News, tips and tricks from Google's Gmail team and friends.

More on today's Gmail issue

September 1, 2009

Posted by Ben Treynor, VP Engineering and Site Reliability Czar

Gmail's web interface had a widespread outage earlier today, lasting about 100 minutes. We know how many people rely on Gmail for personal and professional communications, and we take it very seriously when there's a problem with the service. Thus, right up front, I'd like to apologize to all of you — today's outage was a Big Deal, and we're treating it as such. We've already thoroughly investigated what happened, and we're currently compiling a list of things we intend to fix or improve as a result of the investigation.

Here's what happened: This morning (Pacific Time) we took a small fraction of Gmail's servers offline to perform routine upgrades. This isn't in itself a problem — we do this all the time, and Gmail's web interface runs in many locations and just sends traffic to other locations when one is offline.

However, as we now know, we had slightly underestimated the load which some recent changes (ironically, some designed to improve service availability) placed on the request routers — servers which direct web queries to the appropriate Gmail server for response. At about 12:30 pm Pacific a few of the request routers became overloaded and in effect told the rest of the system "stop sending us traffic, we're too slow!". This transferred the load onto the remaining request routers, causing a few more of them to also become overloaded, and within minutes nearly all of the request routers were overloaded. As a result, people couldn't access Gmail via the web interface because their requests couldn't be routed to a Gmail server. IMAP/POP access and mail processing continued to work normally because these requests don't use the same routers.

The Gmail engineering team was alerted to the failures within seconds (we take monitoring very seriously). After establishing that the core problem was insufficient available capacity, the team brought a LOT of additional request routers online (flexible capacity is one of the advantages of Google's architecture), distributed the traffic across the request routers, and the Gmail web interface came back online.

What's next: We've turned our full attention to helping ensure this kind of event doesn't happen again. Some of the actions are straightforward and are already done — for example, increasing request router capacity well beyond peak demand to provide headroom. Some of the actions are more subtle — for example, we have concluded that request routers don't have sufficient failure isolation (i.e. if there's a problem in one datacenter, it shouldn't affect servers in another datacenter) and do not degrade gracefully (e.g. if many request routers are overloaded simultaneously, they all should just get slower instead of refusing to accept traffic and shifting their load). We'll be hard at work over the next few weeks implementing

these and other Gmail reliability improvements — Gmail remains more than 99.9% available to all users, and we're committed to keeping events like today's notable for their rarity.

Labels: Google Apps Blog