



Case Study



InBoxer Accelerates Critical Analysis with the Digipede Network

Better Rules Make Better Filters

Email is a critical tool for internal and external communication for every modern business. As spam has grown from a trickle to a flood, it has become a serious drain on productivity. Email filters must improve continuously to keep ahead of increasingly sophisticated spammers.

The Company

InBoxer develops award-winning email filters for individuals and enterprises. Its business depends on classifying email efficiently and accurately, according to each user's preferences. Applications include filtering "spam" from inbound email, and detecting inappropriate or confidential content in outbound email. InBoxer employs sophisticated statistics to validate classification algorithms that can distinguish between different kinds of email. These algorithms get better as they are exposed to more email.

The Challenge

To provide the best possible "out of the box" experience for its clients, InBoxer trains its algorithms on a library of more than a million email messages. By evaluating each message based on many different attributes, InBoxer provides remarkably accurate classification—and that accuracy improves with further use.

Scoring millions of email messages on dozens of attributes is, however, extremely compute- and data-intensive. Jobs that test new settings of InBoxer's filters were running all night on a single server, delaying results analysis and limiting the productivity of InBoxer's research team. "We were at a point where we were paying people to wait for our server to finish these jobs," said Sean True, CTO of InBoxer. "That had to change."

InBoxer began exploring alternatives for increasing the speed and scalability of its algorithms, while maintaining the progress of its other development projects. These criteria ruled out several grid computing applications, which require dedicated, full-time staff to implement. "We needed a solution that could scale with us, and improve performance as we continue to grow," said True. "At the same time, we couldn't afford to reallocate our technical resources to install and maintain a complex solution."

The Solution: More Power, Scotty!

InBoxer's choice was the Digipede Network, which is designed to provide dramatically increased performance for compute-intensive, data-intensive, and transaction-intensive Windows applications. By harnessing the power of both dedicated and shared Windows resources, the Digipede Network can provide order-of-magnitude performance increases without extensive application redesign.

Scaling out InBoxer's classification research applications to take advantage of existing computers seemed like a logical approach, but this

"The Digipede Workbench made getting jobs distributed across our network a snap!"

— Sean True, CTO of InBoxer



Case Study



exercise would be pointless if it took weeks or months of coding by the company's in-house team. By design, the Digipede Network minimizes the modifications necessary for effective distributed execution of a broad variety of application types, and modifying InBoxer's code took very little time: just a few days, including learning the fundamentals of grid computing.

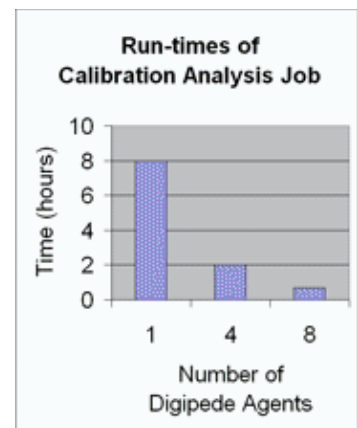
InBoxer's decision to use the Digipede Network did not require a long implementation project, extensive retraining, or expensive consultants. The Digipede Network is radically easier to buy, install, learn, and use than competing grid solutions. No Digipede staff visited InBoxer's office during the implementation; even using beta versions of the software, InBoxer staff had the system up and running quickly with a bit of telephone and email support.

The Results: From Overnight to Over Coffee

Installing the [Digipede Network Team Edition](#) at InBoxer had an immediate positive impact. With a single Digipede Server and four Digipede Agents running on existing Windows XP workstations (licensed from Digipede for less than a thousand dollars), InBoxer saw a four-fold improvement in application performance.

Overnight jobs were now running in less than two hours. Sean True said, "We managed this improvement without learning new scripting languages, or moving datasets around manually; the Digipede Workbench made getting jobs distributed across our network a snap."

True was so impressed by the improvement in speed that he went shopping for several new Windows servers, and licensed additional Digipede Agents to reduce run-time even further. Complex statistical analysis jobs now run in under 40 minutes—a duration that could be cut further, simply by adding more commodity hardware and Digipede Agents.



While the numbers are impressive, the biggest impact is really on human productivity. When you reduce the turnaround time for jobs by an order of magnitude, people work differently and progress comes faster. With so much processing power on tap, InBoxer also expanded its analysis. True added, "We are now running jobs that are far larger than any we were ever able to complete on a single machine, and making a better product as a result."

For years, grid computing has been deployed only in organizations with large IT resources. While a few large organizations benefit, this important technology has been out of reach for small and medium businesses and departments. InBoxer's experience shows that the Digipede Network brings the benefits to grid computing to a far wider audience.