting existing open and interoperable software interfaces, and developing new ones as required through the help of existing public standards organizations. The Open Geospatial Consortium (OGC) and the International Organization for Standardization (ISO) are two of the leading standards bodies that play an important role in the development of the CGDI.

Standards have been endorsed for data visualization (Web Map Server) and the companion presentation (Styled Layer Descriptor) and storage (Web Map Context) of geospatial data.

Data manipulation is defined by the Web Feature Service, and data querying is handled by the Filter Encoding specification.

Other CGDI-endorsed standards include: the Geography Markup Language for coding the transport and storage of geographic information; the Geodata Discovery Service that searches geospatial data resource registries; and Metadata for Geodata that provides a standard way of describing geospatial data.

The CGDI architecture continues to evolve, and a number of other specifications are making their way from discussion paper through to recommendation paper, and endorsed specification.

Operational View of the Architecture

The CGDI will be implemented as a network of cooperating physical servers, providing services, and data via these services, in such a way that an application can be developed that makes use of these services. This permits savings in development time, operating costs, and allows for uniformity of data.

CGDI services are provided by a broad spectrum of suppliers. There is no real "centre" to the CGDI. As the operational architecture develops, it further defines who implements the services described in the technical architecture and provides guidance as to how organizations go about implementing services in ways that foster interoperability. Other key operational issues for consideration include the degree of "looseness" of the infrastructure in terms of hierarchies of servers and registration authorities, as well as the bindings or linkages that make systems part of the CGDI.

Security and authentication will also play a major role in geospatial data processing and operations. The need for sound security and authentication mechanisms increases when sharing information in an open and interoperable fashion. The CGDI must provide for the authentication and authorization of clients when it is required. The goal is to provide harmonization of identification, certification, authorization and authentication across jurisdictions, for both private- and public-sector users.

The Future of the CGDI

The CGDI has achieved a level of technical maturity that allows for further exploration of the issues and opportunities that will shape its future form. Great strides have been made over the past five years, however, there is a new need for refinement and for greater participation by partners, to ensure that there is a critical mass of geospatial data, services, and applications. Without continued investment and contributions from all the partners it is unlikely that the CGDI will be able to provide current and relevant geospatial information to assist in policy development and decision making by governments, industry, and the not-for-profit sector.

Leadership and technology have been the focus and strength of the CGDI to this point in time. Looking ahead, a shift in focus to a service-oriented approach is required with a continued need for strong leadership. With a renewed emphasis on leadership, and a focus on user-driven service delivery, the technological foundation of the CGDI will bloom.

1 Introduction

The Canadian Geospatial Data Infrastructure (CGDI) makes Canada's geospatial databases accessible on the Internet through the use of technology, standards, access systems and protocols. In more concise terms, the CGDI is made up of *data, services and applications* that enable *users* to make use of geospatial information on-line. As is the case for a majority of data infrastructures, the Internet has had a profound effect on how the CGDI is being built and used. Networked access makes data more readily and freely available. This premise is the driving force for building an Internet-based, distributed geospatial data infrastructure.

In 1999 the GeoConnections program was created to oversee the development of the CGDI. Since its inception, the work of GeoConnections has centred on the management and coordination of the partners contributing to the infrastructure's technical foundation. The CGDI is an open information technology infrastructure that is based upon publicly available standards and specifications. The architecture is designed to enable the implementation of systems to support **service providers, data providers and application developers**, using interoperable and re-usable components. This is achieved largely through the use of internationally recognized specifications. The ability to connect computers nationally and globally to access geospatial data is a powerful proposition. Canada's commitment to the CGDI has placed it at the forefront of the global geospatial community.

In the past few years, Web services have emerged as a leading mechanism in delivering Internet-based functionality and are at the core of the CGDI. Since 1999, GeoConnections' development efforts have focused on the implementation of standards-based Web services. In this respect, the CGDI has successfully leveraged existing and emerging technologies and standards to adopt a general-purpose service-based platform. This approach allows services to be efficiently deployed and invoked within an infrastructure where geospatial data can be shared and accessed.

The CGDI has achieved a level of technical maturity that will enable an exploration of the issues and opportunities that will shape its future form. This document discusses the architectural premise from which the CGDI was born, and aims to describe the evolution in technology and user requirements over the past five years and what is projected for the future.

1.1 Alignment with a Federated Architecture

In the *Government of Canada's Federated Architecture, Iteration One* there are five distinct business drivers derived from the analysis of the government's operating environment. Although these drivers are specific to a federated architecture, they are also key success factors for the Canadian Geospatial Data Infrastructure.

1. Accessibility to Government Programs

Government information and services will be accessible regardless of location, time, method of access and group (e.g. language, culture, age, and ability).

2. Security, Authentication and Authorization

Access to information and services will be authenticated to the degree required for specific information and services. Information will be protected to the level required both internally and externally. The availability and integrity of on-line information and services will be assured.

3. Government as an Integrated Enterprise

The Government of Canada will provide a single point of access for coherent government information and services (i.e. "no wrong door") spanning departments and other levels of government.

4. Client-centric Service Delivery

Government information and services will quickly respond to the client's changing expectations (e.g. clustering, complete service fulfillment, speed, reliability, responsiveness and ease of use, timeliness, customization and privacy). A comprehensive Help service will be provided for all methods of access.

5. Efficient and Effective Service and Information Delivery

Delivery costs will be reduced and the quality of services and information will be improved.

1.2 Architecture Framework

The Canadian Geospatial Data Infrastructure (CGDI) strives to build an ideal architecture framework based on the Canadian Treasury Board Federated Architecture principles, and the ISO Reference Model for Open Distributed Processing (RM-ODP).

The RM-ODP is an ISO/IEC standard that defines views of a distributed system of systems. Those views encompass business model (i.e. Enterprise viewpoint), information content and system behavior (i.e. Information viewpoint), components, interfaces and constraints (i.e. Computational viewpoint), infrastructure and mechanisms for component distribution (i.e. Engineering viewpoint), and implementation and deployment environment using technologies, standards and products of the day (i.e. Technology viewpoint). The RM-ODP describes architecture in terms of the viewpoints described in the following table:

RM-ODP Viewpoint	Areas of concern
<u>Enterprise viewpoint</u> Articulates a "business model" that should be understandable by all stakeholders; focuses on purpose, scope, operational objectives, policies, enterprise objects, etc.	Purpose and scope Policies Responsibilities Business process along with use cases
<u>Information viewpoint</u> Focuses on information content and system behaviour (i.e. data models, semantics, scheme)	Information processing semantics Information of the system
<u>Computational viewpoint</u> Captures components, interfaces, interactions and constraints without regard to distribution	Functional decomposition Interfaces Operations Binding rules
<u>Engineering viewpoint</u> Describes infrastructure and mechanisms for components distribution, distribution transparency	Infrastructure required to support distribution

and constraints, and binding and interaction	
<u>Technology viewpoint</u> Defines implementation and deployment environment using technologies, standards and products of the day – best of breed	Choice and suitability of technology to support system distribution

Source: GOS Implementation Architecture, Open Geospatial Consortium® document, 2003

The RM-ODP framework recognizes that an architecture is complex and cannot be described in a single representation; hence the approach to describe an architecture from a number of "viewpoints". The five viewpoints chosen by the RM-ODP (enterprise, information, computational, engineering, and technology) indicate the wide breadth of scope that architectures must consider to lead to successful results. The viewpoints cover a range of issues from organizational through to selection of technologies, in a carefully organized and hierarchical manner.

The CGDI architecture is essentially a blueprint that describes the structure of components, their relationships, and the principles and guidelines governing their design and evolution over time. It endorses the RM-ODP, and describes elements that can be used to support the information, engineering and computational RM-ODP viewpoints. The complete application of RM-ODP is left as the responsibility of agencies collaborating in the CGDI.

2 Overview

2.1 A Geospatial Data Infrastructure

A distributed infrastructure that provides on-line access to geospatial data offers a significant opportunity for a wide range of stakeholder groups. These stakeholders range from groups who will use the Canadian Geospatial Data Infrastructure (CGDI) to produce goods and services to citizens who will benefit through their availability.

Established networks, such as the Public Switched Telephone Network, provide great benefits to society, and the CGDI offers the same type of potential. However, before we derive the myriad of societal benefits possible, a natural evolution must take place. A networked geospatial data infrastructure is the starting point. With an established CGDI, the impact of geospatial data will continue to grow, penetrating almost every facet of daily life.

The development of the CGDI will focus on the following five key thrusts:

- 1. Data Access:** developing the technical components so that Canadians can access geospatial data and services over the Internet, working across federal agencies, provincial and territorial governments, the private sector and the academic community.
- 2. Framework Data:** providing commonly used geospatial data sets required for analysis, allowing users and agencies to add value, develop applications and create more detailed geospatial databases.
- 3. Geospatial Standards:** using standards that simplify access, improve data quality and integration, and encourage commercial development of internationally compliant software technology.
- 4. Partnerships:** creating collaborative agreements and projects among the federal government, provincial and territorial governments, industry and academia to make geospatial data more available by sharing new developments and capitalizing on emerging technologies.
- 5. Supportive Policies:** simplifying government policies, and harmonizing the access and use of geospatial data.

2.2 The CGDI Architecture

The CGDI is based on a single data infrastructure. From an architectural perspective, the notion of a single infrastructure is a key design principle. There are numerous benefits to a single data infrastructure. From a cost-saving perspective, a properly designed and deployed architecture that is sufficiently robust, to meet the needs of its users will avoid the costs associated with parallel infrastructures. In addition to cost savings, one CGDI will provide the most comprehensive single point of access to geospatial information for Canada.

Implementing a single architecture is a significant undertaking. The success in its design and development will ultimately determine its future sustainability. Cooperation and

collaboration among contributors to the CGDI is the foundation for present and continued success.

2.2.1 The Business Rationale

The CGDI is intended to provide access to geospatial data in a distributed and networked fashion. For the architecture to be successful, however, it must effectively serve the business needs of its users.

The market potential for geospatial data is only limited by the barriers that exist around the information. Access to accurate, timely and affordable geospatial data will provide economic opportunities across all private and public sectors.

The CGDI has the potential to address a wide range of decision-making requirements and foster new business opportunities. An open networked data infrastructure will create greater equality among end-users looking to leverage geospatial data to serve a business need that was formerly not an alternative. As geospatial data becomes practically available and consequently more integrated in everyday society, the importance of the CGDI and the path to its self-sustainability become clearer.

Despite the development efforts that have occurred to date, in relative terms, the CGDI is still at a very early stage of its life cycle. As we move forward, momentum from the community of users will propel the development of the CGDI down the maturity curve. For this momentum to be realized, though, the user capacity must reach a certain level of maturity. Therefore, building user capacity is one of the key success factors for the future of the CGDI.

2.3 *Characteristics of the CGDI*

To support the vision of a geospatial data infrastructure that enables access to authoritative and comprehensive sources of geospatial information, the CGDI will:

- Enable universal access to any kind of geospatial information, anywhere, anytime;
- Enable applications to discover and access remote online information through a distributed network;
- Enable integration of disparate geospatial information to provide seamless views;
- Enable the seamless chaining of applications, data and services or combinations of these;
- Provide geospatial update and exchange capabilities, enabling collaborative activities;
- Promote the sharing of geospatial semantics to make integration of information easier;
- Enable wide-scale interoperability by adhering to common and open information standards and specifications;
- Facilitate the development of effective partnerships with regional and sector-specific spatial data infrastructures (SDIs)
- Facilitate international linkages with other national SDIs to form the Global Spatial Data Infrastructure (GSDI);

- Protect against system failure through implementation of a robust and reliable infrastructure with appropriate redundancy; and
- Provide for the authentication and authorization of clients when required. Access to information and services will be authenticated to the degree required by the specific information and services.

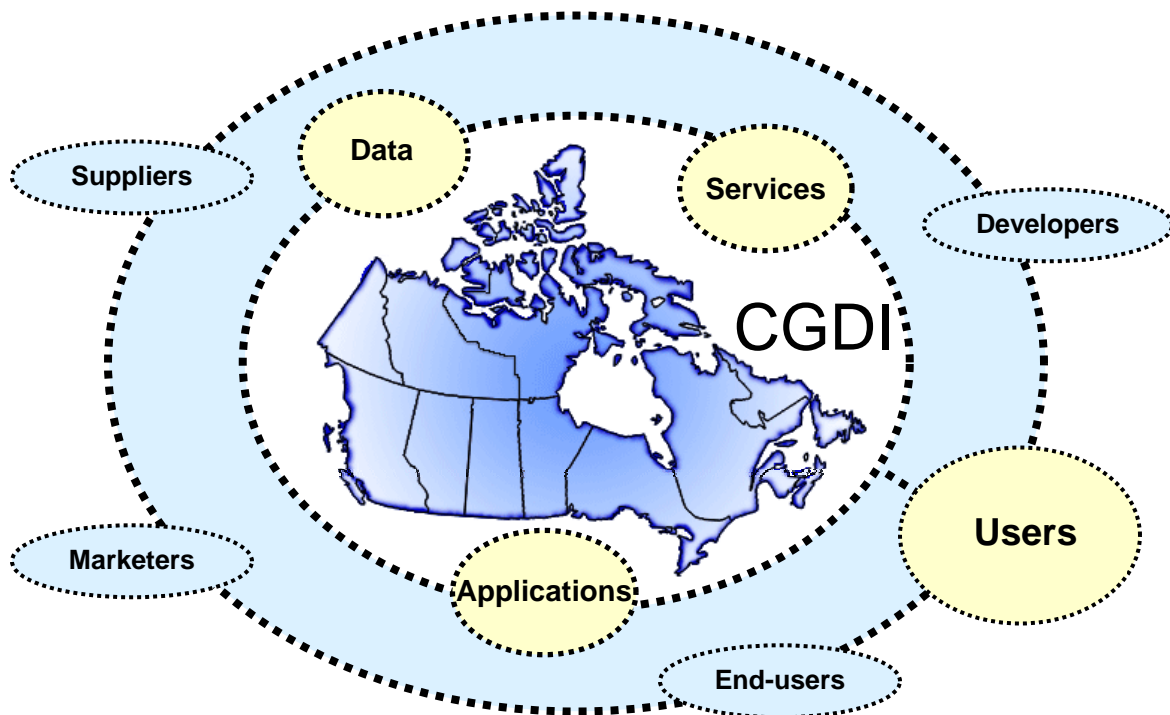
2.4 Conceptual Architecture Elements

The CGDI architecture is made up of four key elements: **data, services, applications** and **users**. From a conceptual perspective, the architecture consists of data and service providers and consumers who utilize applications to access geospatial information.

The data infrastructure is the highway traveled by all players participating in the CGDI initiative. It builds upon federal, provincial, territorial, municipal and industrial geospatial capabilities with the goal of presenting a collaborative framework for global access to geospatial information.

The following figure illustrates the four key elements and their relationship to the CGDI. It further delineates the users into four distinct categories: **suppliers, developers, marketers** and **end-users**.

Figure 1 – CGDI Conceptual Architecture



Data – Data is at the core of the CGDI. Timely and secure access to accurate and up-to-date data is the key to the CGDI's success.

Services – The CGDI is based on open Web services that provide access to geospatial data.

Applications – Applications use data from Web services to provide users with the ability to produce and analyze geospatial information to make informed decisions.

Users – Users are the consumers of geospatial data. Users fall into four categories:

- 1 **Suppliers:** Providers of geospatial data and Web services.
- 2 **Developers:** Creators of applications for other user groups to facilitate interaction with the CGDI.
- 3 **Marketers:** Sell/support geospatial applications for end-users.
- 4 **End Users:** Consumers of geospatial data.

2.4.1 Data

“Geospatial data” is defined as data that is registered to a specific point on the Earth's surface. It answers the question “where on Earth is it?” Geospatial data includes details about characteristics, relationships to other things or ideas, and the dimension of time as it relates to all of these.

Building on the concept of geospatial data, “data sets” are defined as collections of information that can be mapped, or *located*, so that people can use them to research, analyze and plan. Geospatial data sets are much more than maps: they allow users to mix and combine different types of data to produce more relevant and valuable results.

The CGDI provides access to thematic data sets that describe the characteristics of geospatial features, providing information on specific topics, such as rainfall, geology, or population. Thematic data attributes are geospatially referenced so they can be tied to locations on the Earth and be used in applications.

The Atlas of Canada provides a collection of maps and related information about Canada. It offers thematic data, such as population, via CGDI Web services. The Discovery Portal is the best place to discover thematic data. Searching on a particular subject (e.g., vegetation or snowfall) yields thematic results. Most thematic data sets can be distributed via the CGDI and it is the sharing of these sets that will enable the most powerful CGDI applications.

Framework Data

Framework data is intended to reduce duplication and improve the interoperability of data sets.

CGDI framework data is a set of continuous and fully integrated geospatial data that provide context and reference information for the country. Framework data is expected to be widely used and generally applicable, either underpinning or enabling most geospatial applications. Although continuity and cross-linkages are specific goals, information in this form is not immediately available, and most framework data will have to evolve in order to reach this definition.

Framework data take three principal forms:

Alignment layers include geometric controls required to position geospatial information adequately. By themselves, these layers do not have representations of physical, economic or social phenomena, as will other framework layers or application specific layers, but these layers are critical to the reliability and use of all other layers.

Land Feature/Form layers contain well-defined and readily observable natural or manmade physical features that are not subject to interpretation or speculation. These layers include many of the same features that are visible on topographic maps, such as roads, rivers and elevation. Although useful for some applications by themselves, they are also used to provide reference information for conceptual layers.

Conceptual layers are the frameworks that society develops and uses to describe and administer the country. These layers complement a vast amount of application-specific data. They are often interpreted from observations of physical, economic or social factors, and include features such as municipal boundaries, federal electoral districts, and ecological areas. Inclusion of any specific type of boundary under this layer is subject to its availability over large areas of the country, its geometric integration to alignment layers, and a consensus among major stakeholders on the importance of the ubiquity of the data.

2.4.2 Services

Web services have emerged as a key technology in the delivery and exchange of information on the Internet. In the context of the CGDI, services allow for access to and exchange of geospatially-referenced information. From a technical perspective, services can be defined as a collection of operations, accessible through one or more interfaces, that allows a user to invoke a behaviour of value to that user. Most descriptions of Web services highlight the following three characteristics:

1. They are discrete applications handling a specific set of tasks.
2. They are built using XML interfaces, which allow different Web services to be combined within an application.
3. Data and functionality are separate from the user interface, so users see a seamless application even though data and functionality may come from many different Web services.

The underpinning of a distributed architecture is the means by which independent systems communicate and collaborate with one another. Web services based on open standards provide the basis for interaction across the Internet that will allow users to contribute, access, and exchange geospatial data.

Although the concept is straightforward, the success of Web services requires a commitment to the on-going development of open standards and specifications. CGDI-endorsed

standards provide the data infrastructure with the basis for deployment of services and applications. The CGDI defines services that enable access to geospatially-referenced information, and defines a set of interfaces to these services. One or more of these services is invoked by every interaction that a user has with the CGDI. Users access most of these services through the help of client software applications. The visible part of client software applications is comprised of well-known and “open” sets of software interfaces. This approach places the CGDI on the leading edge of the global geospatial landscape. *(Note: Please see the Appendix for details of an architecture based on Web services.)*

2.4.3 Applications

Applications are commonly referred to as computer programs that use instructions to perform specific tasks. The CGDI relies on the development of applications to satisfy the requirements of the geospatial user community. These applications range from those that facilitate end-user access to geospatial information to those that meet the requirements of a data provider who wishes to distribute its data efficiently to all clients of its own community.

A rich set of client software applications based on core CGDI components, interfaces and services, built by CGDI developers and providers, will ultimately deliver most of the anticipated benefits to Canadians.

Applications use one or more Web services to access geospatial data or process geospatial information. Utilization includes, but is not limited to, viewing, publishing, editing or discovering geospatial data. Some general types of client applications include:

- **Viewer client:** Displays graphics that come from map server components.
- **Discovery client:** Enables users to query geospatial data. There are several kinds of discovery mechanisms in the CGDI, for example, provincial, commercial, thematic and national (e.g. the Discovery Portal). The discovery client sends user requests to the data provider’s search server in a structure defined by the search protocol.
- **Publisher client:** Enables data providers to select how to distribute their data to users.
- **Editor client:** Allows data providers to add, remove or update their shared data.

2.4.4 Users

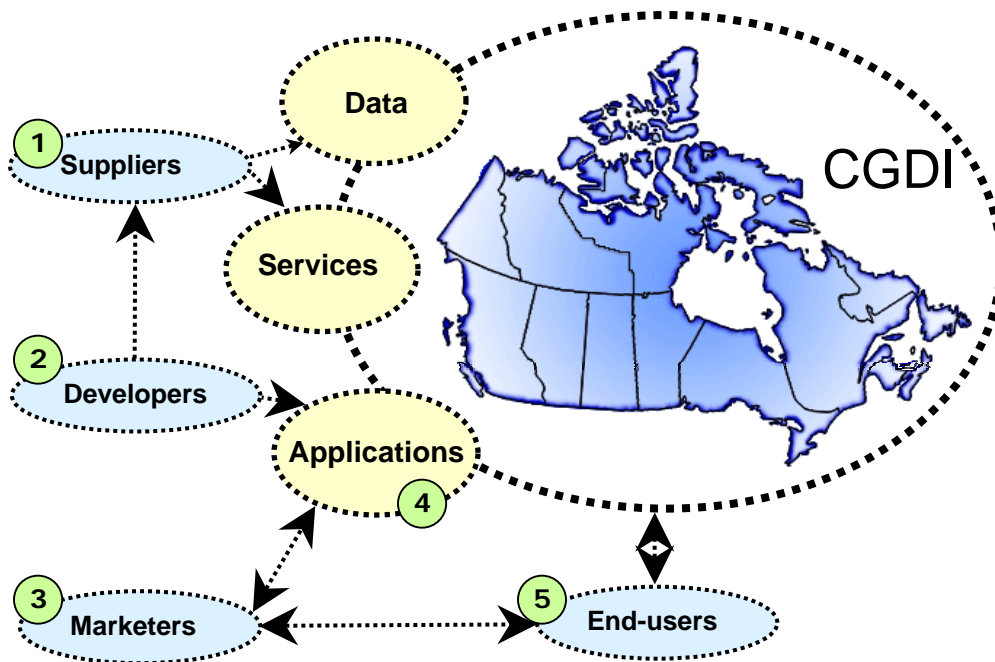
CGDI users include the full range of groups or individuals who participate in the CGDI. They can be represented in the following categories:

- **Suppliers** provide geospatial data and Web services to the CGDI. This category of user operates at the core of the CGDI, providing the building blocks necessary for the development of applications.
- **Developers** create applications that facilitate interaction with the CGDI for users. Applications are developed based on Web services.

- **Marketers** sell and/or support geospatial applications predominantly for end-users. These applications are marketed based on their ability to meet a demand for access to geospatial information.
- **End-users** are consumers of geospatial information for the purpose of making decisions. For the most part, this category of user relies on applications to produce outputs for consumption.

The following figure provides an illustration of a typical relationship between the four categories of users.

Figure 2 – CGDI Users



In this illustration:

1. Suppliers provide data and services to the CGDI.
2. Developers use these inputs to produce applications.
3. Applications are marketed to end-users.
4. Applications can be developed to meet supplier needs.
5. End-users can access the CGDI directly.

3 Architectural Components

3.1 *Objects*

Geospatial objects describe real-world entities that are employed in client applications. The CGDI makes the interfaces to geospatial objects available to clients and providers with seamless views of the information.

The highest-level objects can be represented as packages containing collections of related objects that can be distributed across the infrastructure. The objects can serve multiple functions. Some are at the client level for visualization, some work at the application server level for processing, while some exist at the database level for storage and access.

A representative list of fundamental geospatial objects includes:

- Geographic Features
- Geographic Coverages
- Geographic Measurements/Observations
- Spatial Reference Systems
- Geographic Projects
- Geographic Events
- Geographic Transformations
- Map Styles and Symbologies

3.2 *Open Standards and Specifications*

The CGDI is based on open and shared standards and specifications for operational transactions and information exchange. In distinguishing between a standard and a specification, the authority of a specification rests on its inherent technical excellence and on the breadth of its acceptance in the marketplace. On the other hand, the authority of a standard derives from the authority of the standard-setting organization sponsoring it.

The endorsement and adoption of international or national standards ensures the infrastructure is interoperable with other infrastructures around the world. The process of standardization requires consensus and acceptance of a technical solution by those people who influence the direction and development of a specific technology. To obtain the best results, developing sound standards involves building upon existing work, and refining and improving the standards and specifications. In developing the CGDI, there has been and will continue to be a focus on identifying and promoting existing open and interoperable software interfaces, and developing new ones as required through the help of existing public standards organizations.

Two leading standards bodies play an important role in the development of the CGDI: the Open Geospatial Consortium (OGC) and the International Organization for Standardization (ISO). In general terms, the emphasis of the OGC is on specifications for implementation, while the ISO concentrates more on abstract standards.

3.2.1 The Open Geospatial Consortium

The Open Geospatial Consortium (OGC) began in 1994 in an effort to create a technical committee that would agree on open interfaces for network interoperability of systems. The OGC is a cooperative effort by industry, academia and government to promote and develop open standards based on distributed spatial computing solutions.

Since 1999, the OGC has been developing open Web-based mapping and related specifications in cooperation with a large community of public organizations and software vendors. OGC specifications allow software vendors to implement their products using open and interoperable interfaces that provide users, such as developers and end users, with a larger set of tools for geodata access and related geoprocessing services.

To date, the OGC has developed a number of software interface specifications, including:

- Web Map Server interface specification;
- Web Feature Server interface specification; and
- Geography Markup Language specification.

The OGC has a conformance and testing program for the specifications they develop. This practical bottom-up approach by industry and its vendors develops specifications as a result of implementation and interoperability scenarios. This is in contrast to typical top-down standardization efforts that rely on the hopes that the industry will implement many of the resulting specifications.

3.2.2 The International Organization for Standardization (ISO)

The International Organization for Standardization is a world-wide federation of national standards bodies established in 1947. The mission of ISO is to promote the development of standardization and related activities in the world with a view to facilitating the international exchange of goods and services, and to developing cooperation in the spheres of intellectual, scientific, technological and economic activity.

ISO/TC 211 Geographic information/Geomatics is responsible for the ISO geographic information series of standards. This work aims to establish a structured set of standards for information concerning objects or phenomena that are directly or indirectly associated with a location relative to the Earth.

Unlike previous ISO technical committees, ISO/TC 211 has the unique distinction of beginning a culmination of work that includes the concurrent development of an integrated set of 20 standards for geographic information. While the development of singular or stand-alone ISO standards occurs at a faster rate, the carefully developed ISO/TC 211 set of integrated standards results in interoperability among the standards.

ISO/TC 211 provides standards for:

- Geospatial data, including imagery and information;
- Geospatial data administration and information management; and
- Geospatial services, including location-based services.

ISO/TC 211 has a relationship with OGC whereby OGC specifications are tabled and are eventually formalized as ISO standards.

3.3 CGDI-endorsed Standards

Open standards and specifications that have been endorsed for the purpose of CGDI-deployed services and applications are called “**CGDI-endorsed standards**”. These may include specifications such as those of the OGC and formalized standards such as those of ISO.

Developing geospatial applications using CGDI-endorsed standards provides more than functional requirements for data and services, it also allows for:

- Interoperability among all stakeholders;
- Increased availability, access and sharing of geospatial information;
- Efficient use of digital geospatial information and associated hardware and software systems;
- Promotion of existing and evolving Canadian technology and expertise;
- Ensuring compliance of Canadian data products and services with international standards;
- Collaboration on international activities; and
- Solutions for global delivery of geospatial information.

In terms of day-to-day activity, using CGDI-endorsed standards achieves the following benefits:

- Reduces development costs;
- Minimizes redundancy;
- Hides the complexity of components; and
- Permits GIS practitioners and developers to benefit from “plug and play” components.

CGDI-endorsed standards fall into the following categories (*actual specification names in italics*):

- **Data visualization:** The *Web Map Server (WMS)* interface specification defines services to create and display superimposed map layers of geographic data from multiple sources.
- **Data presentation:** The *Styled Layer Descriptor (SLD)* specification is a companion specification to the WMS interface specification. It provides some control over how geographic features are to be portrayed, or symbolized.
- **Data storage:** The *Web Map Context (WMC)* document specification allows for the creation and storage of information (metadata) about maps for re-use. It is also a companion specification to the WMS interface specification.
- **Data manipulation:** The *Web Feature Service (WFS)* specification defines data manipulation operations on geographic features such as water bodies, infrastructures, etc.

- **Data encoding:** *Geography Markup Language (GML)* is an XML coding for the transport and storage of geographic information, including both the geometry and properties of geographic features.
- **Data query:** The *Filter Encoding* specification provides a standard XML-based means of restricting the records that are returned from a query. Filter encoding can handle both spatial and non-spatial aspects of a query.
- **Data access:** The *GeoData Discovery Service* searches geospatial data resource registries, and retrieves metadata about geospatial data. Gazetteers look up geographic regions based on secondary reference systems such as postal codes and place names (e.g., Canadian Geographic Name Service or the National Topographic Name Service).
- **Data description:** The *Metadata for Geodata* specification provides a standard way of describing geospatial data.

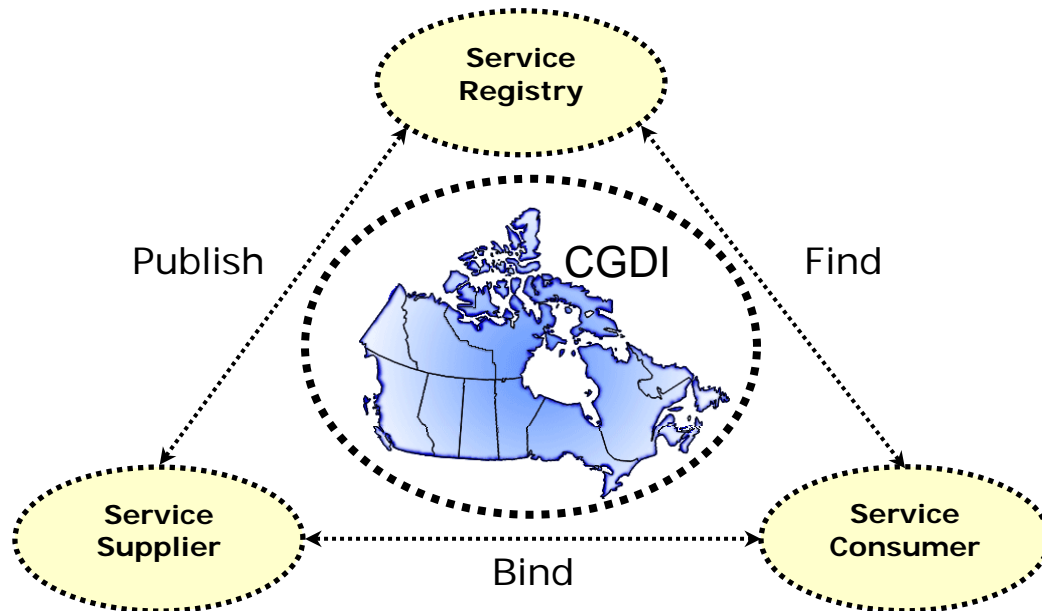
Compliance with the CGDI implies that an application or service must use or support the relevant CGDI endorsed-standards.

A number of other specifications have been identified as recommendation papers or discussion papers, but have not yet been endorsed as CGDI standards. Developers are encouraged to examine these specifications, which are identified on the CGDI Architecture Webpage at

<http://www.geoconnections.org/en/communities/developers/architecture/fa=architecture.welcome>

3.4 Registries

In a Web services architecture, the service registry serves as the key link between the supplier and the consumer. Once a supplier has published metadata about its offering to the registry, a consumer can verify with the registry before connecting to the supplier. Furthermore, when changes to the supplier's offering occur, the registry can redirect the consumer to the new location, or present alternative services from similar suppliers. Visually this relationship can be depicted as follows:

Figure 3 – Registries within a Web Service Architecture¹

To facilitate service discovery and access, the CGDI will contain one or more interoperable Service Registries that will provide a mechanism to:

- Populate or publish information (metadata) that describes CGDI-compliant services.
- Search and retrieve detailed information (metadata) that describes the CGDI-compliant services, and allows the consumer to evaluate the suitability to meet a need.
- Search and retrieve detailed information (metadata) that describes the CGDI-compliant services, and allows CGDI developers to evaluate a service for potential integration into an application.

3.5 Metadata

Metadata has become an essential element in the creating, finding, sharing and applying of geospatial information. Metadata, or data about data, describes the origin of the data or service, and tracks any changes. Metadata answers the who, what, where, when, why, and how about every facet of the data or service being documented. This includes details regarding ownership, quality, time of collection or update, attribute information, and how it can be accessed and obtained.

In terms of functionality, metadata allows a user to discover, evaluate and access data. Using metadata, GIS practitioners are able to:

- Determine what geospatial data is available;
- Evaluate the suitability of the data for use;
- Access the data;

¹ Programming Web Services with SOAP, James Snell, Doug Tidwell, Pavel Kulchenko, December 2001.

- Transfer and process the data; and
- Accomplish these things in the order appropriate for them.

3.5.1 The Importance of Metadata

Metadata serves many important purposes. It is a vital foundation for understanding, collaborating and sharing resources with others. It allows people to determine what the best resources are for their individual needs by permitting them to see the details of the data itself and its history. It benefits data-producing organizations by ensuring that data holdings are well documented over time so their value for the data holder and user is maintained. Metadata is important in the creation of a spatial data clearinghouse, where potential users can search, find and compare data in all its detail.

Metadata can be organized into several levels, ranging from a simple listing of basic information about a data collection, to complex and detailed documentation about an individual data set. The key benefit of metadata is that it provides the user with a complete description and history of the data or service. Structured and complete metadata empowers users to seek geospatial attributes using specific fielded and geographic parameters in their search as opposed to simple free text. This will return a search result that is more specific to user needs as opposed to a search result filled with many unrelated items.

3.5.2 Metadata Standards

A metadata standard is a common set of terms and definitions that describe geospatial data. A level of conformance is important to ensure everyone can discover, understand and share data by finding and comparing common details regarding the data. A metadata standard outlines the characteristic properties to be recorded, as well as the values the properties should have. Such standardization of the vocabulary makes information sharing more reliable and universal. The CGDI endorses two metadata content standards: FGDC CSDGM and ISO 19115.

3.6 *Security and Authentication*

Security and authentication are critical for global geospatial data processing and operations. Even a simple service such as rendering a map layer has security implications. Different sources of geographic data may have different levels of sensitivity in terms of who is allowed to see them. The need for security and authentication mechanisms increases with the need to share information in an open and interoperable fashion, particularly in those operations that create or update data.

The CGDI must provide for the authentication and authorization of clients when required. Ideally, the security model will provide harmonization of identification, certification, authorization and authentication across jurisdictions, for both private- and public-sector users. Access to information and services will be authenticated to the degree required by the specific information and services.

CGDI security and authentication is an enabling layer that can be implemented with all documented CGDI Web services. A secure infrastructure will provide protected, verified and authorized access to services and data:

- **Protected access:** Interactions between components are private (prevents eavesdropping) and integrity is ensured (prevents tampering).
- **Verified access:** Communications are authenticated (to prevent impostors by confirming identity or role) and signed (to be non-deniable).
- **Authorized access:** Access to services and data is controlled by the verified identity and/or role of the requesting user or client.

3.7 Robustness and Redundancy

Some of the services and registries in the CGDI are critical to the ongoing operations of the infrastructure. The services that support access to critical functionality are considered “core” services.

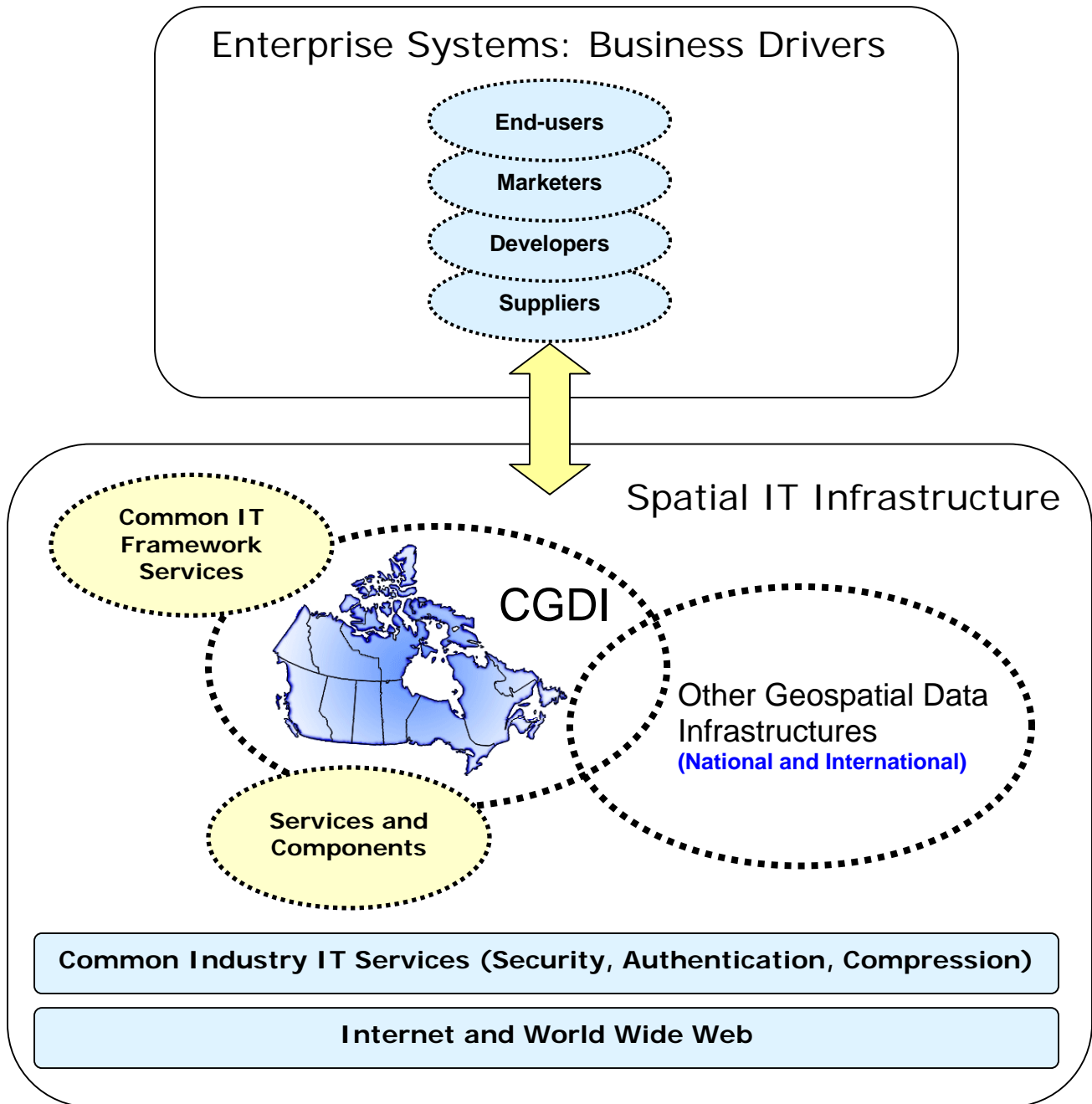
Some core services will be implemented by identifying a “Master” implementation that is mirrored on other identical servers, but most services will be implemented on multiple systems for one or more of the following reasons:

- Robustness arising from redundancy (so that portions of the architecture will continue to work while others are out of service or inaccessible).
- Efficiencies that may be obtained by coupling local datasets with local service implementations (so custodians of information can provide access to information from their own systems), and
- Benefits arising from local management and maintenance of service implementation.

4 Architectural Requirements

The CGDI is a spatial information technology infrastructure driven by a variety of key business applications from both the private and public sector. Implementation of the CGDI will offer software components, services and data that will reduce the difficulty, cost and time of application development and provide value that can come only from a network of clients and providers. This premise is illustrated in the following figure.

Figure 4 – Architectural Requirements



The principal architectural requirements for the CGDI include:

- Enable online access to a wide range of geographic information and services;
- Enable integration of geographically distributed geospatial information;
- Enable collaboration by multilateral information exchange and synchronization;
- Allow autonomous organizations to develop interdependent relationships in a distributed environment; and
- Facilitate the definition and sharing of geospatial semantics.

Each requirement is key to the success of the CGDI, yet they are enabling features rather than functional features. They are achieved through organization and structure rather than through individual component capabilities.

4.1 Interoperability

Common to each of the requirements stated above is the need for interoperability. Interoperability is the ability of a system or component of a system to access a variety of heterogeneous resources by means of a single, unchanging operational interface.

"To be interoperable, one should actively be engaged in the ongoing process of ensuring that the systems, procedures and culture of an organisation are managed in such a way as to maximize opportunities for exchange and re-use of information, whether internally or externally."

Paul Miller (2002)

Interoperability facilitates information sharing, and provides the freedom to mix and match information system components without compromising overall success. Interoperability allows users to:

- Find information, services, and applications when they are needed, independent of physical location;
- Understand and employ the discovered information and tools, no matter what platform supports them, whether local or remote; and
- Evolve a processing environment for commercial use without being constrained to a single vendor's offerings.

Interoperability between systems and components of systems has several aspects:

- **Network Protocol Interoperability** allows basic communications between components.
- **Standard Interface Specifications** allow client applications to execute procedures on remote systems.
- **Data Transport Interoperability** allows for sharing of data and services through transparent access, regardless of any proprietary data storage formats.
- **Semantic Interoperability** refers to applications interpreting data consistently in the same manner in order to provide the intended representation of the data.

5 Development Strategy

5.1 Overall Approach

The architecture of the CGDI is founded on the assumption that enough domain knowledge exists to start the development of core services and the implementation of systems based on those core services. As the infrastructure continues to develop, an increasing number of services and their relationships will be specified, and existing services will be refined.

For data infrastructures comprised of standards-based services it is possible to describe that architecture very efficiently by specifying the action and interface for each of the services. The prerequisite is that the behaviour of the services and the required interfaces between those services are clearly understood.

The CGDI is comprised of a nucleus of standards-based services in the following areas:

- Searching registries and spatial databases;
- Rendered maps;
- Geographic features;
- Registry data;
- Geospatial metadata; and
- Geographic Feature Encoding

Developing a complete description of a distributed architecture is an extensive and time-consuming activity requiring the concentrated efforts of a team of architects and engineers. The CGDI has adopted a spiral development approach to its design, based on the specification of standards-based services.

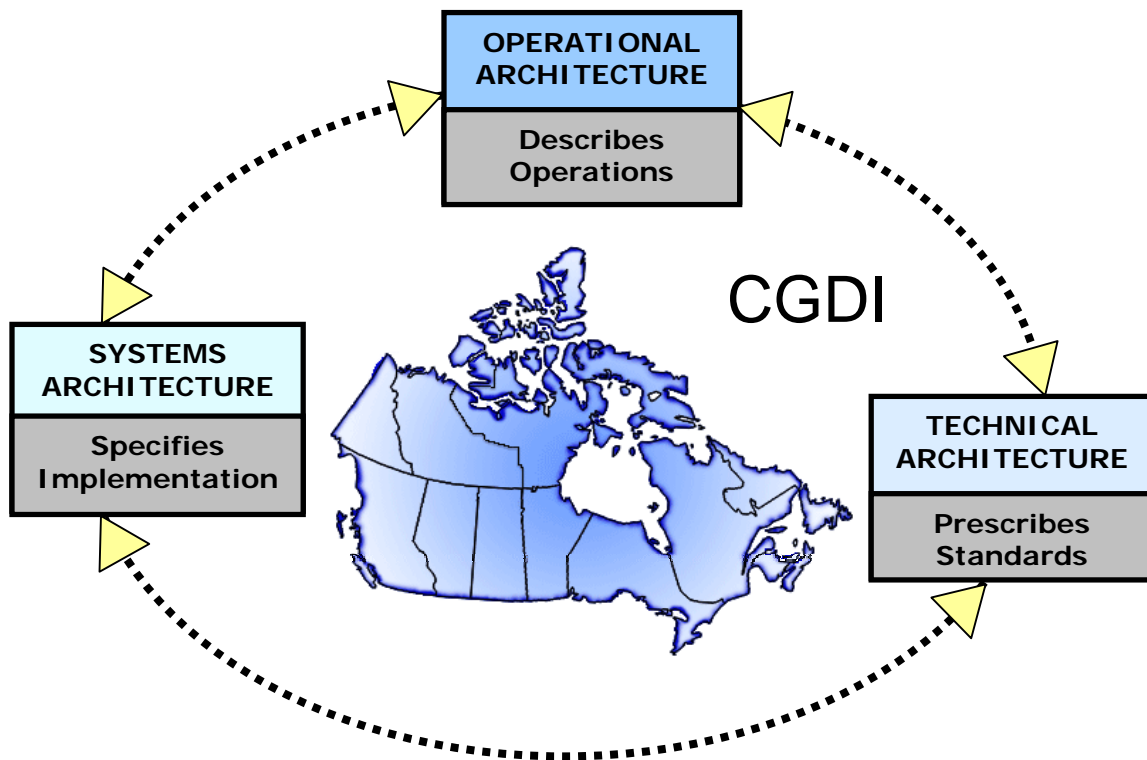
5.2 Architecture Perspectives

The CGDI has multiple stakeholders that interact with each other across distributed facilities. The architecture of the infrastructure and its supported capabilities must be clearly described and available to those who interact with it. Descriptions of the architecture from different perspectives are used to enable the communications between the diverse stakeholders: suppliers, developers, marketers and end-users.

Three different perspectives of the CGDI architecture are illustrated in Figure 5. The major components include:

- **Operational Architecture:** Describes operational concepts and defines how operational requirements of the enterprise clients and users are realized.
- **Technical Architecture:** Standards, profiles, and specifications that govern and constrain the design of components that must interoperate in the system environment.
- **Systems Architecture:** Design and implementation descriptions of components and systems.

Figure 5 – CGDI Architecture Perspectives



The **Operational Architecture** provides a view from the perspective of users in the language of the business enterprise or application domain. It includes use cases that describe the operational functions of the infrastructure.

The **Technical Architecture** provides a view of the infrastructure from the perspective of the architect concerned with interoperability, adaptability, maintainability, availability and quality of service.

The **System Architecture** provides a view of the infrastructure from the perspective of the individual systems engineers and developers who design, integrate, test and maintain components and operational systems.

The development of the Operational Architecture and the Technical Architecture falls within the umbrella of the CGDI. The Systems Architecture design is left to each of the individual participating organizations. The specification of services and interfaces that form the Technical Architecture will allow organizations to develop and/or implement compliant components or systems that offer and/or use these services. The CGDI is described in the Technical Architecture in terms of standards-based services that are well defined and whose behaviour and relationships are well understood. With the knowledge and understanding of the underlying IT architecture, geospatial components and systems can be designed and implemented to either make use of CGDI services or extend existing services.

5.2.1 Operational Architecture

The operational architecture view is a description of the tasks and activities, operational elements, and information flows required to accomplish or support an operation.

It contains descriptions (often graphical) of the operational elements, assigned tasks and activities, and information flows required to support the CGDI users. It defines the types of information exchanged, the frequency of exchange, which tasks and activities are supported by the information exchanges, and the nature of information exchanges in detail sufficient to ascertain specific interoperability requirements.²

The CGDI is being implemented as a network of cooperating physical servers, providing services (and data via services) in such a way that:

- An application can be developed that makes use of these services. This permits a saving in development time and costs, and operating costs, and allows for uniformity of data; and
- A subset of geographically related problems can be solved in an *ad hoc* way by using publicly available user interfaces built on top of the underlying services. These interfaces provide discovery of, access to, and general "geospatial" functionality on top of publicly available data.

CGDI services are provided by a broad spectrum of agencies. There is no real "centre" to the CGDI. As the operational architecture develops, it will further define:

- Who implements the services defined in the technical architecture;
- How organizations go about implementing services in ways that foster interoperability;
- The degree of "looseness" of the infrastructure in terms of hierarchies of servers and registration authorities;
- The bindings/linkages that make domain infrastructure and/or systems part of the CGDI;
- The role of common or centralized services;
- Who can access common/centralized services, and who pays; and
- The service level guarantees in terms of reliability and response time.

5.2.2 Technical Architecture

The technical architecture view is the minimal set of rules governing the arrangement, interaction, and interdependence of system parts or elements, whose purpose is to ensure that a conformant system satisfies a specified set of requirements.

The technical architecture view provides the technical systems-implementation guidelines upon which engineering specifications are based, common building blocks are established, and product lines are developed. The technical architecture view includes a collection of the technical standards, conventions, rules and criteria that govern system services, interfaces,

² C4ISR Architecture Framework Version 2.0, C4ISR Architecture Working Group, 1997.

and relationships for particular systems architecture views and that relate to particular operational views.³

The technical architecture continues to evolve. This area is undergoing tremendous and rapid change, and the next 2-3 years will likely see huge leaps in enabling technologies, with adoption trailing perhaps a year or two behind. There are a host of services available and others currently under development. As the CGDI continues to mature, the number of available services will continue to increase.

5.2.3 Systems Architecture

The systems architecture view is a description, including graphics, of systems and interconnections providing for, or supporting, CGDI functions.

For a domain, the systems architecture view shows how multiple systems link and interoperate, and may describe the internal construction and operations of particular systems within the architecture. For the individual system, the systems architecture view includes the physical connection, location, and identification of key nodes, circuits, networks, platforms, etc., and specifies system and component performance parameters (e.g., mean time between failure, maintainability, availability). The systems architecture view associates physical resources and their performance attributes to the operational view and its requirements per standards defined in the technical architecture.⁴

5.3 *Using the CGDI*

When using the CGDI, clients use services that obtain and manipulate data. This is done using applications that access the physical implementations of these services. These physical implementations are called "servers".

Developers use the service and interface specifications to develop compliant components, and deploy them to offer CGDI-compliant services. For example, a CGDI data provider will provide metadata through a server that complies with the specifications for the Geographic Data Discovery Service. Either that server could be based on a software component obtained from another organization, or it could be based on a custom component developed entirely in-house.

Application developers also use the service and/or interface specifications in order to develop end-user applications. It is possible to chain together any number of CGDI services in order to provide a wide variety of specialized systems and applications.

³ Ibid

⁴ C4ISR Architecture Framework Version 2.0, C4ISR Architecture Working Group, 1997.

Glossary of Terms

The terms and definitions used in this document are described below. Although significant effort has been made to be consistent in the use of terms, the terminology continues to evolve.

Term	Definition
Application	The combined set of software programs that perform a specific function directly for a user. Further, a CGDI application is the utilization of CGDI technology (e.g., tools and/or services) and CGDI data by a given user or community of practice to address a specific issue.
Architecture	The organizational structure and operating environment of the CGDI, including the relationships between its parts, and the principles and guidelines governing their design and evolution.
Canadian Geospatial Data Infrastructure (CGDI)	An Internet infrastructure comprised of the developments of the federal, provincial, territorial and private-sector partners who are creating the technology, standards, access systems and protocols necessary to harmonize Canada's geospatial databases, and make them accessible on the Internet.
Client	A software component that accesses a service. The Guide to the CGDI distinguishes between a client (an inanimate part of the process) and a user (an individual who uses a computer, program, network or related service).
Conceptual Architecture	An overview of the services, data, technology and institutional environment of the CGDI. It describes, in general terms, both what the CGDI will include, and how it will operate.
Data	Distinct pieces of factual information, especially information organized for analysis or used to reason or make decisions. Data is usually formatted in a special way, and exists in a variety of forms. Data in the CGDI comprises maps, satellite images, publications and other geospatial data provided by Canadian and international sources.

Event	An occurrence of interest to users or developers of the CGDI. Events can be things such as the adjustment of a feature in a framework data layer, a flood in the Red River basin, or the release of a new specification for a CGDI service.
Framework Data	The set of geospatial data that provides the reference framework for all other CGDI data.
Gazetteer	Dictionary of instances of a class or classes of features containing some information regarding position.
Geodata	Georeferenced spatial data such as a road network or a satellite image. Geodata explicitly describes the spatial extent of a set of features or describes a measurable surface. It includes both geospatial data and geolinked data.
Geographic Information System (GIS)	A computer system for capturing, storing, checking, integrating, manipulating, analyzing and displaying data related to positions on the Earth's surface. A GIS can be used for handling various types of maps. These might be represented as several different layers where each layer holds data about a particular kind of feature. Each feature is linked to a position on the graphical image of a map, and layers of data are organized to be studied and to perform statistical analysis.
Geographic Markup Language (GML)	An open XML grammar specification used to transfer geographic features via the Internet.
Geolinked Data	Data that is referenced to an identified set of geographic features without including the spatial description of those features. Geolinked data is normally attribute data in tabular data (such as population counts) that refers to a known framework (such as provinces), where the elements (the provinces) are referred to by their unique identifier (such as the province name). Geolinked data refers to all attribute data that is not directly attached and bundled with the geographic coordinates to which it applies.
Geospatial	

	Referring to location relative to the Earth's surface. "Geospatial" is more precise in many GIS contexts than "geographic," because geospatial information is often used in ways that do not involve a graphic representation, or map, of the information.
Geospatial Data	Data with explicit geographic positioning information included, such as a road network from a GIS, or a georeferenced satellite image. Geospatial data may include attribute data that describes the features found in the dataset.
Geospatial Information	Information about entities and phenomena that includes their location with respect to the Earth's surface. Frequently used as a synonym for "geodata", but technically geodata are "dry" digitally represented facts or recorded observations, which on their own have no meaning. They become information when interpreted and put in context by humans.
Infrastructure	A reliable, supporting environment, analogous to a road or telecommunications network that facilitates the access to geographically related information using a minimum set of standard practices, protocols and specifications.
Interface	A specification for a set of operations that are made externally available by a component to other components. The state and functionality of a component is hidden, and is only made externally accessible through the interfaces of the components. The interfaces are the only "public" or "visible" part of the component. The same interface may be provided by several components and used by many components or applications.
Metadata	Information about data. Metadata describes how and when and by whom a particular set of data was collected, and how the data are formatted. Metadata is essential for understanding information stored in data warehouses.
Reference Architecture	A technical blueprint that identifies and defines the services that comprise the CGDI, and specifies the interfaces to those services.

Registry	A listing of the individual datasets, services, or other things made available by an organization to users of the CGDI.
Server	A physical installation of a component that delivers a service and provides the realization of its operations.
Service	A collection of operations, accessible through one or more interfaces, that allows a user to evoke a behaviour of value to that user. A service is delivered by a server.
Specification	A document written by a consortium, vendor, or user that specifies a technological area with a well-defined scope, primarily for use by developers as a guide to implementation. A specification is not necessarily a formal standard.
Standard	A document that specifies a technological area with a well-defined scope, usually by a formal standardization body and process.

References

- C4ISR Architecture Working Group. C4ISR Architecture Framework Version 2.0. 1997.
- CGDI Online Training.
http://www.geoconnections.org/publications/training_manual/e/index.htm
- CGDI Service Registry Service. Version 0.8.6. July 7, 2003.
- General Description of the CGDI Service Registry Service (SRS). Version 0.8.6. July 7, 2003.
- GeoConnections. A Developers' Guide to the CGDI: Developing and Publishing Geographic Information, Data and Associated Services. February 2004.
- Miller, Paul. Interoperability. What is it and Why should I want it?
www.ariadne.ac.uk/issue24/interoperability/intro.html
- Security and Authentication Strategy. Draft 2.0. November 7, 2001.
- Snell, James, Tidwell, Doug, Kulchenko, Pavel. Programming Web Services with SOAP. December 2001.

APPENDIX: AN ARCHITECTURE BASED ON WEB SERVICES

The Canadian Geospatial Data Infrastructure (CGDI) is based on a Web service architecture. This appendix defines the needs that led to the development of Web service architectures, and describes the attributes of Web service architectures in general and the CGDI in particular, with a description of the benefits and an example of how to develop an application using this architecture.

A1.1 The Need for Web Service Architectures

Web service architectures arose from the need for computer-to-computer communication between distributed organizations. The Internet has vastly increased person-to-person communication. It has also fuelled the demand for computer-to-computer communication, but progress in this area has lagged due to the lack of a ubiquitous communication mechanism.

This mechanism must accommodate all who wish to participate, no matter what computer platform they use. And, the barriers to participation must be low. Historical solutions to this need, e.g. CORBA, have had too high an entry barrier. These needs are paramount for B2B e-commerce, and the market is now fuelling the development of Web service architectures.

The same needs are also imperative to the establishment of distributed infrastructures such as the CGDI. The CGDI is purposed to increase the on-line availability of geospatial data and services, and foster new geomatics applications. Open computer access to distributed data and other geomatics services is crucial to meet the goal of the CGDI.

A1.2 Web Service Architecture Attributes

Web service architectures provide a distributed environment in which services can be deployed and invoked using standard Internet protocols. A service is a reusable encapsulation of cohesive functionality with a well-defined programmatic interface.

The key concepts in this definition are:

- **Internet Protocol:** Web services invocation is based on ubiquitous Internet transport mechanisms such as HTTP and XML.
- **Distribution:** Web services can be deployed on any computer that is connected to the Internet, and can potentially be invoked by any other computer connected to the Internet.
- **Reusability:** The services are units of functionality that are potentially reusable in many different applications.
- **Cohesion:** All the functions performed by the service are related to one another, and are not tightly bound to other services.
- **Programmatic Interface:** The service has a well-specified contract that it obeys, and this interface is suitable for programmatic access.

A Web service architecture must fundamentally define a distributed computing platform in which Web services can be deployed and invoked. Architectures that are more sophisticated provide additional capabilities such as service discovery, security and quality of service. There are a number of emerging Web service architectures such as IBM's Web Service Conceptual Architecture and Microsoft's .Net

A1.3 The CGDI architecture

The CGDI is a Web service architecture instance purposed for the Canadian geospatial domain. The CGDI, as a Web service architecture, must provide an environment in which Web services can be deployed and invoked. The CGDI could define its own platform, but instead leverages the efforts of other endeavours, such as B2B e-commerce and the Open Geospatial Consortium (OGC), and adopts a general-purpose and widely supported platform when available. In the absence of a widely supported general-purpose platform, the CGDI today is based on the platform described in the draft OGC Basic Service Model.

This platform is summarized as follows:

- a service is deployed via a URL accessible through a Web server;
- services are invoked via an HTTP "get" or "post";
- when using a "get", parameters are passed using the CGI conventions;
- when using a "post", parameters are XML encoded; and
- the service must support a GetCapabilities operation that returns service metadata in XML form.

The OGC is evolving this platform and working to align it with other emerging Web service architectures. One recent aspect of this alignment has been the adoption of Web Services Description Language (WSDL) for service definition.

Standard Web service platforms do not, nor will not, provide some capabilities that are crucial to the CGDI. The missing capabilities relate mostly to the fact that, within a geospatial infrastructure, **data** deserves the same prominence as **services**. A general-purpose Web service platform allows **services** to be discovered and invoked, but the CGDI must also allow **data** to be discovered.

CGDI **data** can, of course, be discovered through **services**, but these **services** are sufficiently important to be considered part of the CGDI architectural platform (just as service discovery services are part of a standard Web services platform). The CGDI Web services platform thus provides services to register geospatial data and to discover geospatial data.

The CGDI also spawns the identification and specification of Web services specific to the geospatial domain.

A1.4 Benefits of a Web Service Architecture

Web service architectures leverage off the pervasive Web, providing a ubiquitous distributed computing platform. Applying distributed computing no longer means a heavy financial and training investment in technologies such as CORBA. Many methods now exist to publish legacy applications to the Web.

A service can also be made Web accessible no matter how it is implemented, or what platform it executes on. Applications can be easily built from services running on heterogeneous platforms in any location. These benefits are vital for large organizations (corporations or governments) whose distributed divisions, running different computing platforms, need to inter-operate. Inter-organizational infrastructures, such as the CGDI and B2B infrastructures, become possible through Web service architectures.

Applying Web service architectural principles to development yields applications composed of loosely coupled, distributed (or distributable), and reusable services. Complex applications are decomposed into simpler entities that can be independently developed. The development adds to the pool of services that become available for use in new, even more sophisticated applications. And, the cost of application development is reduced, making application development more cost effective.

A1.5 How to Exploit The Web Service Architecture

This example demonstrates how an application can be built using a Web service architecture. For example, consider a traditional Web application that accepts place names, and returns a map showing a road route between the two places.

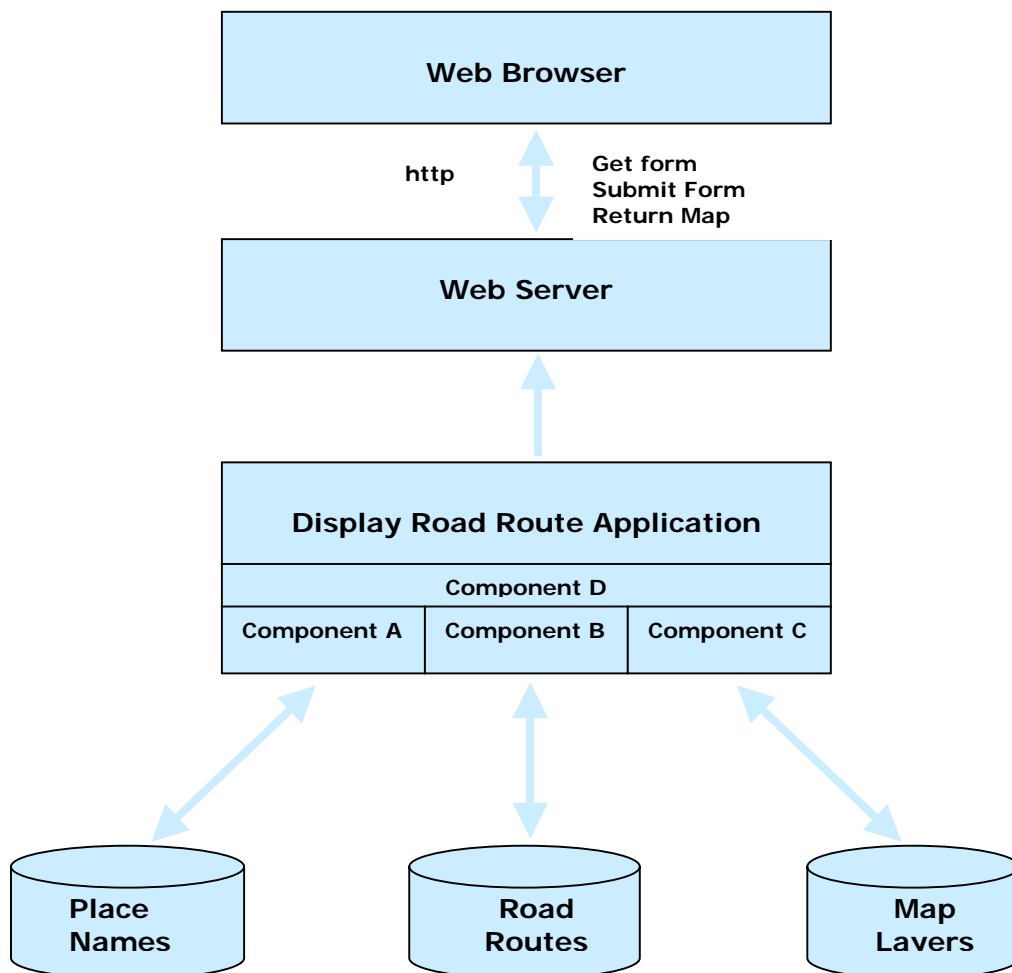


Figure 1: Traditional Application Design

The user invokes the application through a Web browser, and enters the place names. The Web server then invokes an application, which performs the necessary processing (accessing a database of place names, a road route database, and a database of other map features), and returns a map to the browser as a GIF image. For this discussion, the important aspect is the way in which the application is designed, particularly the partitioning of the software into components and the interfaces between the components. Traditional design techniques would structure the software into logical units with well-defined interfaces, but would deploy the application as a monolithic entity, not suitable for distribution.

Applying Web-service based techniques to this application would decompose the application into a number of Web service components, which are deployed on standalone servers with an HTTP interface.

The user's access to the Web browser and the browser-Web server interaction is unchanged, and hence is not shown. The application has been constructed to make use of the services provided by distinct servers:

- A gazetteer server provides services to convert a place name into a geographic location;
- A route server provides services to get the road route between two geographic locations;
- A map server provides services to get a base map on which the route will be displayed;
- An overlay server provides a service to overlay the route on the map; and
- A display route server makes use of the above to convert a road route request into a map displaying the route.

The services are deployed in the Web service platform and can potentially be located anywhere on the Internet; though for an initial deployment, the services might all reside on a single host computer.

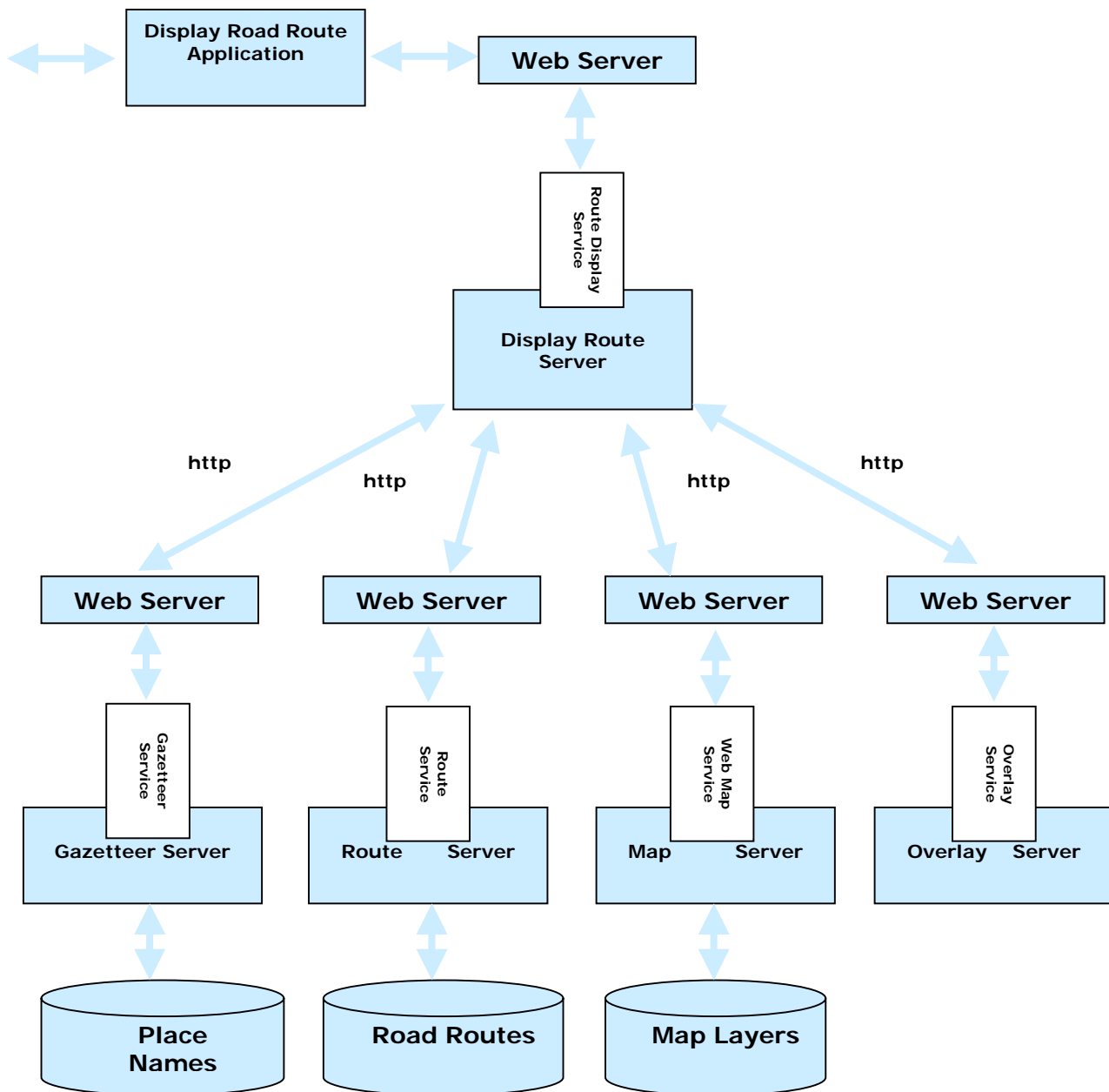


Figure 2: Application Design using Web Services Architecture

The benefits of this architecture are not necessarily apparent when one considers this application in isolation. However, they are immediately obvious when the application is considered as part of a distributed infrastructure such as the CGDI:

- The development of this application has resulted in the deployment of a number of reusable services that can bootstrap the development of other applications;
- The architecture is scalable - even if all of the services were initially deployed on a single computer and heavy use swamped the server, services could easily be migrated to different computers.;

- The back-end services can be provided by the most suitable participant (e.g. the Atlas of Canada can provide the server that provides the base map).
- Additional benefits are realized when the service interfaces become "standards" (formal or *de facto*), and there are multiple deployments of the same services. The design allows the application to make use of new and better implementation of the services as they emerge.