A Cloud-based Solution to Today's Wireless Connectivity Challenges

By Connect, Inc.

Contents

Background

When the real-time use of hand held mobile devices (scanners, printers, and mobile computers) was first introduced into large scale data-transfer over RF networks, it was quickly learned that wireless network bandwidth and the processing capacity of the mobile devices were limiting factors on the functionality of the devices as well as the number of devices that could be added to a network. The nature of wireless also introduced stability problems not seen in wired networks. These challenges meant, overall, that the mobile device and wireless network could not achieve their full potential of delivering the productivity and accuracy enhancements to the companies that implemented them.

The solution at that time was to introduce a network control unit, which was a coupled hardware and software solution that sat in the middle of the Spectrum One or WLAN network, between the host application and the mobile computers, to manage the network sessions and traffic. By reducing and directing data traffic on the network, the control unit technology optimized the performance and stability of the mobile session and, therefore, the mobile computer. Connect Inc., a start-up company at the time, was the company to deliver this solution and patent the technology that made it all work. The impact of this technology in the market place was the rapid adoption of wireless mobile computers for industry.

The first adoption of the network control units, which later evolved to middleware server software, was in the early 1990s. In the subsequent ten years, Connect's technology was the prevalent wireless connectivity solution used by global distribution centers, warehouses, retail centers, etc. Gradually, as wireless bandwidth increased and mobile computers added more memory and processing power, the server technology became optional, particularly for lightly-loaded networks.

Today, many large scale enterprise-wide users of mobile computers are seeing diminishing returns on their mobility investments. Increasing use of web-based applications for industry, the growing usage of multiple platforms in a single enterprise application, and the increased competition of a global recession, has again called for an innovative connectivity solution.

This paper outlines the specific challenges brought on by the new mobility environment and describes Connect's innovative solutions that ultimately restore wireless productivity, protect the value of past mobility investments and enable the adoption of new technologies and mobile trends.

Challenges in the Mobile Enterprise

Performance

While some mobile applications are unaffected by response time latency, many are subject to very specific latency thresholds. A fork-lift driver unloading a truck can tolerate several seconds of latency after scanning the pallet before the destination must be known, due to the time required to physically move out of the trailer and onto the receiving dock. By contrast, the tolerance for latency by a stock picker filling orders on a conveyor belt is measured in milliseconds. Failure to meet the minimum threshold is not just a productivity issue; it can make the application unusable.

In the transition to new applications, response time challenges tend to increase. Web-based applications, for example, require multiple, high volume network exchanges to complete one logical transaction, introducing latency and heavy traffic congestion. Even older applications often generate high volumes of segmented data exchanges in the course of completing a single transaction.

The wireless environment presents challenges that are not present in the wired environment. To increase bandwidth for wired networks additional hardware can be added with relatively little expense. This is a much more difficult and expensive undertaking with wireless, especially when, as is often the case, the area of coverage is small. Wireless is also subject to interference that results in a greater frequency of retransmissions, creating even more competition for limited bandwidth.

Maximum wireless performance dictates the smallest possible transaction size and the fewest number of network transactions per application transaction. To some extent, applications can be developed to supply these characteristics. In practice, however, developers have limited control over network performance. For decades, application development tools have been migrating to increasingly higher levels of abstraction, under the reasonable assumption that network technology would compensate for inefficiencies at the lower levels.

In the typical high speed wireless network, the volume of data for a small number of devices is trivial. However, the latency incurred with each message exchange accumulates rapidly. In the case of web-based applications, the sheer number of two-way exchanges required to complete an application transaction quickly aggregates until response times of several seconds are common, even with small numbers of active devices on the network.

In the case of telnet applications it is common to transmit far more information than is necessary. It is also common for the application to break the transaction into segments, each of which introduces latency. But because the segments don't require a response, the effect is not as severe. Telnet applications therefore tend to be more forgiving, and larger populations of devices can generally be accommodated.

Both telnet and web-based applications benefit from the introduction of an intermediate "fixer" that compensates for the respective weaknesses of each technology. The solution for telnet applications is a fixer that collects the segments, tracks what the device already knows, and transmits only what the device needs to know, in one message.

The web solution is more complex. The modern web application requires the exchange of multiple documents to complete a transaction. There is at least one HTML document that provides the basic layout, usually at least one cascading style sheet (CSS) document that further refines the appearance and provides visual continuity, and usually at least one JavaScript document to handle all of the things HTML and CSS cannot. And of course there is the potential for one or more image documents.

This "house that Jack built" design tends to improve with successive version of the HTML standard, but even then it won't be practical to rewrite every web application to meet the latest standard, and it won't always be practical to expect every device on the network to have the latest browser software. This is especially true in BYOD environments.

The traditional method for reducing latency in web applications is to perform a "pre-fetch" of all of the documents the browser is likely to request. However, that doesn't address the underlying wireless problem associated with sub-transactions. Not only must the "fixer" pre-fetch the documents, it must combine them into a single document in such a way that the browser is not aware of the merge.
All browsers implement caching, which is to say that they don't have to repeatedly download documents they already have. However, in the case of multiple documents, browsers do query the server to determine if a cached document is current; this requires yet another sub-transaction.

The browser documents (HTML, CSS, JavaScript, and images) are all subject to compression. Most application development tools, however, do not enforce compression. Image documents are already in compressed format and can only be reduced in resolution, which is ideal for mobile devices but undesirable for desktops sharing the same application.

Summarizing, the browser application 'fixer' must perform at least two functions if response times are going to be reduced to a minimum: 1) Collect and merge all of the documents related to the transaction into one document and compress it, and 2) reduce the resolution of image documents, or else selectively discard them.

With the proper "fixer" technologies in place, average response time improvements of over 100% should be expected. For the operations manager, this translates into increased productivity. For the IT manager, it reduces the need for complex and expensive wireless network expansion. For the CIO, this translates into reduced IT costs and higher profits.

Hybrid Application/Device Environments

As application and device technologies constantly evolve, homogeneous environments are disappearing. There are many good reasons to maintain a mix of the old and the new, both in terms of applications and devices. The trend toward BYOD amplifies the challenge.

Legacy applications remain the majority almost two decades after the introduction of web technology, and for very good reason: They are thoroughly debugged and comparatively efficient, having been designed for low speed networks.
The introduction of an intermediary that allows browser devices to be used with legacy applications prolongs the life of legacy applications and reduces the pressure on IT departments to convert to newer, relatively inefficient and comparatively untested web-based applications. Decisions to migrate to new applications can be based on business model changes, rather than device requirements.

Enabling a BYOD environment translates into reduced capital expenditures for the CIO. The ability to support multiple application architectures, including viable legacy applications, reduces development and maintenance costs for the IT manager. The provision of a consistent user interface across all devices on a network reduces training costs for the operations manager.

Fault Isolation

Given the high performance nature of modern wireless applications, combined with the coexistence of mixed application and device technologies, a means for isolating and correcting the myriad poten-

tial individual sub-system faults is essential. Simply knowing the status of a session or the operational statistics of a device is not enough. System administrators must have a uniform, easy-to-use mechanism for quickly identifying and analyzing application, network, device, and operator faults. In some cases, fault isolation is nearly impossible without access to the data level.

The business cost of not having a comprehensive diagnostic platform is similar to that of not having a fault tolerance. First, "living" with operational inefficiencies for an extended period of time costs just as much, if not more, in lost productivity than a complete one-time shut down. Second, the effort required to isolate sub-system faults, such as application logic errors or operational errors, can stretch over weeks and even months, with frequent meetings involving IT staff, operations staff, and external vendors. All the while productivity is suffering and profits are lost.

A cloud-based diagnostic platform that provides access to the data level of each device session, the ability to isolate the host from the wired network, and the ability to isolate the wired network from the wireless can reduce a thousand-hour on-site effort to a few hours of a remote technician's time.

Wireless Session Interruptions

Wireless networks are subject to frequent transient session interruptions not found in the wired world. Rarely is it practical to engineer applications or device software to deal with this inherent characteristic. The ideal solution is an intermediate "proxy" server that maintains the host connection throughout the interruption, and automatically restores the wireless side of the session as transparently as possible.

Standard telnet software requires an initial telnet session negotiation when an interrupted session is reestablished. However, an intermediate proxy server that is capable of imitating the role of the server and re-establishing the session with the host provides an efficient and transparent solution. This approach does not require custom software on the mobile device.

Telnet applications converted to browser applications (as described in the previous section) are far less sensitive to session losses. The intermediary isolates the mobile device from the underlying telnet protocols, so restoration of the host connection is completely transparent to the operator. And no custom software is needed on the mobile device to assist in the restoration; standard browser refresh controls suffice.

Numerous studies detail the high costs of session interruptions in terms of productivity losses and operator frustration, both of which have significant impacts on the top and bottom lines.

Fault Tolerance

Virtualized cloud-based architectures have the inherent characteristic of fault tolerant operation, limiting equipment-related downtime to mere seconds. The effects of eliminating downtime are obvious and well documented. At the CIO level, the impact on the top line is substantial. A system outage of even a few minutes can wreak havoc on all aspects of the business.

Connect's Solution

Connect is providing solutions to these challenges by building on the NCU technology it has been perfecting, in the field, for more than 20 years. By compiling this ultra-stable, bug-free solution into a virtual appliance the "single point of failure" argument that has been used against the "hard" NCU is no longer relevant. Fault tolerance is an inherent characteristic of virtual appliances.

Additionally, because there is no hardware, and virtual appliances come with pre-built, self-contained operating systems, IT implementation and maintenance costs for this technology are eliminated.

The architecture of the core technology is modular. This makes it easy for Connect to adapt this core to the latest mobility challenges by adding a module. Such a platform empowers users with a cost-effective and seamless migration path as new challenges are introduced by the evolution of technology.

CloudMax

CloudMax is a Virtual Interface and Performance (VIP) Platform: Virtual – it is a virtual appliance; Interface - it enables any type of application to connect to any type of mobile device platform; Performance - it significantly improves the performance and functionality of mobile devices being used in real-time environments; Platform – its modular architecture enables additional capabilities to be added as its users' challenges evolve.

CloudMax comprehensively resolves the challenges of performance, hybrid environments, fault isolation, wireless session interruption and fault tolerance.

Performance

For telnet based applications CloudMax combines the multiple sub-transactions generated by the host into a single transaction for transmitting to the mobile device. Additionally, many techniques are used to reduce the data volume of each of those transactions as well. On average, data traffic is reduced by 75% and data volume is reduced by at least 95%. The result is a substantial improvement in response times and a dramatic reduction in traffic over the network.

The techniques for web-based applications are somewhat similar. As mentioned earlier, a web-host and the browser will open multiple sockets to render one screen update. Among other optimization techniques CloudMax maintains a single socket connection, merges all of the documents into a single HTML document, compresses the data and reduces the resolution of any graphics.

In both cases, telnet and web, the level of optimization can be manually adjusted to fit the behavior of the application.

Hybrid Application/Device Environments

CloudMax supports hybrid environments by enabling legacy applications for web-based devices and enabling web-applications for character-based devices. Additionally, it can enable a 3270 or 5250 application for a VT terminal emulator.

Fault Isolation

Connect resolves the challenge of fault isolation with CloudMax. In some cases, it can be impossible or extremely costly to identify the root cause(s) of faults without data-level access to the transactions between the host application and the device. This "transaction aware" characteristic makes CloudMax an unbiased observer and powerful fault isolation tool. In other cases, the ability to troubleshoot faults requires the ability to segregate the wired and wireless portions of the network.

The CloudMax web interface allows for easy remote access to the necessary data logs and traces. The user simply takes a "snapshot" of the diagnostic traces and CloudMax automatically uploads them to Connect's FTP site where they are analyzed. This allows Connect to efficiently identify the root cause(s) of the problem(s) without requiring the hassles of getting approval for VPN access to customers' networks.

The ultimate benefit to the customer, especially in multi-vendor environments, is a significantly reduced time to problem resolution. The benefit to the vendors is the ability to quickly identify that their own technology is not the root cause, and if it is they can fix it before it costs the customer in any significant way.

Wireless Session Interruptions

CloudMax transparently restores lost, dropped, or casually connected sessions. This is ideal for field services and enabling legacy applications for outside the four walls.

Fault Tolerance

As a virtual appliance, CloudMax integrates into high availability and load balancing environments. When it is moved seamlessly to another physical server there is no interruption to the operators as the IP address remains the same. For non-virtualized environments the virtual appliance can be copied to any server on the network for multiple redundancies. However, the IP address will change. If ClientMax (Connect's proprietary client-side software) is being used it knows to look for the next available CloudMax upon reboot of the device.

Case Studies

Customer: Worldwide supplier of industrial products

Problem: Capacity Shortage

Description:

This customer was opening new warehouses with the goal of deploying 300 RF devices per site. In testing, when they reached 25 RF devices, the entire system became unstable and sluggish. Users were experiencing long delays, time-outs and disconnections.

Connect diagnostics uncovered that the application was extremely inefficient in its operation. Typical screen updates were using 50KB of emulation dialog with up to 25 separate sub-transactions.

This inefficient operation caused all of the available RF bandwidth to be used, which was the source of

the capacity problem. The options at this point were to rewrite the application or to expand RF capacity. Both options were expensive and time consuming.

Instead, they implemented CloudMax and they were immediately able to get all 300 devices operating with no stability or performance problems. CloudMax made this possible by reducing network latency to less than 50 milliseconds.

How? CloudMax reduced screen updates from 50KB and 25 sub-transactions to 20 bytes and 1 transaction. This was accomplished with zero changes to the application or existing infrastructure.

Customer: European Ferry Operator

Problem: Session loss while ferries are out to sea

Description:

This customer uses fixed mount RF devices on each of the car and people ferries it operates. They connect to Wi-Fi networks while in port. Once they are out to sea, they do not need to use the devices until they reach the next port.

The pain point was that during the voyage the session was occasionally being lost. When they got back into port they had to reboot the RF device, log back into the application and do a recovery procedure to get back to where they left off. This caused operator frustration, loss of data and increased operational time.

Connect diagnostics uncovered that three conditions were causing the session losses. The first was TCP/IP "Keep Alive" probes, the second was unsolicited host screen updates and the third was power loss to the terminal.

Our recommendation to this customer was to implement a CloudMax system configured for session persistence. Now when ferries are out to sea, sessions are maintained by CloudMax. CloudMax responds to the "keep alive" probes in proxy to the terminal, holds up delivery of host updates until they are back in port and restores lost sessions due to power loss. Now when they reach the next port, the session is active and at the exact spot in the application as it was at the last port.

CloudMax, in effect, changed this customer’s ferry application to an "outside the 4 walls" application without any code or infrastructure changes.

Customer: Cable TV Operator

Problem: PC application did not work well on RF devices

Description:

This customer wanted to implement portable RF devices in its service dispatch operation. They had been using a PC-based application using fixed station PC's. The operation had grown to a point where

this was no longer practical, so they needed portable devices to quickly get the service vehicles equipped and out the door for their daily install and service activities.

They installed the required infrastructure and RF devices and configured them to run the PC application. This turned out to be a disaster for their operations. They found the accuracy had decreased and the time had increased, which was the exact opposite of what the goal was.

With the data from CloudMax traces sent to Connect via the snapshot feature, we found that the operators were entering information into the wrong fields, spent a lot of time "scrolling" the display around and other operational problems.

The options they considered were:

- Redesign and expand the garage area with more fixed station PCs
- Redesign the PC application to work with the new RF devices

Both options were expensive and had very long implementation times.
We recommended the implementation of CloudMax. They used the integrated CloudMax tools to make the existing PC application more RF friendly. They used screen formatting to rearrange the screens, dialog language to automate tasks and mobility features to improve operations.

The customer was able to make the PC application work extremely well with the RF devices. All of this was done in 30 days without any changes to the host application or infrastructure.

Customer: U.S. Auto Manufacturer

Problem: Session disconnections

Description:

The original diagnosis by company IT personnel was an RF coverage or infrastructure problem. The disconnections were occurring when the users were in one specific area of the warehouse. The company brought in the RF, terminal and router vendors to run tests.

This process took about 4 months as tests and upgrades were done. Some of the upgrades created new problems which then needed to be worked out as well. At the end of 4 months, disconnections were still occurring at the same rate.

Using the data from CloudMax, Connect found that the problem occurred while they were in the receiving application (which is only done in one specific area). The trace showed there was a brief error message about a "record lock". A few transactions later a disconnection occurred. The message was displayed so briefly that none of the operators had ever seen it.

Once the company knew of the error it allowed them to correct a problem with the application. This eliminated the error and the disconnections that it led to. Total time was 2 days to identify the root cause and 2 days for this company to correct. All diagnostic services were done remotely and had no

impact on normal operations.