Duo Security is now part of Cisco.

Guide to Business Continuity Preparedness

Version 2.2 Published November 27, 2019

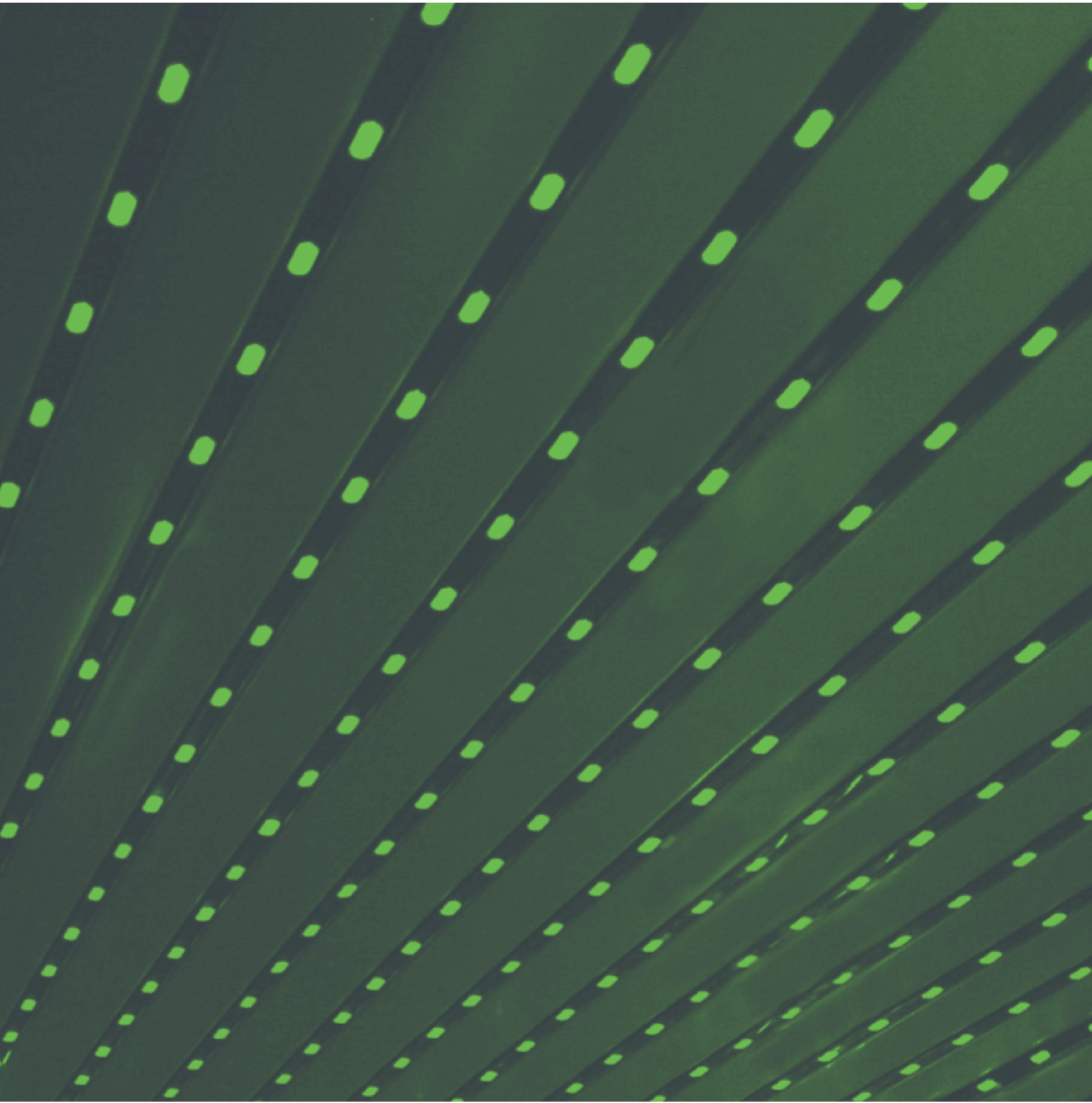

Table of Contents

Overview

Why Do I Need This Guide?

Even robustly engineered solutions can occasionally experience a disruption of service. Duo has maintained uptime of greater than 99.99% for more than four years, which still leaves a small window in which the Duo service may be unavailable.

Outages impact the productivity of your workers and have the potential to temporarily weaken your security posture. As your trusted access provider, we want you to be prepared for any situation that may arise and ensure you have a plan in place to respond to potential outages.

More details on Duo's service, and how our cloud architecture and product development processes are designed to ensure high availability, are available in our Service Reliability whitepaper.

Understanding Outages and Your Environment

Consideration should go into how your Duo-protected applications and IT organization will react if the Duo service is unavailable.

Being prepared to navigate potential service disruption scenarios will ultimately ensure a better experience with Duo.

This guide will help you:

- Understand the two categories of outages
- Understand Duo's failure modes and how to decide on Fail Safe vs. Fail Secure
- Understand how your applications will respond to different types of outages
- Message your users during an outage

Planning for Success

Once you have read this guide and understand outage scenarios and your application's failure mode behaviors, we highly recommend you create application-specific disaster recovery (DR) plans. This planning should include:

- Understanding the processes required to block or bypass Duo's cloud service if failmode behavior is not invoked as expected.
- Procedures for removing Duo from the authentication workflow for **each protected application**.

Types of Outages

Duo service disruption can generally be placed in one of two categories: when the Duo service is **unreachable** or when there is a Duo service **degradation**.

The distinction between Duo being unreachable and a service degradation is important because it affects how Duo-protected applications behave, and ultimately what your end-users experience.

Duo service
unreachable

(failmode invoked)

- Duo's service temporarily unavailable
    - Duo service fails to display the Duo Prompt or respond to Auth API requests within 180 seconds
- Network connectivity preventing customer applications from reaching Duo's service. These typically occur on the customer's local network, but can sometimes occur at a regional or global level
    - Construction company digs through fiber
    - Firewall or DNS issues
    - ISP connectivity or routing problems
- Application/integration issues
    - Duo-protected application is misconfigured

Duo service
degradation

(failmode not invoked)

- Two-factor method failure
    - Duo's telephony provider (SMS/phone call) experiences an outage of their own
    - APNS (Apple Push Notification Service) fails to deliver push notifications to users iPhones*
- Authentications not completing
    - Duo pushes bad code preventing users from successfully logging in
    - Duo authentication latency and failures increase but service remains partially available

*Users frequently forget to allow notifications on their phones—this does not qualify as a Duo service degradation. Users can launch their Duo Mobile app and will see a pending authentication request. For additional end-user troubleshooting see our iOS or Android guides.

How Do I Know When an Outage is Occurring?

If you or your users are experiencing issues that you suspect may be related to an outage, first check status.duo.com for any news about potential outages. If you believe you are experiencing an outage or have a technical issue not related to a service disruption, contact Duo Support.

Additional options for monitoring Duo's cloud service are documented in this knowledge base article.

Duo Failure Modes: Duo Service Unreachable

Failure mode (often referred to as “failmode”) is invoked when the Duo service is unreachable or a critical problem is detected. Duo has error detection mechanisms that trigger failmode based on detected errors.

To ensure the continued reachability of Duo's service, customers must allow communication to all Duo's IP ranges. A list of IP ranges with additional considerations is available in this article.

Sometimes service disruptions manifest in other ways. For example, the Duo service could be reachable, but authentication is failing for other reasons. For more on this, scroll down to Duo service degradation.

Failmode can be configured to behave in one of two ways:

1. **Fail safe** (also known as “fail open”) - if Duo service is unreachable, users will be **ALLOWED** access to Duo-protected applications if they pass primary authentication.

   - This weakens your security posture, as two-factor authentication is temporarily removed.
   - It causes less pain for users and does not interrupt workflow—employees can still log in and work.
   - For example, in an authproxy.cfg file, this will be indicated in a Server section by the following syntax: For more information on Authentication Proxy configurations, click here.

   ```
   failmode=safe
   ```

2. **Fail secure** (also known as “fail closed”) - if Duo service is unreachable, users will be **DENIED** access to Duo-protected applications *even if* they pass primary authentication.

   - This is the most secure option.
   - It can be the most disruptive option with regard to daily workflow as it denies the user access to the app.
   - For example, in an authproxy.cfg file, this will be indicated in a Server section by the following syntax:

   ```
   failmode=secure
   ```

See the Configuration Decisions section below for help on deciding whether you should enable fail safe or fail secure mode.

IMPORTANT: Not all integrations provide a mechanism to control the failmode behavior.

- Integrations that use the Authentication Proxy or Duo Access Gateway **do have the option** to specify a failmode.

- The majority of Duo developed integrations allow for failmode configuration during the installation process. For example, the failmode for Duo Authentication for Windows Logon and RDP is configurable in the installer. Most Duo application packages also offer a way to modify the failmode behavior post-installation (such as with Duo Unix and Windows Logon).

- Duo integrations created by third parties such as Thycotic, Ping Federate, and LastPass may not offer a way to control the failmode and may default to fail secure. Please consult the vendor's documentation to review any failmode capabilities.

- The WebSDK integration does not include failmode checking logic. More details in the "Understand How Your Applications Will Respond to Different Types of Outages" section.

- Azure Conditional Access will fail closed if Azure's cloud service cannot reach Duo's cloud service.

- The Duo Network Gateway (DNG) will fail closed if it cannot reach Duo's cloud service

Duo Failure Modes: Duo Service Degradation

In this situation, the Duo Authentication Proxy, Duo Access Gateway, or other Duo-protected application is able to reach the Duo service, but authentication is unable to complete. Failmode is not being invoked.

Possible Scenarios

- Authentication not completing
  - Real world example: Duo pushed code that broke authentication for customers using a legacy version of the Duo Prompt. Users were failing to login after having successfully passed primary authentication and approved the secondary authentication request.
- Failure of one or more authentication methods
  - Real world example: Duo's SMS service was not successfully delivering text messages to users, leaving them unable to authenticate.

Solutions

Some solutions may not be viable for all customers in all scenarios. For example, a firewall rule change may be more burdensome than messaging your users. Consider the solutions best for your organization. Degradation issues are typically resolved within 30 minutes. If you do employ any of the following solutions, **be sure to revert the changes following the resolution of the issue**.

- Apply an Authentication Policy to bypass two-factor authentication while the service degradation persists. Customers with paying editions of Duo can utilize Authentication Policies. In the event of the Duo service being unreachable, this solution can also be used.
  - How: Create and temporarily apply an application-level Authentication Policy.
  - What it does: Allows users to access a specific application without completing two-factor authentication. Access can be restricted or enabled based on a user's group membership.
- Inform users of the interruption and offer workarounds if Duo has posted any on status.duo.com.
  - How: Refer to the End-User Messaging Templates section below. Navigate to status.duo.com to see if Duo has identified any temporary workarounds.
  - What it does: Ensures users that your organization and Duo are both aware of the problem and are working to fix it.
- Move all or some users into a group set to "bypass" status.
  - How: Manually or bulk update users to move them to a group if they are not in one already. That group needs 1) access to the protected application and 2) to be set to "bypass" status.
  - What it does: This will bypass two-factor authentication for any user in the group.
- Revert configuration/profile on applications so as to not invoke Duo.
  - How: Consult your specific application's documentation on duo.com/docs.
  - What it does: This will remove two-factor authentication from the authentication workflow.
- For applications that use the Duo Authentication Proxy, use the Primary Only Mode feature.
  - How: This feature was introduced in Authentication Proxy version 2.14.0 and is triggered by running a command on the proxy server.
  - What it does: This temporarily (default is one hour with a four-hour maximum) skips Duo authentication for all logins to RADIUS or LDAP configurations that use the default "fail safe" behavior.

- Manually block Duo's service via a firewall rule to effectively create an unreachable outage scenario.
    - How: Block *.duo.com and *.duosecurity.com on TCP port 443.
    - What it does: This will invoke failmode. If fail safe is configured, access to the application will be granted without two-factor authentication. If fail secure is configured, access will be blocked. Monitor status.duo.com closely to know when this change can be reversed.

Understanding Your Application's Failmode Behavior

Failmode configuration options and behavior during an outage can differ depending on the application that Duo is protecting. In this section, we will examine important distinctions and details about Duo-developed applications, Duo's WebSDK, and popular third-party developed applications so you can better understand how your Duo-protected applications will be affected by an outage.

Duo-Developed Applications That Provide Failmode Control

The following table lists the Duo-developed and supported applications that provide failmode control along with additional details on how failmode may or may not be invoked in different outage scenarios. To ensure you have all features and security improvements, we always recommend updating to the latest available version.

- Duo Authentication Proxy
- Duo Access Gateway (DAG)
- Duo for Windows Logon/RDP
- Duo Unix
- AD FS 2.X
- AD FS 3/4
- OWA
- RD Web/Gateway
- Oracle Access Manager
- Shibboleth < 3.3 (Github)

Duo-Developed Applications That Do Not Provide Failmode Control

- Duo Network Gateway (DNG)
- Microsoft Azure Active Directory (Conditional Access)
- Duo Single Sign-On

WebSDK

Duo's WebSDK does not have a built-in mechanism to trigger failmode or otherwise automatically validate that the Duo service is reachable from your WebSDK application. If Duo's cloud service becomes unreachable, the WebSDK alone will not allow users to authenticate without successfully completing two-factor authentication.

It is very important to carefully program the conditions under which the application will fail safe (open) to avoid creating an unintentional 2FA bypass scenario. To monitor the service, you can use Duo's Auth API ping endpoint to implement a liveness check for the Duo service (which doesn't require any Duo integration information), and then use the Auth API check endpoint (recommended) to verify the integration information and signature. Learn more in our documentation here.

If you do develop a custom fail safe behavior, please ensure that you thoroughly test the conditions that invoke failmode behavior. As always, fail secure (closed) remains the most secure option in all scenarios.

Third-Party Developed

While Duo attempts to work with as many third parties as possible to ensure integrations follow best practices, Duo does not require third parties submit developed integrations for review or notify us. As a result, Duo is not necessarily aware of all the third-party integrations and how those integrations might provide failmode control.

Below is a list of popular third-party developed Duo integrations and if/how they support failmode:

- LastPass
  - No configurable failmode
  - Users can trust a device and then not have to perform MFA again for a period of time
  - In DR scenarios, admins need to login to LastPass and remove Duo from the authentication workflow
- 1Password
  - No configurable failmode
  - Does not require MFA for offline access or standalone vault
- Okta
  - No configurable failmode
  - In DR scenarios, admins need to login to remove Duo from the authentication workflow
- OneLogin
  - No configurable failmode
  - In DR scenarios, admins need to login to remove Duo from the authentication workflow
- Ping Federate
  - Offers configurable failmode. Documentation here.
- Shibboleth > 3.3 (Native Duo)
  - No configurable failmode
  - Supports the addition of custom auth flow code that could add failmode detection. Additional documentation here.
- CAS
  - Offers configurable failmode. Documentation here.

Configuration Decisions

Carefully consider which failmode configuration you should use for each application (if available).* Your choice will likely hinge upon:

- Policy and compliance factors
- The type of data contained within protected applications
  - Health records, financials, Personally Identifiable Information (PII), Intellectual Property (IP), etc.
- Groups of users with varying levels of access
- The need to balance security with usability

*Consult your specific application's documentation on duo.com/docs to see whether it features a configurable failmode.

With regard to failmode configurations and action plans in the event of service degradations, there are generally three main categories an organization's application can fall under:

Restriction level	Unreachable (failmode invoked)	Degradation (failmode not invoked)
Most restrictive Contract, law, policy, or sensitivity of data contained within the protected application **requires two-factor authentication, without exception**.	Fail secure	Users and Groups should NOT be switched to "bypass" status, as this will skip two-factor authentication.
Restrictive Some subsets of users are always required to pass two-factor authentication, without exception.	Fail secure	Users and Groups who without exception are required to pass two-factor authentication should NOT be switched to "bypass" status, as this will skip two-factor authentication. Users and Groups for whom it is tolerable to access this application without two-factor can be switched to "bypass," allowing them to skip two-factor authentication.
Less restrictive Contract, law, policy, or sensitivity of data contained within the protected application **does not mandate that two-factor authentication be used in all circumstances**.	Fail safe	Users and Groups for whom it is tolerable to access this application without two-factor can be switched to "bypass," allowing them to skip two-factor authentication.

*Take consideration if a Group of users has access to more than one application, as putting them in "bypass" will skip two-factor authentication for all applications they have access to. If this must be done for a Group to access another application, that Group's access to this application should first be removed.

End-User Messaging Templates

Consider at what point during an incident your organization is comfortable with messaging end-users. It could be as soon as your users begin reporting problems, it could be after Duo posts a notification on status.duo.com but before your users have reported anything, or it could be only if an incident has remained unresolved for 20 or more minutes.

Time of day

- An incident at 11 a.m. on a weekday may require immediate messaging to users.
- An incident at 11 p.m. on a weekend may not need to be messaged to users immediately.

Time of quarter

- An incident during the last week of a quarter may require immediate messaging to users, regardless of the time of day.

Criticality of access

- An incident affecting access to a critical application may require immediate messaging regardless the time, date or other factors.

Duo Service Unreachable or Degraded

If you're using fail safe when Duo is unreachable or invoking failmode manually during degradation:

SUBJECT: Authentication problems - In Progress

BODY: Duo is reporting problems with their service. As a temporary workaround, we are lifting the requirement of Duo two-factor authentication. Once the problem is resolved, two-factor authentication will be reinstated.

If you're using fail secure:

SUBJECT: Authentication problems - In Progress

BODY: We are experiencing problems with Duo two-factor authentication. Due to the nature of data contained within <your application>, access will be denied until this problem is resolved.

Authentication Method-Specific Issue

SUBJECT: Authentication problems - In Progress

BODY: Duo is reporting problems with their <push/SMS/phone> service. As a temporary workaround, please use <sms/push/phone callbacks>. Expect another update when the issue has been resolved.

Frequently Asked Questions

Can I be notified when failmode is invoked?

Failmode is configured and invoked locally in your Duo Authentication Proxy or Duo-protected application. We recommend using a monitoring tool or SIEM solution to watch for a failmode transaction.

You can determine whether failmode has been invoked by examining your Authentication Proxy's logs. The default directory for storing logs is C:\Program Files (x86)\Duo Security Authentication Proxy\log on a 64-bit Windows machine and C:\Program Files\Duo Security Authentication Proxy\log on a 32-bit Windows machine.

Below are two examples of Duo Authentication Proxy logs showing when failmode has been invoked.

1. **fail safe** log example

```
!! Auth Proxy versions 2.11.0 and later

2018-11-09 16:53:57-0500 [-] (('10.0.1.99', 54698), 146): Failmode Safe - Allowed Duo login on
preauth failure
2018-11-09 16:53:57-0500 [-] (('10.0.1.99', 54698), 146): Returning response code 2: AccessAccept
2018-11-09 16:53:57-0500 [-] (('10.0.1.99', 54698), 146): Sending response

!! Auth Proxy versions 2.10.1 and earlier

2016-03-22 16:52:22-0400 [-] (('10.2.4.121', 43031), 170): Allowed Duo login on unexpected failure
2016-03-22 16:52:22-0400 [-] (('10.2.4.121', 43031), 170): Returning response code 2: AccessAccept
2016-03-22 16:52:22-0400 [-] (('10.2.4.121', 43031), 170): Sending response
```

2. **fail secure** log example

```
!! Auth Proxy versions 2.11.0 and later

2018-11-09T16:57:51-0500 [-] (('10.0.1.99', 43674), 63): Failmode Secure - Denied Duo login on
preauth failure
2018-11-09T16:57:51-0500 [-] (('10.0.1.99', 43674), 63): Returning response code 3: AccessReject
2018-11-09T16:57:51-0500 [-] (('10.0.1.99', 43674), 63): Sending response

!! Auth Proxy versions 2.10.1 and earlier

2016-03-22 16:54:38-0400 [-] (('10.2.4.121', 53573), 109): Denied Duo login on unexpected failure
2016-03-22 16:54:38-0400 [-] (('10.2.4.121', 53573), 109): Returning response code 3: AccessReject
2016-03-22 16:54:38-0400 [-] (('10.2.4.121', 53573), 109): Sending response
```

Below is output from a SIEM-friendly authevents.log:

1. **fail safe** log example

```
!! Auth Proxy versions 2.11.0 and later

{"username": "testuser1", "status": "Allow", "client_ip": null, "server_section":
"radius_server_auto", "timestamp": "2018-11-09T21:53:57.950000Z", "auth_stage": "Secondary
authentication", "factor": null, "msg": "Failmode Safe - Allowed Duo login on preauth failure"}

!! Auth Proxy versions 2.10.1 and earlier

{"username": "testuser1", "status": "Allow", "client_ip": null, "server_section":
"radius_server_auto", "msg": "Allowed Duo login on unexpected failure", "timestamp":
"2018-04-17T21:39:13.416000Z", "auth_stage": "Secondary authentication"}
```

2. **fail secure** log example

```
!! Auth Proxy versions 2.11.0 and later

{"username": "testuser1", "status": "Reject", "client_ip": null, "server_section":
"radius_server_auto", "timestamp": "2018-11-09T21:57:51.326000Z", "auth_stage": "Secondary
authentication", "factor": null, "msg": "Failmode Secure - Denied Duo login on preauth failure"}

!! Auth Proxy versions 2.10.1 and earlier

{"username": "testuser1", "status": "Reject", "client_ip": null, "server_section":
"radius_server_auto", "msg": "Denied Duo login on unexpected failure", "timestamp":
"2018-04-17T21:38:11.822000Z", "auth_stage": "Secondary authentication"}
```

Are Duo deployments built with high availability or designed as “active/active?”

Yes, Duo deployments should not be considered a single node. Duo utilizes numerous independent clusters or “deployments” of its technology to scale to support customer growth, as well as ensure that the impact of a failure in any one deployment is minimized. The underlying infrastructure components of each of these deployments are backed by real time replication between the multiple physical data centers that make up Amazon Web Services (AWS) availability zones. Duo also replicates customer data in real time to at least one additional AWS region for each individual deployment.

How does Duo protect deployments against DDoS attacks?

AWS takes steps to transparently mitigate DDoS attacks against Duo's infrastructure utilizing their proprietary AWS Shield DDoS prevention technology. In the event of an attack against Duo services that is not automatically mitigated by AWS or Duo's own hardened infrastructure, Duo personnel would be alerted to the issue immediately and respond as necessary. This response could include the systematic black-holing of specific IP addresses/netblocks, or even the relocation of customer traffic to unaffected infrastructure and/or IP addresses.

Can my account(s) be moved over to another deployment in the event of an outage?

The technology used to move customers between deployments is designed to minimize impact on both the source and destination deployment, and can often be a relatively lengthy process as it runs in the background. This process is not suitable to be exercised as part of a failure scenario. Additionally, in some cases, moving customer workloads to alternate infrastructure may not resolve the underlying issue and could even increase the impact of an outage. In some outage scenarios, moving impacted customer accounts to another deployment can bring the underlying issue along for the ride, potentially extending the timeline of the outage.

We suffered an outage. Can we be moved to another deployment?

All Duo deployments are created equal and share the same high availability properties and historical track record for excellent uptime. For these reasons, moving off of your current deployment does not inherently reduce the risk of an outage.