

Making Everything Easier!™

Amazon Web Services™

**FOR
DUMMIES®**
A Wiley Brand

Learn to:

- Get acquainted with the AWS® toolset
- Implement a website that utilizes AWS
- Manage AWS usage and costs
- Use Amazon Glacier™ to store large quantities of data

Bernard Golden

Author of Virtualization For Dummies



Get More and Do More at Dummies.com®



Start with **FREE** Cheat Sheets

Cheat Sheets include

- Checklists
- Charts
- Common Instructions
- And Other Good Stuff!

To access the Cheat Sheet created specifically for this book, go to
www.dummies.com/cheatsheet/amazonwebservices

Get Smart at Dummies.com

Dummies.com makes your life easier with 1,000s of answers on everything from removing wallpaper to using the latest version of Windows.

Check out our

- Videos
- Illustrated Articles
- Step-by-Step Instructions

Plus, each month you can win valuable prizes by entering our Dummies.com sweepstakes. *

Want a weekly dose of Dummies? Sign up for Newsletters on

- Digital Photography
- Microsoft Windows & Office
- Personal Finance & Investing
- Health & Wellness
- Computing, iPods & Cell Phones
- eBay
- Internet
- Food, Home & Garden

Find out “HOW” at Dummies.com

*Sweepstakes not currently available in all countries; visit Dummies.com for official rules.



Amazon
Web Services[™]
FOR
DUMMIES[®]
A Wiley Brand

***Amazon
Web Services™***
**FOR
DUMMIES®**
A Wiley Brand

by Bernard Golden

**FOR
DUMMIES®**
A Wiley Brand

Amazon Web Services™ For Dummies®

Published by: **John Wiley & Sons, Inc.**, 111 River Street, Hoboken, NJ 07030-5774, www.wiley.com

Copyright © 2013 by John Wiley & Sons, Inc., Hoboken, New Jersey

Published simultaneously in Canada

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without the prior written permission of the Publisher. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at <http://www.wiley.com/go/permissions>.

Trademarks: Wiley, For Dummies, the Dummies Man logo, Dummies.com, Making Everything Easier, and related trade dress are trademarks or registered trademarks of John Wiley & Sons, Inc. and may not be used without written permission. Amazon Web Services is a trademark of Amazon Technologies, Inc. All other trademarks are the property of their respective owners. John Wiley & Sons, Inc. is not associated with any product or vendor mentioned in this book.

LIMIT OF LIABILITY/DISCLAIMER OF WARRANTY: THE PUBLISHER AND THE AUTHOR MAKE NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE ACCURACY OR COMPLETENESS OF THE CONTENTS OF THIS WORK AND SPECIFICALLY DISCLAIM ALL WARRANTIES, INCLUDING WITHOUT LIMITATION WARRANTIES OF FITNESS FOR A PARTICULAR PURPOSE. NO WARRANTY MAY BE CREATED OR EXTENDED BY SALES OR PROMOTIONAL MATERIALS. THE ADVICE AND STRATEGIES CONTAINED HEREIN MAY NOT BE SUITABLE FOR EVERY SITUATION. THIS WORK IS SOLD WITH THE UNDERSTANDING THAT THE PUBLISHER IS NOT ENGAGED IN RENDERING LEGAL, ACCOUNTING, OR OTHER PROFESSIONAL SERVICES. IF PROFESSIONAL ASSISTANCE IS REQUIRED, THE SERVICES OF A COMPETENT PROFESSIONAL PERSON SHOULD BE SOUGHT. NEITHER THE PUBLISHER NOR THE AUTHOR SHALL BE LIABLE FOR DAMAGES ARISING HEREFROM. THE FACT THAT AN ORGANIZATION OR WEBSITE IS REFERRED TO IN THIS WORK AS A CITATION AND/OR A POTENTIAL SOURCE OF FURTHER INFORMATION DOES NOT MEAN THAT THE AUTHOR OR THE PUBLISHER ENDORSES THE INFORMATION THE ORGANIZATION OR WEBSITE MAY PROVIDE OR RECOMMENDATIONS IT MAY MAKE. FURTHER, READERS SHOULD BE AWARE THAT INTERNET WEBSITES LISTED IN THIS WORK MAY HAVE CHANGED OR DISAPPEARED BETWEEN WHEN THIS WORK WAS WRITTEN AND WHEN IT IS READ.

For general information on our other products and services, please contact our Customer Care Department within the U.S. at 877-762-2974, outside the U.S. at 317-572-3993, or fax 317-572-4002. For technical support, please visit www.wiley.com/techsupport.

Wiley publishes in a variety of print and electronic formats and by print-on-demand. Some material included with standard print versions of this book may not be included in e-books or in print-on-demand. If this book refers to media such as a CD or DVD that is not included in the version you purchased, you may download this material at <http://booksupport.wiley.com>. For more information about Wiley products, visit www.wiley.com.

Library of Congress Control Number: 2013942773

ISBN 978-1-118-57183-5 (pbk); ISBN 978-1-118-65198-8 (ebk); ISBN 978-1-118-65226-8 (ebk)

Manufactured in the United States of America

10 9 8 7 6 5 4 3 2 1

Praise for Amazon Web Services For Dummies

This is a great resource for anyone considering the jump into cloud computing. Golden accurately explores the roster of AWS services while clearly illustrating ways for developers to make applications easier to build and manage. He manages to address both business requirements and technical content in a way that will appeal to almost any audience.

— Jeff Barr, Sr. Technology Evangelist,
Amazon Web Services

One of the challenges Bitnami users face is understanding the breadth and power of AWS. *Amazon Web Services For Dummies* helps our users build a great foundation of AWS skills. Anyone who is new to AWS and wants to be successful should start with this book.

— Erica Brescia, COO and co-founder of Bitnami

Netflix is all-in on AWS. We believe it is the richest, most scalable, most innovative cloud platform in the industry. Building AWS skills is critical for careers today — and *Amazon Web Services For Dummies* is the best resource I know of to learn AWS from the ground up. Buy this book to learn what your future will look like.

— Adrian Cockcroft, Netflix Cloud Architect

Contents at a Glance

<i>Introduction</i>	<i>1</i>
<i>Part I: Getting Started with AWS.....</i>	<i>5</i>
Chapter 1: Amazon Web Services Philosophy and Design.....	7
Chapter 2: Introducing the AWS API	25
Chapter 3: Introducing the AWS Management Console.....	37
<i>Part II: Diving into AWS Offerings</i>	<i>53</i>
Chapter 4: Setting Up AWS Storage	55
Chapter 5: Stretching Out with Elastic Compute Cloud.....	101
Chapter 6: AWS Networking	135
Chapter 7: AWS Security	155
Chapter 8: Additional Core AWS Services	183
<i>Part III: Using AWS.....</i>	<i>209</i>
Chapter 9: AWS Platform Services.....	211
Chapter 10: AWS Management Services	241
Chapter 11: Managing AWS Costs	279
Chapter 12: Bringing It All Together: An AWS Application.....	289
<i>Part IV: The Part of Tens</i>	<i>329</i>
Chapter 13: Ten Reasons to Use Amazon Web Services.....	331
Chapter 14: Ten Design Principles for Cloud Applications	339
<i>Index</i>	<i>347</i>

Table of Contents

<i>Introduction</i>	1
<i>Part 1: Getting Started with AWS</i>	5
Chapter 1: Amazon Web Services Philosophy and Design	7
Cloud Computing Defined.....	8
IaaS, Paas, SaaS	8
Private-versus-public cloud computing	10
Understanding the Amazon Business Philosophy	10
Measuring the scale of AWS	12
Checking the bottom line.....	14
The AWS Infrastructure	14
Making hard hardware decisions	15
Examining Amazon’s software infrastructure strategy.....	16
The AWS Ecosystem.....	18
Counting Up the Network Effects Benefit	20
AWS versus Other Cloud Providers	22
Getting Ready for the 21st Century	24
Chapter 2: Introducing the AWS API	25
APIs: Understanding the Basics	25
Benefiting from Web Services	27
An Overview of the AWS API.....	29
Web services: SOAP or REST?	30
The AWS API.....	31
The AWS API in real-world use.....	32
AWS API Security	34
Chapter 3: Introducing the AWS Management Console	37
Setting Up Your Amazon Web Services Account.....	38
Accessing Your First AWS Service	45
Loading Data into S3 Buckets.....	48
S3 URL Naming Conventions	51
Last Words on the AWS Management Console.....	52



***Part 11: Diving into AWS Offerings* 53**

Chapter 4: Setting Up AWS Storage 55

Differentiating the Amazon Storage Options	56
Object storage	58
Distributed key-value storage	58
Storing Items in the Simple Storage Service (S3) Bucket	60
S3 storage basics	61
Managing Volumes of Information with Elastic Block Storage (EBS)	67
EBS reliability	69
EBS scope	69
EBS use	70
EBS performance	71
EBS snapshots	72
EBS pricing	73
An EBS example	75
Managing Archive Material with the Glacier Storage Service	78
Glacier in action	80
Glacier scoping	81
Glacier pricing	82
Scaling Key-Value Data with DynamoDB	85
Key-value versus relational databases	86
DynamoDB characteristics	87
Using DynamoDB	88
DynamoDB read consistency	90
DynamoDB scope and availability	90
DynamoDB cost	90
A DynamoDB example	91
Selecting an AWS Storage Service	96

Chapter 5: Stretching Out with Elastic Compute Cloud. 101

Introducing EC2	101
Seeing EC2's Unique Nature	103
Understanding images	104
S3-backed images	106
EBS-backed images	107
EC2 instance types	109
EC2 image sizes	111
EC2 scope	112
EC2 pricing and deployment options	113
Creating new EC2 images	121
Working with an EC2 Example	124

Chapter 6: AWS Networking	135
Brushing Up on Networking Basics	136
Virtual LANS — keeping data private	137
The Amazon alternative to VLANs	139
AWS Network IP Addressing	140
AWS IP Address Mapping	145
AWS Direct Connect	146
High-Performance AWS Networking	148
AWS Elastic IP Addresses	149
Elastic IP address pricing	150
Elastic IP addresses and AWS network scope	151
AWS Instance Metadata	151
Instance IP Address Communication	152
Chapter 7: AWS Security	155
Clouds Can Have Boundaries, Too	156
The Deperimeterization of Security	158
AWS Security Groups	161
Security groups	162
Using Security Groups to Partition Applications	166
Security group scope	168
Security group cost	169
Security Group Best Practices	169
AWS Virtual Private Cloud (VPC)	170
VPC overview	171
How VPC works	172
VPC scope	179
VPC cost	179
Using VPC	179
AWS Application Security	179
Chapter 8: Additional Core AWS Services	183
Understanding the Other AWS Services	183
Deciding whether it makes sense to use other AWS services	185
Working with Identity and Access Management (IAM)	188
Elastic Load Balancer (ELB)	190
Route 53	193
CloudFront	196
Relational Database Service (RDS)	200
ElastiCache	202
Integrating Additional AWS Services into Your Application	205
Choosing the Right Additional AWS Service Integration Approach	207
Dealing with AWS Lock-in	207

Part III: Using AWS..... 209**Chapter 9: AWS Platform Services211**

Searching with CloudSearch	212
Managing Video Conversions with Elastic Transcoder	215
Simple Queue Service.....	219
Simple Notification Service	223
Simple E-Mail Service	228
Simple Workflow Service	231
Dealing with Big Data with the Help of Elastic MapReduce	233
Redshift	237

Chapter 10: AWS Management Services 241

Managing Your AWS Applications	242
Watching the cloud with AWS CloudWatch	242
The wonders of AWS Auto Scaling.....	247
Introducing AWS Elastic Beanstalk.....	251
Using Elastic Beanstalk	253
AWS CloudFormation	257
AWS OpsWorks	265
Which AWS Management Service Should I Use?	277

Chapter 11: Managing AWS Costs 279

AWS Costs — It's Complicated	280
Taking Advantage of Cost and Utilization Tracking.....	281
Managing Your AWS Costs	287

Chapter 12: Bringing It All Together: An AWS Application 289

Putting the Pieces Together	290
Creating your own AWS account	290
Enabling access on your security group.....	291
Locating and launching an appropriate WordPress Amazon Machine Image (AMI).....	291
Vertically partitioning your WordPress application	303
Creating a new Amazon Machine Image (AMI) from your WordPress application.....	316
Improving Application Robustness with Geographical Redundancy ...	318
Horizontally partitioning the RDS database	319
Launching a second WordPress application tier instance in a separate availability zone	322
Creating an Elastic Load Balancer	323

Part IV: The Part of Tens* 329*Chapter 13: Ten Reasons to Use Amazon Web Services 331**

AWS Provides IT Agility	331
AWS Provides Business Agility	332
AWS Offers a Rich Services Ecosystem	332
AWS Simplifies IT Operations	333
AWS Spans the Globe	334
AWS Is the Leading Cloud-Computing Service Provider	334
AWS Enables Innovation	335
AWS Is Cost Effective	336
AWS Aligns Your Organization with the Future of Technology.....	336
AWS Is Good for Your Career.....	337

Chapter 14: Ten Design Principles for Cloud Applications 339

Everything Fails All the Time	339
Redundancy Protects Against Resource Failure	340
Geographic Distribution Protects Against Infrastructure Failure	340
Monitoring Prevents Problems	341
Utilization Review Prevents Waste.....	342
Application Management Automates Administration	343
Security Design Prevents Breaches and Data Loss	343
Encryption Ensures Privacy	344
Tier-Based Design Increases Efficiency	345
Good Application Architecture Prevents Technical Debt.....	346

***Index* 347**

Introduction

This book is designed with one purpose in mind: to make it easy for you, the reader, to understand and begin using Amazon Web Services (AWS) — an emerging technology platform that is profoundly disrupting the technology industry and enabling hundreds of thousands of individuals, businesses, and nonprofit organizations to gain easy access to on-demand computing resources.

About This Book

In a sense, this book is an extension of my earlier book *Virtualization For Dummies* (Wiley Publishing), which has a chapter describing “The Future of Virtualization.” In my research to identify which direction virtualization would take, I came across Amazon Web Services, a then-new offering was referred to by Amazon employees as *Infrastructure as a Service*. To indicate how briefly this new type of computing has been available, the term *cloud computing* was still more than a year away when *Virtualization For Dummies* was published.

As I spoke to Amazon representatives about the company’s new offering, I experienced the same reaction I had when first exposed to open source software — a visceral response that made me ask out loud: “If this service is available to users, who will stick with the old way of doing things?”

Nothing in the subsequent years has changed my mind — in fact, that experience strengthens my conviction that cloud computing in general, and Amazon Web Services in particular, will transform the way applications are designed and built. I’ve worked with people from many companies who have resigned themselves to the length of the usual IT resource provisioning process — taking six weeks or more to obtain a virtual machine. When I demonstrate the ability of AWS to provision an instance (Amazon’s term for a virtual machine) in ten minutes or less, these people regard what they’re seeing with disbelief, staggered that the conventional (lengthy) provisioning process isn’t somehow set in stone.

Amazon continues to challenge the incumbent community of technology vendors, releasing new services and cutting prices at an unrelenting pace.

I fully expect that a decade from now, AWS will be one of the top two or three global technology vendors, and that a number of today's giants will be gone, driven out of business, or into forced mergers by their inability to compete on Amazon's terms.

But (there's always a *but*, isn't there?) how to get started is a challenge that many people face when they consider using AWS. AWS documentation is quite thorough, but you won't find there a general guide for beginners to start from scratch and develop new skills.

For this reason, I proposed this book to the publisher. I've heard from many people who are excited about using AWS but frustrated about how to learn about and use AWS. The Powers That Be at Wiley and I agreed that an introductory book about AWS that helps newbies begin using it productively would be extremely useful — and so we set to work to create the book that you now hold in your hands. I hope that you'll find it a useful and helpful roadmap for your AWS journey.

Using This Book

This book contains a mix of text, URLs, and terminal commands that you can execute. Please note these stylistic tidbits:

- ✓ Text that you type just as it appears in the book is in **bold**. The exception is when you're working through a step list: Because each step is bold, the text to type is not in bold.
- ✓ Web addresses and programming code appear in `monofont` type. If you're reading a digital version of this book on a device connected to the Internet, you can click the web address to visit that website, such as this one: www.dummies.com.

Foolish Assumptions

This book is designed to address a range of readers. Part I is an overview of AWS and an introduction to how the service works. It's appropriate for executives, project managers, and IT managers wanting to gain a basic understanding of the service so that they have a context for the benefits their organization can realize by using AWS. No particular technical background is assumed or necessary in Part I.

If you plan to work with AWS in a hands-on manner, Parts II and III provide a comprehensive review of all AWS offerings. I devote a full chapter to the use of the AWS technology, with a set of exercises that begin with a simple example and progressively build into a more complex application that leverages a number of AWS products. A technical background is necessary to comprehend Parts II and III; however, none of the information or exercises is particularly difficult from a technology perspective.

Icons Used in This Book



The Tip icon marks tips (duh!) and shortcuts that you can use to make using Amazon Web Services easier.



Remember icons mark information that's especially important to know. To siphon off the most important information in each chapter, just skim these icons.



The Technical Stuff icon marks information of a highly technical nature that you can normally skip over.



The Warning icon tells you to watch out! It marks important information that may save you headaches.

Beyond the Book

The technology industry continues to invent and evolve rapidly — and that goes double for cloud computing. It's important to have up-to-the-minute information on important new technology trends, and we're committed to providing new information as AWS evolves over time.

Here are three places you can look for information and help outside of this book:

- ✓ **Cheat Sheet:** You can find the Cheat Sheet for this book at www.dummies.com/cheatsheet/amazonwebservices. It describes the family of AWS services and provides guidelines for using them. Given how complex AWS is turning out to be, a general set of recommendations is useful indeed!

- ✓ **Dummies.com online articles:** Be sure to check out www.dummies.com/extras/amazonwebservices for additional online content dealing with AWS. Not everything I wanted to say could fit within the pages of this book, so I parceled out some content for the World Wide Web.
- ✓ **Updates:** Amazon Web Services continues to evolve rapidly. Amazon rolls out new services extremely quickly. I'll post updates about new AWS services at www.bernardgolden.com. Look there to learn the latest about AWS.

Unlike a novel, which requires you to begin at the beginning and carry on methodically throughout the book, *Amazon Web Services For Dummies* is designed to support what I like to call “random access” — if you hear about a particular AWS product and want to find out more, well, dig right in to that section of the book. If you want to understand the phenomenon of AWS, read the first part and then pick and choose among other areas that seem intriguing. This book supports your learning pattern and imposes no “official” reading approach. Dive in anywhere that makes sense to you.

Part I

Getting Started with AWS



Visit www.dummies.com for great Dummies content online.

In this part . . .

- ✓ See how Amazon designed Amazon Web Services from the beginning to be extremely scalable, modular in design, and highly robust.
- ✓ Find out how AWS reflects Amazon's unique approach to operating its business.
- ✓ Get an introduction to AWS, its business and technology underpinnings, and even get a small taste of hands-on use.
- ✓ Visit www.dummies.com for great Dummies content online.

Chapter 1

Amazon Web Services Philosophy and Design

In This Chapter

- ▶ Figuring out the cloud
 - ▶ Watching Amazon grow from retailer to the world's first cloud provider
 - ▶ Understanding the foundation of Amazon Web Services
 - ▶ Introducing the Amazon Web Services ecosystem
 - ▶ Seeing how the network effect helps you
 - ▶ Comparing Amazon Web Services to other cloud computing providers
-

You may be forgiven if you're puzzled about how Amazon, which started out as an online bookstore, has become *the* leading cloud computing provider. This chapter solves that mystery by discussing the circumstances that led Amazon into the cloud computing services arena and why Amazon Web Services, far from being an oddly different offering from a retailer, is a logical outgrowth of Amazon's business.

This chapter also compares Amazon's cloud offering to other competitors in the market and explains how its approach differs. As part of this comparison, I present some statistics on the size and growth of Amazon's offering, while describing why it's difficult to get a handle on its exact size.

The chapter concludes with a brief discussion about the Amazon Web Services ecosystem and why it is far richer than what Amazon itself provides — and why it offers more value for users of Amazon's cloud service.

But before I reveal all the answers to the Amazon mystery, I answer an even more fundamental question: What *is* all this cloud computing stuff, anyway?

Cloud Computing Defined

I believe that skill is built on a foundation of knowledge. Anyone who wants to work with Amazon Web Services (AWS, from now on) should have a firm understanding of cloud computing — what it is and what it provides.

IaaS, PaaS, SaaS

As a general overview, cloud computing refers to the delivery of computing services from a remote location over a network. The National Institute of Standards and Technology (NIST), a U.S. government agency, has a definition of cloud computing that is generally considered the gold standard. Rather than trying to create my own definition, I always defer to NIST's definition. The following information is drawn directly from it.

Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.

This cloud model is composed of five essential characteristics:

- ✓ **On-demand self-service:** A consumer can unilaterally provision computing capabilities, such as server time and network storage, automatically as needed without requiring human interaction with each service provider.
- ✓ **Broad network access:** Capabilities are available over the network and accessed via standard mechanisms that promote use by heterogeneous thin or thick client platforms (such as mobile phones, tablets, laptops, and workstations).
- ✓ **Resource pooling:** The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There's a sense of so-called *location independence*, in that the customer generally has no control or knowledge over the exact location of the provided resources but may be able to specify location at a higher level of abstraction (by country, state, or data center, for example). Examples of resources are storage, processing, memory, and network bandwidth.
- ✓ **Rapid elasticity:** Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be appropriated in any quantity at any time.

- ✓ **Measured service:** Cloud systems automatically control and optimize resource use by leveraging a metering capability at a level of abstraction that's appropriate to the type of service (storage, processing, bandwidth, or active user accounts, for example). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service.

Cloud computing is commonly characterized as providing three types of functionality, referred to *IaaS*, *PaaS*, and *SaaS*, where *aaS* is shorthand for “as a service” and *service* implies that the functionality isn't local to the user but rather originates elsewhere (a location in a remote location accessed via a network). The letters *I*, *P*, and *S* in the acronyms refer to different types of functionality, as the following list makes clear:

- ✓ **Infrastructure as a Service (IaaS):** Offers users the basic building blocks of computing: processing, network connectivity, and storage. (Of course, you also need other capabilities in order to fully support IaaS functionality — such as user accounts, usage tracking, and security.) You would use an IaaS cloud provider if you want to build an application from scratch and need access to fairly low-level functionality within the operating system.
- ✓ **Platform as a Service (PaaS):** Instead of offering low-level functions within the operating system, offers higher-level programming frameworks that a developer interacts with to obtain computing services. For example, rather than open a file and write a collection of bits to it, in a PaaS environment the developer simply calls a function and then provides the function with the collection of bits. The PaaS framework then handles the grunt work, such as opening a file, writing the bits to it, and ensuring that the bits have been successfully received by the file system. The PaaS framework provider takes care of backing up the data and managing the collection of backups, for example, thus relieving the user of having to complete further burdensome administrative tasks.
- ✓ **Software as a Service (SaaS):** Has clambered to an even higher rung on the evolutionary ladder than PaaS. With SaaS, all application functionality is delivered over a network in a pretty package. The user need do nothing more than use the application; the SaaS provider deals with the hassle associated with creating and operating the application, segregating user data, providing security for each user as well as the overall SaaS environment, and handling a myriad of other details.



As with every model, this division into *I*, *P*, and *S* provides a certain explanatory leverage and seeks to make neat and clean an element that in real life can be rather complicated. In the case of IPS, the model is presented as though the types are cleanly defined though they no longer are. Many cloud providers offer services of more than one type. Amazon, in particular, has begun to provide many platform-like services as it has built out its offerings, and has even ventured into a few full-blown application services that you'd associate with SaaS. You could say that Amazon provides all three types of cloud computing.

Private-versus-public cloud computing

If you find the mix of *I*, *P*, and *S* in the preceding section confusing, wait 'til you hear about the whole private-versus-public cloud computing distinction. Note the sequence of events:

1. Amazon, as the first cloud computing provider, offers *public cloud computing* — anyone can use it.
2. Many IT organizations, when contemplating this new Amazon Web Services creature, asked why they couldn't create and offer a service like AWS to their own users, hosted in their own data centers. This on-premise version became known as *private cloud computing*.
3. Continuing the trend, several hosting providers thought they could offer their IT customers a segregated part of their data centers and let customers build clouds there. This concept can also be considered private cloud computing because it's dedicated to one user. On the other hand, because the data to and from this private cloud runs over a shared network, is the cloud truly private?
4. Finally, after one bright bulb noted that companies may not choose only public or private, the term *hybrid* was coined to refer to companies using both private and public cloud environments.

As you go further on your journey in the cloud, you'll likely witness vociferous discussions devoted to which of these particular cloud environments is the better option. My own position is that no matter where you stand on the private/public/hybrid issue, public cloud computing will undoubtedly become a significant part of every company's IT environment. Moreover, Amazon will almost certainly be the largest provider of public cloud computing, so it makes sense to plan for a future that includes AWS. (Reading this book is part of that planning effort, so you get a gold star for already being well on your way!)



If you want to drill down further into cloud computing definitions, check out NIST's full description at <http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>. The U.S. federal government has been an early adopter of, and hard charger in, cloud adoption, and NIST has been assigned to create this (excellent) government-wide cloud computing resource.

Understanding the Amazon Business Philosophy

Amazon Web Services was officially revealed to the world on March 13, 2006. On that day, AWS offered the Simple Storage Service, its first service. (As you may imagine, Simple Storage Services was soon shortened to S3.) The idea

behind S3 was simple: It could offer the concept of object storage over the web, a setup where anyone could put an object — essentially, any bunch of bytes — into S3. Those bytes may comprise a digital photo or a file backup or a software package or a video or audio recording or a spreadsheet file or — well, you get the idea.

S3 was relatively limited when it first started out. Though objects could, admittedly, be written or read from anywhere, they could be stored in only one region: the United States. Moreover, objects could be no larger than 5 gigabytes — not tiny by any means, but certainly smaller than many files that people may want to store in S3. The actions available for objects were also quite limited: You could write, read, and delete them, and that was it.

In its first six years, S3 has grown in all dimensions. The service is now offered throughout the world in a number of different regions. Objects can now be as large as 5 terabytes. S3 can also offer many more capabilities regarding objects. An object can now have a termination date, for example: You can set a date and time after which an object is no longer available for access. (This capability may be useful if you want to make a video available for viewing for only a certain period, such as the next two weeks.) S3 can now also be used to host websites — in other words, individual pages can be stored as objects, and your domain name (say, `www.example.com`) can point to S3, which serves up the pages.

S3 did not remain the lone AWS example for long. Just a few months after it was launched, Amazon began offering Simple Queue Service (SQS), which provides a way to pass messages between different programs. SQS can accept or deliver messages within the AWS environment or outside the environment to other programs (your web browser, for example) and can be used to build highly scalable distributed applications.

Later in 2006 came Elastic Compute Cloud (known affectionately as EC2). As the AWS computing service, EC2 offers computing capacity on demand, with immediate availability and no set commitment to length of use.



Don't worry if this description of AWS seems overwhelming at first — in the rest of this book, you can find out all about the various pieces of AWS, how they work, and how you can use them to address your computing requirements. This chapter provides a framework in which to understand the genesis of AWS, with details to follow. The important thing for you to understand is how AWS got started, how big of a change it represents in the way computing is done, and why it's important to your future.

The overall pattern of AWS has been to add additional services steadily, and then quickly improve each service over time. AWS is now composed of more than 25 different services, many offered with different capabilities via different configurations or formats. This rich set of services can be mixed and matched to create interesting and unique applications, limited only by your imagination or needs.

So, from one simple service (S3) to more than 25 in just over six years, and throughout the world — and growing and improving all the time! You’re probably impressed by how fast all of this has happened. You’re not alone. Within the industry, Amazon is regarded with a mixture of awe and envy because of how rapidly it delivers new AWS functionality. If you’re interested, you can keep up with changes to AWS via its What’s New web page on the AWS site, at

```
http://aws.amazon.com/about-aws/whats-new
```

This torrid pace of improvement is great news for you because it means that AWS continually presents new things you can do — things you probably couldn’t do in the past because the AWS functionality would be too difficult to implement or too expensive to afford even if you could implement it.

Measuring the scale of AWS

Amazon is the pioneer of cloud computing and, because you’d have to have been living under a rock not to have heard about “the cloud,” being the pioneer in this area is a big deal. The obvious question is this: If AWS is the big dog in the market and if cloud computing is the hottest thing since sliced bread, how big are we talking about?

That’s an interesting question because Amazon reveals little about the extent of its business. Rather than break out AWS revenues, the company lumps them into an Other category in its financial reports.

Nevertheless, we have some clues to its size, based on information from the company itself and on informed speculation by industry pundits.

Amazon itself provides a proxy for the growth of the AWS service. Every so often, it announces how many objects are stored in the S3 service. Take a peek at Figure 1-1, which shows how the number of objects stored in S3 has increased at an enormous pace, jumping from 2.9 billion at the end of 2006 to over 2 trillion objects by the end of the second quarter of 2012. Given that pace of growth, it’s obvious that the business of AWS is booming.

Other estimates of the size of the AWS service exist as well. A very clever consultant named Huan Liu examined AWS IP addresses and projected the total number of server racks held by AWS, based on an estimate of how many servers reside in a rack. Table 1-1 breaks down the numbers by region.

Figure 1-1:
Counting S3
objects over
the years.

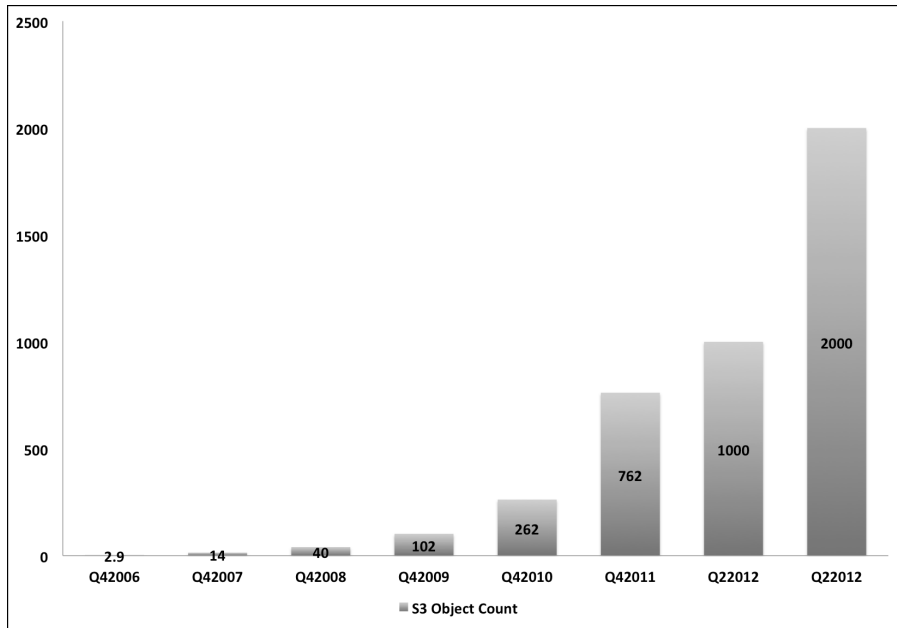


Table 1-1

Total AWS Servers

<i>AWS Region</i>	<i>Number of Server Racks</i>	<i>Number of Servers</i>
US East	5,030	321,920
US West (Oregon)	41	2,624
US West (California)	630	40,320
EU West (Ireland)	814	52,096
AP Northeast (Japan)	314	20,096
AP Southeast (Singapore)	246	15,744
SA East (Brazil)	25	1,600
Total	7,100	454,400

That's a lot of servers. (To see the original document outlining Liu's estimates, along with his methodology, go to <http://huanliu.wordpress.com/2012/03/13/amazon-data-center-size>). If you consider that each server can support a number of virtual machines (the number would vary, of course, according to the size of the virtual machines), AWS could support several million running virtual machines.

Amazon publishes a list of public IP addresses; as of May 2013, there are over four million available in AWS. This number is not inconsistent with Liu's estimated number of physical servers; it's also a convenient place to look to track how much AWS is growing. If you're interested, you can look at the AWS numbers at <https://forums.aws.amazon.com/ann.jspa?annID=1701>.



If you're not familiar with the term virtual machines, don't worry: I describe AWS technology in depth in Chapter 4. For an even more detailed discussion of virtual machines and virtualization proper, check out *Virtualization For Dummies*, by yours truly (published by John Wiley & Sons, Inc.).

Checking the bottom line

Though Amazon doesn't announce how many dollars AWS pulls in, that hasn't stopped others from making their own estimates of the size of AWS business — and their estimates make it clear that AWS is a very large business indeed.

Early in 2012, several analysts from Morgan Stanley analyzed the AWS business and judged that the service pulled in \$1.19 billion in 2011. (You gotta love the precision that these pundits come up with, eh?) Other analysts from JP Morgan Chase and UBS have calculated that AWS will achieve 2015 revenues of around \$2.5 billion.

The bottom line: AWS is big and getting bigger (and better) every day. It really is no exaggeration to say that AWS represents a revolution in computing. People are doing amazing things with it, and this book shows you how you can take advantage of it.

The AWS Infrastructure

If what Amazon is doing with AWS represents a revolution, as I describe in the previous section, how is the company bringing it about? In other words, how is it delivering this amazing service? Throughout this book, I go into the specifics of how the service operates, but for now I outline the general approach that Amazon has taken in building AWS.

First and foremost, Amazon has approached the job in a unique fashion, befitting a company that changed the face of retail. Amazon specializes in a low-margin approach to business, and it carries that perspective into AWS. Unlike almost every other player in the cloud computing market, Amazon has focused on creating a low-margin, highly efficient offering, and that offering starts with the way Amazon has built out its infrastructure.