

Traditional Data Problems Graphic Text Description

- Your back end consists of an RDBMS with a table of that schema and web server.
- Whenever someone loads a web page being tracked by your application, the web page pings your web server.
- Web server increments the corresponding row in the database.
- The web analytics product is a huge success, and traffic to your application is growing.
- You start getting lots of emails from your monitoring system: "Timeout error on inserting to the database."
- Problem is obvious: the database can't keep up with the load, so write requests to increment pageviews are timing out.
- You realize that it's wasteful to only perform a single increment at a time to the database.
- The DB updates are more efficient if you batch many increments in a single request.
- Instead of having the web server hit the database directly, you insert a queue between the web server and the database.
- Whenever you receive a new pageview, that event is added to the queue.
- You then create a worker process that reads 100 events at a time off the queue, and batches them into a single database update.