BRINGING THE CLOUD TO THE EDGE

Terence Johny
Solutions Architect
EMC
Terence.Johny@emc.com

EMC²

Table of Contents

The Premise

There has been a paradigm shift in how humans interact with technology today. We have embraced the prevalence of the omnipresent smart devices that control important aspects of our daily lives. With smart wearables', driverless cars, talking refrigerators, and more, the **Internet of Things (IoT)** has changed the landscape of the relationship between humans and technology.

The **"Cloud"** is considered the powerhouse that will fuel and support the expansion of IoT, which continues to gain momentum as vendors and enterprises begin to embrace the opportunities this market presents. According to new research from International Data Corporation (IDC), **the worldwide Internet of Things market will grow from $655.8 billion in 2014 to $1.7 trillion in 2020 with a compound annual growth rate (CAGR) of 16.9%** [11] with devices, connectivity, and IT services taking a majority stake in the IoT market. This emerging wave of end-computing deployment requires mobility support, geo-distribution, location awareness and, most notably, very low latency. Will the Cloud be able to provide these features? Or maybe, the right question to ask is if it will be able to sustain the expected growth of IoT, with billions of devices communicating over data shared across inter-clouds while providing the quality of service that we have come to expect over time.

In this paper, we pore over the existing Cloud Computing landscape and contemplate its place in the **"Things to Come"** era of computing. We look at new hierarchical distributed architectures that extend from the edge of the network to the core of the cloud and delve into the idea of extending the cloud and bringing it to the edge of the compute endpoints. Something that is now being called, **"The Fog".**

The Present Perfect Continuous

The cloud symbol was initially used in the 1970's to represent a variety of computing equipment and network components in the original ARPANET, one of the early predecessors to the internet itself. While the word "Cloud" would continue to float around making occasional infamous appearances, it would be almost over three decades before its ascension into the limelight.

Since its establishment as the *de rigueur*, highly scalable computing platform, the Cloud has emerged as a behemoth technology/service model that has disrupted the traditional meaning of technology services. Increasing user demand for elastic provisioning of resources coupled with the perennial and on-demand access to data, cloud computing has been recognized as the emerging technology to meet such dynamic contingencies.

Today we live in the world of mega clouds, mini-clouds, hybrid clouds, and even inter-clouds. Making things even more interesting is the revitalization of the industry with a plethora of subset and superset trends enabled by the maturation of several emerging technologies. This potentially breakthrough amalgamation of technology, processes, services, and raw data and their interaction with people has opened a new frontier of a data-driven future that is currently pushing the tech industry into one of its most disruptive phases in history.

One of these has been the recent surge of connecting endpoints into the internet and expecting intelligent data services around otherwise mundane devices, such as a thermostat. This we are now calling the era of "**Internet of Things" (IoT)**.

The IoT can be simply explained as a future state where internet-enabled smart objects connect to each other providing the ability to talk and share valuable information linking their environments into a global mesh of information. It can even be considered as the future of the internet itself.

The IoT Math problem . . . .

As of early 2016, the Earth's population is close to 7.4 billion and growing. There were approximately an estimated 13 billion connected "things" in 2015 alone.

Figure 1: "The World Population Counter" as of February 08, 2016, 1:39:25 PM

By **2020**, we expect by some estimates close to over **200 billion** connected things.

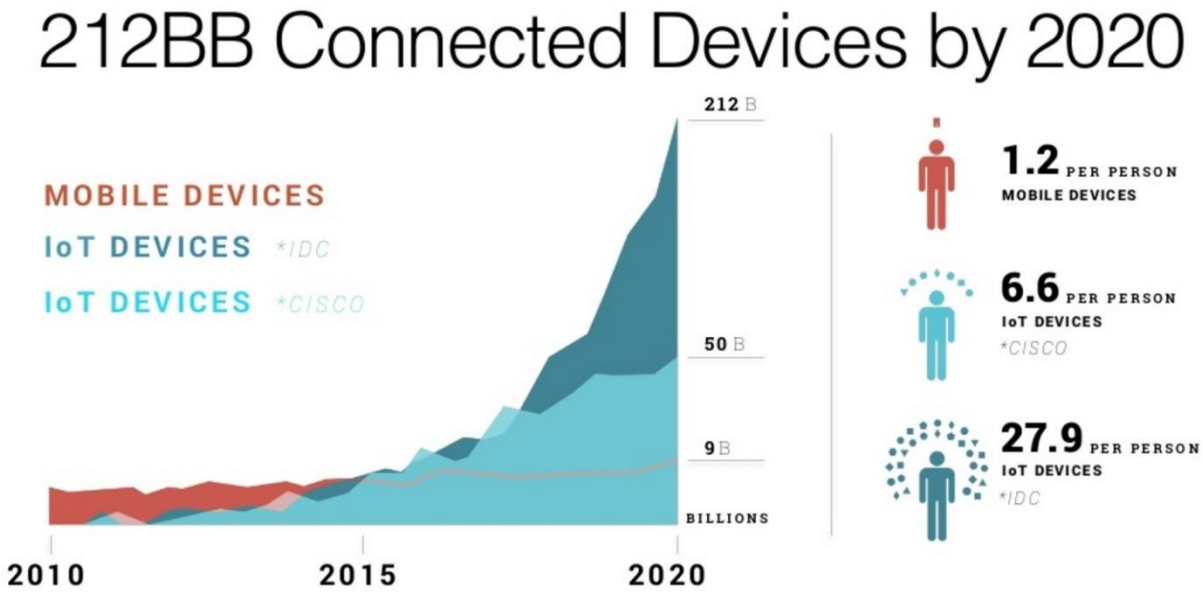


Figure 2: Internet of Things (IoT) and its predicted scalability.[19]

One of the outcomes down the line for IoT will be its gradual metamorphosis into the **"Internet of Everything"** where not only would "things" be interconnected but also people, applications, and processes would blend into an amorphous network of interconnectivity.

The IoT market continues to forge ahead with growing interest and momentum propelling growth exponentially over the next several years. However, we are moving into a complex ecosystem, with multiple layers and numerous market players.

The IoT trend and potential has been rightfully identified as the perfect storm that will engulf technology this coming decade. While other transformative forces in cloud, big data, and wireless technologies will shape their own mini-storms, it will be the incredible rate of evolution and the level of innovation that IoT will come to demand that will set it apart. This further adds evidence to the causality of its eventual convergence into a single transformative force that will shape technology's future.

Depending on the context and the personality, IoT is almost always defined in different ways. However, the underlying theme has always been about an all-encompassing global phenomena linking a myriad of aspects of life - from smart homes to smart cities, smart cars to smart roads, to devices that can track the vital signs of the human body, transmit that over a network and send the data to a doctor's cell phone as an emergency action notification. IoT has unleashed a vivid, imaginative process that now enables its architects at each level to dream up a million possibilities of its application in real time, always connected, always knowing intelligent circuits that are designed to enhance the human experience.

Some advocate the mobile computing platform as the "eyes and ears" of the applications connecting the Internet of "Everything" while some speak of a class of innovations that do not exist today as an indication of things to come.

Further down the Rabbit hole . . . . .

It is very easy to limit IoT to things and people and miss the point by a mile. How does IoT distinguish itself from merely being a rebranding of the existing Machine To Machine (M2M) market today? M2M solutions have been around for a while and already use remote connections to and between devices and the Internet but that's where the similarities end. While M2Ms' are siloed ecosystems, the IoT represents an evolution trying to bridge disparate systems into an expansive fabric that adds a value that is missing from the existing M2M stack.

IoT in its evolution…

Most products today come with embedded processors mostly with "command and control" functionality, making objects "smart." Once a product or device becomes smart, its next evolutionary phases are charted out that lead to the eventual ecosystem it will come to live in, very similar to evolutionary biology. The next logical step for an intelligent device is to build the capability to reach out to other smart devices in its vicinity through remote communication. This is an endeavor that is most natural, in order to take intelligence to the next level which ushers in data sharing. Once you can communicate and control remotely, automation is the next big leap. The goal is to be able to automate things based on set preferences and make things happen without any intervention.

IoT in its culmination…

Picture this; you're driving home after work, your estimated time of arrival is calculated by your smart car based on location tracking, traffic, past routes, and your driving style. The output of this calculation is posted automatically to interested parties that you have authorized based on preferences set in the past. Your smart watch has been tracking your body temperature all day and has synched its analysis with your car that then sets the climate control to a comfortable temperature before you even get into it. It also checks the local weather around your neighborhood and has sent this data to the thermostat in your living room. The thermostat has your ETA and has adjusted the temperature to be at a preferred setting by the time you are to arrive. Once in the proximity of your home, your garage door has detected your car's unique ID a few feet away and has already begun opening. As you walk past the garage to your front door, the sprinklers have been delayed by an interrupt from the motion sensors around the house avoiding getting you drenched in a pre-shower. Your front door lights are now turned on. Your door detects the encrypted key generator app on your smartphone and has cross-checked for any unknown fingerprints that may have accessed your phone to ensure it really is you carrying the phone in your pocket and automatically opens up. Inside the temperature is perfect, your

favorite Beatles track is playing in the background, and your beer is just the way you like it. All this without any repeated human interaction.

This augmented reality is maybe approaching us faster that we want to believe. But for this vision of IoT to achieve maturity and truly create value we would need to agree on a common platform for cross-functionality. While it's anybody's guess as to how IoT will shape the application space with hundreds of applications being considered and identified by different industries, mostly it can be boiled down to two primary primitive roles that have been around since the Neanderthals. **"The Hunters" and "The Gatherers"** [T]

The Hunters [or Trackers] – In this role, applications will primarily provide remote tracking, control, and routing between millions of interconnected "aware" devices, interacting within their ecosystem, extending the automation and "Machine-to-Anything" (M2X) communications that can help simplify people's lives.

The Gatherers – Applications in this role would be all about leveraging the data that gets collected by the end nodes and mining for trends and behaviors that can generate business value. Disciplines like Data Mining, Big Data Analytics, and Data Science would be heavily invested in generating this potential value.

While the possibilities are exciting for both industrial and consumer spaces, deployments will need further thought since today IoT is mainly comprised of disparate verticals. While systems may share platforms or protocols, we still have to acknowledge the lack of a horizontal connection that can provide easier cross-application integration.

The “Horizontal-Vertical” Dilemma [T]

Rapid IoT deployments will stipulate neutral, cross-platform integrations that are very organic in nature. As this ecosystem builds out it becomes important to address market approach and distribution.

Furthermore, most IoT offerings currently are delivered through tightly coupled systems that follow a vertical, domain-specific approach to service delivery. This is not completely unexpected as vertical growth is mostly the initial route new technologies and service delivery models take because it is more focused in scope and helps in the initial acceleration of market adoption.

But sooner rather than later we have to be able to supplement these with a horizontal structure. Look at how the internet was born or even the growth of the M2M market. But understanding the need for horizontal and vertical application balance is the key to sustaining the exponential growth that is expected.

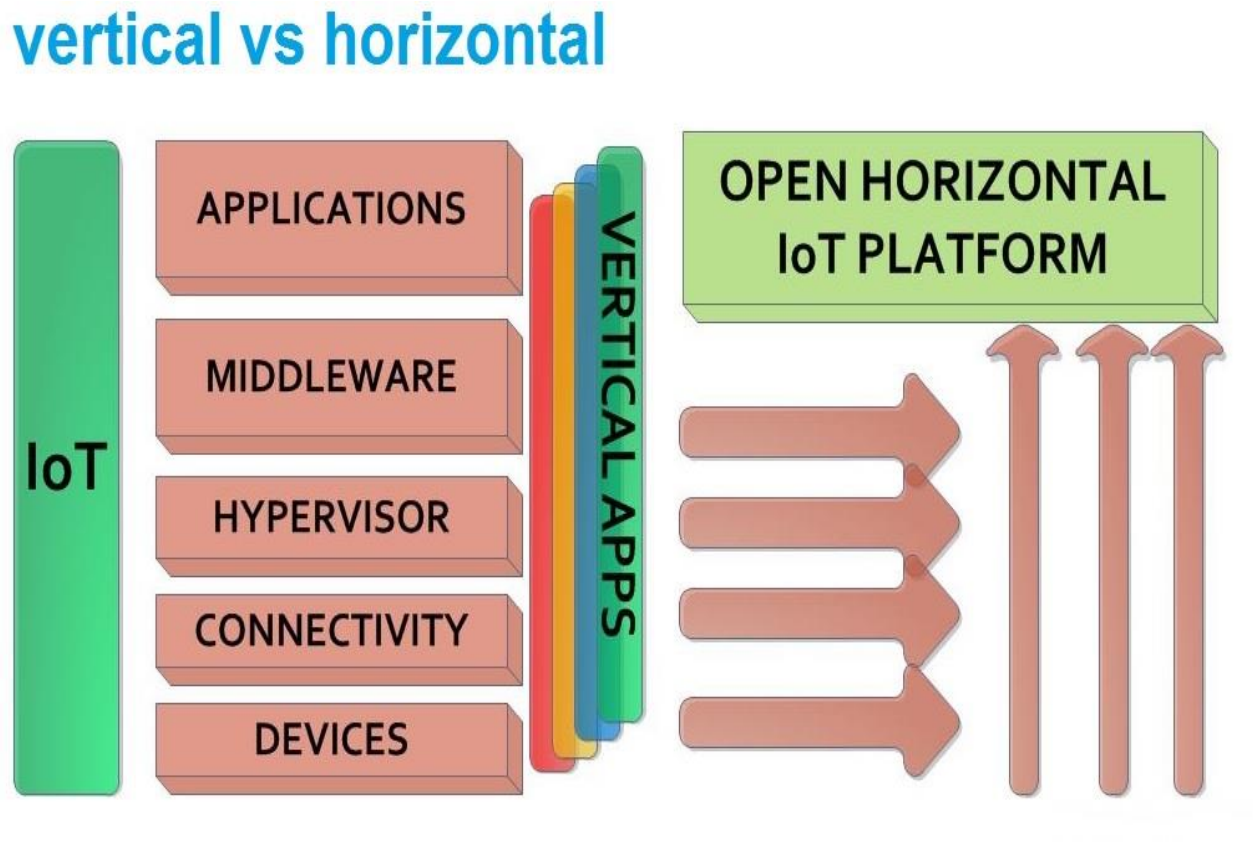


Figure 3: Vertical vs Horizontal IoT [T]

IoT has gained significant vertical traction in the last couple of years. There has been a surging vertical sprawl with domain-specific players developing one-stop-shop solutions so that they have something to generate a revenue stream with, today, while the market matures. However, the biggest drawback with these verticals is that they are only as good as their use case.

Suppose that a user wants cross-platform integration. For example, an IoT-enabled "smart-home" owner wants the temperature control system to cross-integrate with the security system, both that exist on separate IoT vertical in the present, so that a sudden rise in room temperature can be detected by the former system and could indicate a fire and alert the security system to dial emergency services. Well, with vertical-only IoT, this won't be on the horizon anytime soon.

Ergo, horizontal models become a necessity in the IoT world, so that multiple solution providers can capitalize on a common framework that can allow a common messaging bus for devices and services to share information and resources easily. The horizontal approach makes innovation easier and allows for rapid proliferation of new applications and businesses, but it needs to gain considerable traction before it can pay off on its assurances.

This is also, to some extent, the same discussion as "Open Source vs Proprietary" IoT platform development. Some hail open source as the key to the development of IoT. While the Open Source IoT Platform will definitely spur the rate of innovation and IoT markets mature over the next few years, achieving the Horizontal-Vertical balance in time, will ultimately lead us to the promised land of the grand IoT vision of the future.

DATAVERSE [T] - The Digital Universe

Billions of interconnected things, billions of application instances, billions of intercommunication processes and threads, with billions of interconnected **species**, yes, species, inevitability.

It won't be just humans on this information superhighway. Imagine your pet dog wearing a "smart" collar that can track his vitals, check anxiety levels, and syncs periodic statistics of his physical activity levels to your phone. Or a pet fish with an electronic skin tag that checks its health, the water's oxygen quality and bacteria build up in its environment and lets you know it's time to clean the fish tank.

Now extrapolate these scenarios to ocean life or insects that carry minute data sensors that can gather information about air quality or pollen pollution. Some of these we have been doing for a couple of decades now, but with obvious limitations.

A picture speaks a thousand words and here it is …

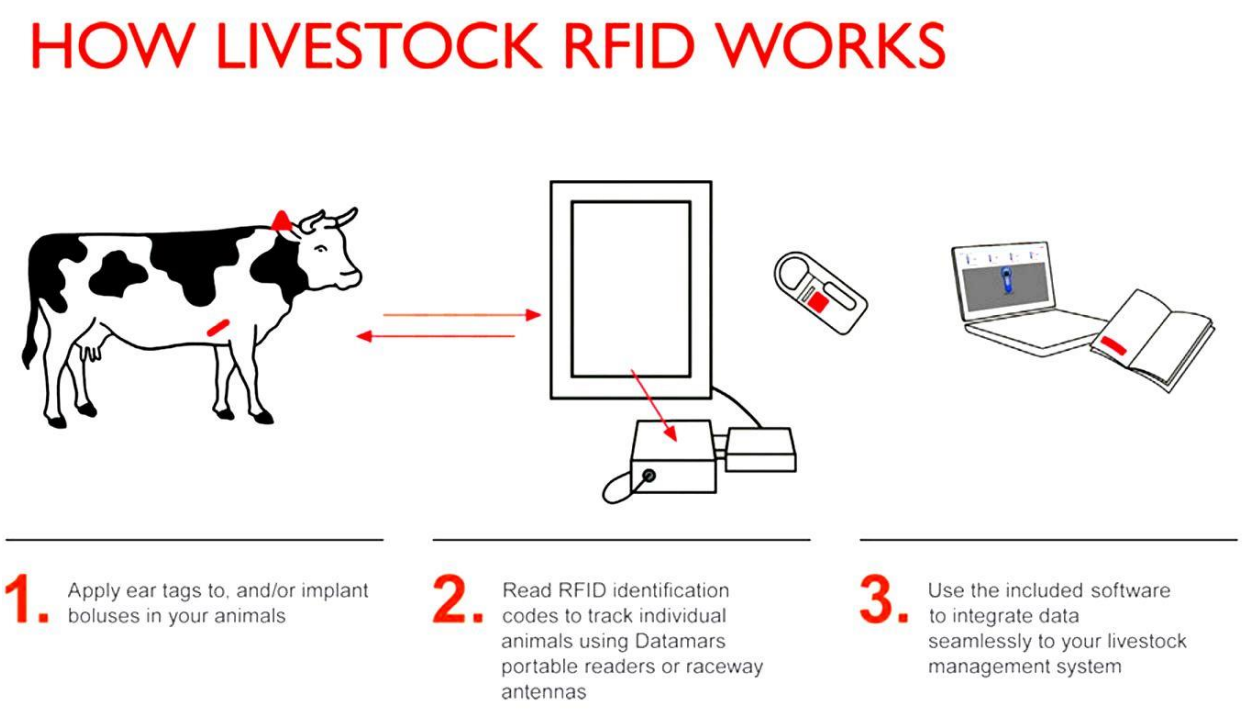


Figure 4: The current state of Animal Tracking using RFIDs [16]

We would truly be living in a digital universe- a **"dataverse"**

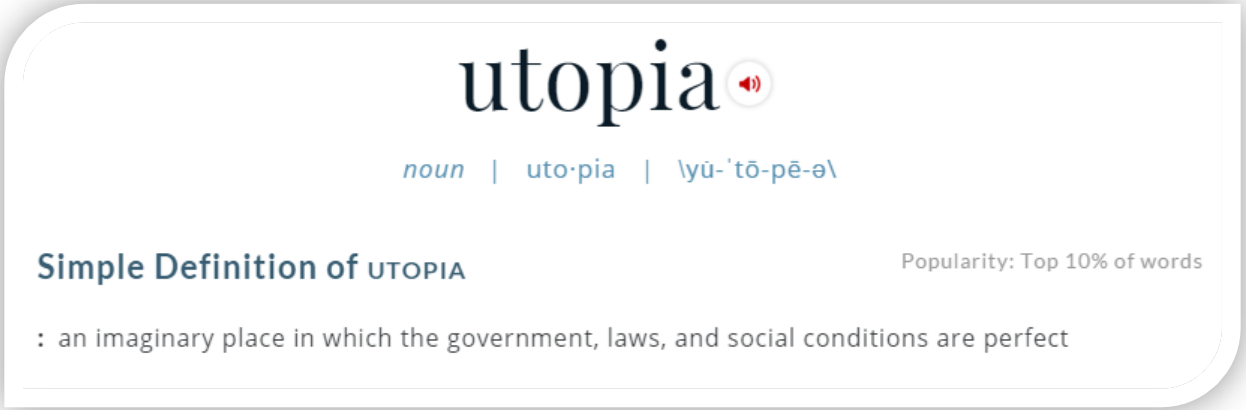


Figure 5: Coined in *Greek: οὐ* ("not") and *τόπος* ("place") and ironically means "*no-place*"

This **Utopian** dataverse will by definition generate data on a massive scale; "**Big Data**", characterized by the **"four Vs' " Volume, Variety, Velocity,** and **Veracity**. In that, it will come in large amounts (volume), will be treated or untreated (variety), arrives at the speed of light (velocity), each providing a spectrum of varied value (veracity).

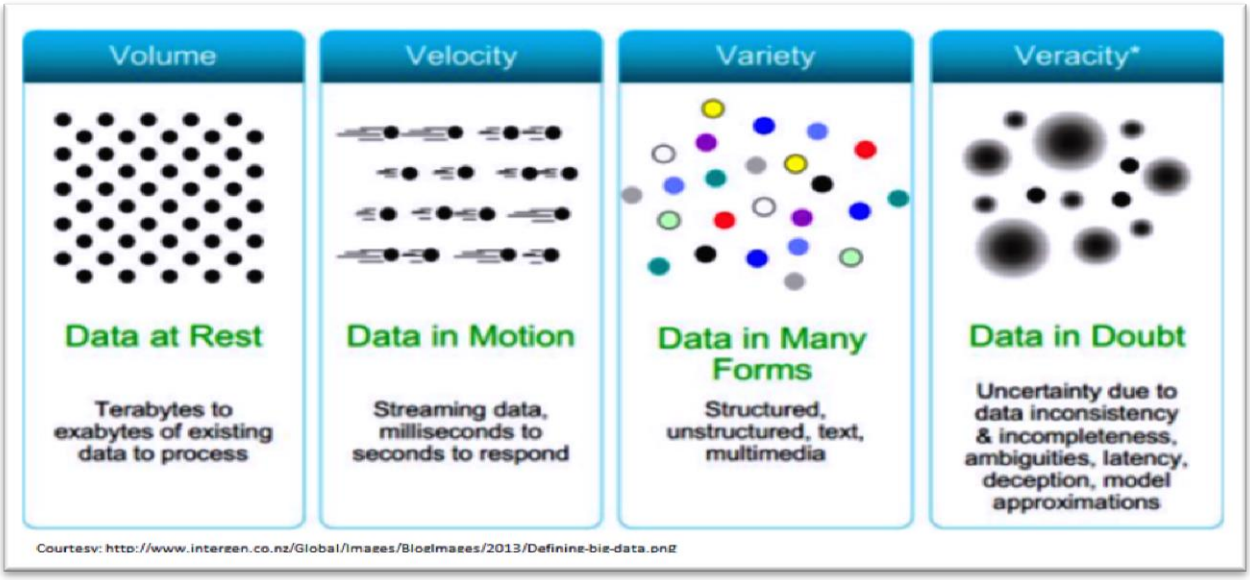


Figure 6: Illustration of the Big Data 4 "Vs' "

This type of data will be contemptible with traditional data management systems and has been for a while now, which is why a range of new platforms with additional intelligence has evolved to service this segment. Hadoop, for example, deals with this massive data processing through a hierarchy of job trackers that track tasks that in turn perform on the data nodes. All this architected into a very neat multi-node, parallel-processed, open-source framework that can be run on clusters of commodity hardware. It provides massive storage for any kind of data, enormous processing power, and the ability to handle virtually limitless concurrent jobs.

Undoubtedly, IoT and Big Data are intimately connected. Some estimates project just the data generated from sensors and similar embedded systems to grow from 2% currently to over 10% in the next 5 years alone. The key challenge, of course, would come down to separating **"the data"** from the data. The true ingenuity lies in the ability to identify data that is valuable within this diverse dataverse. The target data will have to be of high-value with a low footprint that serves to be insightful and transformative and can be converted to actionable business value.

The Visualization challenge: It would be sacrilegious at this point to not hover over the fact that many times, businesses and technologist are so consumed by methodologies and processes that accounting for perspective is missed. The power of visualization is not to be underestimated. What good is it if you have all the data and it makes sense, but only to you?

In today's hyper-competitive environment, to find and analyze the relevant data is only half the battle. Shaping that data into sensible decisions is equally critical. Data visualization is becoming an increasingly important component of analytics. You have to be able to rapidly process sheer volumes of data parsed at increasingly challenging degrees of granularity while providing near-realtime insights. Proper domain expertise is crucial to set it in the right context while ensuring the data is clean. Finally, the most overlooked part is being able to display meaningful results that can clearly communicate trends and outliers.

Venturing into the FOG – The 4th Platform

When all is said and done, IoT will come down to speed and real time results. If you peruse through the umpteen number of IoT use cases being floated around the internet today, most, if not all, are not of much value if the decision-making apparatus – that is, the point of value-creation – has a delay that exceeds the threshold beyond which the solution simply erodes.

An autonomous STLS (smart traffic light system) is no good if it has to traverse a signal path that does not provide the latency requirements to be able to function as a real-time accident prevention monitoring hub.

Referring to the previous example of augmented reality in the "IoT's culmination" section, the whole cyclic **feedback loop effect** that a smart watch gets, by its interplay with the connected ecosystem it operates within, is void, if it's unable to generate a context of the analysis and its locational awareness is affected by the fact that pre-processing has to traverse four or more layers upward toward the cloud.

The IoT Reference Model emphasizes the importance of the Edge Computing layer.

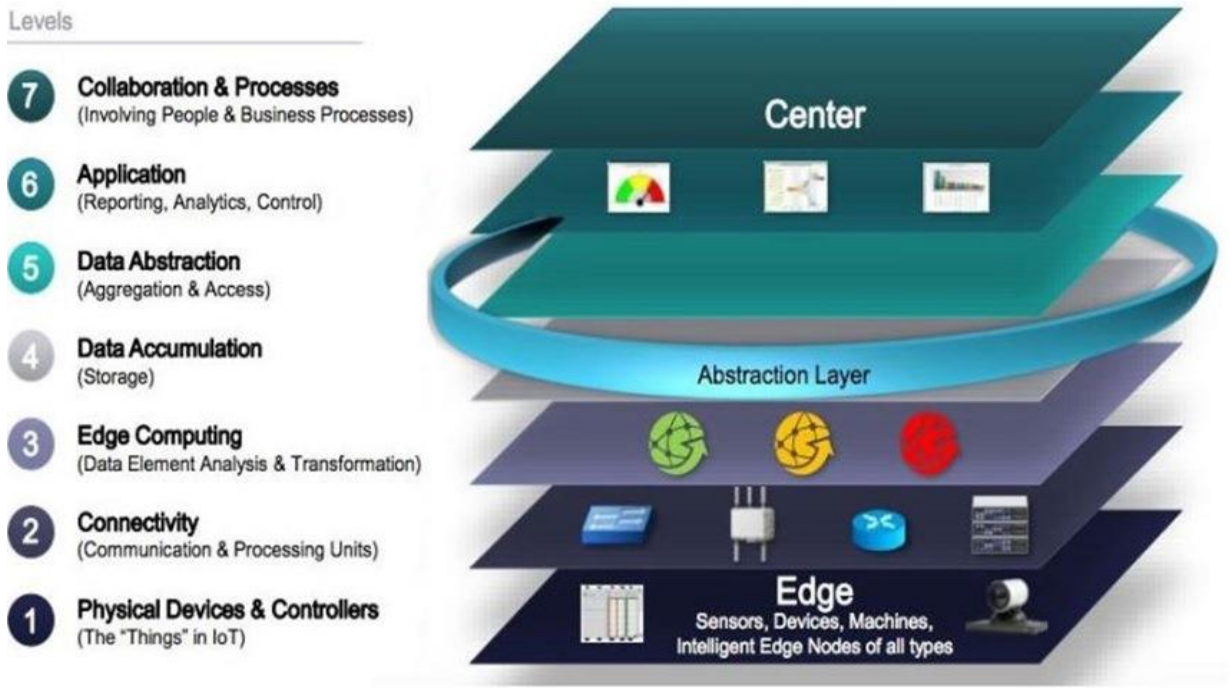


Figure 7: The IoT Reference Model [Source:Cisco]

The IoT decision-making process is as local as it is centralized. This means that IoT processing has to depend on decision making in two distinct places. [14]

One of these places will definitely be the Cloud where ideally data traversing through multiple filters from multiple sources is analyzed for insight and value. The second place more recently has been identified as localized decision making “PIN”s’ (Point in Network); which can pre-process and filter the events within its locality and can make quicker decisions, much faster than the cloud repository alternatively, in many cases, eliminating the need for the cloud to be involved in the pre-processing phase of the data and utilizing it for more advanced analytics in its post data clean up phase. Figure 8 illustrates the machine learning algorithm used for supervised learning.

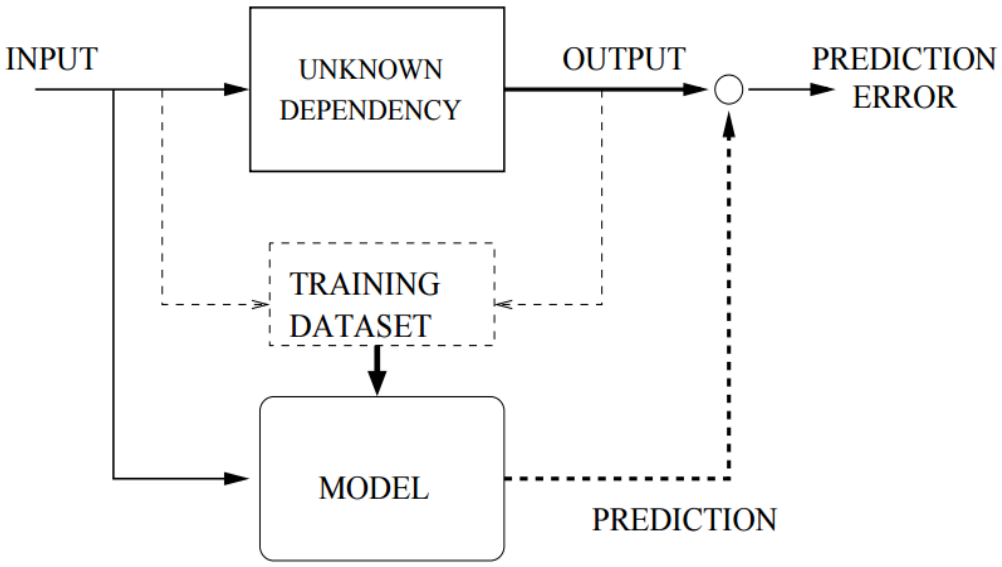


Figure 8: Machine Learning for prediction: Supervised Learning[3]

Statistical machine learning is the discipline that deals with the problem of supervised learning, where we have to infer from historical data the possibly nonlinear dependence between the input (past embedding vector) and the output (future value).[3]

The model predicts dependencies that in real-time can be inputs for the next iteration of predictive calculation thus closing the margin of error. Noise and jitter have to be accounted for within such models as well.

A basic system like a fire alarm typically evaluates a fixed threshold in a narrow data space and reacts in near real-time. With Edge Intelligence, the same systems can also track and record events that can be correlated against the businesses' operational execution. This, in turn, can offer a new perspective to improve business and operational models.

A significant differentiator for IoT/IoE, when compared to centralized on-premise processing like Cloud or Big Data, is that it will primarily be bi-directional. The **"'trackers and gatherers"**[T] will have to be tied into a closed loop at least for the initial immediate value creation process.

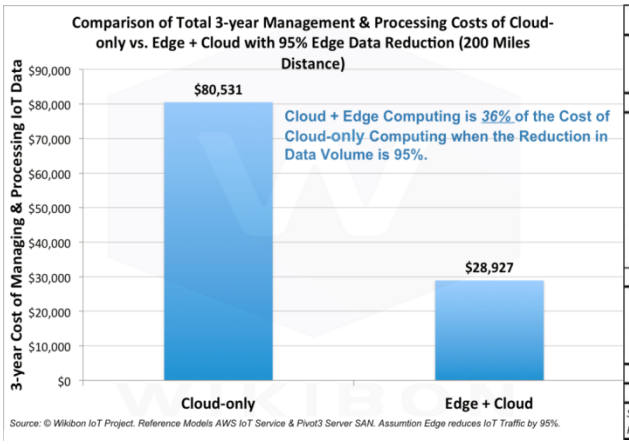
Edge Intelligence will accelerate the response and awareness and can act as an intermediate layer between the IoT device and the cloud. By its proximity to the input it can be a facilitator for deeper and faster business insights leading to better service levels. This edge intelligence will have to be a notch higher than the traditional data conduits in M2M systems. In addition to in-stream analytic capabilities, these intelligent edge devices will have to provide context awareness, geo-distribution, and real-time response to the ingested data.

Fog computing's intrinsic value will be in creating this intermediary layer that can sit in close proximity to the edge of the network and extend up to aggregation layers that can plug into the core cloud services and features thus providing a seamless extension of the cloud from an end-user perspective.

Making the Business Case

How does a cloud-centric architecture to manage IoT compare versus an Edge computing model. Research published by Wikibon[12], designed to compare the cloud-only architecture approach with a low-cost converged "edge computing" architecture, concludes that IoT systems will be safer, more reliable, and lower cost while being more functional if using a hybrid of Edge computing and Cloud.

The case study used was of a remote wind-farm with security cameras and sensors. The model was designed based on certain assumptions; the distance from the wind-farm to the data center ~ 200 miles, video quality from the security cameras compressed by about 100:1 which is in line with the minimum quality requirements for the security cameras, the sensor gathered a small amount of data ~20 bytes every three seconds from 100 sensors.



Cost Comparison of Cloud-only Management & Processing vs. Edge + Cloud

		Cloud-only Processing			Edge + Cloud Processing		
		# Sensors	Video (MPEG)	Total	# Sensors	Video (MPEG)	Total
	# Inputs	100	2		100	2	
Edge Transport Costs	Readings/minute	20	60		20	60	
	Bytes/Read	20	672		20	672	
	GB/Month	1.7	3.5	5.2	1.7	3.5	5.2
	3-year GB	62	125	188	62	125	188
	Local Consolidation	0%	0%		95%	95%	
	Miles	200	200		200	200	
	$/GB/Month/Mile	$2	$2		$2	$2	
	Monthly Cost	$691	$1,393		$35	$70	
	Edge 3-year Transport Costs	$24,883	$50,165	$75,048	$1,244	$2,508	$3,752
Cloud Costs	AWS $/Million Messages	$5	$5		$5	$5	
	AWS Message Size (Bytes)	512	512		512	512	
	AWS Cloud Cost/Month	$86	$10		$4	$1	
	AWS 3-year Cloud Costs	$3,110	$373	$3,484	$156	$19	$174
	On-site Equipment 3-year Costs			$2,000			$25,000
	Total 3-year Costs			$80,531			$28,927

Source: © Wikibon IoT Project. Reference Models for Cloud is AWS IoT Service. Reference Model for On-site Server SAN is Pivot3 with an Open Source Time Series Database On-site & in the AWS Cloud

Figure 9: Cost Comparison: Cloud vs Hybrid (Edge + Cloud) Architectures

The research compares the 3-year costs of an industry leading cloud-only solution compared with an Edge + cloud approach. With a distance of 200 miles between the wind farm and the cloud, and with an assumed 95% reduction in traffic from using the edge computing capabilities, the total cost is reduced from about $81,000 to $29,000 over 3 years.

The cost of Edge + Cloud is about 1/3 the cost of a Cloud-only in this case.

The Cloud-Fog Interplay

OK, let's state the obvious. Fog and Cloud are separate and do not emulate each other per se. The Fog is more like a vehicular enabler for the modern geo-distributed intelligent ecosystems like IoT that are evolving around us. The delta is significant in terms of the topology, communication end points, and latency requirements to name just a few.

With Cloud Computing, there is data center endpoints that are well defined and data center interconnect protocols (DCI) are well understood. Applications are specifically designed with these aspects in mind. Data center application developers have worked under the presumption that the applications begin at a certain securitized level due to the infrastructure and devices that they would cater to.

With Fog Nodes in play, this paradigm changes, and there is a strong demand for cooperative exchange between the Edge and the Core.

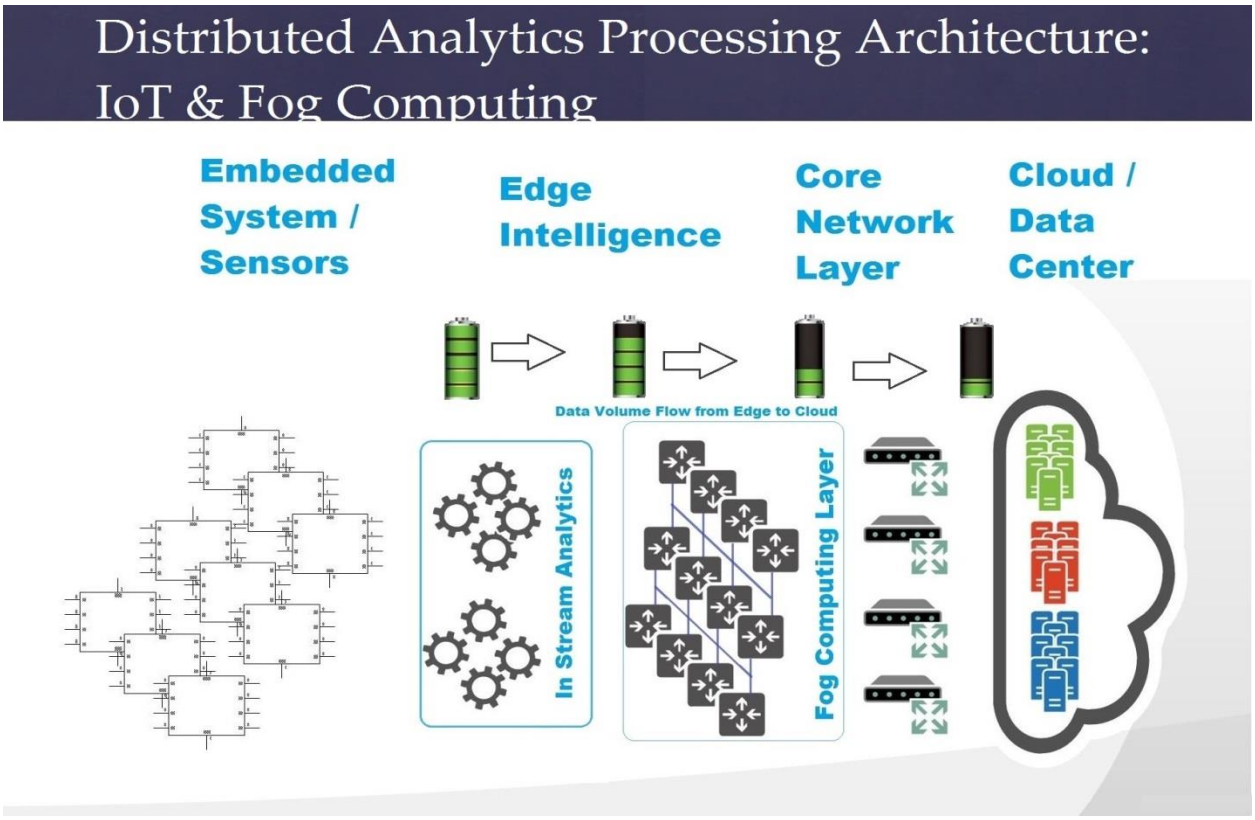


Figure 10: Hierarchical Framework for IoT and Fog Computing [T]

As an illustration, Figure 10 depicts the hierarchical architecture for distributed analytics processing. It tries to capture at a high level, the interplay between the fog and the cloud layers. The depiction illustrates the new hierarchical framework that most IoT- and Fog-enabled architectures will resemble. The endpoints talk to the fog layer in-between that relays the data in all directions based on design and need. The volume of data that will get generated and the amount that each layer will cater to will taper going toward the cloud spreading the workload to lower tiered devices at the edges with varying levels of intelligence, processing power, and security.

Fog Nodes implement local processing, storage, and networking, bringing intelligence closer to the data source. As defined by Cisco [2], “Fog Computing is a paradigm that extends Cloud computing and services to the edge of the network”. Proximity to end-users, dense geographical distribution, and support for mobility are the main distinguishing characteristics of Fog Computing.

The Cloud will still continue to be the powerhouse that will fuel the greater expansion and analytical prowess of the larger system-wide function. A cloud-first mentality may even be preferred in specific cases. Fog Computing will be complementary to the Cloud and will act more as a bridge to the end nodes. Applications that demand features like low and predictable latency, rapid mobility, and that are distributed geographically across wide areas will be most apt to adopt the fog computing model. Multiple fog nodes spread across the application's functional domain space will be able to provide the near/real-time feedback that will be the norm that future smart systems will come to require.

A smart car that is traveling at 65 mph will have traversed multiple wireless zones and dead spots. If modern IoT is all about providing real time value of the smart car experience, the fog node architecture will be able to provide the coverage required for high speed moving targets with the analytic results required for its decision loops. This experience won't be as feasible with data moving in and out of the cloud to provide these results and that's where edge intelligence will excel in the future.

Fog Computing Ecosystem and Architecture

The discussion so far warrants the pursuit of trying to define, albeit at a high level, a software architecture that can elucidate the fog node and its interaction with its ecosystem. We have already looked at where in the system hierarchy a fog node would fit in within the fog computing model. Now it would make sense to see if we can define what that node at the fog layer can be architected like. Figure 11 is a representational diagram of the components in a fog node and its interplay across the end points.

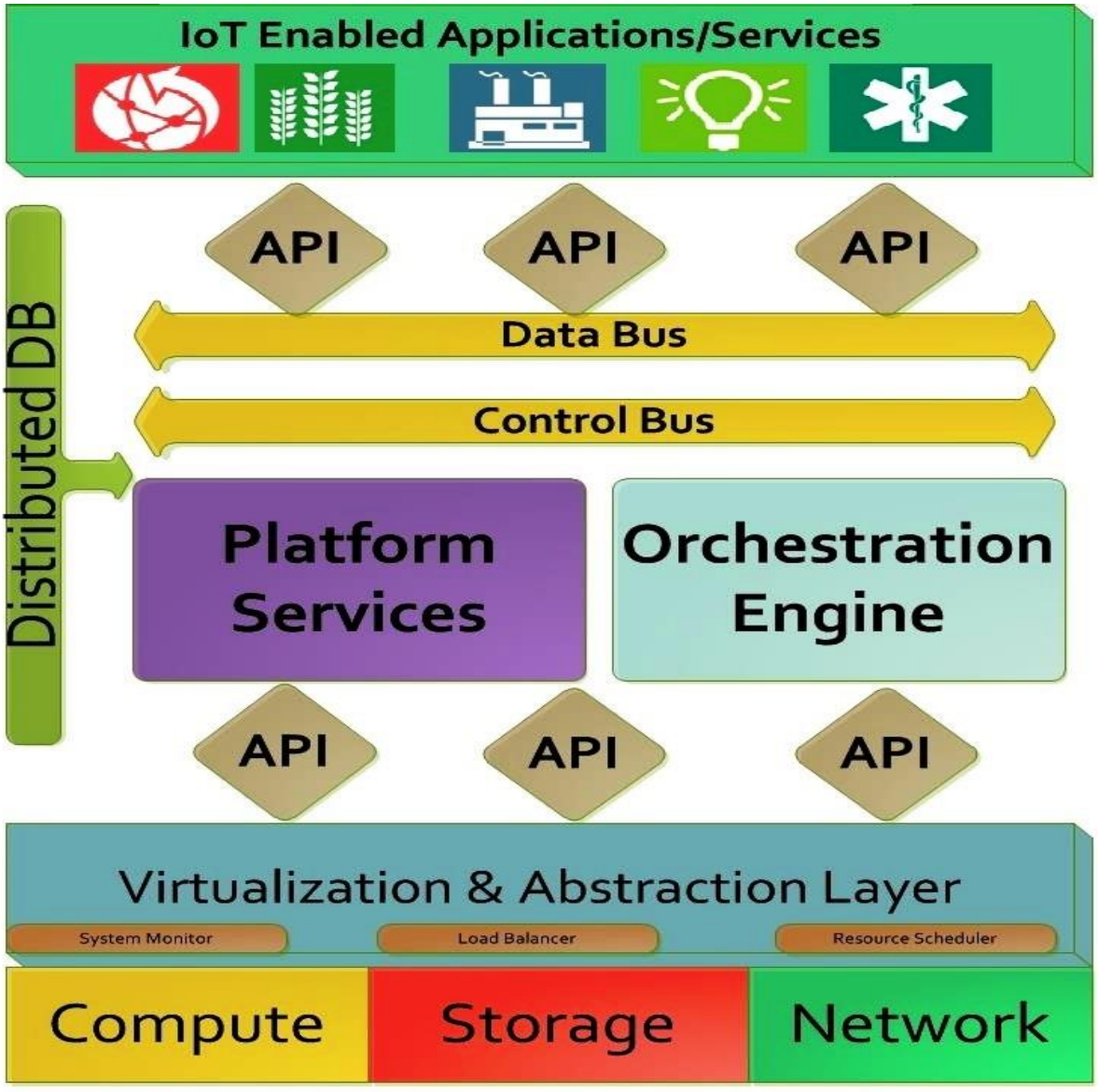


Figure 11: Fog node software architecture and ecosystem [7]

Physical Layer: Fog nodes will be deployed in a variety of environments. This would mean that they would have to be scalable in terms of their computational performance, network characteristics, and storage capacity. They can be net new standalone boxes or nascent compute, network, and storage resources available in existing network elements that can be leveraged to create fog nodes as ancillary units within them like switches, routers, storage appliances, or edge servers. The Fog Node network infrastructure has to be versatile as well, and there has to be an open platform architecture that lets fog nodes utilize accessible interface ports on switches as network access points while at the same time be flexible to multiple connectivity options, ranging from high-speed links connecting to cloud or on-premises data centers or wireless access technologies.

Hypervisor Layer: A highly virtualized layer with enhancements and security extensions built into a capable hypervisor. It will have inbuilt modules geared toward software management, network-function virtualization (NFV), file systems, database management systems, and cryptography processors. This layer abstracts the fog node physical resource layer and exposes generic APIs that can interface with the hypervisor services for monitoring and provisioning. Virtualization will enable multi-tenancy which is critical for the isolation for tenant-specific allocated resource shares. This enables the multi-tenant hosting services on a single fog node instance providing security. The abstraction of the physical layer is important for the fog nodes' dynamic capabilities, to spin up and down fog instances or container services based on real-time resource requirements, and provide short-term storage, analytics, and decision support.

Platform Services: This layer can be a virtualized instantiation of the fog node that can provide platform services that enable intelligent and interconnect services to the fog node within or across nodes. It can interface with APIs to support policies for different tenant services. This software-defined layer will provide fog services like authentication, location service, service management, data management, analytics engines, cloud agent, and communication services providing a generic framework on how data is shared across the platform. In the fog node architecture provided in Figure 11 the platform services and orchestration engine are decoupled from the functionality that they cater to. This is important to support the core function of the fog node; to act as a pre-processing hub and provide intelligent services, thus filtering out what the cloud is tasked with.

Orchestration engine: Will be the power-horse of the fog node providing the core analytics and intelligence services. In addition to supporting the management of distributed services on a large volume of fog nodes, the orchestration engine will be responsible for securing the interchange through strong policy-based service lifecycle management of the fog services. The engine needs to be architected to act as the data probe functioning as the analytic engine and decision maker, planning the allocation and management of resources while providing security for distributed service management based on defined policies that can be enforced based on aggregated decisions.

Database, Message, Data Bus

A distributed database implementation can augment the scalability and fault-tolerance of the Fog nodes. It should be designed to support a distributed service architecture providing faster data storage and access.

The Fog nodes will have a message and data bus separating the control and data path. It will use encrypted channels to satisfy multi-tenancy and security requirements. The buses will glue the different physical and virtual components into a cohesive node.

IoT Applications & Services

The Fog node and its software layers interface with applications and service through external API functions that can provide control and data exchange between Fog nodes and external services. REST API can be a good starting point.

Key Value Adds with Fog Computing

Extends & Complements Cloud Capabilities: Fog computing addresses congestion and latency issues that arise due to data traffic in the cloud data pipes. It can save bandwidth as data handling happens at the edge and the cloud assumes more general computing roles.

- ***Acts as a Data Filter:*** Suppresses data that is non-consequential and extraneous and prioritizes what is important. Select data makes it into the central repositories in a data center or cloud for further processing. It can be seen to some extent addressing the issue of "Data Gravity" that will need to be contended at some point in the overall discussion.
- ***Heterogeneous endpoint Integration***: The Clouds homogenous resource centralization architecture will not sustain the proliferation of endpoints. The distributed infrastructure of IoT will comprise of heterogeneous resources needed to be managed and supported for mobility and interoperability. Fog computing components can interoperate and federate across domains.
- ***Endpoint Intelligence:*** A device that acts as an endpoint in IoT ecosystem usually has functional and physical constraints that need to be taken into consideration. If we try to enhance the intelligence of the endpoint beyond a limit, the cost vs. feasibility benefit may be overrun. There is only so much computational power that can be concentrated into a defined space. Fog nodes are optimal in this regard solving the intelligent endpoint challenges. Concentrating intelligence in endpoints while trying to use the data network between as "dumb pipes" may work for a subset of use cases, but the fundamental lack of modularity at endpoints can only be solved by fog nodes.

- ***Shifts the "Flow of Data":*** Existing Cloud Computing paradigm promotes a more innards data flow and processing model that's mostly unidirectional. Fog changes the way data flows in and out of the cloud and what role each layer focuses on. Fog provides the data element analysis at the Edge while the cloud will play a more post-data-cleansing role. Figure 12 depicts this below.

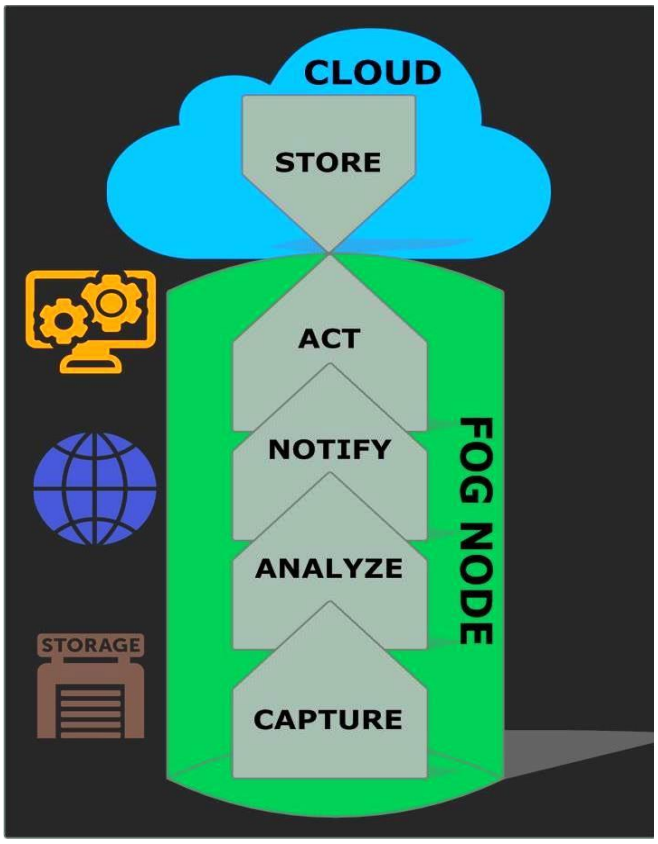


Figure 12: Data-Flow with Edge Intelligence [7]

- ***Security & Data Privacy:*** Fog computing can provide significant improvements to existing resource-prohibited end devices. It can enhance security by incorporating strong privacy preserving algorithms, robust intrusion detection, and access control mechanisms with the Fog node or instances. Shifting the data locality into Fog nodes can address the endpoint security holes due to lack of strong authentication, crypto support, or vulnerability to physical breaches.

In search of Data Harmony [T]

Prediction – "IoT/IoE" will cause more damage to people, organizations and/or communities than tsunamis or earthquakes in the next decade or so. This is an extremely macabre outlook to have and I hope I am wrong on all counts, but the prevention of such ghastly outcomes in a "smart" future is what exactly this chapter is all about.

The motivation is to try and explore how we can secure this proliferation of technology. It is not the intention to discuss security in detail and explore the various crypto-encryption algorithms or draw parallels to what we have now in terms of security with the cloud or the internet. Instead, we will try to explore what foundational framework will be needed and what scope will it have to span. While being ever mindful of the fact that there is no such thing as absolute security; hence, "in search for data harmony".

"Data Harmony [T] is defined as the balanced, congenial, and coherent existence of information at all levels of its representation and manifestation that fosters a state of consistency and is free of corruption".

Achieving harmony, data or otherwise, is no small endeavor, even if you can figure out where to start. Treating security as an afterthought at the IoT scale will be a recipe for disaster. We can agree that devices like sensors, actuators and other high-volume, low-footprint end nodes can only be equipped with so many security features before it tips the cost vs risk benefit to one side. The other important consideration is that some risks are difficult to prevent, for example, insider attacks that need to be met with appropriate detection techniques if there stands a chance to prevent it. The imminent billion end devices at the bottom most layer of the IoT hierarchy are all vulnerable points of entry into the system due to the physical and practical constraints that bind them to certain maxims of design. Fog Computing will provide a special architectural position within this hierarchy that can provide the prominent advantage of having a "Fog Node" in between this scheme of "things" and "everyone".

The security paradigm shown in Figure 13 is an attempt at elucidating at a high-level the security framework that will have to be in place as IoT systems scale. It shows how the **cost vs. risk** benefit rises as we scale up to the top of this inverted pyramid. The level of security at each layer has to cater to the each segment's risk appetite while ascending toward more robust and mature security methodologies. Many end devices that act as the input signals upward on the wire toward the Fog nodes will be placed in unsecured locations with minimal physical security

controls. These system-on-chip units are vulnerable throughout their power cycle from their boot-up process to the transient running states.

The Fog nodes and their instances a layer above will have connections running in all directions, north-south and east-west with objects, machines, and infrastructure forming a global neural network enabling services across distributed localities.

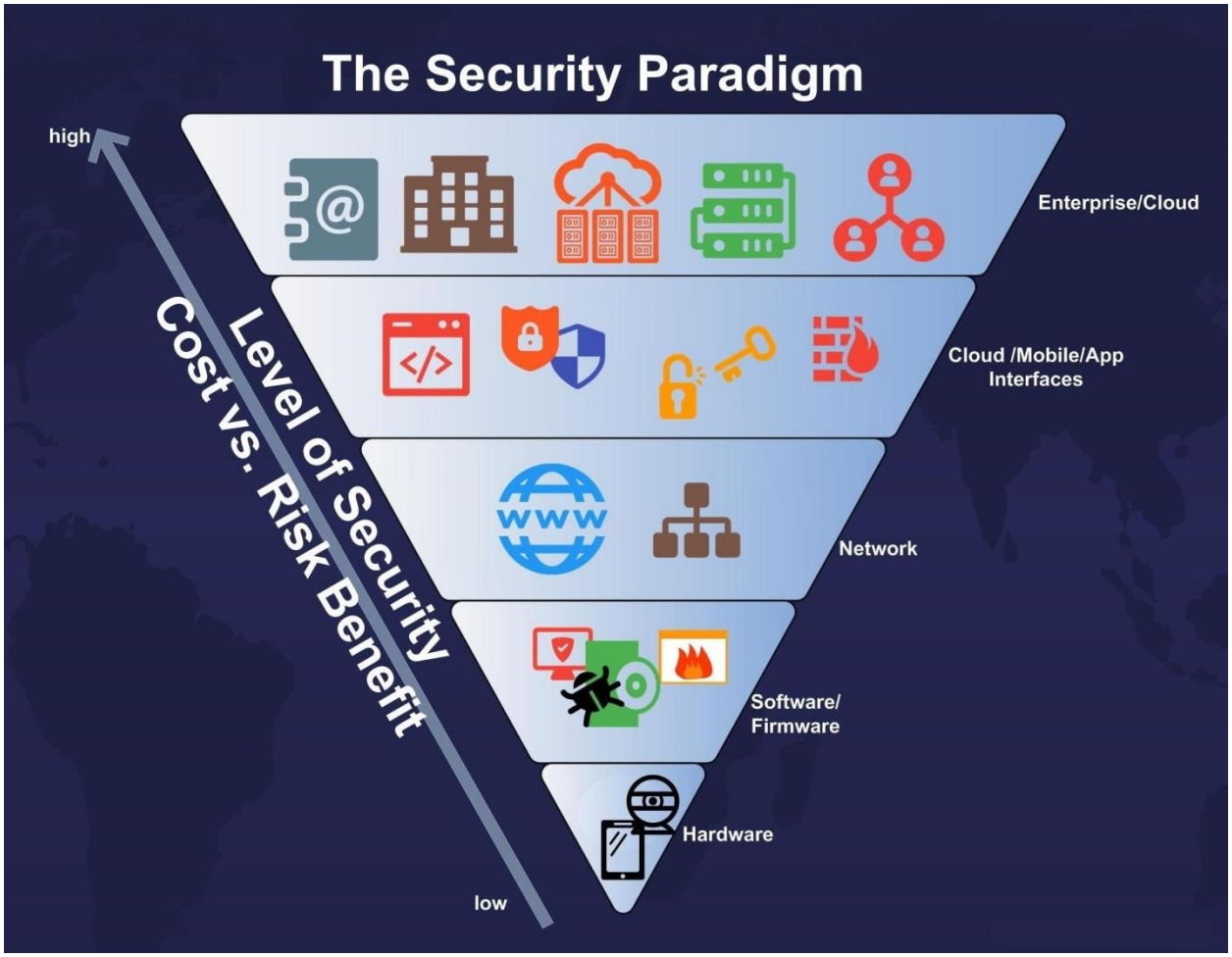


Figure 13: Various layers to consider for an end to end security solution in IoT [T]

These fog nodes can be used as proxy connections with a rich and flexible set of security features rooted in the hardware. They can also be leveraged as authentication and access point verification controllers between various fog components, fog instances, gateway nodes, and end devices while ensuring policy specifications are met for multi-tenant applications.

Fog nodes can act as security gateways that can facilitate secure data transport north of the hierarchy or even sideways, across geo-distributed nodes.

Generic Security Characteristics in a Fog Node

The fog node can be thought of as a self-reliant computation system within itself. Its security foundation depends on instantiating within a trusted computing environment with a secure policy-based mechanism. The node points will have to be hardened to be able to connect to **not-so-secure** heterogeneous endpoints while being able to provide the basic network security fundamentals of **confidentiality, integrity, and availability.**

A fog node's trust has to be able to be built from the root up extending it up the chain of execution. The hardware boots to a static predetermined trusted state with a secure peripheral device connected to the central messaging bus. Fog nodes will have to be heavily virtualized in hardware and software. The hardware-based virtualization provides a layer of security by inherent QoS controls and resource isolation important for multi-tenancy.

Firmware over the Air **(FOTA)** with differential compression provides vulnerability and patch management for endpoint security. Data security through encryption for data at rest and data in motion have to be granular at a filesystem or object level, for isolation between multiple tenant data on common physical storage.

Network interface interconnects need to have encrypted communication at line rate and support software defined network features with NFV-like capabilities. Features like netflow analysis for pattern behaviors and anomaly detection significantly fortify network layer security. There has to be mechanisms in place that can control and guarantee **"end to end" trust relationships** for remote services and orchestration flows based on identity management and secure key-management.

Finally, physical bias to the node itself has to be taken into account. Unlike data center routers or blade equipment these fog nodes are locality based and may exist very close to the exposed edge. Appropriate intrusion detection and anti-tampering features through motion sensors, self-erasing code that can wipe any data or clear memory buffers based on policy-execution mechanisms, may be present.

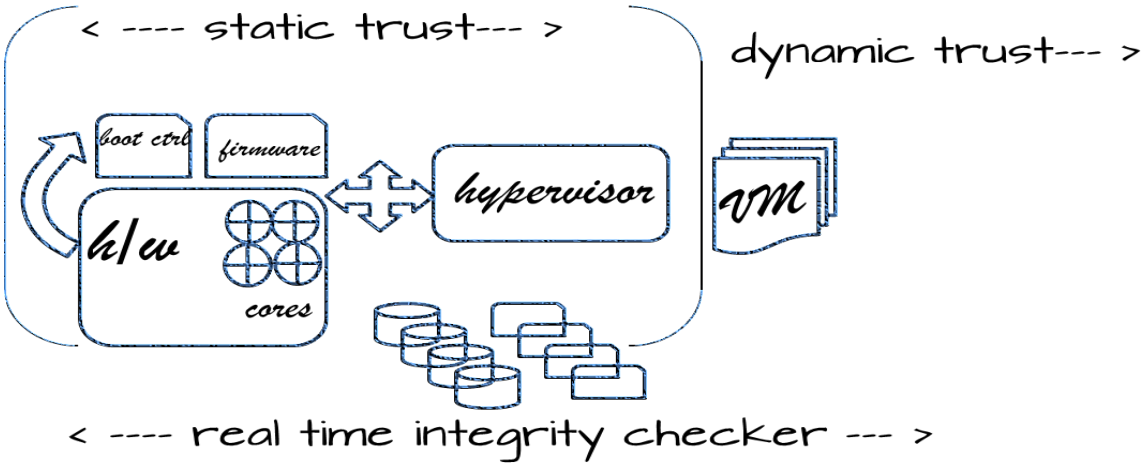


Figure 14: Fog node execution environment [T]

The Fog Computing security discussion is a double-edged sword. On one hand, we alleviate the security problems on resource-crunched endpoints by shifting it to the fog nodes while on the other hand; we introduce a whole new layer that needs to be additionally fortified. Proximity of Fog nodes to end users means it can collect more sensitive data faster and provide a context to decode that data based on path trajectories and usage patterns.

The true security and privacy implications are profound in the world of IoT/IoE and require a very serious discussion built into its foundation.

Back To the Future – The Final Chapter

"Fog Computing" has the potential to be the **4th Platform** that will solve some of the challenges arising with the grand vision of IoT and its subtle, yet profound transition into IoE. The Cloud is not going anywhere, **au contraire**; the interplay between "the fog" and "the cloud" will fuel a whole new generation of applications, services, and business models that will rake in the **"/ˈmo͞olä/".**

Fog computing's foundation will be built on leveraged capabilities of cloud and virtualization platforms, thus avoiding re-inventing the metaphorical **"tech-wheel"**, while acting as an **"intelligent bridge [T]"** that complements and extends the cloud to the edge. As IoT passes through the various stages of the technology hype cycle, it becomes evident that Cloud Computing alone may not be able to pull it off. The most compelling argument, at least for me has been; why should zillions of bits traverse across the entire hierarchy from an end point across multiple network hops up to a centralized depository like the Cloud?

It was 1977 and mainframes were going to be big forever. Then the client/server model came in and disrupted everything. More than three decades later, we decided to go back to a centralized Cloud model that solves the scalability and availability issues of the traditional client/server. IoT and Fog will change this again and propel toward a flatter hierarchy with one leg up in the cloud. It feels like we keep riding around in this "merry-go-round" of technology evolution.

According to ancient Vedic Cosmology recorded in India (~12th century B.C), it is believed that the universe follows an endless cyclic rhythm of creation and dissolution spanning 311.04 trillion years- An **oscillating universe** that renders **"creation"** timeless, much like the cycle of **"innovation and disruption"** that charts the evolution of the technology dataverse around us.

What cosmic or earthly events will unfold in the future paving the way to the direction IoT/IoE will take is yet to be revealed. But for now, **"Fog Computing"**, if thought through carefully, can be an **intelligent bridge** that will mesh the centralized and distributed worlds that will come to be.

Bibliography

1. *“Analytics, Alarms, Analysis, Fault Detection and Diagnostics” -John Petze*
2. *“A Few Useful Things to Know about Machine Learning” -Pedro Domingos,*
3. *“Machine Learning Strategies for Time Series Prediction” –Gianluca Bontempi,*
4. *“A hierarchical distributed fog computing architecture for big data analysis in smart cities” -Gerald Heffernan, et al.*
5. *“Fog Computing and Smart Gateway Based Communication for Cloud of Things” - Mohammad Aazam, Eui-Nam Huh Kyung Hee,*
6. *“The Evolution of the Internet of Things” -Jim Chase, Texas Instruments”*
7. *“Fog Computing: A Platform for Internet of Things and Analytics” -Flavio Bonomi, Rodolfo Milito, Preethi Natarajan and Jiang Zhu*
8. *“Video-aware Wireless Networks-Improving Video Performance with Edge Servers in the Fog Computing Architecture” -Chan et al.*
9. *"Fog Computing and Its Role in the Internet of Things" -Flavio Bonomi et al.*
10. *“Internet Connected Objects for Reconfigurable Eco-systems” -Raffaele Giaffreda, R. Venkatesha Prasad, Michael Koster*

Web Links

11. *http://www.idc.com/getdoc.jsp?containerId=prUS25658015*
12. *http://wikibon.com/the-vital-role-of-edge-computing-in-the-internet-of-things/*
13. *http://machinelearningmastery.com/a-tour-of-machine-learning-algorithms/*
14. *http://srinivasansundararajan.sys-con.com/node/3268761*
15. *https://pando.com/2013/05/17/for-the-internet-of-things-will-it-be-verticals-or-horizontals/*
16. *http://www.datamars.com/markets/livestock-id/*
17. *www.flaticon.com – free Icons used for schematics*
18. *www.sas.com*
19. *http://embt.co/iot-you*

[T]. Original Material Source: Definitions, Terms & Schema developed by the author for this work.

Appendices

Appendix A: IoT Architecture: 10,000 Foot view

In the work presented at the International Conference on Future Internet of Things and Cloud 2014, Mohammad Aazam et al. provide a 5-layer architectural view of the IoT. This excerpt is exploring the idea presented. “The architecture of IoT is usually considered to be 3-layer, having Perception layer, Network layer, and Application layer, but some add two more layers: Middleware layer and Business layer”. [5]

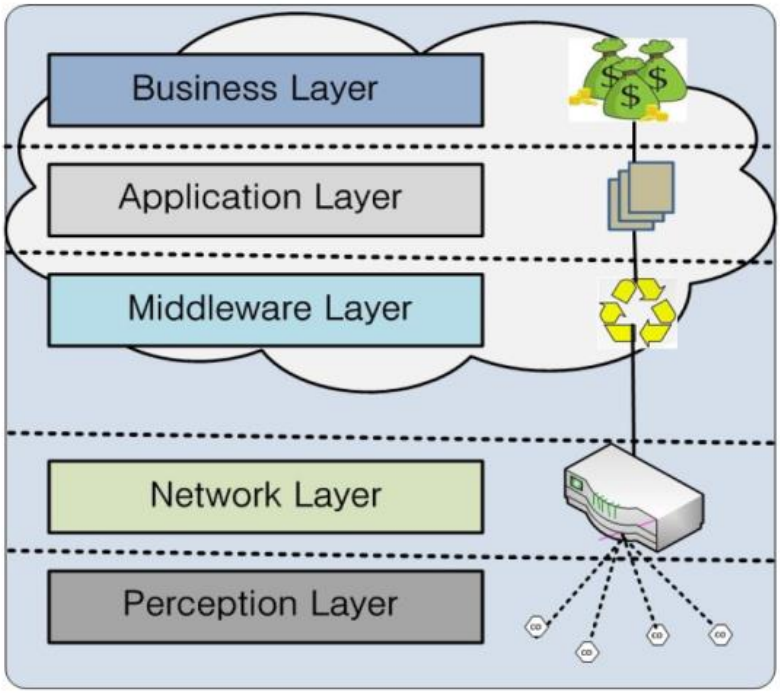


Figure 15: This five layer architecture for IoT.

Perception layer is the lowest layer in the IoT architecture. As the name suggests, its purpose is to perceive the data from the environment. The data collection and data sensing part is done on this layer. Sensors, bar code labels, RFID tags, GPS, and camera, lie in this layer. Identifying object/thing and gathering data are the main purposes of this layer.

Network layer collects the data perceived by the Perception layer. The Network layer is like the Network and Transport layer the of OSI model. It collects the data from the lower layer and sends it to the Internet. Network layer may only include a gateway, having one interface

connected to the sensor network and another to the Internet. In some scenarios, it may include network management center or information processing center.

Middleware layer receives data from Network layer. Its purpose is service management and storage of data. It also performs information processing and takes decisions automatically based on results. It then passes the output to the next layer, the Application layer. Application layer performs the final presentation of data. Application layer receives information from the Middleware layer and provides global management of the application presenting that information, based on the information processed by the Middleware layer.

Depending upon the type of devices and their purpose in Perception layer and the way they have been processed by the Middleware layer, according to the requirement of user, Application layer presents the data in the form of: smart city, smart home, smart transportation, vehicle tracking, smart farming, smart health, and other many kinds of applications.

The Business layer is all about how the service or model works. While it is about making money from the service being provided, non-profit and government owned efforts involved in IoT may also be part of it. Data received at the Application layer is molded into a meaningful service and then further services are created from those existing services. Furthermore, information is processed to make it knowledge and further efficient means of usage make it wisdom, which can earn a good amount of money for the service provider.

Appendix B: FOG COMPUTING - USE CASE I

Hierarchic Distributed Fog Computing Platform for Smart Cities [4]

The proposed 4-layer Fog computing architecture:

Layer 4: edge of network: consists of millions of non-invasive, highly reliable, and low cost sensors; generating massive data streams that are geospatially distributed.

Layer 3: edge compute node: comprised of many low-power and high-performance computing nodes.

Layer 2: intermediate edge node: consists of a number of intermediate computing nodes, each of which is connected to a group of edge devices at Layer 3 and associates spatial and temporal data to identify potentially hazardous events.

Layer 1: cloud: the data analysis results at Layer 2 &3 are reported to the top layer providing city-wide monitoring and centralized control.

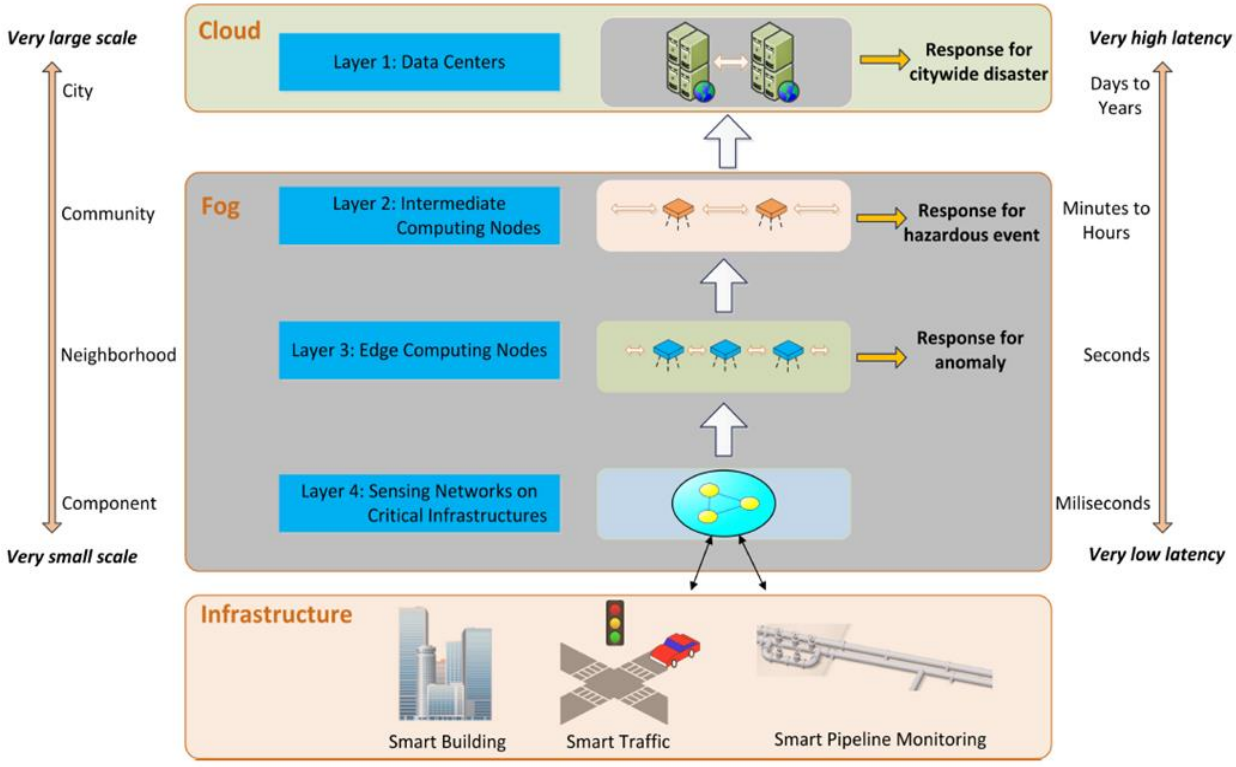


Figure 16: The 4-layer Fog Architecture for smart cities in which scale and latency sensitive apps run near the edge

Appendix C: FOG COMPUTING - USE CASE II

Virtual Desktop Infrastructure (VDI) as an application

Performance Enhancement for Interactive Applications – Cisco Systems [8]

In this use case, demonstrated by Cisco Systems, the presence of a fog proxy can also significantly benefit other interactive applications. In this section, Virtual Desktop Infrastructure (VDI) as an application example is used.

One key challenge with VDI deployments has been the rendering of images and videos to the end user thin clients with graphics-centric workloads like web pages, PowerPoint, Flash, etc. This can be painfully sluggish to the point of hindering user productivity. The main issue is the high latency-low bandwidth of WAN or mobile links that connect the server to the end users and optimizing VDI performance over these already constrained networks.

In VDI, graphic components are rendered together with the desktop image at the VDI server, to be sent across the WAN to the user. When there are no rasterized graphics, the remote desktop components can typically be transported to the end user at a relatively low data rate (for example, via vector graphical commands).

However, if graphics are involved, like when the user plays a YouTube video, the VDI server requests the video from its own location, combines the video stream with the remote desktop composite image, and sends it to the user; repeating each time a new video frame is played.

The resulting data rate for continuously updating the remote display can easily overload the WAN link. Consequently, the video may end up being rendered in a very choppy manner. Cisco proposes a solution to mitigate this problem by separating the graphical and non-graphical elements of the remote desktop as show in Figure 17.

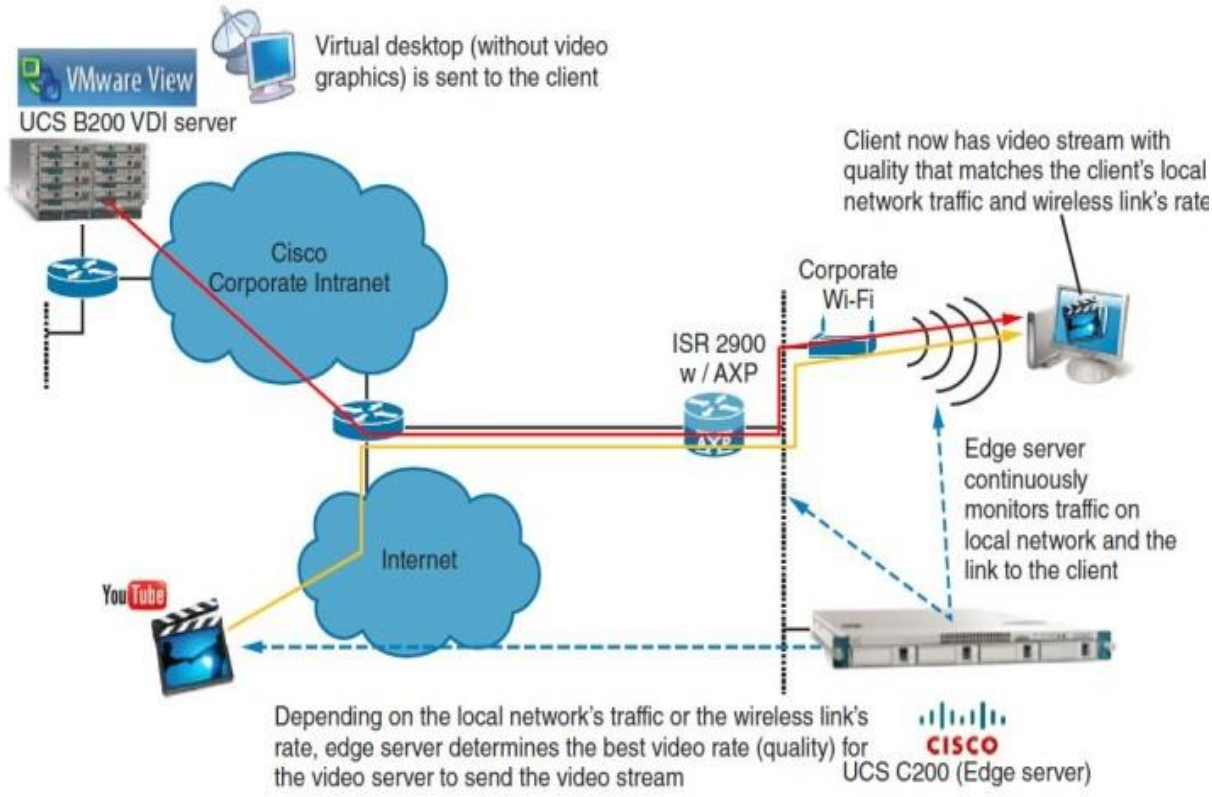


Figure 17: Fog node as a proxy for VDI application performance case study

The proposed architecture introduces a local server at the access edge, like a fog node, that can render at the remote VDI server across the WAN. This new architecture proposed in this use case fits well into Cisco's vision of the fog computing paradigm and aligns well with the principles and architecture discussion in this paper as well.

The VDI server lies at the cloud layer. It will only send non-graphical elements across the network across the WAN while the localized fog nodes and gateway implementations can fetch and render the graphical elements thus avoiding the graphical data from traversing the WAN twice as is the practice currently in conventional VDI implementations. A detailed study of this use case can be referred to in the whitepaper *"Video-aware Wireless Networks- Chan et al.[8]*

Appendix D: Gartner Hype Cycle 2015

Every year Gartner Inc. publishes **"The Hype Cycle for Emerging Technologies"** report. This report provides a cross-industry perspective on technologies and trends.

The year 2015 saw "The journey to digital business" as the key theme of the "Hype Cycle for Emerging Technologies, 2015". The prediction focus is on the impact of the evolution of the digital business and pays the closest attention to the IoT since it has introduced new concepts for identity management and users. Further, as the smart wearables market continues to grow and evolve, Gartner predicts that by 2017, 30% of smart wearables will be completely unobtrusive to the eye.

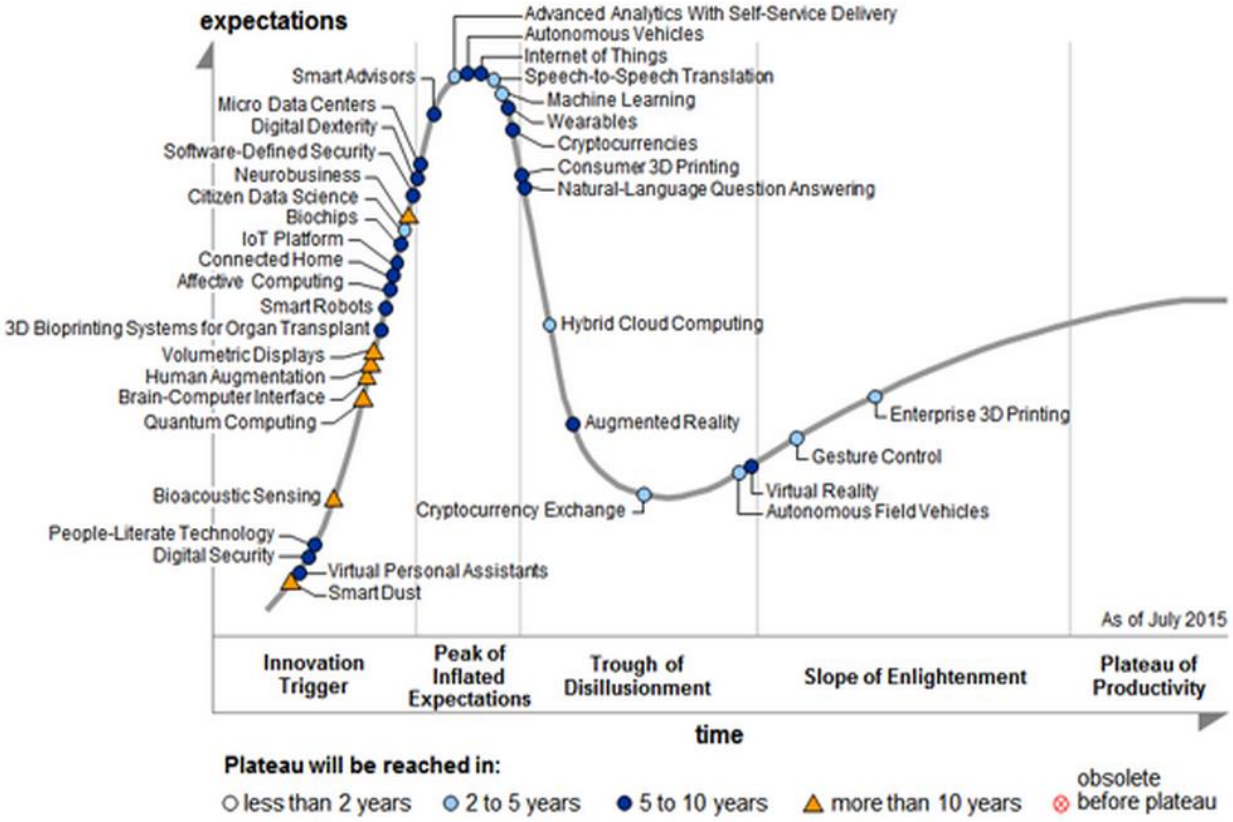


Figure 18: Gartner Hype Cycle for Emerging Technologies 2015

Appendix E: Machine learning Mind Map

Below is an interesting and handy mind-map showing over 60 different machine learning algorithms' for supervised learning organized by type.

Figure 19: Machine Learning algorithms Mind Map created by Jason Brownlee - machinelearningmastery.com