Enterprise Computing—System-Level Thinking, Much More than an IT Data Focus, and Pedagogical and Career Imperatives to Include Such Topics in IT Curricula

Lawrence W. Hill
Information Sciences and Technology Department
Golisano College of Computing and Information Sciences
Rochester Institute of Technology
152 Lomb Memorial Drive
Rochester, NY 14623-5603
(585) 475-7064
Lawrence.Hill@rit.edu

ABSTRACT

What is, "Enterprise Computing"? Most people would say something about a large-scale computing focus, others might say something cloud referenced, still others might include datacenters and massive networking requirements and some will mention a mainframe (or enterprise server) focus. They'd all be correct in that there is no agreed-upon definition of the term but indeed the term is all inclusive of the above. When pressed to explain their definitions, more than likely most will probably focus upon providing applications and application data reliably and securely to end users or customers. No doubt this is the primary business requirement of large-scale computing but very few non-systems thinkers ever consider the underlying support structure that underpins the very existence of large-scale systems, without which those systems could not exist. This author will demonstrate the need for our graduates to be those systems-thinkers.

This paper focuses upon various infrastructure topic areas and how they relate to, and play extremely important roles in, the operation, functionality, and maintenance of modern data systems with the ultimate goal of providing excellent reliability, availability, and serviceability (IBM's mantra for decades) [1] in the delivery of applications and data to end users. Failure in any one (or more) of those critical infrastructure areas has the potential of rendering the system useless with a corresponding loss of data resources to customers. As has been heard too many times to count in the information age's vernacular, "It's all about the data!" and this paper will, hopefully, convince its readers that, indeed, infrastructure topics should be included in IT curricula such that *the data* will continually flow and be available when and where it's needed, knowledgably and efficiently managed and maintained by our graduates.

1. INTRODUCTION

Undergraduate IT computing programs generally follow the commonly accepted simplistic IT-focused view of large-scale computing because the traditional focus of IT programs is centered on five professionally defined pillars of computing. Those pillars are generally accepted as (in no order, but arguably of equal importance) database, programming, networking and data communications (to include systems administration topics at certain schools), human-computer interaction, and web systems [2], notably none of which focuses upon the infrastructure required to make it all possible (maybe we add a sixth pillar?). In fact the accepted ACM model IT curriculum deprioritizes infrastructure.[3] The closest pillar to infrastructure is networking, but that function is considered a commodity in current times by users of these systems. It's ubiquitous, like plugging a vacuum cleaner into the wall and expecting it to work, with no attention paid to how that AC power got to that wall in the first place until, of course, a blackout occurs. Not being able to use your vacuum cleaner is mere nuisance, but not having access to your data can be catastrophic and that underscores the need for students to grasp the importance and basic requirements of such infrastructure as they venture out into their chosen careers after graduation.

Incumbent among all who purport to be IT educators, especially those who teach topics in enterprise computing, is the need and commitment to including an infrastructure mini-focus. It is not this author's intent nor would it be appropriate to change IT education to include enough of this infrastructure focus to enable our graduates to become system designers and/or engineers. After all, that is the purview of other established disciplines and subsequent career paths. However, this author feels strongly that our graduates should be at least conversant in system infrastructure details with the appropriate personnel responsible for the design and implementation of system infrastructure, and have the basic understanding of how systems function in their totality. In conversing with graduates having a year or two working experience, it is very common to hear that they wished they had some exposure to these infrastructure topics in their undergraduate experiences and they felt they would have been far more confident, valuable, and productive had these topics been included in their studies. See the appendices at the end of this paper to get an idea of what two students thought they learned in this author's datacenter operations course. Those presentations were a response to an assignment that asked, "What is datacenter ops?"

The areas described, and their importance demonstrated, are the use of mainframe servers, basic electricity/electrical principles, AC/DC power generation and distribution, electrical safety, project/configuration management and documentation (above and beyond system administration CM requirements) HVAC fundamentals, weather and its relationship to cooling requirements, datacenter construction (hot aisle/cold aisle and airflow/water cooling), fire suppression systems, physical equipment and facility access and security, and special attention given to data backup/restoration/recovery scenarios (beyond what is taught in standard IT courses). Each area has a vital role in the operation of data systems.

For the reasons outlined, and the reality that enterprise (mainly mainframe) talent in the global employment pool is diminishing (voluntary termination, retirement, and unfortunately, declining health) at an alarming rate (generally, anyone that knows mainframe technology are of the baby-boomer generation), much faster than new graduates could replace even if they had the talent, it is imperative that IT educators and the programs in which they teach address this issue of real overall systems thinking that include infrastructure considerations. There are a wealth (10's of thousands)[4] of high paying and rewarding employment opportunities awaiting those graduates with the skills necessary to fill those positions, and it is the IT teaching profession's purpose for existence, that is, to see that everything within reason is done that culminates in our graduates attaining their degrees with meaningful and rewarding employment opportunities available to them as a result of our efforts and expertise.

2. BODY

Contrary to popular belief, MAINFRAMES ARE NOT DEAD! In fact, they are thriving and becoming an even more critical component in today's complex systems that deal with the massive data explosion and subsequent data analytics that drive modern IT systems. The utilization of the mainframe platform can and does respond well to almost all of the aforementioned infrastructure topic areas/requirements and has overwhelming industry evidence demonstrating their vital role in the enterprise over decades of reliable service in countless business domains. Again, mainframes are not dead, evidenced by the fact they are utilized in virtually all Fortune 500 companies domestically and globally. Almost all major banks, insurance companies, and healthcare systems have mainframes for their exceptional speed, security, and extreme memory and storage capabilities, and most importantly, a history of the most reliable service of any computing platform on the planet. IBM's Watson^{R} platform, famous for its AI implementation would simply not be possible if not for its hosting by mainframe and supercomputer platforms. Such systems will play an ever important role in today's knowledge-retention and access systems.

As previously stated, people with mainframe knowledge are part of a diminishing labor pool, and this reduction in talented employees that operate, program, maintain, and administer these systems needs to be addressed in IT curricula. Integration of such content is easier than it might first appear in that IBM offers its Academic Initiative program free of charge to any faculty at accredited institutions for the purpose of fostering an interest and teaching the basic skillset to students.[5] Students exhibiting this system skillset literally hold the future of large scale computing in their hands, without which the industry would suffer irreparable harm.

AC mains Power availability is a constant concern among datacenter facility designers. The fact is, there is not enough power availability on the planet to support the growing needs of data-related services and infrastructure. Many otherwise ideal locations for such systems have to be rejected because of a lack of sufficient operational power. Many new datacenter projects must supply much of their own power locally, using wind, solar, and geothermal systems, in addition to whatever is available from the local grid. An example is the Apple datacenter project in North Carolina [6] where more real estate is devoted to power than the datacenter itself using solar (photovoltaic), geothermal and biogas.

The IT staff member(s) on-site must provide the power budget for the racks of IT equipment in use to the infrastructure power people. They must be in constant communication to ensure that enough power is available for current as well as future processing capability. To that end, basic AC power generation and distribution techniques, Ohm's Law, basic AC and DC circuits, transformer operation, and electrical safety topics should be included in some course or courses that will round out the students' understanding of infrastructure areas. Extreme amounts of watt-hour capacity are present within the power switching and battery rooms of datacenters and it is vital that anyone with access to these physical areas within the plant are aware and have knowledge of the necessary safety equipment required to protect life and limb when close to the power distribution equipment in these locations. Arc flash was interesting topic to cover during one of this author's courses as it shocked (no pun intended!) them into thinking about something previously unknown to them. Some discussion about the use of DC conversion and distribution internal to the datacenter facility should be included as this topic is starting to be utilized within datacenters to improve efficiency of local power usage.

Cooling is as important as power availability when considering datacenter operations. The processing of IT information results in the generation of massive quantities of heat within the equipment that must be quickly and efficiently removed from the premises. The cooling methods utilized over time have varied but the two most common methods of heat removal are air and water. Water is many more times heat efficient than air, and was used well into the past. It is not without its own issues, and was later replaced on a large scale by air conditioning. That lasted until the recent past where the newest processing environments generate so much heat (mainly due to their increased computational speed and density) that water is again becoming the method of choice. Regardless of the cooling methodology used, it becomes the responsibility of the IT professional to understand this heating load in his/her system and how to communicate with the HVAC professionals to ensure the facility has adequate cooling capacity for the current usage as well as later expansion. How datacenter equipment physically sheds its heat load is very dependent upon how that equipment is racked and positioned within the equipment rooms inside the datacenter. Students should be able to assist in the physical layout of the data processing equipment, and assist in the design of the topology such as hot aisle/cold/aisle partitioning and location of cooling ductwork. The design of the wiring and raised floor technology is taken for granted by many until, for example, the cooling fails only to find out that there is so much junk from previous implementations under the raised floor plenum that airflow is essentially cut off to the bottoms of the racks.

Part and parcel to cooling is the consideration of local weather considerations in terms of seasonal and short-term temperature and humidity variations as both play an important part in capacity and implementation of cooling systems. As an example, students must understand why the HVAC professionals included a water injection system in their facility that was put in place to increase the humidity and subsequently the heat content of the moving air to augment the system's heat removal capacity. Weather also plays a vital role in disaster scenarios and will be mentioned later.

Other facility systems such as fire suppression and physical access barriers need to at least be understood by IT professionals. These systems will be used in every datacenter in one manner or another. Many building codes require a water sprinkler system to be used regardless of the fact that the room is full of electrical equipment that will be completely destroyed after a sprinkler event. One needs to be in a position to deal with local code enforcement to mitigate this potential hazard to IT equipment, as an example. Physical barriers to unauthorized access must be considered. How do they work? How are people authenticated allowing restricted access? That is most likely another IT application that must be managed by the IT professionals.

IT datacenters are massive projects in their own right that demand a process-oriented approach in their conception, design, implementation, commissioning, operation, upgrading and maintenance, and documentation of all the above. ITIL [7] forms the basis of the configuration management of IT, datacenters, information handling, and the daily procedures used in the operation of data processing facilities. The basics of ITIL, and the understanding of configuration management in IT systems is vital for IT professionals to document their systems and have successful careers in doing so.

Disaster mitigation and recovery also has many definitions. Any definition would have to begin with the mantra of ensuring the availability of user data wherever, whenever, and on any device, by authorized users of that data. Period! Disasters range from data breaches and hacking, to hurricane damage or bad geographical location (the weather connection), to fire, flood, loss of mains power (a simple blackout), to the entire datacenter being swallowed by a sinkhole. IT professionals must be aware of the potential risks when commissioning new facilities and be able to propose mitigation strategies for response to any adverse events. The use of geographically-displaced alternate datacenters using real time data replication application hardware and software (nothing does that better than mainframes) must be recognized as a viable solution to many disasters by IT professionals. These technologies can be very complicated and could be covered by a dedicated course. Breadth of knowledge in these techniques should be adequate for an IT professional to be conversant with the respective personnel in these domains.

3. CONCLUSION

IT is many things to many people. Enterprise computing shares the same downfall. No one really knows what they are. What is known is the systems that make today's information processing possible are very complex (and becoming very more so) and requires those responsible for their reliable and error-free operation to wear many hats. The IT professional of old, greater than 10 years ago, could concentrate on the data applications and the configurations of the systems used to make those data and applications available to end users, and would have been an accomplished professional in doing so. With today's increased complexity, massive data storage, increased focus upon data and application delivery 24x7 anywhere and on any device fixed or mobile, one must now wear many hats to manage this complexity. This complexity requires system-level thinking and knowledge of non-traditional IT themes. The addition of systems infrastructure topics is one way send our graduates into the working work with increased confidence, efficiency, and self-esteem, resulting in more accomplished professionals than in the past.

4. REFERENCES

[1] Introduction to the New Mainframe:zOS Basics. Ebbers, Mike, et. al.,2009, IBM.

[2] IT Discipline, chart 1, ACM SIGITE website. http://www.sigite.org/?page_id=22

[3] Information Technology 2008, Curriculum Guidelines for Undergraduate Degree programs in Information Technology. Lunt, Barry M., et. al.. pg.19, 2008 ACM SIGITE website. http://www.acm.org//education/curricula/IT2008%20Curriculum.pdf

[4] Bureau of Labor Statistics, 2016, website,

[5] IBM Academic Initiative website. https://www-03.ibm.com/systems/z/education/academic/

[6] The Apple Data Center FAQ website. http://www.datacenterknowledge.com/the-apple-data-center-faq/

[7] ITIL Processes & Best Practices. BMC website. http://www.bmc.com/guides/itil-introduction.html

5. APPENDICES

These two presentations were done by two of this author's students in the spring, 2016 term at The Rochester Institute of Technology

1. "What I learned in datacenter ops". Steinberg, Peter. 2016
2. "Datacenter Operations & Me, What I Learninated [sic] in NSSA 425". Harvey, Morgan. 2016

What I learned in Data Center Ops is…(blankety...blankety...blank)
By: Peter Steinberg
Physics
Transformers: voltage vs current

Power (it ain’t clean)
Flywheel/Battery Backup
Redundancy
More Redundancy

Safety
Fire suppressants
Arc Flashes
Cooling
Hot-Aisle/Cold-Aisle
Raised floors
Sensors

Temperature
Location Location Location
Dew Point and Relative humidity: monitor it

- Datacenter Operations & Me
- What I learninated in NSSA 425.

By: Morgan Harvey

- Math!
- One thing you always wonder in class:
    - ‘When the hell will I ever need this?’
- Being able to math is a very important skill/tool in datacenter operations!
    - Practice makes perfect.

- Physics!
- Knowing proper energy requirements to operate the datacenter is crucial.
- Ensuring the facility and equipment can handle fluctuations in power is required to ensure continual operation.
- More math!
- Thermal Dynamics!
- A lot more thought and planning needs to go into cooling a datacenter than just ‘lets get the heat out.’
- Hot aisles and cold aisles!
    - Where do you want to keep your pizza warm?
- What type of cooling do you want?
    - Air?
    - Water?
    - Refrigerant?
- Meteorology!

- Weather matters!
- Knowing what kind of fronts are en route will allow you to ensure the facility is ready.
    - This could include:
        - Raising/lowering the HVAC output.
        - Getting (de)humidifiers prepped.
        - In extreme circumstances, bracing for impact!

- Geography!
- Location! Location! Location!
- Where you put your datacenter is as important as what you put inside it.
- All locations will have some sort of natural disaster potential.
    - Know what kinds of disasters can happen where your datacenter is located and take the appropriate measures to be ready for when they hit.
- Nephology!
- No matter where you look, there will be clouds!
- There are many ways for datacenters to scale out into the cloud.
    - SaaS, IaaS, PaaS, DRaaS, XaaS….
- Just because it’s there doesn’t mean you have to use it.
- Safety!
- Arc Flashes are (not) fun!
    - Especially in Russia!
- Make sure you are always following safety code.
    - They are there for a reason!