Preparing for Data-Driven Systems and Architectures – Edge, Cloud, and Core

IRDS Systems and Architectures Team

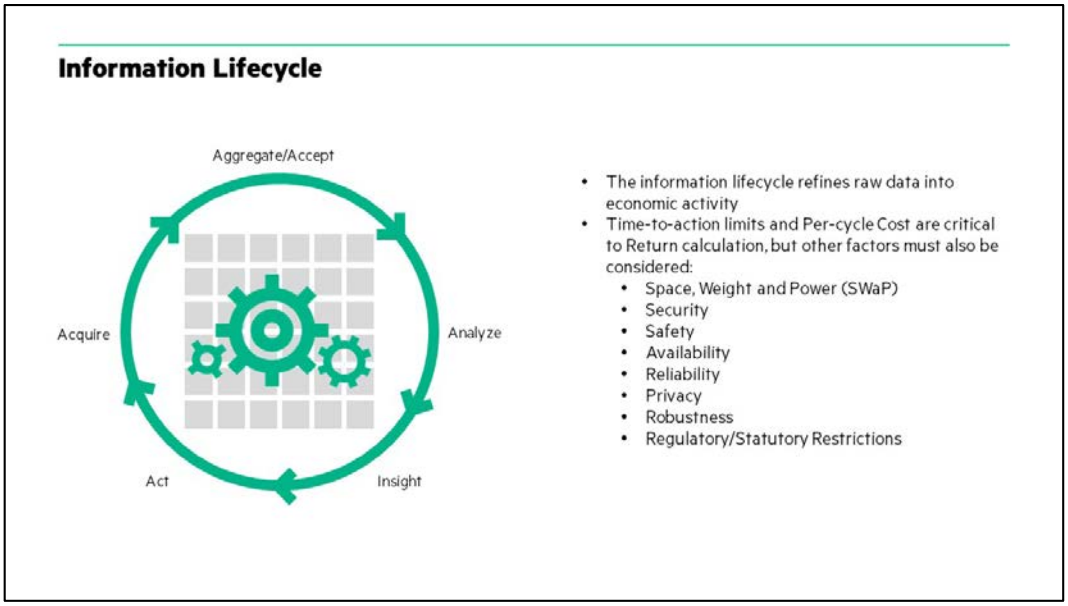
In IRDS 2017, established a top-down system-driven roadmap for key market drivers based on a division of systems into four categories: mobile device, IoT edge device, cyber-physical system (CPS and data-center/cloud. The market drivers were focused on existing workloads and applications in each category and assumed continuation of the conventional System-on-Chip (SoC) integration of additional functions around Industry Standard Architectures (ISA) cores.

During the IRDS 2018 Update, the SA decided to re-examine the rapidly changing application space driven by novel data-driven applications such as Artificial Intelligence/Machine Learning (AI/ML), real time enterprise-scale in-memory graph and data analytics where end-to-end solutions are creating cross-category optimization demands.

In light of this, the SA team will focus on the IRDS 2019 full document update. This document summaries the application, end-to-end solution and system-level the SA team is factoring into their chapter update.

Shifting Towards Data-Driven Systems at every scale

Whether it is the data flowing through the simulation of an exascale compute cluster, the petascale structured and unstructured data pools that underpin the core enterprise application estate of a Fortune 50 corporation, or the sensor fusion platform of a connected and soon to be autonomous vehicle, at every scale and interlocking from edge to cloud to core, data-driven applications are emerging that will stress system architectures in ways that must be accommodated down through the stack to their



Information Lifecycle

- The information lifecycle refines raw data into economic activity
- Time-to-action limits and Per-cycle Cost are critical to Return calculation, but other factors must also be considered:
  - Space, Weight and Power (SWaP)
  - Security
  - Safety
  - Availability
  - Reliability
  - Privacy
  - Robustness
  - Regulatory/Statutory Restrictions

Figure 1 - Information Lifecycles

underlying components. By analyzing the drivers of the constituent Information Lifecycles, we can being to map computational, storage, and communications components at each scale. The dominant factors in the mapping are the time constant of the lifecycle, how long can the use case tolerate, and the information capacity required to achieve the insights that leads to action that restarts the cycle. With the advent of AI/ML as leading analytic techniques, these lifecycles become interlocked and fractal as the analytic tools themselves the output of continuously operating information lifecycles.

Note that only a short while ago, this lifecycle would have included a visualization step because human analytics and decision authority would have been assumed. As time constants for CPS systems incorporated into critical social infrastructure drop from seconds to microseconds, human perceptive lag can no longer be tolerated. As the global data corpus continues to grow exponentially with a two year doubling period and with that growth to come disproportionately in distributed CPS and IoT/edge systems, time of flight through wired or wireless communication systems in addition to bandwidth limits will continue to hold data at the edge, necessitating ever efficient computations capabilities to flow outward to the data.

Design Envelopes from Embedded to Exascale

Of the four categories examined in the IRDS 2017, Mobile Devices are self-contained. The remaining three categories, while the represent distinct optimization choices, are themselves composed of an overlapping continuum of design envelopes from embedded devices up through to data center scale systems for either cloud or scientific HPC workloads.

Beacon Sensor	IoTe	CPS		Trusted data sources; 2.5D/3D integration of sensors, memory, accelerators, computation, and comms; Energy Harvest with inducted power boost modes; SRoT/Blockchain trust mechanisms
Access point				Unified 5G/WiFi/IoT radio access point; Identity, Activity, Locality triangulation; ML/AI augmented operation;
Aggregation Point				Robust environmentals; edge local secure hosting of containerized workloads; Static composition; smallest IT/OT Blended Platform target
Edge Hardened				Robust environmentals; Legacy PXI/AXIe plus next gen modular FF; Static composition; Robust IT/OT Blended Platform target at several capacity points.
Single System Flex				OPC/Rack/Tower systems with next gen modular FF option bays and electrical/optical Gen-Z expansion; Static fabric configurations between reboots; low cost point-to-point expansion
Enclosure Composable			Data-Center	Blade Enclosure augmented with next gen modular FF and memory fabric at the enclosure and rack level; Enclosure level switching of fabrics; Static/Dynamic fabric configurations.
Rack Scale				Dense next gen modular FF enclosures with integrated switching; Superdome Flex enclosure as endpoint; Dynamic fabric configuration; Dematertialized and legacy free; Design for Flex Capacity, Co-Lo, aaS Consumption models. Containers on memory fabrics.
Aisle/Pod Modular				Dense next gen modular FF enclosures with integrated switching; ToR switch; Dynamic fabric configuration; Dematertialized and legacy free; ; Design for Flex Capacity, Co-Lo, aaS Consumption models. Petascale HPC and Petascale Enterprise in-memory DB/Analytics
Exascale HPC				DC scale memory-semantic fabric over photonics; all liquid/conduction cooling environmentals; Aisle/Pod modular for I/O nodes; 2.5D/3D integrated CPU/GPU/Memory modules;

Table 1 – Design Envelope Characteristics

Ten System-Level Technology Inflection Points

Adoption of Data-Driven systems in science, engineering, enterprise governance, and social infrastructure will be accelerated by the confluence of key technology inflection points.

1. From Programming to Training and Inference

   The shift here is driven by the combination of open source software frameworks and the rise of AI machine learning frameworks capable of creation of very effective models based on statistical inference. Unsupervised learning techniques can comb over huge volumes of structured and unstructured data to find correlations independent of expert blind spots. Intelligence craves data and artificial intelligence is no exception.

   This creates a shift in the economic potential from those who create code to those who create the data without which those code stacks are not useful. This also challenges us because the utility of these AI systems is limited not by the ingenuity of the human programmers but instead by the degree which we have engineered systems to admit as much data as possible into training regime as our physics and our legal and security systems will allow.

   Creation of models is only half of the challenge, deploying and utilizing the model, gathering anomalies from operation to fuel of continuous integration and continuous deployment also demand infrastructure and innovation.

2. From One Physics to Many

   Through the first two ages of semiconductor scaling, there have been incredible advances in the other aspects of computer science: algorithms, programming languages, storage and communications technologies all contributed, but they were fundamentally modulated by the CMOS transistor. Innovations were tested against the cost and performance improvements predicted by Moore’s law and if they didn’t have the exponential growth characteristics they were not admitted. Even the obvious defects in security source to the conceptual basis of software models based on 1960s threat landscapes failed to be fixed at the source because of dominance of architectures with the tailwind of CMOS advances.

   Now as CMOS advancement transitions from equivalent scaling to 3D Power scaling, novel computational approaches are increasingly competitive. The work that might spring most quickly to mind, Quantum Computing, along with Cryogenic Computation, emerged as a particular area of focus but it is not the only one. Other areas include novel switching technologies, such as carbon nanotubes; adiabatic and reversible computing that operate at the limits of thermodynamic information theory; neuromorphic and brain inspired computing that draws inspiration from biological systems but, much as with aerodynamics, utilizes materials and energies not available to their biological analogs; Networks of organic and inorganic materials whose behavior calculates desirable functions at breakthroughs in space, weight, and power; finally systems created in our own image which are designed primarily to host intellect which offer computation as a byproduct of intelligence.

3. From Data Centers to Data Everywhere

Today, 90% of information which the enterprise, public or private, cares about is housed in a data center. By its very name, it describes the actions that we have undertaken. In order for data to enter into economic activity, it must be centered, either because it was created there or it had to be transported there.   But, with the advent of so many rich, high definition sensors housed in the ever proliferating number of mobile devices, in as few as five year that ratio may shift drastically to as much as 75% of enterprise information never being housed in a data center. It's not that the data center footprint will shrink, although it will continue to coalesce into clouds both public and private, but that data will grow exponentially and disproportionately at the edge, in distributed social infrastructure, in edge devices personal, public, and private, in all those intelligent things.

There are two forces which keep data at the edge: physics and law. The exponential growth of recorded data, currently a two year doubling period, means that even with the advent of 5G communications and massive communications backbones, there will never be enough bearer capacity to centralize all the data and even if there was, Einstein's limit of the speed of light means that at even metropolitan distances our fastest communications will fail to meet the demands of autonomous vehicles or 5G communications.

The second force is law. There is no common global norm on privacy and the relation and responsibility of the individual to the larger society, which means that there will not be a single regulatory regime that spans the globe. Just like citizens and goods today, data needs to obey the imposition of boundaries. Will frameworks like GDPR continue to offer the protections that they strive to when the vast majority of data will never be in a data center, when the very term data center will be an oxymoron?

The question to ask here is what will it take to admit as much data as possible into economic activity. The first answer is to exploit the asymmetry of the query versus the data to be analyzed. Instead of moving the data to the compute, move the compute to the data. This requires us to understand where we position potentially shared computation resources proximal to the data, in sensors, edge devices, distributed edge compute enclosures, autonomous vehicles. The second requirement to admit data to activity is security in the broadest sense: protection, trust, and control. The protection of robust and energy efficient cryptograph so that query and response are demonstrably safe and correct. The trust of provenance back by secure supply chains, silicon roots of trust, and distributed ledger systems with low energy consensus functions so that every byte flowing into an enterprise can be trusted. The control of data embedded unforgeably in the data, so that down to the byte and the access cycle, all stakeholders in a computation can have their rights verified and protected.

4. From Imperative to Declarative

Imperative control systems rely on enumeration of conditionals and responses, the classic if-then-else diamonds of the flowchart. The problem with Imperative control is that the systems we are creating, social, technical and economic, are two complex to be enumerated. No matter how long we spend, we never can catch the corner cases, there are always exceptions and that means we need to guard band and that means inefficient use of resources, whether it's spectrum allocations or transportation capacity.

Declarative management instead relies on systems which expose their operational state and control surfaces to goal seeking algorithms, such as reinforcement learning. Instead of enumerating all the "ifs" and "thens," set goals to be achieved and let the system strive to maximize those goals. This approach has the added benefit that it does not suffer from the human bias of presupposition of causality preventing us from finding correlations hiding in plain sight. Unsupervised learning and autocorrelation will naively, blindly discover those relationships we cannot precisely because it cannot presume it knows better.

5. From Scarce Memory to Abundance

A decade after Alan Turing created the mathematic theory of computation, John von Neumann was realizing that theory as an operational feat of engineering in his 1946 outline of EDVAC. What von Neumann noted then and what has remained true is that the fundamental limiter to computation is how reliably and cheaply the memory can be made that can keep up with computation. Computation performance has always advanced faster than memory performance. But that is changing. As we enter the age of 3D Power Scaling, memory is advancing faster than computation. The regular rows and columns of memory, the inherent shared, redundant, and repairable structures of memory, the low power dissipation of memory mean that it can growth in the Z axis in a way that may never be possible for the high power and random logic of computation. Layers within a die, die within a module, modules within a package, memories can scale.   At that point the switch to photonic communications can allow the scaling to continue at the enclosure, rack, aisle and data center scale.

A second scalability of memory is scalability in energy. All of the novel memory technologies looking to replace the transistor memory, phase-change, resistive, spin torque, magnetic, all have a degree of persistence. They cost energy to write, they cost much less energy to read, but they cost no energy to maintain their contents. This is what can allow all of those zettabytes of data into unsupervised learning which we can now afford the energy to hold it all in memory. It also reintroduces a technology older than electronic computation, the lookup table. The table of numerical functions used to be the constant companion of the scientist or engineer. Energy was expended to calculate numbers one time, to write those numbers one time, and then those costs could be amortized in perpetuity.   From the 1970s onwards, it has been cheaper to recalculate a result than to remember and recall it. But with persistent memories applied to immensely complex calculations like machine learning routines, incredible volumes of information can be distilled into insights that can be taken to the most energy starved environments like interplanetary space.

6. From Hindsight to Foresight

If we consider all of the information technology infrastructure of a Fortune 50 company, the alphabet soup of HR, CRM, ERP, GL systems, we'll find a system of hindsight knowledge. That's because what represents the state function of the enterprise, the operational data of all of those systems, is spread over petabytes in thousands of relational databases connected by hundreds of thousands of asynchronous updates, and much of that data would be copies.   In order to evaluate the state function of the enterprise, we need to go through a ritual of reconciliation. We need to "close the books", take a snapshot of all of those systems and painstakingly reconcile them.   It is only then that the leadership team has a value of the state function of the enterprise, but it is at best days, most likely weeks old and represented a single moment in time, the instantaneous close of the period.

If, instead, we were able to hold all of that operational state in a unified memory, evolving as a time varying graph, then we can achieve insight. The system function of the enterprise can be evaluated instantaneously and continuously, which means that we can also take its derivatives with respect to time and understand velocity and acceleration, gradient and curl. Now the leadership team can ask any ad hoc question and the enterprise can answer.   We've extended the concept of a digital twin from its origins in physical systems management and extended it to economic systems management.   But what's more, we can unleash unsupervised learning and anomaly detection tools to audit and analyze the data, looking for the telltale signs fraud or inefficiency.

But we can also extended the preventative maintenance concepts to this new economic model. While machine learning gives us powerful statistical inference tools to find in data the patterns we've seen before, techniques like graphical inference and belief propagation allow us to predict behaviors we haven't seen. From hindsight "what has been happening around here" we gain insight "what is happening right now" and then foresight "what most likely to happen next".

7. From General Purpose to Built-for-Purpose

"log2(X)*24". Traditionally, that's how a point innovation has had to survive in months. If you expect an advantage of "X" times the state of the art today, then the log base 2 is how many doublings it will take to match. The Moore's Law doubling period of 18~24 months has set the timeframe for innovation, especially when Dennard scaling was still available. Faster, cheaper to make and cheaper to use is a triple word score. Unfortunately since Dennard scaling ended 15 years ago the straightforward way to continue to reduce power and increase performance has been to make larger and larger die. We're at the point now of "dark silicon" which means that we can make more transistors than we can deliver power to. If all the circuits on a die were active, the heat couldn't be removed fast enough and the chip would fail. Add one more law, Rock's Law, the observation that each successive chip fab costs twice as much. Moore – Dennard + Rock is the recipe for consolidation at every level: the number of companies that can compete to the number of competitive architectures.

But during this transition period between Equivalent Scaling and 3D Power Scaling, may be a period when the tide will shift back to the economic value of novel accelerator design. As the double period extends, there is more time for innovation to remain competitive. Also, as there is more capacity available, there will be more access and lower costs. Open, photonic based interconnects and new manufacturing techniques to allow smaller, innovative designs to be quickly brought together as an ensemble at every level from embedded to exascale will re-admit precision as more advantageous over general purpose and this in turn will enable admission of the new security and energy efficiency innovations that the general purpose has kept at bay for so long.

8. From Proprietary to Open

The Open Source development and collaboration model has proven incredibly effective in software, not only in the complexity of systems which can be delivered, but also in the diversity of those who are enabled to participate. This creates the virtuous cycle where internationalization and localization occur as primary efforts coincident with innovation rather than after the fact, creating greater diversity of representation which again fuels greater inclusion in the economic and social benefits of innovation.

The same guiding principles of open source software development are being extended down the stack. As an example, Gen-Z is a memory-semantic fabric driven by an industry consortium applicable to every level of integration from embedded to exascale. It has been open for review by the open source software community during the entire draft period and lowers the barrier to innovation for novel computational, memory, and communications devices. Regardless of whether it maximizes the potential of conventional CMOS or enables new physics to accelerate a particularly onerous computation, lowering the barrier to innovation and breaking the cycle of improvement solely through consolidation is the antidote for today's technical monoculture.

RISC-V is an Instruction Set Architecture with an open governance model which fully embraces the open source development model in that it freely extensible and licensable. This is a unique new proposition which simultaneously allows for a sustained core software development model that also allows innovation and customization that can be realized in custom or programmable silicon. When coupled with the emerging capacity of from multiple foundries of relatively competitive logic processes, this again enfranchises an ever increasing number of innovators everywhere.

9. From Central Authority to Distributed Systems

   Whether they are economic (cryptocurrency and public ledger), power (microgrids), or communications systems (mesh networks), distributed systems are more complex than centralized systems. But they are more sustainable, more available, more secure, and more equitable, which in turn makes them arguably more just.

10. From Data as Cost Burden to Data as Opportunity

   From its inception, information technology has been dominated by the mechanical advantage and error reduction of automation of human calculations, affording an incredible increase in productivity. Coupled with this productivity increase is the inevitable desire to contain the associated costs. All of the other effects combined yield the more transformative effect, the shift of information technology from a cost center to a profit center by simultaneously increasing the return on processing information while reducing the cost of information. In fact, given the predictive capability of these systems and the efficiency at which

ML/AI systems can operate themselves, everywhere there is data, every manufacturing step, every business operation, every customer interaction casts off information continuously, potentially at a greater level of return than the underlying process itself. The hypercompetitive business relentlessly and sustainably turns raw data to economic advantage via process improvement, investment strategy, customer satisfaction, market expansion, warranty reduction, and direct monetization.

Intelligent Mobility – A Case Study

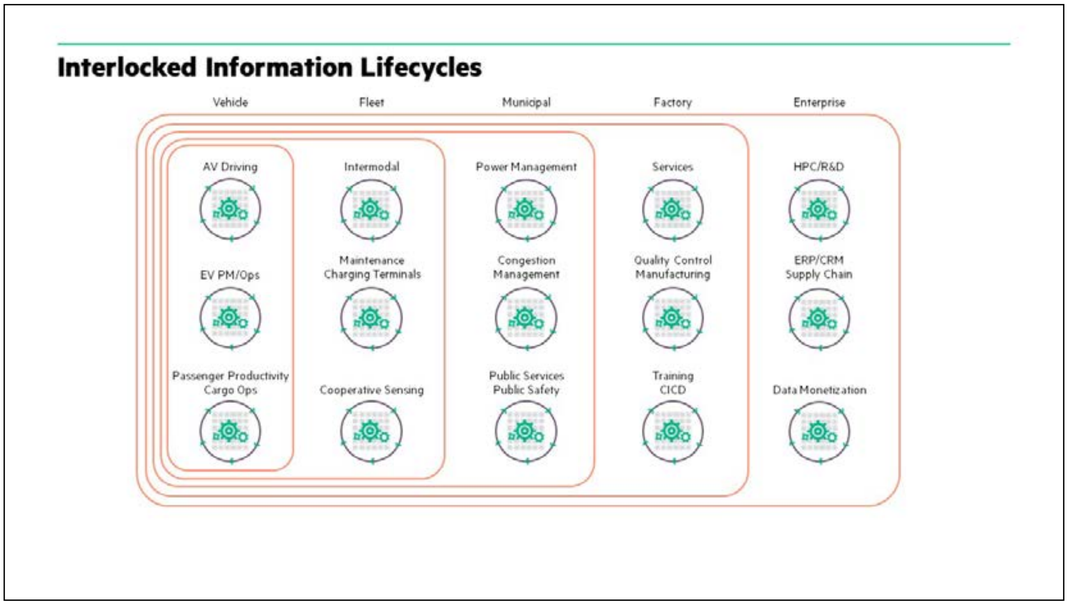


Figure 2 – Interlocked Information Lifecycles in Intelligent Mobility

It can be useful to view this through a specific lens. Looking at enabling an end-to-end intelligent mobility solution that can capitalize on the intersection of the replacement of internal combustion engines with electric motors, the rise of autonomous navigation and the increasingly urban landscape will connect successive edges to clouds to enterprise cores. An analysis of this complex but immanent system captures the interplay of interlocked information lifecycles of acquisition, assurance, analysis, insight and action.

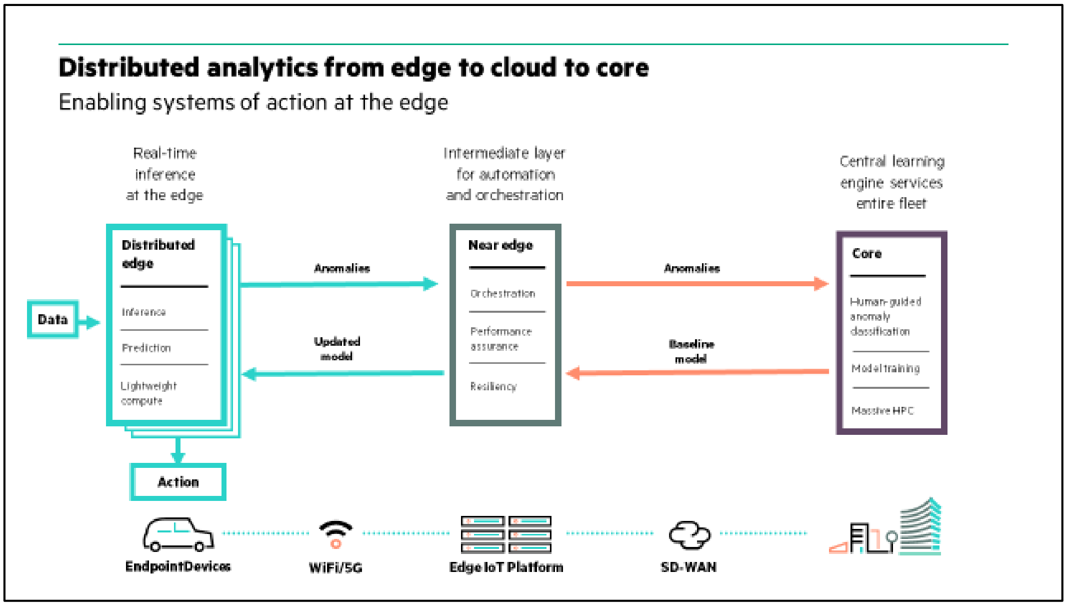

Distributed analytics from edge to cloud to core
Enabling systems of action at the edge
Real-time inference at the edge
Intermediate layer for automation and orchestration
Central learning engine services entire fleet
Data
Distributed edge
Inference
Prediction
Lightweight compute
Action
Anomalies
Updated model
Near edge
Orchestration
Performance assurance
Resiliency
Anomalies
Baseline model
Core
Human-guided anomaly classification
Model training
Massive HPC
EndpointDevices
WiFi/5G
Edge IoT Platform
SD-WAN

Figure 3 – Edge/Cloud/Core end-to-end architecture

Vehicle, Fleet, Municipality, Factory and Enterprise. At each level of aggregation of information lifecycles the demands of life cycle time constant, information capacity and flow along with capital and operational costs plus secondary considerations go into the optimization of when data should move to computation and when computation must happen in-situ with the data. Holistic security (protection, trust, and control) that can be attested back through fabrication will be critical and will place new demands on the entire supply chain.

Combining the opportunities and the constraints and potential economics returns from each information lifecycle with the constraints of the applicable design envelopes that each can admit helps to provide guidance as to which of the technology inflection points will likely be of interest in creating the entire solution.