Google Apps - Gmail
Incident Report
February 24, 2009
Prepared for Google Apps Premier Edition Customers

Summary

Between approximately 9AM to 12PM GMT / 1AM to 4AM PST on Tuesday, February 24, 2009, some Google Apps Gmail users were unable to access their accounts. The actual outage period varied by user because the recovery process was executed in stages. No data was lost during this time. The root cause of the problem was a software bug that caused an unexpected service disruption during the course of a routine maintenance event. The root cause of this unexpected service disruption has been found and fixed.

Additional Details

A few months ago, new software was implemented to optimize data center functionality to make more efficient use of Google's computing resources, as well as to achieve faster system performance for users.

Google's software is designed to allow maintenance work to be done in data centers without affecting users. User traffic that could potentially be impacted by a maintenance event is directed towards another instance of the service. On Tuesday, February 24, 2009, an unexpected service disruption occurred during a routine maintenance event in a data center. In this particular case, users were directed towards an alternate data center in preparation for the maintenance tasks, but the new software that optimizes the location of user data had the unexpected side effect of triggering a latent bug in the Gmail code. The bug caused the destination data center to become overloaded when users were directed to it, and which in turn caused multiple downstream overload conditions as user traffic was automatically shifted in response to the failures. Google engineers acted quickly to re-balance load across data centers to restore users' access. This process took some time to complete.

Improvement Actions

We received thoughtful feedback from customers, partners, industry analysts, and our own employees both during and after this outage. Below is a summary of the feedback and the actions that we're taking to make things better:

1. Given the risks associated with maintenance events, we understand that it's a traditional IT practice to limit maintenance events to weekends and evenings. This being said, Google's large distributed global infrastructure makes it impossible to mimic this traditional model because complex maintenance events cannot be completed to fit every user's off-hours. Our goal is therefore to innovate on the technology and process fronts to make our systems as self-healing and self-managing as possible. We feel that we run a very reliable system, but we also believe that there's always room for improvement. To that end, Google engineers work around the clock to make our production systems better.
2. It's critical to proactively communicate with customers when outages occur. We understand that we need to provide information quickly during an outage. On this front, we are launching a Google Apps status dashboard very soon. This dashboard will provide information both during and after an outage. During an outage, we will quickly acknowledge the problem, provide a best estimate of when service will be restored, and offer useful workarounds as available. After an outage, we will post an incident report after the issue is resolved. We will also respond to special requests to participate in internal post-mortem calls with large customers.
3. It's critical to prevent long outages. We understand that our customers rely on our products to run their businesses and outages are very disruptive. We see the effect of outages first-hand because we run our own business on Google Apps. In this particular case, though the total duration of the outage was 3 hours, the actual outage was less for most of the users because our systems are designed to enable recovery to take place in stages.

   Google engineers take system outages very seriously. This commitment is demonstrated in our drive to build resiliency into everything that we develop. Despite this commitment, we're not perfect, and we don't always get it right the first time. Please rest assured that we monitor our systems 24 x 7, we have engineers available to analyze and resolve production issues 24 x 7, we are staffed to respond quickly to problems, and we develop ongoing improvements to our systems to proactively make them better and to prevent recurrence of problems.

We are very sorry for the inconvenience that this incident has caused. We understand that system problems are inconvenient and frustrating for customers who have come to rely on our products to do many different things. One of Google's core values is to focus on the user, so we are working very hard to make improvements to our technology and operational processes so as to prevent service disruptions. We are confident that we will achieve continuous improvements quickly and persistently.

Once again, we apologize for the impact that this incident has caused. Thank you very much for your continued support.