NTE Commerce Performance Assessment

NTE Commerce Architecture

NTE Commerce is based on top of three basic design principles: the architecture is component-based, the code is platform independent, and there are no third-party licensing fees. This allows for a grid-style implementation where any single part of the processing could be separated off onto its own hardware (or even to multiple machines) if needed in order to maintain acceptable levels of performance.

Component-Based Architecture

All of the major operations of the NTE Commerce system are broken into a set of self-contained components. For example, inbound communications are handled by three major components: inbound file handling (pulls files from different inbound communication sources, such as FTP or AS2, and delivers them to the translation process), inbound file translation, and business rule processing (processes the translated XML document into the database). Outbound processing has similar components (Event handling, translation and delivery). Lastly, there are the physical communications servers - FTP and AS2 – and the enterprise-wide database. Any one of these components could be split off onto its own physical hardware if it started to use up a prohibitive amount of the machine's processing resources.

Platform Independence

All of the components have been built or acquired with platform-independence as a requirement. This allows NTE to be unrestricted in its choices of hardware and operating system platforms. If a particular component, for example one of the communications servers, was found to run more efficiently on a particular operating system, it could be moved over while leaving the other components alone. Going forward, NTE will not need to be concerned with portability when considering choices between Windows, UNIX or Linux platforms.

No Third-Party Licensing

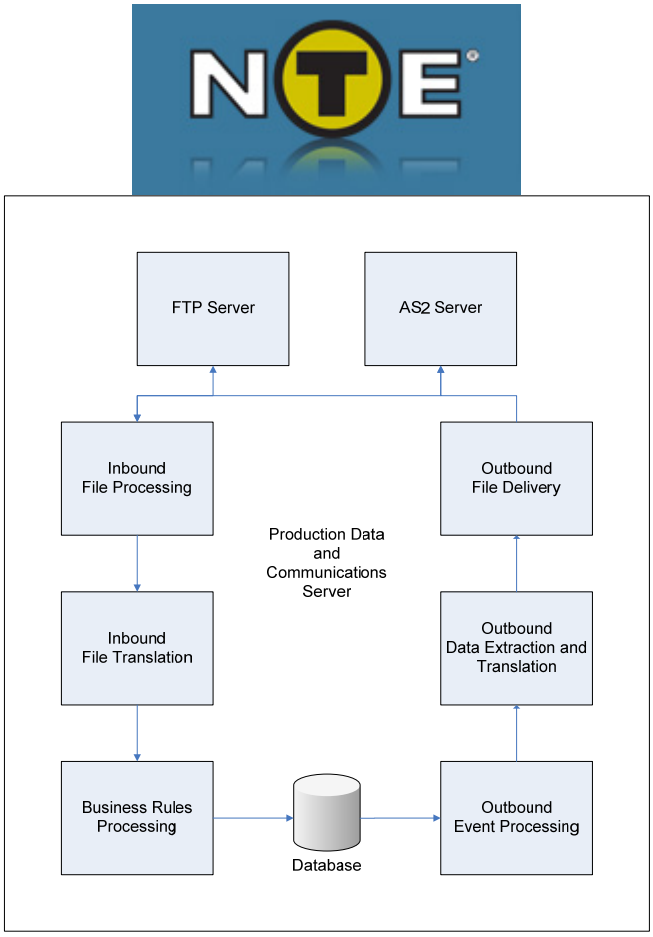
With the exception of the central database, for which there is already an enterprise-wide license, there are no licensing concerns for any of the NTE Commerce components. So, if the AS2 server eventually became over-run, causing business partners to fail in acquiring connections, a second server could immediately be added in for no cost to run on a second communications port. Similarly, any number of translation nodes, the most processing-hungry of the components, could be created across the grid.

Performance Migration Example

The following is an example of how NTE Commerce could evolve as communications volume continues to increase over time.

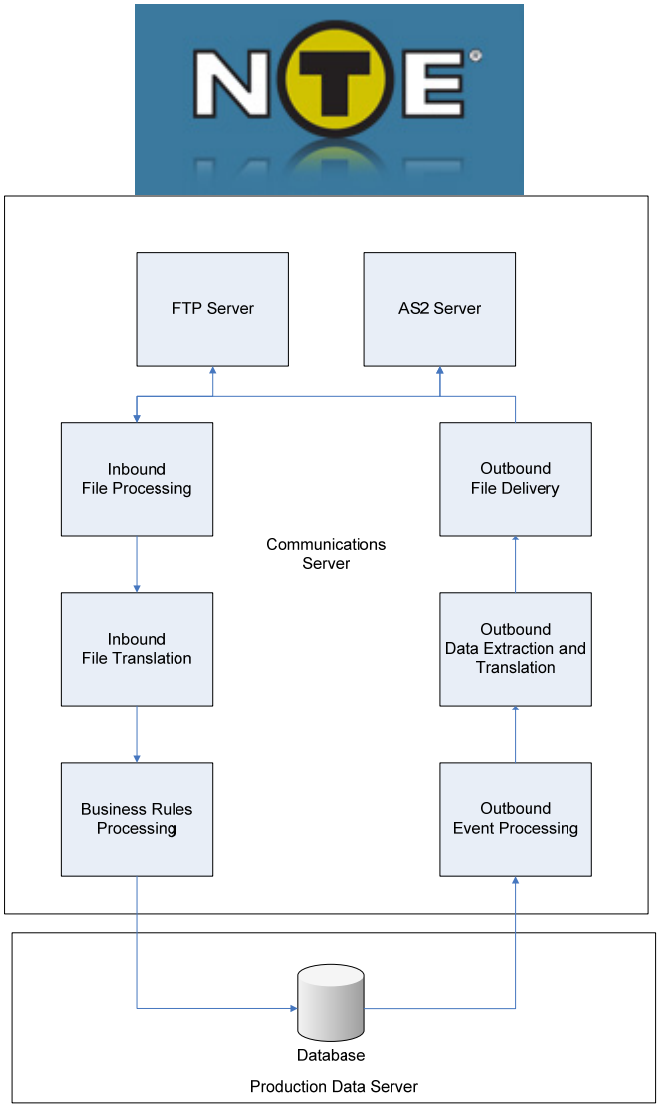
Current Implementation

In the current implementation, all of the components are still on a single, central server along with the enterprise database. Performance has not yet hit a point that would require any of the more complicated implementations. The layout of the system is simply as follows:



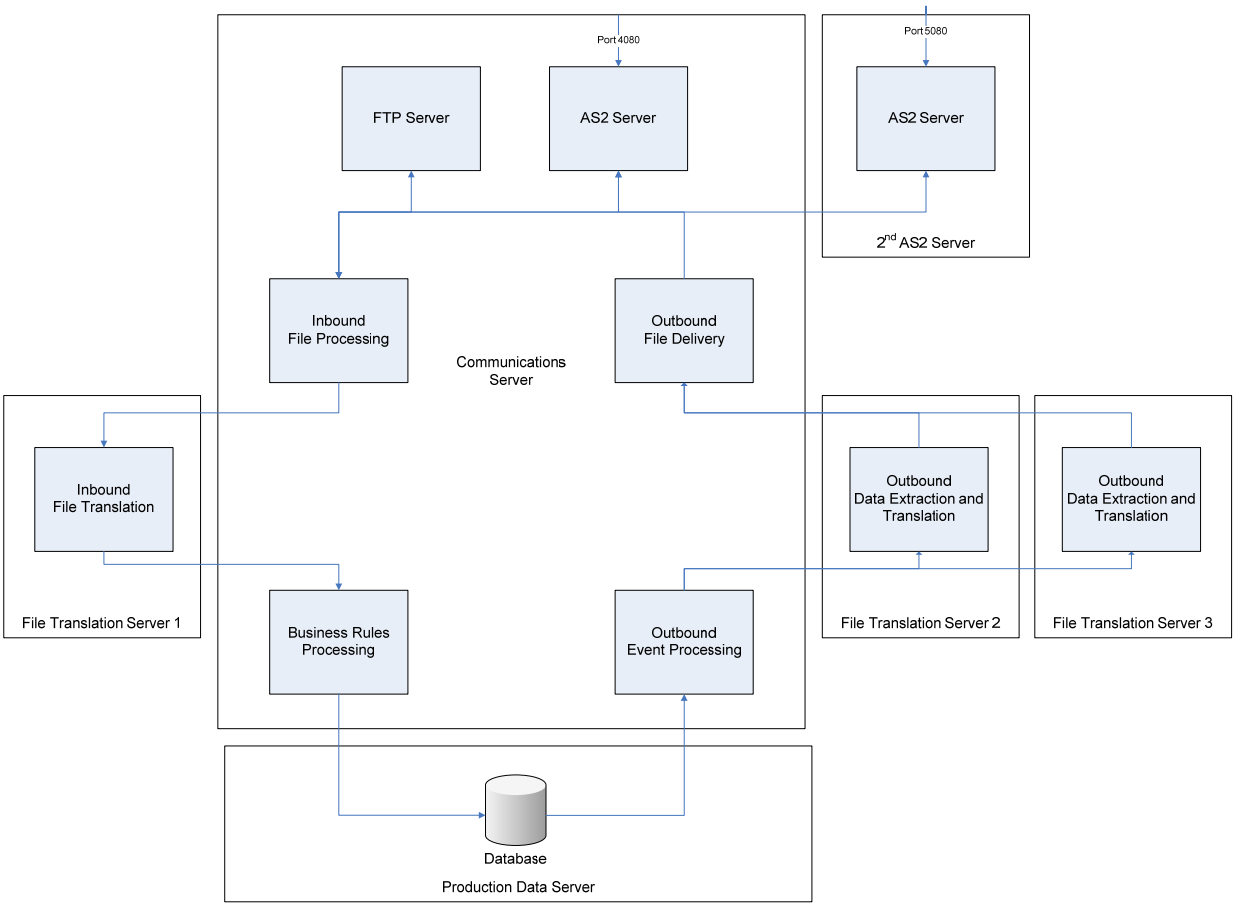
Initial Migration: Separate From the Central Database

The single largest user of resources on the server is the central database. The database not only handles communications, but all NTE transactions including web traffic. If CPU utilization hits an unacceptable point, then communications could just move off to its own machine, as so:



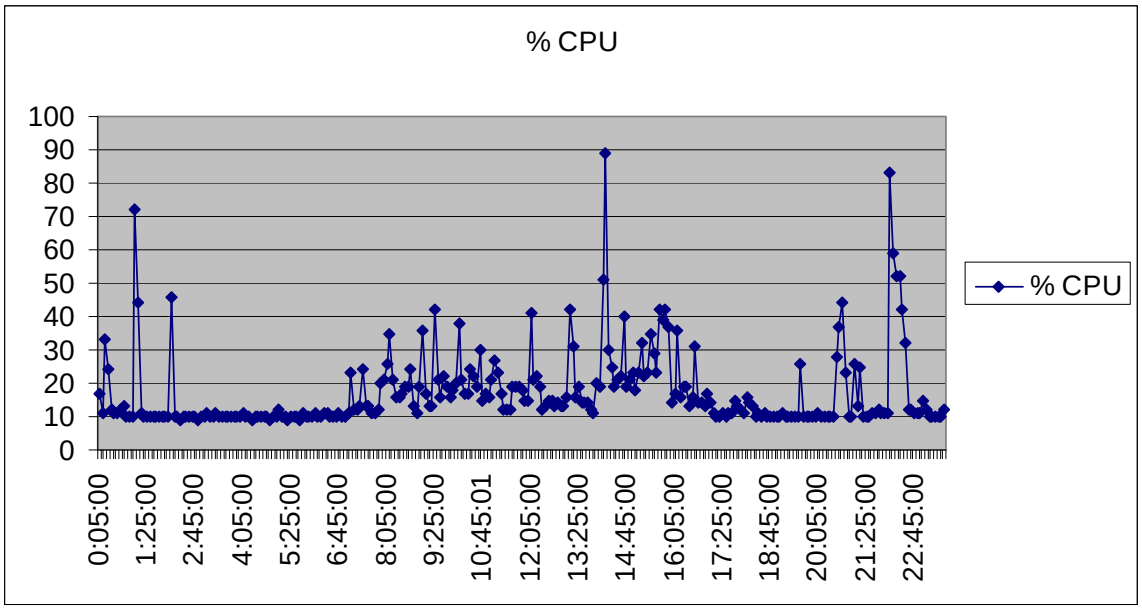
Full Grid-Based Implementation

If volume continued to climb to a point where CPU utilization was being maximized on even the communications server, a full assessment of the load for each component would be done to find out how much power would be needed to run each area. Then, the components would be assigned to the amount of hardware needed to maintain an acceptable level of performance. The following is a theoretical large-scale implementation:



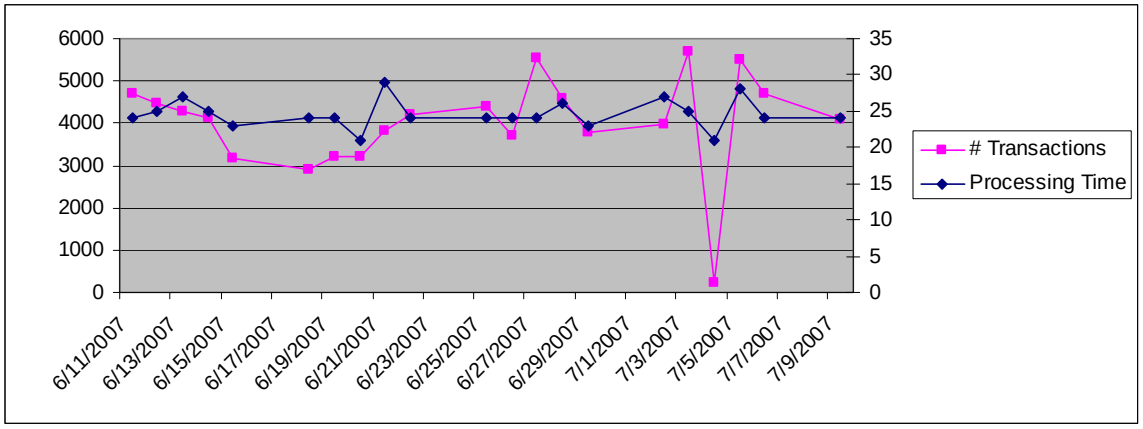
Current Performance Situation

As mentioned, there has been no need to separate any of the components onto its own hardware. The following is a graph of CPU utilization on a typical day:



The spikes at the beginning and ending of the day are all due to daily/weekly background jobs (database and file backups, archiving, etc.). The one large spike in the middle of the day was due to the receipt of several hundred EDI 204 shipment tenders and several thousand EDI 754 routing instructions from a number of different business partners all at the same time. All of these were processed within a few minutes, and the CPU still did not hit the maximum.

Similarly, processing times for individual transactions have not reflected any physical resource limitations. Below is a graph of the average time (across a 24-hour period) that it takes for an outbound communication to be sent from the time that the triggering event happens.



Most of the times are in the 20 to 30 second range (since polling intervals for translation and delivery are 10 seconds, the actual processing time is about 10 to 20 seconds). Processing time does not have a high correlation with transaction volume. These times have held up at the height of the busy season, during which several thousand outbound and several thousand inbound transactions could be sent/received in less than 30 minutes.

Looking Ahead

NTE keeps an eye on CPU utilization and average processing times on a regular basis. As performance limitations near, the NTE Commerce architecture is prepared to adapt in advance of any performance problem to ensure that an acceptable level of performance is maintained into the future.