Intercloud

Solving Interoperability and Communication in a Cloud of Clouds

Jazib Frahim, CCIE® No. 5459
Venkata Josyula, CCIE® No. 13518
Monique J. Morrow
Ken Owens

ciscopress.com

Intercloud: Solving Interoperability and Communication in a Cloud of Clouds

Jazib Frahim, CCIE No. 5459
Venkata Josyula, CCIE No. 13518
Monique J. Morrow
Kenneth Owens

Cisco Press
800 East 96th Street
Indianapolis, Indiana 46240 USA

Intercloud: Solving Interoperability and Communication in a Cloud of Clouds

Jazid Frahim

Venkata Josyula

Monique J. Morrow

Ken Owens

Published by:
Cisco Press
800 East 96th Street
Indianapolis, IN 46240 USA

Printed in the United States of America

First Printing June 2016

Library of Congress Control Number: 2016905366

ISBN-13: 978-1-58714-445-5

ISBN-10: 1-58714-445-X

Warning and Disclaimer

This book provides an overview of Intercloud technologies. Every effort has been made to make this book as complete and as accurate as possible, but no warranty or fitness is implied.

The information is provided on an “as is” basis. The authors, Cisco Press, and Cisco Systems, Inc. shall have neither liability nor responsibility to any person or entity with respect to any loss or damages arising from the information contained in this book or from the use of the discs or programs that may accompany it.

The opinions expressed in this book belong to the author and are not necessarily those of Cisco Systems, Inc.

Trademark Acknowledgments

All terms mentioned in this book that are known to be trademarks or service marks have been appropriately capitalized. Cisco Press or Cisco Systems, Inc., cannot attest to the accuracy of this information. Use of a term in this book should not be regarded as affecting the validity of any trademark or service mark.

Special Sales

For information about buying this title in bulk quantities, or for special sales opportunities (which may include electronic versions; custom cover designs; and content particular to your business, training goals, marketing focus, or branding interests), please contact our corporate sales department at corpsales@pearsoned.com or (800) 382-3419.

For government sales inquiries, please contact governmentsales@pearsoned.com.

For questions about sales outside the U.S., please contact intlcs@pearson.com.

Feedback Information

At Cisco Press, our goal is to create in-depth technical books of the highest quality and value. Each book is crafted with care and precision, undergoing rigorous development that involves the unique expertise of members from the professional technical community.

Readers' feedback is a natural continuation of this process. If you have any comments regarding how we could improve the quality of this book, or otherwise alter it to better suit your needs, you can contact us through email at feedback@ciscopress.com. Please make sure to include the book title and ISBN in your message.

We greatly appreciate your assistance.

Publisher: Mark Taub

Portfolio Manager: Brett Bartow

Business Operation Manager, Cisco Press: Jan Cornelssen

Executive Editor: Mary Beth Ray

Managing Editor: Sandra Schroeder

Development Editor Communications: Box Twelve

Senior Project Editor: Tracey Croom

Copy Editor: Barbara Wood

Technical Editor: Linda Strick

Editorial Assistant: Vanessa Evans

Cover Designer: Chuti Prasertsith

Composition: codeMantra

Indexer: Cheryl Lenser

Proofreader: Chuck Hutchinson

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV
Amsterdam, The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at www.cisco.com/go/offices.

About the Author(s)

Jazib Frahim, CCIE No. 5459, is a Principal Engineer in Cisco Security Solutions. He has been with Cisco for about 17 years, with a focus on cyber security and emerging security technologies. Jazib is also responsible for guiding customers in the design and implementation of security solutions and technologies in their networks. He leads a team of solutions architects to manage the lifecycle of solutions development. Jazib has also been engaged in the development of a number of customer-focused services, such as advanced threat analytics, hosted identity services, bring your own device (BYOD), and many others.

Jazib holds a bachelor's degree in computer engineering from Illinois Institute of Technology and an MBA from North Carolina State University. In addition to CISSP and CISM, Jazib also holds two CCIEs, one in routing and switching and the other in security. He has presented at many industry events, such as Cisco Live, Interop, and ISSA, on multiple occasions. He has also authored and coauthored numerous technical documents, whitepapers, and books, including the following Cisco Press titles: *Cisco ASA: All-in-One Firewall, IPS, Anti-X, and VPN Adaptive Security Appliance* (First, Second, and Third Editions), *Cisco Network Admission Control, Volume II: NAC Framework Deployment and Troubleshooting,* and *SSL Remote Access VPNs*.

Venkata (Josh) Josyula, CCIE No. 13518, attended the University of Miami, Florida, where he completed MS and PhD degrees in engineering. His undergraduate education is from Andhra University, Waltair, India. He has about 30 years of IT and telecommunications experience working at Cisco as Distinguished Engineer and AT&T/ Lucent–Bell Laboratories as Distinguished Member of Technical Staff.

Josh led and built the first Cisco cloud solution, known as Virtualized Management Data Center (VMDC), and holds a patent on cloud containers that is widely used in deploying clouds. In addition, Josh has published over 60 technical papers, articles, and books on network management, OSS/BSS, and cloud. He is coauthor of the Cisco Press book *Cloud Computing: Automating the Virtualized Data Center*.

Josh recently retired from Cisco and is doing consulting work on cloud solutions for enterprises and service providers. He lives near Princeton, NJ, with his wife.

Monique Morrow is the Chief Technology Officer (CTO), New Frontiers Engineering, at Cisco. Monique has a track record of co-innovating with customers across the globe from North America to Europe and Asia. Her specialties are in networking technology; grid, cloud computing, Intercloud federation, Internet of Things; M2M security and e-health; semantic web; and business development. Under Cisco's Office of the CTO, both as an individual contributor and manager, Monique built a strong leadership team in Asia-Pacific. Her specific geo-area targets were China and India. Monique's role in these important regions drove Cisco's globalization and country strategies and met all of her targeted goals.

Monique is a staunch advocate for women in technology and was selected as one of the top 50 inspiring women in technology for Europe 2016, and a finalist for Digital Woman of the Year for Europe 2015. She is spearheading the "Internet of Women" global collaborative movement with a goal to develop a new social science for women in technology to sustainably transform this industry.

Monique has also been published in IEEE and other journals and speaks frequently at conferences, and she has coauthored three books, including *MPLS and Next-Generation Networks: Foundations for NGN and Enterprise Virtualization* by Cisco Press.

Ken Owens is Chief Technology Officer (CTO), Cloud Platforms Engineering, at Cisco Systems. Ken is responsible for creating and communicating technical/scientific vision and strategy for Cloud Infrastructure Services (CIS) business. He brings a compelling view of technology trends in enterprise IT (infrastructure, computing, SaaS, virtualization, and cloud) and evangelizes the technology roadmap for the business.

Ken started MANTL.io, a fully integrated and automated microservices infrastructure, and Cisco Shipped, an automated development SDLC. He is responsible for Cisco's cloud-native development engineering and Technical Committee representative to the Cloud Native Computing Foundation (CNCF).

Before joining Cisco in 2014, Ken spent over seven years at Savvis as the Chief Scientist, CTO, and VP Security and Virtualization Technologies. Prior assignments include five years as a network security architect at A. G. Edwards & Sons, Inc., and Edward Jones brokerage firms in St. Louis, MO, and 10 years in the design and architecture of communications systems and components for Erlang Technology, Tellabs, and Williams Telecommunications (WilTel).

Ken holds bachelor's and master's degrees in electrical engineering from Missouri University of Science and Technology.

About the Technical Reviewers

Linda Strick has been with the Fraunhofer Institute FOKUS in Berlin, Germany, since 1987. She works in the application domain of innovation and technology transfer. Her main areas of work are cloud computing, distributed systems, telecommunications, service-oriented architectures, security and privacy, cloud certification and pre-commercial procurement of cloud services, and e-government. She has been working with national and international projects in standardization organizations and has published several papers.

Dedications

Jazib Frahim:

I would like to dedicate this book to my lovely wife, Sadaf, and my two lovely and adorable children, Zayan and Zeenia, who have patiently put up with me during the writing process.

I would also like to dedicate this book to my parents, Frahim and Perveen, who have always supported and encouraged me in all my endeavors.

Finally, I would like to thank my siblings, including my brother Shazib and sisters Erum and Sana, sister-in-law Asiya, brothers-in-law Faraz and Nabil, my handsome nephew Shayan, and my adorable nieces Shiza and Alisha. Thank you for your patience and understanding during the development of this book.

Venkata (Josh) Josyula:

I would like to dedicate this book to my mother, Saraswathi Josyula, who passed away on February 23, 2013; to my wife, Dr. Leela Sai, daughter Deepa, and son Vikram, who work very hard every day and are an inspiration to me. I would also like to thank many friends at Cisco whom I look up to every day.

Monique J. Morrow:

This book is dedicated to my mother, Odette G. Morrow, who passed away on September 17, 2013; to my father, Samuel A. Morrow Sr.; Veronique Thevenaz; Andre C. Morrow; Samuel A. Morrow Jr.; and Michelle M. Kline. You have all been my shining lights! Thank you for your love and encouragement.

Ken Owens:

I would like to dedicate this book to my wife, Dr. Christine Owens, sons Kenny and Nathan, and daughters Carolyn and Abbey, who inspire me every day to be the father and man that I was created to be. I would also like to thank my coauthors and many friends at Cisco who have encouraged me.

Acknowledgments

Jazib Frahim: I would like to thank the technical editors for their time and technical expertise. I would also like to thank the Cisco Press team, especially Brett Bartow, Mary Beth Ray, Christopher Cleveland, and most of all Jeff Riley, who challenged his authors to create an impactful product in this book. Thank you!

Many thanks to our Cisco management team, including Bryan Palma, James Mobley, and Russell Smoak, for their continuous support. They enthusiastically encouraged us throughout this project.

Venkata (Josh) Josyula: I would like to thank my coauthors for their friendship, dedication, and willingness to complete this book. Also, I want to thank Sunil Kripalani, VP, Cisco Services, for reviewing earlier versions of my work on this book. Also I thank the technical reviewers and Cisco Press staff for editing this book.

Monique J. Morrow: I would like to thank my coauthors, Josh, Jazib, and Ken, for their valuable insights in developing this book and for never giving up! When I think about the vision of this book, I must call out Mike Geller and Guy Daley, who worked with me to develop key concepts. My gratitude goes to David Ward, SVP and Chief Architect at Cisco, for his commitment to my personal success and for instilling in me a self-confidence to be better. Finally, a special thank-you to the Cisco Press team and to our technical reviewers for shaping this book.

Ken Owens: I would like to thank the technical editors and the Cisco Press team. Their support, encouragement, and comments made this process seamless. I would also like to thank Faiyaz Shahpurwala, SVP Cisco Intercloud, and Biri Singh, CTO and Platforms, for their encouragement.

Contents at a Glance

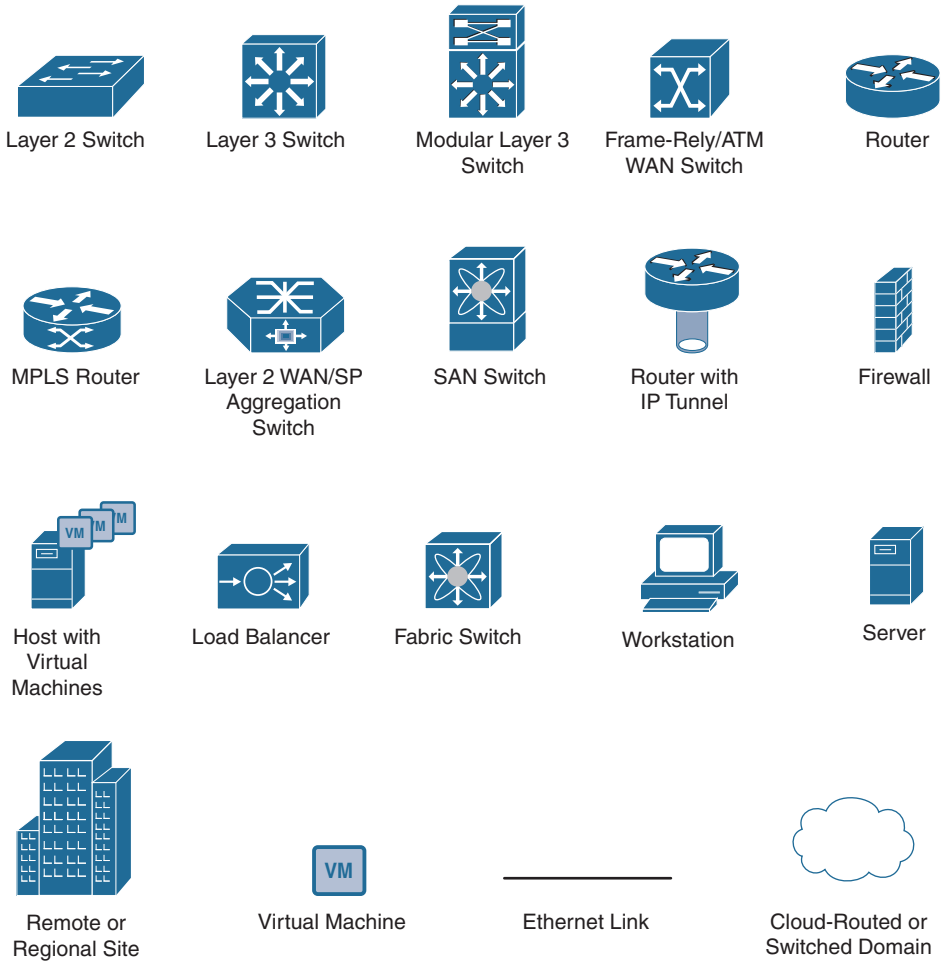
Reader Services

Register your copy at www.ciscopress.com/title/9781587144455 for convenient access to downloads, updates, and corrections as they become available. To start the registration process, go to www.ciscopress.com/register and log in or create an account*. Enter the product ISBN 9781587144455 and click Submit. Once the process is complete, you will find any available bonus content under Registered Products.

*Be sure to check the box that you would like to hear from us to receive exclusive discounts on future editions of this product.

Contents

Icons Used in This Book



Command Syntax Conventions

The conventions used to present command syntax in this book are the same conventions used in the IOS Command Reference. The Command Reference describes these conventions as follows:

- **Boldface** indicates commands and keywords that are entered literally as shown. In actual configuration examples and output (not general command syntax), boldface indicates commands that are manually input by the user (such as a **show** command).

- *Italic* indicates arguments for which you supply actual values.
- Vertical bars (|) separate alternative, mutually exclusive elements.
- Square brackets ([ ]) indicate an optional element.
- Braces ({ }) indicate a required choice.
- Braces within brackets ([{ }]) indicate a required choice within an optional element.

Foreword

The information technology (IT) industry is again on the cusp of a major transformation. Business digitization is a trend that is impacting companies across industry verticals, from healthcare to manufacturing to retail and more. Digitization, defined as the leverage of IT to improve the experiences of a workforce and customer base while streamlining operational processes, is bringing the opportunity for organizations to fundamentally transform themselves and create new lines of business and revenue.

To realize the promise of this new era of IT in which we live, companies will need to embrace new ways of thinking about their networks and overall IT strategy: machine learning, harvesting the data that will come from the billions of new devices to be securely connected via the "Internet of Things," and last, the megatrend that has arguably made the most significant impact on the IT industry since its inception: cloud.

Throughout my career, at both startups and Fortune 500 companies, I have experienced firsthand the positive business impact of moving to a cloud-first model. Cloud computing has forever changed how businesses think about building, consuming, and deploying their products. This book establishes a foundation for understanding the next transformation in cloud computing called Intercloud. Intercloud enables enterprises to interconnect and network between public clouds and private clouds to take the benefits of the cloud computing paradigm to a new level.

This book provides an in-depth description of the components of Intercloud for the business and IT professional and addresses the security concerns of businesses that leverage this new platform.

Zorawar Biri Singh
Cisco Systems
April 2016

Introduction

"Cloud computing" is a term applied to various computational and IT services provided on a "utility" basis, typically hosted in large geographically distributed data centers. The services provisioned within these data centers are highly automated and take seconds to deliver upon a subscriber's request. Additionally, these data centers typically leverage virtualization technologies, such as hypervisor-based virtual machines, as an abstraction technique to deliver cloud services. Within cloud computing there are service layers such as infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS) that can be used in a private, public, or hybrid deployment model.

Because of hardware commoditization and growing competition among cloud providers, enterprises are looking for cost-efficient providers to host their applications. In some cases, certain applications reside in one cloud provider while others might reside in a different cloud provider. This trend is leading the market to define an emerging "Intercloud" approach. The concept of a cloud operated by one service provider or enterprise interoperating with a cloud (or clouds) operated by another is a powerful idea. In recent years cloud interoperability has been limited to use cases where code running on one cloud explicitly references a service on another cloud. There is no implicit and transparent interoperability.

Furthermore, the external interfaces, failure domains, scalability limits, orchestration, assurance, and billing/remediation services used by cloud service providers are disparate and by design not interoperable or compatible in any way. While standards bodies are discussing some aspects of this domain, most of the gaps are not defined in these areas and are not being addressed at all. This is a concern for enterprises that are trying to provide agility and flexibility to their businesses.

Many cloud providers have built data centers all over the world. So why do we need Intercloud? As providers have built their cloud instances, it became apparent to them that no single provider can have the flexibility and capacity to provide everything consumers need. This has led to the idea of sharing resources among providers. This concept is not new, and it is similar to the cooperative relationships (peering) used by mobile providers. The fundamental idea behind the Intercloud deployment model is that resources in multiple cloud domains can be cooperatively engaged in a seamless and transparent way. The most important benefits of Intercloud are application portability, IT flexibility, reduced cost, and faster response to market and business needs. The Intercloud challenges faced at each of the service layers (IaaS, PaaS, SaaS) are different, but they can be addressed with consistent themes such as API standardization, workload definition and mobility, and network security. Solutions to these problems are actively being developed, and a nascent Intercloud ecosystem is coming into existence.

Today, cloud services are deployed via technology and business architectures at various "places in the network" by service/cloud providers and consumed by all type of customers, from large enterprises to small consumers. These "places in the network" include the customer premises, the virtualized network edge, and the data center/cloud.

Consumers demand a dynamic mashup of such services from different providers. In order to make such cloud-based services a reality, where different cloud providers provide different sets of services, user-to-network service contracts (UNCs)—similar to user-to-network interfaces (UNIs)—network-to-network user contracts (NNCs)—similar to network-to-network interfaces (NNIs)—security (which includes policy and identity), and management profiles for these early Intercloud services must be defined.

An alternative way to think of Intercloud is to picture a world where each cloud/service provider hosts an "exchange" where services are delivered from their respective clouds. A network is needed to connect these "exchanges." The NNC then defines the information and structure of the "peering" relationship between the two services and/or cloud providers. The UNC would then define how the set of services are presented to the customers who eventually consume them. The entire network of cloud providers becomes the next-generation e-commerce platform where services are bought, sold, and otherwise transacted.

This book provides a business and architectural advantage to cloud-based service providers and consumers in the Intercloud framework. Readers will obtain an understanding of the microservices architecture and overlay required for applications to be ported across the Intercloud.

In discussing Intercloud, there are certainly adjacent and/or contributing technologies such as grid computing, distributed computing, service-oriented architecture, and virtualization. The question for the reader becomes: How do these technologies fit together into overall cloud computing? Further, if we could imagine an open "Internet of clouds," what technologies, business models, and standards are required to enable such a paradigm? This book addresses these concepts, touching on the role of cloud's adjacent and contributing technologies, and focuses on the new technologies and business challenges creating the Intercloud.

At a high level, the plan takes the three primary cloud service layers (IaaS, PaaS, and SaaS) and adds to them a new dimension paralleling the evolution of today's networks (and network topology/connectivity paradigms). Intercloud exchange/marketplace, service fulfillment, service assurance, billing, identity, security, visibility, control, optimization, application peering, and federations are just a few of the many technology and business components to be analyzed in this book. Early use cases for the new utility include delivery of applications and services via a "public cloud." Also, "private clouds" are illustrated, where an enterprise extends its own data center dynamically in house. Hybrid clouds are explained, in which private clouds are extended through the services of a provider, eliminating the need to inefficiently build out new capacity for each application and thus leading to a peak utilization strategy.

The technical and business perspective on Intercloud is valuable, but it is somewhat incomplete without placing the Intercloud in the context of specific use cases. These use cases are based on the evolution of service delivery and consumption in today's multiprovider IT services market. We cover in detail the data models and business case for the cloud network-to-network interface (peering and cloud/service provider dynamics enabling connectivity between clouds at the various layers) and for the cloud user-to-network interface (service delivery models and dynamics describing the elements used to deliver services in a modular fashion to those who consume them).

Who Should Read This Book?

The primary audience for this book includes CTOs, IT managers, network managers, security architects, service provider product managers, application owners, and service architects who are responsible for assessing technology and architecture as a basis of secure service and solutions deployment. Industry analysts, focusing on telecommunications, hosting, and cloud computing, constitute a secondary audience for the book.

This is an intermediate-level book, albeit with an unusually broad technology scope. Readers should have a basic knowledge of internetworking principles and are assumed to know how basic TCP/IP and routing protocols function. Likewise, readers should be familiar with the basic building blocks of computing and storage, as well as fundamental security practices. Regardless of technical knowledge, a layperson will benefit from the background, technology overview, and business impact discussions that are provided in the book.

This book provides the reader with a complete overview of the nascent Intercloud, discussing both existing and developing enabling technologies as well as giving the reader an overview of adjacent technologies that support and lead to Intercloud. The reader will learn how to build Intercloud and provide services, linking a variety of cloud computing technologies and business model examples. The book provides many use cases, case studies, and a helpful description of emerging standards.

How This Book Is Organized

PART I: Introduction to the Current Cloud Landscape

You have to know where you've been to know where you're going. Part I delivers a brief synopsis of where we've been and how what is current is the initial step forward to the Intercloud.

Chapter 1: Cloud as We Know It Today

This chapter provides a brief review of current technology and business methods that are the foundation of the Intercloud. A cursory examination of the current cloud landscape provides a brief review of the material for the reader. This chapter may also act as a brief technical introduction, encouraging cloud service providers to deploy the foundational elements required to be active on the Intercloud, the next platform for business over the Internet. This chapter covers

- Cloud computing concepts
- Cloud taxonomy
- Evolution of compute
- Evolution of cloud computing to Intercloud

- Definition of Intercloud
- Definition of federation
- Definition of exchange
- Definition of broker

Chapter 2: Intercloud Architecture and Technologies

For the last couple of years, cloud providers have set up data centers at many locations to service their customers. For example, Amazon, Google, Microsoft, Salesforce.com, and others have established data centers for hosting cloud application services such as social networking, gaming portals, and business applications. In many of these social networks, the service components run in separate virtual machines that may be hosted in data centers owned by different cloud computing providers. In order to link the services provided by different providers at different locations around the world and to provide quality service to a consumer, it is necessary to build an Intercloud to bring the services together. This chapter details

- Intercloud architecture model
- Intercloud use cases
- Intercloud deployment model

Chapter 3: Intercloud IT Management Strategy

Businesses that thrive are the businesses that understand customer needs and deliver those needs seamlessly, meeting the QoS and SLAs. Services are the building blocks for IT service management, but how do you determine which services are the "right" services to develop and deliver to customers?

In this chapter, we discuss areas that are essential to deliver services and meet most customer needs in an Intercloud environment:

- Guiding principles
- Technology Business Management (TBM)
- Cisco lifecycle for IT services and management
- Fundamental service building blocks for cloud
- Intercloud services use case(s)

PART II: Evolution of Intercloud Management and Security

Evolution means that IT needs to manage workload portability while maintaining service levels, security, and governance in more complex and important cloud environments.

Part II of this book discusses details of the Intercloud exchange/marketplace, service provisioning, service assurance, billing, security, and platform APIs.

Chapter 4: Intercloud Architecture(s) and Provisioning

This chapter discusses intercloud architectures and provisioning using OpenStack:

- Cisco Intercloud strategy
- Standards-based cloud architecture(s)
- IEEE Intercloud architecture
- The Cisco vision of Intercloud
- Cisco Intercloud IaaS provisioning using OpenStack

Chapter 5: Intercloud Service Assurance

The previous chapter discussed how to provision a subscriber on Intercloud. Now, we have to maintain customer satisfaction by ensuring that the services promised are delivered using appropriate service assurance strategy and architecture and maintaining customer SLAs. This chapter discusses all of the service assurance aspects of Intercloud:

- Intercloud assurance strategy
- Intercloud assurance architecture
- Intercloud end-to-end monitoring flow

Chapter 6: Intercloud Accounting and Billing

The users connect to cloud from their end devices (smartphone, iPad, laptop, or other device) and enjoy video on demand, movies, voicemail, and other content. These services are not necessarily provided by the cloud providers and may be provided by the content providers. The Intercloud pricing, charging, and billing present challenges in terms of how the customer bill will be divided among the providers. The cloud provider most likely handles authentication, authorization, and accounting services and provides a single bill to the users. The Intercloud billing model will be handled by a collection of middleware that collects and processes the accounting data. This chapter provides details related to Intercloud services billing:

- Accounting and billing in the Intercloud
- Intercloud accounting and billing architecture
- Key performance indicators (KPIs)
- Business intelligence (BI)

Chapter 7: Intercloud Security

This chapter discusses best practices in security for cloud and Intercloud and their various implementations. In the context of the Intercloud, identity and security elements are discussed in detail. We examine the application of key industry groups like the Cloud Security Alliance and how the security models apply to today's public clouds, private clouds, and hybrid clouds that set the foundation for the Intercloud. We present a security framework that leverages two important security principles: asset/data visibility and security controls. We also discuss how the security framework can be used to build a comprehensive security program to address cloud adoption challenges. We take a deep dive into the security architecture of Intercloud and discuss the identity architecture of a federated Intercloud partner ecosystem.

Chapter 8: Cloud Operating Systems

This chapter discusses the function of a cloud operating system that can be browser based and includes a service abstraction capability. The chapter also addresses cloud APIs and their role in the cloud operating system:

- Cloud API landscape
- Standards bodies—their goals and current work efforts
- The "network up" view and APIs
- The "application down" view and APIs
- The Intercloud and application-level peering: introduction to the data model and API structure

Part III: Intercloud Case Studies

In Part III of the book we bring together concepts laid out in Part I and the evolution of Intercloud management details discussed in Part II to build Intercloud use cases. Also, we discuss challenges, standards, and data models.

Chapter 9: Use Cases for Building Intercloud Hybrid Clouds

This chapter introduces a case study based on cloud services; identifying the business model; and value proposition and challenges as the service evolves toward models delivering a dynamic service demanded by today's consumers and business people. Specifically, private clouds and hybrid cloud using OpenStack are discussed.

Appendix A: Intercloud Standards Organizations and Industry Bodies

Appendix B: Acronyms and Abbreviations

Chapter 2

Intercloud Architecture and Technologies

This chapter covers the following topics:

- **Cisco Intercloud Architecture:** This section describes the underlying architecture for the Cisco Intercloud solution. Each layer and component will be described and interfaces between components composed. In addition, this section provides two use cases to show how both enterprises and service providers can benefit from the Intercloud.
- **The Cisco OpenStack Platform and Underlying Infrastructure:** This section gives an overview of the OpenStack platform and the underlying infrastructure consisting of physical, virtual, and automation components.
- **Cisco Cloud Services (CCS) and Products:** This section gives an overview of the layer above the OpenStack platform and describes services and capabilities that are added by the Intercloud to enable advanced services around networking, security, network functions virtualization (NFV), data, database, load balancing, and application policy.
- **Cisco Application Enablement Platform as a Service:** This section provides an overview of the Intercloud's primary use case around enabling application portability and interoperability across a world of clouds.
- **Cloud Operational Support Systems (OSS):** This section provides an overview of the OSS functions and how they relate to the Intercloud's architecture.
- **Cloud Back-Office Support Systems (BSS):** This section provides an overview of the BSS functions essential for accounting, billing, and service fulfillment.
- **The Intercloud Marketplace:** This section introduces the curated set of products and services from Cisco and its Intercloud partners that enable enterprise customers to consume products and services in their enterprise catalog and across the global Intercloud product catalog from a simple, easy-to-use marketplace.

The Intercloud is a connection of global "standalone clouds," similar to the Internet which is a connection of "networks." The Intercloud is based on scenarios where a single cloud has no infinite physical resources or ubiquitous geographic footprint. If a cloud saturates the compute and storage resources of its infrastructure, or is requested to use resources in a geography where it has no footprint, it would still be able to satisfy such requests for service allocations sent from its clients. The Intercloud scenarios would address such situations where each cloud would use the infrastructures of other clouds. This is analogous to the way the Internet and telephony work, in that a service provider, to which an endpoint is attached, accesses or delivers traffic from/to a source/destination address outside of its service area with other service providers with whom it has a prearranged exchange or peering relationship.

Organizations are being asked to adapt to an increasingly competitive business environment to quickly create differentiation for the business. To capture the opportunity enabled by the IoE, companies need to consider how this impacts their current business model for customer reach and ease of deployment. Failing to adapt can have significant consequences; some forecast that 40% of Fortune 500 companies won't exist in 10 years. At each turn, organizations are asked to deliver faster innovation, drive revenue growth, and transform the business into a digital company. Every country, city, industry, and business is becoming digital to leverage the unprecedented opportunities brought about by the IoE. Becoming a digital business requires rethinking core business processes. It means implementing a fast IT and open-source technology foundation that supports both traditional enterprise and cloud-scale workloads. It means embracing new security, cloud, mobility, social, and analytics technologies to empower developers and the business to quickly and securely launch and evolve new services. This is the goal of the Intercloud.

The Intercloud is architected to enable standalone clouds to work as one; it involves not just connecting clouds but also accelerating development for cloud and unifying workload management across clouds. It ensures that network and security policies follow the workload, harnessing an expanded ecosystem of best-of-breed partners and service offerings. It takes advantage of global data while meeting local and regional requirements. Cisco has led a global market change just like this before. Just as Cisco did for the Internet by connecting and integrating isolated PC LAN networks, Cisco is doing the same for cloud.

To help lead organizations through this digital transition and IT transformation, Cisco and its partners are building the Intercloud, a globally connected network of clouds capable of delivering secure cloud applications and infrastructure everywhere in the world. The Intercloud will deliver choice with compliance and control to empower customers to innovate and transform their businesses. It is made possible through Cisco and its partners' leadership in cloud infrastructure, applications, security, open source and open standards, and consulting services and is built with Cisco Intercloud enabling technologies.

Cisco Intercloud Architecture

Although cloud computing has achieved broad acceptance for the promises it makes, enterprises are still having a difficult time adapting to it. There are several reasons for this, but the top concerns are security, increased complexity, inconsistent methodologies,

and traditional mindsets. Existing cloud offerings have failed to deliver solutions that meet customers' expectations and business objectives and achieve the agility and flexibility the enterprise requires. There are primarily two meta-issues with the current cloud deployment models. The first issue is that cloud does not equal IaaS, which has led to the increased complexity and mismatched expectations with the business. The second issue is that development methodologies are in the midst of a transformation.

Many cloud solutions on the market today are nothing more than Managed Hosting 2.0 with limited automation of traditional systems and processes. This situation causes the greatest concern with the adoption of cloud because it's very complex and requires cloud consumers to understand the following factors:

- The underlying compute capabilities with CPU/memory combinations/constraints and memory management/ballooning concerns.
- The underlying network capabilities with vNIC (virtual network interface card) configurations, limited server load balancing, very basic single-network private space, and public space tied to providers' IP address blocks.
- The underlying security capabilities—access lists, security groups—must be configured, and these are tightly coupled to a single deployment space and not easily replicated across sites or regions.
- The underlying storage capabilities must be managed and maintained separately, typically with different domain spaces; in other words, the object store is separate from disk (local, file, block).
- Consumers of the service must be able to adapt their IaaS architecture to adopting a proprietary code that will transform the code to a workload that operates a script. In its sum, the process can be labeled as a "blueprint."
- Consumers must adopt the user management (identity management) and security controls of their provider.
- Customers must accept the OSS aspects of their provider—ticketing, metering, monitoring.
- Customers must accept the billing and SLAs of their provider.
- Customers must accept the operational practices of their provider.

IaaS LEGO Masters

When considering the current state of IaaS service providers, one cannot help but think about boxes of LEGO blocks. There are so many different shapes, sizes, and colors, and kids get so excited building different creations. But their creations rarely look like what they tried to build. The other issue is that when they try to build the same thing again, they are not able to replicate the exact design. As kids grow older, they get better because they have experience; however, they still have trouble replicating the exact designs. Think of LEGOs as pieces of your business infrastructure; do you really want

to run your business on infrastructure that you have to piece together, that is hard to replicate, requires specialized resources with very specific domain knowledge, and requires detailed knowledge of each interconnecting system?

This is how many existing IaaS solutions function today. You have to be a master builder of the detailed complexities of cloud IaaS components, services, and blueprints available. You must understand how to decompose your services into consumable services while ensuring that the constraints and underlying infrastructure architecture are accounted for. In this current model, any issues are hard to identify because there are so many different components and interdependencies to keep track of, each with its own methodologies and lifecycles of support. The existing systems, processes, and polices do not translate to the provider systems and require you to create custom processes and accept additional risk. In addition, if you're not in lock-in to a single provider and you want to add providers, the ability to add IaaS solutions exponentially increases the complexity and risk.

The flaws in the design of the existing cloud solutions are a mismatch of expectations. Cloud providers offer IaaS building blocks and blueprints, whereas enterprises expect business objectives and business outcomes achieved with agility and flexibility of their services. Cloud is not the same as IaaS. IaaS is a platform for building cloud but not cloud itself. The platform must be based on scale and abstracting all cloud components (compute, network, storage, and security), services, OSS, and BSS components.

LEGO Kits

Continuing the LEGO theme, most application development methodology is written like LEGO kits in that there is a standard and repeatable process (SDLC, software development lifecycle) with repeatable capabilities and services, the ability to create application blueprints consisting of the application components, services, and objectives on a platform designed for scale and agility. With the correct abstraction of the platform, these applications can be deployed over and over again on various platforms, and the result is always consistent. If the result is not, the issue is easy to identify because the processes of application developers are consistent and the issue can be isolated to an application issue or platform issue. With the LEGO kit, the blocks still need to be assembled; however, the design and blocks were created with the desired outcome in mind to ensure that the completed product looks like the picture on the box. The great thing about the kit is that it makes building with LEGOs even more fun because the result is predictable and each step leads to the desired outcome.

Development Methodology Transformation

Most existing (legacy) enterprise applications are not written with an agile cloud-native methodology. Although many enterprise developers have already transitioned or are transitioning toward agile, most are still in a waterfall-agile combination type of development. Thus there is a major disconnect between the tools and processes available within the enterprise and various deployment tools and processes in cloud. The greatest impact has been in the methodologies that have not evolved to take into consideration

the change in deployment to cloud solutions. If you think about the typical methodology, developers develop on their local laptops that have been secured and locked down by the corporate IT department. The code then gets saved to a code repository, where unit testing is performed, and then pushed to quality assurance (QA). This QA environment is within the enterprise lab or data center environment. QA environments are always a scaled-down version of production, so the validation that happened in QA is not 100% guaranteed to translate to production. Once QA is complete, a change window is created and the code is deployed to production. Production is always within the corporate firewall and falls into the IT security policy domain.

Cloud computing completely disrupts this process. Development does not have to happen on corporate laptops any longer, although much still does. The tools and process have changed for easier development, but the security and governance process and policy aspects have not. QA is typically performed in an outside cloud environment; however, the tools and processes for that environment do not follow the same requirements as the QA environments had to follow before. Likewise, production deployments are typically outside the firewall today and do not allow connectivity back through the firewall. And if they do, holes are opened in the security perimeter that put the enterprise at risk.

This change in development and deployment has occurred while transitioning to agile. The industry is commonly in a transition state, so it should not be a surprise that development methodology is continually evolving. What is important to understand here is that the shift needs to be supported in a flexible and hybrid model to support legacy—waterfall, agile, and most importantly the move to continuous integration/continuous delivery (CI/CD)—as part of the cloud migration requirement.

The Intercloud Defined

The Intercloud is a new approach to business and cloud transformation. It is about the continuous delivery of reliable, available, and consistent services over the lifecycle of the application. It is about the business objectives being supported over the lifecycle of the transformation from legacy applications to agile development projects in cloud, to the evolution of applications to CI/CD and beyond. Several key tenets of the Intercloud exist to accomplish this:

- The software-defined infrastructure is fluid and ever evolving to support the ever-changing and growing demands of the business.
- Software-defined services support all aspects of the business application needs:
  - Application: parameters and configuration details of the application including integration, performance optimizations metrics, and service-level objectives (SLOs)
  - Integrated application platform: any underlying dependent or loosely coupled services the application requires (message queue, Doman Name System [DNS], memcache)

    - Network and security: network and security policies and the enforcement of those policy objectives
    - Data: being able to abstract data aspects of the application and provide analytics and business outcome learnings
    - BSS: being able to abstract the BSS aspects to enable complete abstraction to any existing BSS systems in the enterprise or Cisco Powered partner clouds
    - OSS: being able to abstract the OSS aspects to enable complete abstraction to any existing OSS systems in the enterprise or Cisco Powered partner clouds
- SaaS frameworks for Cisco applications and the broad range of partner independent software vendor (ISV) applications comprise broad sets of tools for multiple cloud solutions.
- Public cloud providers are important to support for existing use cases and use cases that make sense.
- Private cloud solutions will be in place for many reasons—compliance, performance, control, and so on. The key here is to support the enterprise in whatever private cloud deployment model it chooses.
- Enterprise is a special case of private cloud where the deployment model is the enterprise data center. Additionally, the enterprise model may want to burst to other deployment models, so managing the policies and processes of the enterprise becomes critical.

The Cisco Intercloud consists of connecting the following:

- Cisco OpenStack clouds to deliver a marketplace for Cisco SaaS, partner, and ISV applications, and IoE applications
- Cisco Powered cloud providers to deliver a platform for IaaS and a marketplace for Cisco SaaS, partner, and ISV applications, and IoE applications
- Public clouds
- Enterprise clouds connecting to the Cisco OpenStack, Cisco Powered providers, and public clouds with policy and control

Figure 2-1 shows how private enterprise clouds, service provider public clouds, partner clouds, and Cisco services and applications are connected to form the Intercloud. Cisco Intercloud not only provides additional capacity and other features required by enterprises, but it also helps service providers to aggregate, integrate, and customize the delivery of their cloud services to meet specific enterprise business needs.

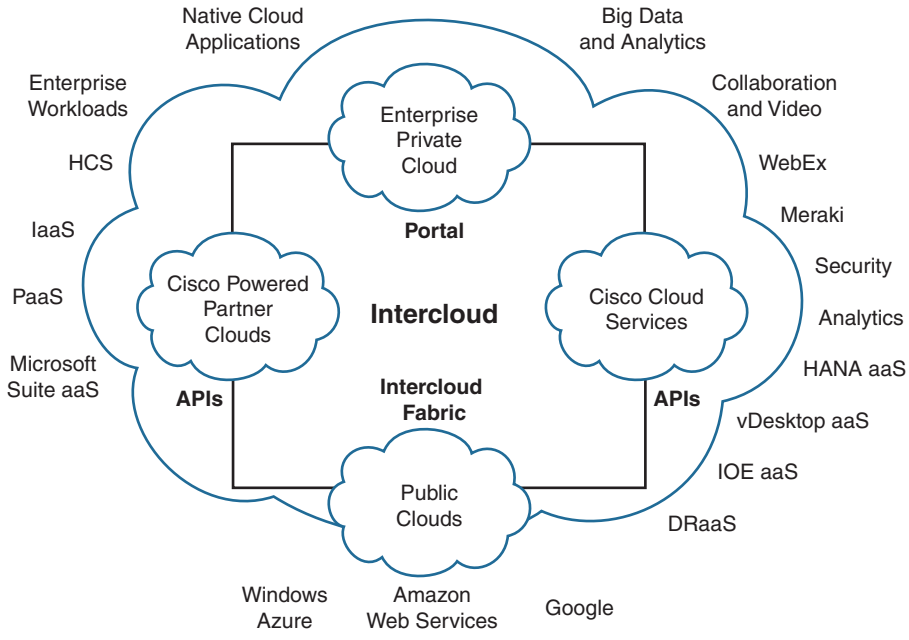


Figure 2-1 *Cisco Intercloud High-Level Architecture*

Cisco Intercloud Architecture Overview

This section provides a high-level overview of Cisco Intercloud architecture, shown in Figure 2-2.

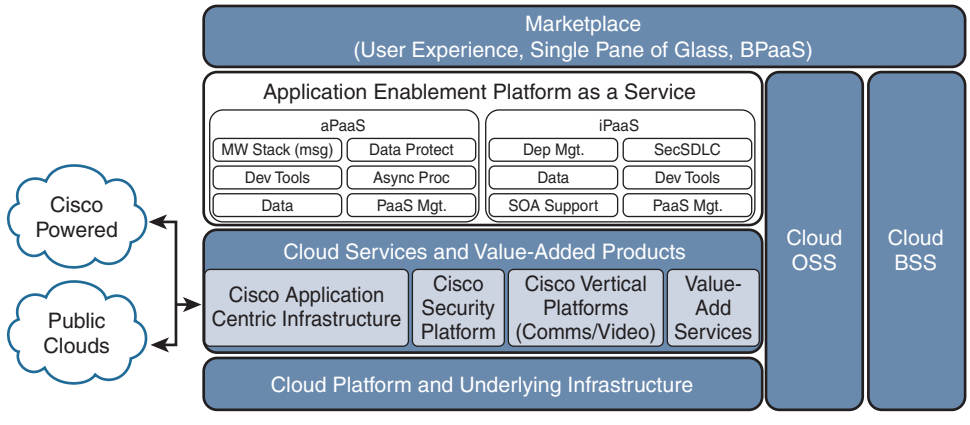


Figure 2-2 *Cisco Intercloud Architecture*

The Cisco Intercloud architecture consists of the following layers, which will be described in more detail in the sections that follow:

- Cloud platform and underlying infrastructure
- Cloud services and value-added products
- Application enablement platform as a service
- Cloud OSS
- Cloud BSS
- Marketplace

Each layer has dependencies and services available to the layer below and above it. This chapter will explore these layers and describe the Intercloud abstraction and interfaces.

The Cisco OpenStack Platform and Underlying Infrastructure

The OpenStack platform and underlying infrastructure consist of physical, virtual, and automation components. The beginning point for any service is the physical data center, controls, and hardware. Whatever services or virtualization you build on top of this infrastructure will only be as good as the underlying physical design and the security controls you integrate. The Cisco Data Center design[1] is a great reference architecture for what is required to achieve the correct level of resiliency and redundancy, including power and cooling for the underlying data center design. In addition, the data sovereignty and compliance regulations vary by geographic region and industry security concerns or last impact of loss (real or perceived). One of the core requirements of the Intercloud architecture is to address customer concerns about cloud adoption.[2]

The Intercloud adds another dimension to this physical design as the location of the data center matters for data sovereignty and security governance. In order to connect data centers from different locations securely, the network and physical connectivity between sites and locations is critical. This requires interdomain standards and cross-provider federation of control and network orchestration protocols.

The underlying platform is critical to consider as it is the foundation of the services that are built and delivered with high quality and reliability. The platform consists of physical servers to provide the underlying compute, memory, and local disk needed to support the infrastructure needs of cloud. The Intercloud node consists of high-performance CPU and memory with large quantities to support all compute needs, from small containers to large nodes to support big data analytics.

The network is another critical aspect that consists of virtual, data center, and WAN components. The virtual network exists within the compute domain. The domain can consist of a single node or group of nodes called a compute cluster. This network creates virtual interfaces that are specific to an organization and scoped to a project level.

The data center network connects all the local compute nodes, keeps all transit traffic within a single data center, and serves as the gateway to all external destinations (Internet, other data center networks).

The WAN connects the data centers together and accesses the Internet. This is the security boundary to the data center and also the intranetwork to connect all sites and other Internet locations. The Intercloud creates a control and management plane that integrates all networking from the local virtual interface of a compute node to the WAN connection across the Internet to another data center or service provider. The Intercloud provides the logical network constructs along with security groups and routing abstraction to enable the end user to create a global mesh of end-to-end connections.

Another important aspect is storage. Storage consists of a variety of different speeds and sizes of Serial Advanced Technology Attachment (SATA), Serial Attached SCSI (SAS), and solid-state (SSD) disks. These disks can be local to the compute node or pooled together in a SAN as a group of storage capacity. This capacity can be tied to different key performance indicators (KPIs) that govern the input/output operations per second (IOPS) or throughput that the SAN can handle.

In developing applications for the platform, the primary considerations are the open interfaces and how to enable these interfaces to remain up-to-date and consistent over time. By leveraging open APIs and interface abstractions, the platform presents an abstraction layer to the underlying infrastructure. This will allow the infrastructure to mature and innovate over time without requiring a rewrite of the application.

Best Practice While developing applications, it is a good business strategy to look at OpenStack as an OS for abstraction and standard open-source integration strategy. This is because OpenStack was written as a cloud operating system and not for an enterprise virtualization layer. This simply means that OpenStack is designed to run as a service, scale to multiple locations and geographies, and support legacy applications as well as cloud-native ones.

Another important strategy to consider is your open-source strategy. Open-source technologies play an important role in innovation and the development of solutions to meet the needs of the ever-changing and expanding marketplace. Open-source technologies can also be a risk to the business if your strategy is not well defined and thought out. The areas of specific concern are security, intellectual property, and support. It is important to have a strategy for open source and to review that strategy often.

Cisco Cloud Services (CCS) and Products

In the layer above the OpenStack platform, services and capabilities are added by the Intercloud to enable advanced services around networking, security, NFV, data, database, load balancing, and application policy. As you move up the stack, it is critical to focus on the API interfaces and capabilities. This layer is also where the Intercloud Fabric (iCF) product resides to enable point-to-point secure Intercloud connectivity.

Intercloud Network Products

The Cisco long-term vision of the Intercloud is a global network of cloud data centers, hosted by an alliance of service provider partners, that allows the collective customer base to run workloads or access valuable enterprise services anywhere across the globe. In order to achieve rapid adoption and increase customer loyalty, Cisco will enable its customers to reach these workloads and services in a secure, highly reliable manner. The users' experience will appear to be a borderless extension of their corporate networks. Additionally, Cisco will equip its network service provider (NSP) partners with methods to easily and securely connect to these services. This approach will allow them to build and offer compelling service bundles to their customers and is called PrivateLink.

PrivateLink is a collection of secure network connection capabilities that provide customers with this dedicated, private, secure connectivity into Intercloud regions without the need to configure or manage any VPN technology in their on-premises networks. This makes it a perfect solution for customers with a large number of enterprise devices or users needing access to their cloud environment. It also gives cloud consumers a way to provide highly available, high-quality connections to their services running in Intercloud regions around the globe.

PrivateLink will play a major role within our broader connectivity strategy of private peering between Alliance Partner networks and creating a seamless mesh of Intercloud regions across the globe. Figure 2-3 shows this capability: a future customer attaches into an Alliance Partner network with PrivateLink and can reach private workload space in any region across the Intercloud.

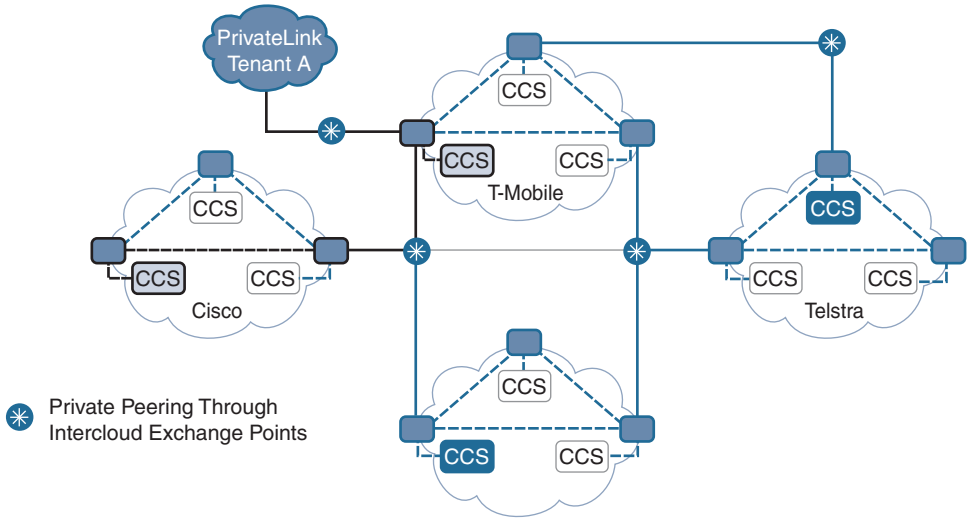


Figure 2-3 *Cisco Intercloud Alliance Partners Connected via PrivateLink*

Intercloud Security

Securing the Intercloud is critically important as security is still the top concern for adoption of cloud services. Additionally, when one considers the global nature and interoperability of the Intercloud, the risk and threat profile warrant extra focus. Intercloud security consists of three layers: foundational security, security services, and security as a service.

Foundational security consists of controls at the infrastructure level to protect the provider and each tenant with a default level of security, segmentation, and access controls. The typical controls in this area consist of identity and access controls as well as network security groups and segmentation. This level is based on industry standards for security and segmentation.

Security services are products that are offered on top of the foundational baseline that enable customers to provide security controls for their deployment. These services typically consist of firewall, intrusion detection, encryption, and other security capabilities that customers need to protect their deployments and applications.

Security as a service consists of security capabilities that are provided in consumption and pay-for-use models as a service. These services are managed by the provider of the service and usually are called “managed services” since the provider manages and provides the security services at an additional cost to its customers through a subscription model.

Intercloud Network Functions Virtualization (NFV)

Networking and the importance of networking are diminishing as the industry considers the x-as-a-service model. This is primarily because of the complexity that network configuration can seem to inflict when the x-as-a-service model extends to public clouds. This has been by far the most-overlooked aspect required for the cloud computing model and is more important than ever before. The network is more critical to application performance and more importantly the customer experience than most any other aspect of cloud. To say the network does not matter is like saying that application behavior and customer experience do not matter. These aspects are intrinsically tied together. The network aspects of the application being delivered need not be overly complex or be explicitly defined at runtime. Application owners or developers do not have to know anything about the network, but they need to be able to define policies and business objectives about the performance, governance, compliance, and data classification.

Recently, SDN and NFV have been receiving attention to address the application performance and customer experience concerns posed by flat network and disconnected network approaches.

Another area that has been evolving over the last couple of years is NFVs. These can vary from network appliances to routers, to firewalls, to content delivery network (CDN) caches. The primary idea is to enable data to be processed in a virtualized manner that is similar to the physical network functions that used to be leveraged in the data center

and service provider. These services can be leveraged in the Intercloud in several ways. The Intercloud Marketplace will provide a way for enterprises to purchase their NFV solution from a marketplace of Cisco partners and industry-leading solutions.

Enterprises that have their own NFV solution of choice and licenses can upload their NFV solution into the catalog and either enable the NFV solution as part of their project or make it available for certain projects to leverage. Last, the Intercloud will enable NFVs as a service within the Intercloud itself. In this model, Cisco will manage, operate, and support the NFV use cases.

Best Practice The networks that meet the business objectives and policies should be well understood by the infrastructure team and virtually or logically defined well in advance of the application development process. This is the optimal use case for SDN. Once the infrastructure team defines the networks and extends them to cloud, the application will be able to run on that SDN and have its business objectives met.

Network services like server load balancing (SLB) with the ability to support elastic scaling of the application services based on application policies are a critical part of the Intercloud's ability to enable better networking and performance for enterprises leveraging the Intercloud.

In many cases, to meet security and compliance requirements, firewall, IPSec (VPN), or advanced services like an intrusion detection/prevention system (IDS/IPS) or web application firewall (WAF) are required. This is a good example of NFV services that the application can leverage over the allocated network by definition of the business policies. Some tuning may be required to take full advantage of these NFV services; however, the general business rules can be deployed as the baseline policy.

Setting up initial and additional access to cloud is a major issue for all cloud deployments. Being able to connect to the "public" network and have your same policies, networks, and access controls in place can take weeks to months. To address this issue, the Intercloud creates an initial direct connection to the "public" network. In addition, the ability to create multiple network segments and leverage your existing private (RFC 1918) space as well as your existing public IP space is supported by default. Extending your internal enterprise network to cloud is as easy as setting up your internal network segment on the Intercloud and connecting it to your internal network through a software-defined gateway and overlay network.

With the Intercloud, Cisco has created a new demilitarized zone (DMZ) for enterprises that lets them securely interconnect from their internal network ("behind the corporate firewall") to the external networks ("public"). This new DMZ is called the cloud edge gateway and provides a secure interconnection between the enterprise data center and the Intercloud.

Intercloud Cloud Services

Intercloud business models are primarily based on three uses cases:

- Infrastructure that is ubiquitous
- Basic network, compute, and security services
- Value-added services that allow the admin or developer to enable applications seamlessly across the Intercloud

Services can be numerous, and dependencies are important to monitor. Many applications require external services, software components, or enhancements to enable their ecosystem to grow and adapt to changes in market conditions.

Enhancements can be in the form of ecosystem partners or industry-leading services (resell or as-a-service models), or developed as part of the service offering. This section will describe the primary use cases that are offered as a service today and how to evaluate their effectiveness and usefulness.

Data services are often an overlooked area due to the complexity of the requirements to manage them. This is a major focus of the Intercloud as it is the core platform that enterprises require to manage their application content and perform analytics to enable business decisions. The Intercloud model provides a ubiquitous infrastructure that consists of the basic compute, network, and storage services that infrastructure as code requires. In addition, the ability to add value-added services and to interconnect across various partner and public clouds to allow flexibility of the services to be consumed where they make the best business sense is preserved.

Database as a service is one such service that makes a great use case for value-added services.

Database as a Service Use Case

Databases are critical to all enterprise services, and today the need to scale out data with NoSQL databases is important for enterprises. The Intercloud supports several deployment scenarios when it comes to databases, the primary one being deploying into the project with the application. This keeps the application and database dependencies related to each other in the same project. The other scenario is to deploy in a hybrid manner where the database is in the enterprise behind the firewall and the application is in cloud. In this model, it is important to consider the networking as well as data latency requirements between the database and application(s).

Trove

Trove is an OpenStack project that provides a user interface (UI) and a command-line interface for periodic maintenance of the database, such as scheduling backups, restoring data from backups, automatically upgrading minor versions of the database software, setting up replication, and so on. The UI is a common front end for managing different databases in the back end. It should also provide customers the ability to monitor and

send notifications when a certain database threshold is exceeded, such as maximum number of connections, storage filling up, and long-running queries, to name a few.

Trove provides a scalable and reliable cloud database-as-a-service provisioning functionality for both RDBMS (relational database management system) and nonrelational engines on top of OpenStack. Figure 2-4 shows a reference architecture of Trove.

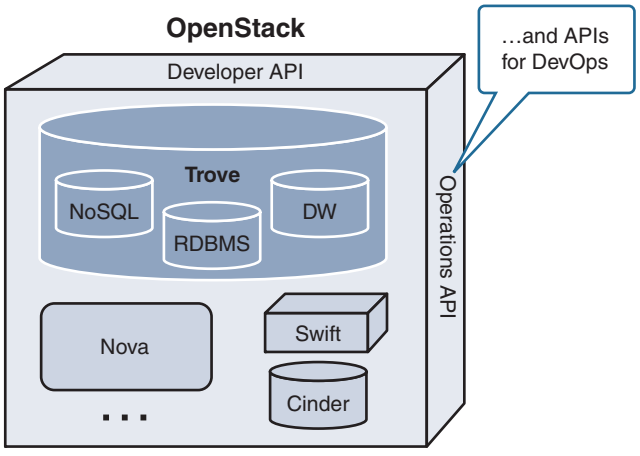


Figure 2-4 *Trove Reference Architecture*

Trove interacts with various OpenStack components for provisioning a VM to set up databases; see Figure 2-5 to better understand how it interacts with OpenStack components.

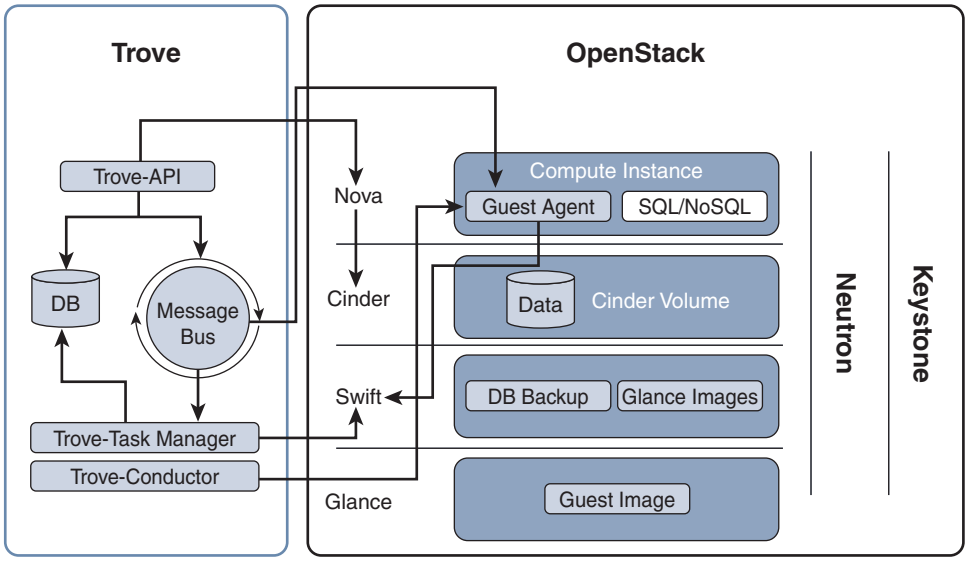


Figure 2-5 *Trove and OpenStack Interaction*

Trove itself is based on a share-nothing messaging system, like Nova. Its components communicate over a message bus and can be run on different servers. It behaves very similarly to Nova in that you send a message over HTTP, that message is translated and sent over the message bus, and actions happen asynchronously. It is currently composed of the following major components:

- API server
- Message bus
- Task manager
- Guest agent
- Conductor
- Scheduler

Figure 2-6 gives a pictorial view of the components and an explanation of each of the components.

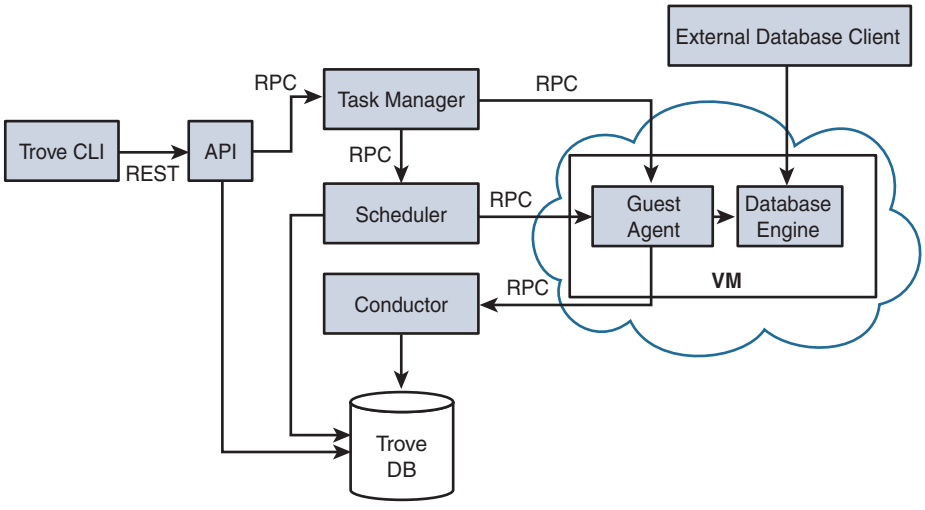


Figure 2-6 *Interactions Among Various Trove Components*

Intercloud Application Policy

A critical architectural component of the Intercloud is the ability to manage applications across a global connected world leveraging business policies. The Intercloud's application policies enable enterprises to have a complete cloud strategy, allowing application visibility and control across private and hybrid clouds for both legacy and new applications. These applications can then be deployed securely and compliantly across hybrid clouds with application awareness, dynamic lifecycle management, and real-time

automation. Through policy-driven management that is independent from infrastructure or systems management, the application-driven policy platform puts the application or business owner back in control.

This platform was developed to bring telco-grade reliability and trust to the data center. It enables IT to deliver the intent of the application anywhere via application-driven policies that are dictated by business needs of providing an abstraction layer between business objectives definition and the enforcement of policies within the application-centric infrastructure, as well as to introduce increased resiliency, providing a clear handoff between developers and operations.

Figure 2-7 presents the intent model as a set of configuration, fault/performance, security/governance, and accounting SLAs.

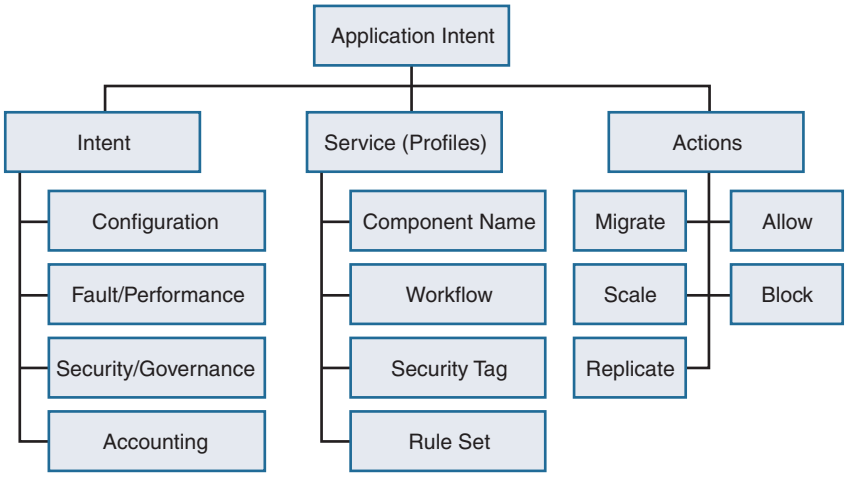


Figure 2-7 *Application Policy Intent Model*

To break down the application intent, it is best to separate the Intercloud innovation around application intentions from the existing policy models (TOSCA, OpenStack Congress, Group-Based Policy, and so forth). In the definition here, we introduce the business goal of sensitivity. Sensitivity is defined as the degree to which the performance and response time of the application influence the end users' perception of the application's performance. It is best to consider it a scale of no sensitivity to high sensitivity that can be adjusted in real time by the perceptions of the performance being measured by the system and end users.

The application intent sensitivities are defined as follows:

- Compute
  - CPU sensitivity
  - Memory sensitivity

- Storage
  - Latency sensitivity
  - Volume sensitivity
- I/O
  - Latency sensitivity
  - Throughput sensitivity
  - Thresholds (optional numerical value—for example, 80 connections/sec)
- Fault/performance
  - Recovery sensitivity
  - Availability sensitivity
  - Scale sensitivity
- Accounting
  - Cost sensitivity

Given these sensitivities, the Intercloud policy system can create an SLA for the business objectives defined here. In addition to these sensitivities, there are attributes that the policy system needs to be aware of. The first one has to do with dependencies:

- Services
  - Service affinity
  - Service anti-affinity
  - Security policies (data classification)
- Placement policies
  - Host affinity
  - Host anti-affinity
  - Availability zones
  - Regions
  - Geography
  - Constraints—noncoexistence

The second has to do with limits and understanding the constraints on the policy:

- Metering limits
  - I/O
  - CPU

- Memory
- Connections per second

- Security governance
  - Organizational constraints (IT, human resources, legal, engineering)
  - Data type constraints (public, sensitive, confidential, top secret)
- Operational constraints
  - Encryption
  - Auditing
  - Log retention

Given the sensitivities, dependencies, and limits, the developer can set the initial application intent, measure the performance of this initial intent, and make changes based on the actual performance. This is one aspect of the Intercloud that is important to understand, as deploying to any environment based on performance, compliance, data sovereignty, and internal enterprise security policy concerns is a first-order problem that the Intercloud addresses.

Cisco Application Enablement Platform as a Service

Every company is becoming a software company. The Intercloud's primary use case is to enable application portability and interoperability across a world of clouds. By virtue of its broad definition, one can throw in various solutions under the PaaS umbrella. For our purposes we will define what PaaS encompasses by looking at what customer problems need to be solved and working backward from there. Our target audience in terms of usage of the PaaS solution on the Intercloud is the enterprise developers and enterprise IT engineering teams on the customer side; the set of decision makers for adoption of a particular PaaS offering will have a mix of technical and business roles. Figure 2-8 shows the traditional differences between infrastructure and PaaS.

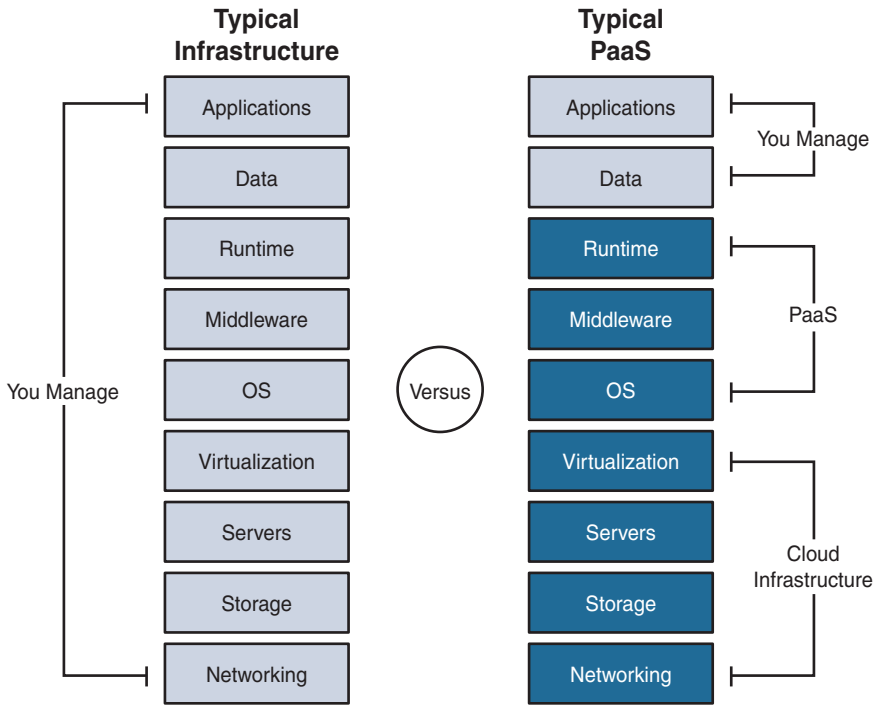


Figure 2-8 *IaaS and PaaS Differences*

We will start by looking at what problems the first group faces and then move on to selecting the PaaS components that will solve those problems. The following three phases clearly distinguish between PaaS and IaaS and how Cisco Intercloud overcomes many of the issues:

- Phase 1 is the development phase, which is when the developers and testers develop, build, deploy, and test the application in their local environment or a preproduction version of the cloud environment they plan to use. One friction point here is that they need to be able to consume the various cloud components from their application code, but doing so involves repeatedly writing code that should ideally be abstracted away in a utility layer. Another pain point at this stage is that the developers need to have confidence that their application that worked correctly in the test environment will continue to do so when deployed to the production environment. The Intercloud addresses these problems by packaging OpenStack APIs (the framework under which the Intercloud is standardized) that let an application request, consume, and release Intercloud resources in software development kits (SDKs) built specifically for the common languages used by customers (for example, Java or Python). The second problem of ensuring consistency through the various deployment stages is best solved by providing a facility for customers to containerize their application so that it deals with the underlying infrastructure only through interfaces previously agreed on between the container and the infrastructure (in this case OpenStack).

- Phase 2 (deployment phase) is the deployment of the built and tested application through various stages, culminating in the production cloud environment. This is similar to the previous phase where developers require consistency in the behavior of their application but in addition are interested in having a deployment experience that is not cumbersome, error prone, and manual in nature. An ideal solution here will provision the required cloud resources, wire them up as needed, deploy the right containers, and provide the customer a report that lists the deployment status, the resources being used, and the endpoints to the applications. All of these would be driven via preconfigured settings to ensure a high level of automation.
- Phase 3 (support phase) starts after the deployment is completed and the application is supporting production traffic (or workloads for an offline application). During this phase it is important to ensure that cloud resources are scaling in an elastic manner, key health parameters are monitored and alarmed upon, and the engineering team is able to quickly fix issues by deploying patches or new versions with minimal downtime. The scalability aspect can be handled by leveraging the autoscaling functionality in OpenStack and ensuring that it works with the container format under which the customer has chosen to deploy. When it comes to monitoring and triggering alarms, the customer should be able to define metrics around system resources (such as CPU usage) and application-specific measurements (such as average latency when communicating with another node). This requirement means there must be a monitoring service that customers can easily integrate into their application and can be viewed in multiple ways (for example, a dashboard for a high-level view, logs for deep diving). Last, the chosen container format should be able to deploy updated containers to a production environment with zero downtime to the service. This monitoring is key to application intent enforcement and improved performance of the application.

Integrated Platform as a Service (iPaaS)

Integrated platform as a service (iPaaS) is defined as components of cloud services enabling development, execution, and governance of integration flows connecting any combination of on-premises and cloud-based processes, services, applications, and data within individual or across multiple organizations.

The primary use case here is to integrate the enterprise PaaS, public PaaS, and application deployments into the development lifecycle. The Intercloud will enable integration between select partner offerings Cisco wants to expose to its customers as “first-class” PaaS offerings and will deeply integrate them into the Cisco ecosystem, including giving them a presence in the Horizon dashboard. An example of this would be running the Red Hat PaaS product OpenShift service in a hosted model that doesn’t require customers to have a dedicated OpenShift controller and will let them choose from a list of “cartridges” that Cisco maintains in its library.

In this same scenario, if customers prefer a single-tenant PaaS controller, they will always have the option to buy OpenShift from the Cisco Intercloud Marketplace for their exclusive use. Another example would be Cisco maintaining its own “hub” and container

repositories for the Docker container format for customers to build and manage their applications. By pursuing this option, Cisco also has the ability to integrate the PaaS solutions to provide a seamless experience for its customers (for example, the Cisco hosted OpenShift can easily fetch containers that a customer owns). When Cisco sees an opportunity to differentiate (such as a unified dashboard for all "first-class" PaaS offerings), it is able to build that and market it exclusively to its customers. It is also worth noting that if a partner's solution is not exposed as a deeply integrated solution on Cisco Cloud Services, the partner is still able to reach the Cisco customers by onboarding on the Cisco Marketplace; this provides Cisco customers with options that suit their workload and interoperability needs.

The availability of competing PaaS solutions in the Cisco Marketplace will also help Cisco maintain its leverage with the vendors it has chosen to supply its "first-class" PaaS offerings. The downside of this option is that it requires Cisco to get buy-in, at a preferred price point, from the partners it wishes to integrate. This can also lead to Cisco having to build and operate parts of the PaaS when it simply cannot find a good external offering.

Application Platform as a Service (aPaaS)

Application platform as a service (aPaaS) is defined as a category of cloud computing services that provides a platform allowing customers to develop, run, and manage applications without the complexity of building and maintaining the infrastructure typically associated with developing and launching an application.

The primary use case here is derived by understanding that the Intercloud is trying to become the cloud-of-clouds platform and the IoE platform. To enable this platform, a curated experience with all the PaaS capabilities will need to be created. Cisco will create a container format that works on all OpenStack environments in a consistent manner and will provide developers a set of SDKs to build against and package toward this container. Cisco will then augment the capabilities of OpenStack's orchestrator, Heat, to deploy these containers to cloud. Next, Cisco will build and operate a service that works in conjunction with Nova (OpenStack's compute module) to ensure autoscaling of the underlying resources. Finally, Cisco will have a homegrown monitoring service that will look for deviation from preconfigured steady states in addition to executing automated healing steps. This is similar to what Amazon Web Services (AWS) has in the form of Elastic Beanstalk, EC2 Auto Scaling, CloudFormation, and CloudWatch.

The benefits of this option are that Cisco can fully control the customer experience, iterate fast on bug fixes and new features, and work with otherwise incompatible systems (for instance, integrate with Cisco NFV modules). Furthermore, by virtue of offering innovative features specifically for its enterprise customers, this option also enables Cisco to differentiate from other cloud offerings more prominently. The downside here is that this will take a larger monetary and time investment to bring to market. Furthermore, this will make it harder for Cisco customers to realize the Intercloud benefits by making it difficult to deploy their applications to other clouds, both private and partner, unless Cisco makes a concerted effort to ensure that its PaaS is available in those clouds.

Use Cases for PaaS

The following use cases for PaaS are covered by the Intercloud:

- **Cloud native:** Most developers are looking at the role of cloud-native development in the next generation of enterprise applications and mobile devices. This model does not mean that all of your development is done in public clouds; it is more about how a developer approaches code development, deployment, and runtime. Containers are becoming very critical to cloud-native developers because of the abstraction layer they provide, and the individual services greatly simplify development. This use case is more likely to use an aPaaS framework as it is more aligned with developers writing to application platforms rather than the underlying infrastructure.
- **Cloud valid:** Several enterprises are on a journey to cloud by enabling many web application and public-facing services. This use case has more to do with the legacy applications in the enterprise and determining which applications should be exposed in clouds and which ones cannot, at least not yet. The primary reasons for not having those applications in clouds mostly have to do with the internal OSS/BSS dependencies, database dependencies, or data security concerns. The Intercloud enables this use case by iPaaS, which enables enterprise customers to leverage the infrastructure APIs in the enterprise and in cloud to create applications that can be deployed behind the firewall or in clouds (public or private), thereby creating a hybrid cloud use case where "hybrid" is defined as application components deployed in the enterprise data center and cloud connected securely.
- **Hybrid DevOps:** Hybrid DevOps is the Intercloud IoE platform for enterprises to start developing today for internal and external deployment patterns that over time will move more and more toward cloud native or the next-generation development pattern (fourth generation of PaaS). DevOps is an enterprise software development phrase meaning a type of agile relationship between development and IT operations. The capabilities of the Intercloud for the hybrid DevOps use case are displayed in Figure 2-9.

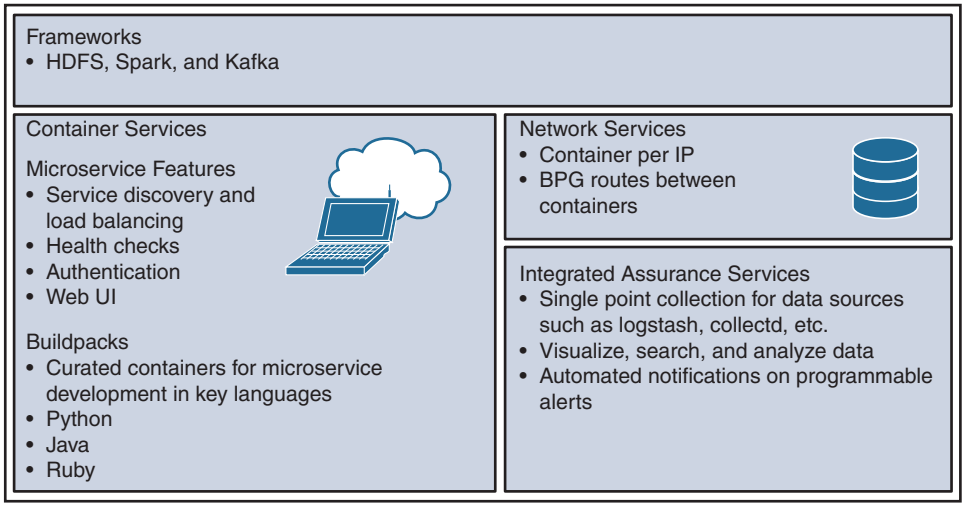


Figure 2-9 *Intercloud Hybrid DevOps Capabilities*

Cloud Operational Support Systems (OSS)

The OSS consist of the following management aspects:

- Change management means managing break/fix and new features to the system following the change process and change windows.
- Configuration management means managing the configuration.
- Capacity management means managing the capacity needs of the physical infrastructure.
- Performance management means managing the expected performance of the systems involved in delivering the service.
- SLA management means managing the SLA thresholds set by configuration and performance management.
- Incident management means managing incidents as they occur, documenting results, and creating tickets.
- Problem management means managing incidents that happen frequently or cause major downtime.
- Service desk is the name for the support staff available to provide assistance.
- Log and event management means managing all the logs and events from the data sources.
- Reporting and analytics means analyzing the events, logs, and data generated by the services to look for incidents or problems and to enhance the performance of the system.
- Service delivery means working alongside the business to deliver the service required.
- Image/lifecycle management means managing software images and patching these images.
- Asset and license management means managing the assets and licenses of the underlying hardware and software.

In the Intercloud, Cisco provides all these elements of OSS as a service. For partners and enterprises that already have OSS capabilities in place, Cisco provides an API into the Intercloud OSS for consumption by the partner/enterprise OSS systems.

Ceilometer

OpenStack has a project for monitoring called Ceilometer. Ceilometer has for several cycles failed to accomplish the performance and scaling that several OpenStack projects require. Ceilometer value can be divided into two major components:

1. OpenStack data collection through notifications (events) and polling
2. Metering API

Monasca

Monasca is the de facto OpenStack project for monitoring. It is not an OpenStack project yet but is targeted to become one at the Liberty release of OpenStack. Monasca strengths are quite complementary to Ceilometer since it has a very solid and scalable storage architecture as well as inline alarming, and it will soon be integrated with Heat, providing a superior solution to Ceilometer for autoscaling. Monasca, though, does not have a full suite of agents and currently does not integrate with events from OpenStack (there is a plan to leverage StackTach V3).

The Intercloud requires a new standard for OSS called Ceilosca, which is a combination of Monasca and Ceilometer. Monasca is really seen as a storage driver for Ceilometer, and its value is in its scalability. Ideally Monasca should be considered a "black-box" component of Ceilometer. Ceilometer will push meters and lately events to the Monasca API. Monasca will store them and provide them back through the same API. Ceilometer will consume meters and samples from the Monasca API and convert them back in the Ceilometer format to be served through the Ceilometer API as shown in Figure 2-10.

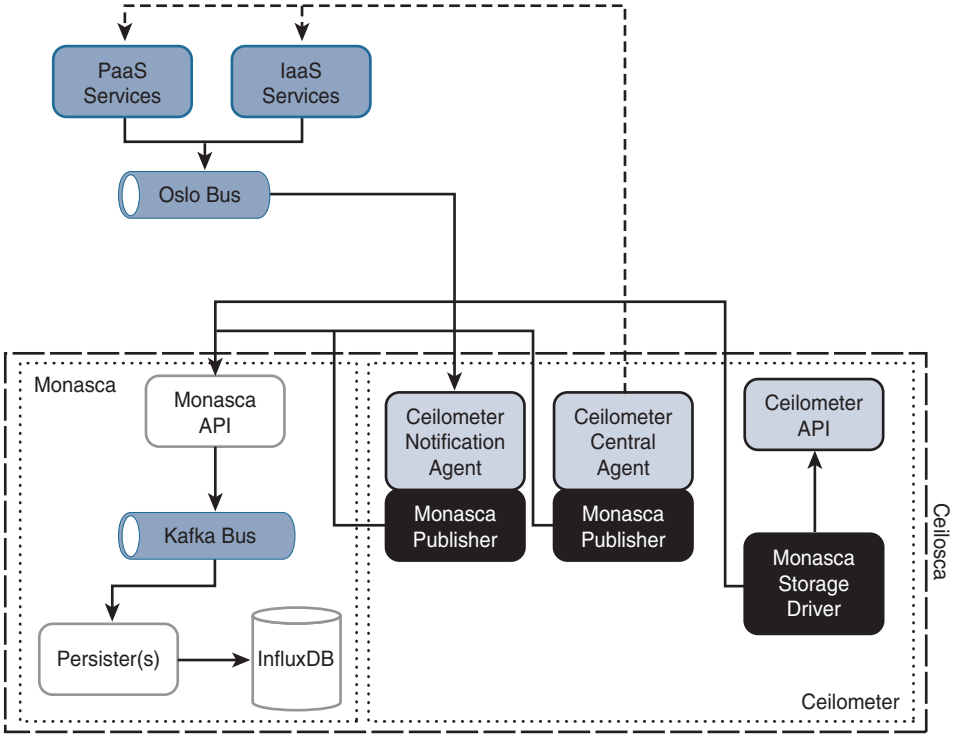


Figure 2-10 *Cisco Intercloud Monasca API with OpenStack*

The system now presents two APIs that can be used to POST or GET data, and they can be interchangeably used to store and retrieve monitoring and metering data. This not only

allows Cisco to store all the data in a single scalable data storage unit but allows other parties to communicate leveraging standard and open APIs.

Intercloud Monitoring

Tenant resources that are in the Cisco and/or partners' Intercloud need to be monitored across clouds, regions, and data centers. The monitoring agents in the partner clouds can POST the relevant subset of the monitoring data to Cisco clouds for the tenants that have instances running on those clouds as well as POST metrics to their internal monitoring system. Similarly, they can consume monitoring data from Cisco clouds if needed for their monitoring solution. Figure 2-11 displays the relationships for Intercloud monitoring.

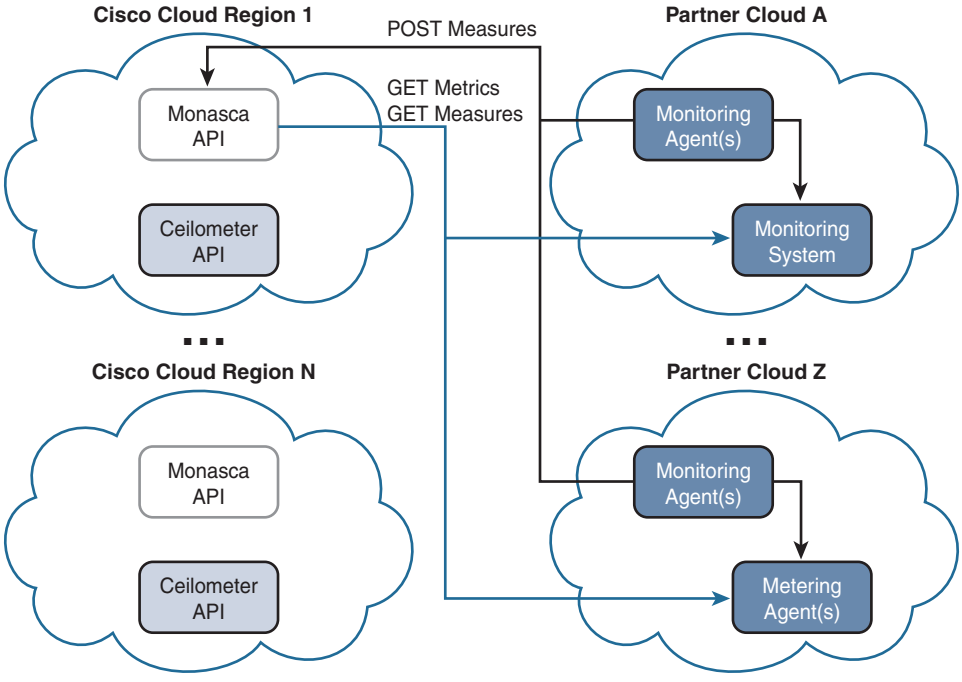


Figure 2-11 *Intercloud Monitoring*

Intercloud Metering

Similarly to the monitoring solution, Intercloud partners can POST samples to the Ceilometer API (and potentially the Monasca API) and collect the metering data and samples. Partners can use data collected from Cisco clouds to provide an overall view of resource usage to tenants belonging to partner clouds as well as perform show-back/charge-back operations, as can be seen in Figure 2-12.

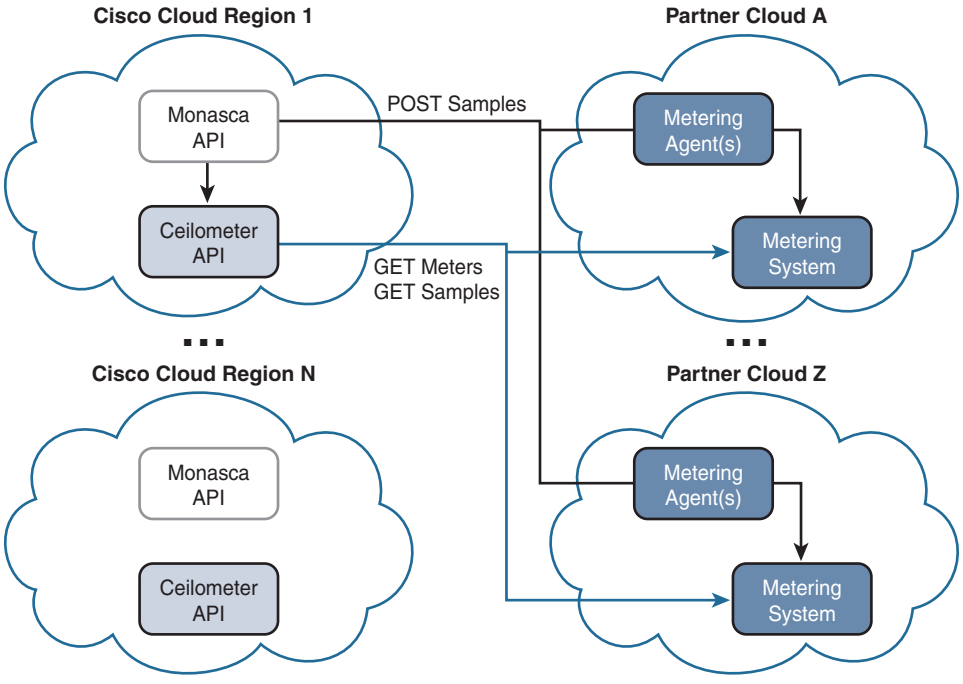


Figure 2-12 *Intercloud Metering*

Cloud Back-Office Support Systems (BSS)

The Intercloud provides the following to BSS for several management functions:

- Accounting provides data for managing the user accounts and departments.
- Contracts and Agreements manages the contracts and support agreements.
- Service Offering Catalog selects what offerings are allowed in the catalog.
- Pricing sets the price for the service.
- Provisioning provisions the services and sets the BSS parameters.
- Orchestration supports the dependencies and order of provisioning services.
- Subscriptions manages the service subscriptions.
- Metering provides usage data.
- Entitlements sets user access levels.
- Order manages the quote to the order fulfillment process.
- Billing creates a single bill and shows the detailed breakdown of the service chargeable components.

- Rating allows the cloud provider to discount the service.
- Clearing and Settlement provides all billing remediation and settlement services.

For partners and enterprises that already have these components in place, Cisco provides an API for easy consumption of these services.

Intercloud BSS Hardware Requirements

The BSS services require the following components, which are intended to run on standard hardware. Recommended hardware configurations for a minimum production deployment are as follows for cloud controller nodes and compute nodes for compute and the image service, and object, account, container, and proxy servers for object storage.

- The Cloud Controller Node (runs network, volume API, scheduler, and image services) requires a 64-bit x86 processor with a minimum of 12 GB RAM and 30 GB disk.
- Compute nodes (running virtual instances) require a 64-bit x86 processor, 32 GB RAM, and 30 GB disk.

Intercloud BSS Software Requirements

The BSS software requires infrastructure, integration, and existing business support enhancements. The Intercloud addresses BSS requirements from the infrastructure and integration of services provided by Intercloud Services, Intercloud Partner Services, and the Marketplace application. The API interface is provided for any enhancements required on the existing BSS systems that support the business.

The Intercloud Marketplace

The Marketplace is the curated set of products and services from Cisco and its Intercloud partners that enables enterprise customers to consume products and services in the enterprise catalog and across the global Intercloud product catalog from a simple, easy-to-use marketplace.

The Marketplace performs the following operations:

- Manage online user ID setup and user directory
- Manage identity authentication and secure site sign-in
- Manage an online directory of applications
- Manage or facilitate the purchase/consumption process
- Meter usage of applications
- Send and receive payments
- Manage the customer feedback system

In performing these operations, the Marketplace may need to use the following:

- The Cisco Infrastructure Services to set up VMs with a machine image provided by the application publisher and communicate via API with application publishers
- Marketplace Platform (Exchange) to facilitate financial transactions
- Service Exchange Delivery Platform (SDP) or Parallels Automation to facilitate OSS/BSS
- Identity Provider (IdP) to provide single sign-on and other services to host a marketplace
- Content delivery network which has 3,000 applications that can be added to the Intercloud Marketplace
- Keystone for authorization; OpenLDAP for identity management
- Application delivery platform to enable cloud-native development on the Intercloud Marketplace
- Metering software

Summary

In the first chapter of this book we took an in-depth look at where we've been and what the current marketplace has to offer in terms of services from cloud. We explored the evolution of compute, IT service architectures and models, network and server virtualization, and the evolution of cloud computing to the Intercloud. In this chapter we looked in detail at the architecture of the Intercloud and the components that make up the Intercloud. The next chapter takes a deeper look into the Intercloud IT management strategy.

This chapter provided an overview of the architecture and technologies that are integral to the Intercloud. The Intercloud architecture is composed of the cloud platform and underlying infrastructure, cloud services and value-added products, application enablement platform as a service, cloud OSS/BSS, and Marketplace. The architecture is defined holistically in this chapter, and the chapters that follow will further develop the components and business requirements of the various layers of the Intercloud.

Key Messages

- The Intercloud architecture is composed of the cloud platform and underlying infrastructure, cloud services and value-added products, application enablement platform as a service, cloud OSS/BSS, and Marketplace.
- The Intercloud connects Cisco Powered cloud providers, enterprise clouds, public clouds, and Cisco Intercloud Services to deliver a robust service across all use cases.
- The Intercloud enables complete application portability while ensuring governance, compliance, operational management, and billing support requirements.

References

1. www.cisco.com/c/en/us/solutions/enterprise/design-zone-data-centers/index.html.
2. www.cisco.com/c/dam/en/us/solutions/collateral/trends/cloud/cloud-security.pdf.

This page intentionally left blank

Index

A

B

C

D

E

F

G

H

I

K

L

M

N

O

P

R

S

T

U

V

W

X

Z