

RESEARCH

Simulation of Medical Device Network Performance and Requirements for An Integrated Clinical Environment

David Arney, Julian M. Goldman, Abhilasha Bhargav-Spantzel, Abhi Basu, Mike Taborn, George Pappas, and Michael Robkin

About the Authors



David Arney is a systems engineer with the Medical Device Plug-and-Play (MD PnP) Interoperability Program. E-mail:

dave@mdpnp.org



Julian M. Goldman, MD, is medical director of Partners HealthCare Biomedical Engineering, director of the

program on interoperability at the Center for Integration of Medicine and Innovative Technology, and founder of the Medical Device "Plug-and-Play" (MD PnP) Interoperability Program. E-mail: jgoldman@mdpnp.org



Abhilasha Bhargav-Spantzel, Ph.D. CISSP, is a security architect at the Healthcare Solutions group at Intel, SC.

Introduction

Healthcare information is increasingly being shared electronically. The adoption of healthcare information exchanges (HIE) and the creation of the Nationwide Health Information Network (NwHIN) are two examples of pathways for sharing health data that help to create the demand for interoperability (Figure 1). Regulated medical devices will play a key role in supplying the data for such systems. These medical devices will now be required to provide data well beyond the point of care, requiring integration of medical device data into the clinical information system. The challenges of interoperability in healthcare were described in a recent Networking and Information Technology Research and Development (NITRD) report¹ that highlighted the need for standards-based interoperability in healthcare. As standards-based interoperable medical device interfaces are implemented more widely, device data can be used to create new decision support systems, allow automation of alarm management, and generally support smarter "error resistant" medical device systems. As healthcare providers connect medical devices into larger systems, it is necessary to understand the limitations of device connectivity on different types of networks and under different conditions of use.

Many groups recognized the need for interoperability and have addressed different aspects of the challenge. For example, Universal Serial Bus (USB) defines a physical and protocol layer.

Similarly, Bluetooth defines the physical layer, protocol within that physical layer, then various application level profiles such as Health Device Profile that help to define where data will sit within the protocol. This full definition of the stack and the application helps to define an environment that makes all devices within that application type interchangeable, and leads to the notion of plug and play (PnP). In healthcare, Health Level 7 (HL7) is one example of an application data exchange format that is quite prevalent. The Continua Health Alliance implementation of portions of IEEE 11073 across multiple physical layers including USB and Bluetooth is another example of healthcare standards-based interoperability that has achieved plug-and-play interoperability.

The modeling done here is based on the Integrated Clinical Environment (ICE) standard ASTM F2761-09² based on work performed by the Center for Integration of Medicine and Innovative Technology Medical Device (CIMIT MD) PnP team.³ The concepts behind the ICE standard go back to the beginning of the CIMIT MD PnP Program, which began as an effort focused on addressing a set of pressing and often critical high-acuity, point-of-care safety problems that emerged during the delivery of care. These include the inability to implement smart alarms, safety interlocks, and closed-loop control algorithms that could be used with devices from multiple vendors and safely deployed in multiple healthcare delivery organizations (HDOs) without excessive

reimplementation or regulatory burden. Specific examples of these applications include patient-controlled analgesia (PCA) safety interlocks^{4,5,6} synchronization of ventilators with X-ray machines or other imaging devices,⁷ and smarter alarm management⁸ to reduce false positive alarms and alleviate alarm fatigue. The problems, many of which had been described for years, were associated with events that differed in terms of clinical and technical details but intersected conceptually in that safe and effective solutions could be devised if the participating medical devices were able to cooperate. The ICE standard specified a hub-and-spoke integration architecture in which medical devices and health information technology (IT) systems are the spokes, and ICE-defined components implement the Hub.

Our hypothesis is that adequate network performance for a large class of interoperating acute care medical applications can be achieved using the standard transmission control protocol/Internet protocol (TCP/IP) over Ethernet even if those medical devices are on a network connected to a hospital enterprise. We devised a series of experiments to evaluate if this is true for applications that have timing requirements measured in seconds. Many medical applications have relatively lax timing requirements compared to classic real-time control applications such as aircraft flight control, and can obtain an acceptable level of performance even on extremely slow or overloaded networks. Our purpose here is to explore the use of

standard networks, which may be shared with other traffic, for the purpose of connecting medical devices to a centralized computer that runs an application using the device data.

Use Cases

We designed our experiments to simulate two realistic use cases with varying network speed and traffic types, including large file transfers, typical workstation network usage, and streaming video. These use cases involve interactions between medical devices, hospital infrastructure including electronic medical records, and the ICE components. In the first use case, we examine the impact of network traffic on a simple closed-loop control system, specifically an infusion safety interlock. The second use case looks at the overall traffic in a complex operating room, and allows us to examine network performance under increasing load levels. We analyzed the performance of a simulated wired segment of the hospital network where medical devices operate (comparing between 1Gbps, 500 Mbps and 100 Mbps bandwidths) and quantified the necessary quality of service levels for assured safe and effective operation of various clinical scenarios. We have defined these quality of service (QoS) levels in terms of the latency of message completion, and deemed the latency to be acceptable if device data is delivered within 1 second and Digital Imaging and Communications in Medicine (DICOM) studies are transferred within 20 seconds.

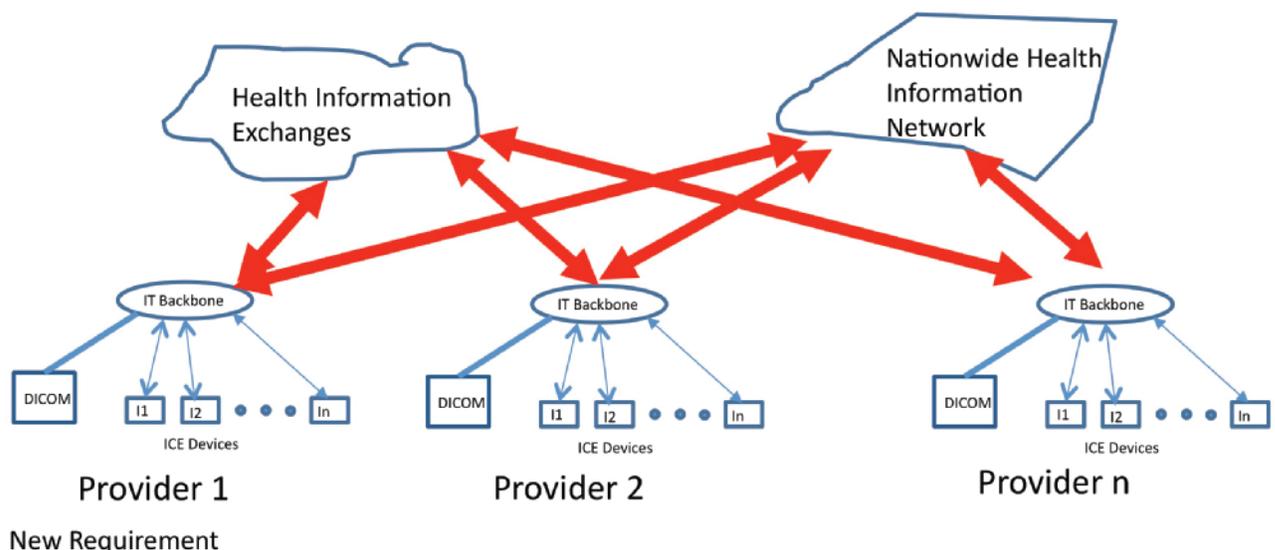


Figure 1. The Hierarchy of the Network Environment

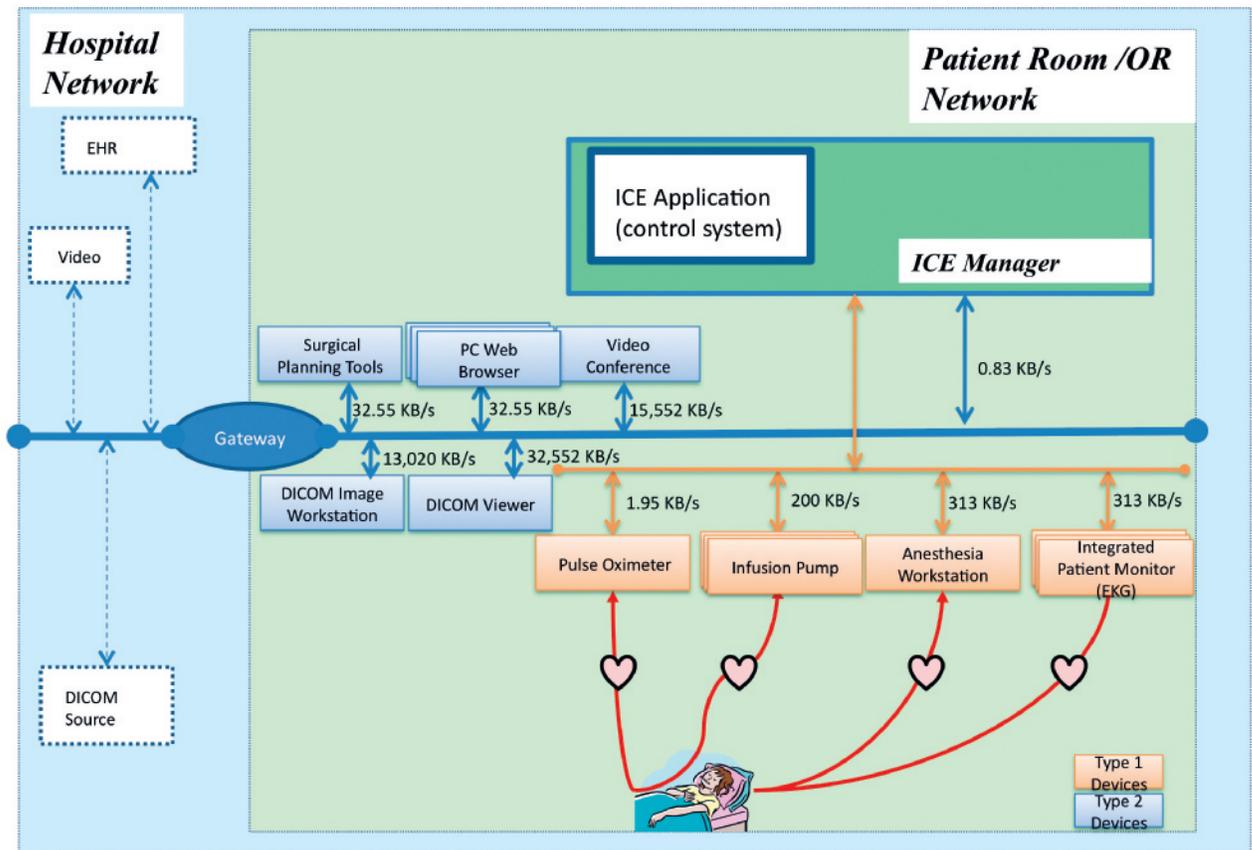


Figure 2. ICE Architecture

Use Case 1

Infusion Safety Interlock

PCA pumps are used to deliver pain medication that affects the patient’s heart rate, respiratory rate, SpO₂, and capnograph waveforms. By monitoring these vital signs, we can detect that a patient may have received too much of an opioid pain medication and automatically stop the pump. Similar systems can be built for

infusions of heparin (monitoring invasive and non-invasive blood pressure) or Propofol (depth of sedation monitors).

To understand the impact of network performance on the PCA example, we look at two categories of network traffic:

Internal operating room (OR) traffic is internal to the isolated network related to the patient’s room or the operating room. In Figure 2, this traffic is between the medical devices and non-medical devices and the ICE manager.

Backend Traffic uses the hospital network to transfer summary data from the ICE to the hospital electronic health record (HER), as well as traffic from the ICE to outside resources, such as a DICOM image server or videoconferencing system.

One important note in this configuration: To avoid a potential bottleneck, we give the ICE manager two different network interface cards with two different IP addresses.

Use Case 2

Advanced Operating Room

This case builds on the infusion safety interlock room set-up and adds additional traffic using

#	Device Name	Data Rate (Kb/second)
1	ICE Manager	6.6
2	ICE Coordinator to ICE Manager	6.6
3	ICE Coordinator to EHR Server	6.6
4	Pulse Oximeter	15.6
5	Infusion Pump	1600
6	Anesthesia Workstation	2504
7	Integrated Patient Monitor (EKG)	2504
8	Surgical Planning Tools	260
9	PC Web Browser	260
10	Video Conference (stream)	124416
11	DICOM Image Workstation (Upload)	104162
12	DICOM Image Workstation (Download)	260416

Figure 3. Devices Used in Experiments

the hospital network to connect to image servers and streaming video systems. These experiments show the impact of large file transfers in the form of DICOM images. We also modeled streaming video, such as is used in laparoscopic procedures or to send video to remote sites for teaching or consultation. Finally, we added networked workstations with a typical office workload, to represent computers that are often placed in operating rooms for documentation and access to network resources including the Internet.

Experimental Design

Our simulations were done using the IxChariot tool⁹ to simulate medical device workloads and measure performance in the above environment. IxChariot is extensively used for simulating network workloads, including use in the medical device industry.¹⁰ Our simulations used physical hardware with a separate computer per simulated device and real network hardware connecting the computers. We simulated representative workloads for medical devices of interest. Workloads for some devices, such as the pulse oximeter, were based on measured values, and other workloads were set to a level that is higher than most current devices, but indicative of the workload that would be associated with a device capable of providing detailed information over its network interface. Our anesthesia workstation, for instance, was assumed to provide pressure and flow waveforms (similar to a critical care ventilator), drug information, and other details about the settings and status of the workstation. For the infusion pump, we examined existing devices and standards such as the draft 11073 Domain Information Model and extrapolated what we felt was a reasonable worst-case data load. In the case of the pump, this meant dozens of state variables, dose history, patient and caregiver identification. Because we intended for this device to provide data to an alarm system, we had it send its 15KB updates every 50 milliseconds, resulting in a rate of 1600Kb/sec. Figure 3 lists the devices that were simulated and their workloads.

We selected a flat network topology when running simulations. There is only one router (switch) that lies between any client and server; therefore, there is only one hop communication for the various traffic types. We simulated three different network bandwidths for our experiments—1Gbps, 500 Mbps and 100 Mbps. Three key domains were setup to emulate the administrative, radiology, and OR environments. The OR domain represented the various operating rooms instances.

The key network performance metrics that we considered were throughput and response time. Throughput measures the amount of data (bits or bytes/second) that can be sent through the network.¹¹ Response time measures the end-to-end roundtrip time required to complete an application-level transaction across the network.¹²

IxChariot has two key limitations that directly influenced our experiment design:

1. All communications are between pairs of endpoints we were

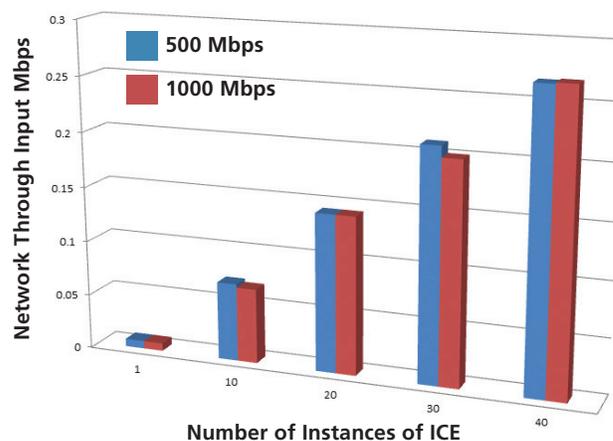


Figure 4. Network Throughput With Increasing Number of ICE Instances

limited to 50 pairs per license. Since we modeled 18 devices in each ICE instance (each of which used one pair to connect to the ICE Manager), we invented a new notational device that aggregated the data from an entire ICE network in order to scale the simulated traffic.

2. IxChariot can only handle a maximum of 10,000 timing records. This made it necessary to aggregate some of the sub-second data packets to a per-second transmission rate (per Figure 3) to keep timing records below the recommended threshold.

Approach

Our experiments progressed through three key stages:

1. After deducing the traffic flow from ICE network within the OR, we followed by simulating the aggregated ICE traffic from the first experiments together with additional traffic between the ICE manager and the EHR database over the hospital network.
2. We progressively added additional traffic types corresponding to other medical devices that are present in the OR to see how they impacted the ICE traffic as well as the other traffic types on the shared network.
3. Once we had the entire complex OR traffic represented on one network, we then grouped different traffic types together to find a distribution of the traffic that achieved acceptable network performance without compromising the response time, safety and efficacy of other medical devices.

Internal OR Traffic

We observed that the average throughput of the ICE devices within the OR (pulse oximeter, anesthesia workstation, infusion pump and EKG monitor) is approximately 35 Mbps and that each of the 14 instances of infusion pumps peaked at around 65 Mbps. The response time for each traffic type was less than 0.05 seconds, with the EKG taking about .03 seconds. These experiments suggest that a single 100Mbps network provides an

acceptable response time for applications that use these types of devices to implement safety interlocks or smart alarms. These response times would also be acceptable for many types of physiological closed-loop control applications, though application developers must take into account many other factors.

Then we scaled from 1 to 40 representations of ICE traffic (outside the OR) and analyzed them with two different network bandwidths – 500 Mbps and 1Gbps. We saw that due to the low usage of the network bandwidth by the ICE traffic, there was minimal difference between the two, and that the ICE traffic can work as effectively in 500 Mbps as in 1Gbps provided there is no other traffic on the network.

Figure 4 shows the network throughput with an increasing number of ICE instances on a single network and we found that even 40 instances use a very small amount of bandwidth.

Backend Traffic

After getting the baseline traffic from the ICE-only experiments, we progressively added additional medical workloads to our environment. Each instance of backend traffic consisted of the following device workload simulations

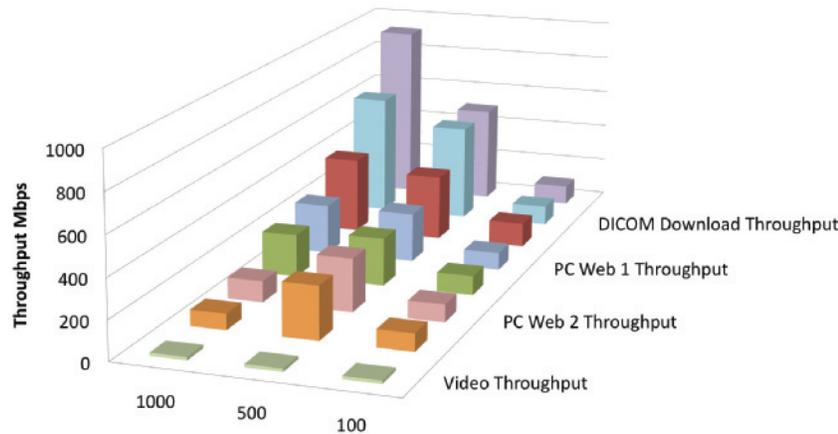
(refer to Figure 3 for workload details):

1. ICE manager
2. Surgical planning tools
3. Two instances of PC Web browsing traffic
4. Video conference
5. DICOM file upload
6. DICOM file download

All data transmissions occurred between the OR domain and other domains (administrative and radiology) residing on the hospital network. The enterprise network was connected to a Complex OR using a common gateway. (Backend traffic path of complex OR devices is identified by blue lines in Figure 2.)

We observed an interesting ordering of the traffic, where the video throughput (video conferencing) was lowest and the DICOM traffic was the highest. This is due to the fact that video traffic used real time protocol (RTP)—a user datagram protocol (UDP) that allows for best effort delivery—set to a high quality bitrate of 15.5 Mbps, whereas the remainder of the transmissions were transmission control protocol (TCP). We experimented with the three different network bandwidths and saw that in a complete complex OR, the DICOM throughput in a 1Gbps network was reaching a choking

Complex OR Traffic Types across different NW Bandwidths



	1000	500	100
Video Throughput	15.928	15.928	15.928
SPT	81.856	269.514	89.788
PC Web 2 Throughput	108.134	271.036	89.704
ICE Throughput	218.186	244.903	97.563
PC Web 1 Throughput	254.375	252.236	88.972
Throughput Mbps	390.057	336.726	123.398
DICOM Download Throughput	628.673	499.949	99.999
DICOM Upload Throughput	923.425	499.954	100

Figure 5. Advanced OR Traffic Types Across Different Network Bandwidths

1G Network – Progressive addition of Medical Traffic Types

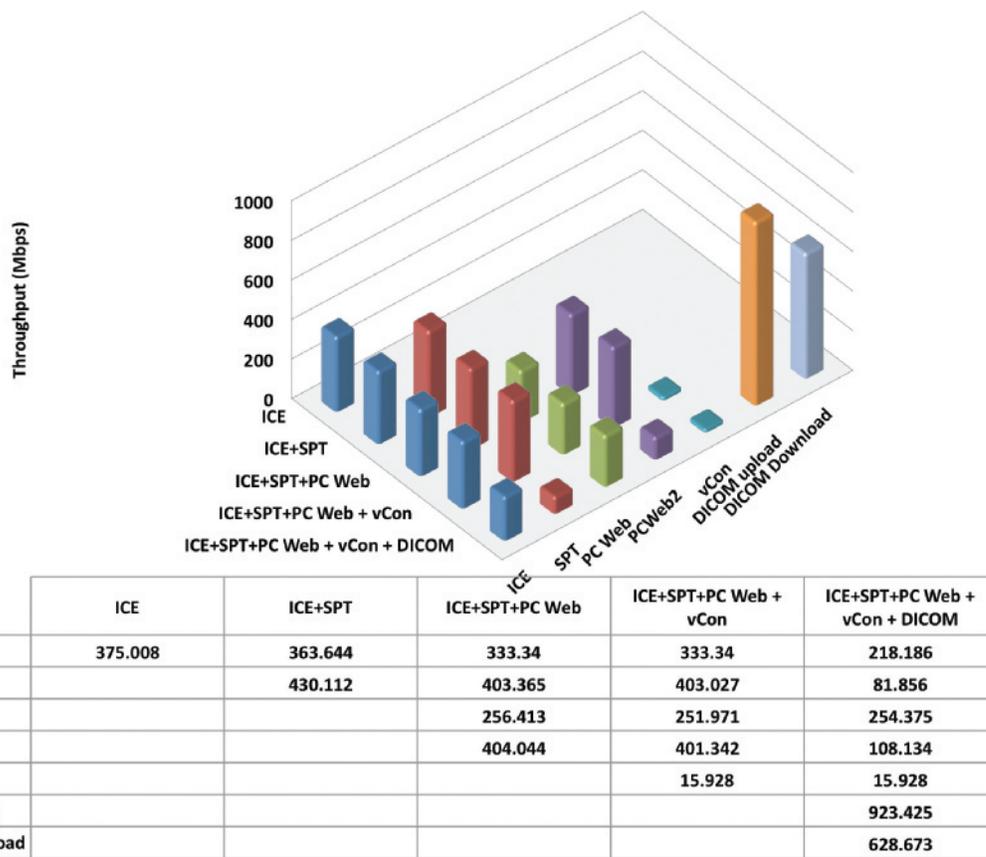


Figure 6. 1Gbps Network—Progressive Addition of Medical Traffic Types

point. Similarly, the thresholds were reached for the average and DICOM traffic for 500 Mbps, and the 100 Mbps could not support any traffic other than video effectively.

When we added the various advanced OR traffic to the background chatter (ICE traffic), we noted that addition of other traffic types (other than DICOM) for a given bandwidth has minimal impact on the throughput of a traffic type (as depicted in Figure 6). With the addition of the DICOM traffic, we see that the SPT and PC Web traffic get impacted significantly.

We also observed how the ICE traffic performs with the addition of different types of traffic with increasing instances of an OR. For a 1Gbps network, we see that there is a significant and sudden loss of throughput with the addition of DICOM traffic when the number of OR instances is around five. We also see a steep dip in the ICE throughput with the addition of the entire complex OR traffic. The corresponding worst-case response times were all below .03 seconds, so were acceptable for the use cases we modeled. We saw similar results for

Avg Throughput (Mbps)	1	3	5	6
PCA	0.007	0.021	0.034	0.04
Non-DICOM	16.733	50.198	83.633	100.398
DICOM	361.312	754.161	867.124	826.864
Complex OR	390.057	867.957	1079.17	948.143
Complex OR + DICOM Var	396.626	811.506	934.654	874.626
Avg Response Time (Seconds)	1	3	5	6
PCA	0.001	0.001	0.001	0.001
Non-DICOM	0.002	0.003	0.003	0.003
DICOM	12.434	23.254	42.644	50.191
Complex OR	3.568	10.198	17.914	18.051
Complex OR + DICOM Var	3.561	10.363	20.41	58.626

Figure 7. Average Network Throughput and Average Response Times of Different Groups of Backend Traffic Over Multiple OR Instances on 1Gbps Network. (Note: Red highlight denote points where traffic is not sustainable due to lack of bandwidth or very high response times.)

the various traffic types and OR instances on the ICE traffic for a 500 Mbps network. Thus, we conclude that while the throughput of the network may be impacted by DICOM traffic, the worst-case response times are not impacted significantly, even with increasing number of ORs (we experimented with up to 10).

Scaling and Grouping Traffic Types

Next we grouped the various traffic types and studied the resultant average throughput utilization of the various groups in a 1Gbps network setting. Based on our previous experiments, we noticed that the DICOM traffic would have the maximum impact, so we added a variation to the type of DICOM traffic. Traditionally, we were sending 4GB files for all DICOM upload and download. In the group depicted by “DICOM Var,” we broke the DICOM packets into a combination of 1GB and 2GB, such that the total data sent is the same as the original DICOM traffic. The results are illustrated in Figure 6. We see that the complex OR takes the maximum throughput, followed by DICOM Var, then DICOM, followed by non-DICOM, and finally PCA. Varying the packet sizes did increase the average throughput use of the 1Gbps network. We see a steep rise in the average response times of Complex OR with DICOM Var for five instances of OR. As seen in previous experiments, the rise in the response time is primarily due to the DICOM traffic.

Using More Detailed Backend Traffic

The advanced OR traffic was next modified to add all internal OR traffic from ICE devices (pulse oximeter, anesthesia workstation, infusion pump, and EKG monitor) and custom bi-directional acknowledgements where acknowledgements were the same workloads pre device returning from receiver to sender. We had lenient requirements for timing—20 seconds for DICOM pair response and 1 second for non-DICOM pair response. This reflects the typical use cases, where clinicians are willing to wait 20 seconds for an image set to load, but device data must be delivered more quickly.

We could not scale this traffic on 500 Mbps or 100 Mbps networks reliably. On a 1 Gbps network, we were able to scale to three simultaneous instances of OR, but since the DICOM pairs did not meet response time expectations,

we concluded that only ORx2 can be realistically supported. Next, we wanted to group DICOM and non-DICOM traffic on two distinct 1 Gbps networks to understand how we could scale the overall OR traffic effectively. This time we were able to scale DICOMx3 and non-DICOMx7 before response time expectations were violated.

The non-DICOM scaling showed an interesting pattern. Beyond non-DICOMx7, the infusion pumpx14 (aggregate of 14 infusion pumps) acknowledgement pair used very little throughput and recorded very high response times. This caused the overall average response times (all pairs) to go up and did not allow non-DICOM instances to meet expectations beyond x7 (1 sec response expectation). So response times degraded much before maximum throughput was reached.

Limitations

The following are the key findings from our simulations:

- This was a simulation of network traffic. While we believe the simulation is valid, any life-critical application requires that these results be confirmed in the future on real networks with real medical devices before being accepted by the industry and implemented in practice.
- Our simulations were done under limitations of the capacity of the modeling tool i.e. IxChariot. This limited (1) the total number of network pairs, necessitating the aggregation of network traffic across devices, and (2) the total number of transmissions, which required the aggregation of many transmissions from a single device.

Conclusions and Future Work

We modeled a patient-centric ICE network in two use cases: an infusion safety interlock, and an advanced OR. We modeled these in a use environment where the devices were isolated on a separate network and another where ICE traffic shared a network with high bandwidth devices, such as videoconferencing and DICOM image transfer. Notably, streaming videoconferencing traffic had very little impact. We found that the throughput and response times measured were adequate for our use cases except where we had a large number of simultaneous DICOM image transfers. This suggests that a network topology that isolates

DICOM traffic (or other large data transfers) may still be necessary. Such isolation could be a separate network or involve simply limiting the bandwidth allowed for those transfers.

However, even when grouping traffic types together (DICOM and non-DICOM), we observed that average response times degrade much before network maximum throughput is encountered. In general, it appears that from a network performance perspective it is feasible to have interoperable medical devices that use both data and control signals on the same network as existing OR and ICU traffic and Hospital Enterprise network traffic provided that the latency requirements of the applications are compatible with the load levels of the network, and that the monitoring of network performance includes latency.

We plan to explore this area further by modeling different network topologies like segmented, distributed, and enterprise in the future. ■

Acknowledgment

This work was supported in part by Award Number U01EB012470 from the National Institute of Biomedical Imaging and Bioengineering. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Biomedical Imaging and Bioengineering or the National Institutes of Health.

References

1. **Networking and Information Technology Research and Development Program.** High-Confidence Medical Devices: Cyber-Physical Systems for 21st Century Health Care. Prepared by the High Confidence Software and Systems Coordinating Group. February 2009. Available at: www.nitrd.gov/About/MedDevice-FINAL1-web.pdf. Accessed July 2012.
2. **ASTM.** ASTM F2761 - 09 Medical Devices and Medical Systems - Essential safety requirements for equipment comprising the patient-centric integrated clinical environment (ICE) - Part 1: General requirements and conceptual model. Available at: www.astm.org/Standards/F2761.htm. Accessed July 2012.
3. **MD PnP Program.** Available at: www.mdnp.org. Accessed July 2012.
4. **Arney D, Jetley R, Jones P, Lee I, Sokolsky O.** Formal Methods Based Development of a PCA Infusion Pump Reference Model: Generic Infusion Pump (GIP) Project. Presented at Joint Workshop on High Confidence Medical Devices, Software, and Systems and Medical Device Plug-and-Play Interoperability, June 2007. Available at: http://repository.upenn.edu/cis_papers/358/. Accessed July 2012.
5. **Arney D, Fischmeister S, Goldman J, Lee I, Trausmuth, R.** Plug-and-Play for Medical Devices: Experiences from a Case Study. *BI&T*. 2009;43(4):313-317.
6. **Arney D, Pajic M, Goldman J, Lee I, Mangharam R, Sokolsky O.** Toward Patient Safety in Closed-Loop Medical Device Systems. Presented at Cyber-Physical Systems (ICCPS 2010) Stockholm, Sweden, April 13-14, 2010.
7. **Arney D, Goldman J, Whitehead S, Lee I.** Synchronizing an X-Ray and Anesthesia Machine Ventilator: A Medical Device Interoperability Case Study. *Biomedicine*. 2009:52-60.
8. **R Moberg, Goldman J.** Neurocritical Care Demonstration of Medical Device Plug-and-Play Concepts. The CIMIT Exploratorium 2009. Available at: www.cimit.org/events-exploratorium.html, Accessed July 2012.
9. **IxChariot tool.** Available at: www.ixchariot.com/products/datasheets/ixchariot.html. Accessed July 2012.
10. **Intelligent Transportation Systems.** Next Generation 9-1-1 (NG9-1-1) System Initiative: Data Acquisition and Analysis Plan. Version 1.0. U.S. Department of Transportation. Washington, DC: March 2008. Available at: www.its.dot.gov/ng911/pdf/NG911_DataAcquisition_v1.0.pdf. Accessed July 2012.
11. **IxChariot FAQ.** How is Throughput Calculated? Available at: www.ixiacom.com/support/ixchariot/knowledge_base.php#_Toc46047794. Accessed July 2012.
12. **IxChariot FAQ.** How Is Response Time Calculated? Available at: www.ixiacom.com/support/ixchariot/knowledge_base.php#_Toc46047804. Accessed July 2012.