DocStore

by aMedia Corp.

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

DocStore

aMedia DocStore allows customers to upload and manage their documents

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

DocStore

aMedia DocStore allows customers to upload and manage their documents

- **Upload and access content from anywhere** – browser and mobile apps
- **PDF, Office docs, images, videos** – text content extracted, indexed, and searchable

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

DocStore

Two Account Types

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

DocStore

Two Account Types

- **Free**

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

DocStore

Two Account Types

- **Free**
  - 5Gb storage
  - Search on name, created/modified date, and tags
    - No full-text indexing
  - All content may be securely downloaded by owner

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

DocStore

Two Account Types

- **Premium**

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

DocStore

Two Account Types

- **Premium**
  - 20 Gb storage
  - Full-text indexing and search for all documents (OCR on images)
  - All content may be securely downloaded by owner

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

DocStore

Two Major Components

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

DocStore

#1: Web Application

- Allow users to upload, manage, and view all documents
- All content uploaded directly to S3

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

DocStore

#1: Web Application

- Both powered by public DocStore API
- Both must scale
- Both must be highly available

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

DocStore

#1: Web Application

- Multi-region deployment for global support
- User can log in to any region
  - Always redirected to “home” region (i.e., region where account was created)

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

DocStore

#1: Web Application

– Work in groups to design architecture that drives the web and mobile applications, including the DocStore API

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

DocStore Requirements

Account Types

- **Free** – 5GB cap, basic document indexing, download original files
- **Premium** – 20GB cap, full-text indexing/search + OCR

Web App/API

- **User authentication**
- **Session state stored off-instance**
- **Powered by API tier**
- **Highly Available**

Content

- **Stored in S3**
- **Efficient upload/download to/from S3**
- **Static assets in one S3 bucket; distributed globally**

Deployment

- **Oregon (us-west-2) and Sydney (ap-southeast-2)**
- **Replicate minimal information (i.e., don't replicate uploaded content)**

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

DocStore

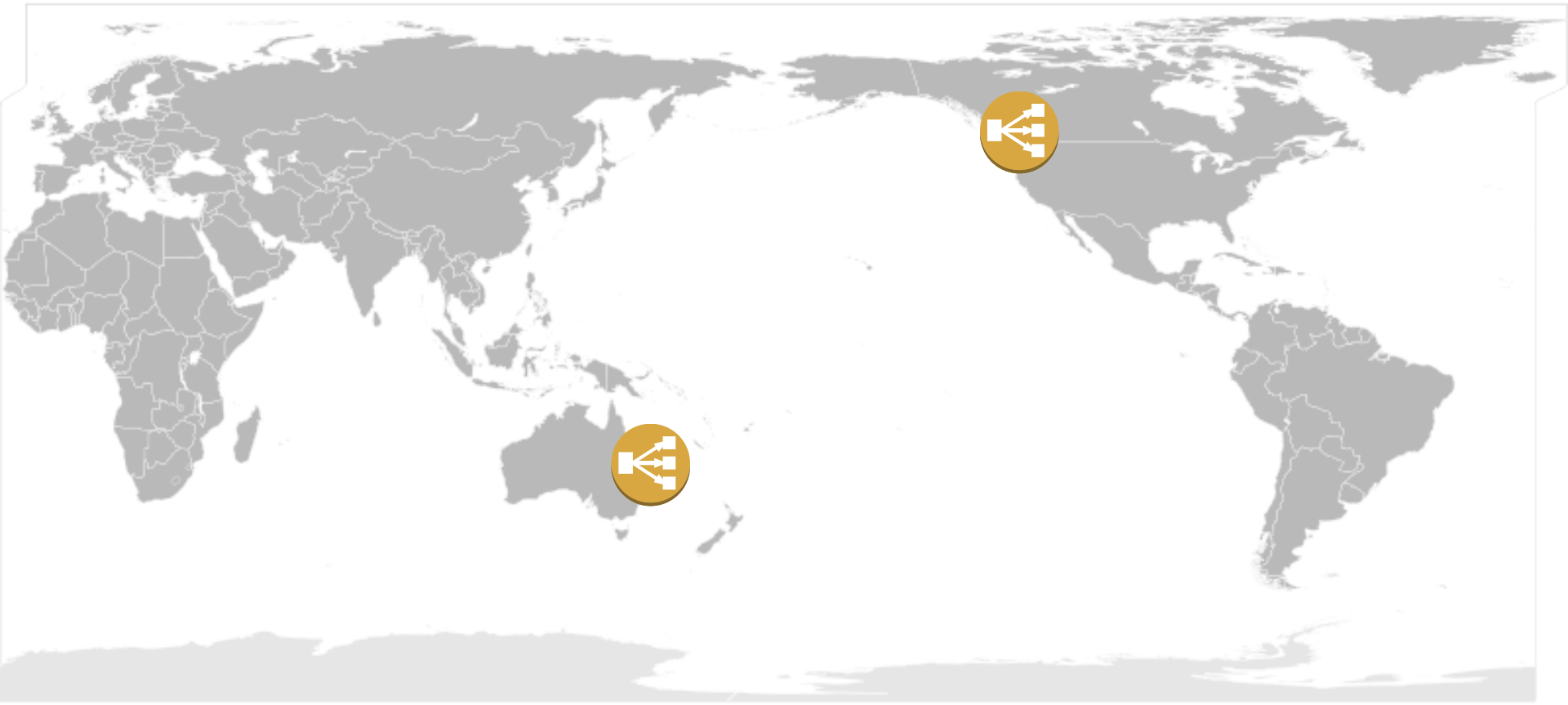
Sample Architecture

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Application deployed in two regions: us-west-2 (Oregon) and ap-southeast-2 (Sydney)

DocStore

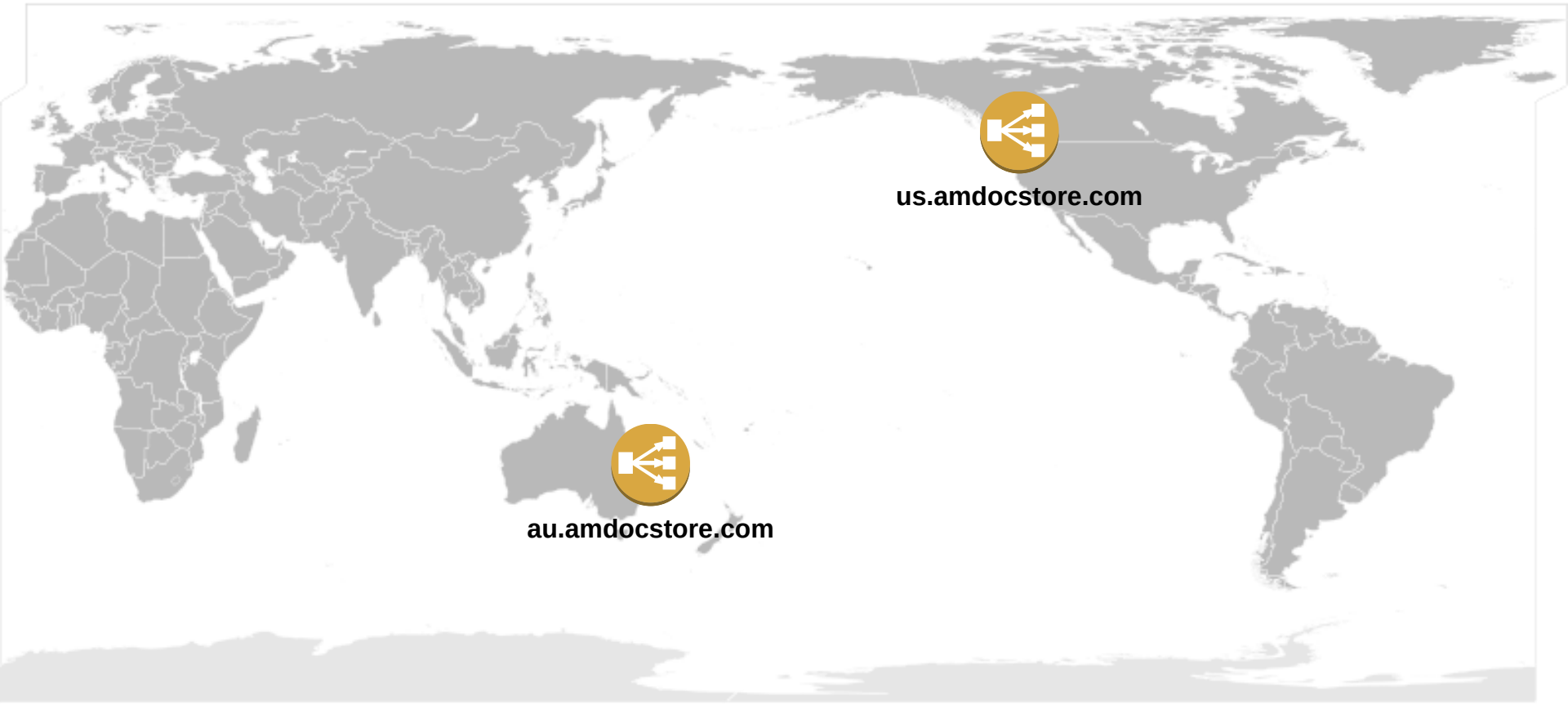


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Each ELB has a region-specific CNAME in Route 53

DocStore



© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Latency Based Routing feature of **Route 53** directs customers to closest endpoint

DocStore

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Closer to Oregon? That's where you go. In Perth? Head to the ELB in Sydney.

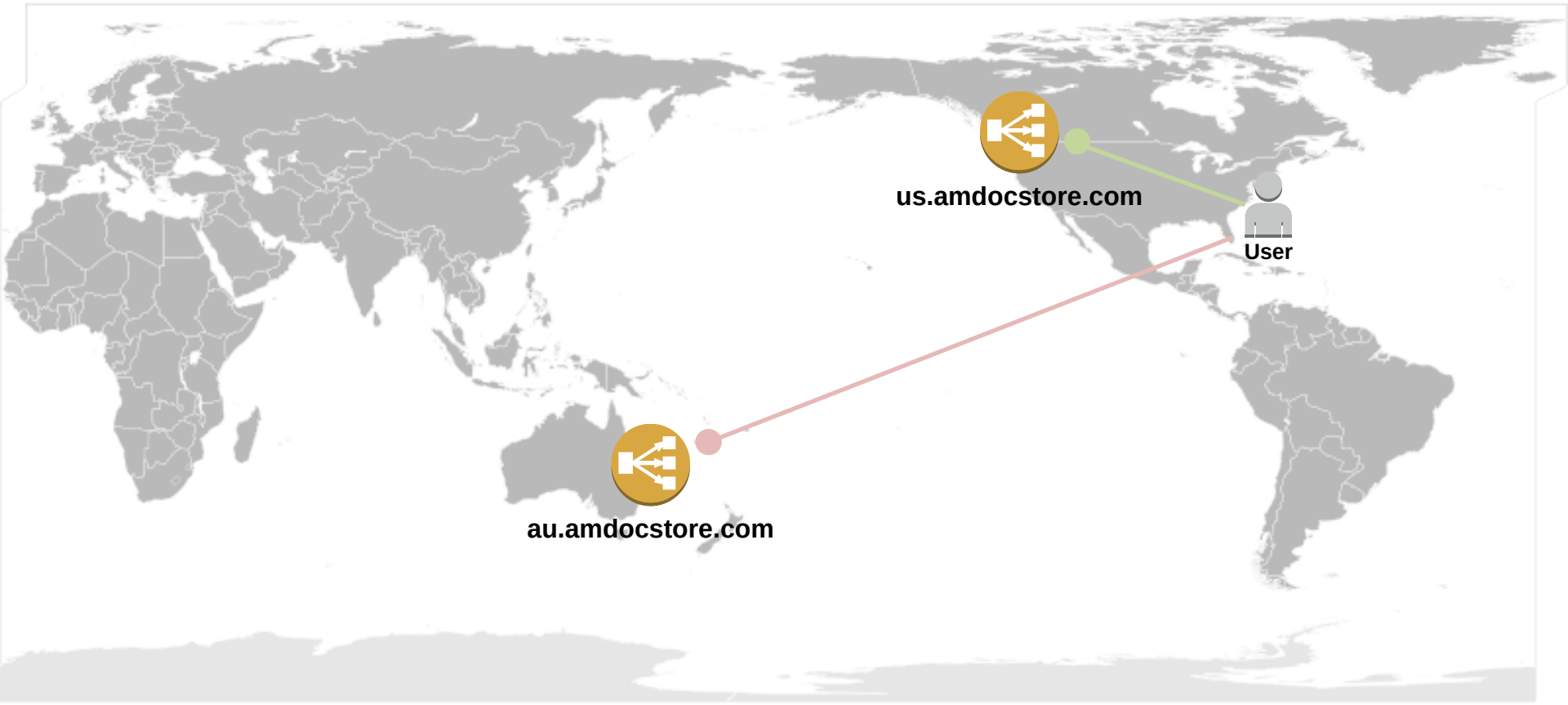
DocStore

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Closer to Oregon? That's where you go.

DocStore

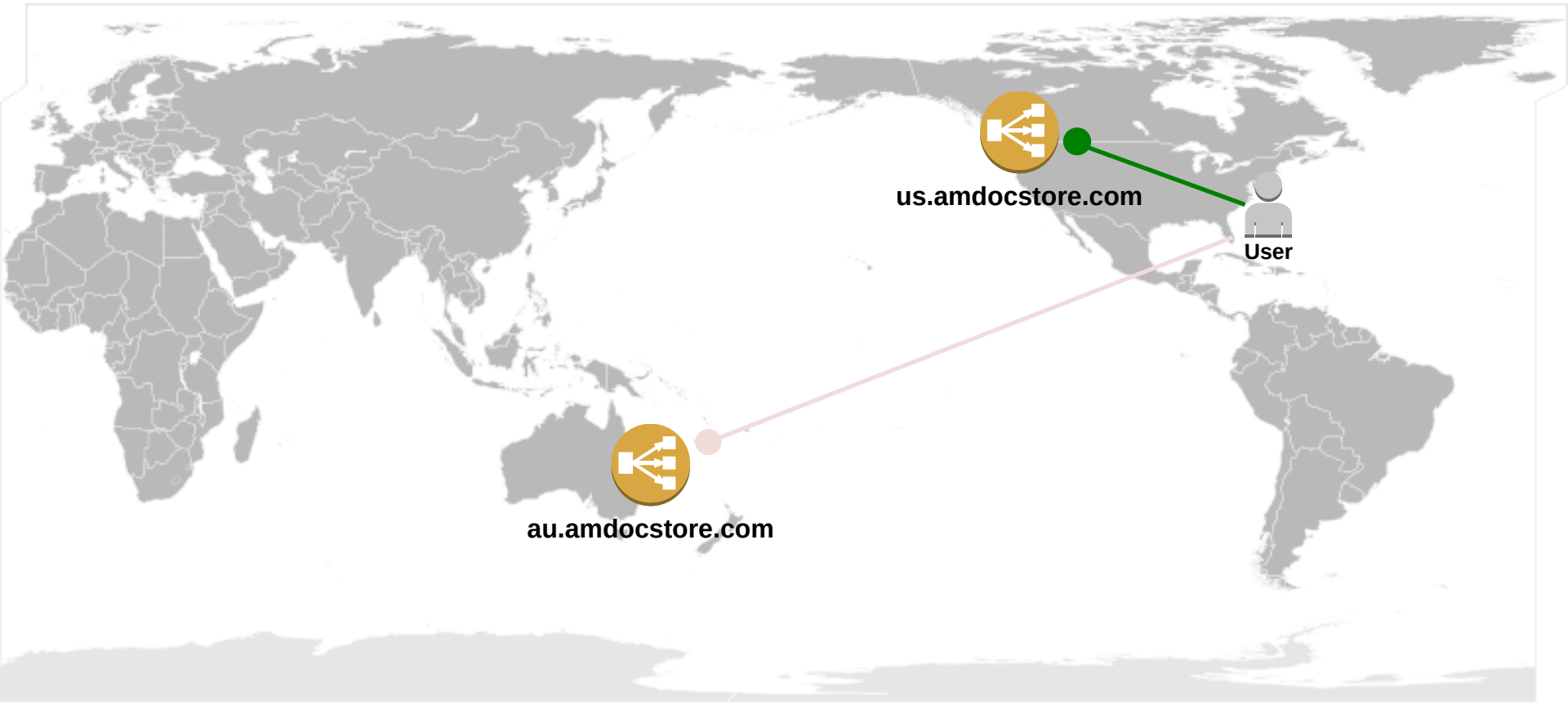


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Closer to Oregon? That's where you go.

DocStore



© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Let’s focus on one region

DocStore

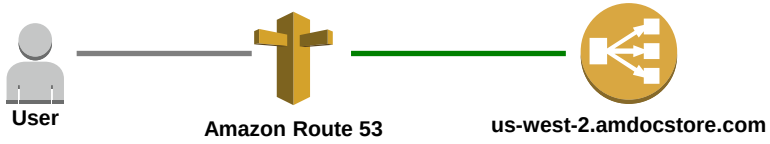
us.amdocstore.com

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Route 53 LBR directs a user to the **nearest Elastic Load Balancer** (in this case, the ELB in us-west-2)

DocStore

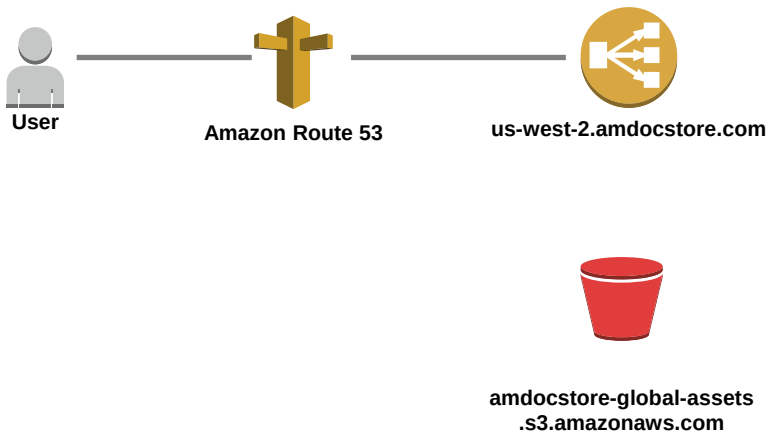


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

DocStore stores all **static web assets** (CSS, images, JavaScript, etc) **in S3**

DocStore

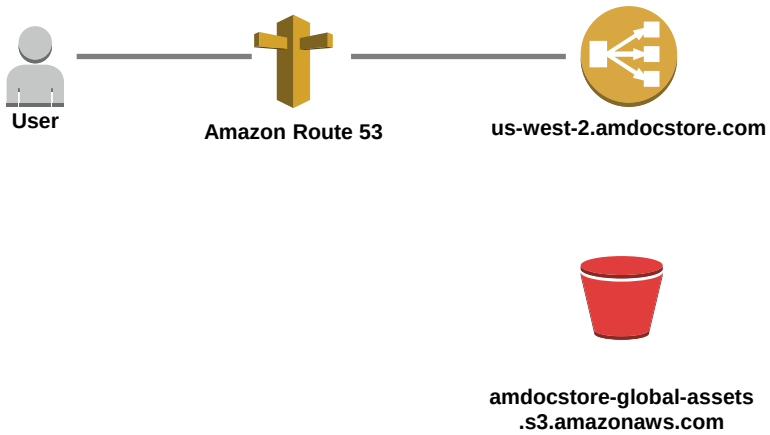


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

This bucket exists in the us-west-2 origin, but is the **canonical source for static asset requests globally**

DocStore

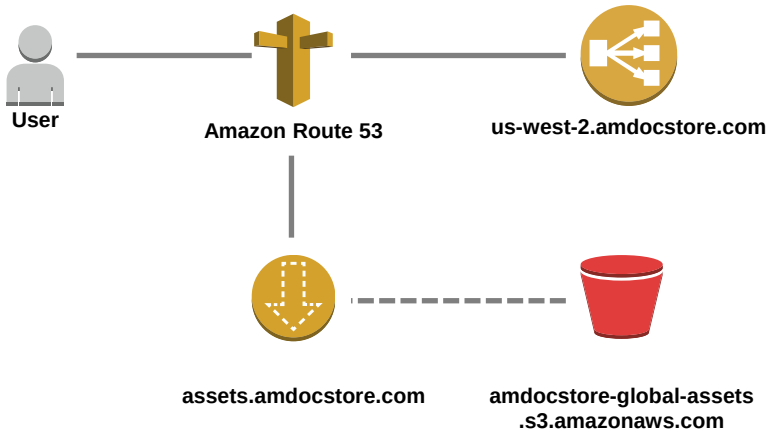


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

CloudFront distributes static web assets in S3 to end users with low latency using a **global network of edge locations**

DocStore

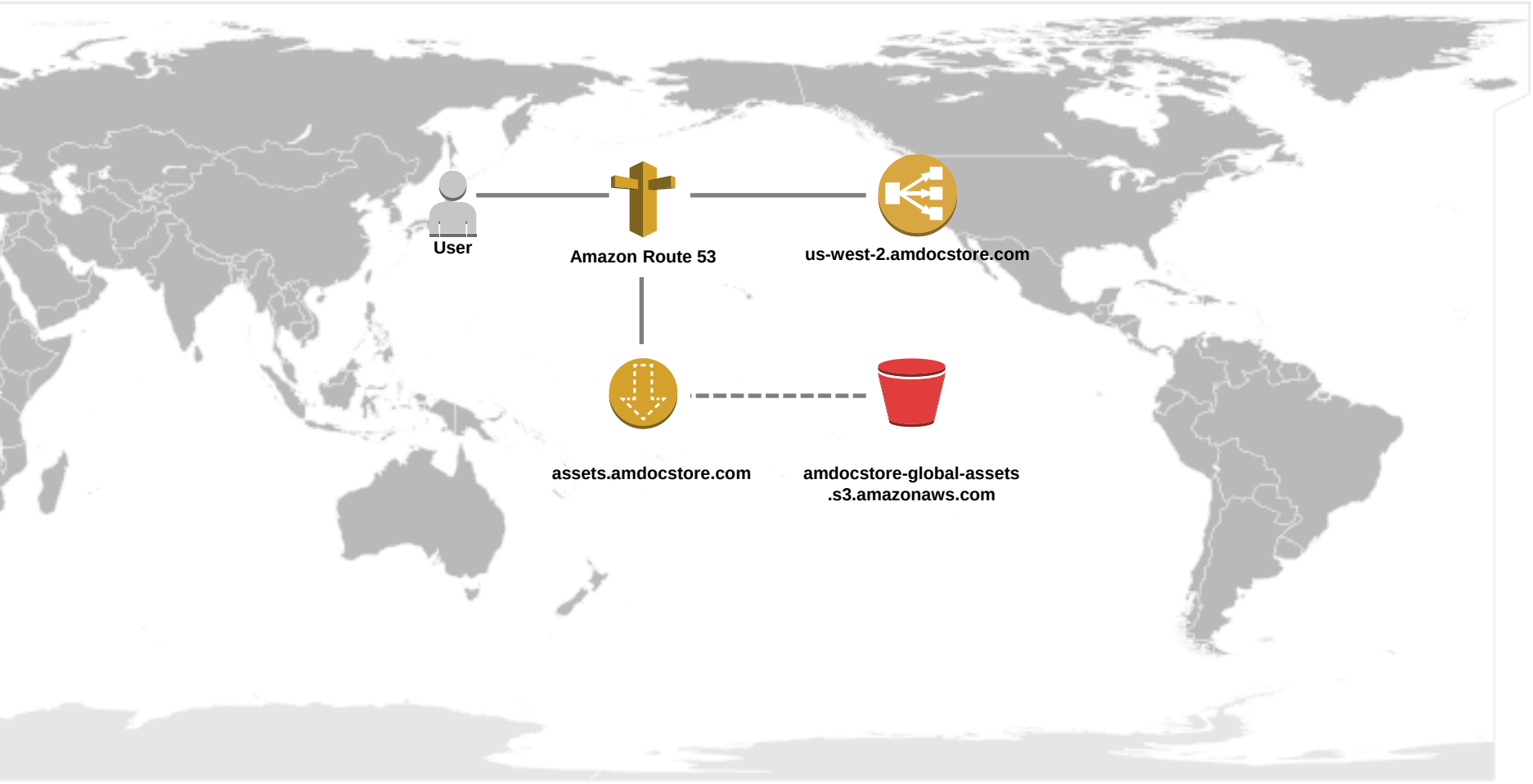


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Let's see how a user in **Tokyo** would access **http://assets.amdocstore.com/css/style.css**

DocStore

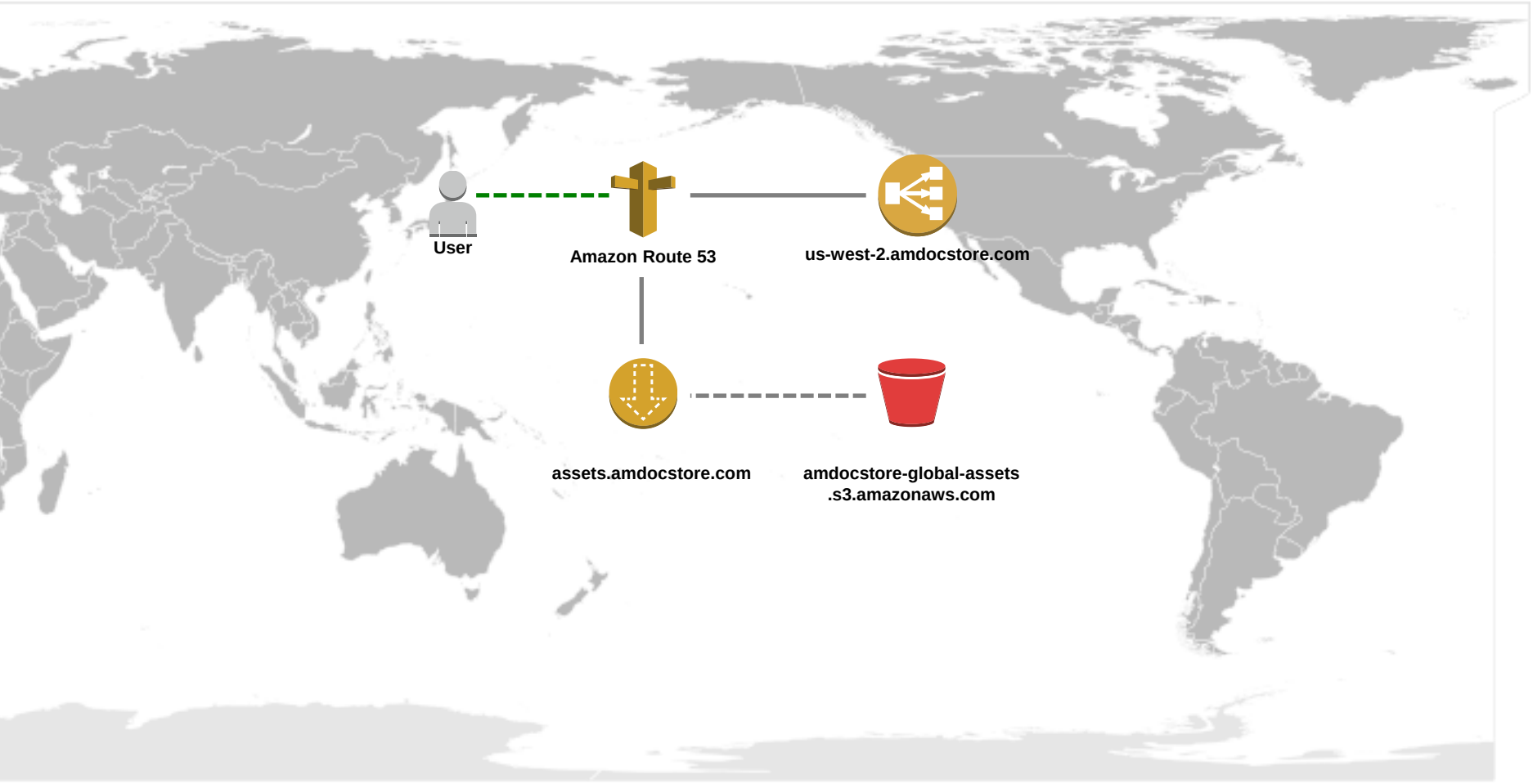


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Any user request for assets.amdocstore.com will **automatically route to the nearest edge location**

DocStore

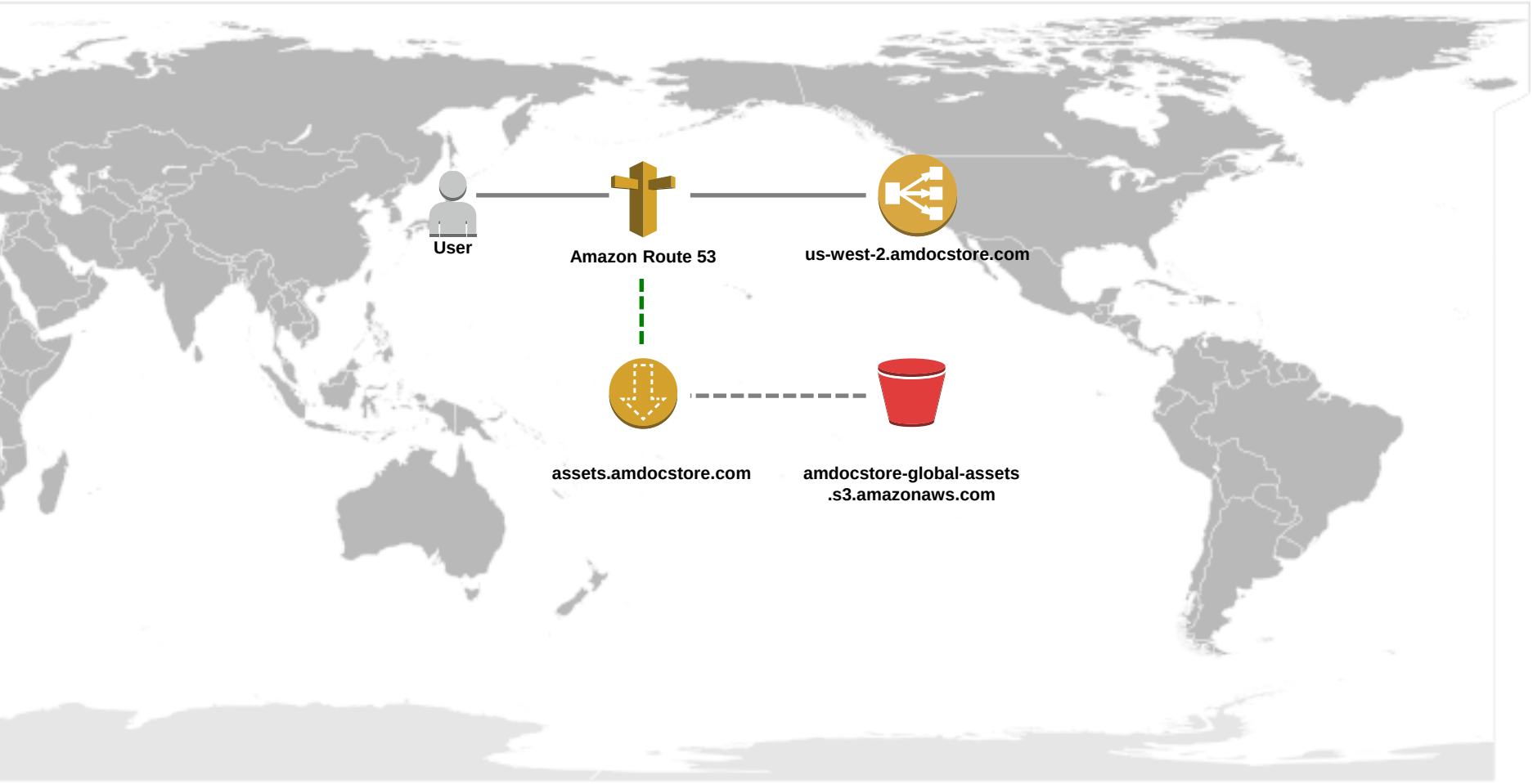


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Any user request for assets.amdocstore.com will **automatically route to the nearest edge location**

DocStore

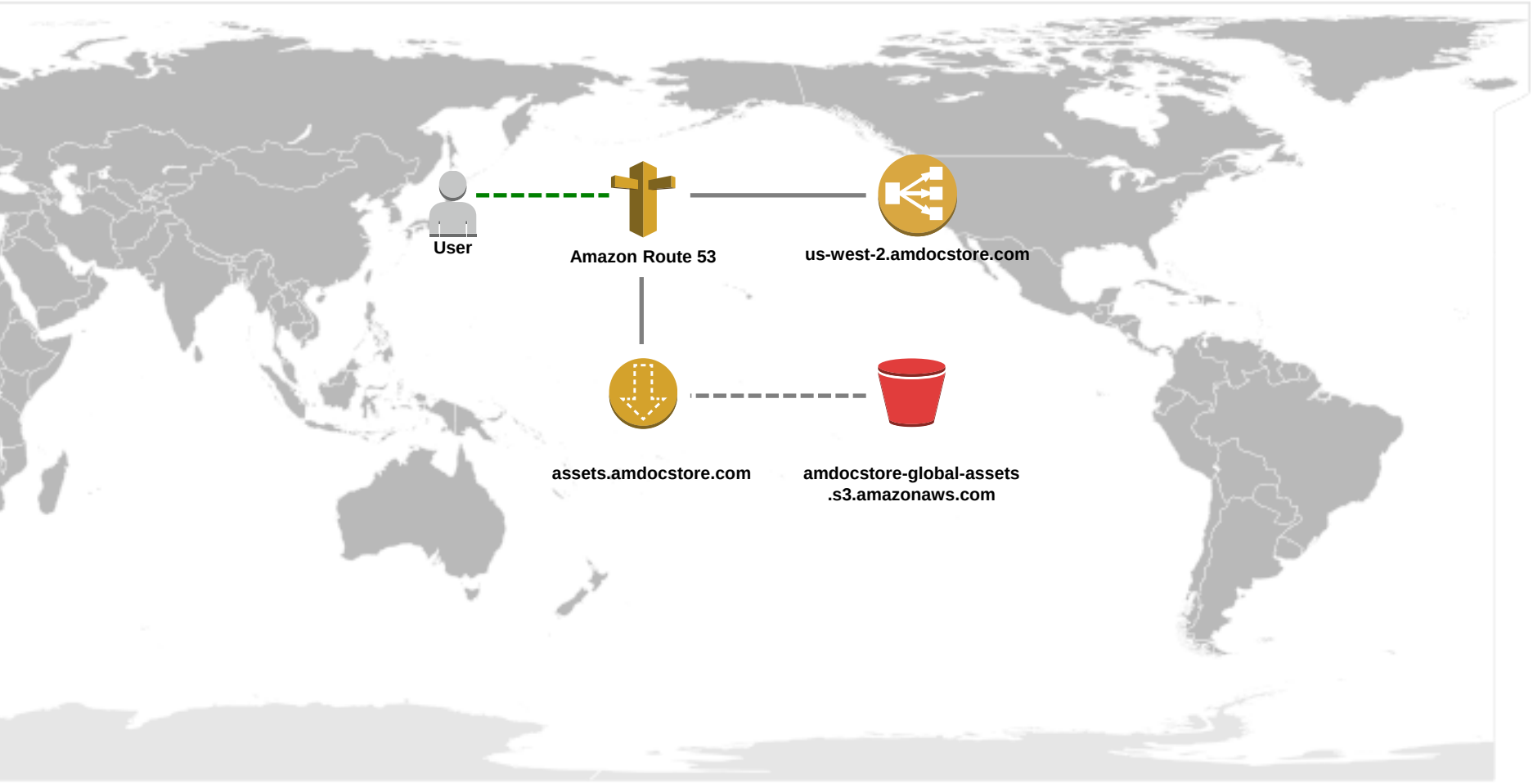


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Any user request for assets.amdocstore.com will **automatically route to the nearest edge location**

DocStore

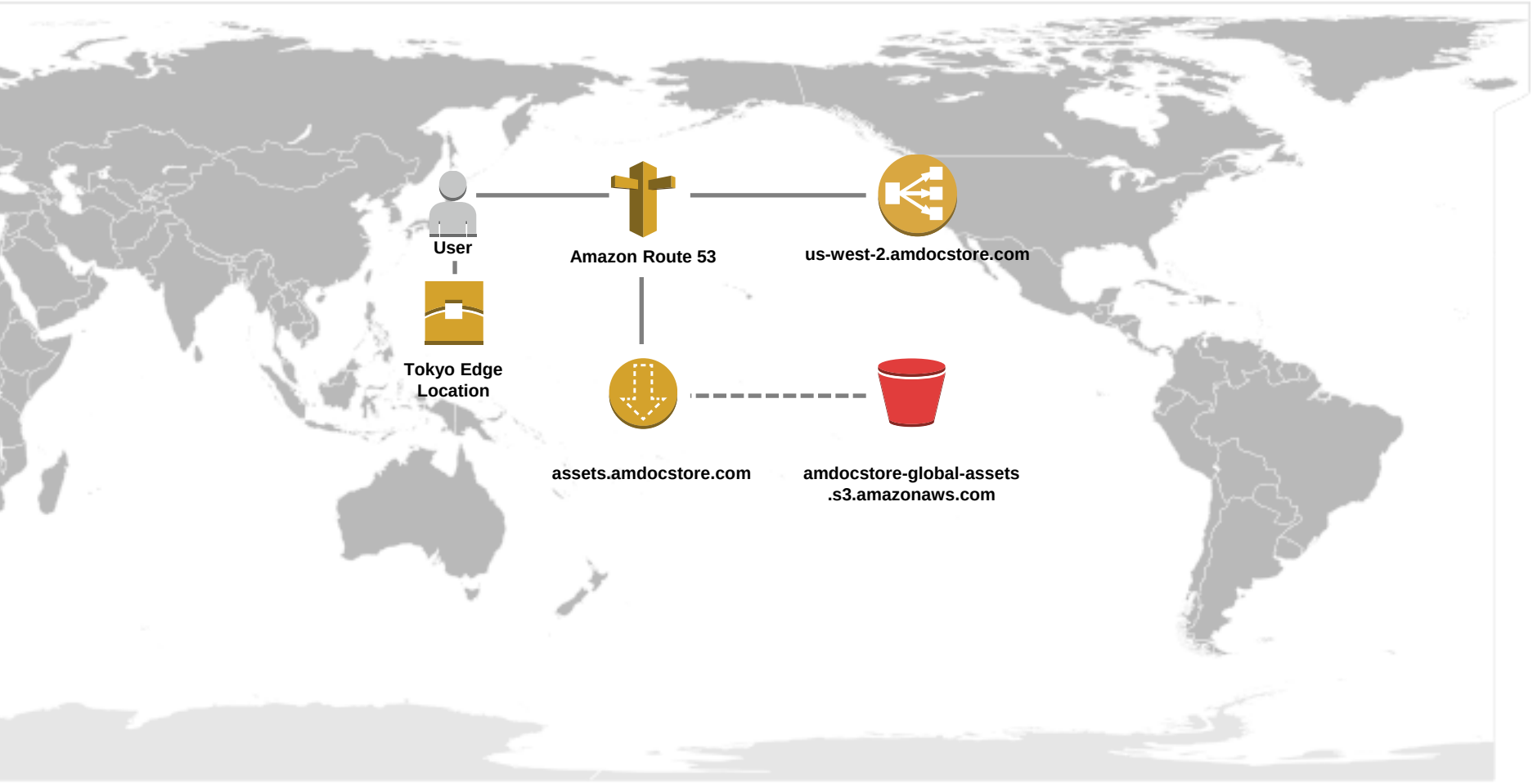


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Any user request for assets.amdocstore.com will **automatically route to the nearest edge location**

DocStore

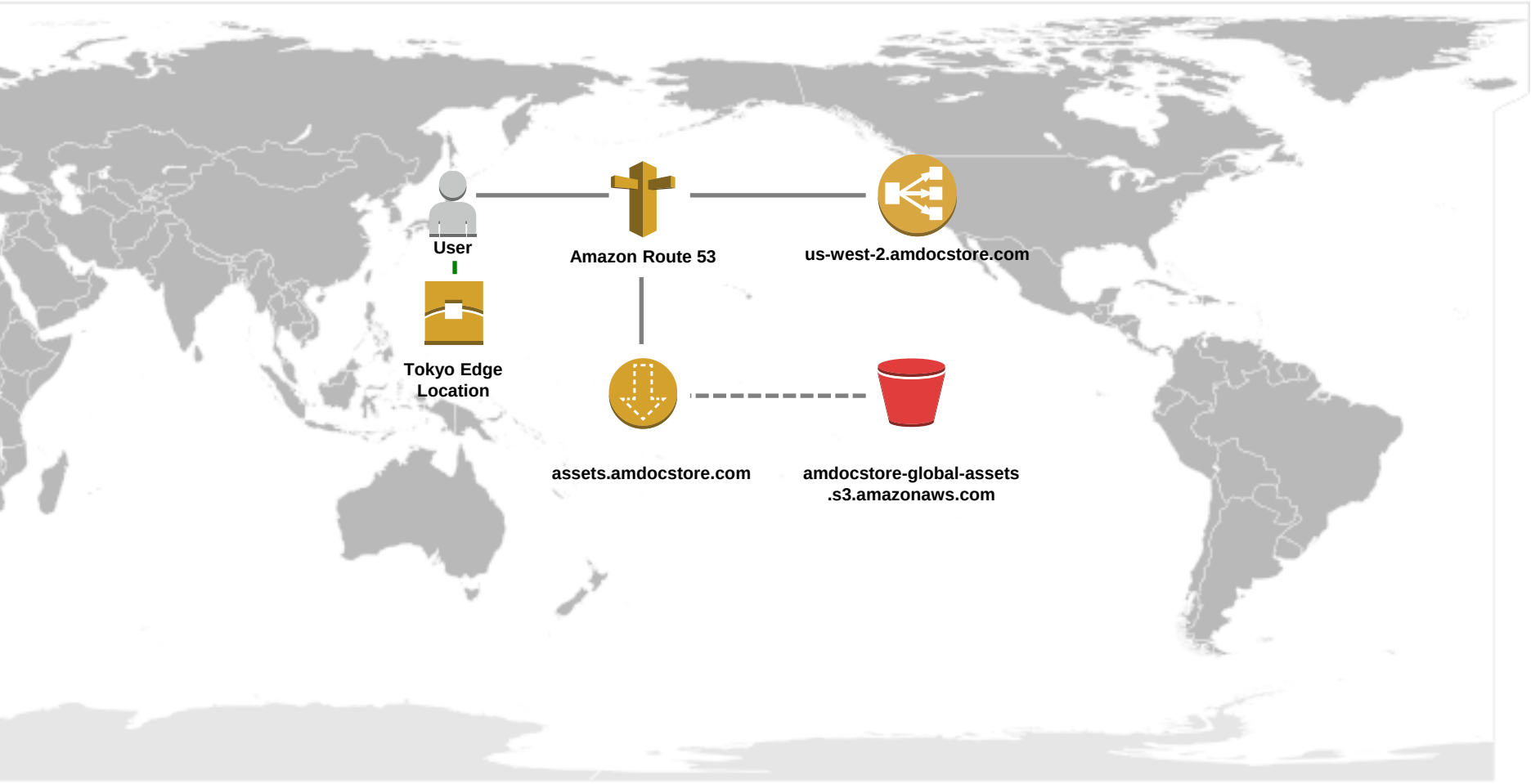


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

The requested **file will be served from the Tokyo edge** if it is in the cache

DocStore

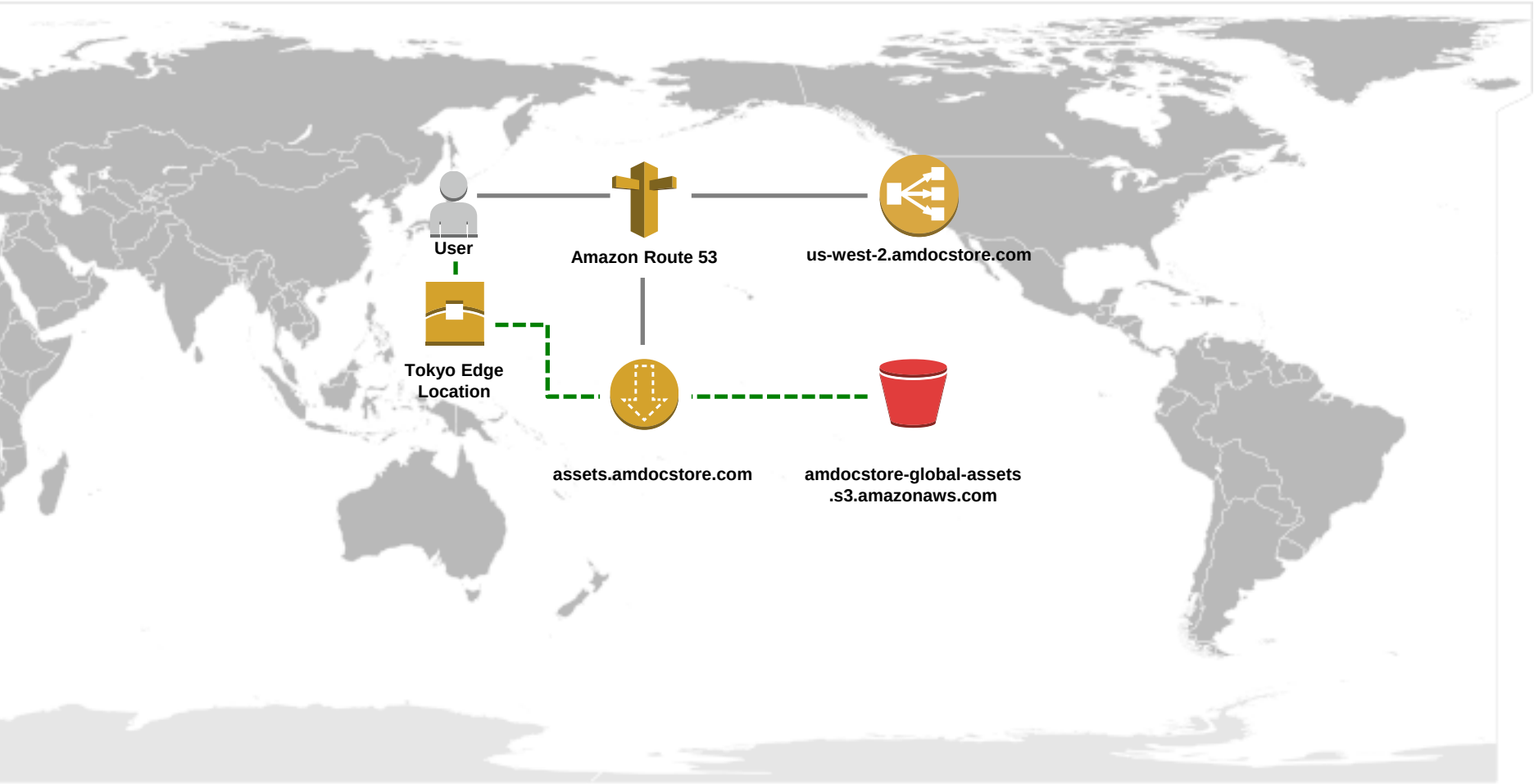


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

The requested file will be served from the Tokyo edge if it is in the cache, or **retrieved from the origin and cached at the edge** for future requests

DocStore

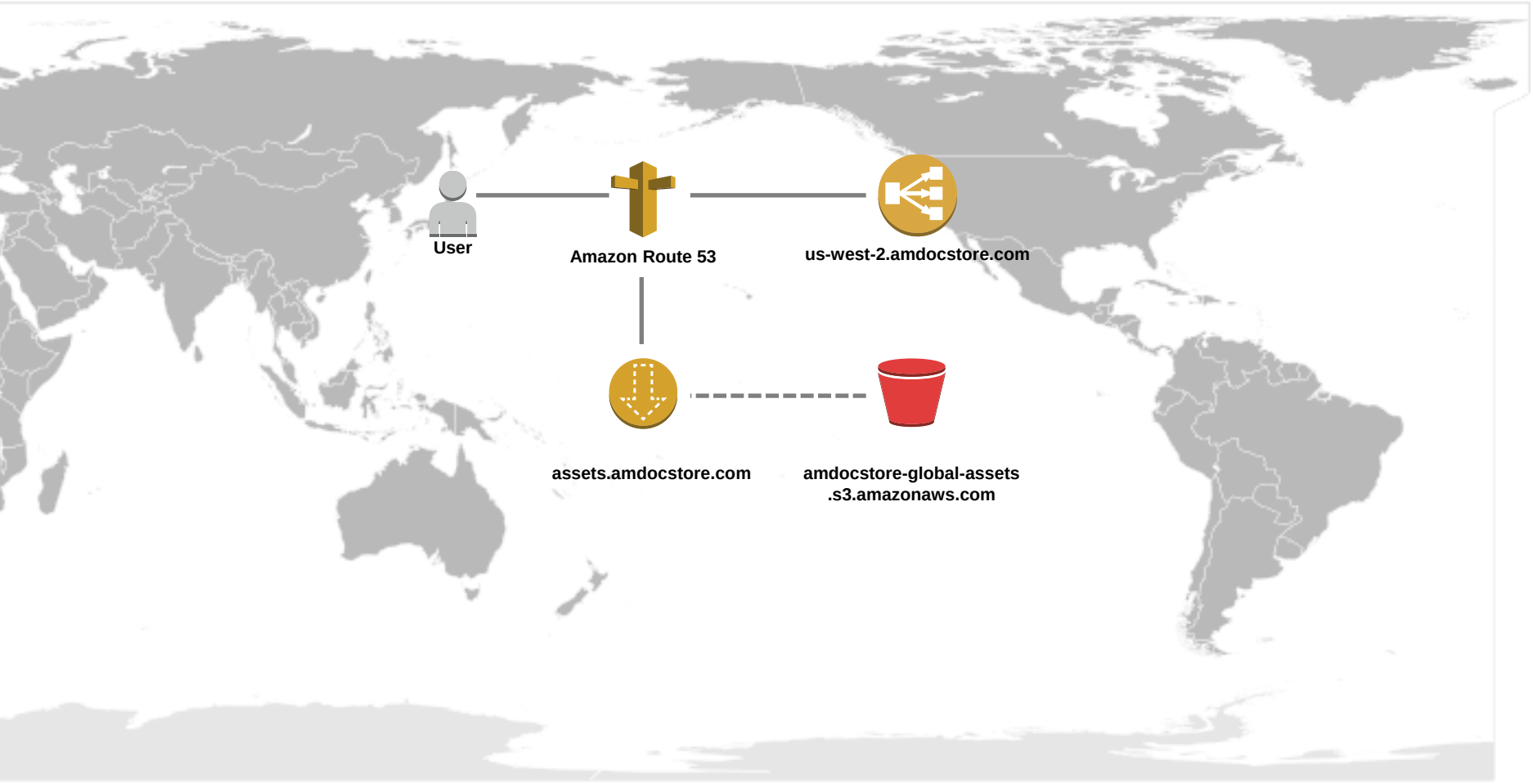


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Let's go back to our application deployment in **us-west-2**

DocStore

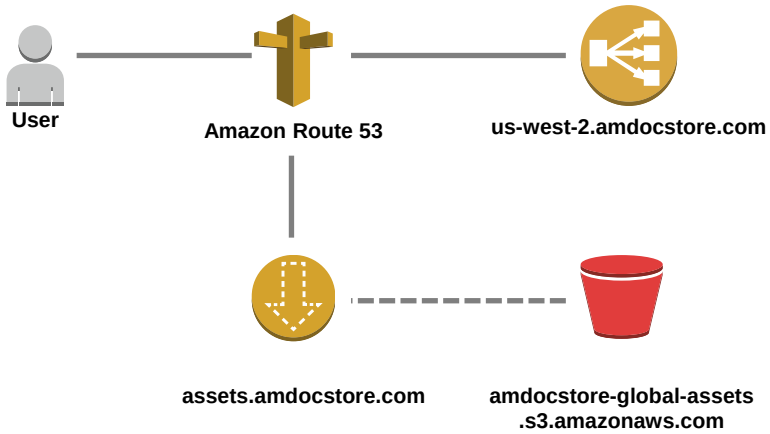


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Let’s go back to our application deployment in **us-west-2**

DocStore

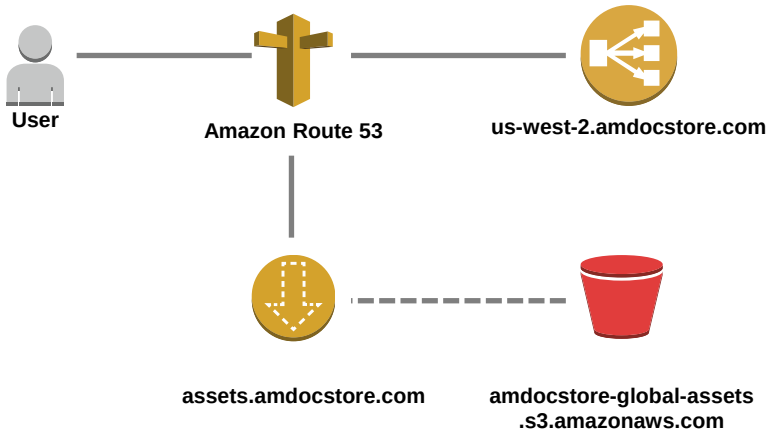


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Let's go back to our application deployment in **us-west-2** and **focus on the browser component of DocStore**

DocStore

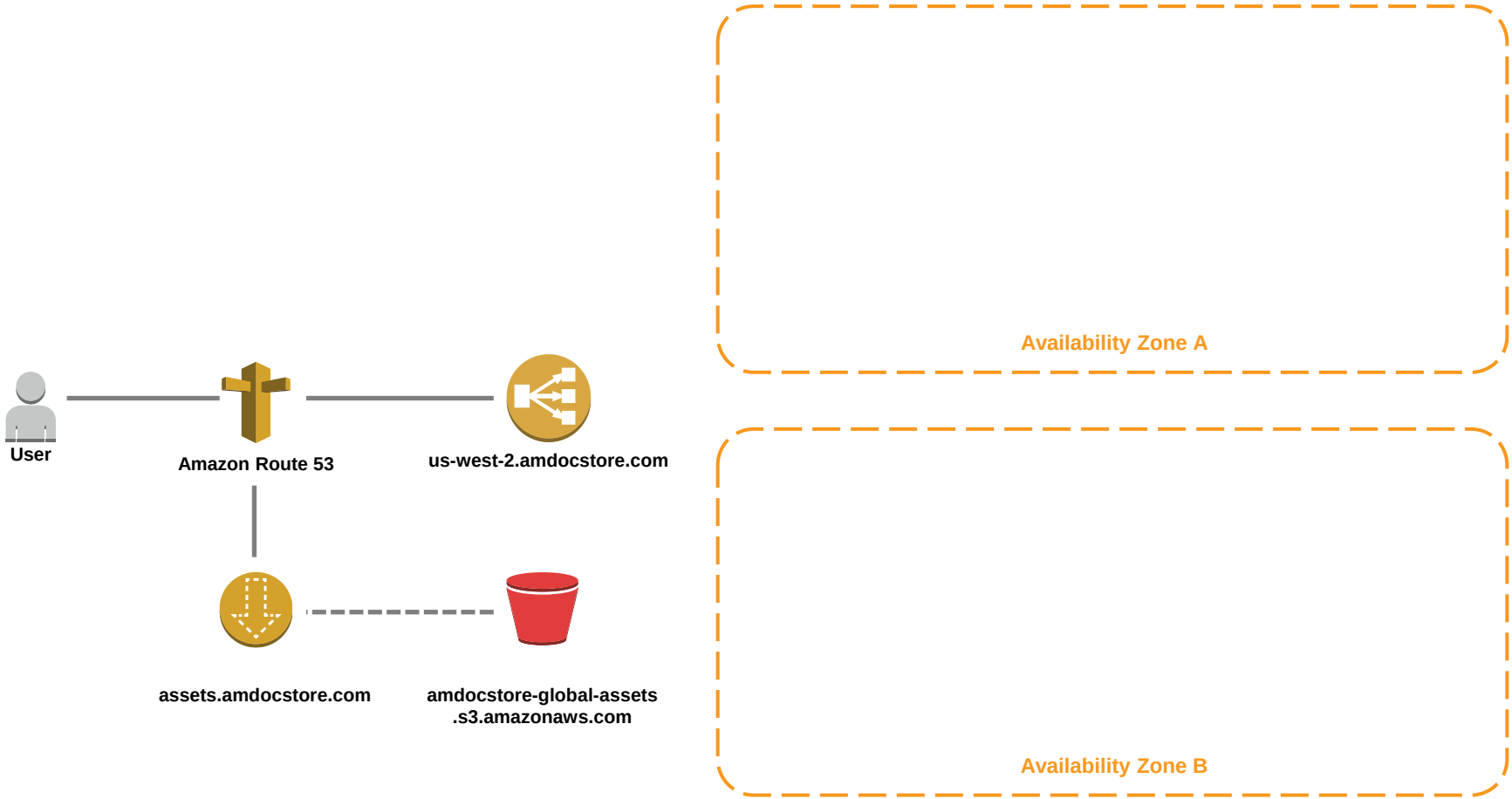


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Web servers will be deployed across multiple Availability Zones using Auto Scaling

DocStore

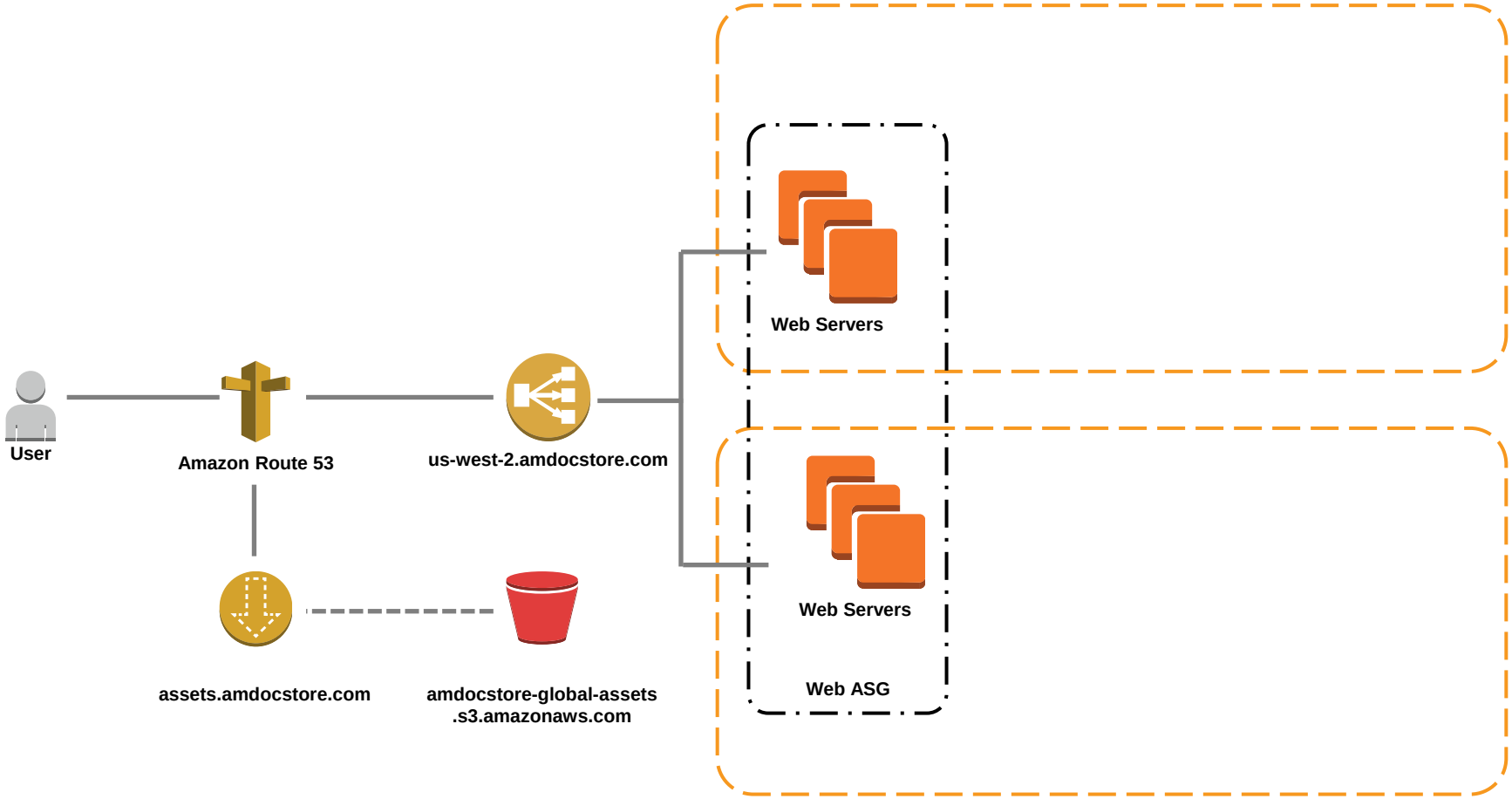


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Web servers will be deployed across multiple Availability Zones using Auto Scaling

DocStore

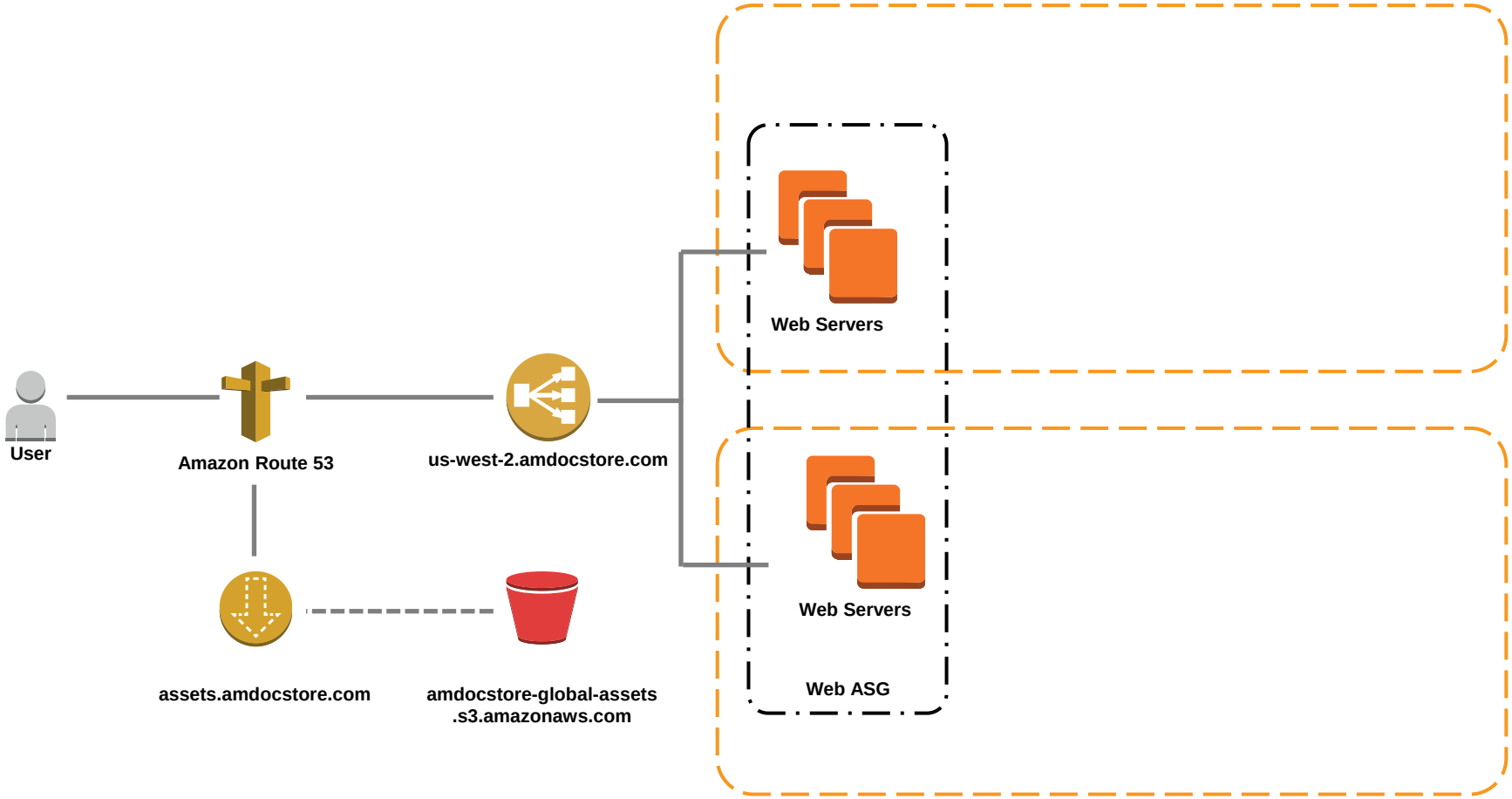


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Web servers host the DocStore interface that users will access with their web browsers

DocStore

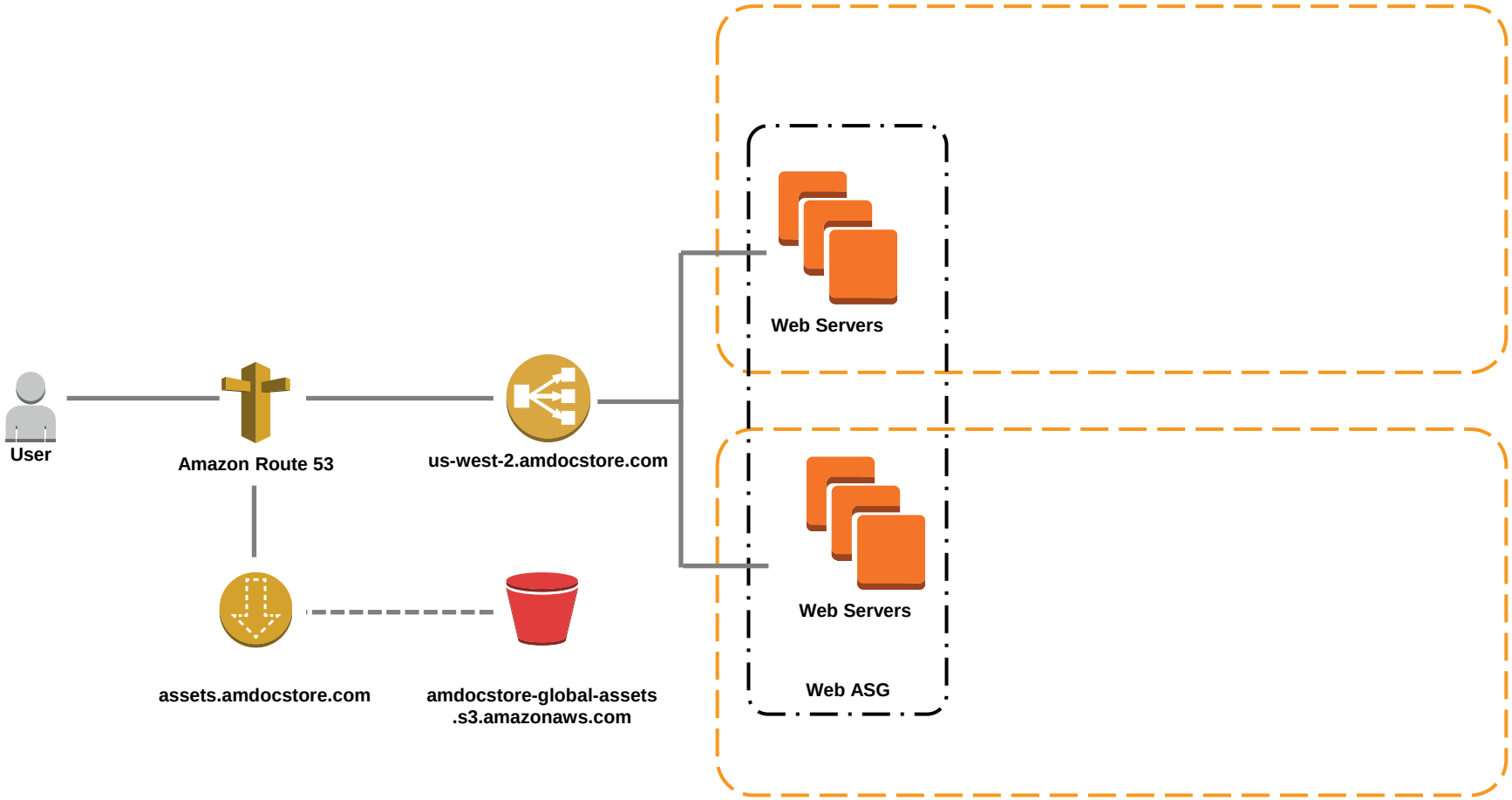


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Web servers might run a Ruby/Rails, PHP, Python, or other tool to render the user's interface/view

DocStore

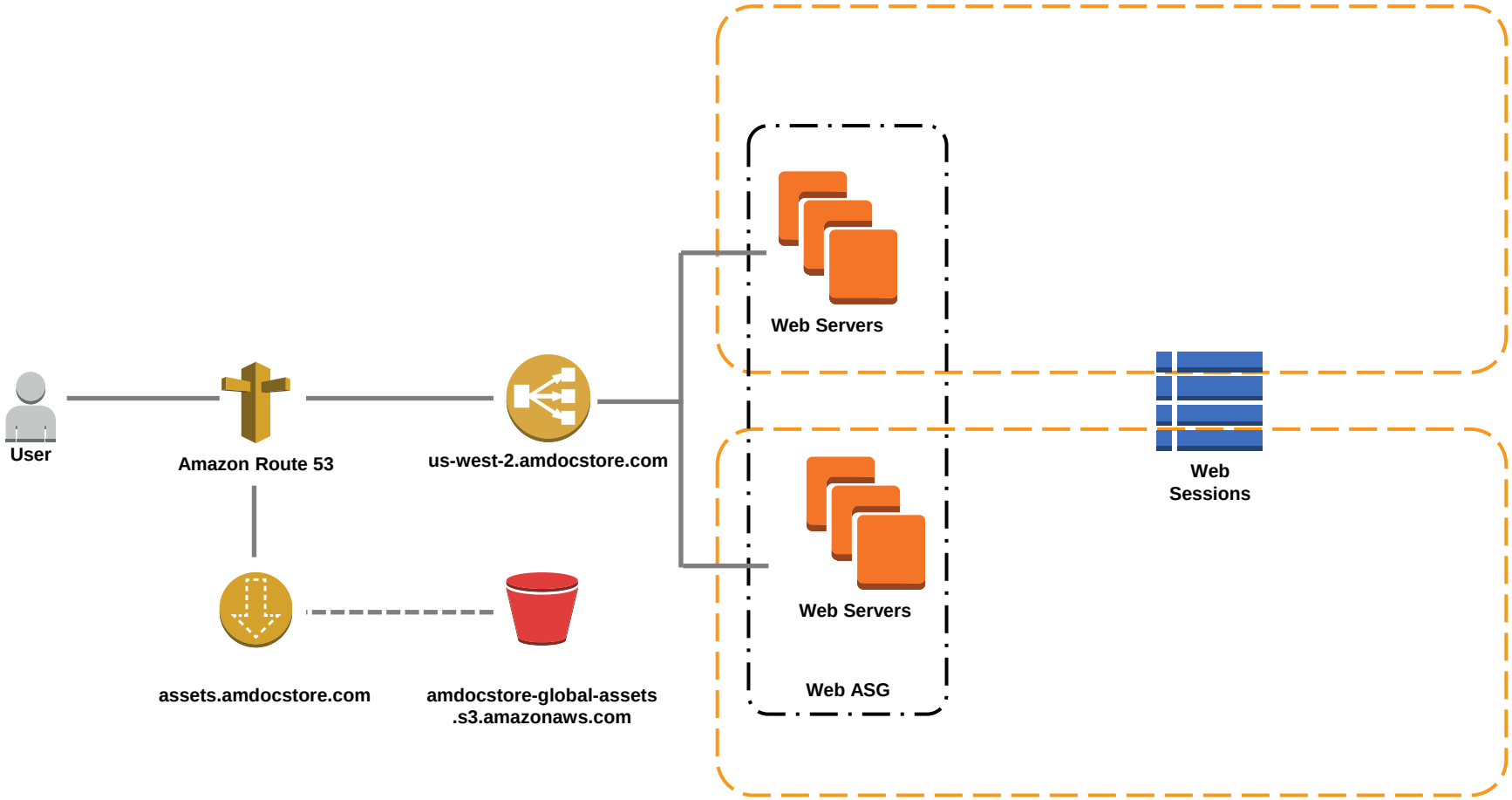


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Session state is stored out of the web tier in a **DynamoDB** table. DynamoDB is a regional service, meaning the data is automatically distributed across AZs

DocStore

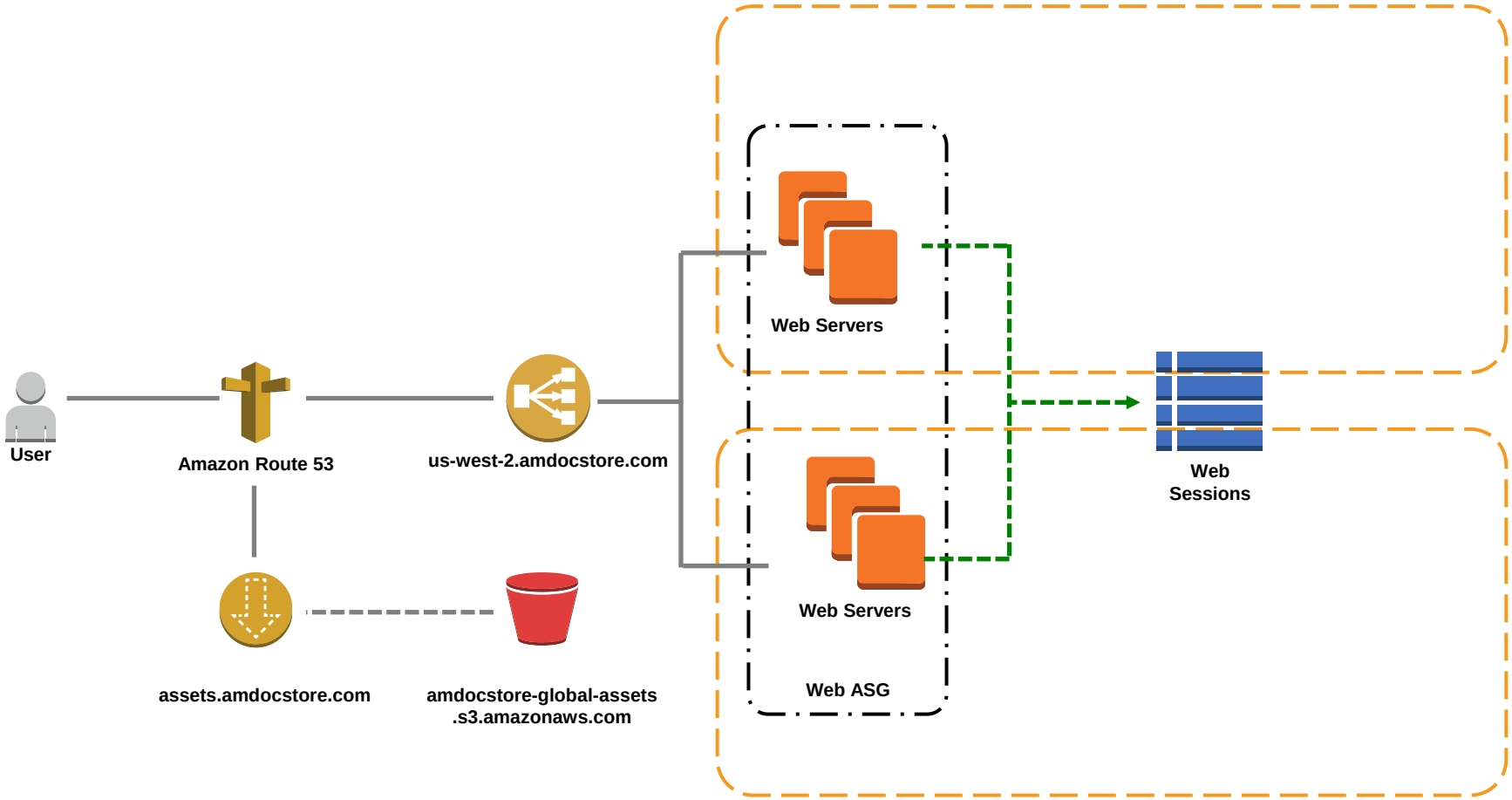


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Session state is stored out of the web tier in a **DynamoDB** table. DynamoDB is a regional service, meaning the data is automatically distributed across AZs

DocStore

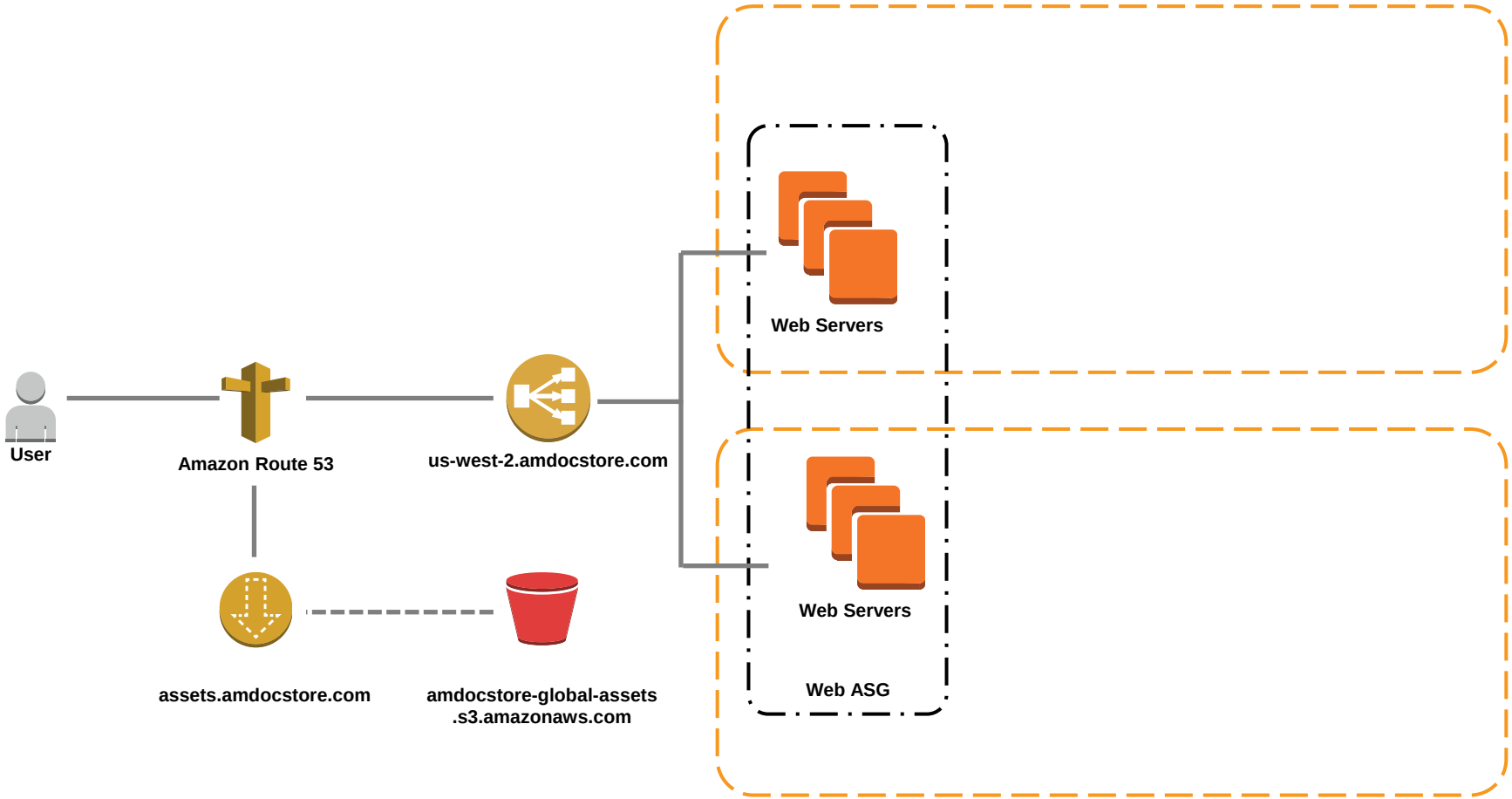


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Core functionality is provided by the **DocStore API.** Web servers consume this API

DocStore

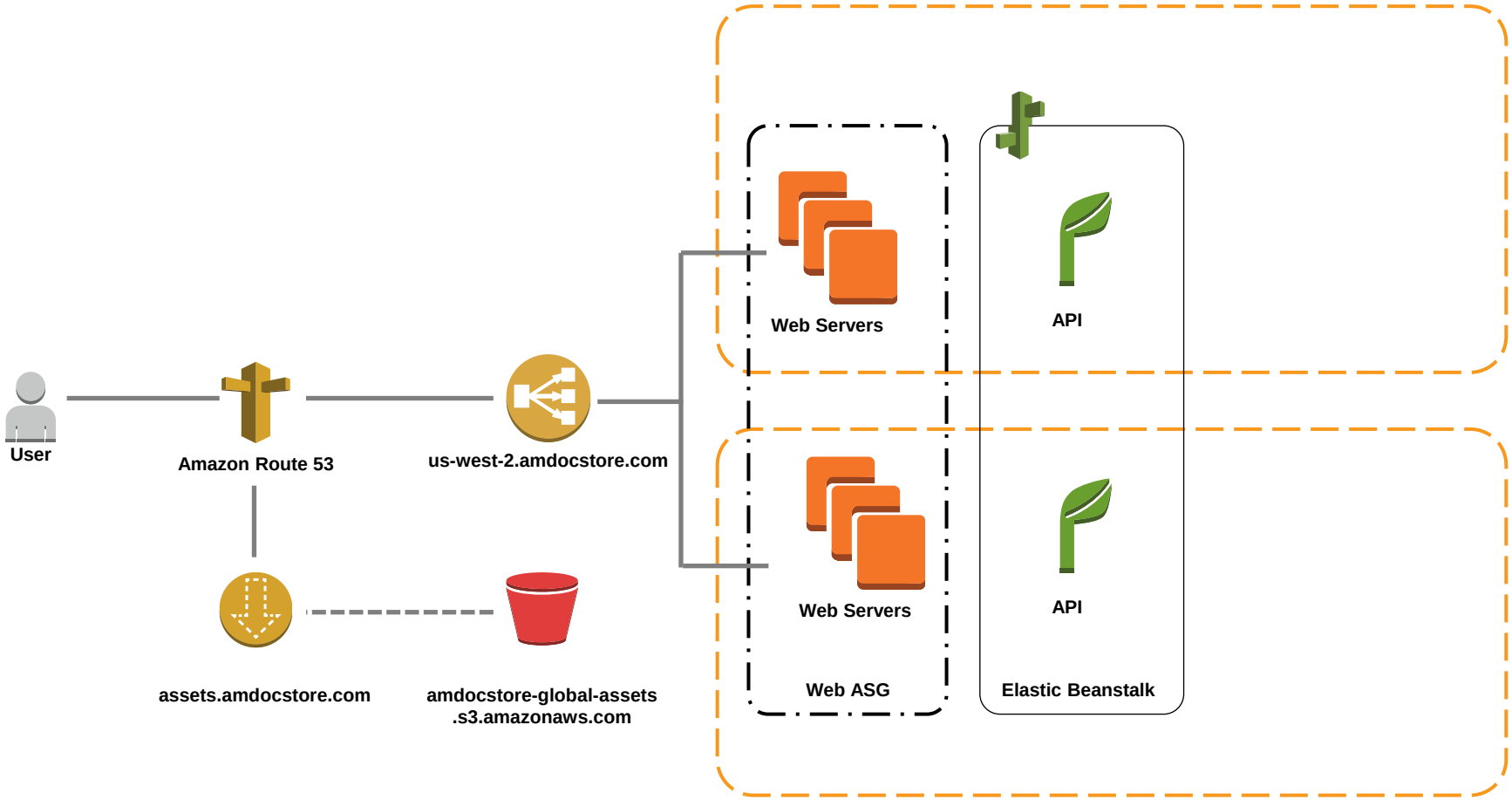


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

The DocStore API will be deployed as an application in an **Elastic Beanstalk container**

DocStore

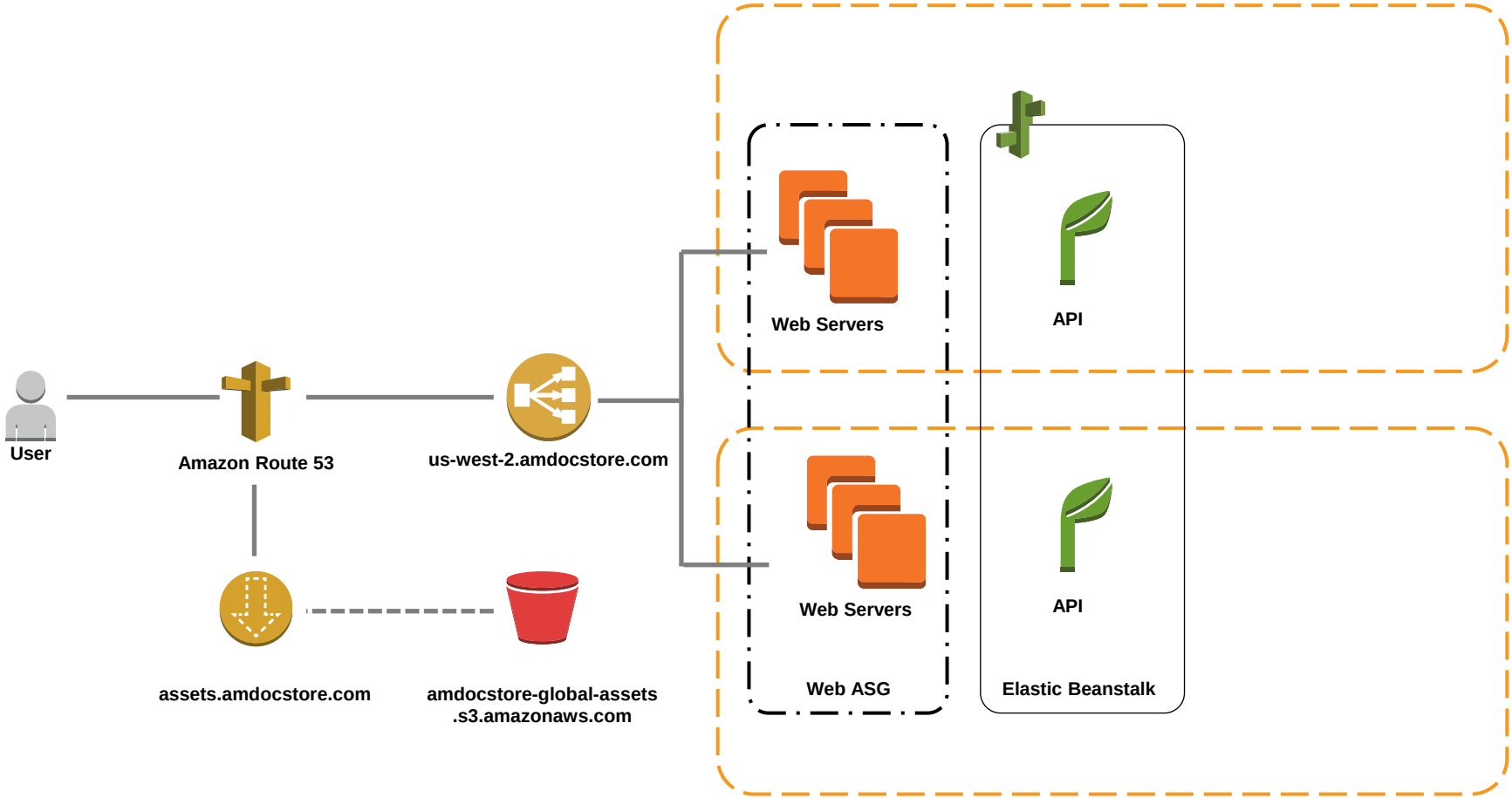


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Elastic Beanstalk provides Multi-AZ deployment, Auto Scaling, and load balancing

DocStore

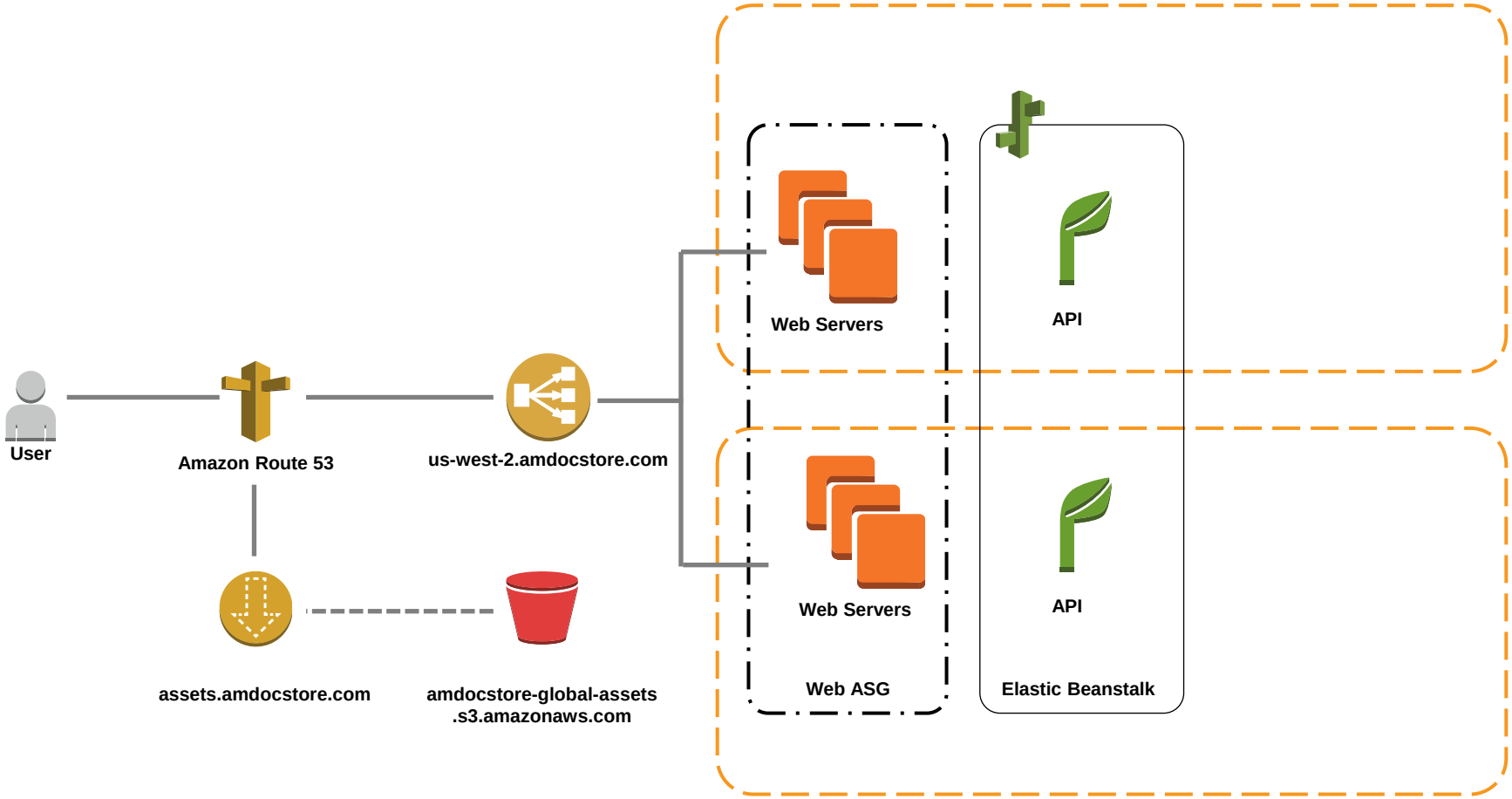


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Elastic Beanstalk supports Java, .NET, Python, PHP, Ruby, and node.js applications

DocStore

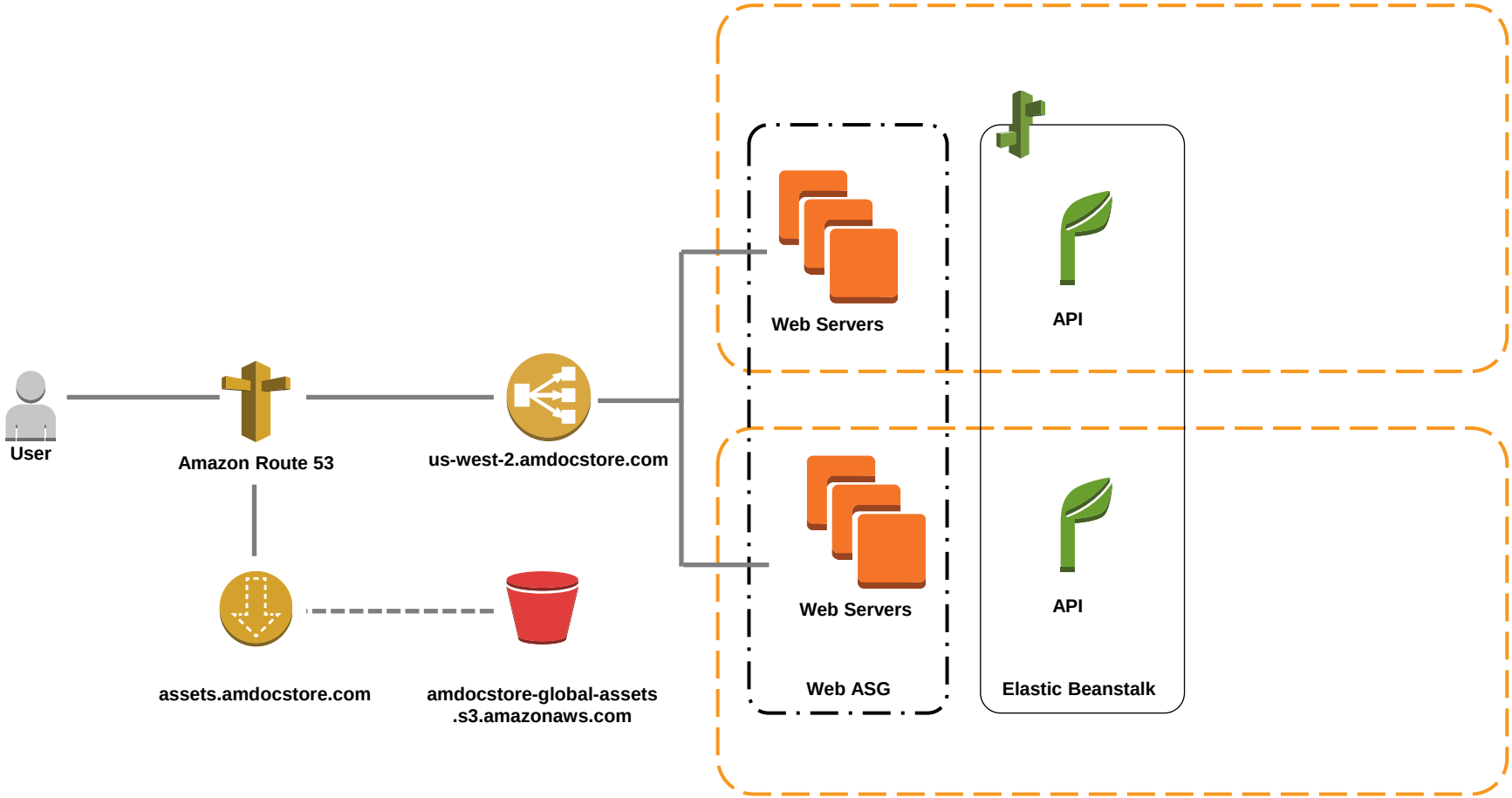


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

The API will use several different data stores for different types of data

DocStore

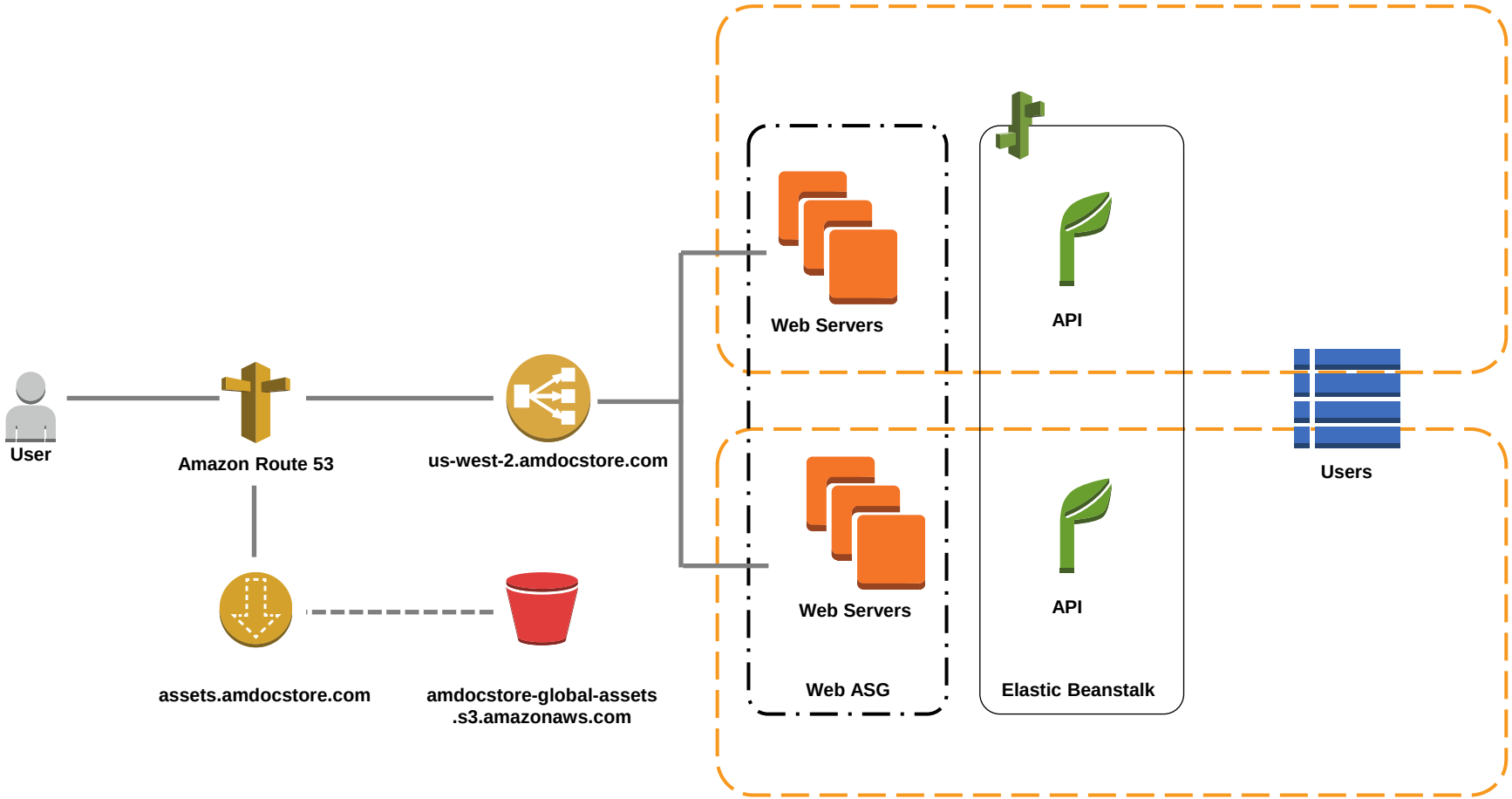


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

User account information will be stored in a **DynamoDB** table

DocStore

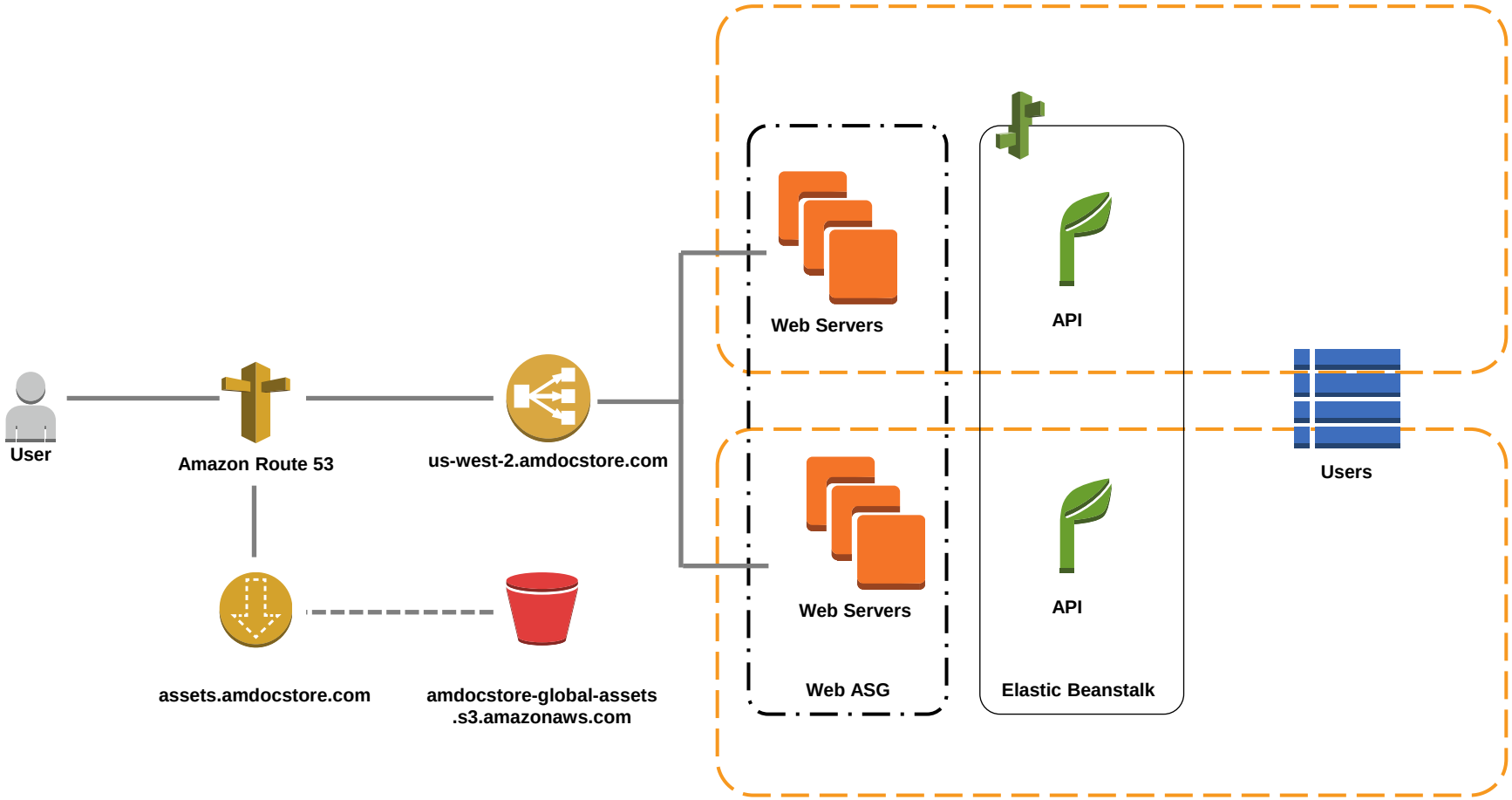


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Let's see what happens when a **new user signs up for DocStore**

DocStore

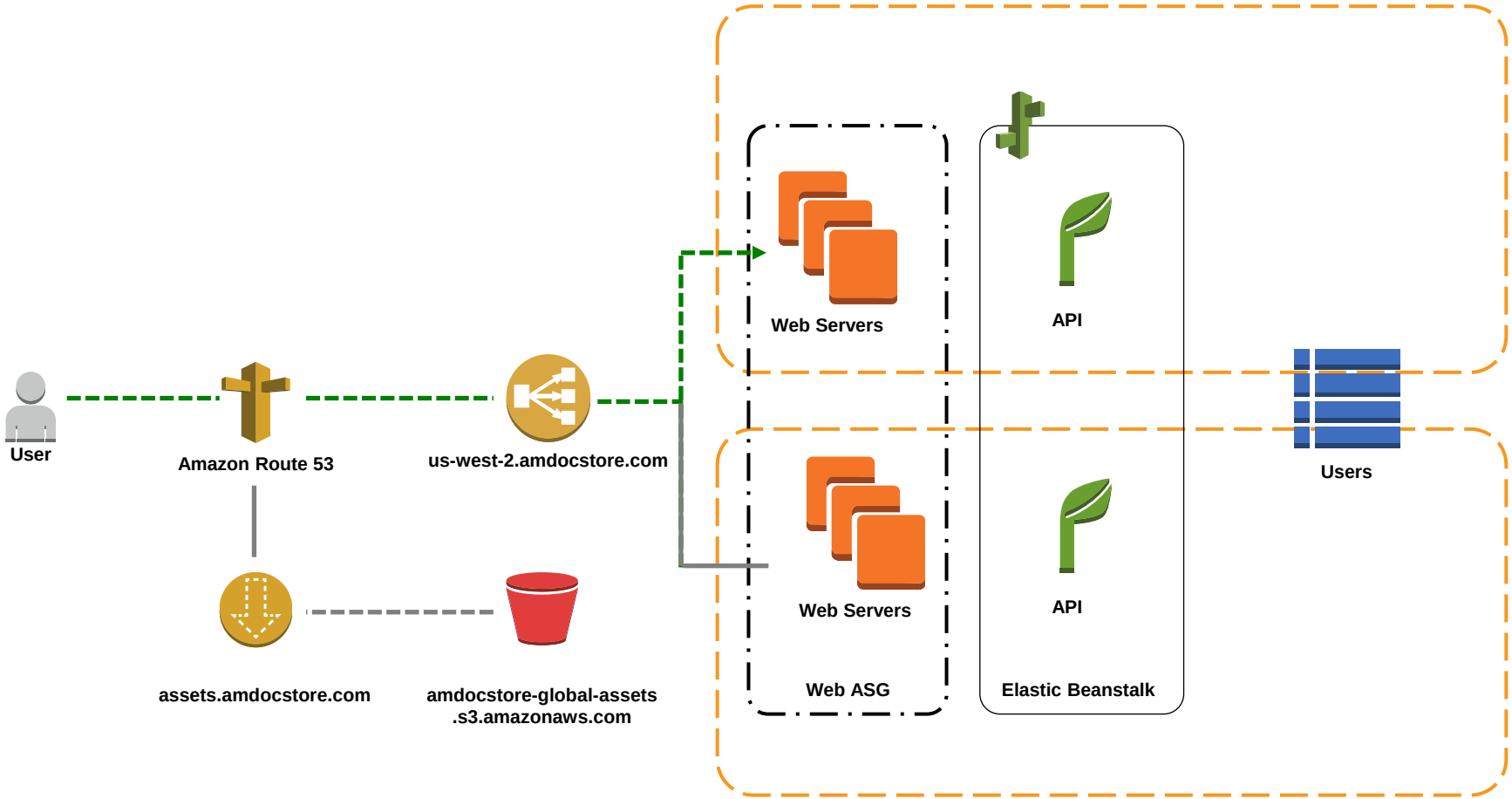


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Create Account: User submits sign-up form to web servers with username, password, etc.

DocStore

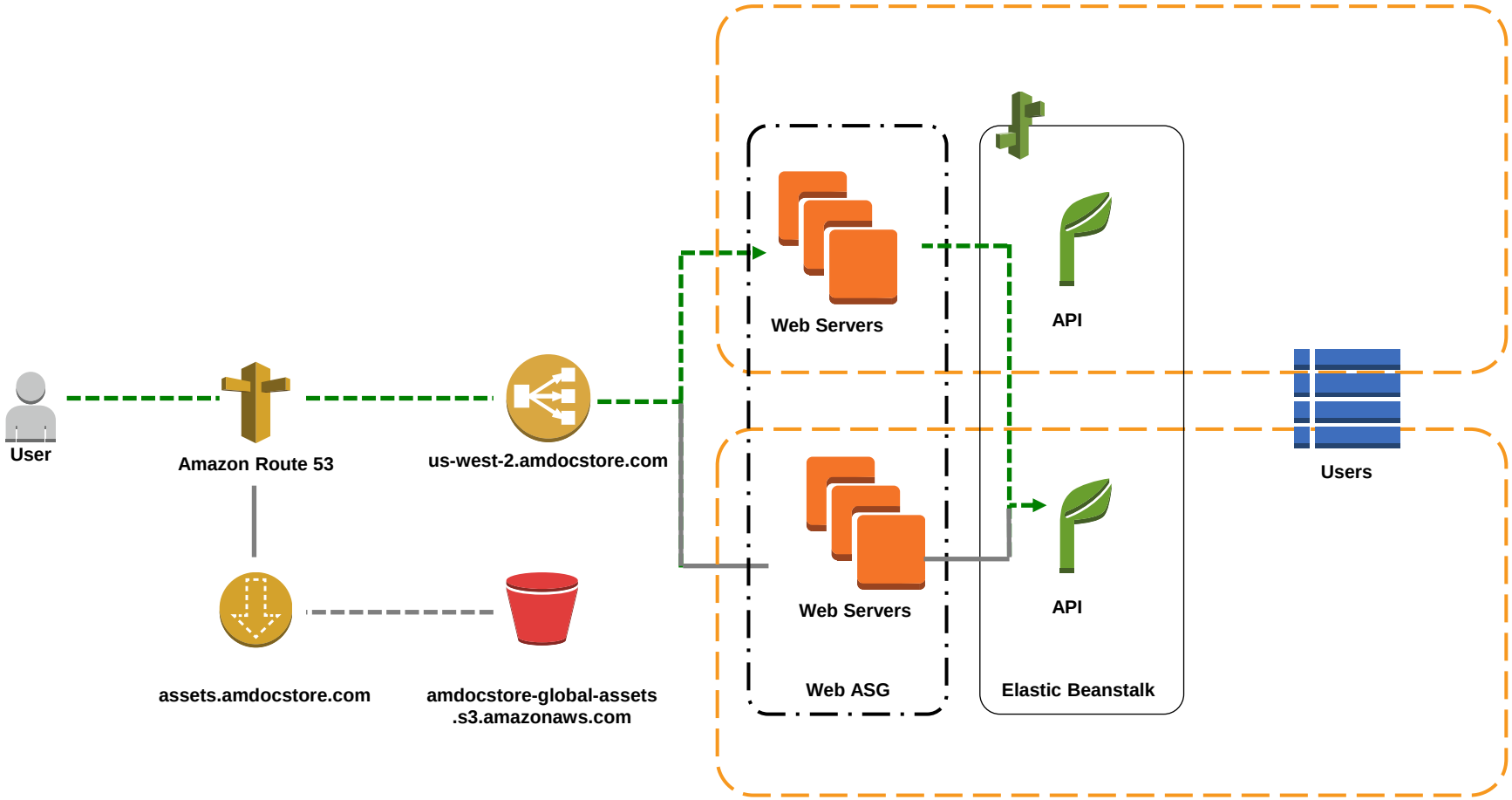


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Create Account: Web servers invoke CreateAccount API

DocStore

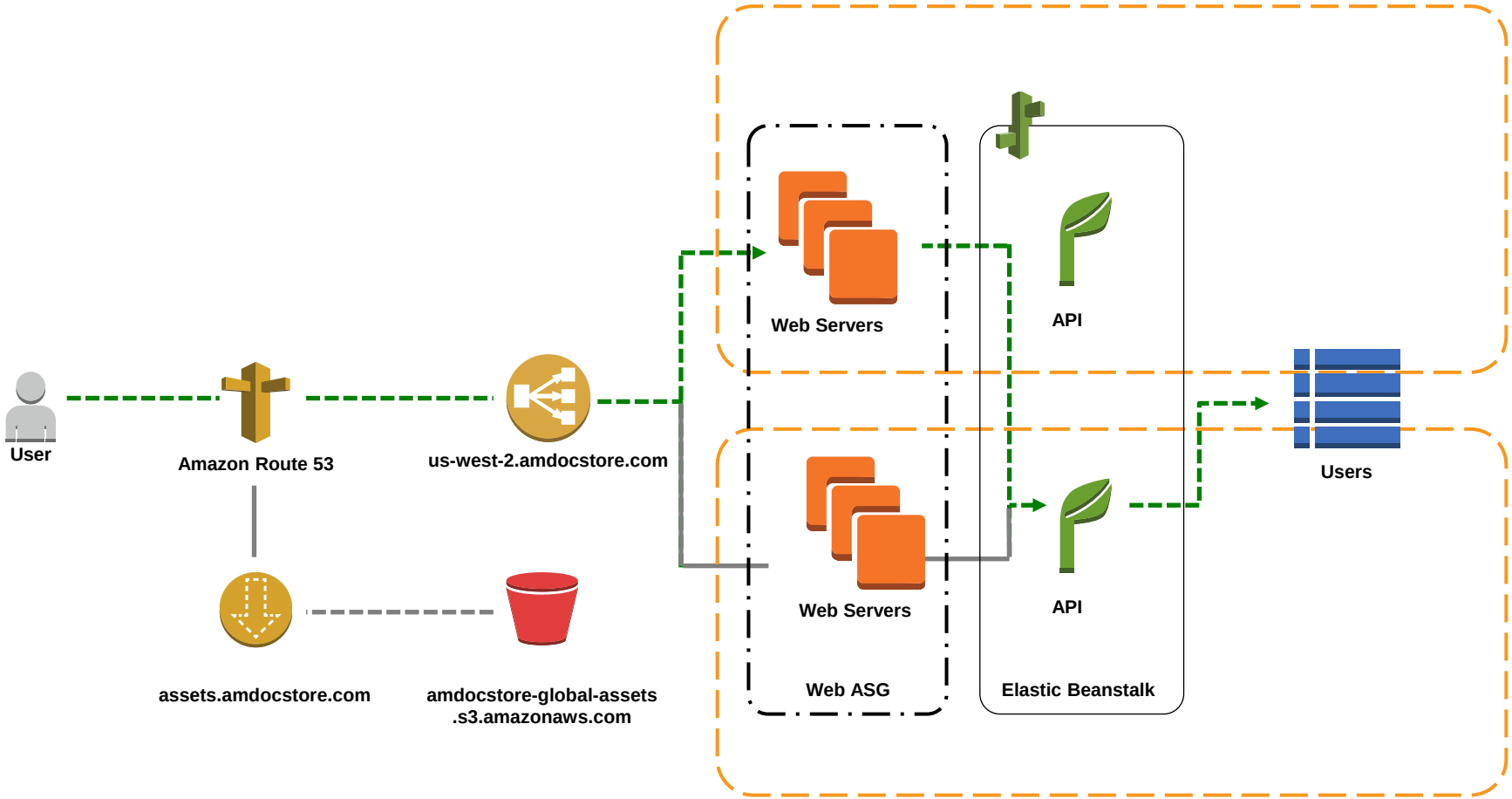


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Create Account: API creates a new item in the DynamoDB table

DocStore

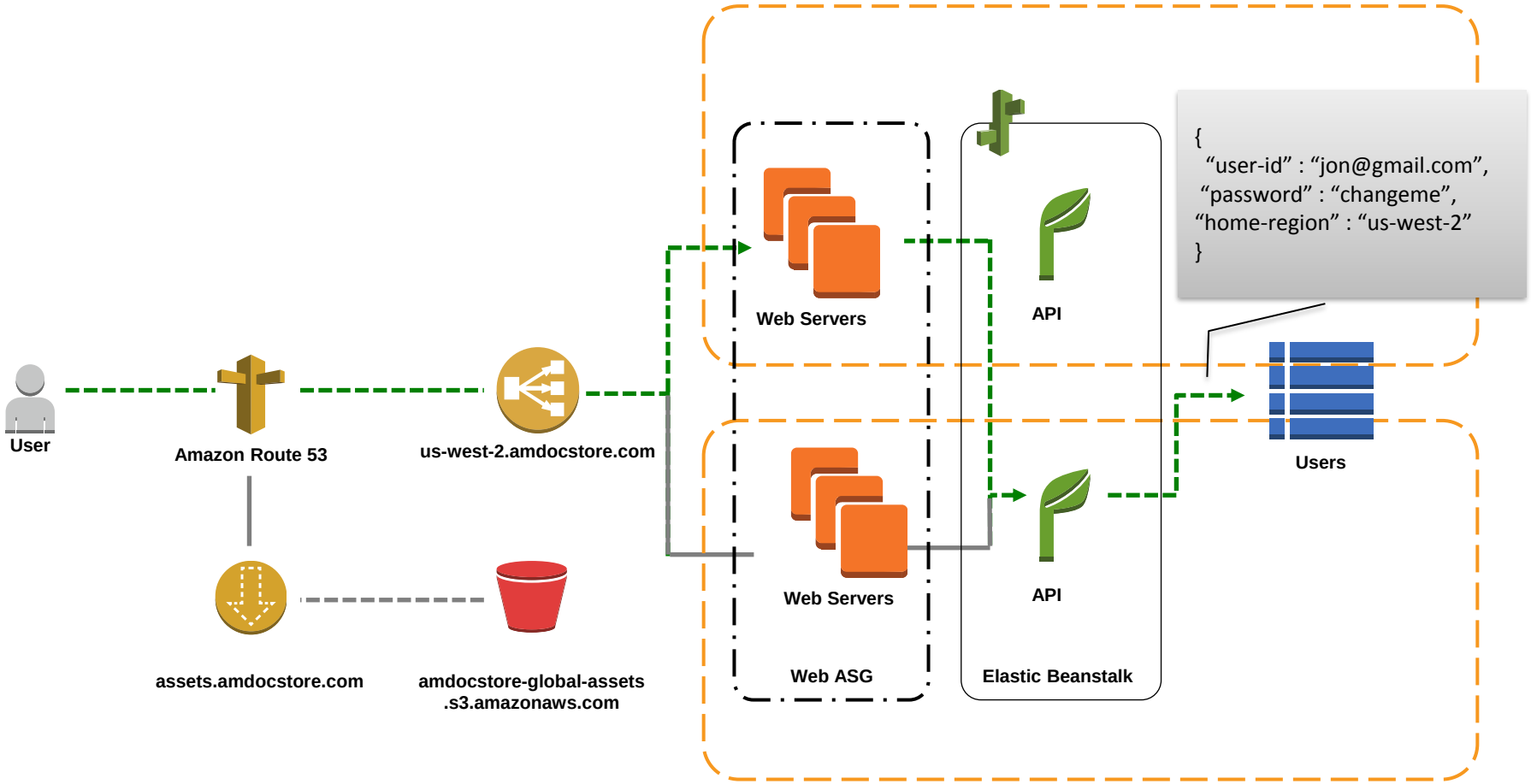


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Create Account: API creates a **new item** in the datastore

DocStore

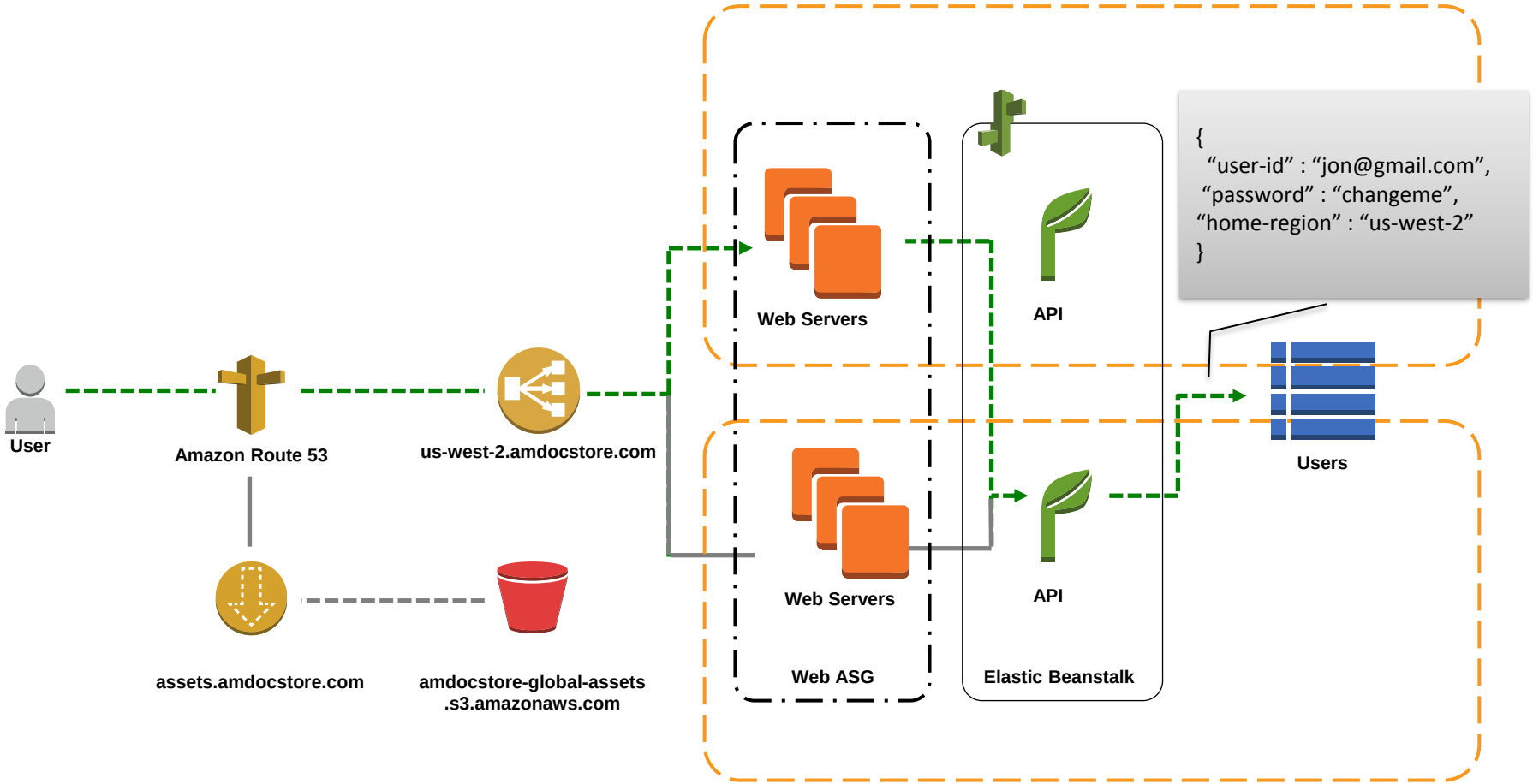


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Create Account: The API app deployed in Elastic Beanstalk has a configuration setting making it aware of the region it is running in

DocStore

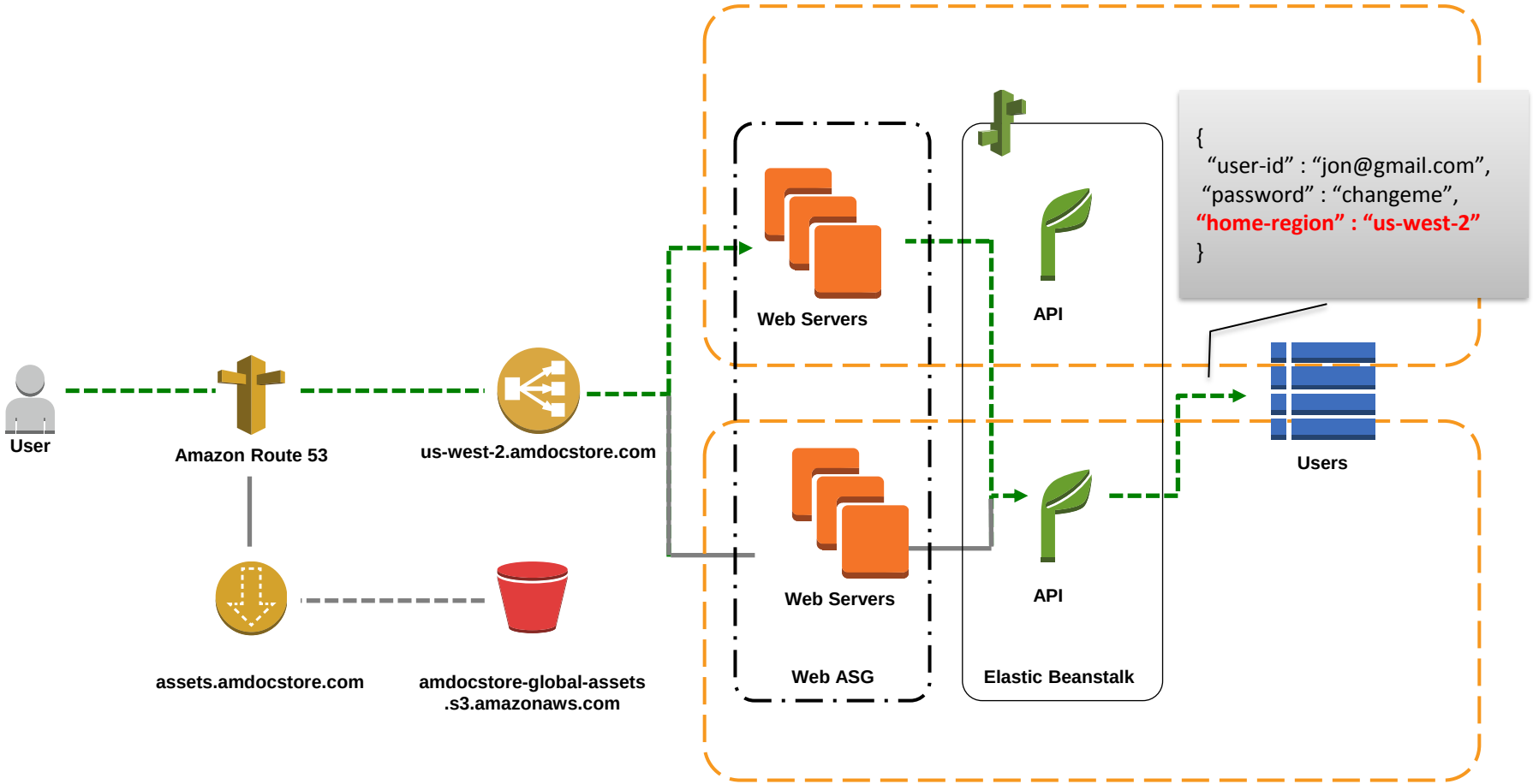


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Create Account: In addition to standard information, the API stores the user's "home region", i.e., the region the user signed up in

DocStore

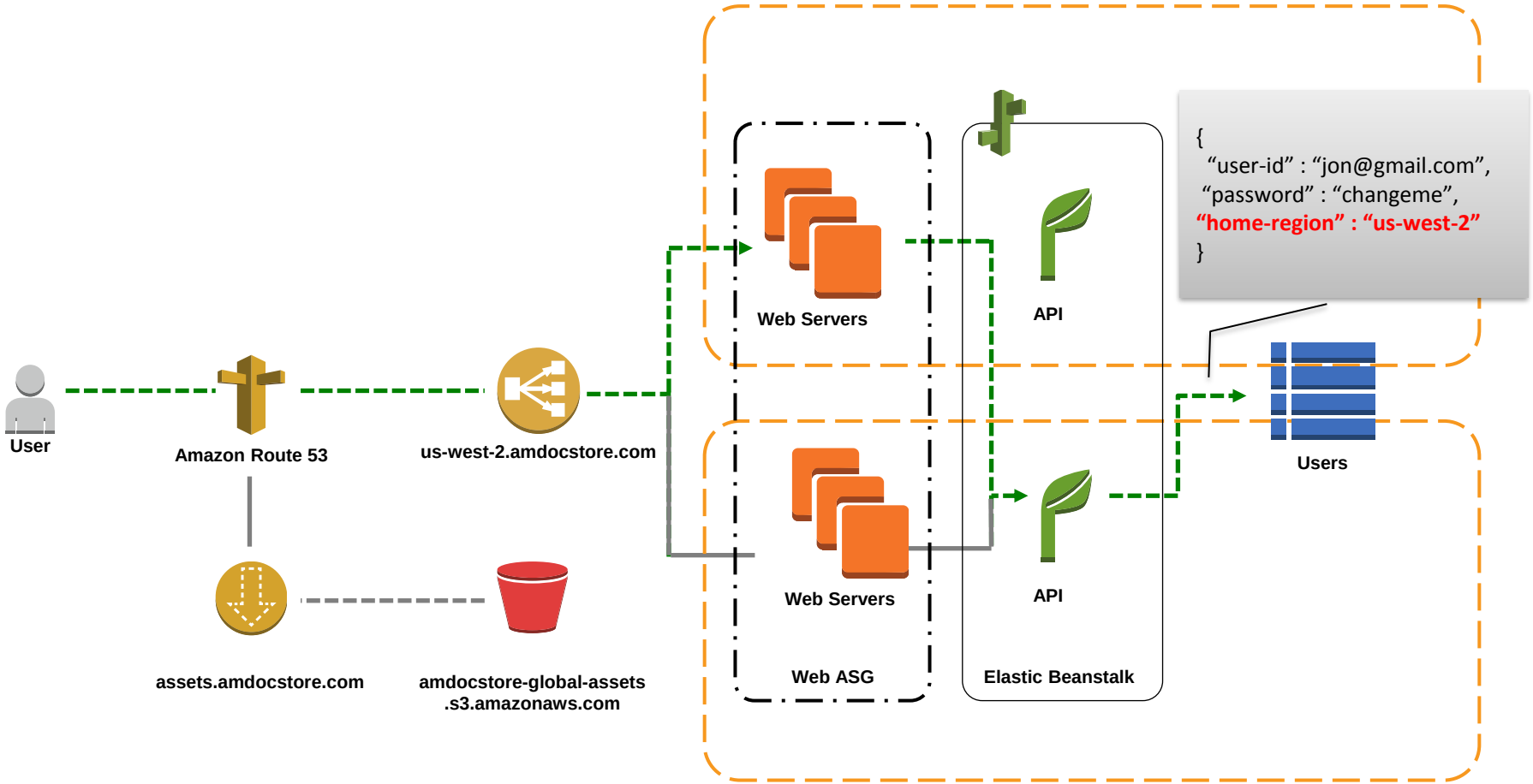


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Create Account: Replicating account data to our Sydney deployment will allow the API in that region to redirect the user to his home region if he ever accesses from there

DocStore

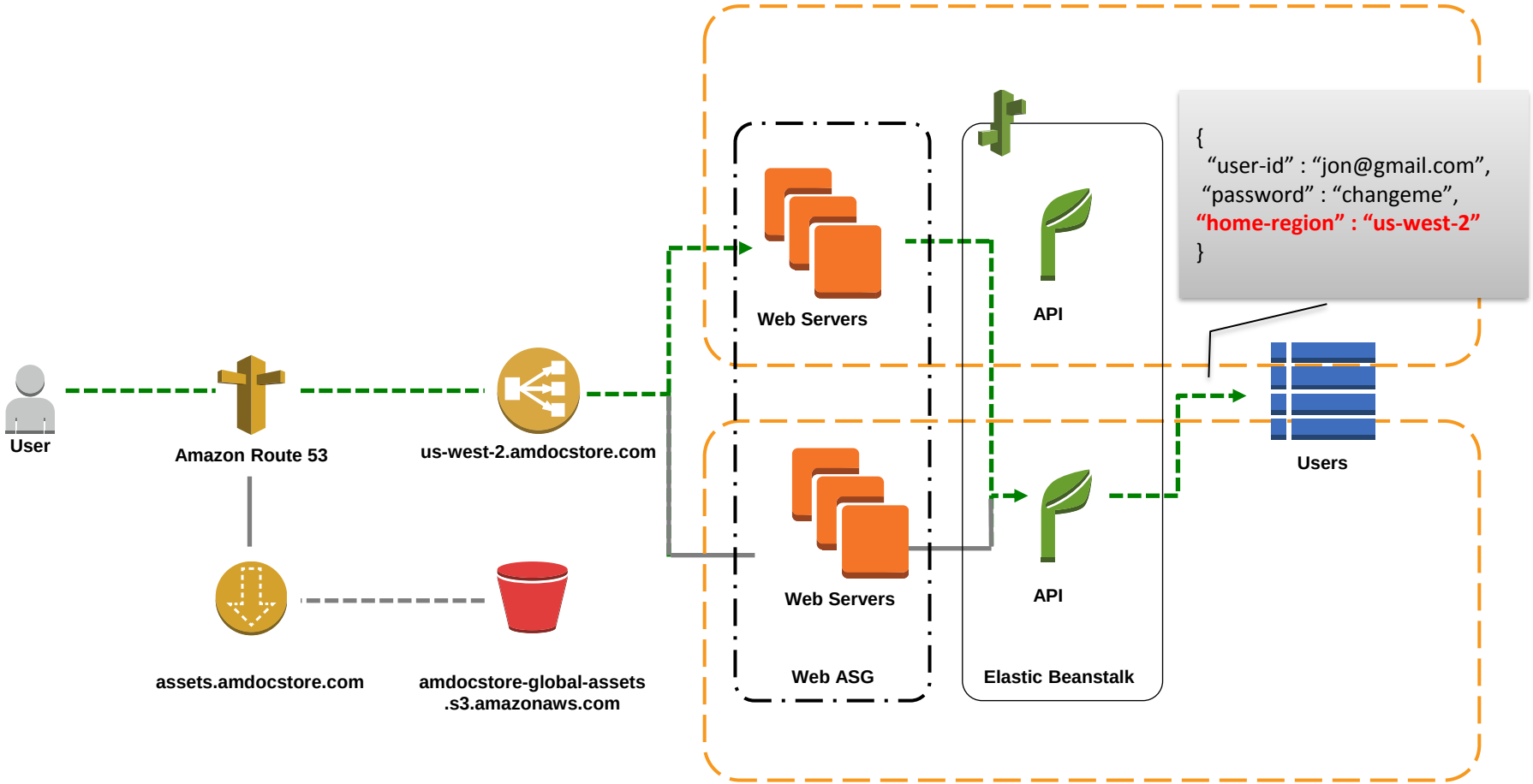


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Create Account: We'll consider **account signup an event** that other components of our application might be interested in.

DocStore

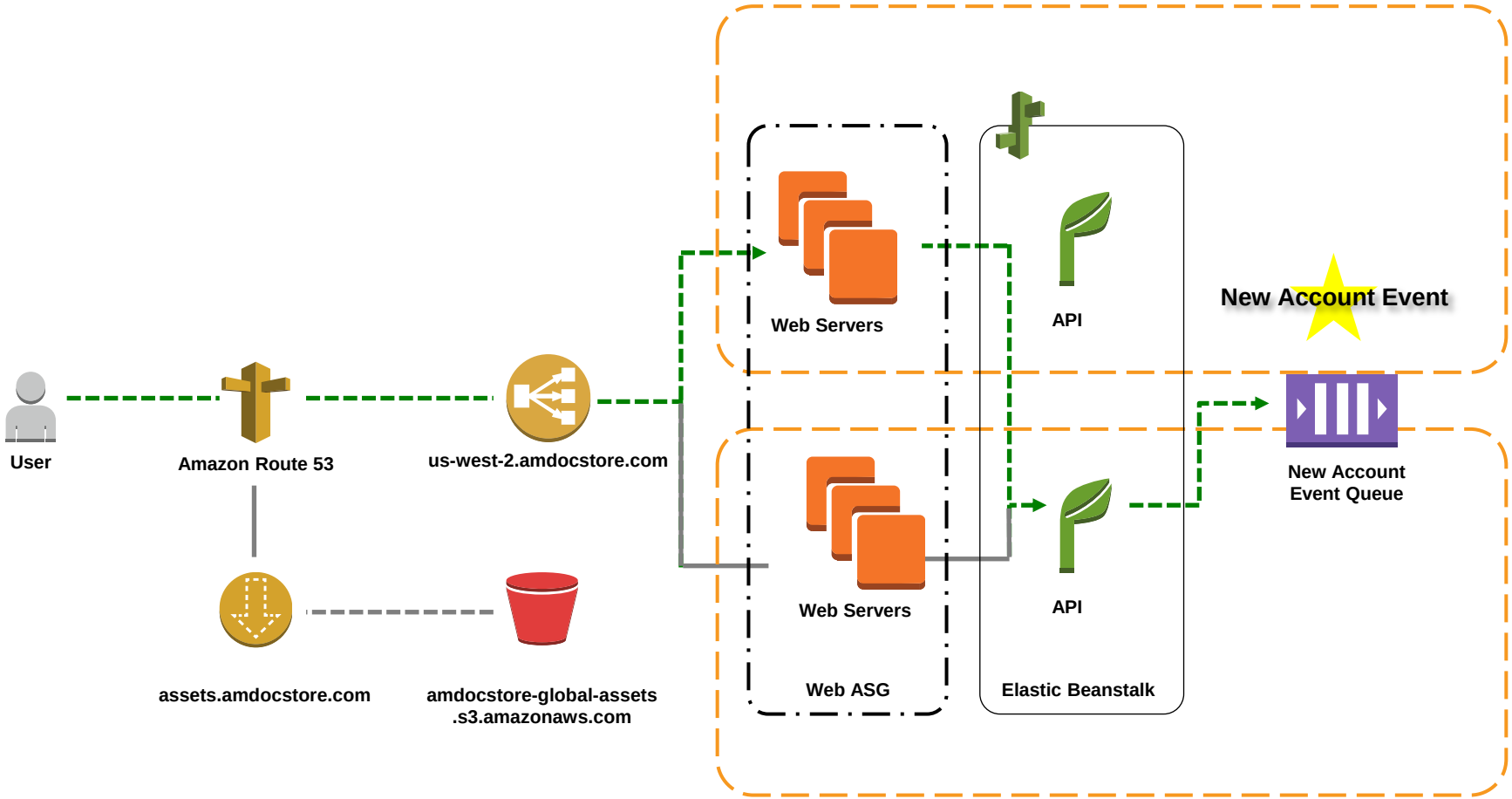


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Create Account: We'll let listeners interested in the event (i.e., a cross-region account replication service) receive the event by putting a message in a queue

DocStore

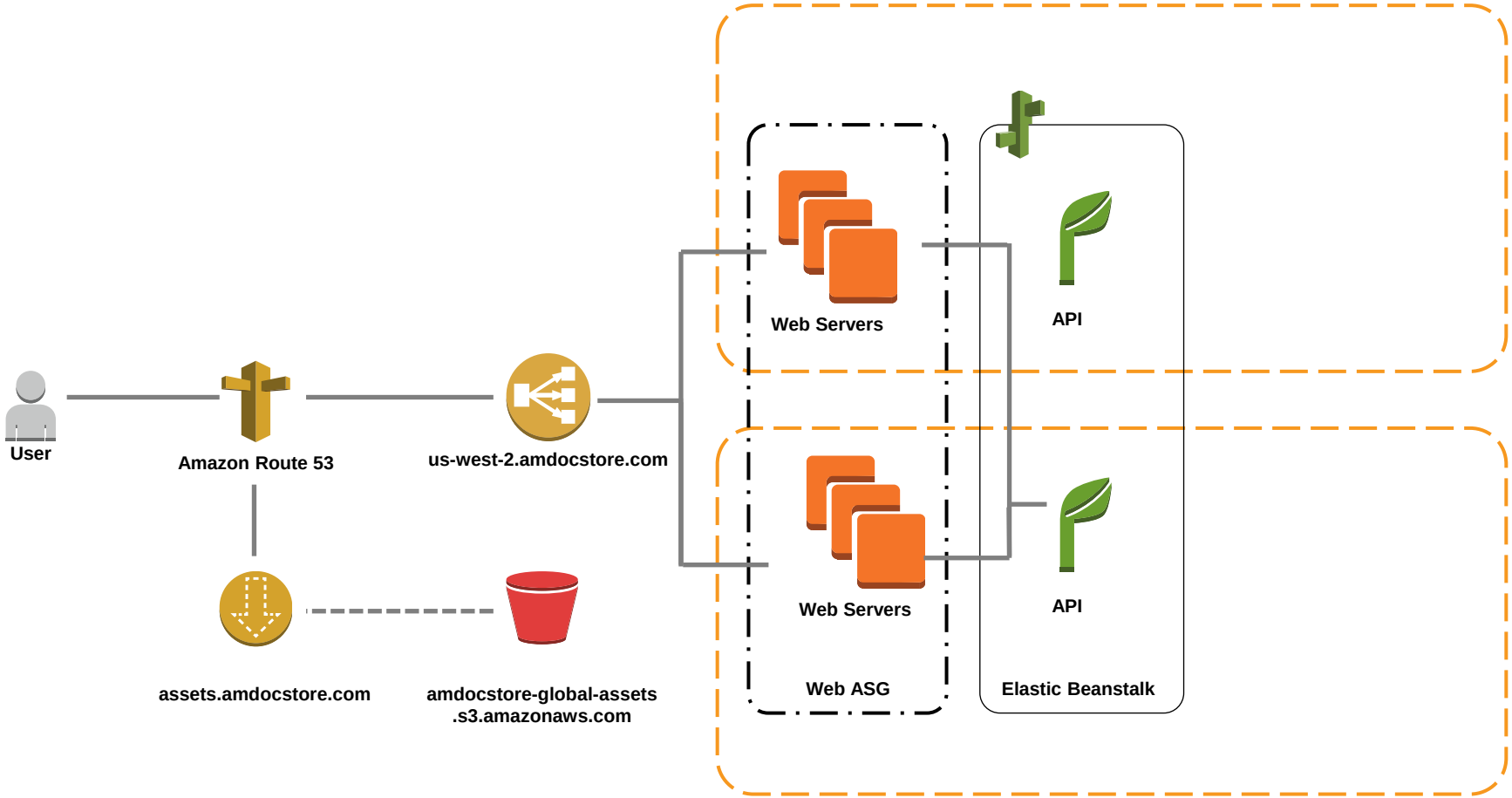


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Documents uploaded by a user are stored in an S3 bucket

DocStore

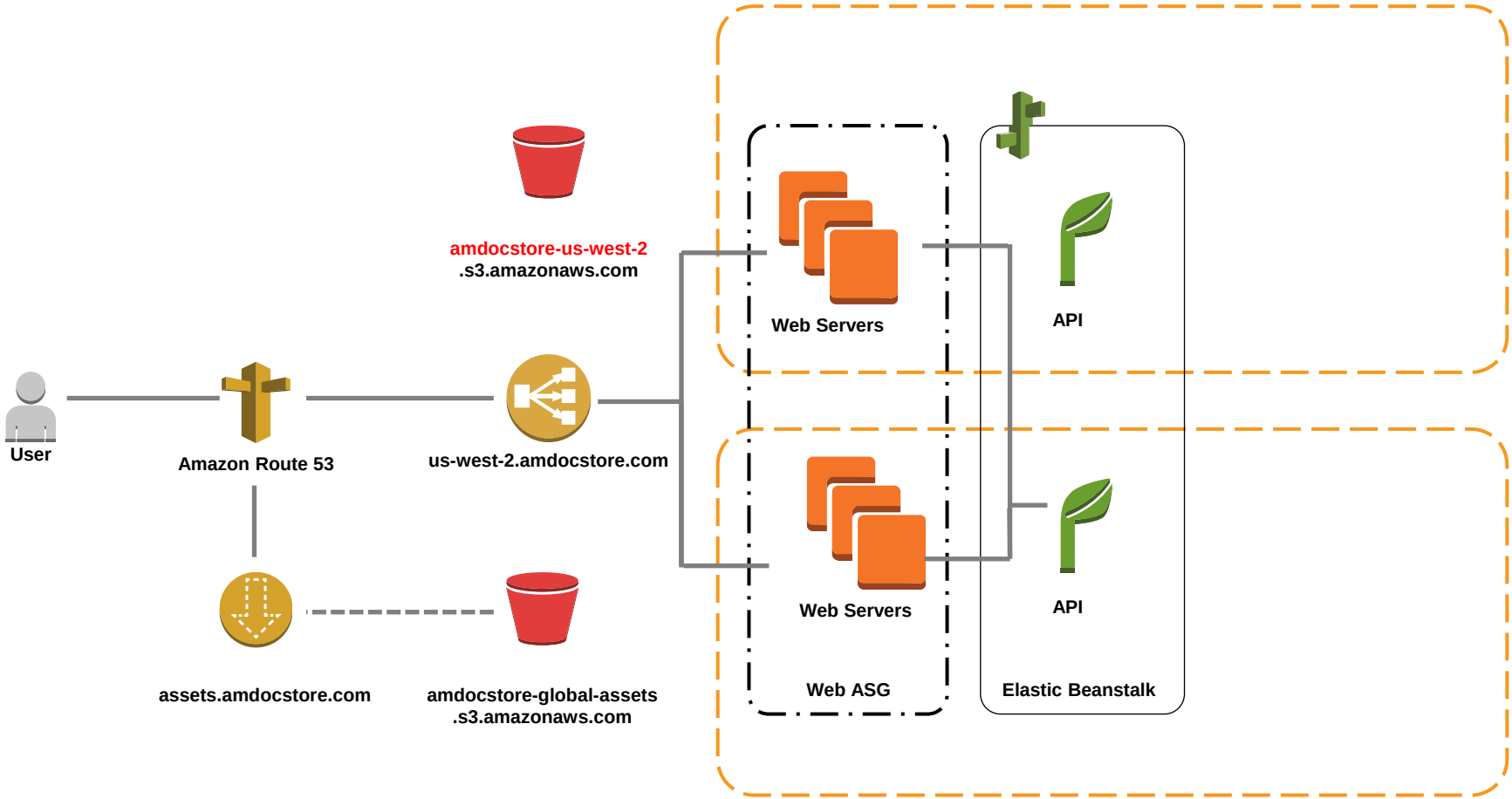


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

The S3 bucket name includes the name of the region it was created it in.

DocStore

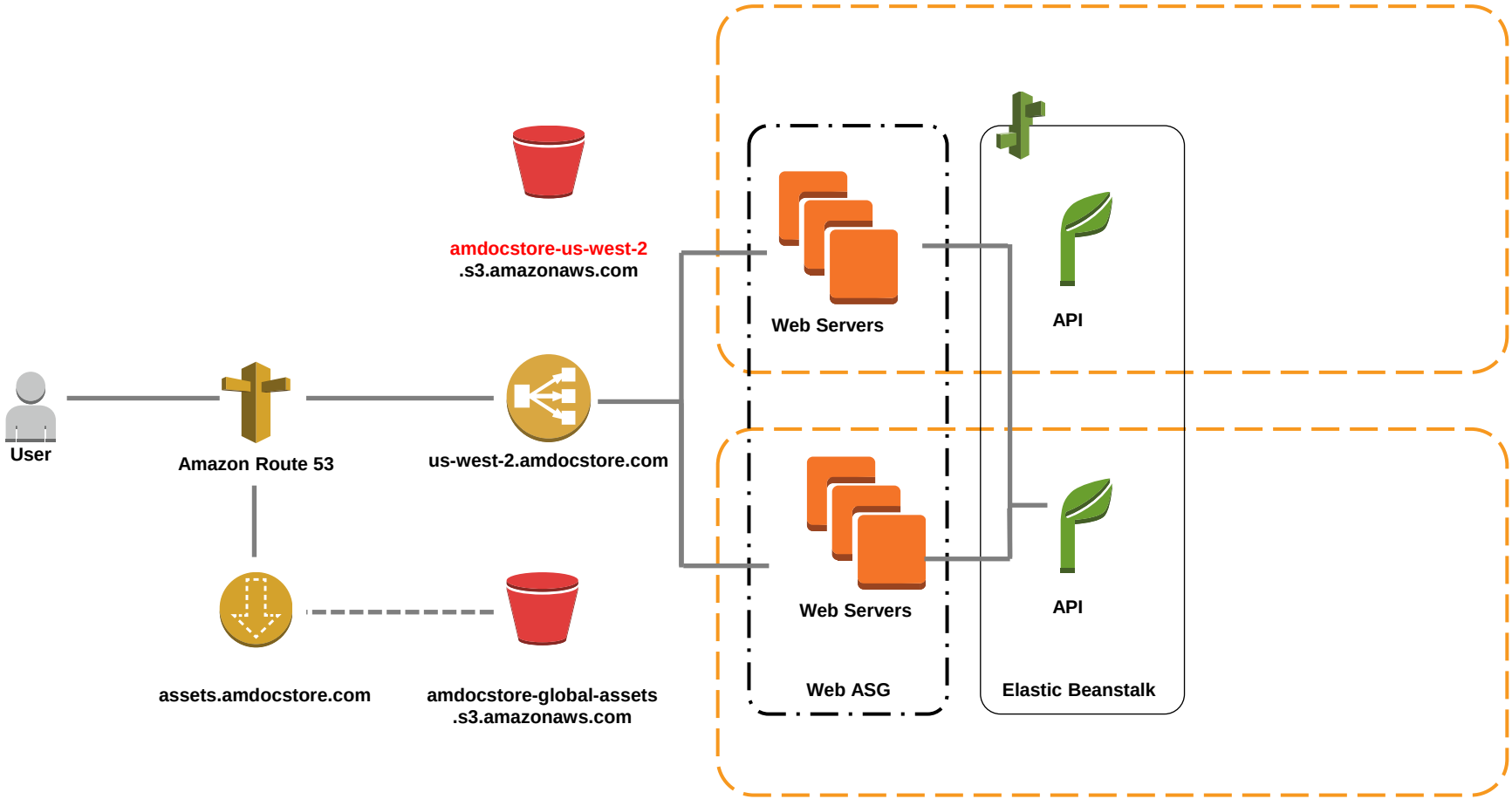


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

The S3 bucket name includes the name of the region it was created it in. Users **upload documents to the bucket in their home region**

DocStore

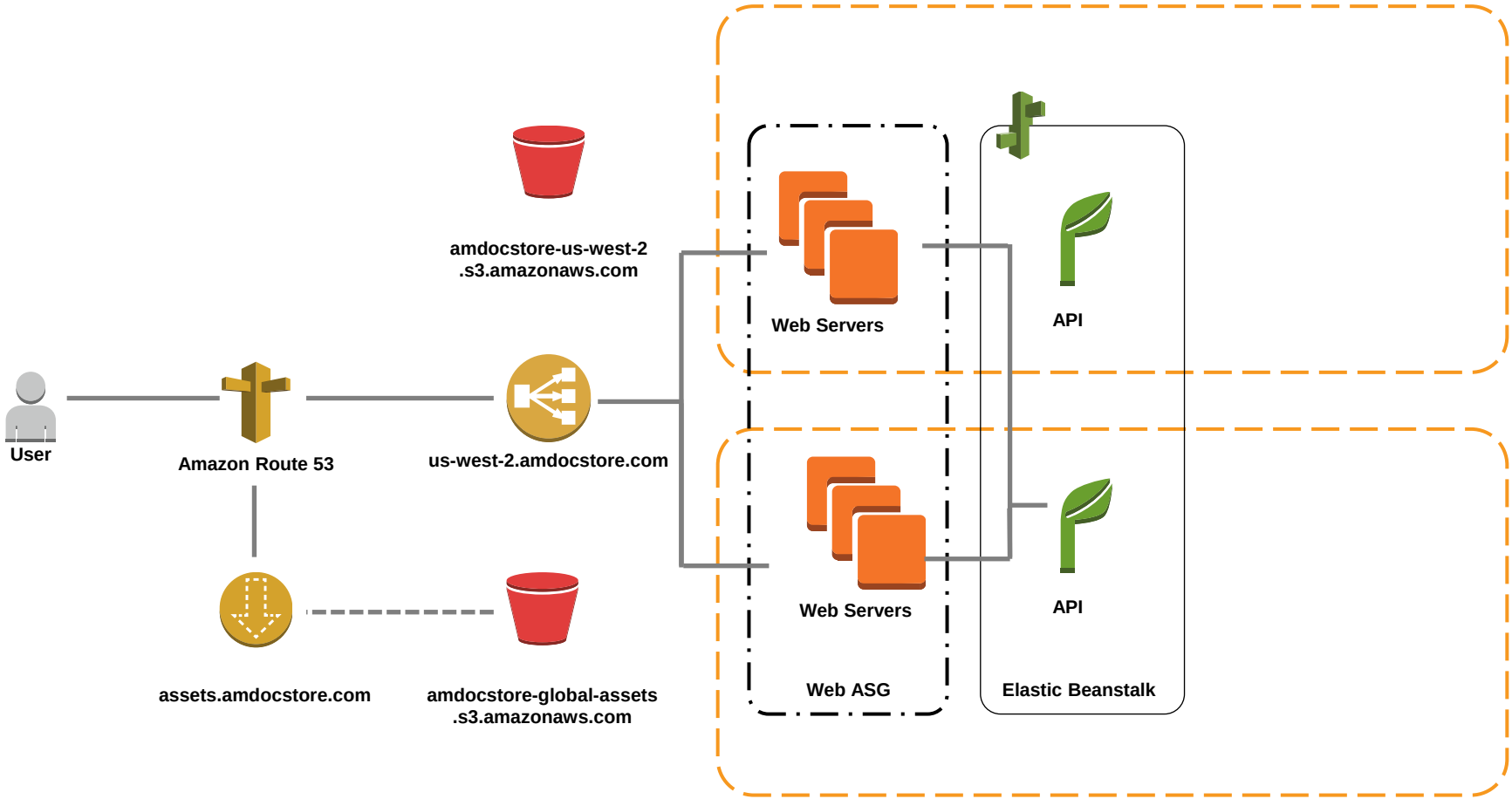


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Let's see what happens when a user wants to **upload a PDF document**

DocStore

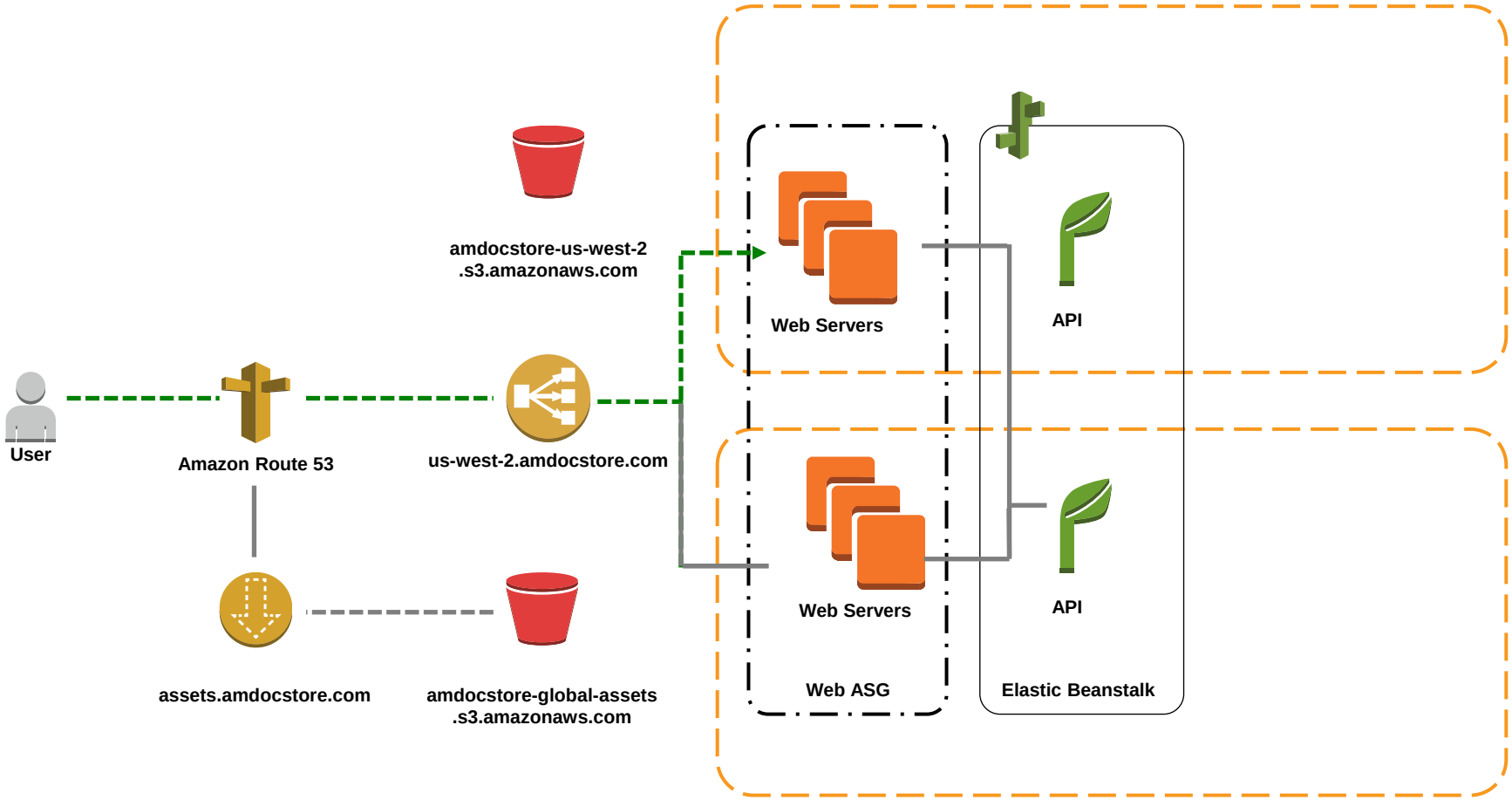


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Upload Document: An authenticated user requests the upload document page from the web servers

DocStore

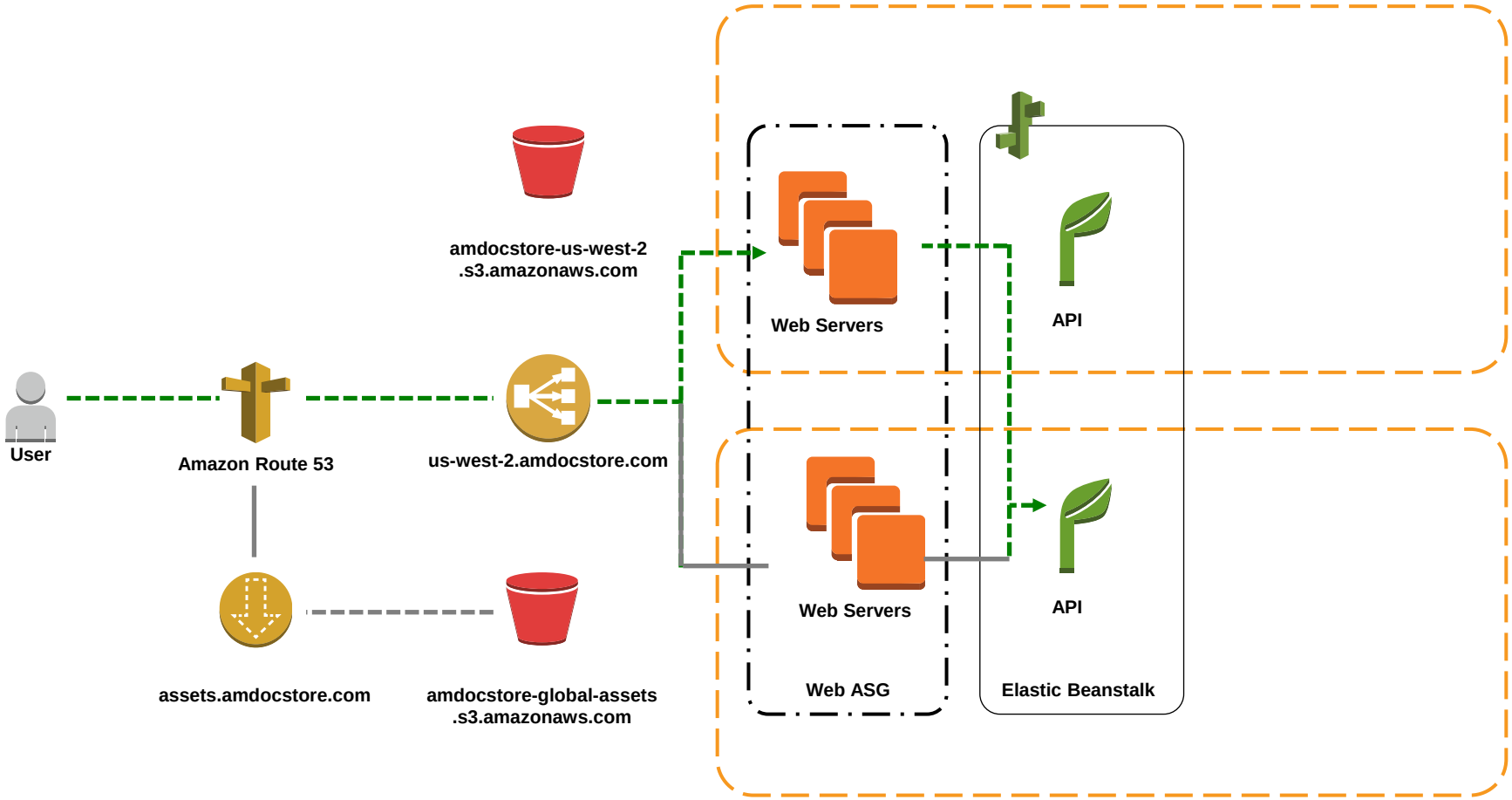


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Upload Document: The web servers request a signed HTML form from the API and return it to the user's browser

DocStore

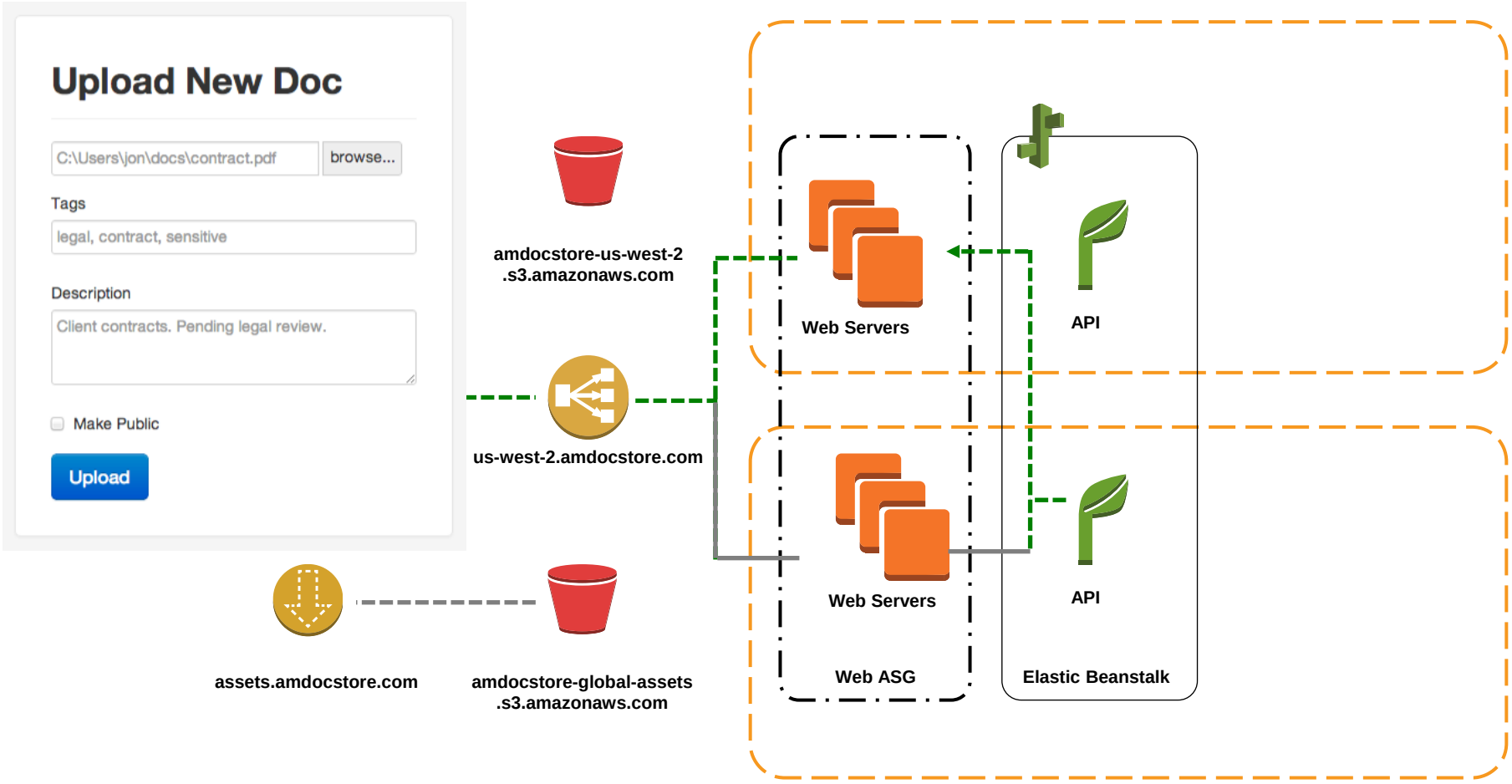


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Upload Document: The signed form is delivered to the user's browser

DocStore



© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

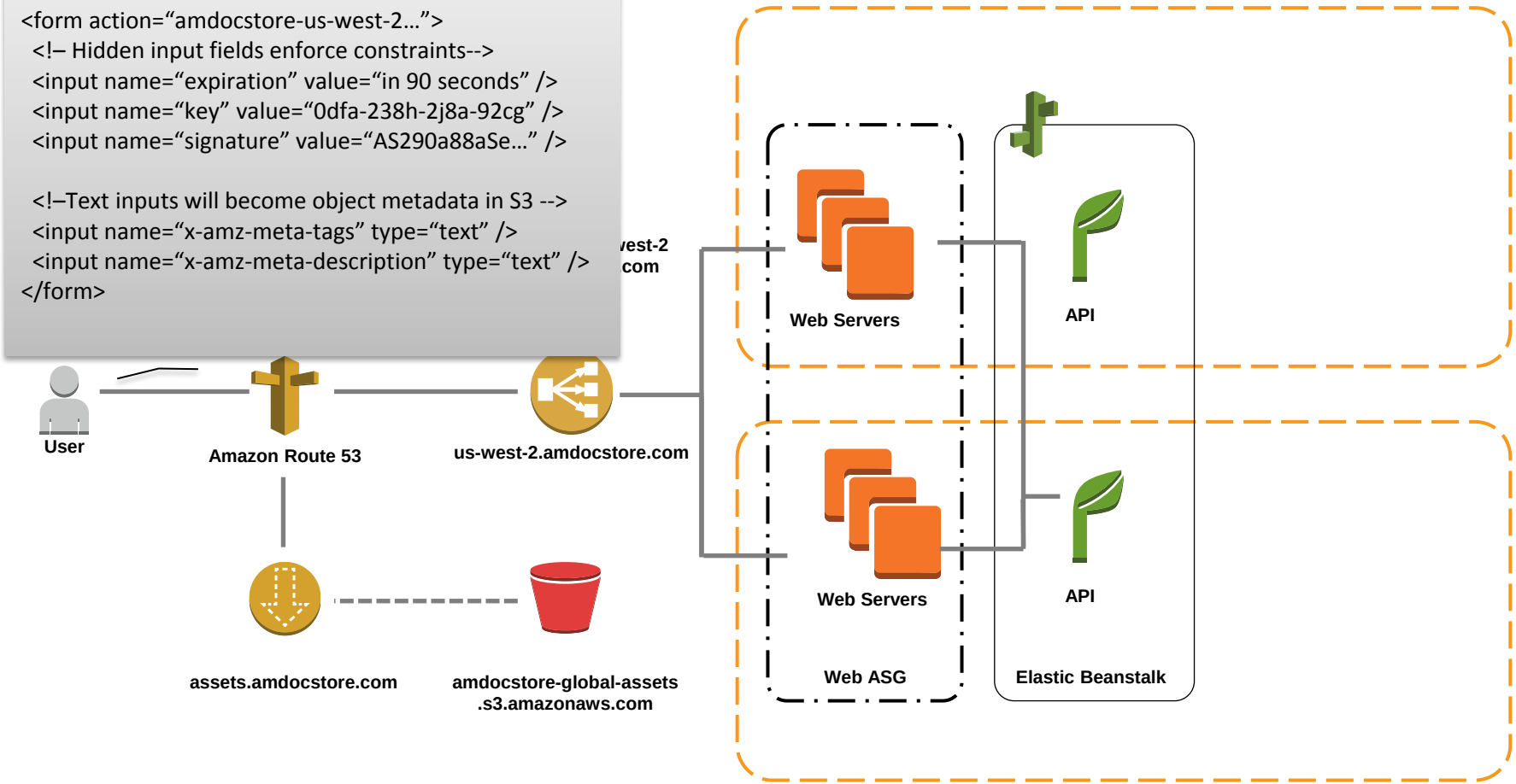
amazon web services™

Upload Document: The signed form allows the user to upload his document **directly to S3**

DocStore

```
<form action="amdocstore-us-west-2…">
  <!– Hidden input fields enforce constraints-->
  <input name="expiration" value="in 90 seconds" />
  <input name="key" value="0dfa-238h-2j8a-92cg" />
  <input name="signature" value="AS290a88aSe…" />

  <!–Text inputs will become object metadata in S3 -->
  <input name="x-amz-meta-tags" type="text" />
  <input name="x-amz-meta-description" type="text" />
</form>
```



© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

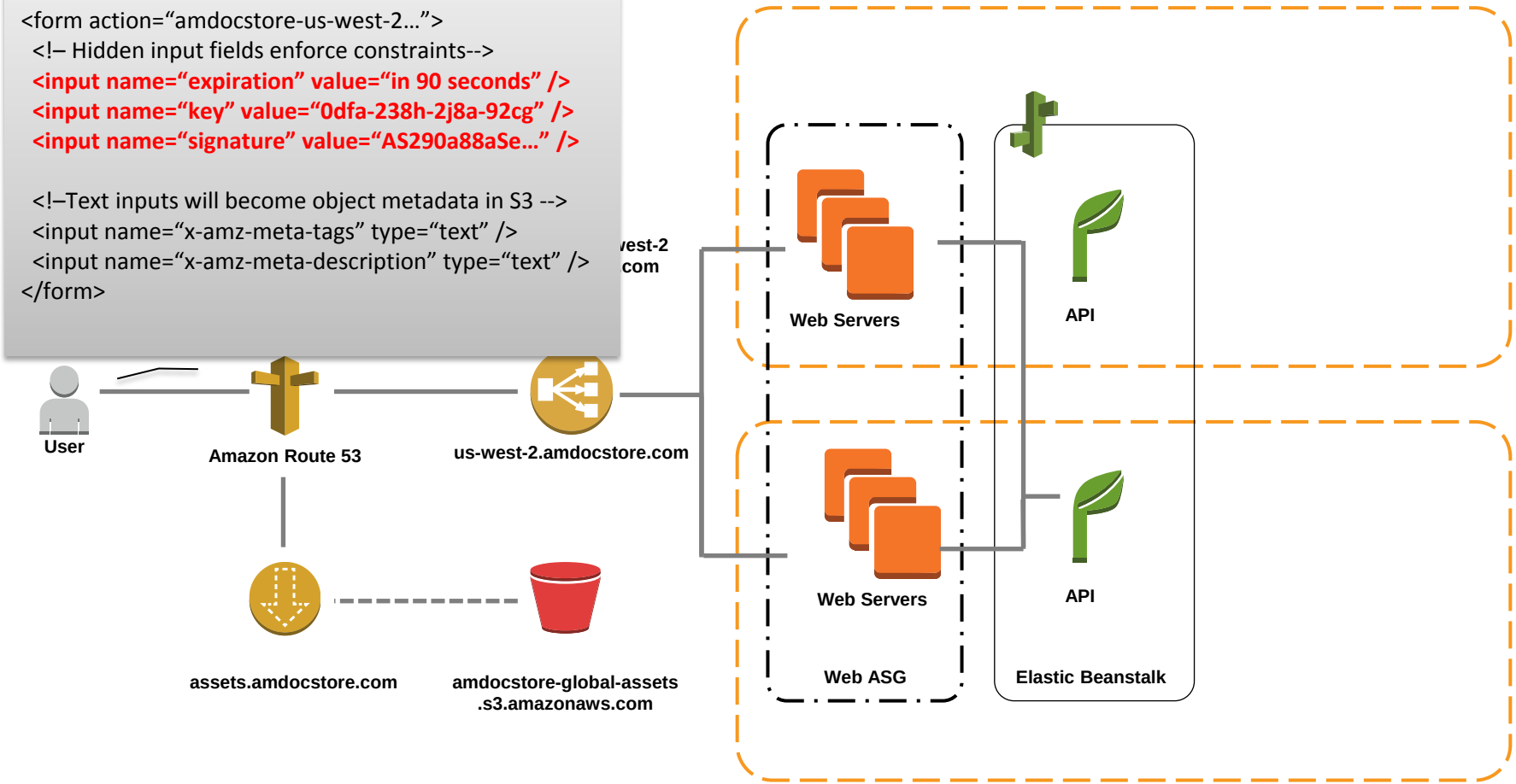
amazon web services™

Upload Document: The form includes some **hidden inputs that constrain** what the user can upload.

DocStore

```
<form action="amdocstore-us-west-2…">
  <!– Hidden input fields enforce constraints-->
  <input name="expiration" value="in 90 seconds" />
  <input name="key" value="0dfa-238h-2j8a-92cg" />
  <input name="signature" value="AS290a88aSe…" />

  <!–Text inputs will become object metadata in S3 -->
  <input name="x-amz-meta-tags" type="text" />
  <input name="x-amz-meta-description" type="text" />
</form>
```



© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

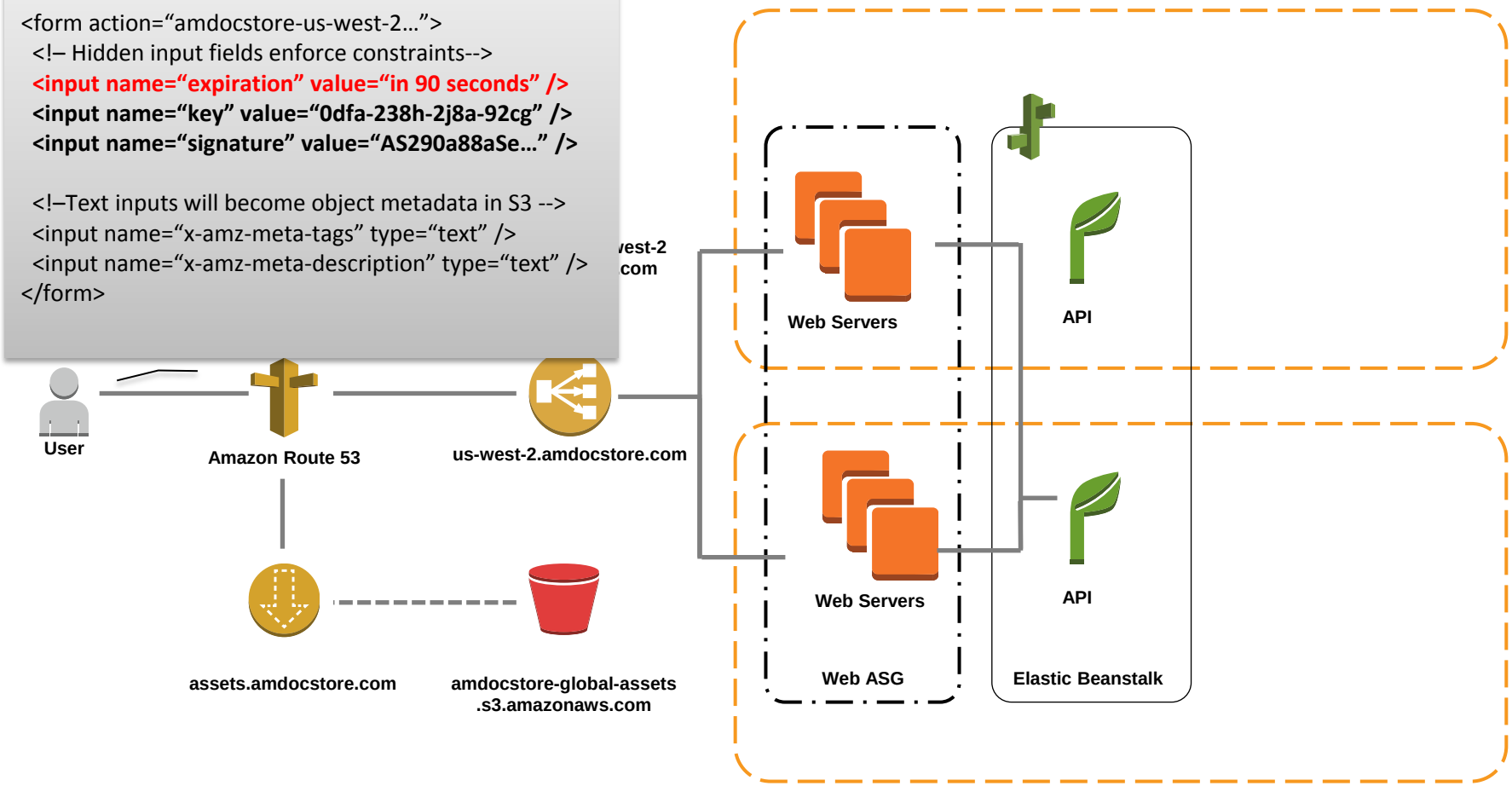
amazon web services™

Upload Document: The form includes some **hidden inputs that constrain** what the user can upload. **The expiration** indicates for how long the form is valid

DocStore

```
<form action="amdocstore-us-west-2…">
  <!– Hidden input fields enforce constraints-->
  <input name="expiration" value="in 90 seconds" />
  <input name="key" value="0dfa-238h-2j8a-92cg" />
  <input name="signature" value="AS290a88aSe…" />

  <!–Text inputs will become object metadata in S3 -->
  <input name="x-amz-meta-tags" type="text" />
  <input name="x-amz-meta-description" type="text" />
</form>
```



© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

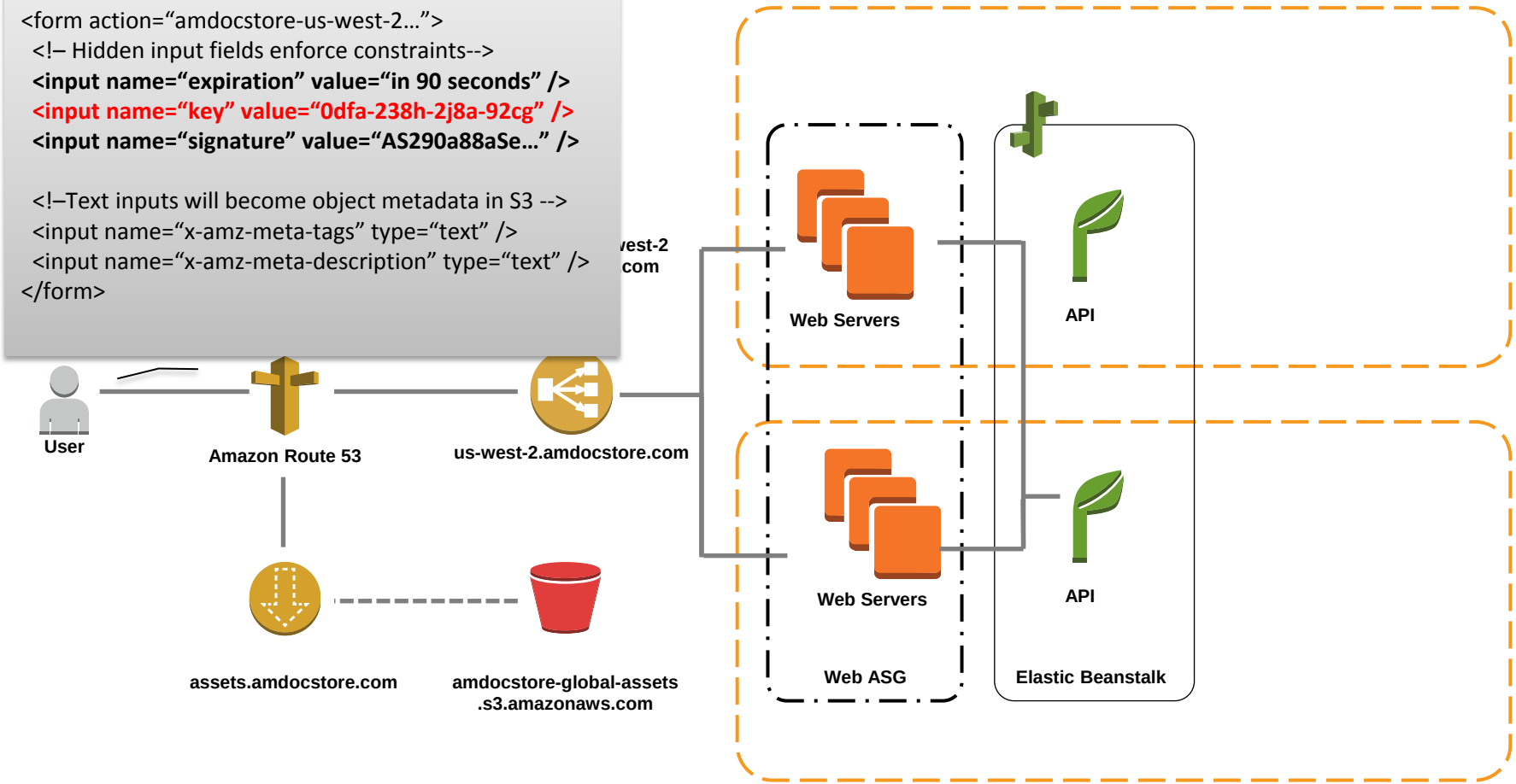
amazon web services™

Upload Document: The form includes some **hidden inputs that constrain** what the user can upload. **The key** allows us to specify what the object will be called in S3 (thus avoiding naming conflicts)

DocStore

```
<form action="amdocstore-us-west-2…">
  <!– Hidden input fields enforce constraints-->
  <input name="expiration" value="in 90 seconds" />
  <input name="key" value="0dfa-238h-2j8a-92cg" />
  <input name="signature" value="AS290a88aSe…" />

  <!–Text inputs will become object metadata in S3 -->
  <input name="x-amz-meta-tags" type="text" />
  <input name="x-amz-meta-description" type="text" />
</form>
```



© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

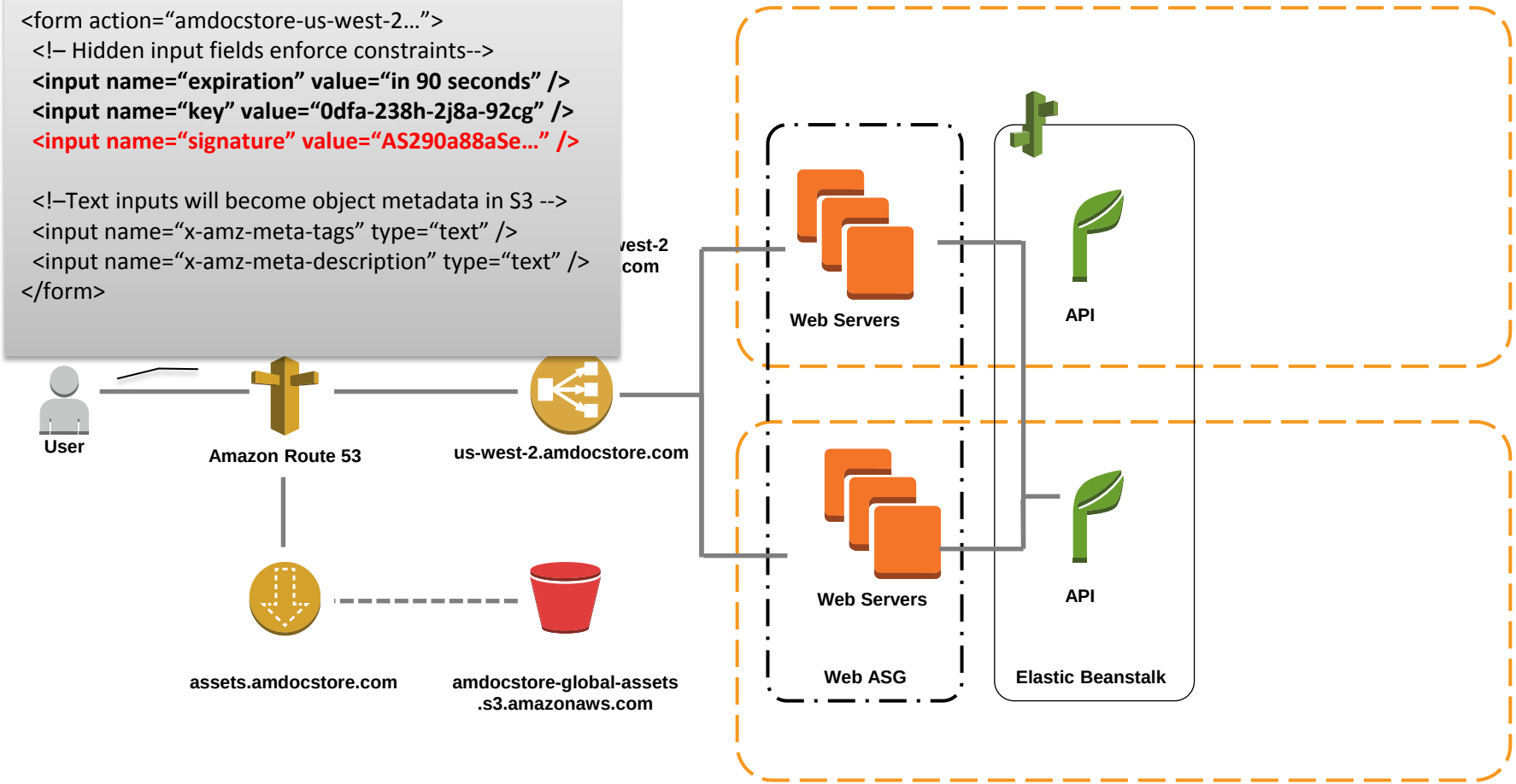
amazon web services™

Upload Document: The form includes some **hidden inputs that constrain** what the user can upload. **The signature** ensures the upload will be rejected if the form is tampered with (e.g., user alters the expiration)

DocStore

```
<form action="amdocstore-us-west-2…">
  <!– Hidden input fields enforce constraints-->
  <input name="expiration" value="in 90 seconds" />
  <input name="key" value="0dfa-238h-2j8a-92cg" />
  <input name="signature" value="AS290a88aSe…" />

  <!–Text inputs will become object metadata in S3 -->
  <input name="x-amz-meta-tags" type="text" />
  <input name="x-amz-meta-description" type="text" />
</form>
```



© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

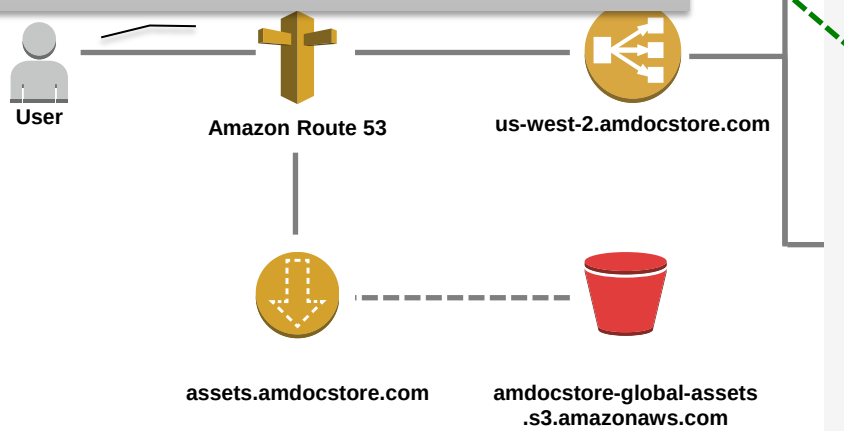
amazon web services™

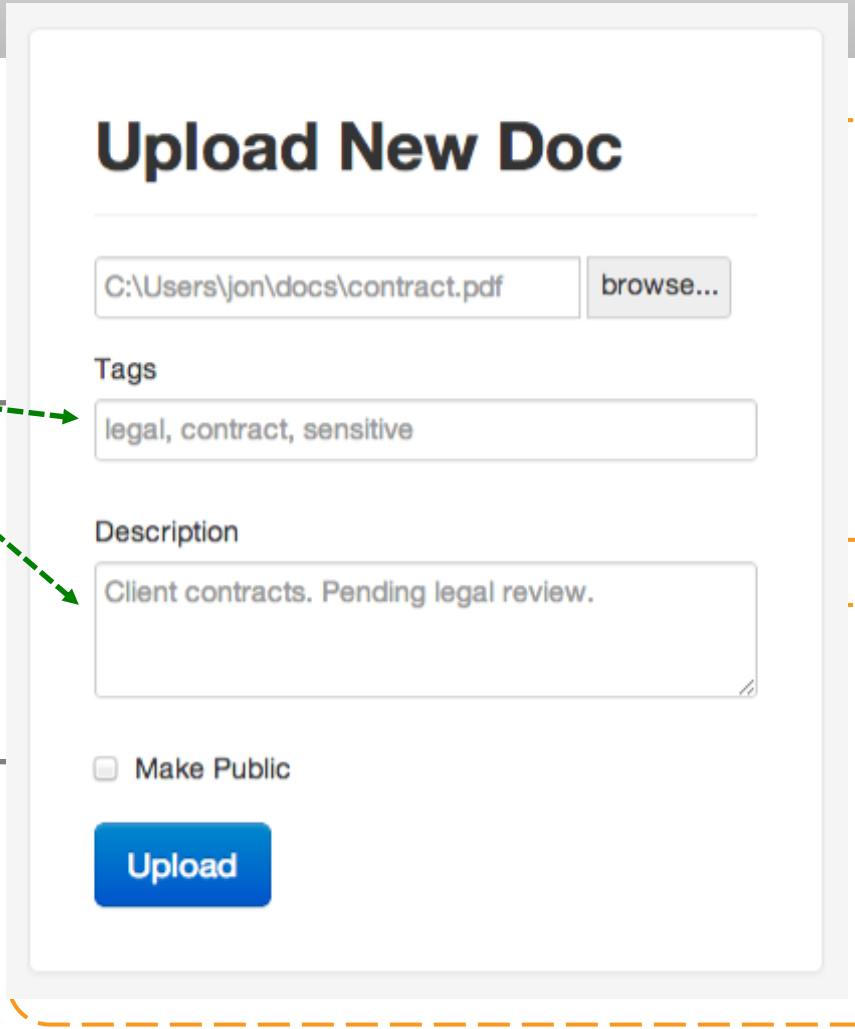
Upload Document: User-editable **text input fields** prefixed with **x-amz-meta-** will be associated with the uploaded object as **S3 object metadata**

DocStore

```
<form action="amdocstore-us-west-2…">
  <!– Hidden input fields enforce constraints-->
  <input name="expiration" value="in 90 seconds" />
  <input name="key" value="0dfa-238h-2j8a-92cg" />
  <input name="signature" value="AS290a88aSe…" />

  <!–Text inputs will become object metadata in S3 -->
  <input name="x-amz-meta-tags" type="text" />
  <input name="x-amz-meta-description" type="text"
/>
</form>
```





© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services

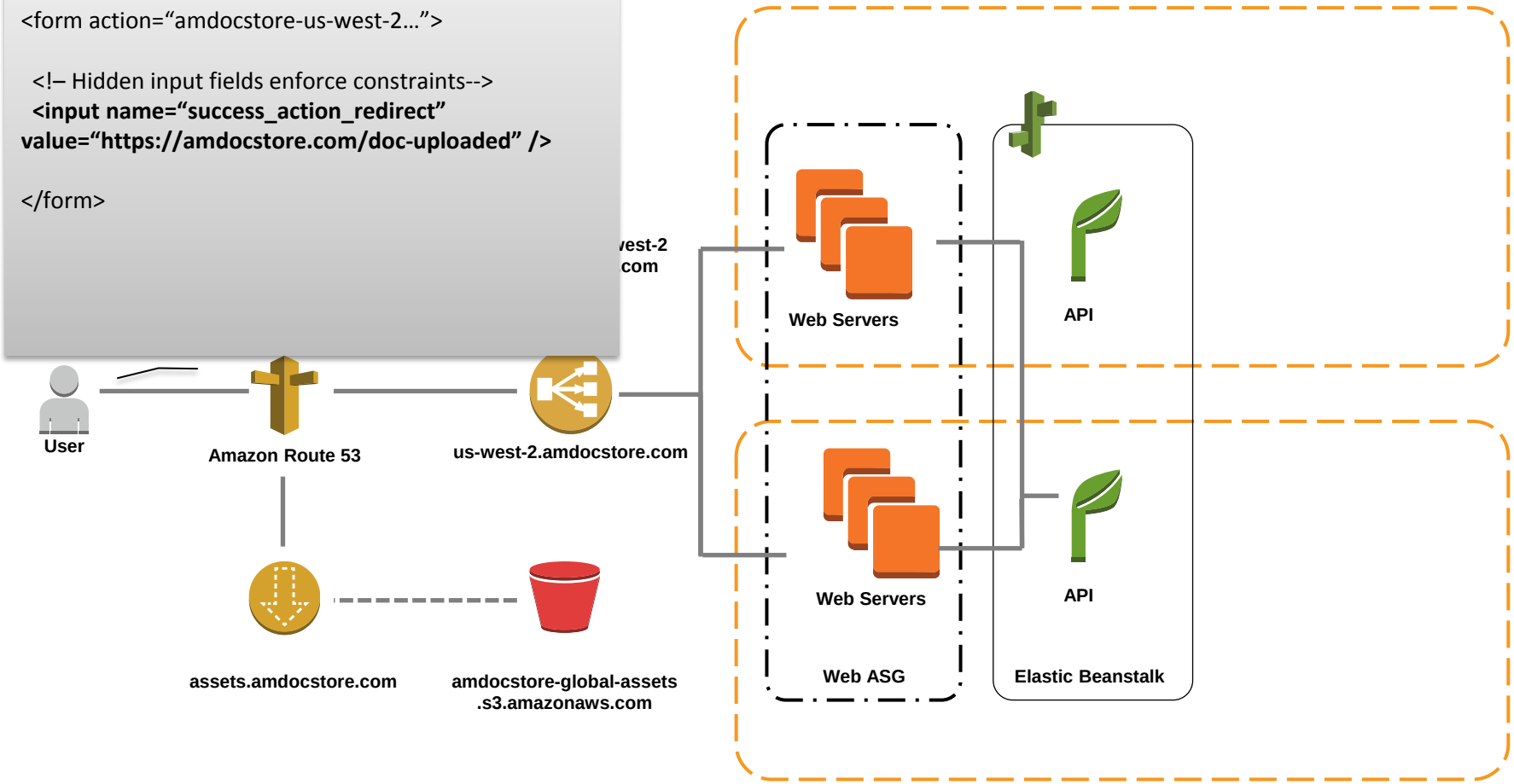
Upload Document: Finally, the **success_action_redirect** specified when the API generated the form tells S3 where to redirect the user's browser when the upload completes

DocStore

```
<form action="amdocstore-us-west-2…">

  <!– Hidden input fields enforce constraints-->
  <input name="success_action_redirect"
value="https://amdocstore.com/doc-uploaded" />

</form>
```



© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

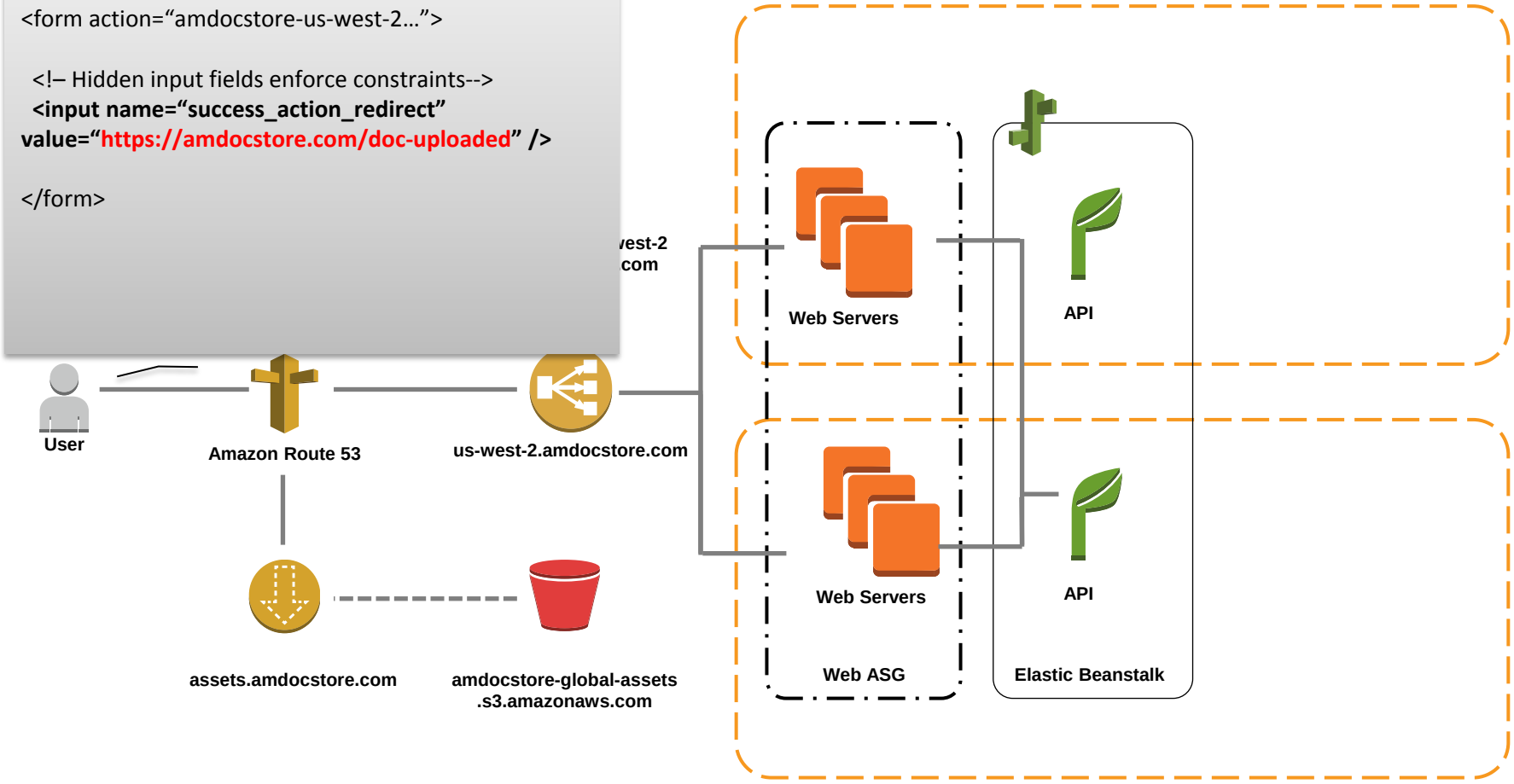
Upload Document: We send the user to a special page on the web servers that will capture this event

DocStore

```
<form action="amdocstore-us-west-2…">

  <!– Hidden input fields enforce constraints-->
  <input name="success_action_redirect"
value="https://amdocstore.com/doc-uploaded" />

</form>
```

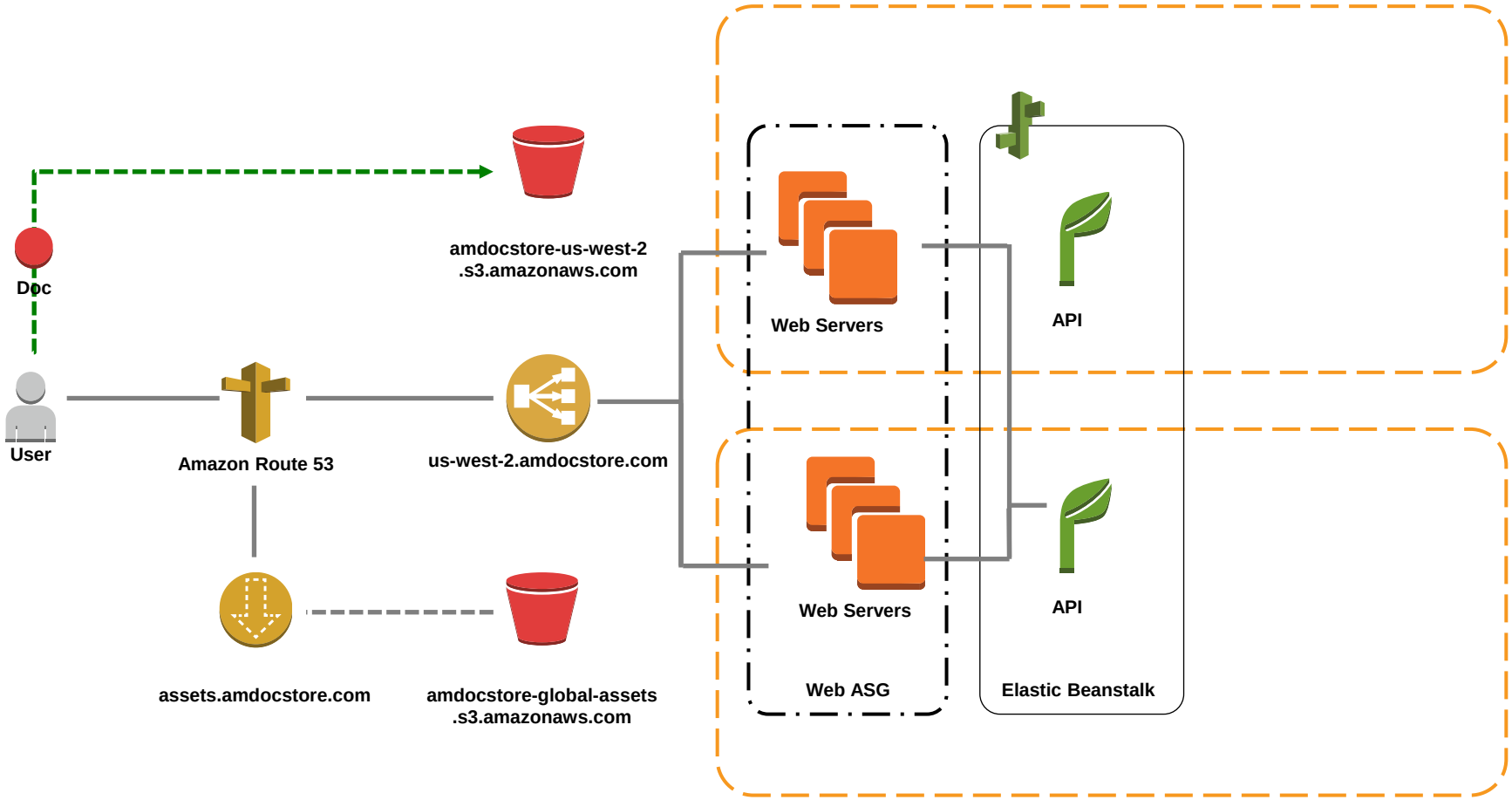


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Upload Document: While uploading the document to S3,

DocStore

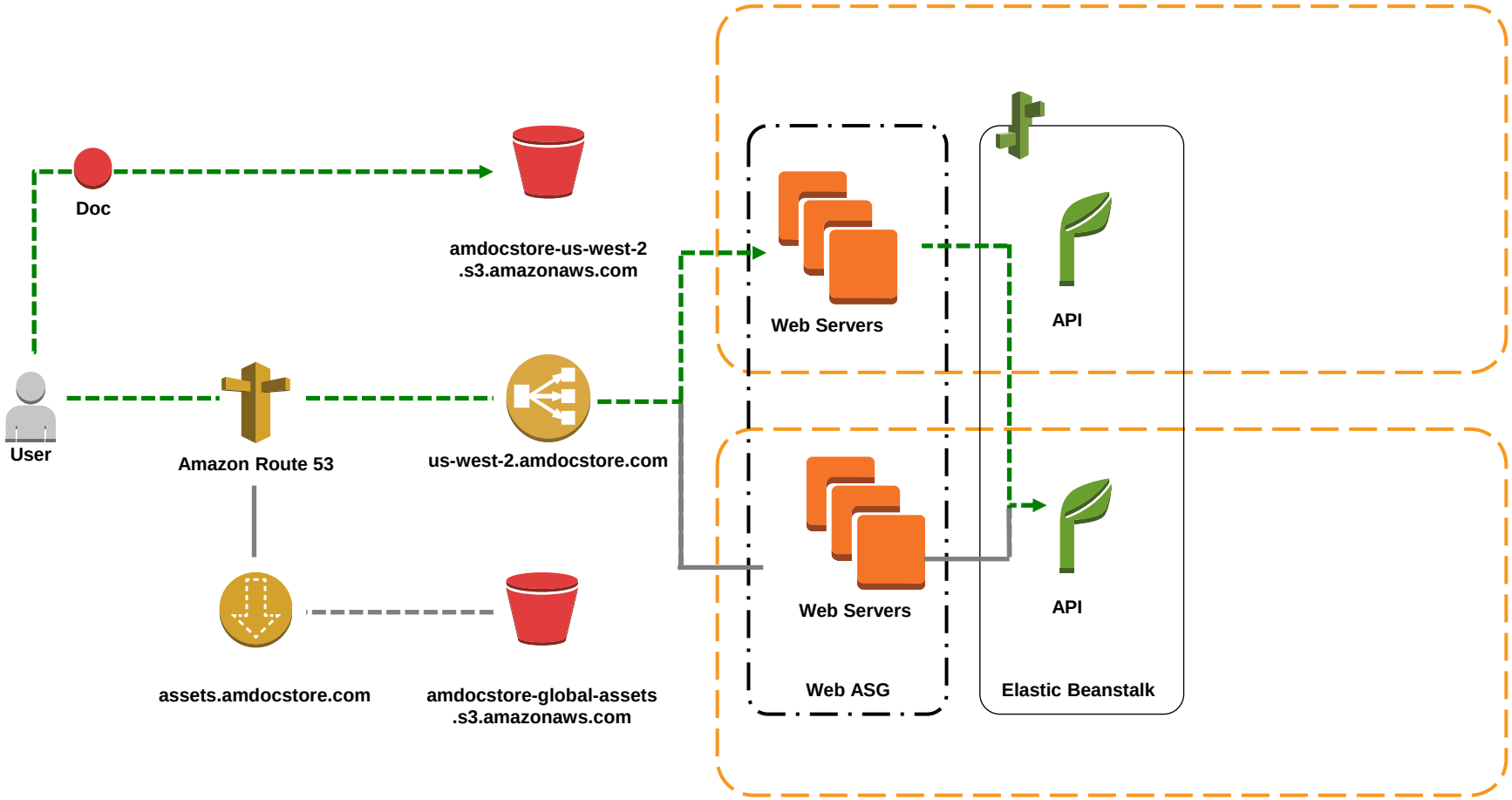


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Upload Document: While uploading the document to S3, we **asynchronously notify** the web servers that the upload has begun

DocStore

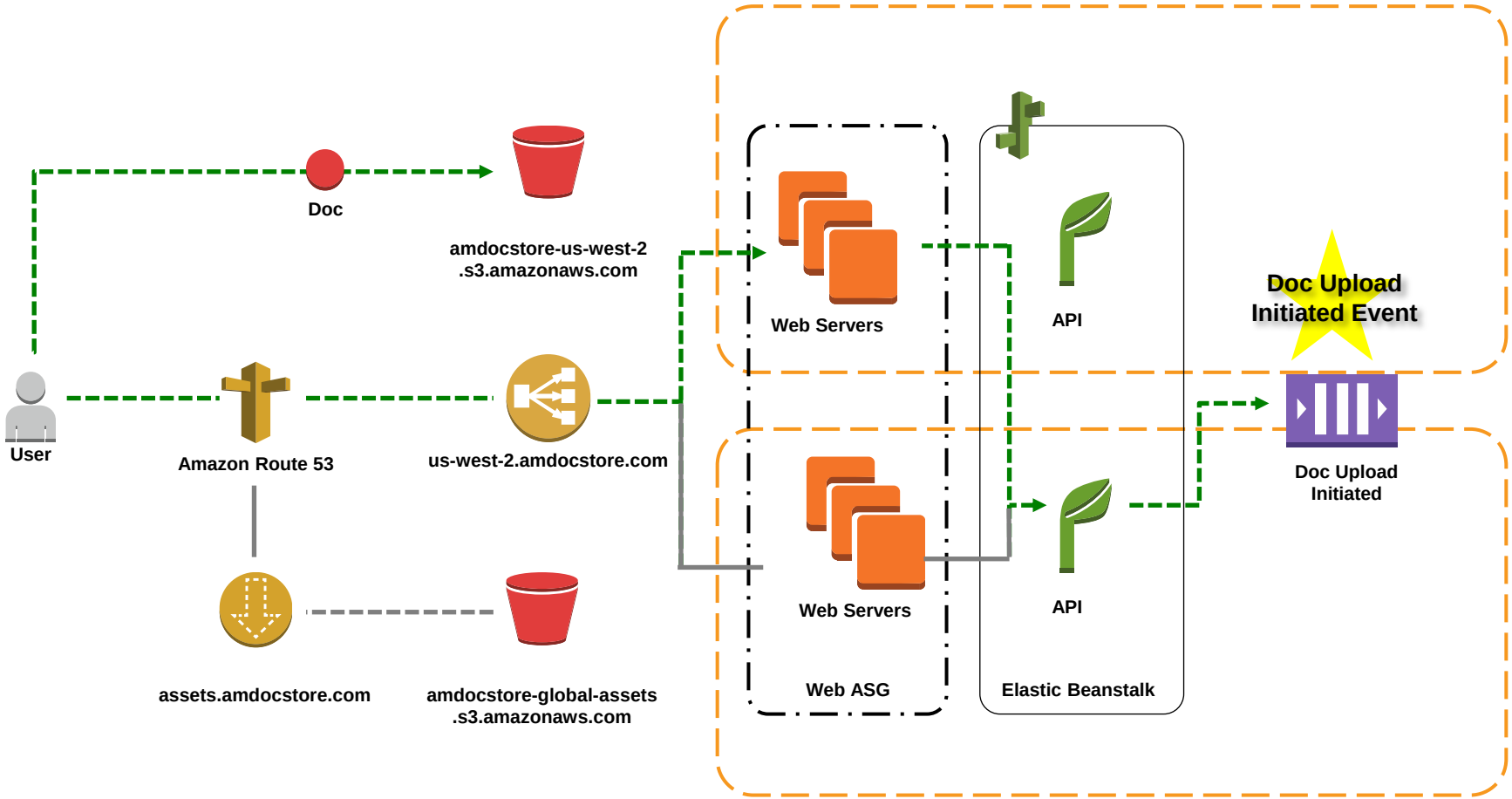


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Upload Document: The API places a message in a queue indicating as much. Other components of DocStore may be interested in this event later

DocStore

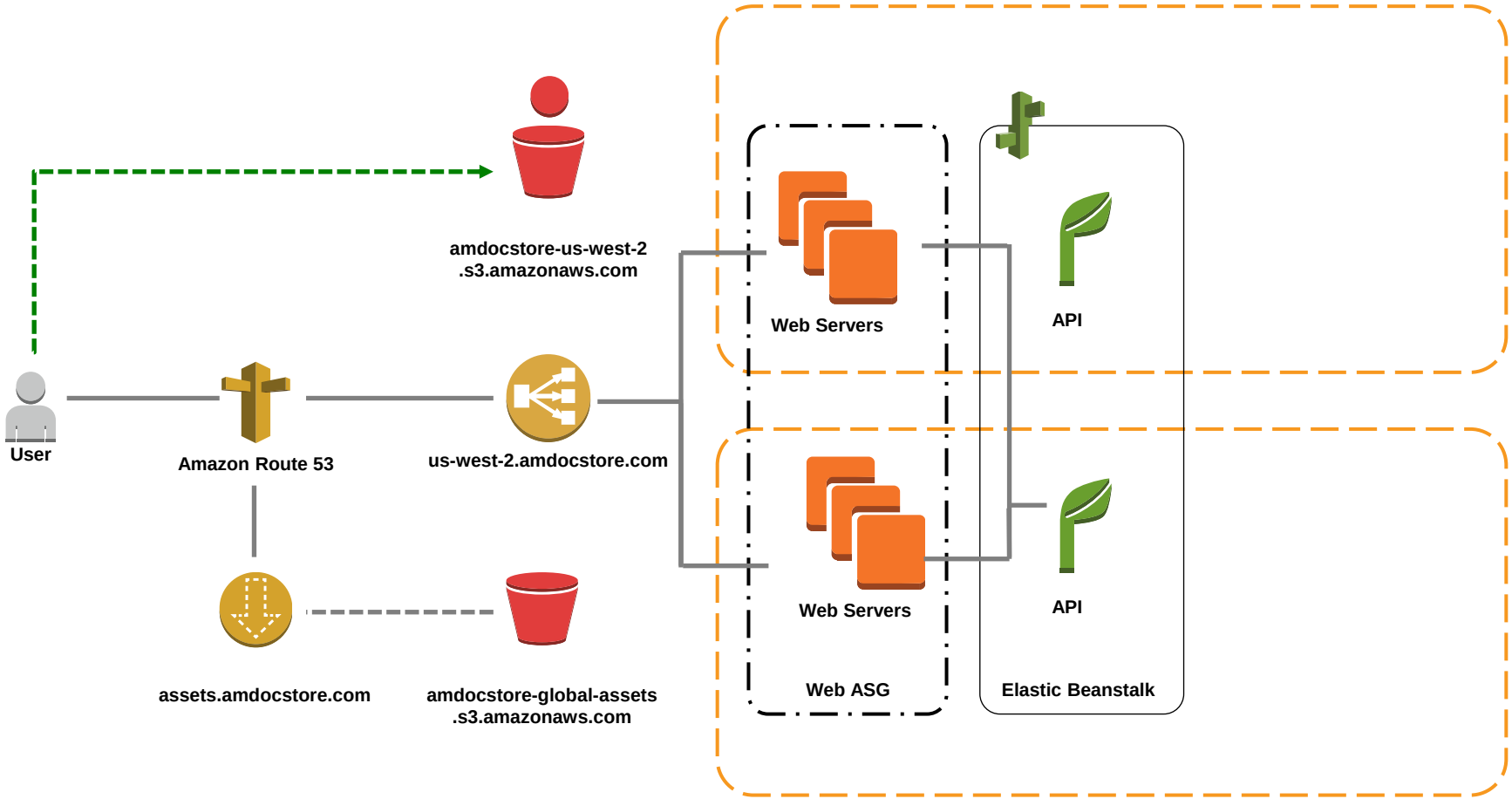


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Upload Document: When the upload completes

DocStore

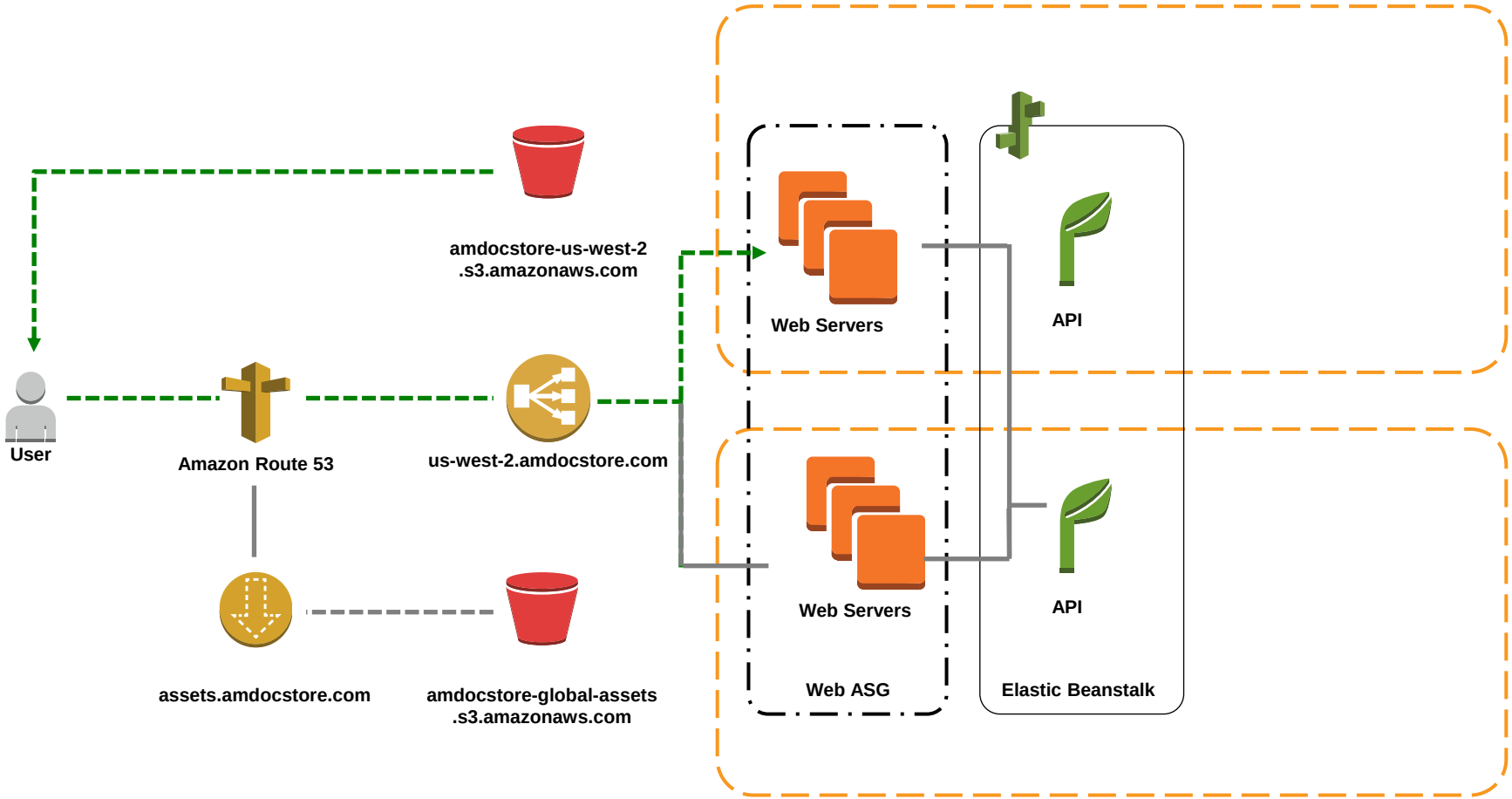


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Upload Document: When the upload completes **S3 redirects the user to the URL we specified earlier**

DocStore

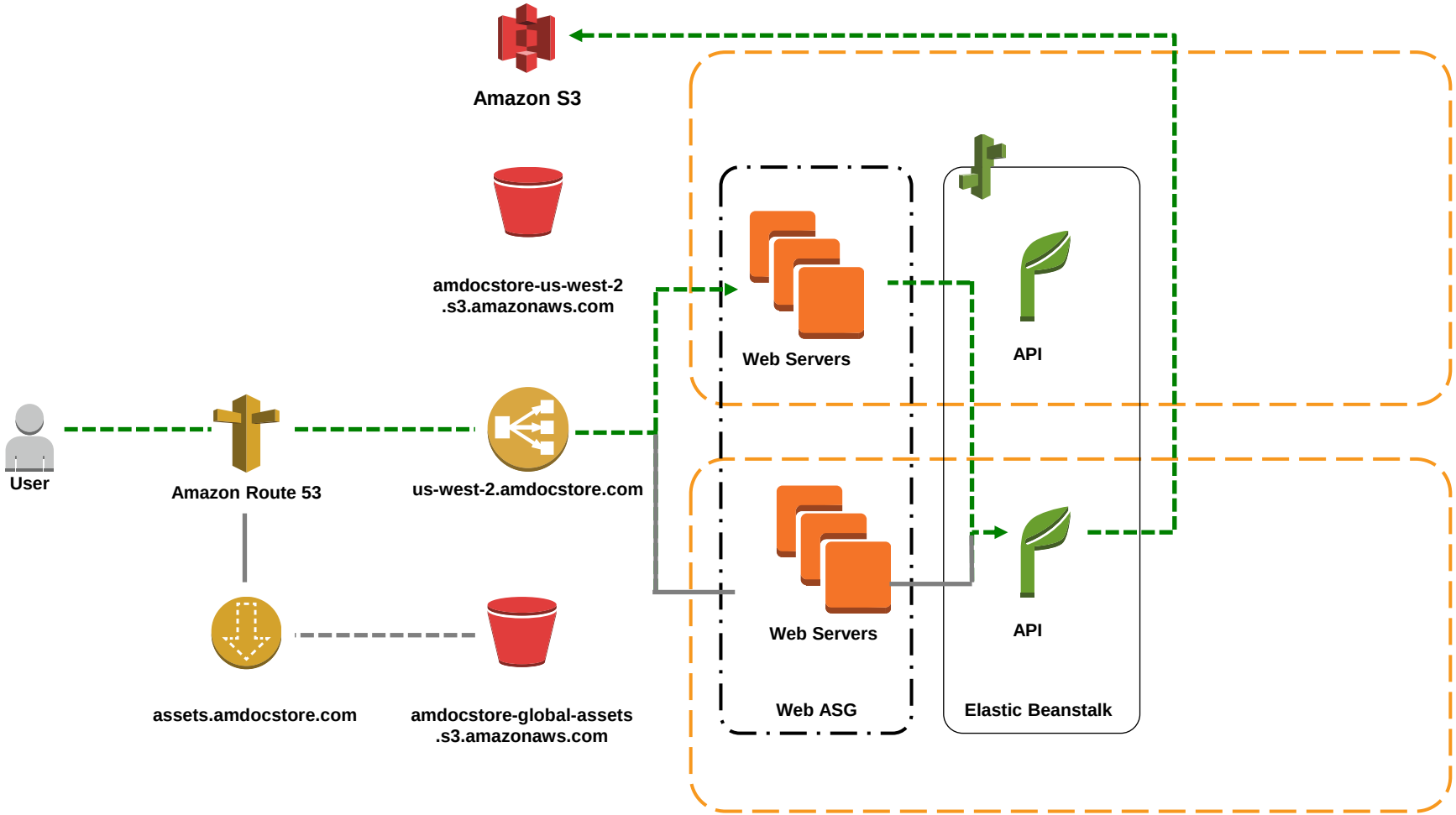


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Upload Document: The API **retrieves metadata about the object** (i.e., size, type, and user-provided custom information like tags and description) from the S3 API

DocStore

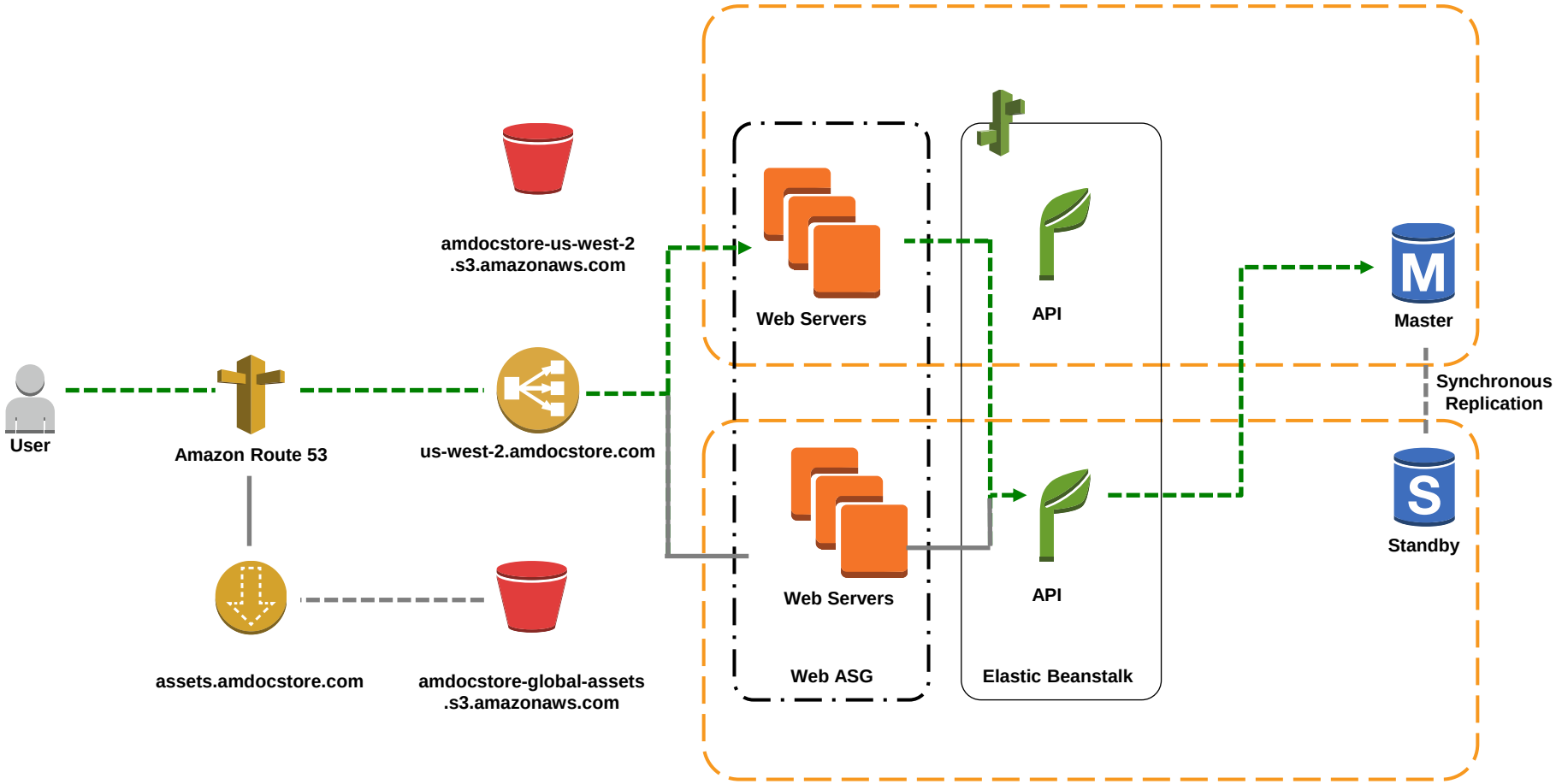


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Upload Document: The object metadata is stored in an **RDS** instance running in **Multi-AZ** mode for **high availability**

DocStore

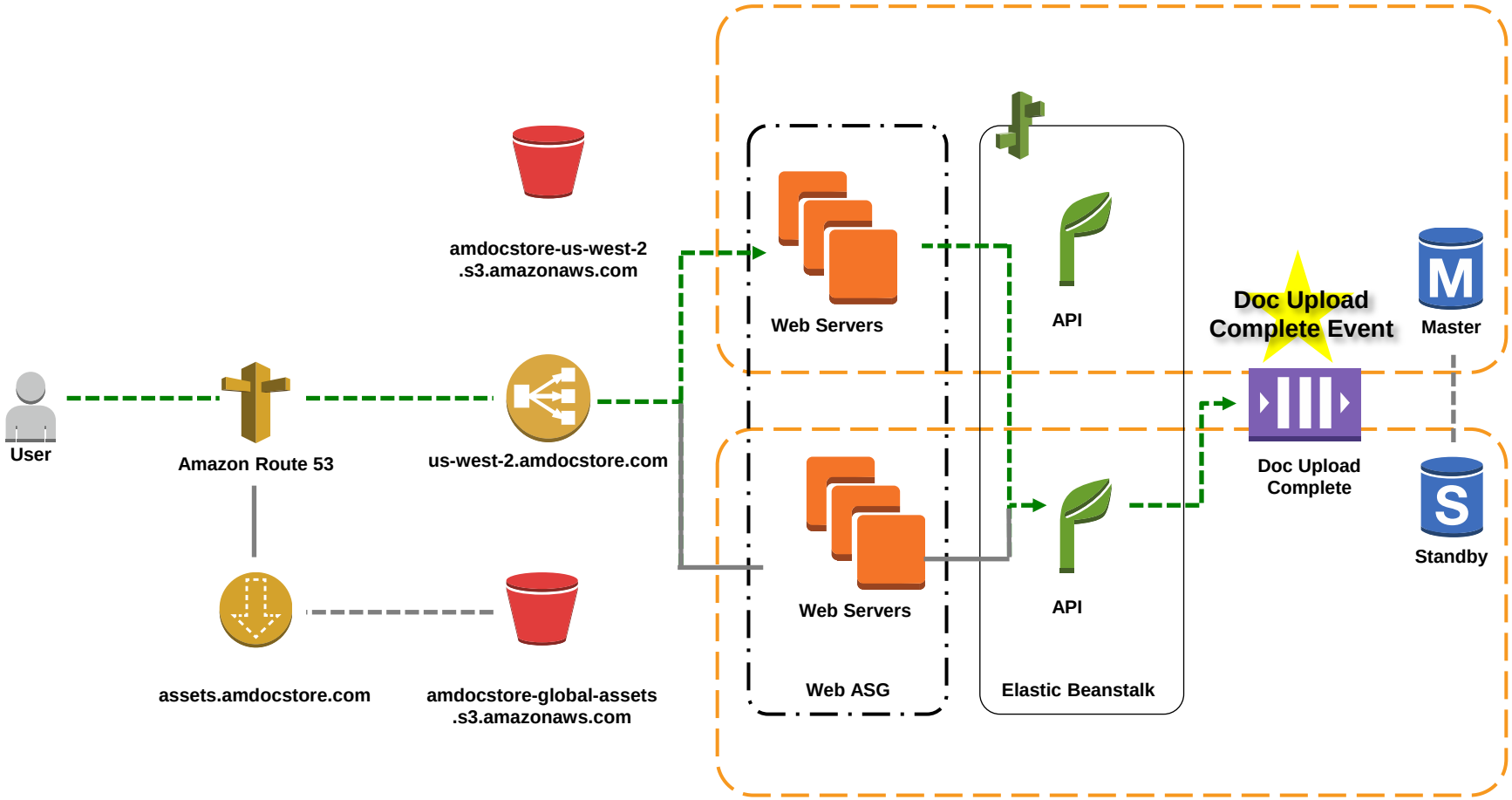


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Upload Document: Finally, the API **places a message in a queue** indicating that the document upload was complete, and includes the ID of the newly uploaded object in S3

DocStore

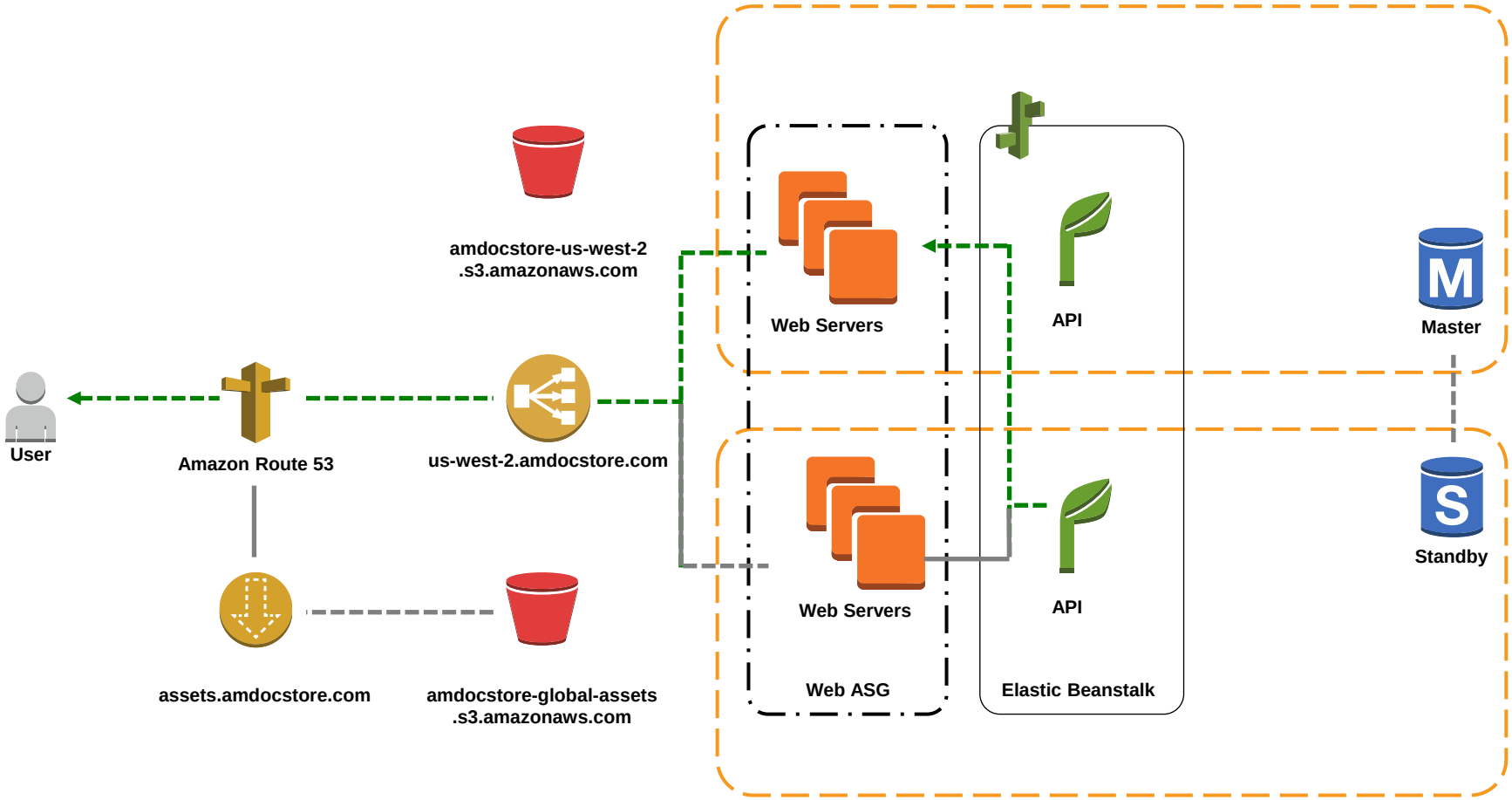


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Upload Document: The user is redirected to a list of his documents

DocStore

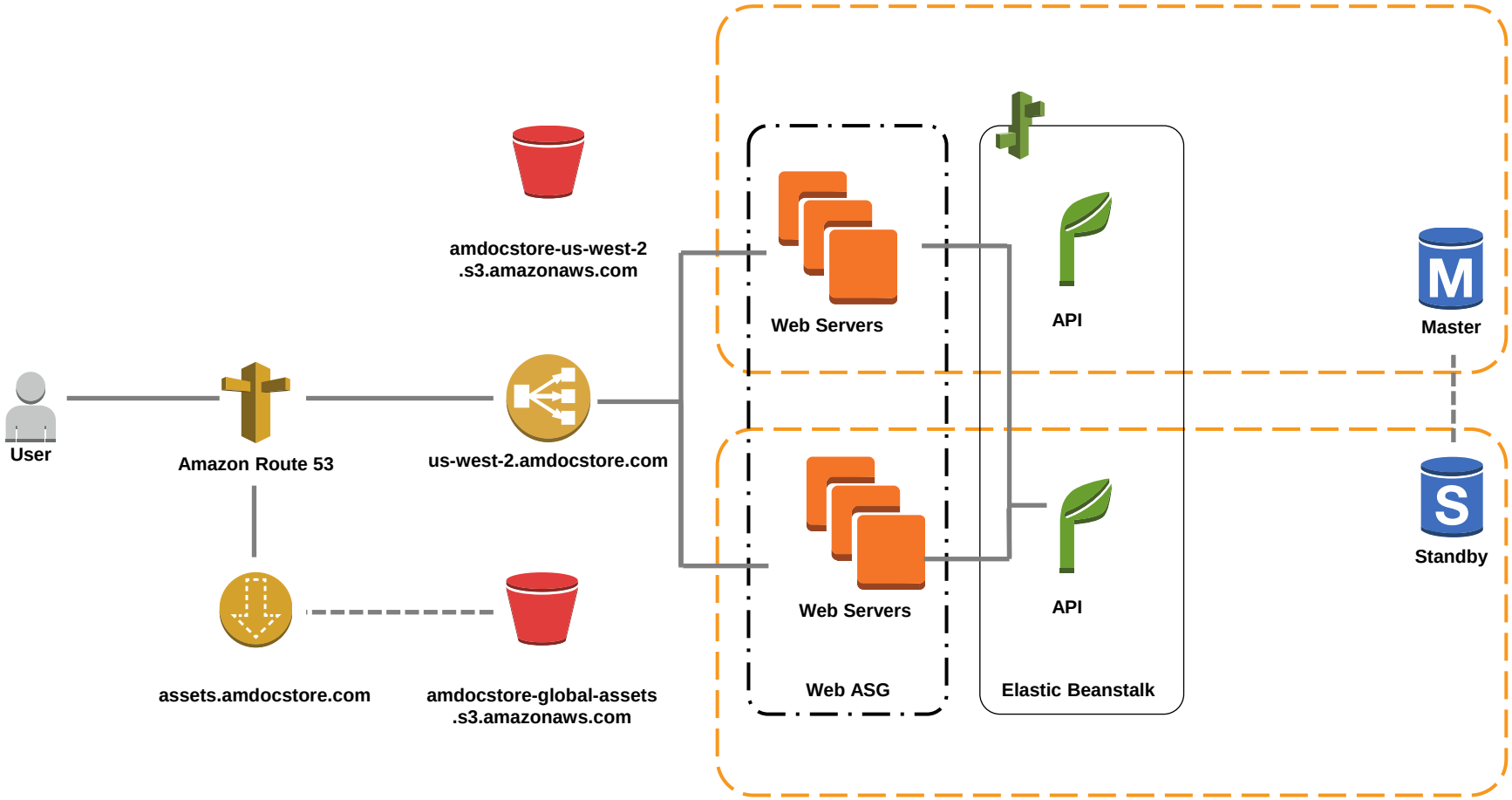


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Let's see what happens when a user wants to **view and download his documents**

DocStore

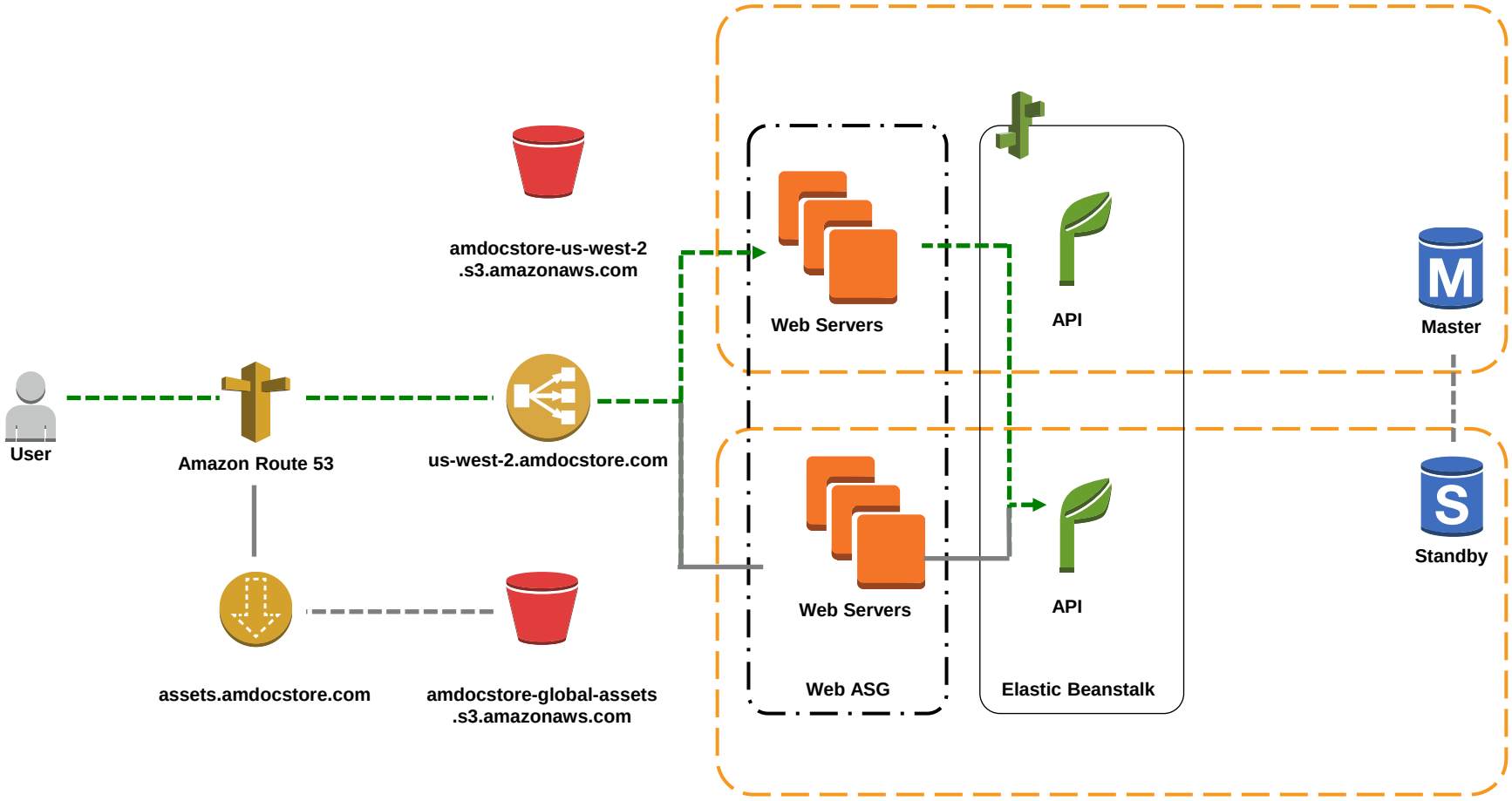


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

View and Download Docs: An authenticated user requests the web site URL to view his docs

DocStore

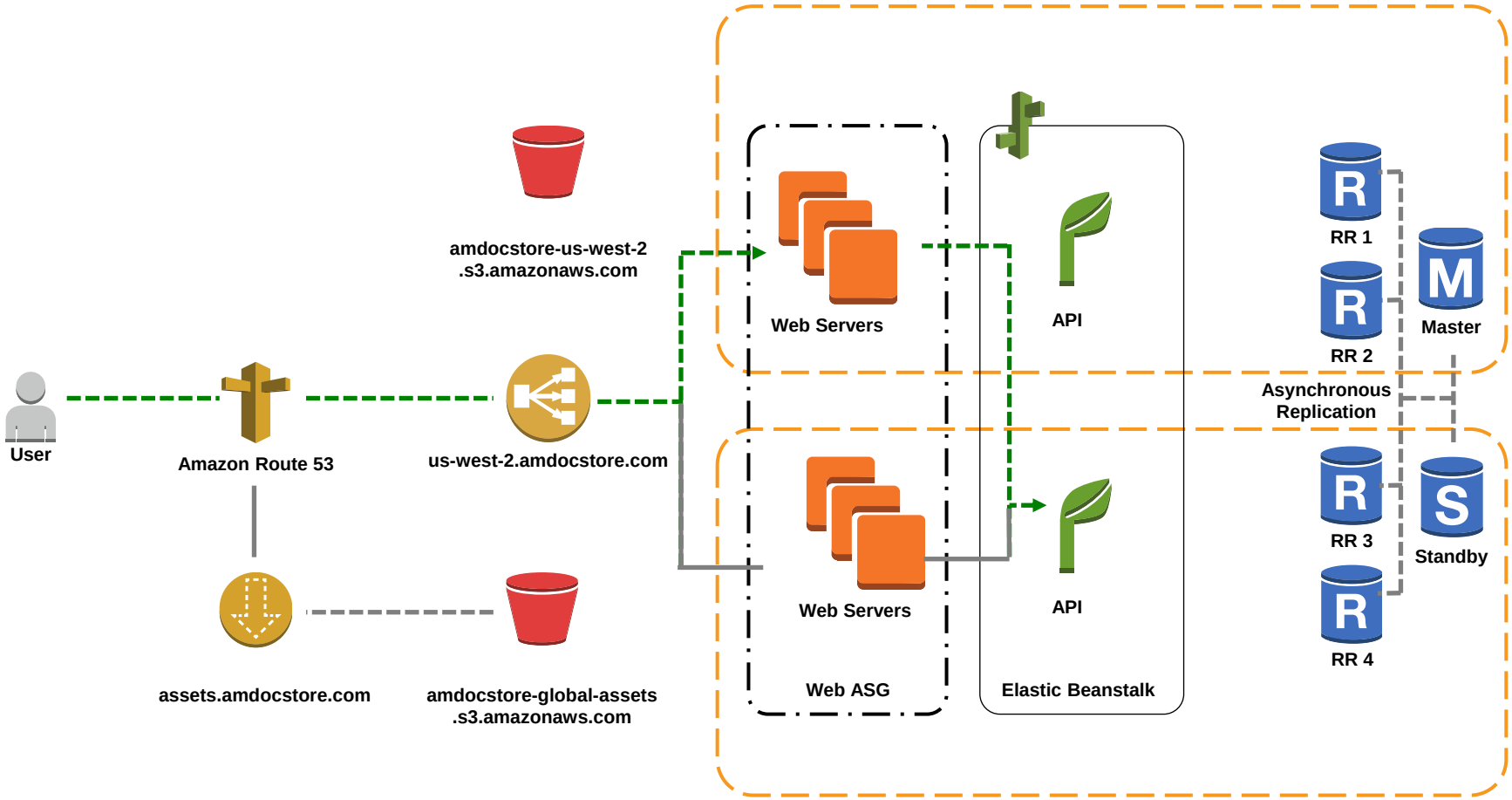


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

View and Download Docs: To **increase read performance** of the relational database, multiple **RDS Read Replicas** are provisioned. Read Replicas are **suitable for read workloads**, like the API selecting a list of documents owned by a user

DocStore

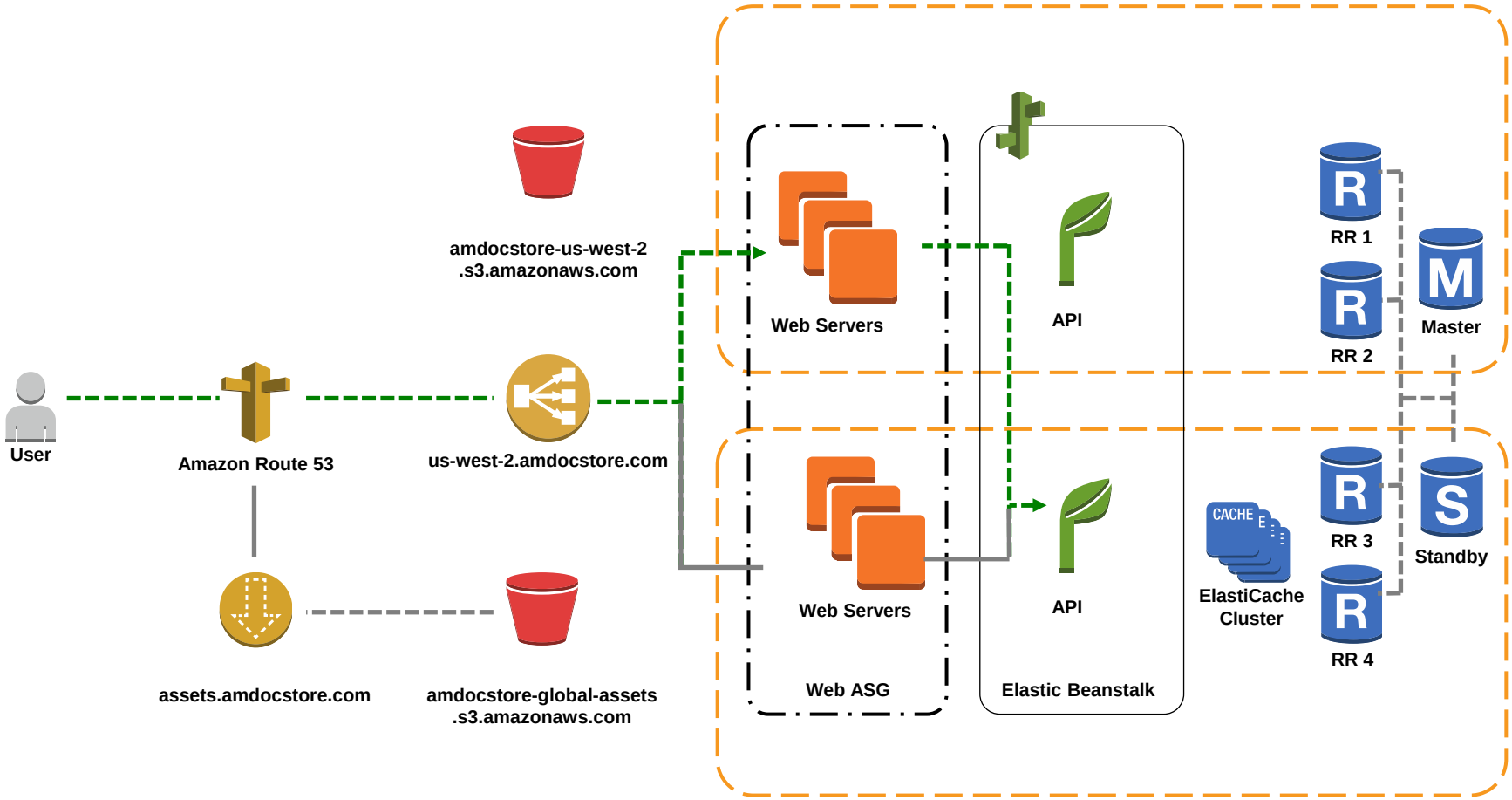


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

View and Download Docs: ElastiCache Cluster in front of the Read Replicas would provide additional performance when listing the user's documents

DocStore

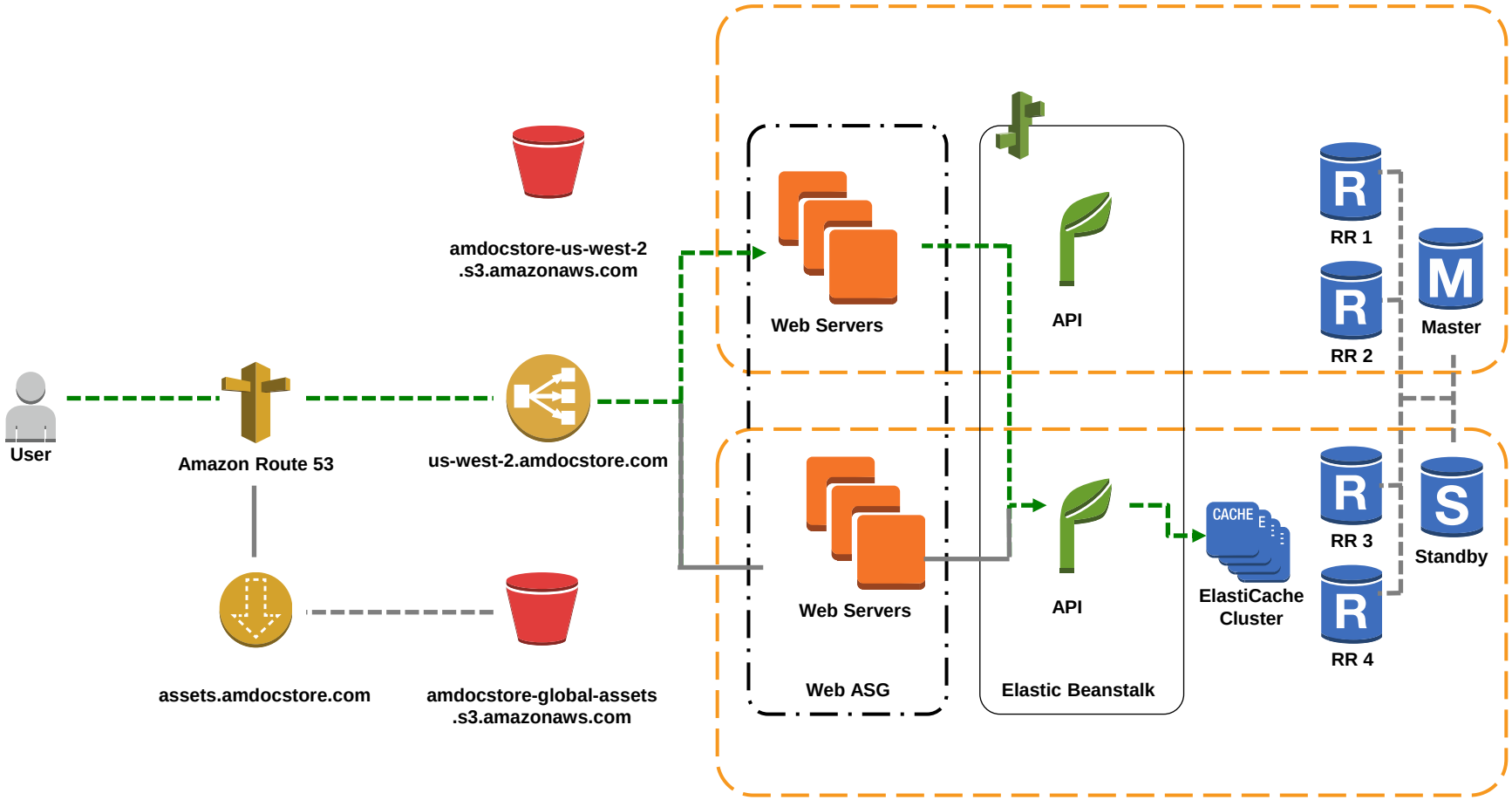


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

View and Download Docs: The API **first checks the cache** for a recent list of the user's documents

DocStore

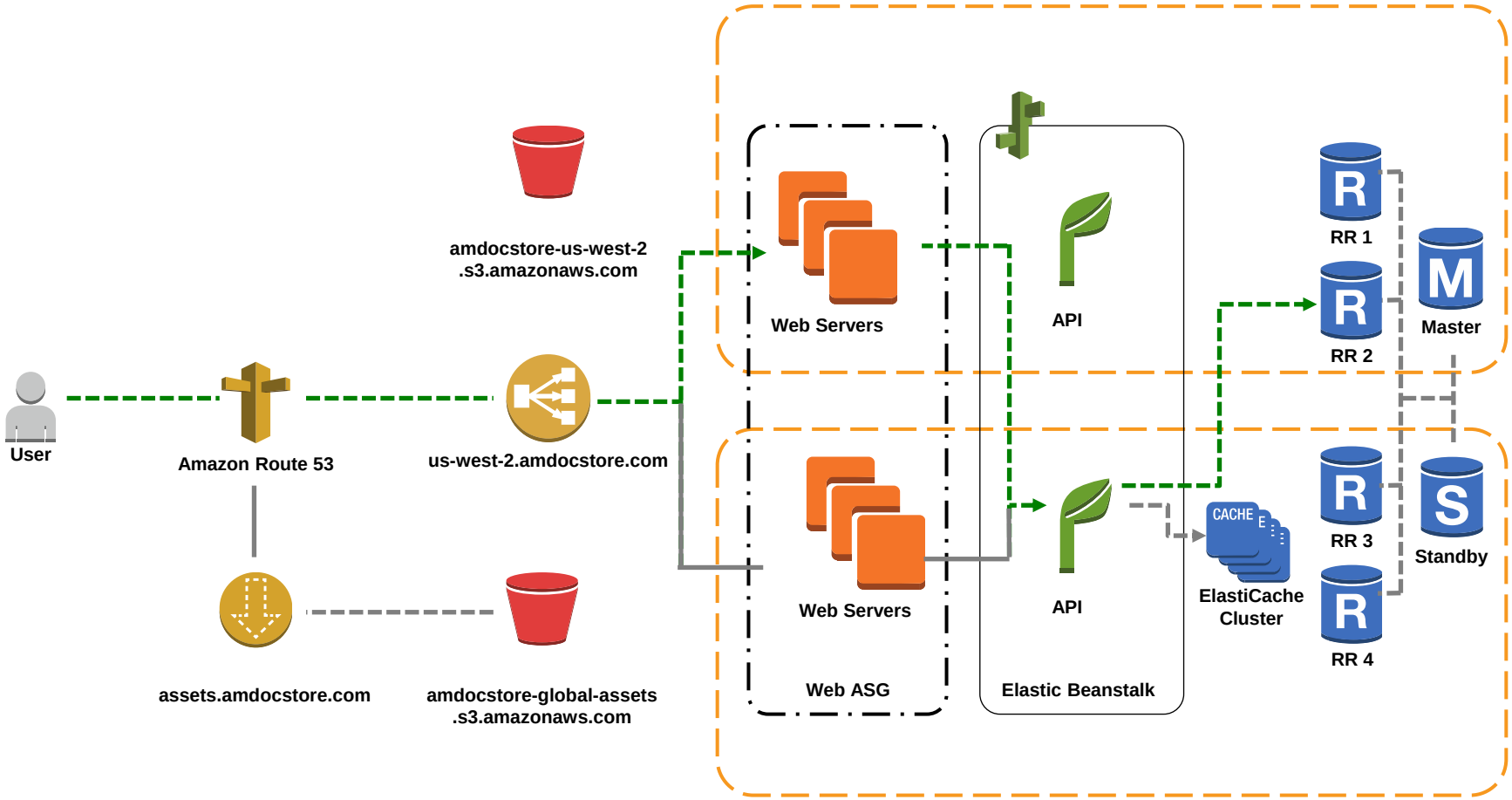


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

View and Download Docs: If no cached data is available, the API chooses one of the Read Replicas and queries for the list of documents

DocStore

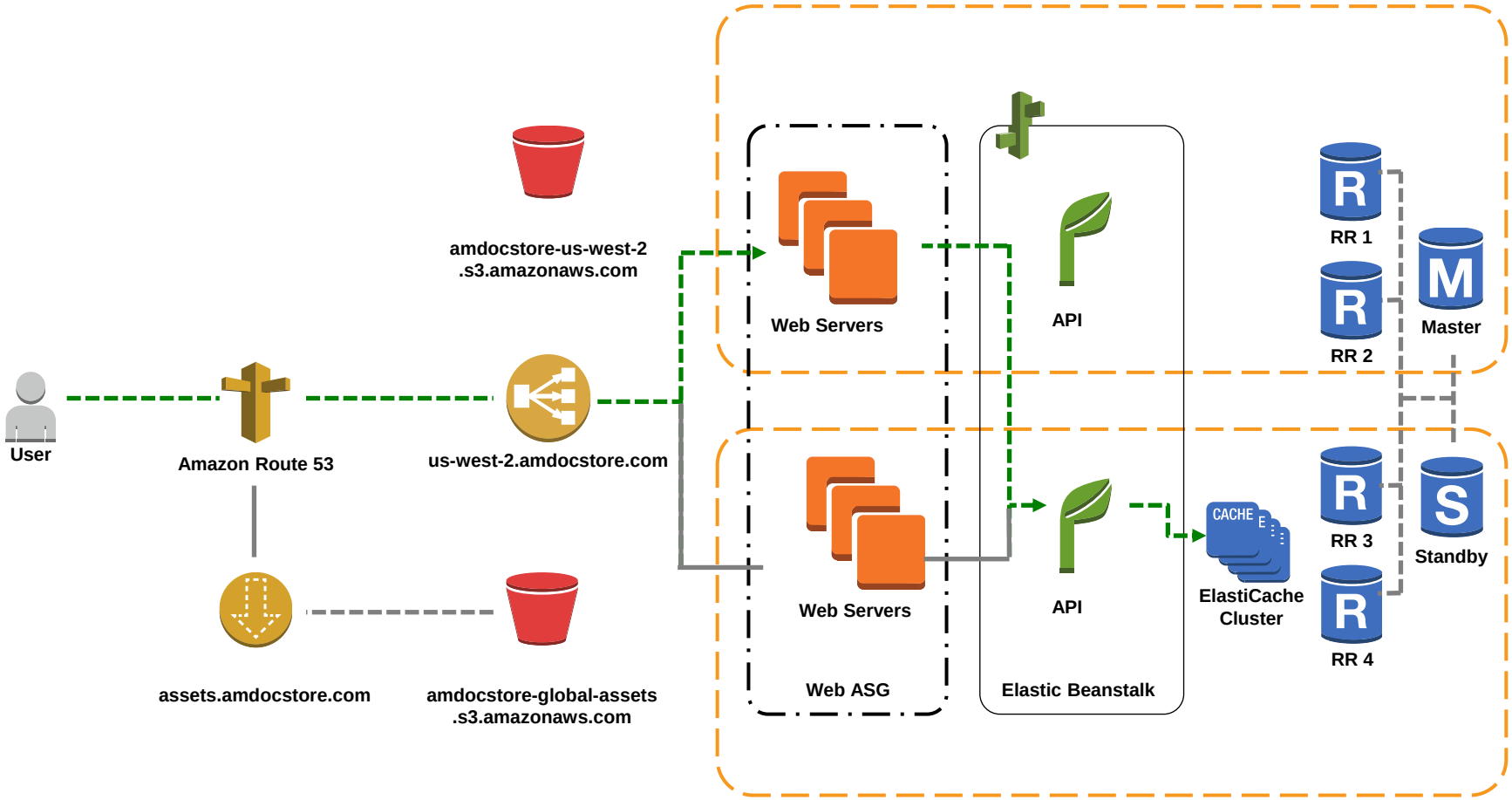


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

View and Download Docs: The result of the query is stored in the cache for future access

DocStore

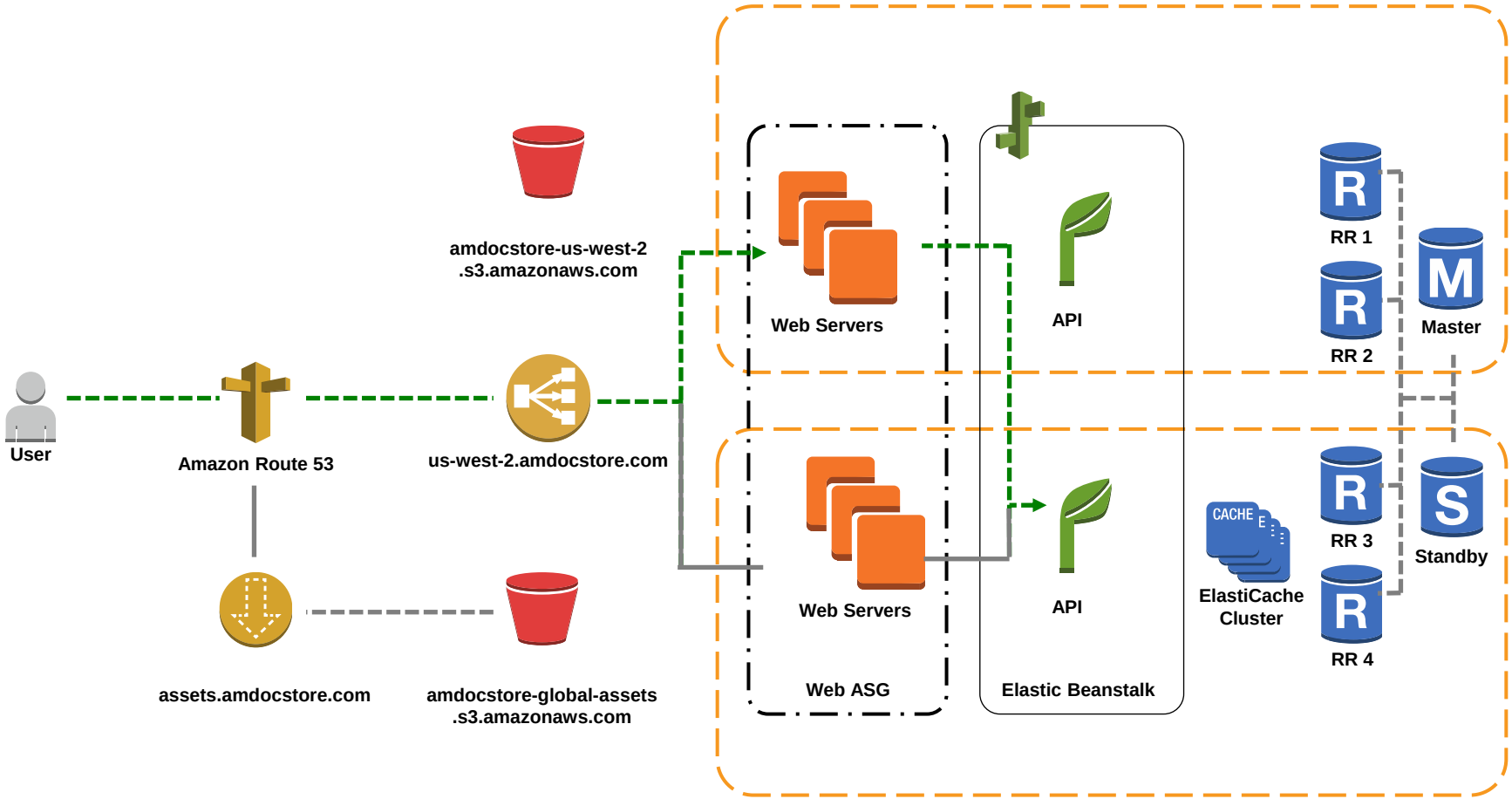


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

View and Download Docs: For each document, the API constructs a secure, auto-expiring, signed URL that will allow the user to download the document directly from S3

DocStore

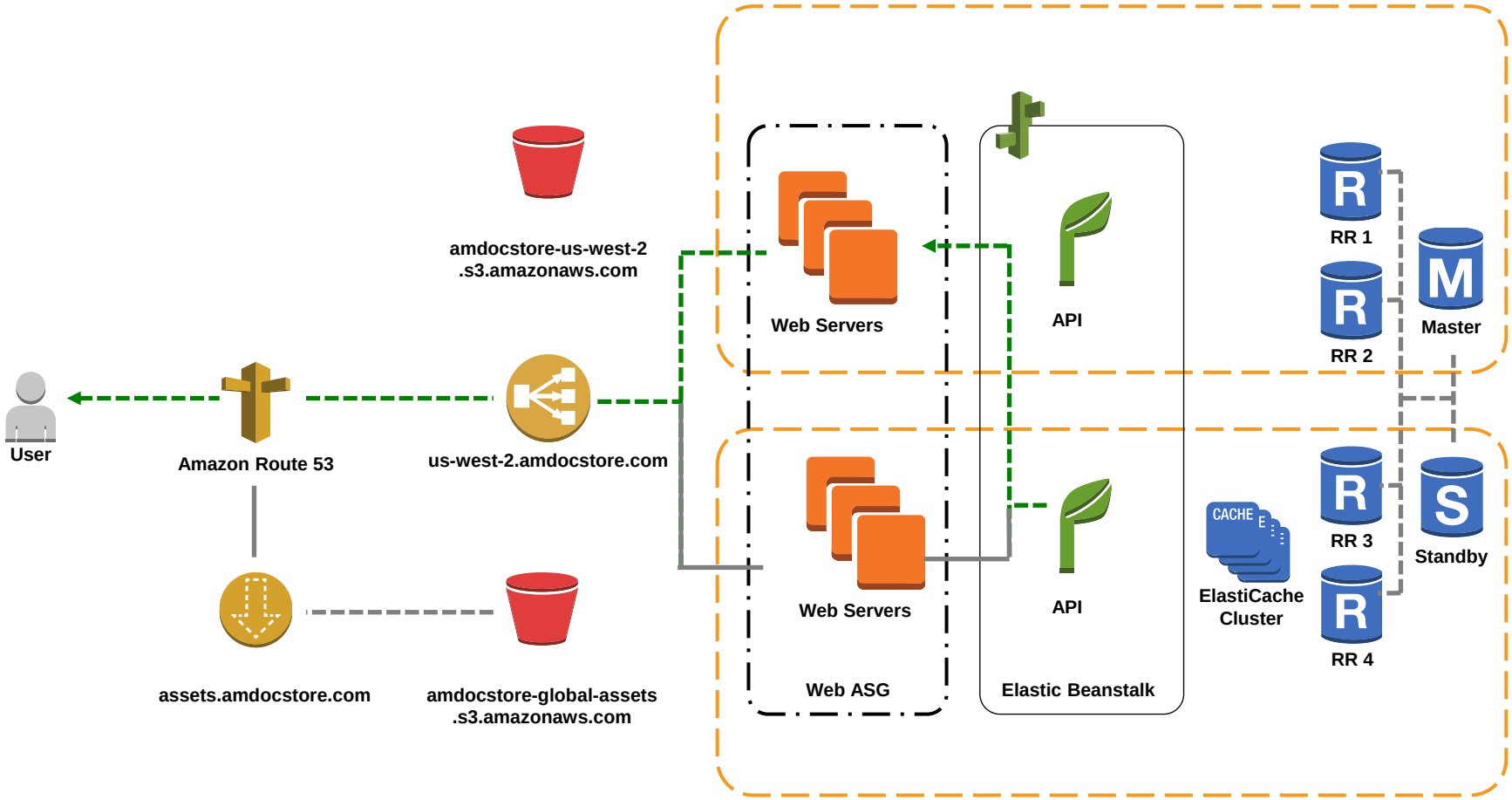


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

View and Download Docs: The list of documents – and links for direct, secure download – are delivered to the user's browser

DocStore

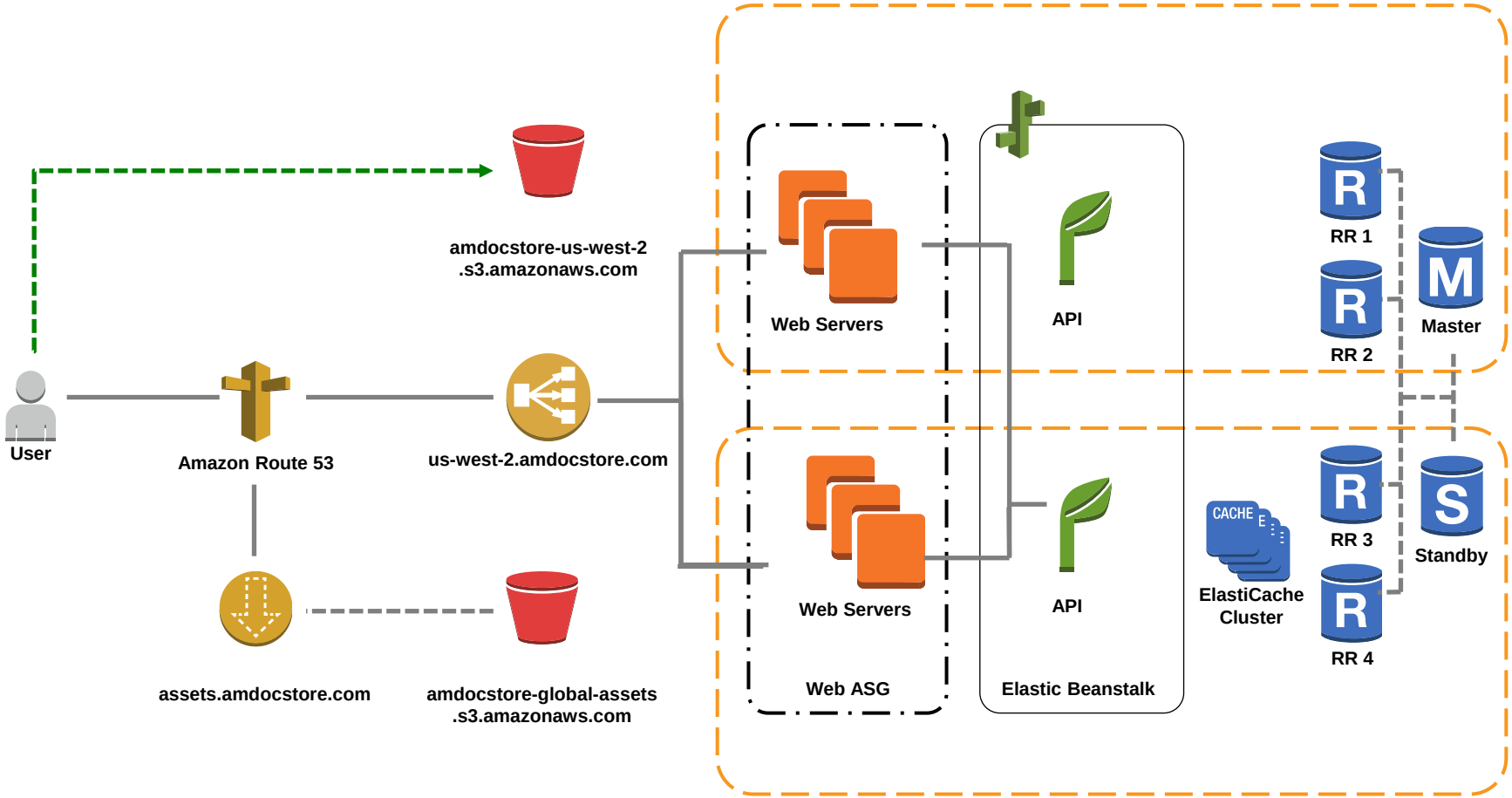


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

View and Download Docs: The user may download documents directly from S3

DocStore

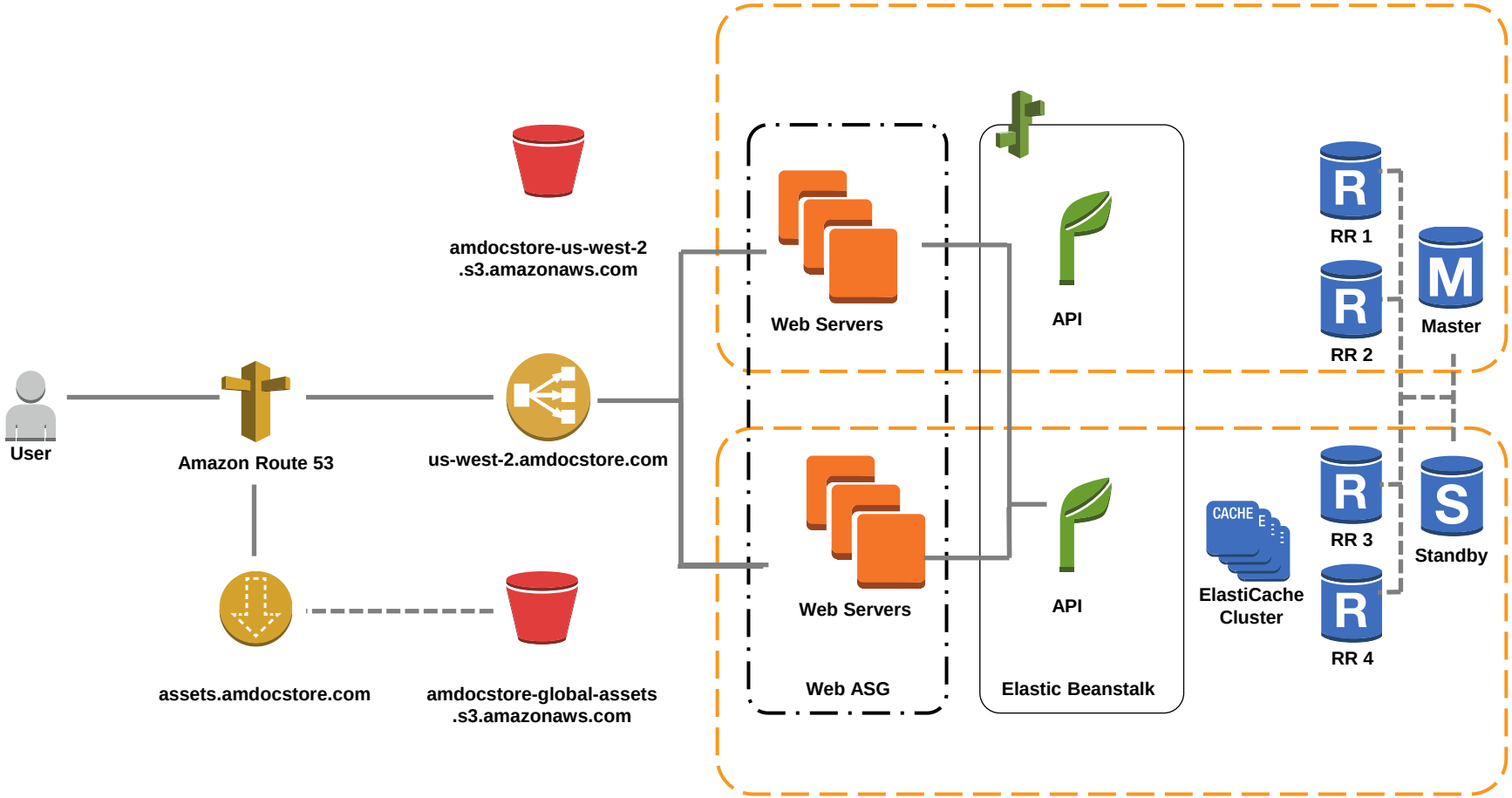


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Let's look at options for **security**

DocStore

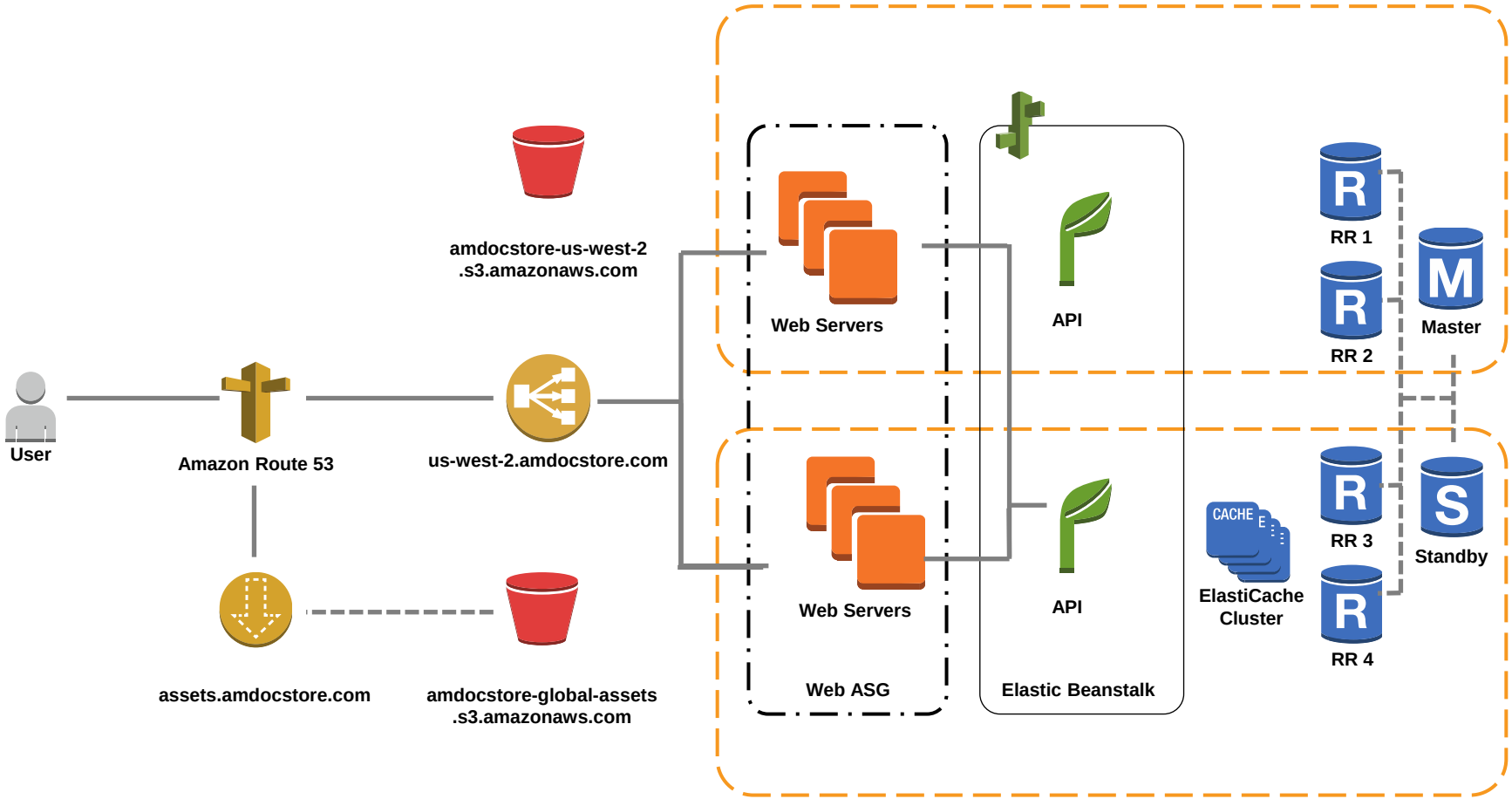


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Security: Where can security measures and controls be implemented in this architecture?

DocStore

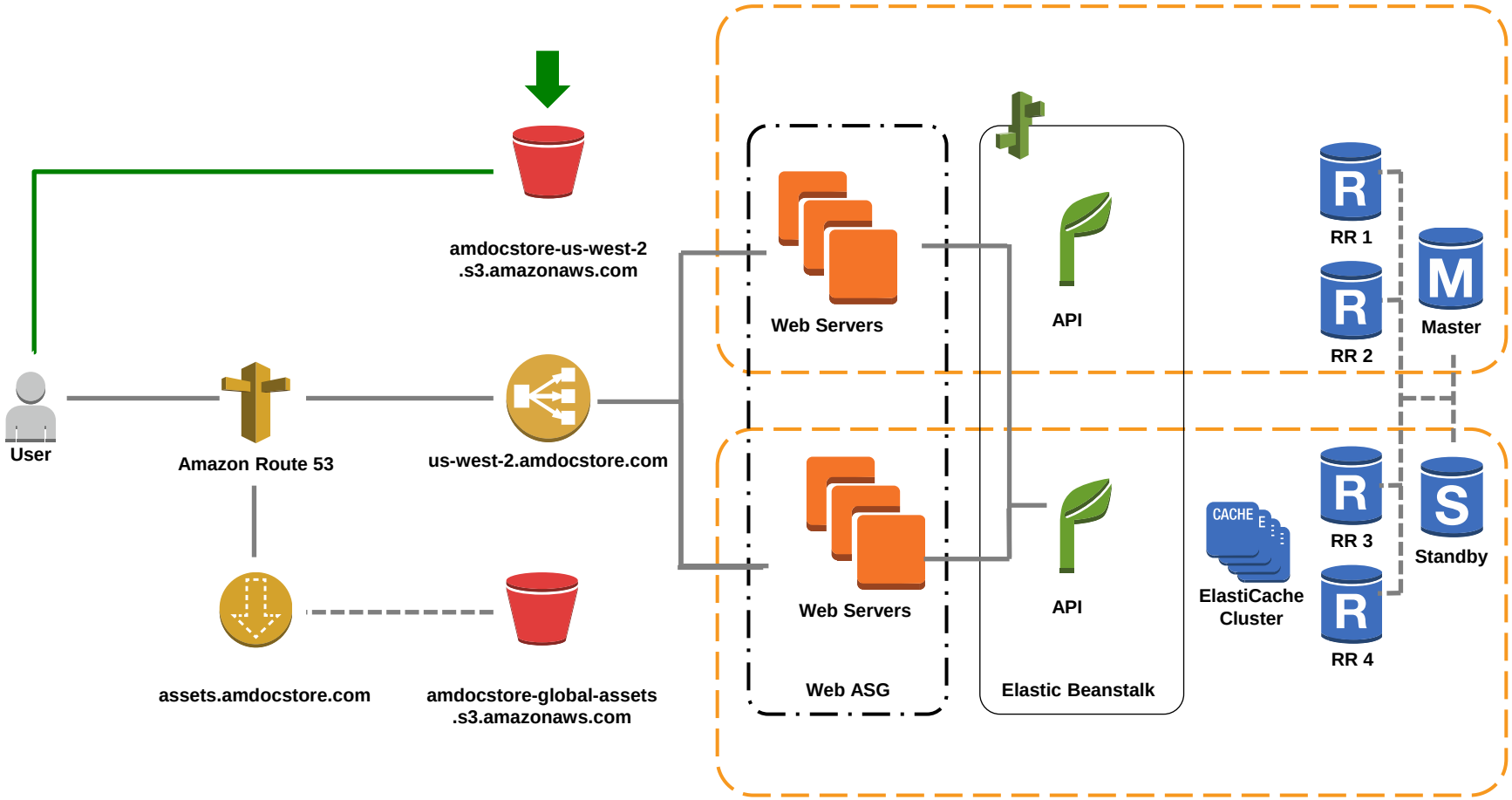


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Security: User downloads documents directly from S3 using a secure, signed URL and over SSL

DocStore

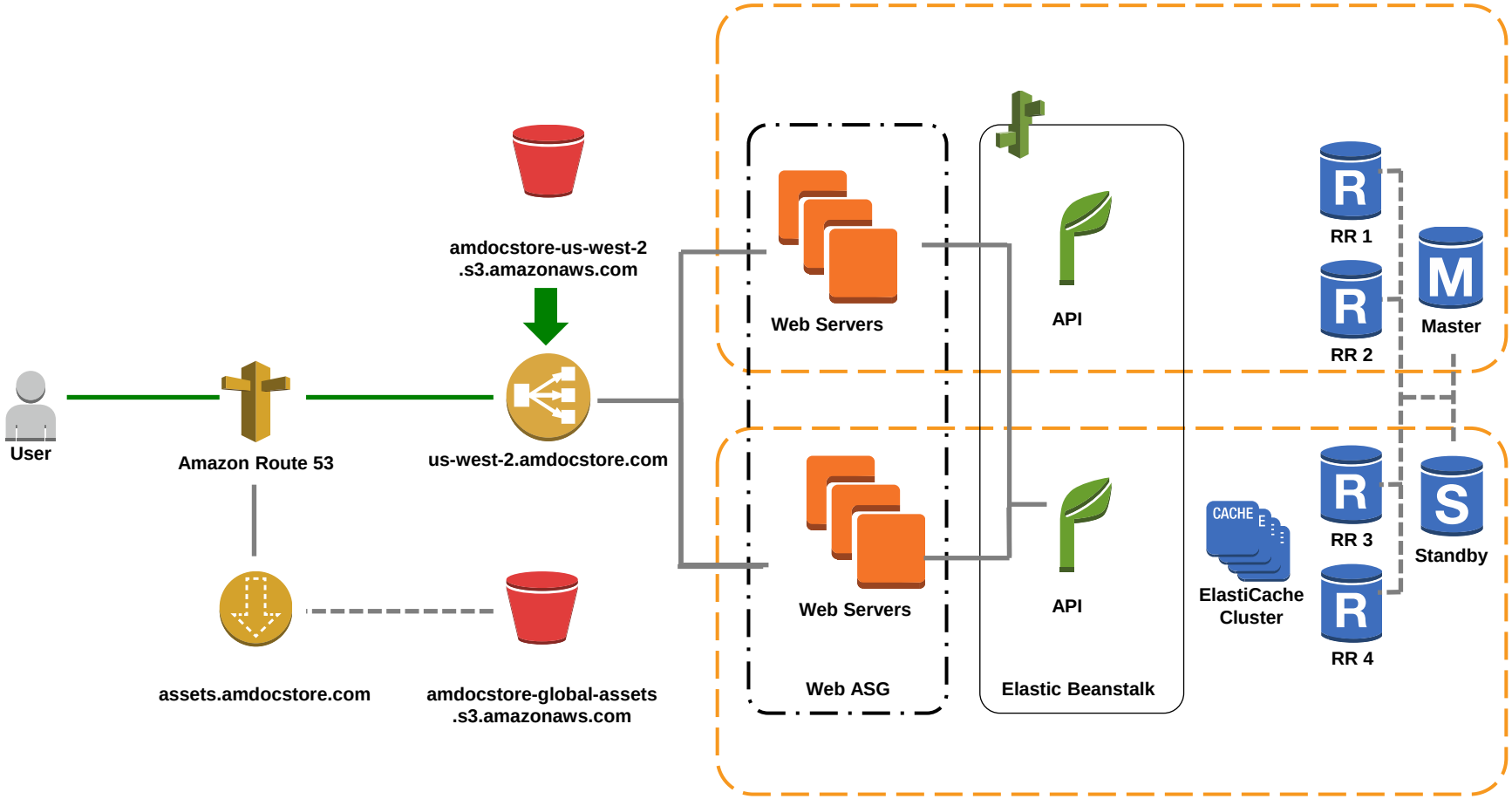


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Security: SSL connection between user and ELB

DocStore

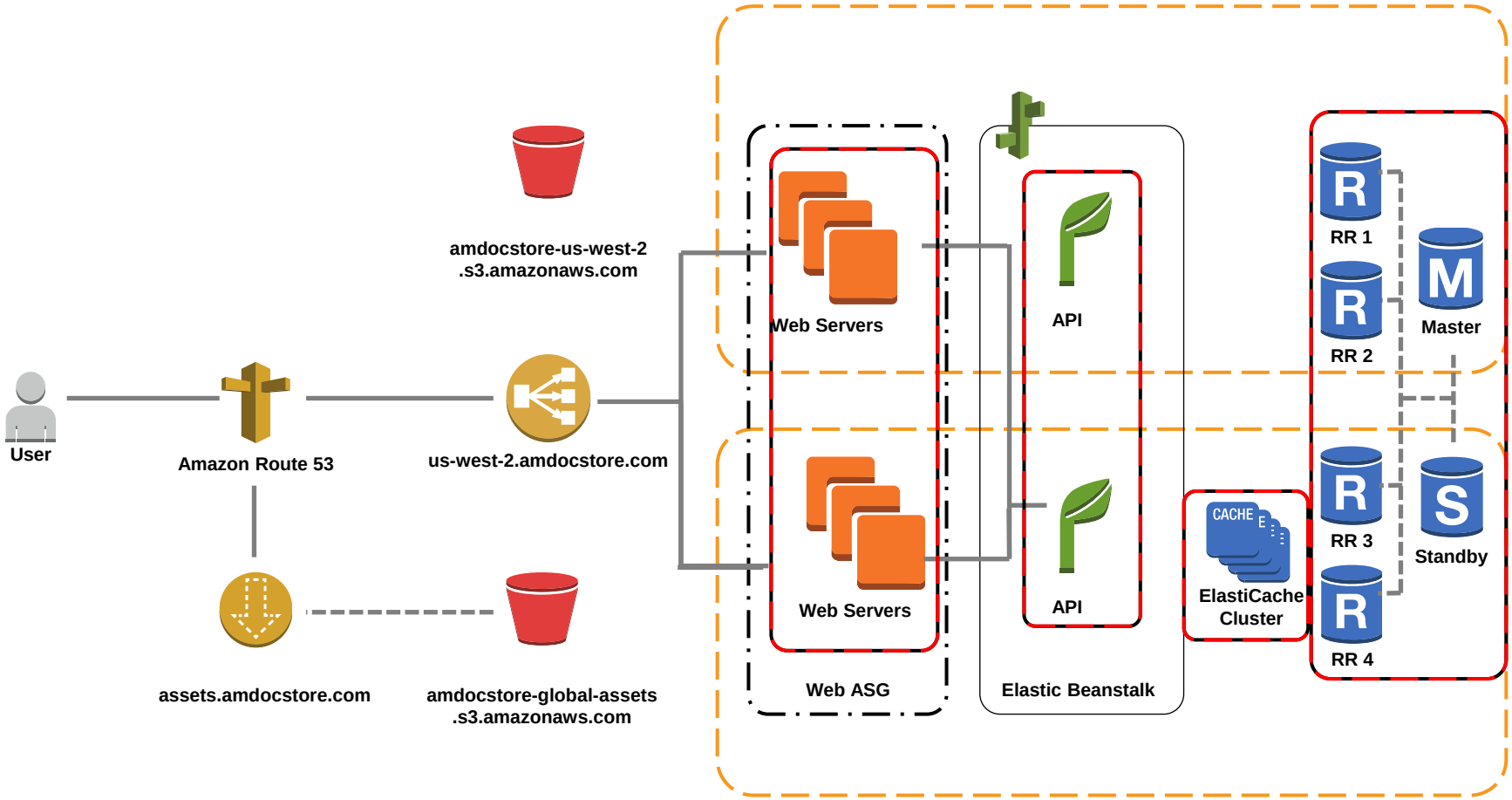


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Security: Security Groups restrict inbound traffic

DocStore

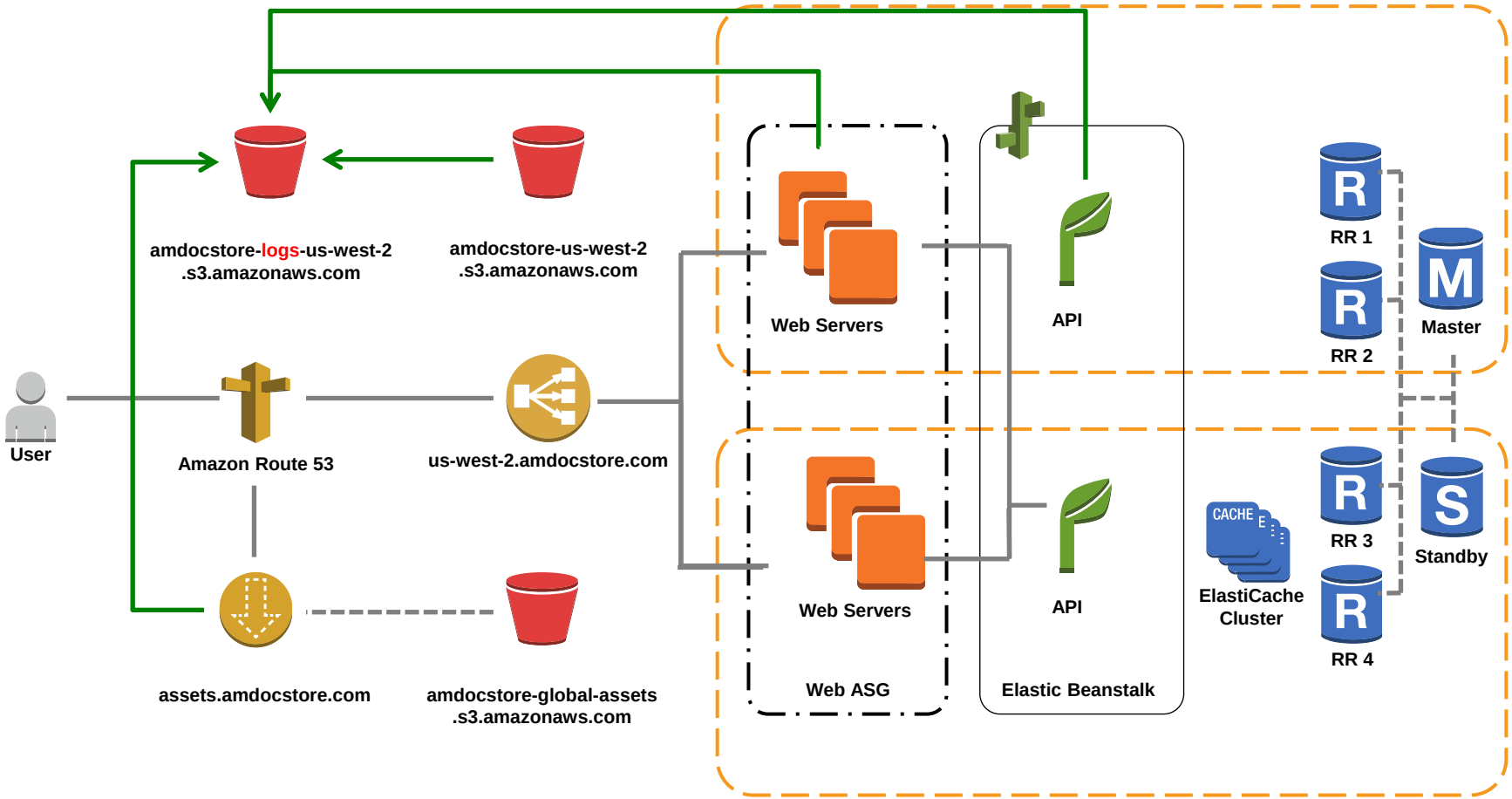


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Security: Access logs from CloudFront, EC2, and S3 are pushed to S3

DocStore

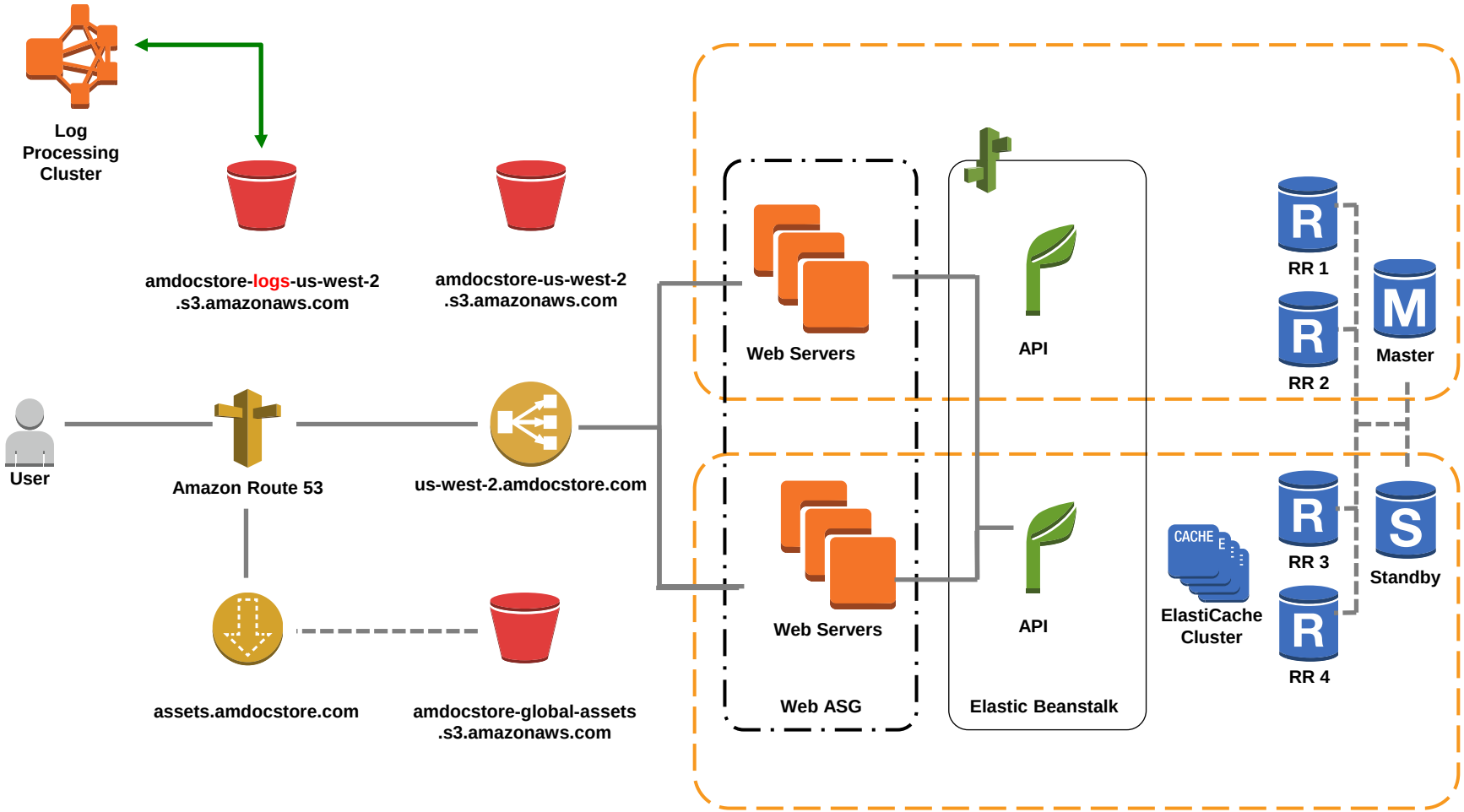


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Security: Logs are processed on a schedule using an Elastic MapReduce cluster

DocStore

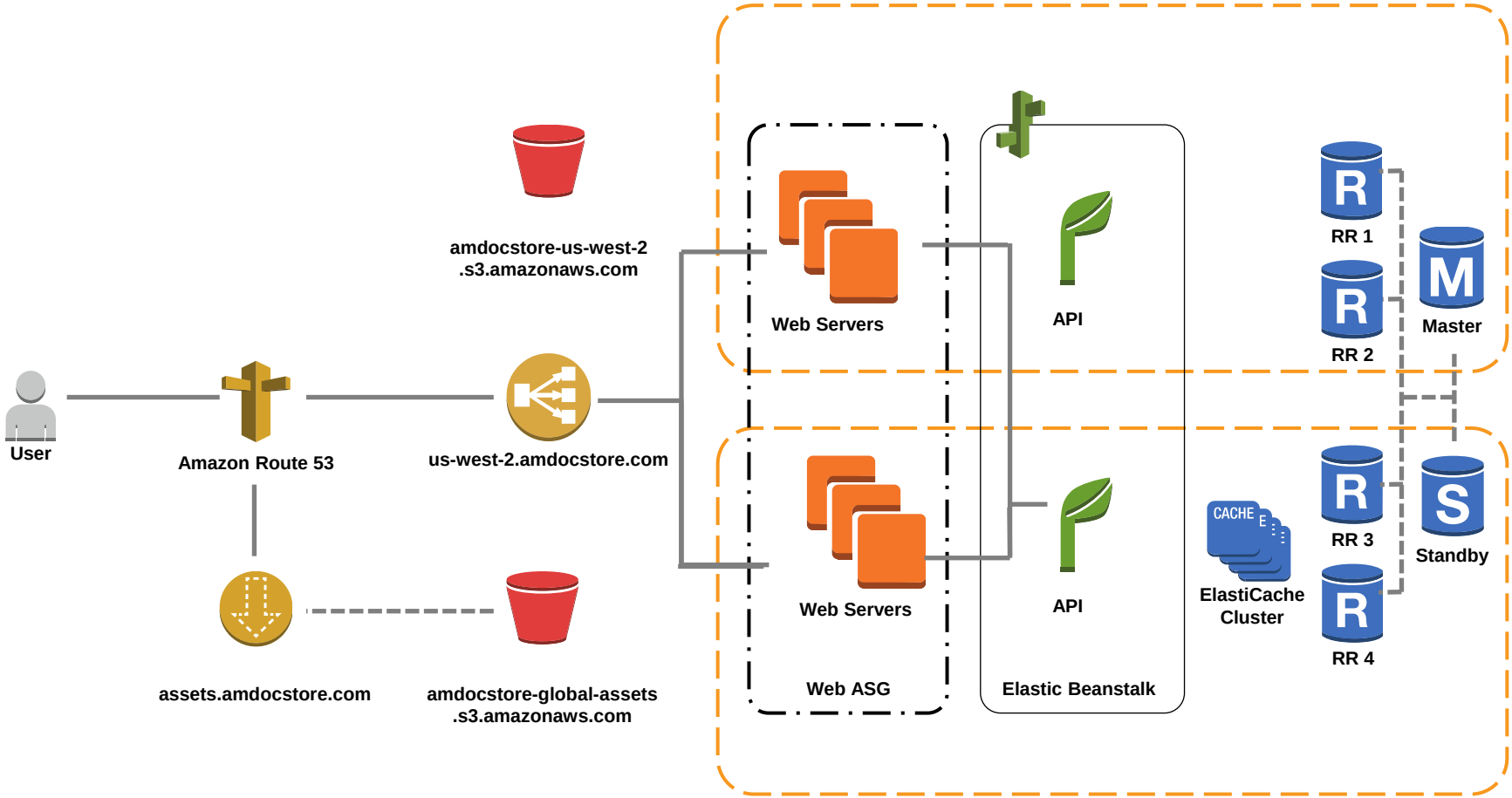


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Let's focus on the networking environment

DocStore

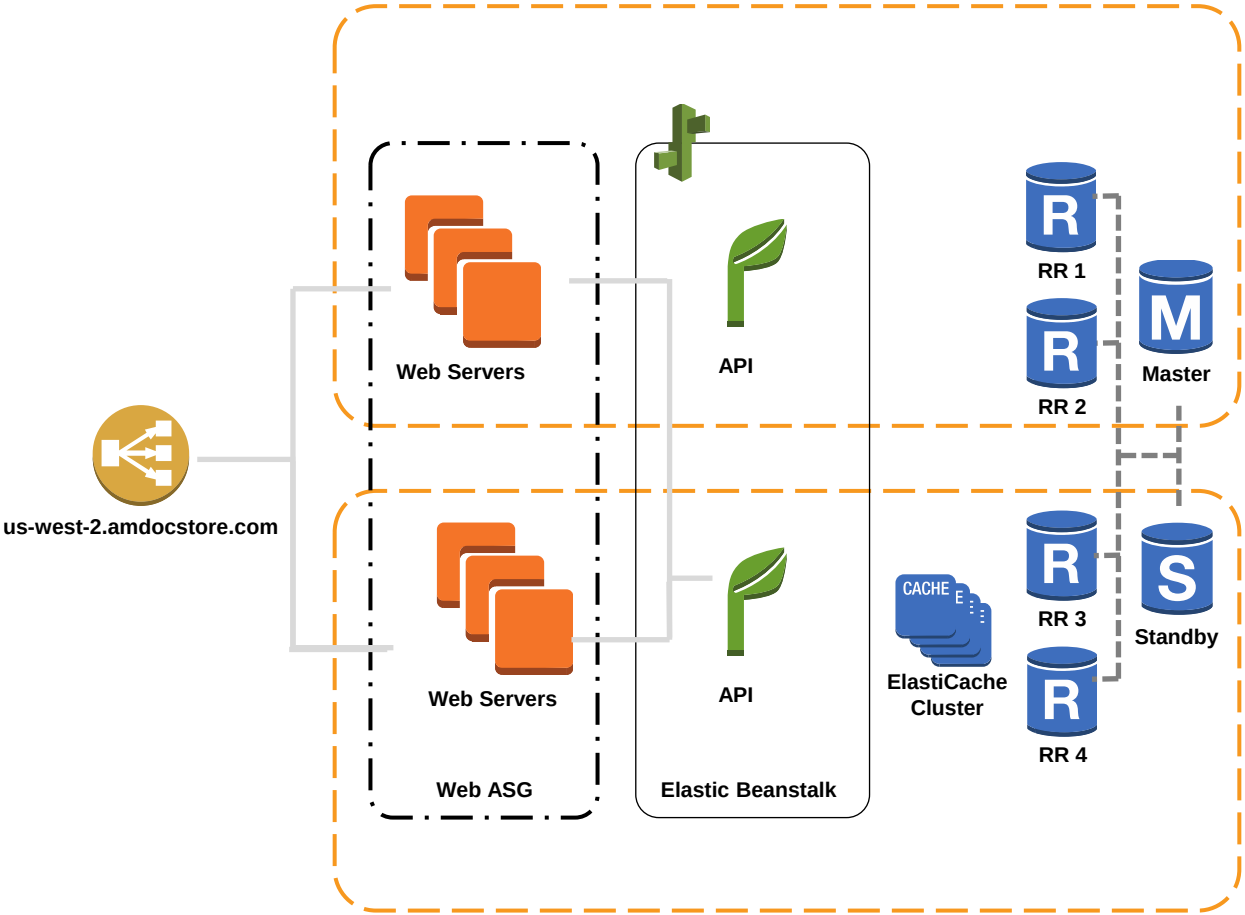


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Let's focus on the networking environment

DocStore

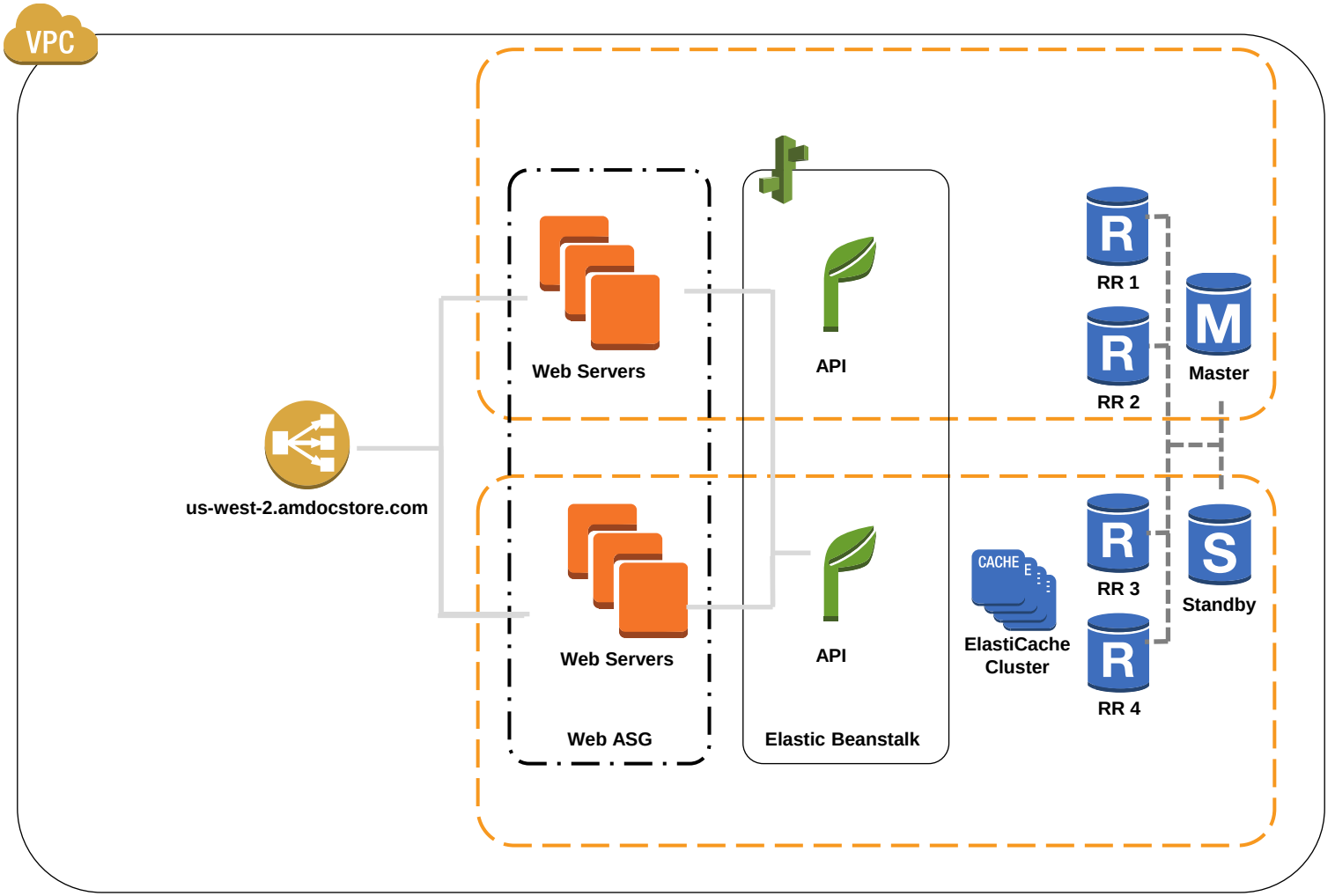


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

VPC: This VPC is in us-west-2 and has a 10.0.0.0/16 CIDR block

DocStore

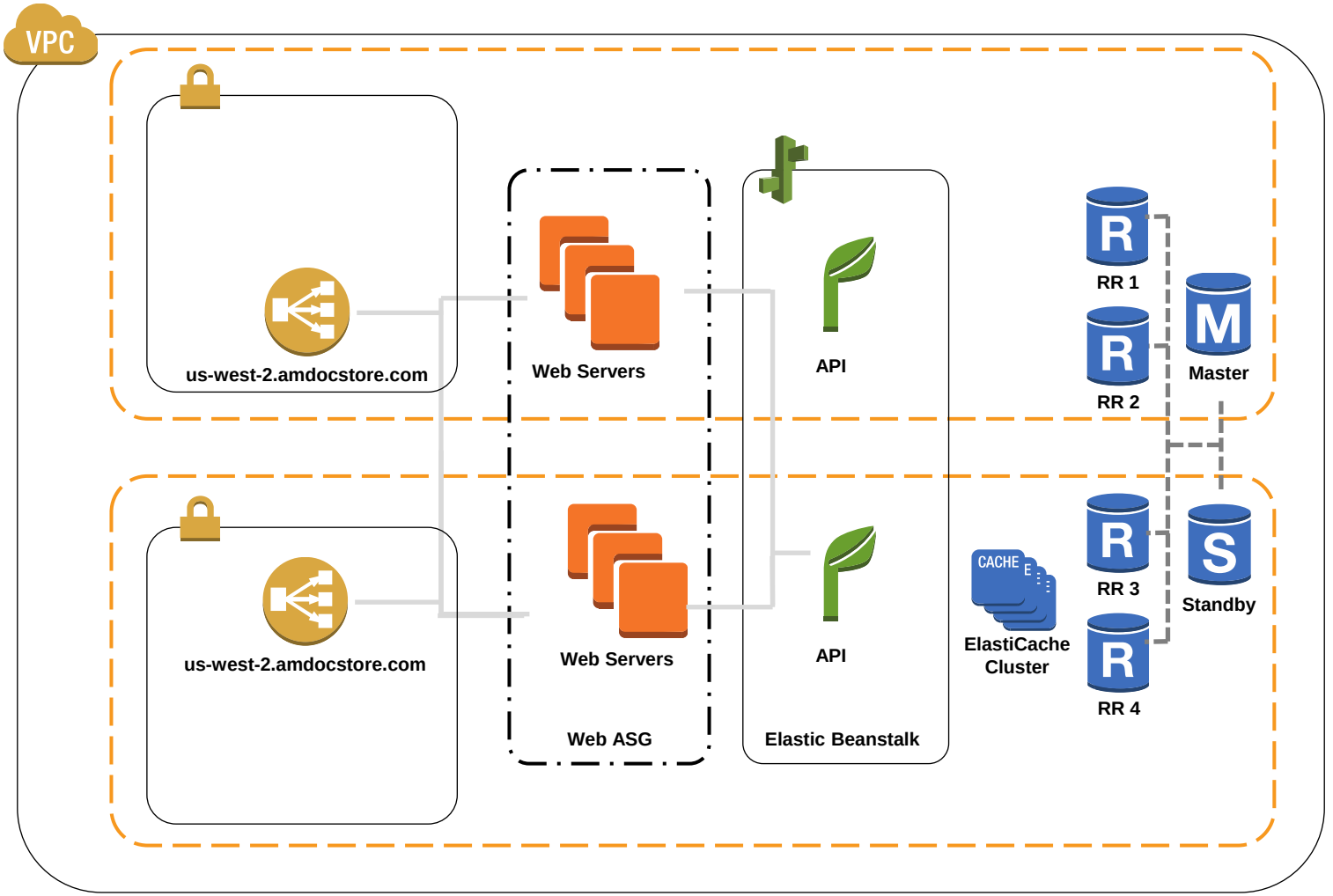


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

VPC: We define two subnets that our ELB will run in (note: although we show the ELB in multiple subnets, this is still a single ELB)

DocStore

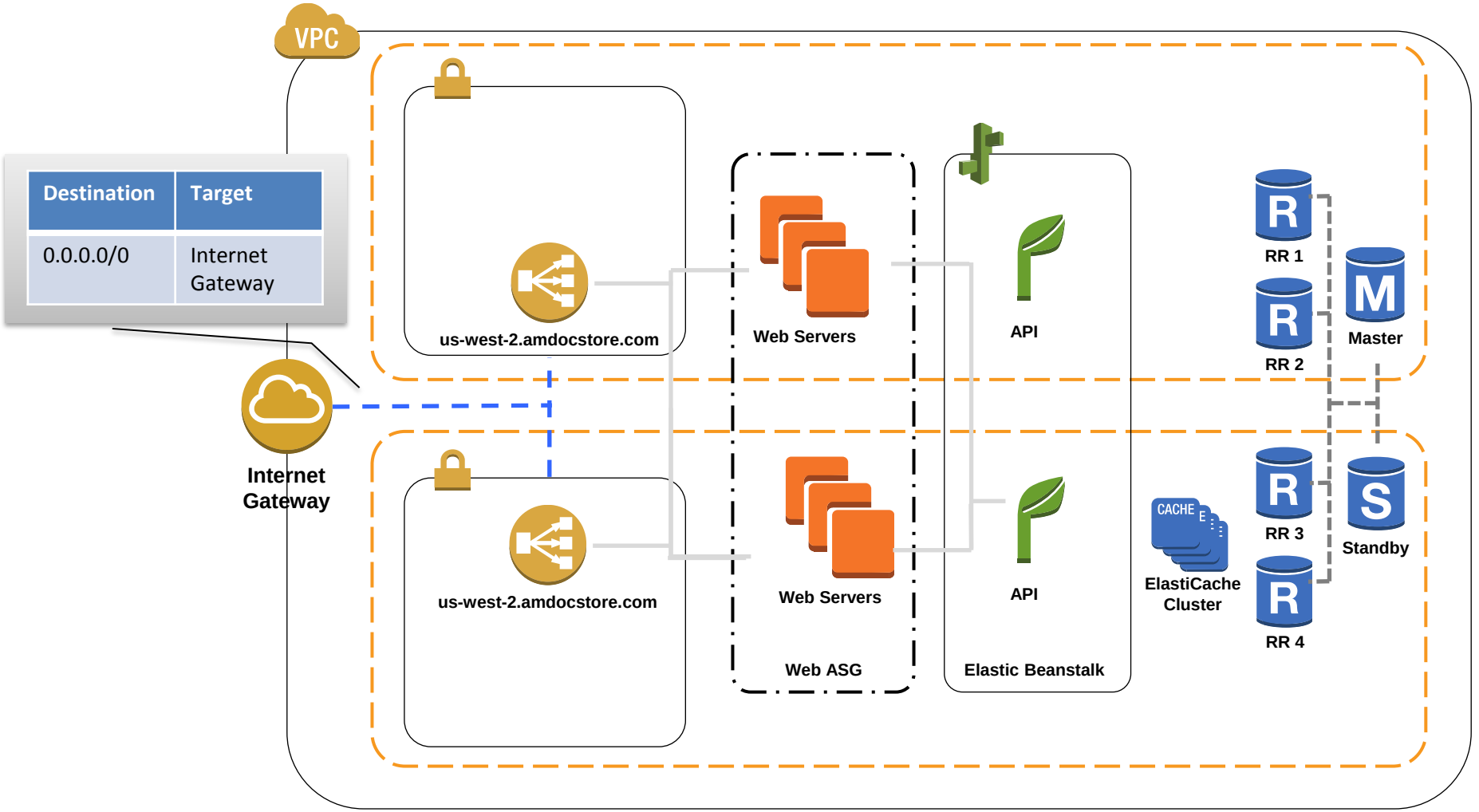


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

VPC: These subnets should be public. We will attach an **Internet Gateway** and **define a Route Table entry** to make them public

DocStore



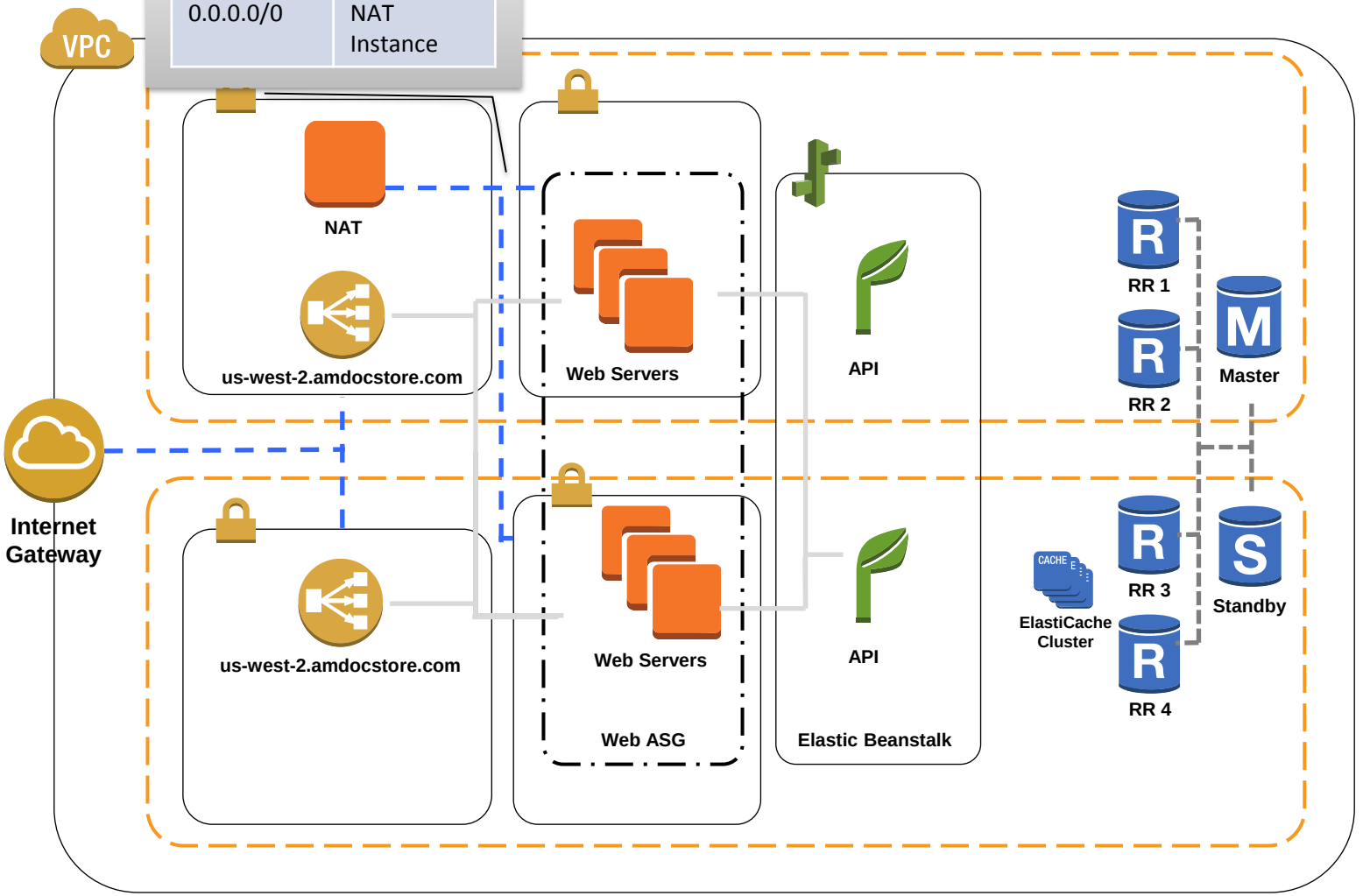
© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

VPC: Web servers run in a private subnet, but have **outbound internet access** via an EC2 instance in the public subnet running NAT

DocStore

Destination	Target
0.0.0.0/0	NAT Instance

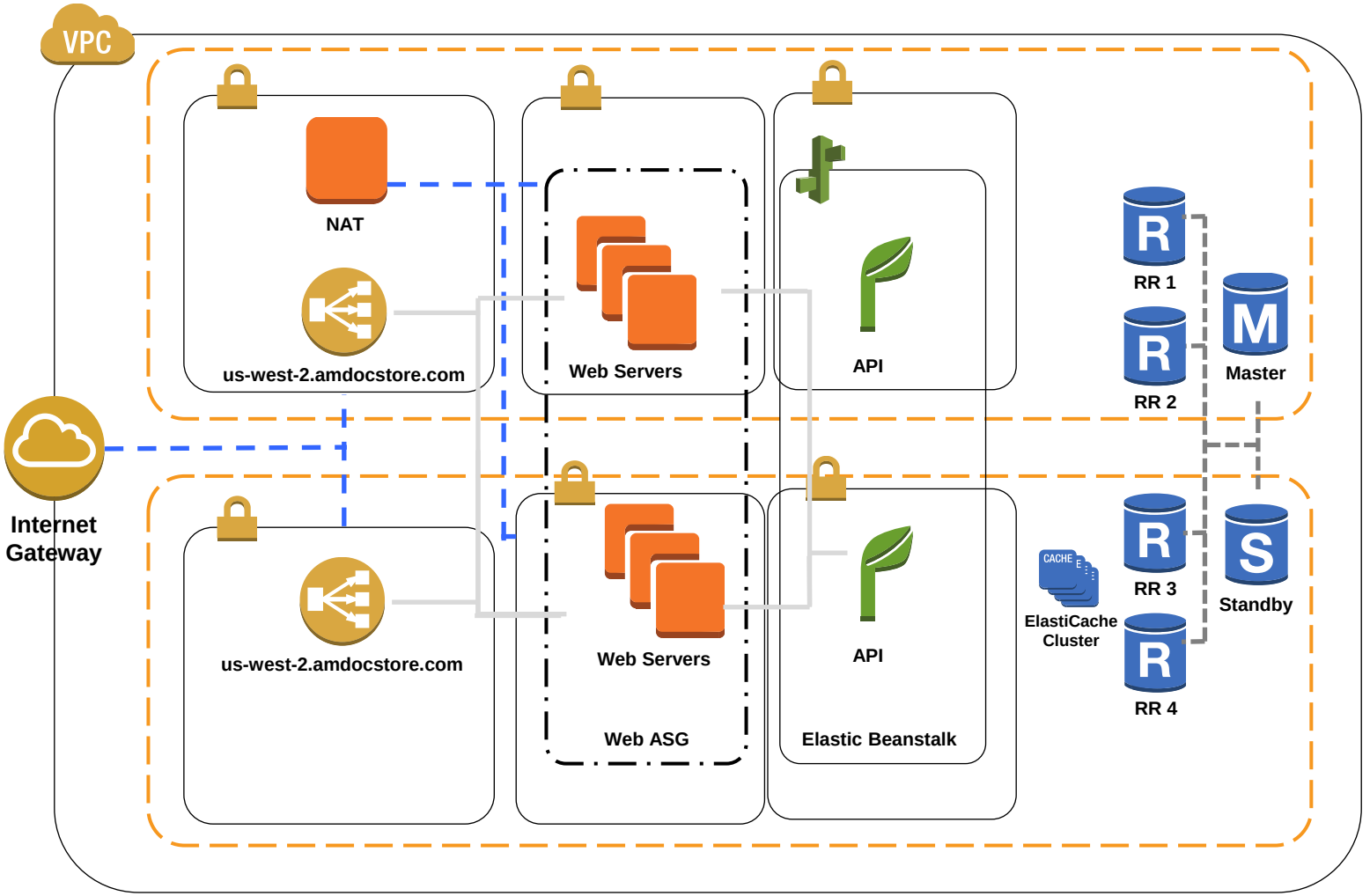


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

VPC: The DocStore API running on Elastic Beanstalk is deployed in both a public (for ELB) and private (for EC2) subnet

DocStore

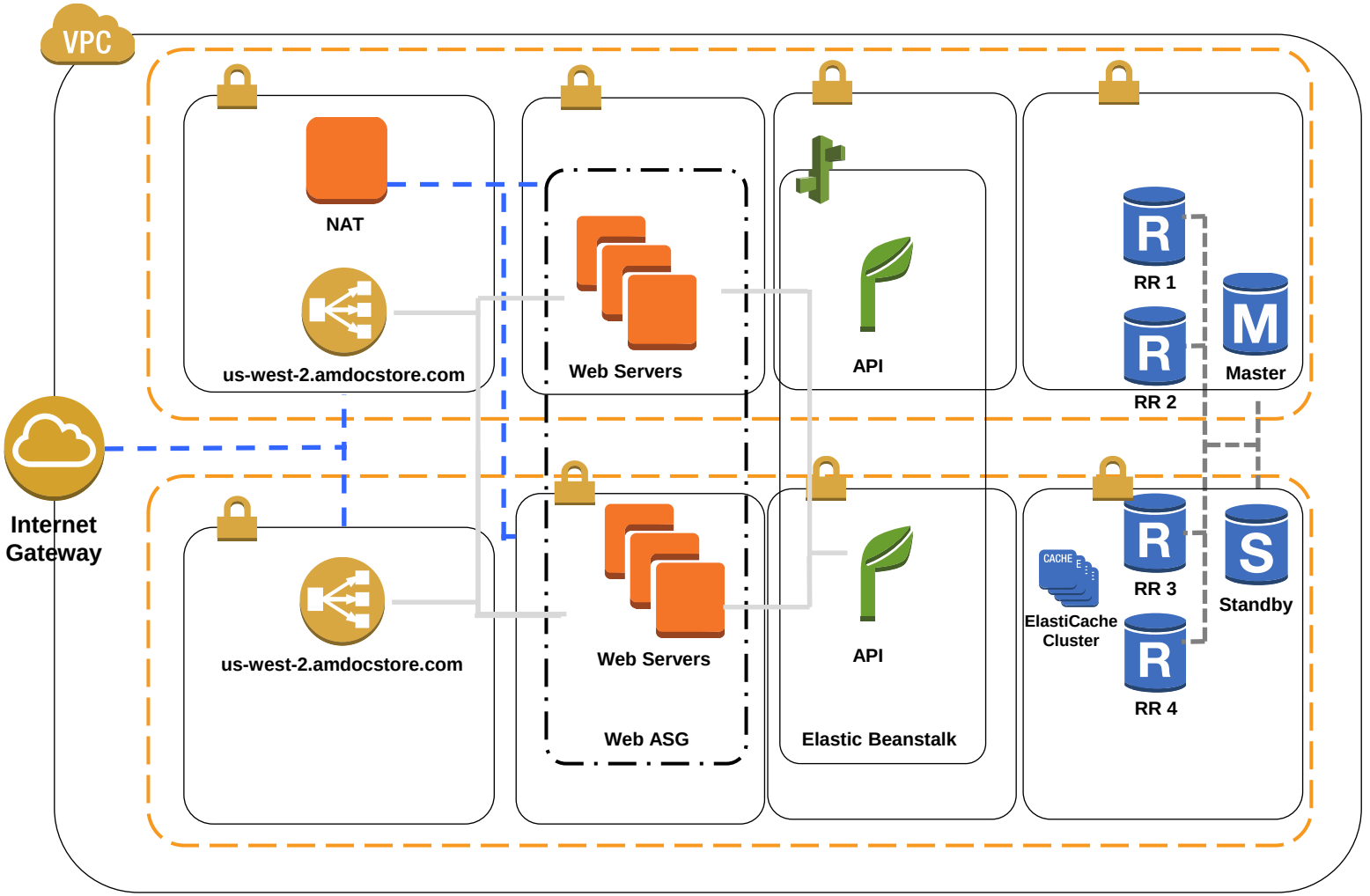


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

VPC: Finally, the RDS and ElastiCache instances are deployed in a private subnet

DocStore



© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

CloudFormation: Entire application is deployed as a stack from a template file

DocStore

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

CloudFormation: The sample template file is used to deploy to us-west-2

DocStore

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

CloudFormation: and ap-southeast-2

DocStore

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

DocStore

#2: Batch Processing Back-End

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

DocStore

#2: Batch Processing Back-End

- Sync user accounts
- Enforce max storage limits
- Extract and index text based on document type (i.e., PDF uses different filter than DOCX, etc.)
- Detailed dashboard with system-wide totals:
  - # docs, avg. doc size, total storage

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

DocStore

#2: Batch Processing Back-End

– Responds to events triggered by Web interface:

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

DocStore

#2: Batch Processing Back-End

– Responds to events triggered by Web interface:

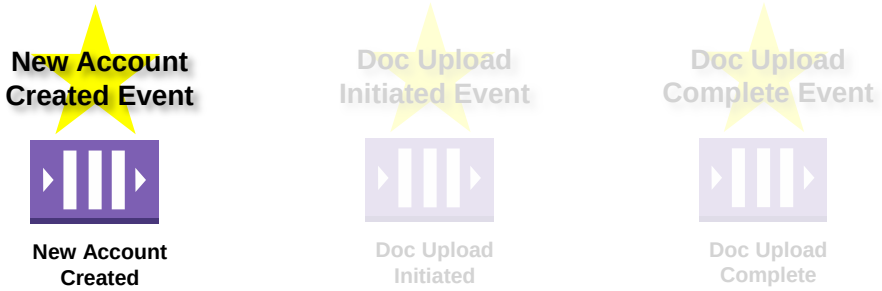
© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

DocStore

#2: Batch Processing Back-End

– Responds to events triggered by Web and Mobile interfaces:

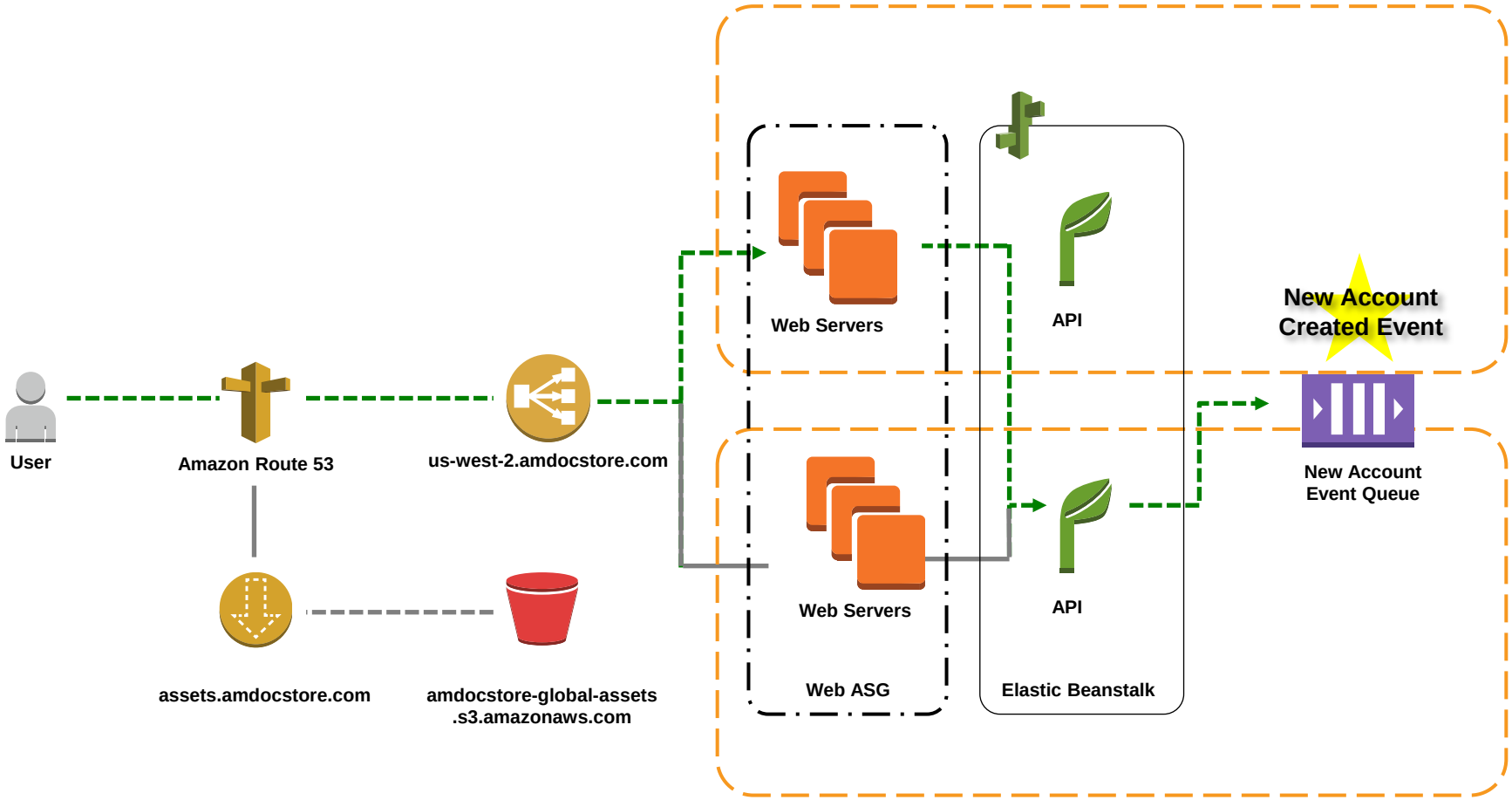


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

New Account Created Event: Triggered when a user creates a new account. Event message in queue includes username, password, and home region.

DocStore

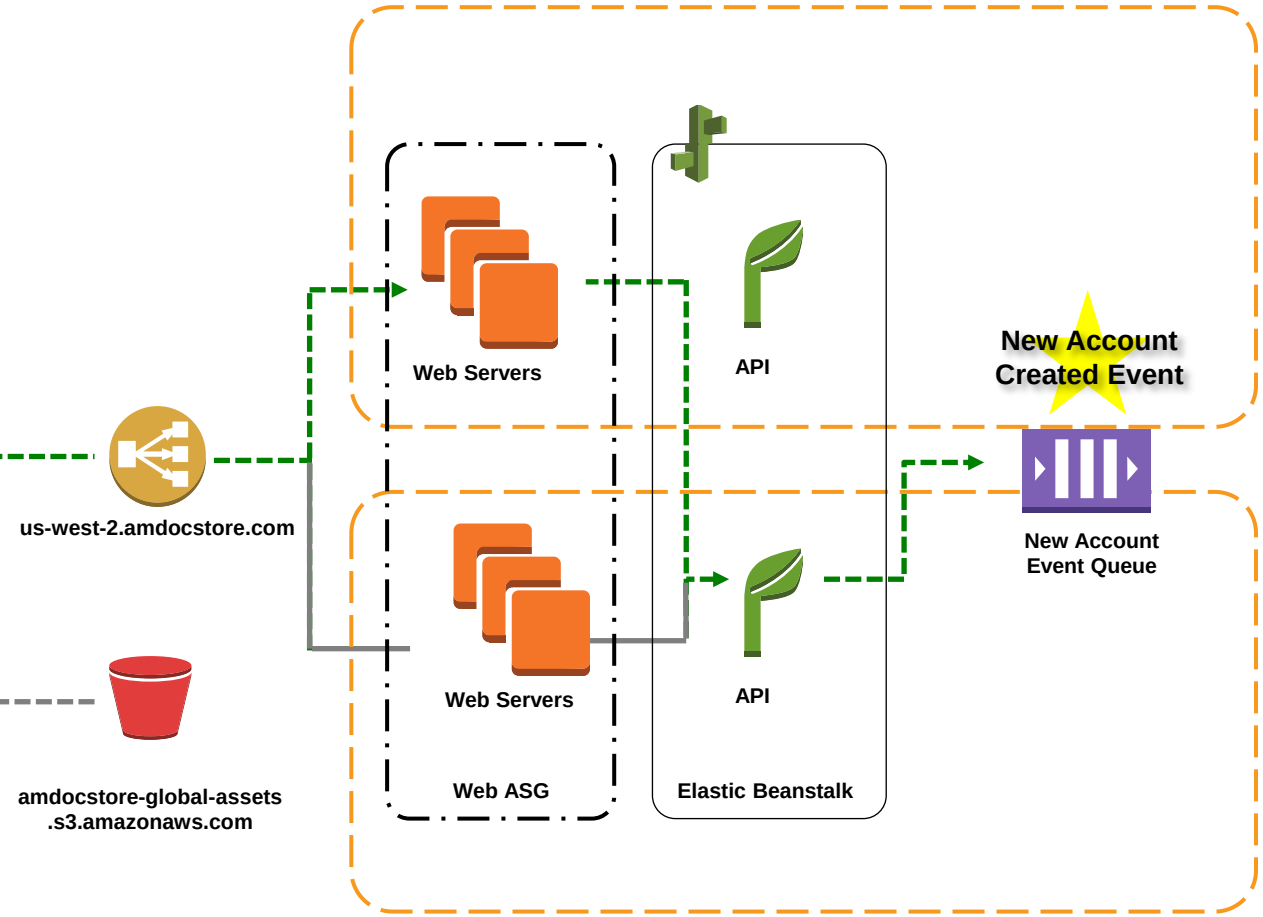


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

New Account Created Event: Triggered when a user creates a new account. Event message in queue includes username, password, and home region.

DocStore

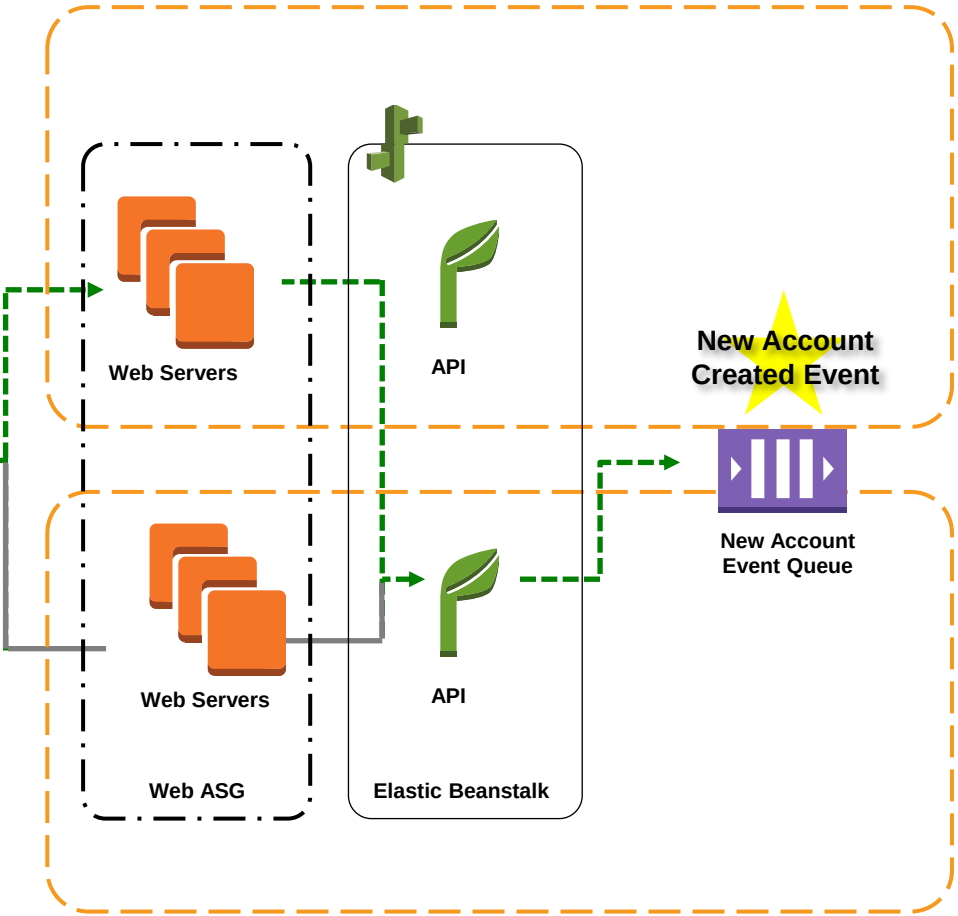


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

New Account Created Event: Triggered when a user creates a new account. Event message in queue includes username, password, and home region.

DocStore

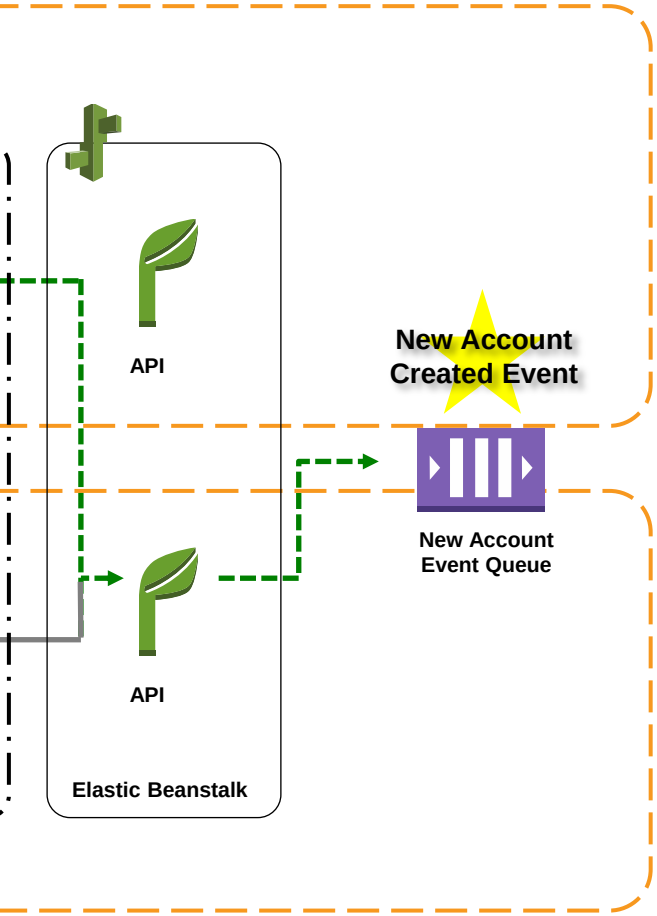


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

New Account Created Event: Triggered when a user creates a new account. Event message in queue includes username, password, and home region.

DocStore

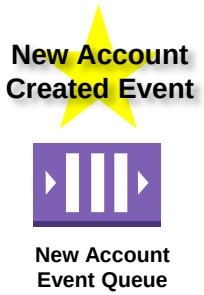


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

New Account Created Event: Triggered when a user creates a new account. Event message in queue includes username, password, and home region.

DocStore

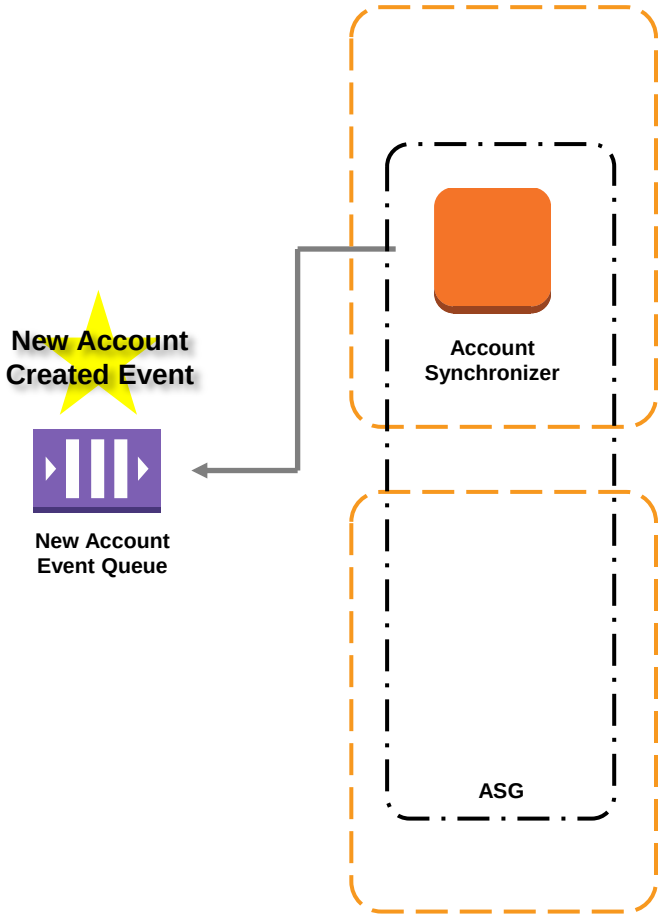


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

New Account Created Event: An instance running the **Account Synchronization servic**e polls the New Account Event Queue

DocStore

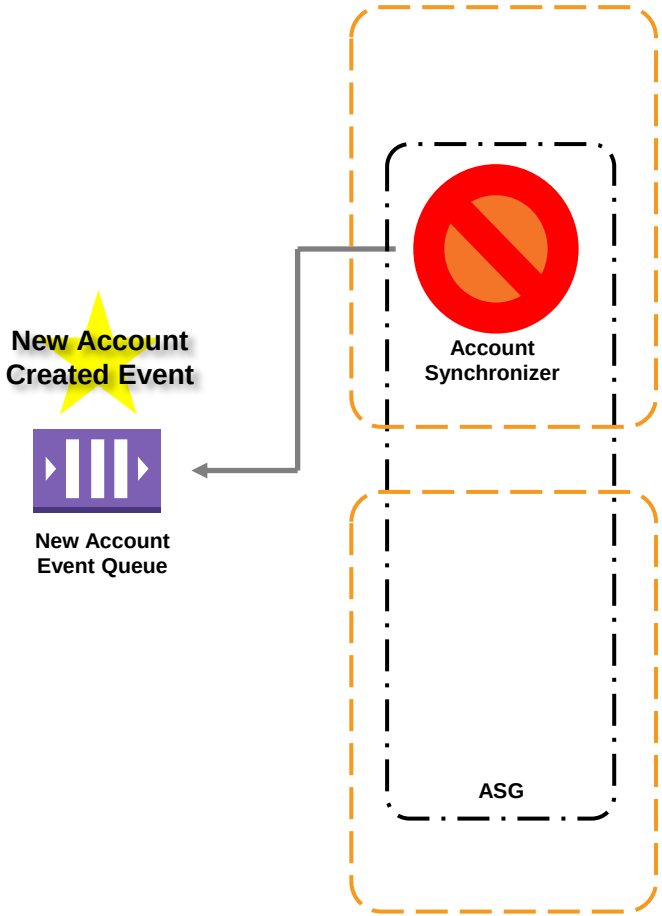


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

New Account Created Event: If the instance fails, Auto Scaling will launch a replacement

DocStore

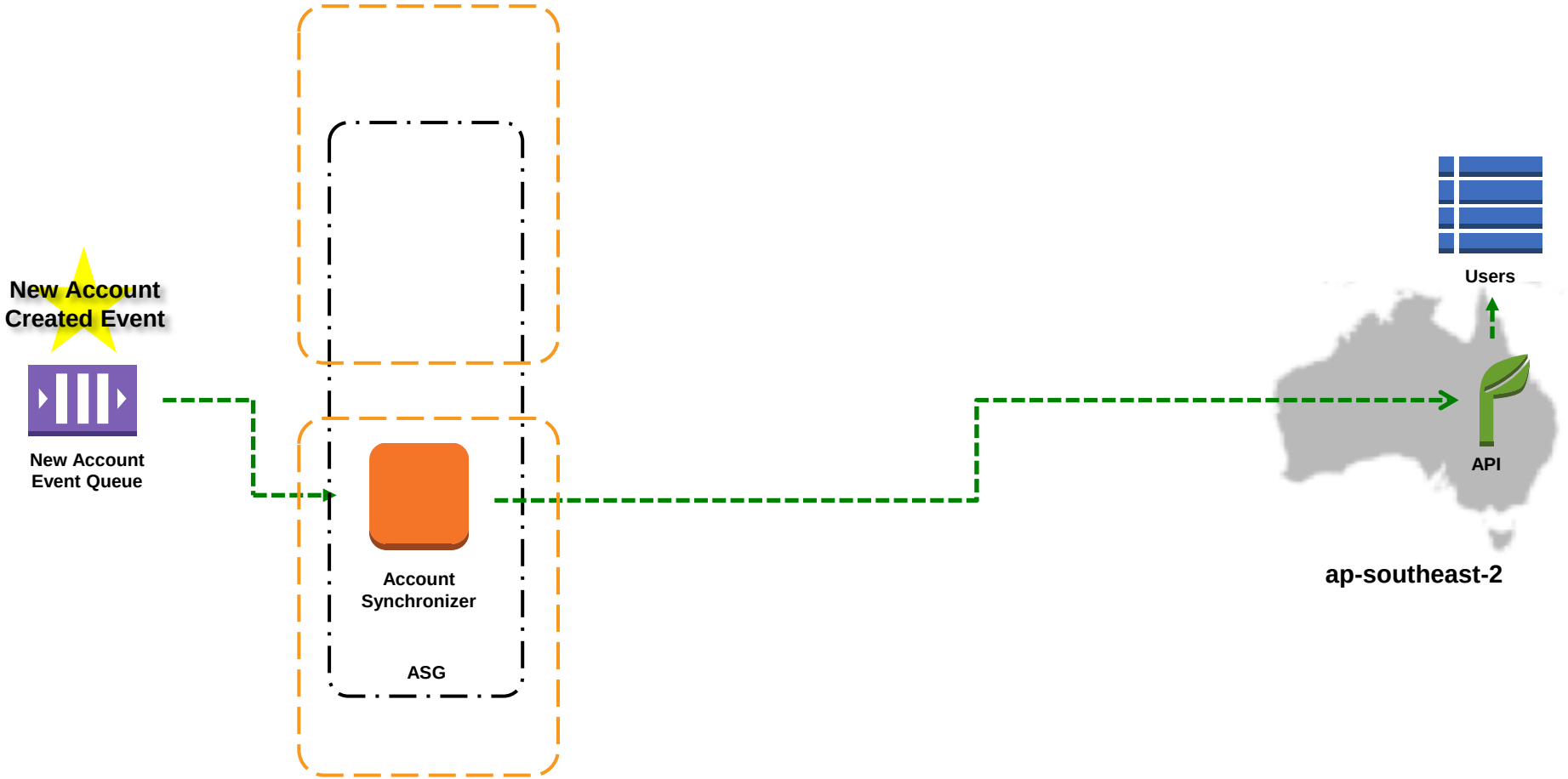


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

New Account Created Event: Account Synchronizer invokes API in ap-southeast-2 to synchronize new user's account data

DocStore

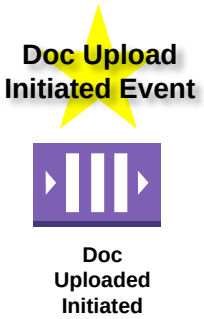


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Doc Uploaded Initiated Event: When a user begins uploading a document to S3, an event is triggered so the upload can be tracked via a workflow

DocStore



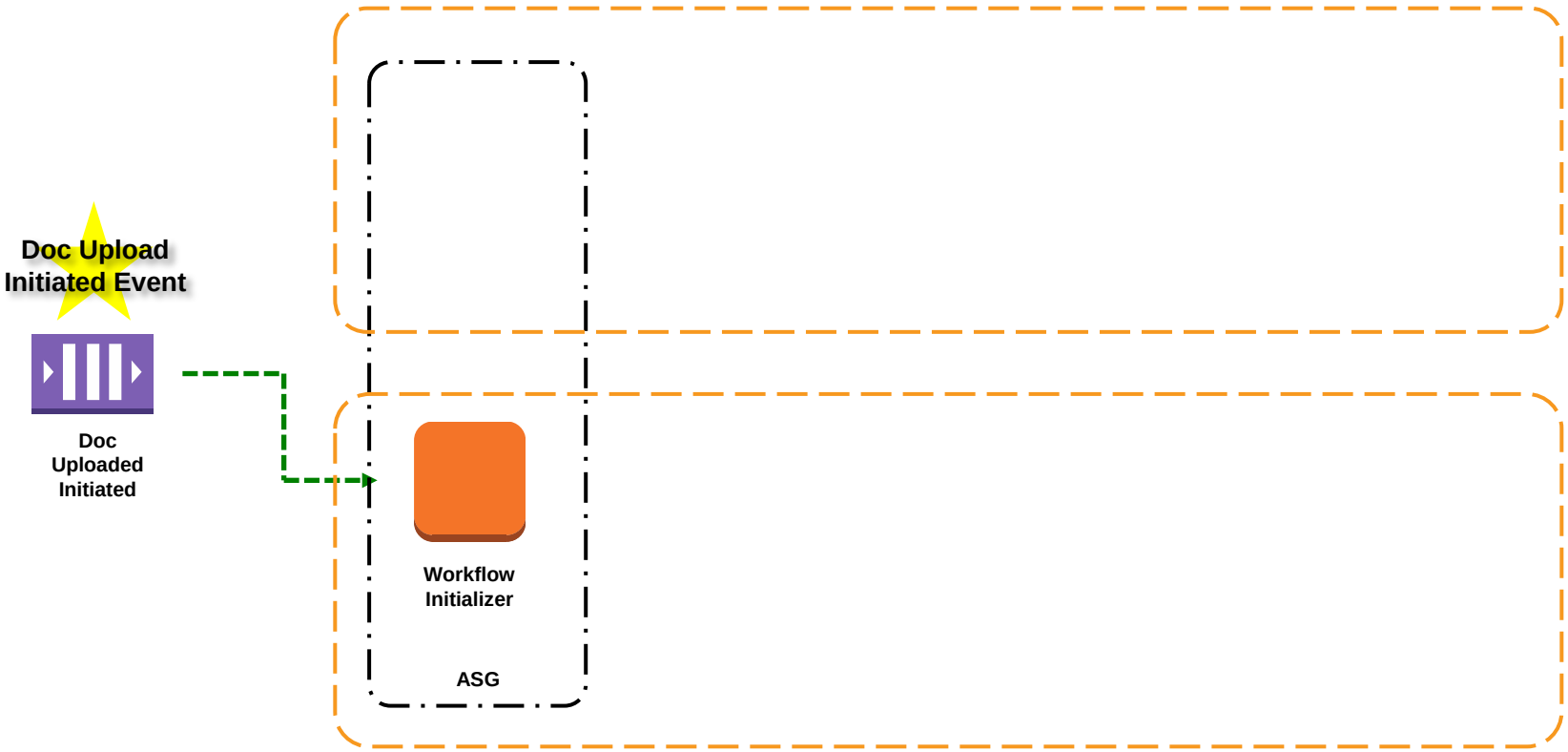
Doc
Uploaded
Initiated

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Doc Uploaded Initiated Event: When an event is received from the queue, the Workflow Initializer kicks off a SWF workflow execution

DocStore

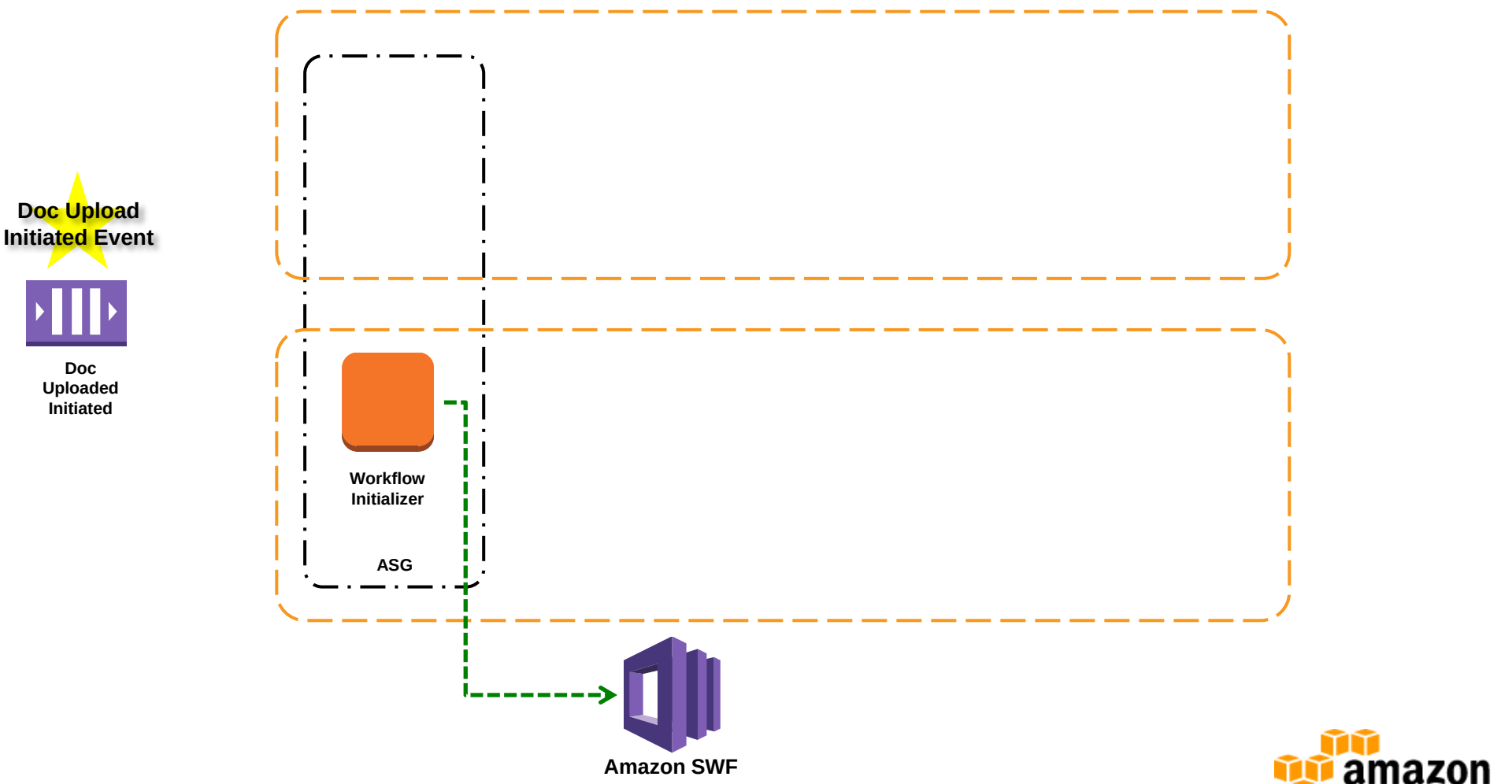


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Doc Uploaded Initiated Event: When an event is received from the queue, the Workflow Initializer kicks off a SWF workflow execution

DocStore

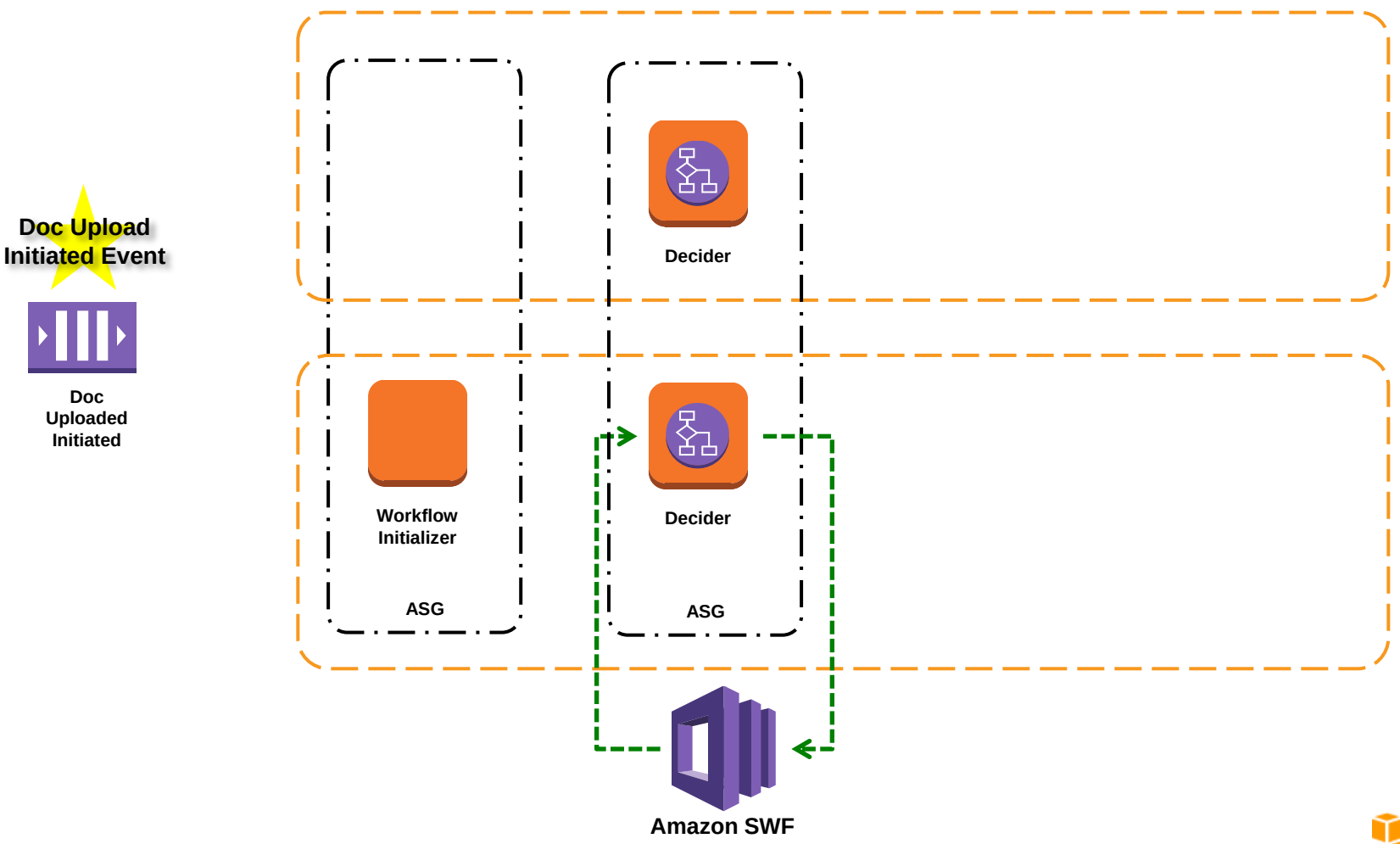


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Doc Uploaded Initiated Event: A an instance running Decider code receives the **workflow execution started** task, and schedules a *Check Upload Status* activity

DocStore

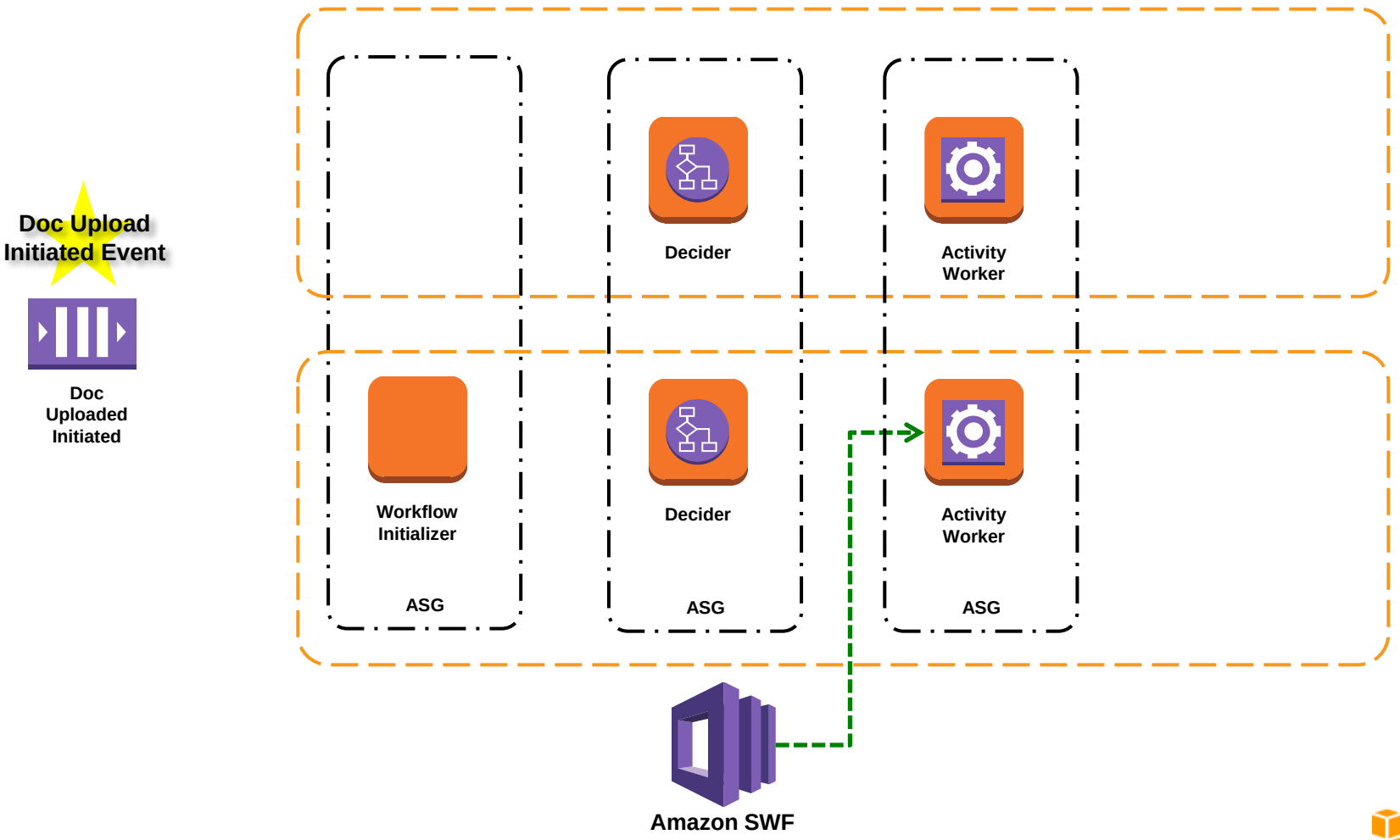


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Doc Uploaded Initiated Event: A an instance running the Activity Worker code receives the *Check Upload Status* task

DocStore

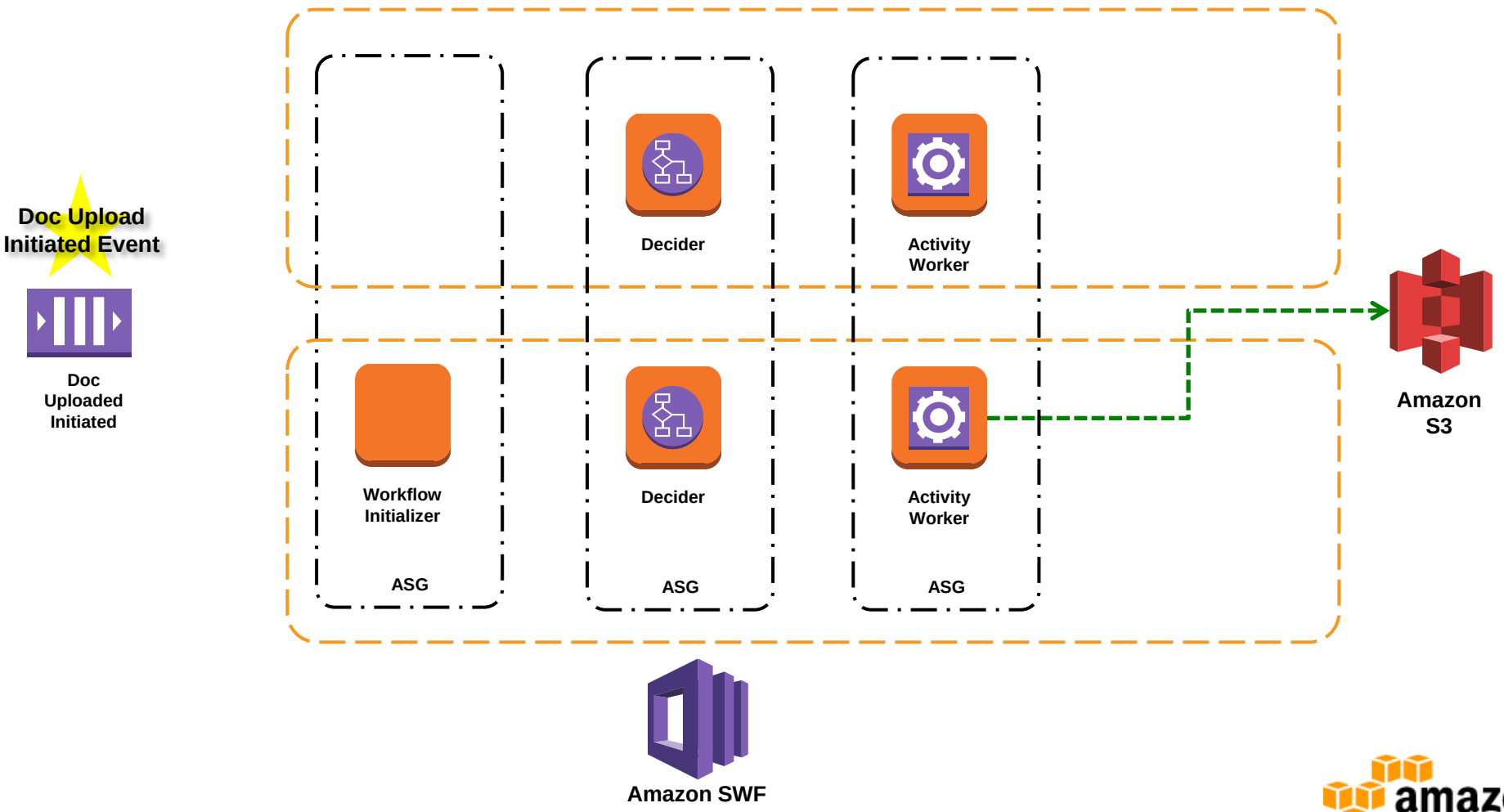


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Doc Uploaded Initiated Event: The Activity Worker periodically checks S3 to determine if the doc upload has completed (i.e., the object exists in S3)

DocStore

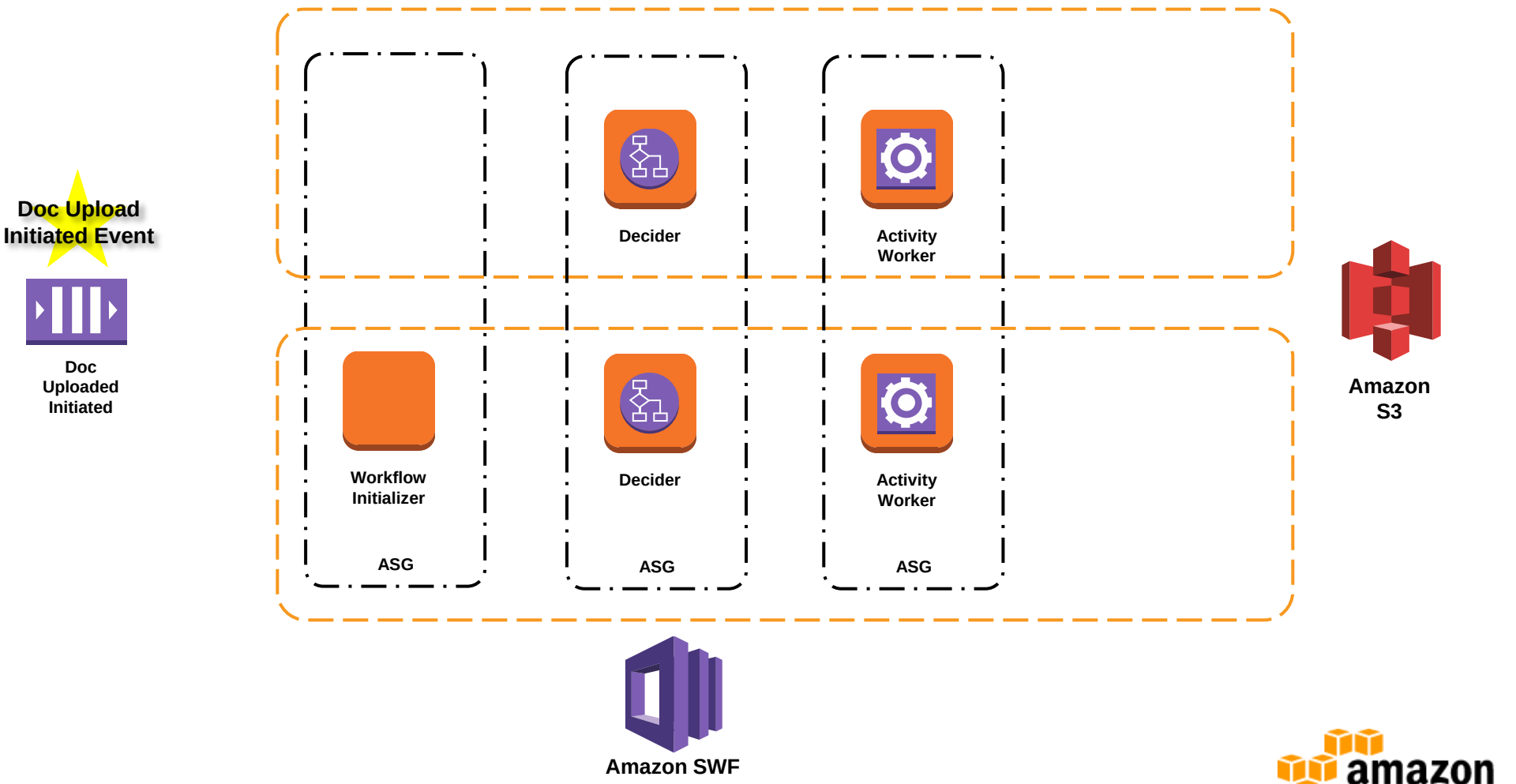


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Doc Uploaded Initiated Event: Once the document is detected in S3, a final task is scheduled in the near future to ensure that the document has been properly processed by the **Doc Upload Complete** workflow

DocStore



© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Doc Uploaded Complete Event: When a document upload has been complete, it needs to be processed

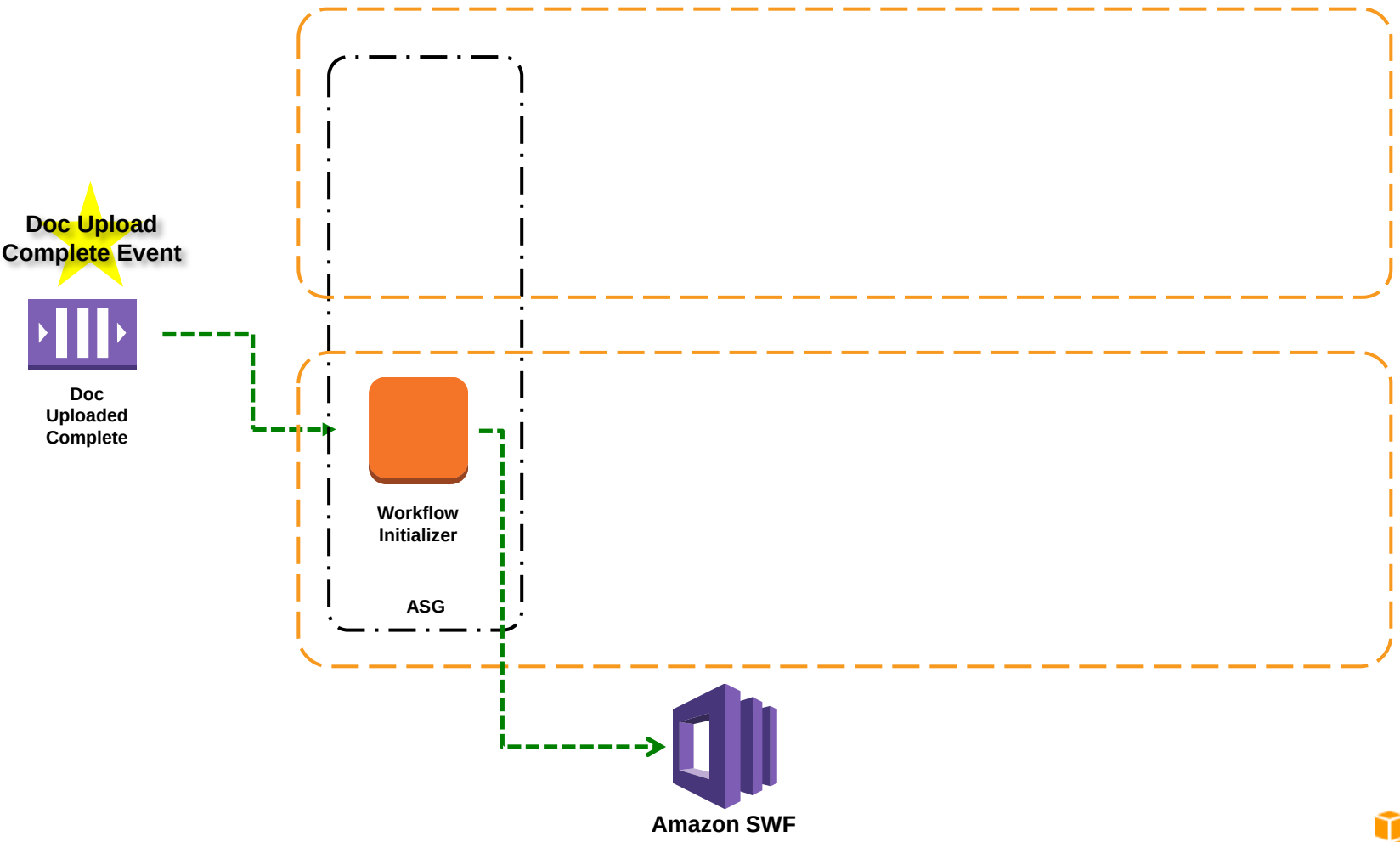
DocStore

© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Doc Uploaded Complete Event: A workflow initiaizer determines the appropriate workflow to initiate (based on customer type, document type, etc)

DocStore

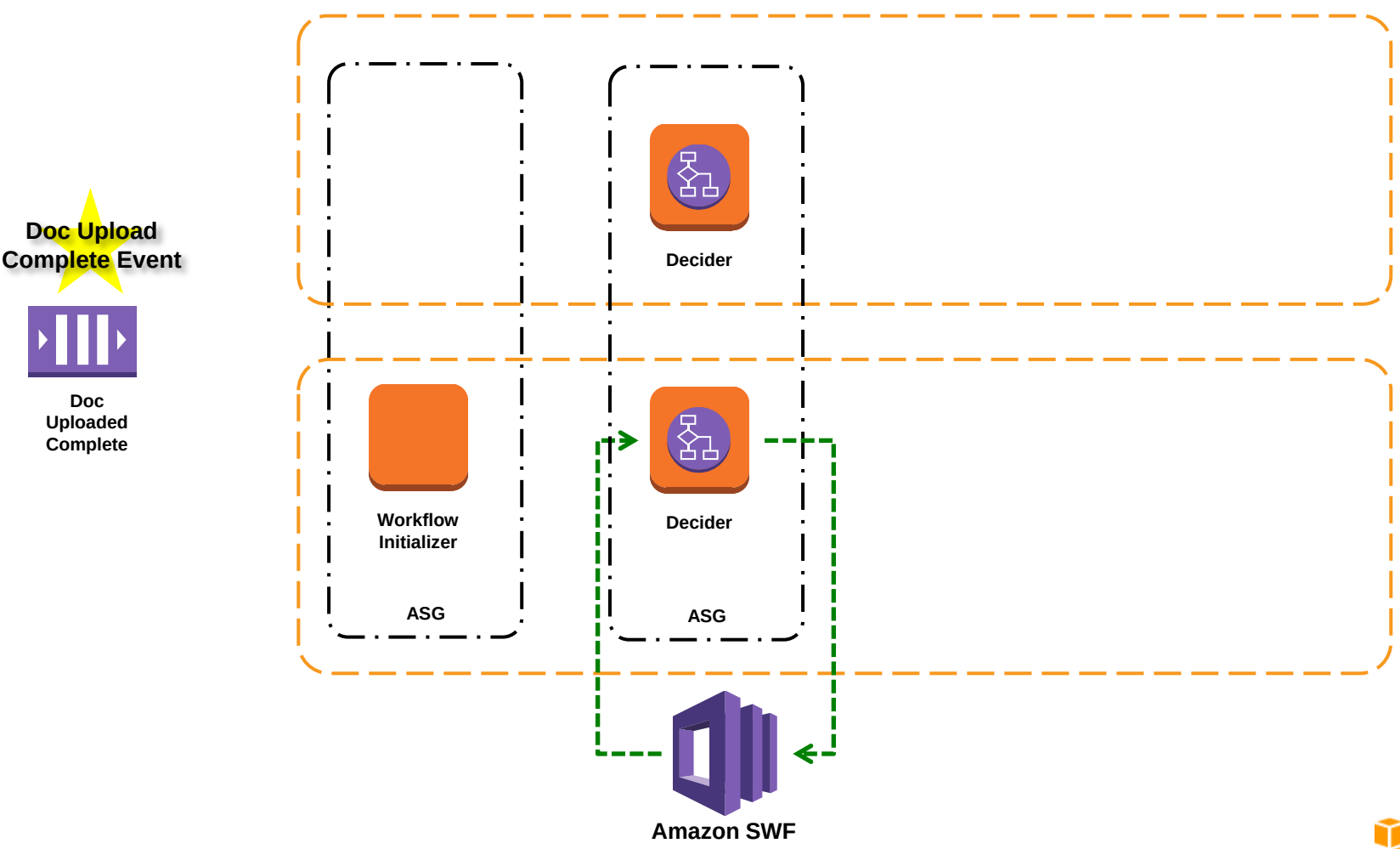


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Doc Uploaded Complete Event: A decider schedules the appropriate activity based on the document type.

DocStore

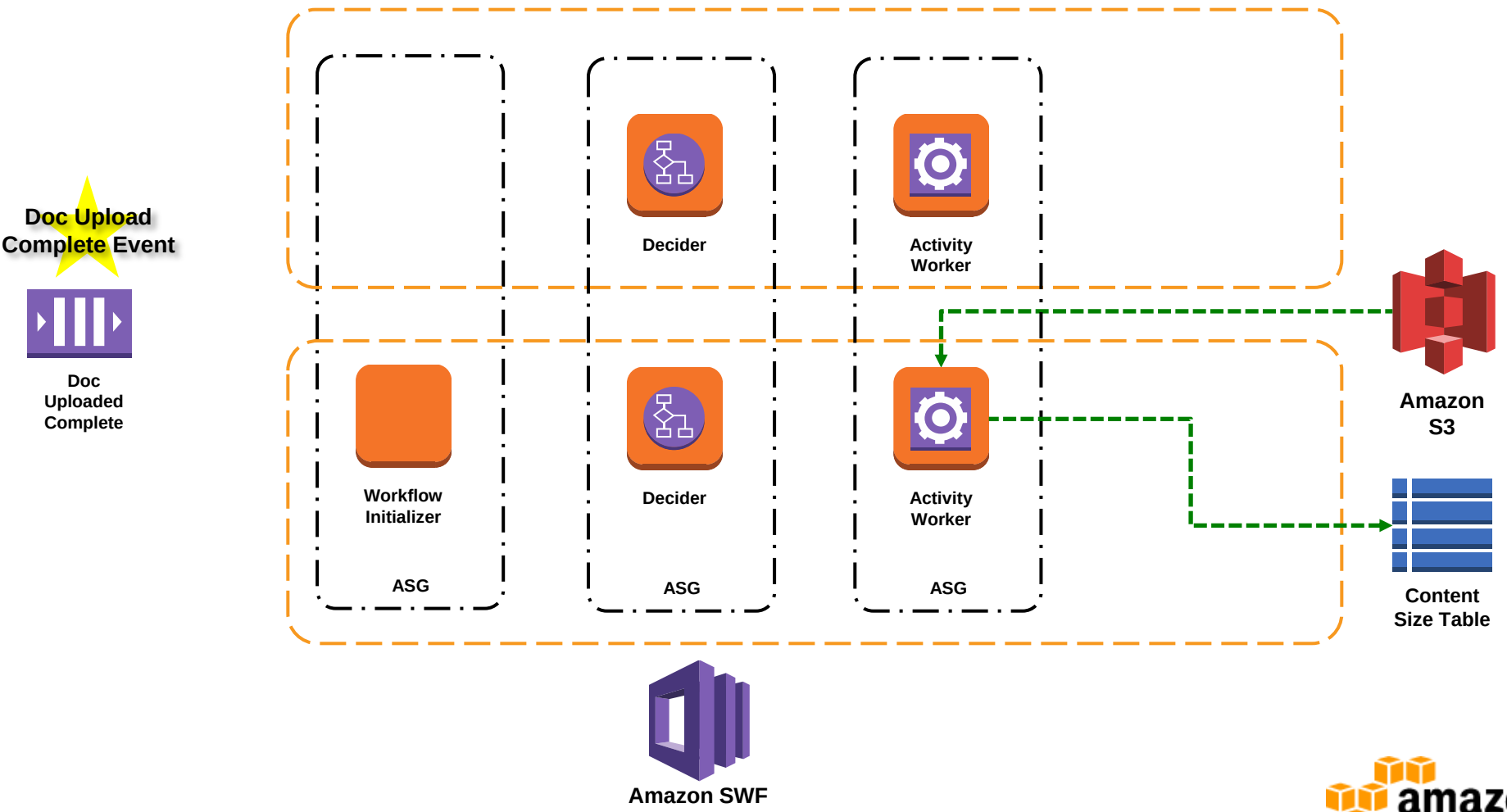


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Doc Uploaded Initiated Event: The **content-length** metadata property of every object should be retrieved from S3 and used to increment a **counter tracking the total size of all uploaded objects**

DocStore

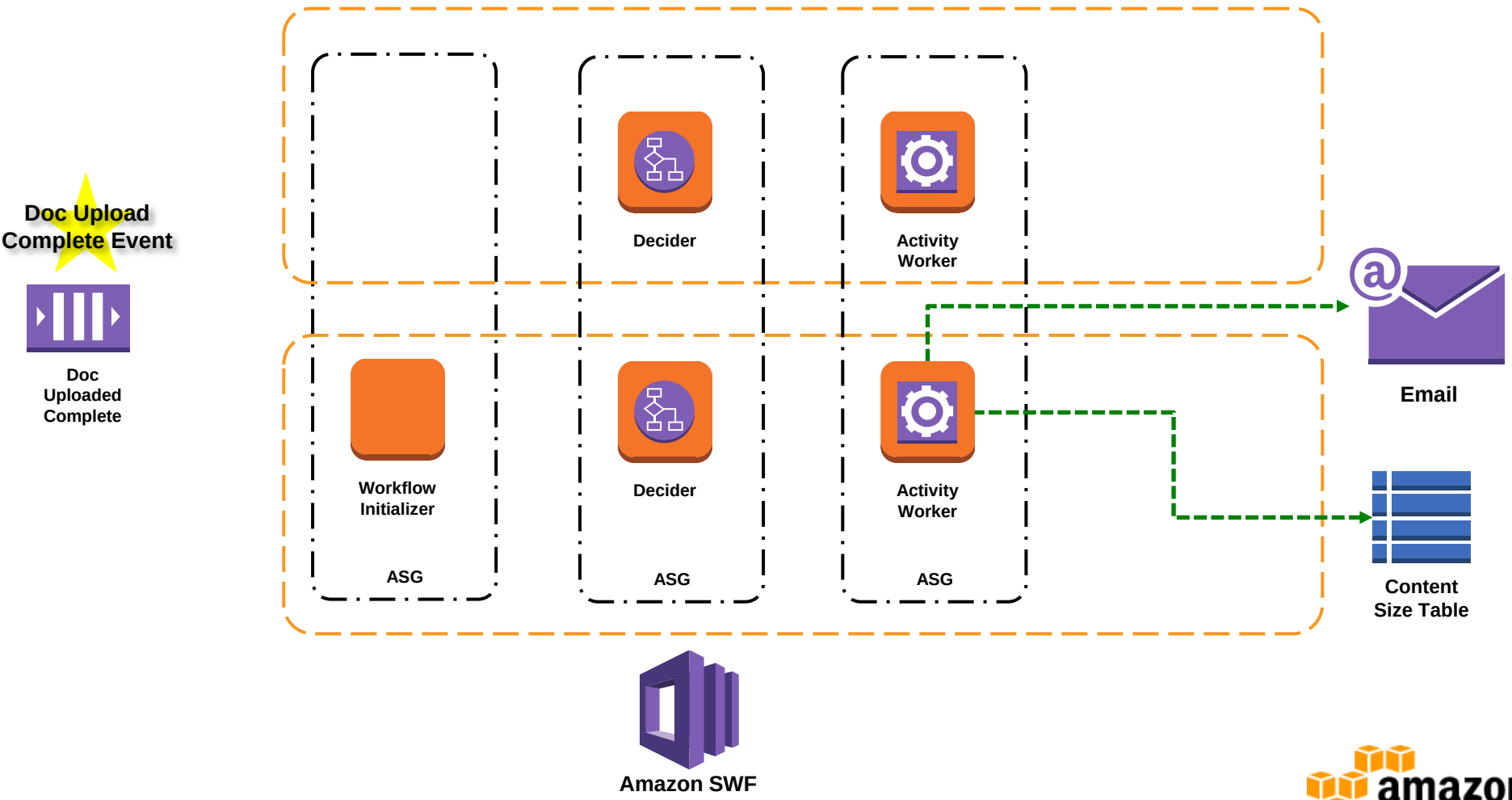


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Doc Uploaded Initiated Event: A counter tracking the total content size uploaded by each user is also incremented. If the new doc causes the user to exceed his allocated amount, a flag is set and the user will receive an email via **SES**

DocStore

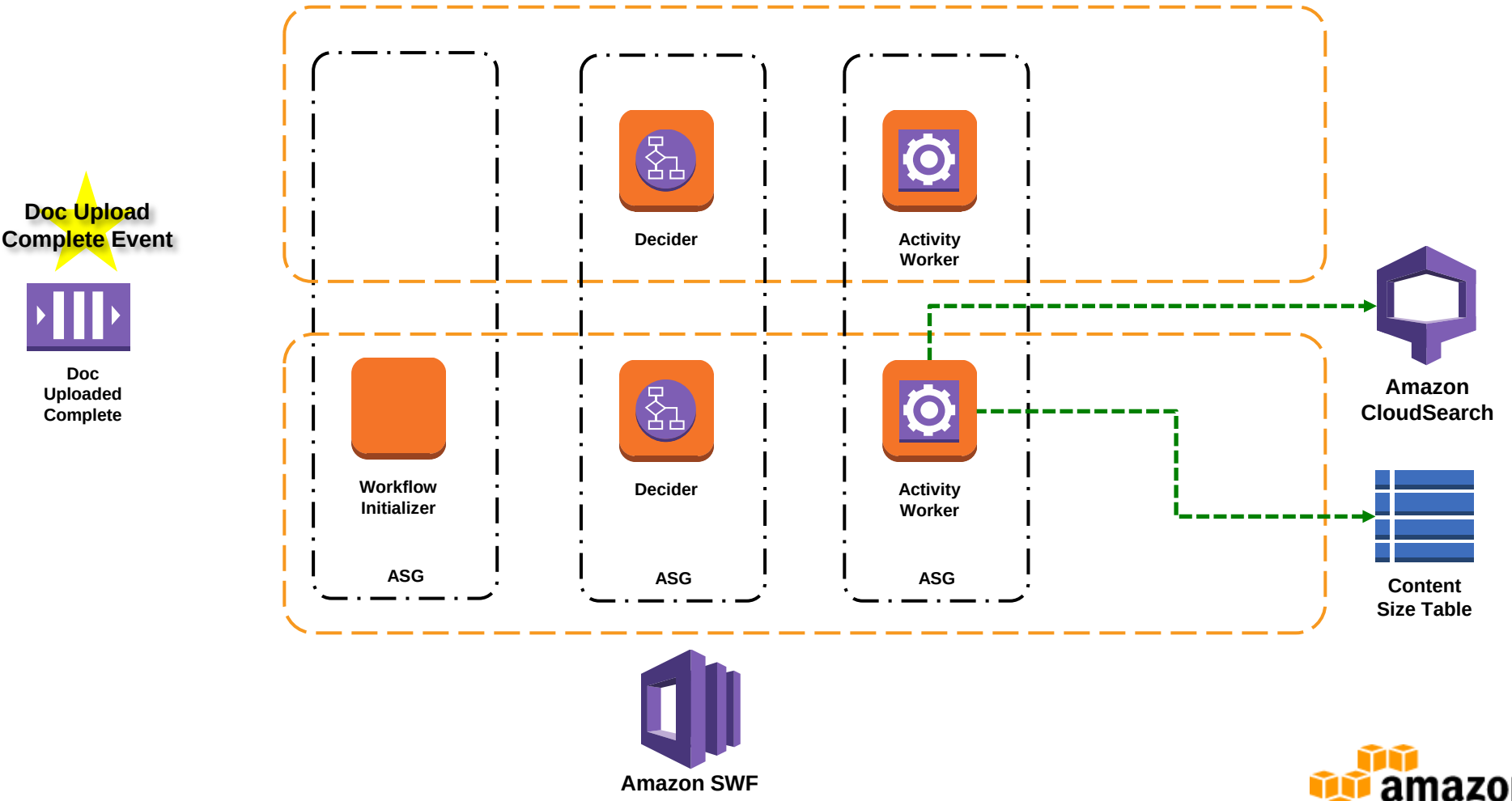


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Doc Uploaded Initiated Event: If the document was uploaded by a premium user and it has text content (e.g., PDF, DOCX, etc), the full-text content will be indexed in **CloudSearch**

DocStore

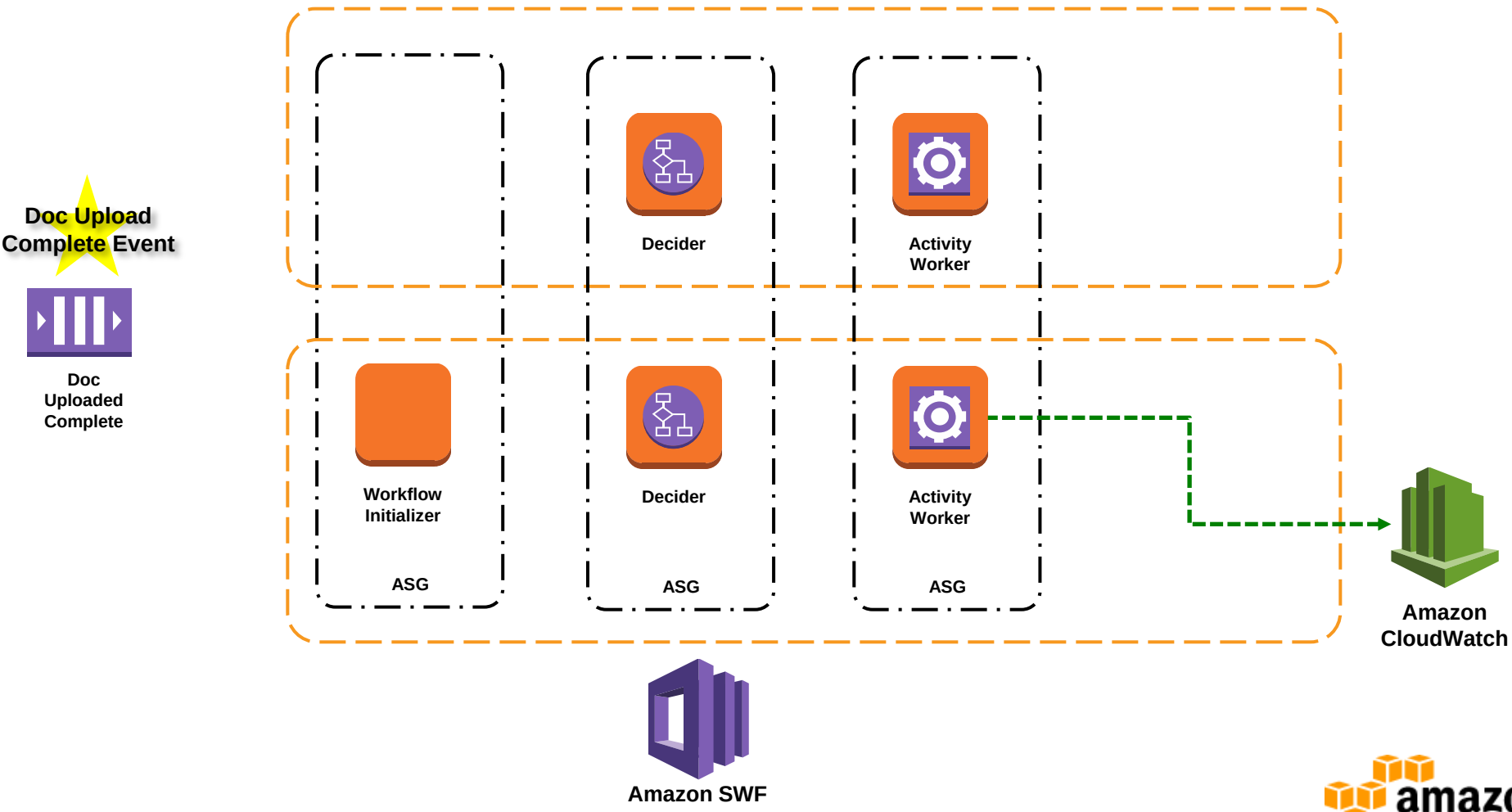


© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™

Doc Uploaded Initiated Event: Finally, several **custom metrics** are also put to **CloudWatch**, allowing us to monitor how many new documents are being uploaded by user type, how big the documents are, etc.

DocStore



© 2011 Amazon.com, Inc. and its affiliates. All rights reserved. May not be copied, modified or distributed in whole or in part without the express consent of Amazon.com, Inc.

amazon web services™