SOLITAIRE INTERGLOBAL

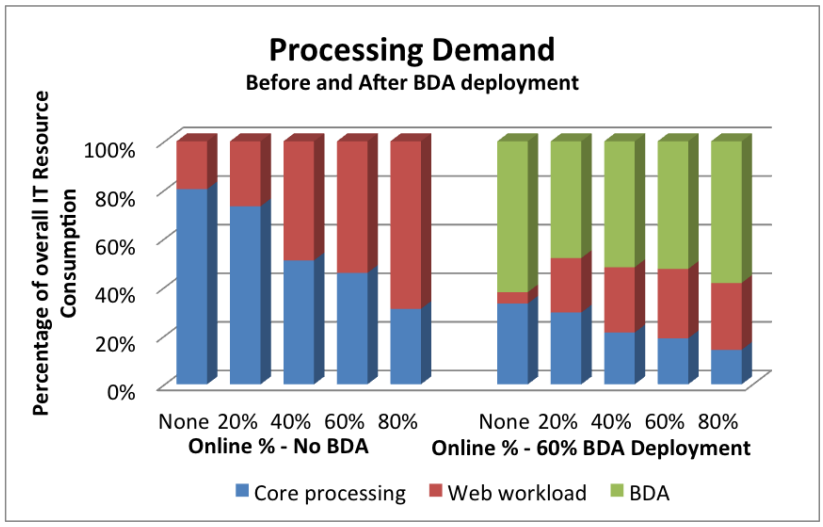
Power Boost Your Big Data Analytics Strategy

Quantitative analysis and business differentiators for business analytics and big data on competitive platforms, including IBM Power

The increase in business analytics and "big data" analysis is an acknowledgement of the growth in overall information that an organization both has at its disposal and with which it must contend. Speed of information arrival, the difficulties of storing that information in reasonable retrieval locations and sheer variety of information content and form make the Big Data Analytics (BDA) challenge overwhelming.

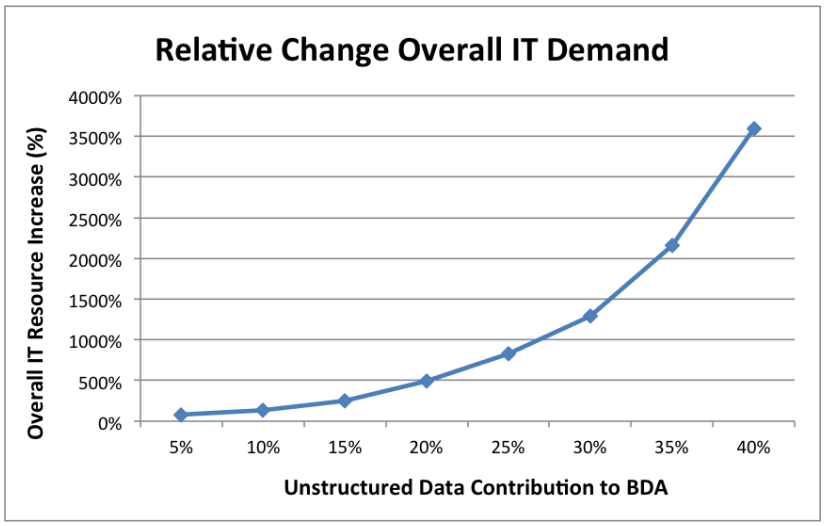
One of the most compelling questions is why does an organization deploy BDA? The short answer is revenue. The whole target of BDA is the maintenance of market share and the increase in revenue. Each of the other goals that organizations state for BDA usage can be tied, directly or indirectly, to this one objective. If this is kept in mind, even things like stronger security, faster time to market, etc. that BDA is typically directed toward can be correlated with revenue targets and opportunities.

The advent of BDA within an organization radically changes the balance of overall capacity planning metrics and the types of resource demand growth in IT infrastructure.



Without BDA, the switch to increasing amounts of processing from online interactions makes the view of performance and capacity using pivotal response and shorter latency transactions key, since those are the metrics that support online user and customer performance requirements. The larger the percentage of online deployment that exists, the stronger the demand for the key IT characteristics of throughput and quick servicing of short transactions becomes. The main capacity control value becomes the number of concurrent users, with processing speed and data trailing far behind. When an organization deploys a significant amount of BDA, critical IT characteristics shift again. BDA workload quickly becomes larger than all of the other workload within the operational domain. This shift also changes the capacity control values to the amount and type of data being touched by BDA. The chart shows the relative capacity demand for organizations with and without BDA. A 60% BDA deployment has been used for a comparison point. This means that 60% of the

overall organizational data, both structured and unstructured, have been included in the BDA domain.



As more unstructured data is included in the BDA efforts, the total processing demand radically increases. The reported BDA deployments within this study show that an organization that deploys 60% of its overall data can see significant shifts in resource demand based on the mixture of structured and unstructured data. The chart shows the reported relative increase in IT demand based on the percentage of unstructured data in the BDA efforts.

All of this makes the platform choice for BDA deployment a difficult challenge for any IT department. The layered resource demand is difficult to project using most capacity planning tools and the explosive use of organizational resources can have a ripple effect of epic proportions. Of the organizations within the study that have fully deployed their BDA practices, the difference between the IT architecture resources before the deployment and after is equivalent to *16 times* the pre-BDA level. The strain on platform I/O, memory, network bandwidth and calculation capacity highlights the increased sensitivity that the post-BDA IT world must face.

> *"We have had three different waves of implementation for our Big Data Analytics over the last 14 months. Little did we know what that REALLY meant in terms of storage, computing power and just the sheer consumption of our platform! For the first wave, we projected a "comfortable" increase of 12%, thinking that we were being overly cautious. That resource level was exceeded in the first month – totally putting us in to hurry up mode to get things running smoothly. We added in over 20% to our server infrastructure just to handle this workload.*
>
> *The second wave was smaller – or so we thought again. So we planned for an additional 20%, again thinking we were safe. The projections that the storage vendor got from you [SIL] were not believable to us. We thought you were just trying to sell their equipment when you projected a whopping 33.5%. So we ignored it. A month and a half after the deployment, we hit our limit and again had to supplement. The final resource level was just a smidge over 33%. You just have to work on that 0.5%!!*
>
> *So this time, when you projected a 51.4% increase of overall resource consumption, we put in 55%. That looks almost exactly on the money. Thanks! However, that means that when we get everything into this process, we will be about 15 TIMES what we started with. Scary, but it is still making us LOTS of money."*
>
> Manufacturing IT Director

The second compelling question is even more basic – "What is BDA?" Many people talk about BDA as if it relates to the *amount* of data that they have or that they are going to process. The lines between BDA, analytics and data mining are easily confused and

benefit from clarification. The first clarification is extremely basic - big data is *not* lots of data. It has no relationship to the amount of data, but instead refers to the vision of the data. The BDA vision incorporates both structured and unstructured data, unlike any other organizational view.

Structured data is information that is stored in databases, files, etc. where the contents of the fields and attributes are known. They have definitions and relationships to events and other data elements. Unstructured data, such as emails, lacks this definition and associative relationships. Its connection to the structured data is murky at best and frequently unknown or undocumented.

The difference among analytics, data mining and BDA rests on the unstructured data. Both analytics and data mining address data that has had inference structures built – how elements relate and how they are populated in time. Analytics is actually the narrowest vision in that it examines the details of what is already processed and known. Further insights are built by coming at the defined structures from new perspectives and derive value from finding previously unknown summations and causations.

Data mining (DM) is slightly broader in perspective, although it is also focused on structured data elements. Data mining itself is a computational process of *discovering* patterns in large datasets, with an overall goal of extracting that information from the data set and transforming it into an understandable and documented structure for further use. BDA takes this to a wider domain with the inclusion of both structured and unstructured data. It adds all of the challenges of the DM processes, with added decryption, isolation, inference and transformation steps. Both processes have a variety of steps, some of which are listed below. Any BDA process encompasses all of the steps, although data mining does not.

- Creation of weighted-interest metrics
- Database interaction (pulling data)
- Fragmentation assembly (BDA only)
- Inference construction
- Initial analysis
- Online update
- Post-processing of tagged structures and elements
- Pre-processing of data components
- Repository creation
- Unstructured element decomposition (BDA only)
- Visualization

Within both data mining and BDA efforts is the automatic or semi-automatic analysis of large quantities of data to extract previously unknown interesting patterns. These patterns can be in a group of records (cluster analysis), identified as anomalies or cognitively dissonant, or associated with data elements that are themselves part of an interesting group or anomaly. If the element itself is only interesting by attachment to another element, the association provides the inclusion in the DM or BDA processing. The biggest difference between the two types of discovery and analysis is the scope and complexity of the process and domain of the data. The inclusion of the unstructured data significantly widens the scope of BDA efforts, with a resulting workload challenge.

The selection of the optimal platform for BDA must include technical and business factors, but the overall decision is a business one. Since the impact of platform for a deployment is a difficult one to quantify on this level, IBM engaged Solitaire Interglobal Ltd. (SIL) to conduct surveys, gather data and perform analysis to provide a clear

understanding of the benefits and relative costs that can be seen when organizations implement IBM Power platforms as part of their BDA IT architecture. This analysis has been primarily directed at the value of that platform use from a business perspective, so that those whose role it is to provide business leadership can understand the benefit of the IBM Power architecture in BDA deployment and evolution.

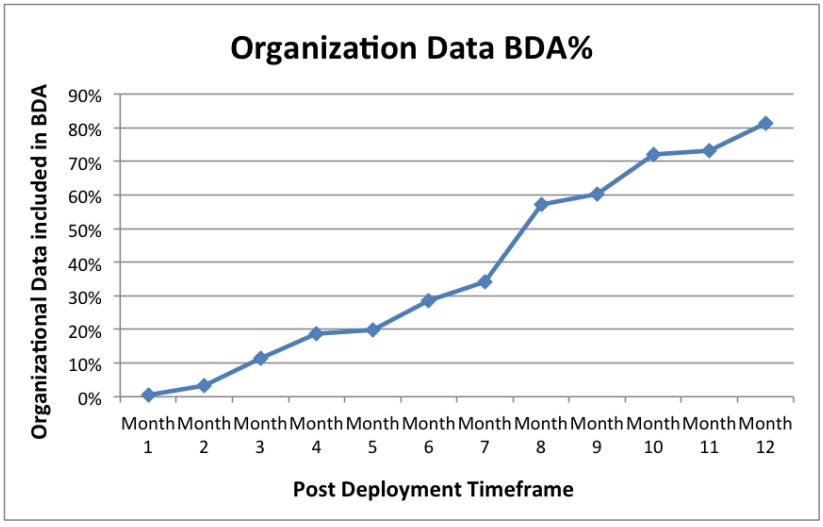
During this study, the main behavioral characteristics of software and hardware were examined closely, within a large number of actual customer sites (31,000+). These customers include organizations that have deployed big data and business analytics as part of their production environments. This group has organizations that maintain deployments that have been customized to support additional functionality and business process, as well as those that have been integrated with third-party or custom components. The information from these customer reports, and the accompanying mass of real-world details is invaluable, since it provides a realistic, rather than theoretical, understanding of how the choice of platform can affect the organization's costs, risk and strategic positioning in the current marketplace.

In the collection and analysis of this data, a series of characteristics were derived. These characteristics affect the overt capacity, efficiency and reliability of the environment and its affects on operational and business performance. These characteristics have been examined within a business framework, since that is the perspective that is needed to make an informed decision. The business perspective encompasses a myriad of factors, including costs, staffing levels and other effects. This is the basis for selection of an optimal platform for an organization's big data and business analytics requirements.

BDA Adoption and Use

The deployment, adoption and use of BDA processing and information is at the core of the study. The processing itself can be viewed as a series of events or significant steps. Each of these points is essential to the BDA effort and must be supported by the deployment platform. The steps are normally executed as listed, but can be reiteratively performed before the initiation of the next step.

- Gather data
- Build repository
- Analyze data
- Develop insight
- Act on insight



In general, BDA follows two possible pathways of usage. The first comes when an organization does not have one of the critical requirements for incorporating BDA results into its strategic decision making. When that happens,

the BDA effort languishes and then is cut. BDA deployments can be expensive for an organization. Without a demonstrable ROI, it can easily be labeled a faddish luxury and is cut from overstrained budgets and staffing.

The other path is taken when an organization successfully incorporates the initial BDA findings into its business processes. When this occurs, the demand for BDA explodes from a capacity demand perspective. The insights and ROI can be literally transformative, and the organizations that realize this, swing fully into BDA efforts. Within the organizations reporting into the study, the adoption of BDA efforts can be seen in the chart above. This chart has plotted the timeframe of the BDA deployment against the amount of organizational data that has been brought into the scope of the BDA efforts.

> *"The implementation of our Big Data practice has been in interesting stages. First we did a bake-off with three hardware and software vendor teams. That looked pretty good as a POC (proof-of-concept) so we started our big data analytics with our marketing folks, who were pretty enthusiastic about the whole thing. We focused on only a couple of things in the marketing area, mostly to develop some customer loyalty actions. That succeeded so well that the rest of the marketing managers were yelling about when they could get into the action.*
>
> *The VP of marketing actually paid for the whole first server for the analytics, and did not fuss at all when we had to get more power three times in 4 months. The results really made a strong enough business case that I did not have to do any arguing or massive presentations.*
>
> *We got a lot of complaining about favoritism from the other departments, so our CEO finally just said that we should make the deployment available to everyone. The extra equipment and resource usage is major…"*
>
> CIO large retail organization

The stages of adoption and the changeover of the usage of BDA within different parts of the organization is also revealing of some significant patterns that need to be considered when evaluating BDA. The stages of adoption can be loosely defined as:

- Initial investigation – This is the quintessential "toe in the water" or proof of concept stage. It is used to develop processes and evaluation points, and serves to show possibilities to management.
- Limited deployment – Normally the real proof of value, rather than the proof of concept for any BDA deployment. In this stage, the BDA efforts are expected to show some value or return.
- Departmental or increased scope – The organizational departments or lines of business that have accepted the value of BDA will normally step in at this point. They typically are ones that can fund the effort. The most prevalent of such areas are marketing and fraud detection.
- Full organizational deployment – Once the overall executive management is convinced of the value of BDA, the efforts are expanded to include core areas for the entire organization. This is funded from an organizational perspective and takes the financial burden out of the individual departments and lines of business. It also may change the control of the BDA efforts.
- Organizational imperative – An interesting continuing evolution of the previous stage, this is marked by the *requirement* to use BDA insights and the inclusion of BDA results in management staff evaluation. This marks a total change in organizational posture from that prior to the initial investigation step.

For each of these stages, the scope of data and the resultant capacity demand change significantly. The following chart shows the percentage changes in resource demand that BDA has consumed, based on the organizations in the study. To normalize the data for comparison, the chart shows the change based on the percentage of organizational data and operational process load.

The significant change in resource demand for organizations that are successfully deploying BDA is startling when looked at from an increased processing perspective. This change can be as much as 1,634.52%, or *16 times* the operational load, and certainly presents a challenge to the deploying organization. With such a large potential explosion of demand, characteristics of the BDA deployment platform become critical from a cost and management perspective. The underlying technical architecture, and its resiliency, speed and resource control can make a substantial difference to the bottom line of the deploying organization.

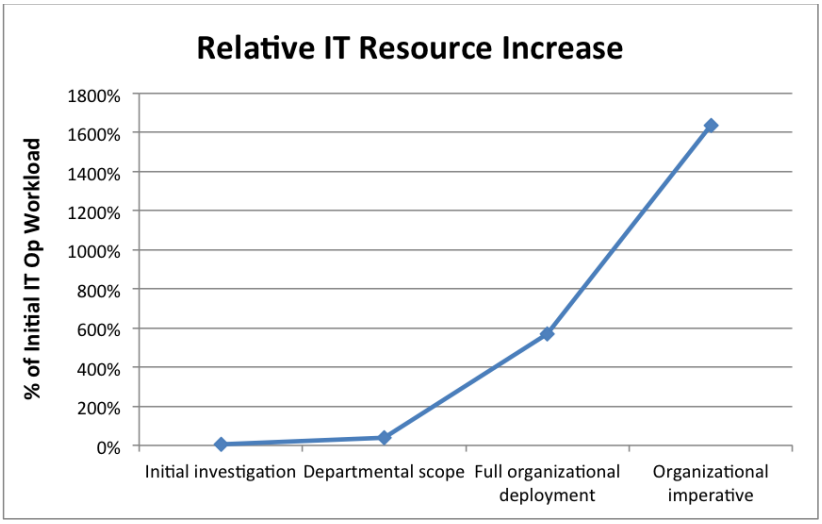
There are several areas of BDA usage and adoption that have shown a high return for deploying organizations. They also have an increased sensitivity to poor performance by the deployment platform, since any shortfall in throughput, availability and capacity have an auditable impact on the organizational bottom line.

Areas of Focus and Business Impact

An increasing leverage point for BDA is the interaction with the customers of an organization. This type of functionality typically builds associations during customer online interactions with catalogs, web searches and purchases. It recommends further purchases, suggests additional considerations and lets customers feel a sense of membership. The BDA component is key in an association that includes unstructured components, tying customer emails, calls, etc. into the interaction. This has been shown to significantly increase customer loyalty and increase the average customer purchase per interaction by a considerable amount. The higher purchase has an average impact of over 41.5% of the individual sales amount, with a corresponding impact on revenue for the study organizations. The top cited metrics in this area are:

- Overall revenue
- Faster time-to-market (agility)
- Rapid response during customer interaction (agility)

Lack of speed and throughput, as well as any shortfall in availability, are highlighted by missed opportunity and adverse revenue impact. A platform that supports strong availability, speed and throughput has a measurable, and highly visible, benefit to the organization deploying the BDA efforts.



Marketing is the largest user of BDA in over 78.4% of the organizations reporting into this SIL study. The marketing and sales departments also have the most comprehensive view of data within an organization, including factors and sources that operational departments often ignore. This expansive view demands flexibility from the BDA deployment, where massive supplemental

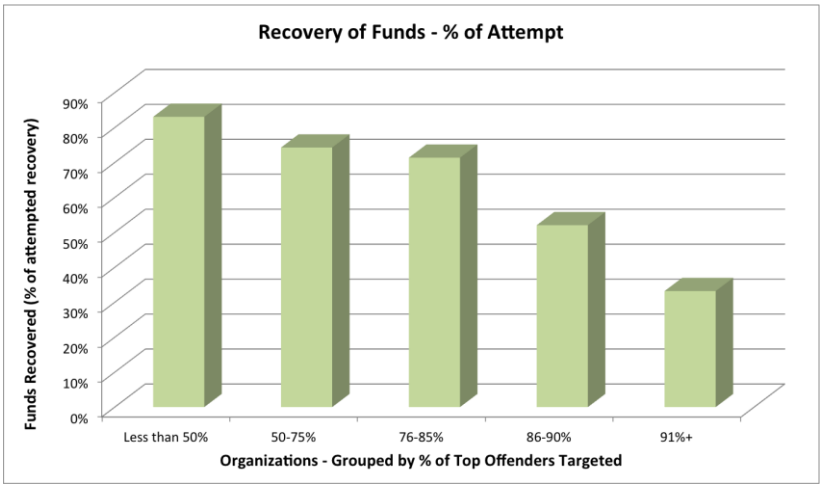
sources of data may be swept into the BDA domain with little or no notice. This fluctuating and growing scope can cause less resilient platforms to become the choke point in critical marketing strategies and product offerings. The ability to handle increasing BDA requirements and to easily expand platform resources to match usage demand becomes critical when looking at successful BDA deployments. The top metrics in this area were cited by more than 97.5% of the study organizations. They are:

- Agility, agility, agility
- Flexibility
- Rapid response during customer interaction

Agility is the supreme metric for marketing. The faster the speed between conceptualization and actualization, the more effective is the sales and marketing effort. Of course, this is normally tracked by revenue goals and changes. BDA becomes intrinsic to the evaluation of successful marketing quickly for those organizations that wholeheartedly deploy BDA processing. The reported BDA deployment average for marketing was 4.1 months, marked from the BDA initial investigation.

The second most reported incentive for BDA deployment is in the areas of fraud, waste and abuse (FWA) discovery. In this area, the deploying organization is normally interested in a dual-faced view – both in the discovery of historical actions, designed for recovery, and in the prevention of future FWA.

Since the discovery of FWA has been reported to be as large as 68.5% of total organizational expenditures for insurance, warranty and other assurance programs, this area has a high visibility. Of course, the discovery of FWA is only one component of effective administration. Other critical components include the management of the investigatory efforts, frequent additional BDA analysis, and established recovery strategy, among others.



Many of the organizations within the study that have assurance programs focused on the reduction of expenditures to fraudulent claimants. These represent 41.3% of the study organizations. Some organizations (22.5%) have secondary objectives of increased service agility, where non-fraudulent claimants get their recompense faster and more effectively. The difference in the two sets of objectives has an underlying impact on platform capacity load and the type of integration into the other organizational processes. Further exploration into this dimension are out of scope for this study, but can be addressed with further inquiry.

The recovery of the fraud expenditures is shown in the chart. It is split into groups based on the recovery threshold set by the organization. Hence organizations that have set a target of the top 80% of offenders are summarized separately from those organizations that have focused on the top 50%, and so on. This summary highlights some of the significant financial impact that BDA can bring an organization.

The percentage of recovery drops dramatically when the target group is 86% or more of the discovered fraud. This stems partially from the fact that smaller fraud occurrences are not as profitable to recover as the larger ones.

The timing of recovery realization, or how long it takes to see recovery funds, varies based on the guidelines set for the investigation and remediation portion of the cycle. If a strategic approach is selected that targets based on ROI, the recovery will start to be noticeable after an average of 4.82 months.

All of the organizations reported accelerated delivery of services, as the ripple effects from the fraud detection and investigation continued. Over an 18-month time period, the reduction in fraud resulted in more accurate client service delivery, with the organizations reporting agility increases that range from 11.3-43.2%. The top metrics in this area were cited by 91.1% of the study organizations. They are:

- Reliability
- Agility
- Scalability

The large potential for financial impact makes the BDA for FWA a highly visible focus for an organization. The sheer size of the return makes interruption or slowing of BDA processing critical for an organization. This translates into the resiliency and availability of the deployment platform being extremely important to the organization goals and objectives.

Another area that leverages BDA is business development and research. Many organizations deploy various levels of BDA in this focus area, with pharmaceutical, bioscience and healthcare companies currently representing the largest group that effectively base their business on BDA. While the bottom line of the BDA process in this segment is again revenue, the identification of new discoveries feed into market position and ROI. Once again, the deployment platform and its characteristics of resiliency, dependability, throughput and scalability are critical underlying factors.

In all forms of BDA deployment agility is an underlying requirement. However, different organizational groups view agility in significantly different ways. A broad split of basic agility questions is helpful in understanding how the various forms of agility needs fit together. The questions that form agility views are:

- Did it get done?
- Did it get done on time?
- Did it get done predictably and dependably?
- Is what got done worth the cost?

These questions are asked in different forms in each of the focus areas for BDA. The table below shows how the focus groups rank and define their agility perspectives.

Focus Area	Agility = done	Agility = done on time	Agility = done on schedule, always	Agility = done with good result	Commentary
Customer interaction	X	X			Executing the BDA on time is critical in customer interactions. The window of opportunity is short and will not come again.
Marketing	X	X		X	Focus is on completion and quickly. If the feedback is delayed, the insight will be lost and creative analysis stymied.
Fraud and abuse	X		X	X	This type of BDA tends to be more massive and scheduled, so the perspective is slightly different than the others.
Research and development	X			X	Done and with good results are the metrics in the area of true research.

While all of the agility needs are based on getting the job done, the definition of the speed and "done" varies. This is an important differentiation to keep in mind when designing the BDA solution, and also when selecting a platform.

There are many other areas of BDA usage. While the focus of the BDA may change, the bottom line revenue effect is frequently the key metric. Exceptions to this, where things such as disease vectors, climate effects, and so on are also extant, but represent a smaller set within the current study.

BDA Metric Perspectives

The varied focus areas of BDA translate into a complex profile for the selection of a deployment platform. The requirements for resiliency, availability and speed make the selection of an architecture challenging. While the normal business perspectives of satisfaction, expense, etc. are still valid, the technical capacity for deployment support becomes extremely critical in the selection process. The underlying platform architecture becomes a significant factor in the profitability of the BDA efforts, shifting the selection from a commodity base to one that is weighted more heavily toward resiliency, scalability and performance. That being said, ultimately, IT and technology management are designed to support business functions. The primary perspective of the study was the view of the technology effects within the BDA space by an organization's business management, both executive and line-of-business. For the purposes of this part of the analysis, the patterns of operations from the study organizations have been grouped into similar categories and then compared to identify their influence on key business metrics. These metrics are:

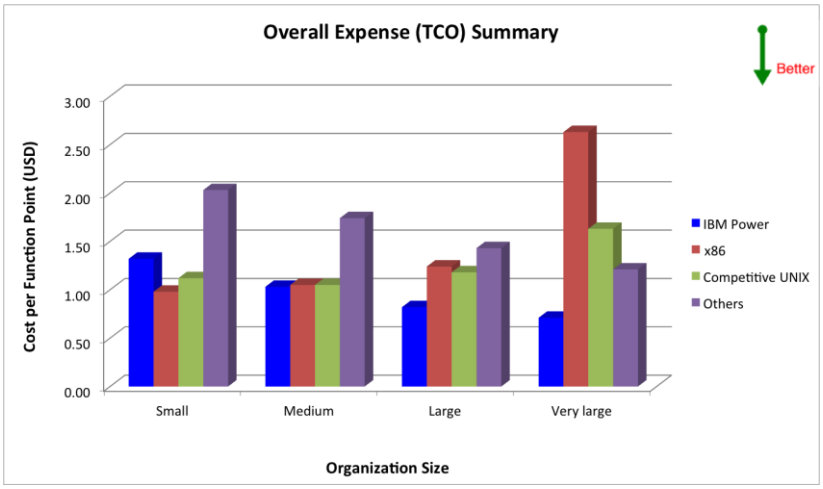
- Cost
- IT risk (security, resiliency, etc.)
- Agility
- Customer satisfaction

Each of these business metrics has measurable and significant differentiation when the projected platforms and platform groups are viewed. For the purposes of this comparison, several architectural groups have been included. They are IBM Power, x86, competitive UNIX and "other". The IBM Power separation is designed to look closely at the effects of the Power architecture, so that it can be compared and contrasted with other architecture offerings.

The more granular business metrics are those measurements that show how a specific measure of success is different in the general population of the implementers versus those that deployed Power. These metrics are fairly broad in coverage and touch on areas of financial consideration, as well as organizational quality. Each metric is presented with a short definition and the focused net effect of IBM Power big data and business analytics deployment. In order to be meaningful across a variety of industries, all of the metrics have been normalized on a work-unit basis[1], and categorized by levels of organization size (small, medium, large and very large). The base measure has been set by the medium company average, so that all other metrics are based on a variance from that standard set point. The implementations included in this study have been restricted to those implementations in production.

Cost - Ownership and Information

This cost perspective looks at the total cost to the corporation of big data and business analytics operations. The expenditures that are included in total cost of ownership (TCO) span many divergent expense categories, including personnel, equipment (i.e., servers, network infrastructure, etc.), utilities, software, and maintenance (facilities, hardware, software, etc.), to name a few. All outlays in these categories pertain to the operation of big data and business analytics, but specifically exclude development and initial production rollout costs. The burden for new enhancement functions is not included, leaving just the summation that identifies the organizational running rate for an installed big data and business analytics system. The TCO financial metric is more comprehensive than a straight operational metric. This metric should not be viewed in isolation, since extraordinary expense patterns for individual organizations may cause minor variance in the exact comparison values. For this reason, the comparison metric should be viewed as indicative and providing a general range rather than an exact value. However, with the large number of contributing organizations, the data is sufficiently large that, combined with the other business metrics, this comparison helps to set an appropriate perspective. The TCO has been normalized in the study based on employee count, sales revenue and legal entity count.



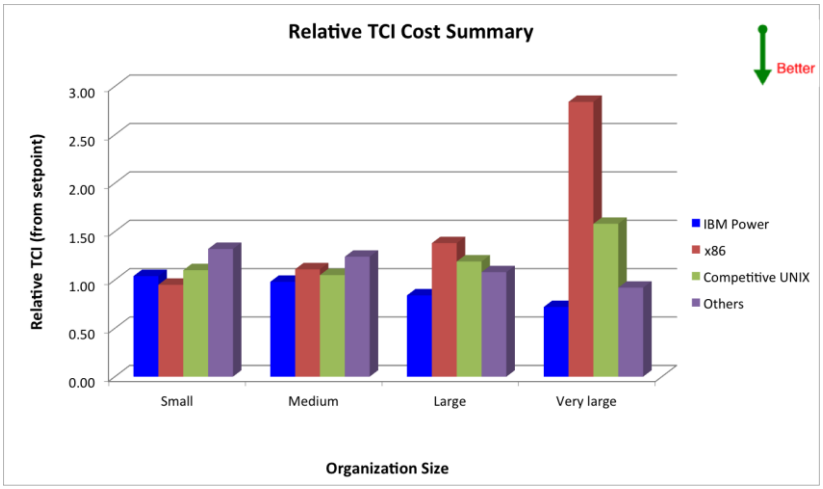
The IBM Power big data and business analytics deployments show lower overall expenses for the larger deployments, with the costs associated with the Power platforms by as little as 73.00% of competitive offerings. The cost of hardware acquisition is

[1] *Work-unit basis has been defined using the published International Function Point User Group standards and are based on function point (FP) analysis.*

frequently higher with the Power platforms than for the smaller Intel platforms. This disparity in cost levels is obviated over time, as the defining expense metric switches from acquisition cost to TCO. This switch happens in all deployments eventually, but is more rapid in the larger installations.

> *"Our big data analysis stuff is on an IBM Power box. We had it running on an Oracle box before, but the costs were pretty high. When we figured out what the cost per model run was, it looked like it would be about 30% cheaper on the Power box. What we are finding is that it is almost 55% cheaper. AND it runs faster. So we have started to transition all of our big data work to the Power platform. "*
>
> Distribution CIO

The difference in TCO among the solutions is based largely on the lower expenses for the efficient deployment and operations of the big data and business analytics implementation and the lower overall cost of the solution, including staffing. This is affected by the scope of the big data and business analytics deployment, with increased expenditure efficiency present as the complexity and size of the deployment increases. These values are limited by current year costs and workloads, and incorporate organizations in various maturities of implementation, although all are in production.

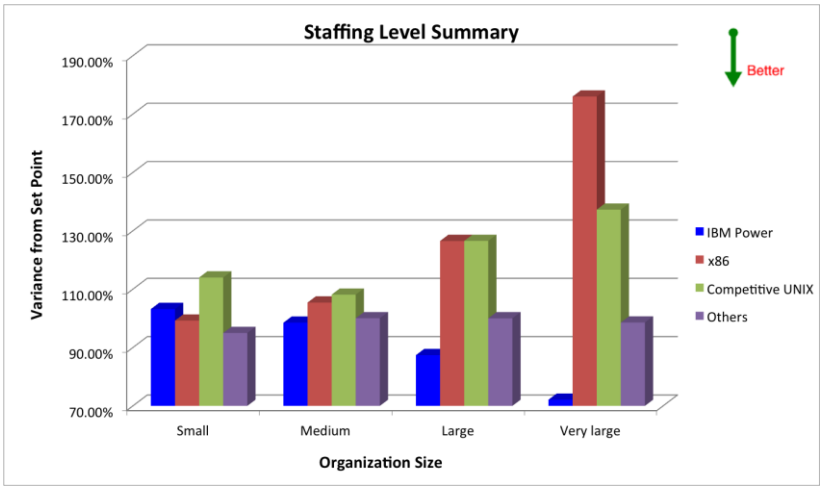


With the amount of work that is being performed in an organization, it is possible to combine the TCO metric with that workload metric to get the total cost of information (TCI). This provides the organization with a unit of work metric that is translatable from platform to platform. It allows comparisons for total costs and capabilities to be built in an objective manner. Using this metric, the trend in cost per work unit for the BDA implementation demonstrates the efficiency of the Power platform. The trend in declining cost per work unit for the increasingly complex large deployments is also indicative of the strength of the platform. This translates into a lower total cost of information (TCI) for Power BDA deployments by as much as 74.65%.

Customers reported a consistent pattern of differentiation for their platforms in:

- Utilization of platforms
- Staffing costs overall (due to tools, stability, etc.)
- Licensing costs
- Datacenter costs (environmental, facility, etc.)

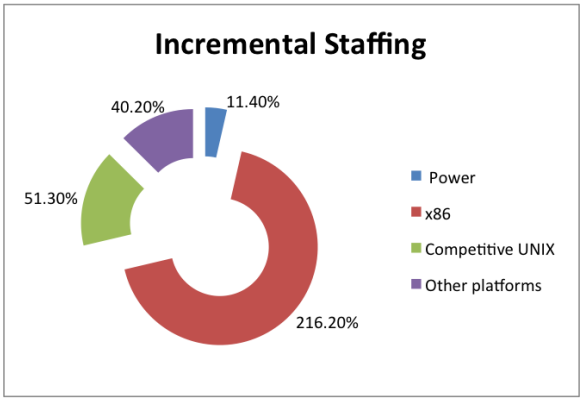
More than 93% of the reporting organizations cited these factors as the most influential factor in their perception of cost.



An underlying factor that shows itself in many other areas is the effectiveness of the interface between the technical user and the infrastructure, including software, hardware and operating system components, and the subsequent effect on staffing. The efficiency of any of the specific components that provide that influence on the user experience are difficult to break down into metrics other than in overly-detailed comparisons that lose their effectiveness by virtue of the degree of detail. Therefore, a general view of the full-time staff position equivalents was reviewed to provide a general metric for the platform comparison. These levels are those required to maintain a "gold standard" environment for each deployment option. In order to provide a common comparison field, the workload on the systems was normalized to identical levels. The set point for comparison was selected as the staff level for a medium-sized organization using Dell x86.

Standards of implementation for big data and business analytics and its best practices have been used to define the rigor of the deployment processes and functions. This is a somewhat murky area, with best practice guidelines very immature and fragmented. Due to this ambiguity, the standards applied to the BDA implementations have been set at a very simple level:

- All data is covered by normal DBA oversight, if contained in a DBMS
- Non-structured data is maintained in some form of fault tolerance, i.e., RAID, etc.
- BDA process statistics are kept for an excess of 6 months
- All BDA statistics are analyzed in reference to workload, timing and effect on operational levels
- All BDA results that are used in strategic or tactical decisions are documented



These standards are extremely simplistic, but are the minimal best practice implementation guidelines possible. It is important to keep the rigor of those standards in mind when reviewing the staffing. The noticeably lower staffing level for Power deployment and use is directly attributable to an efficient unified workflow, as well as a substantially different mechanism to handle the allocation of virtualized resources. The

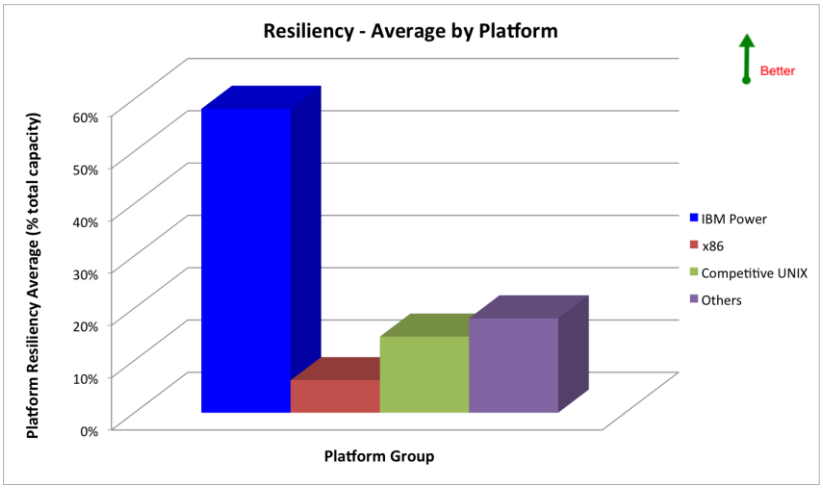
normalized staffing levels for Power deployment are smaller than those for the competitive offerings by as much as 59.02%.

Incremental staffing for those organizations that both have a longer maturity of the BDA practice itself and the BDA integration into decision-making are showing an increasing differentiation on staffing. After six months of BDA integration, the average reported incremental staffing is very indicative of the relative difficulty of implementing BDA on the various platform architectures. The significant variation is indicative of some of the IBM Power workload features, and SIL will continue to monitor this area as more BDA deployments continue to mature.

> *"We were prepared for the explosion in storage requirements and the heavy demand on the servers when we put in our big data stuff. What was not in the plan was the change in staffing. I know now how naïve we were, thinking that we could handle the extra activity with our current staff.*
> *It got pretty rough for a while, but we now have added 7 people to handle the work. That is a lot of people, but with how much good the analysis is doing our company, is a pittance.*
> *Next wave is looming, and we will do much more planning. And also will pay attention when you tell us that we are living in a dream world on the projections! "*
>
> Large service company CIO

Risk

BDA radically increases the workload demand on a platform. The complexity rises in relationship to the amount of an organization's overall data is included in the BDA domain. As the data in the BDA grows, the complexity increases at a geometric rate, making a very challenging environment. The reported complexity increases set the level as much as *11.3 times* as those systems without the BDA component.



The workload is correlated with the complexity, so the increased complexity translates to significantly higher capacity demand from the deployment platform. The ability of the platform to support the rapidly changing requirements is a key attribute for selection of the BDA architecture.

The broad coverage that is inherent with a BDA deployment means that most organizations deploy BDA in phases. These phases can be based on departments, lines of business, or project. Each phase can have a ripple effect on the previously installed operation, with thresholds stressed and timeframes impacted. Since the phases use incremental capacity, the success of the platform to handle the increased volatility of the workload and schedule pinpoints the inherent resiliency of the platform. SIL defines resiliency as the ability of the architecture to absorb periodic spikes in demand without additional platform modifications. In a resilient system, occasional capacity demands can be accommodated without recourse to additional purchase and configuration.

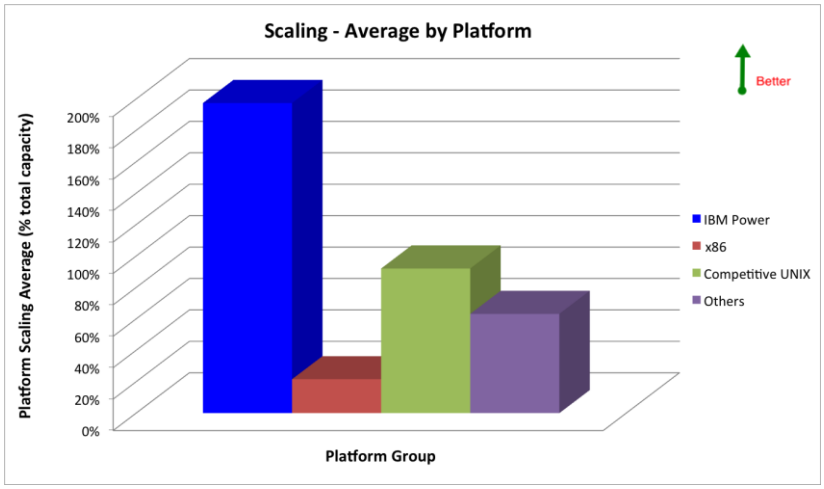
Since in any rollout, the number of transactions, associated data, and processes can significantly increase, workload can develop highly spiked capacity demand. A view of the correlation between BDA rollout events and the study platforms highlights the resiliency of each, based on the number of outages and slowdowns in the month following a rollout event.

The Power platforms clearly show the architectural resiliency in the significantly lower number of outages and negative performance impacts after the rollout events. This difference is as great a difference as 73.16% fewer outages and 92.56% fewer performance problems for Power than the other options. The resiliency of the Power architecture absorbs the impact of the majority of workload spikes, freeing up the operations personnel to handle the more challenging issues that arise on a daily basis.

> *"We have two different big data analysis groups in our company. One of them is deployed on [x86], while the other is on the Power platform. The [x86] one is constantly having to be interacted with, sometimes for reboot, frequently for tuning. It was assumed that the system running on Power was less active, but when we tracked workload, it turns out that it is doing more than 6 times the work, but it seldom needs attention.*
>
> *Even the small variations seem to cause the [x86] system fits. When we look at upgrading that system I am definitely pushing for moving it to Power. I am too old for getting middle of the night escalation calls!"*
>
> Healthcare operations manager

The varied focus areas of BDA translate into a complex profile for the selection of a deployment platform. The requirements for resiliency, availability and speed make the selection of an architecture challenging. While the normal business perspectives of satisfaction, expense, etc. are still valid, the technical capacity for deployment support becomes extremely critical in the selection process. The following chart shows the reported thresholds by architectural group that was reported by the organizations in this study.

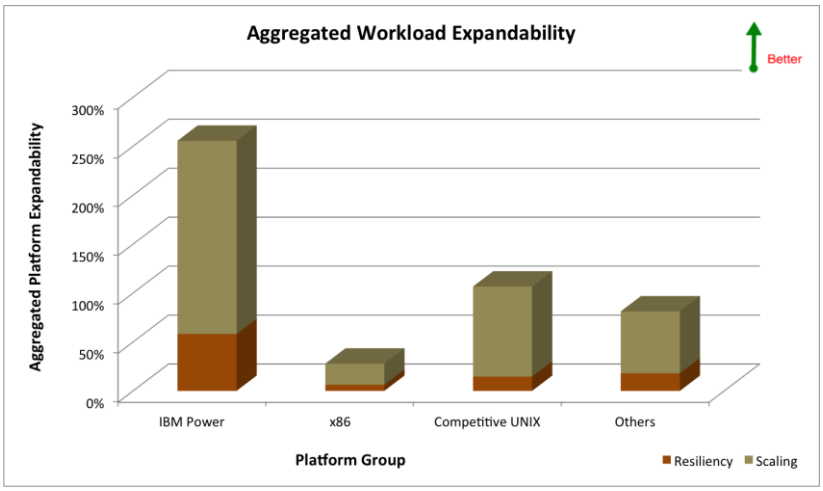
Resiliency is only one dimension of the true impact of the platform in BDA deployment. Given the large growth potential for the analytics, the ability to scale within the same box becomes a significant consideration. While smaller and less expensive platforms can be very attractive from a capital expenditure perspective, the overall cost to agility where time equals money is sensitive to staffing demand, where frequent and time-sensitive infrastructure additions can be prohibitive. This is also a factor when the platform deploying the BDA workload has limitations that require a "forklift" upgrade. The movement of the platform out of the deployment and the insertion of a new one are disruptive to both agility and staffing.



The ability to scale within the installed platform has been shown to contribute to both the lower staffing costs that have been presented before and to the satisfaction of the agility requirements.

The scaling within an installed platform is based on the ability to add further processors, memory and other factors to the platform, allowing it to support more work. In some cases, it means that inactive processors have been turned on, such as the Capacity on Demand functionality that is offered by HP or equivalent product offerings from IBM, Oracle, etc. In other cases, it refers to the ability of the physical structure of the platform to be expanded. The ability to add memory, network channels, etc., along with the tools to allocate those new resources appropriately and smoothly, builds the underlying scaling factor reported by customers within the study.

If both of these factors are aggregated, the platforms and groups can be viewed from the perspective of the flexibility of the platform strategy, rather than any specific model or configuration. This total view of the architecture is useful when balancing the reported rise in overall BDA workload, factored from the base operation workload. From the customer reports within the study, the IT resource demand for a fully deployed BDA implementation is averaging 16 times the pre-BDA operational levels. The huge increase in costs, complexity and scope of the IT architecture is a critical factor in the business case for BDA.



The chart provides a visual perspective on some of the reasons behind general platform practices. The x86 approach of commodity deployment is a reasonably successful mechanism for dealing with underlying bottlenecks in memory access and I/O, as well as compensation for a more inflexible configuration for a "balanced system". That target system is not commonly found in today's market, but is sufficient for the x86 audience and market. However, the lack of expandability is a significant detriment to a BDA deployment, where memory bandwidth and throughput are key characteristics, and where balanced systems do not exist.

The threshold of scaling is therefore an extremely important consideration when choosing an optimal platform for BDA. It affects all aspects of the BDA deployment, from agility through cost. However, the picture of expandability is not yet complete when looking at the different architectural utilization levels. These levels add further differentiation to the potential of an architecture to handle increasing BDA workload. This utilization level creates a multiplier for the expandability, resulting in slightly altered practical limits for the BDA deployments.

The ultimate scaling threshold, where a new physical box is needed, shows a significantly large advantage for the Power platform, with a higher operational threshold of more than 1532.31%, or as much as *15 times*. This encompasses both the resiliency of the platform and the ability to scale up in workload. This translates directly to the bottom line for many of the BDA user group and reflects critical support for that operational advantage. It also contributes to the goals of the different groups when it comes to agility, whatever the form the agility definition may take in a specific organization.

> *"…constantly running around to install, reconfigure and tune more x86 boxes for the big data workload. The demands for more boxes, and of course more money, were happening several times a month at the end. Our users were unhappy, our CFO was VERY unhappy and my techs were exhausted. We had to even get in consultants to help, which caused more slowdown to get them oriented and, of course, cost lots more money. It was a terrible nightmare. We ended up about 840% over budget – which is a really bad place to be.*
>
> *I lost over half my experienced staff. You cannot treat people that way and expect them to stay, but we were in such a crunch time we demanded it. Overall, it was a pretty horrible time.*
>
> *About the only good thing that came out of it is when you guys tell us that we need more resources, more staff, etc., it will be a much more receptive audience."*

Financial services CIO

The security for big data and business analytics systems can be very challenging. By nature, the mixture of operational data, structured data, unstructured data, user profiles, plug-in components and so on, provide many avenues for security incursions. Virtualization is also a factor, since it creates additional vulnerabilities and complexity. If the big data and business analytics application security contributions are assumed to be the same across all the platforms, then the underlying security in the platform architecture can be seen. This forms another layer of protection for customer process, data and intellectual property.

The security arena for BDA has become even more complex in recent months. There are a substantial number of growing threats in this area, with new incursion types being discovered every day. The new dimension of threat hinges on the underlying inclusion of unstructured data. This data easily can contain viral "land mines" that circumvent the security measures around structured data. Much of the unstructured data that is being swept into BDA comes from emails and documents. These data sources are extracted differently, with less scrubbing and scrutiny than data destined for databases. Since the protective processes and filters that are applied to this unstructured data frequently are lacking both controls and governance, threats can be moved closer to the core organizational information assets. When this happens, the security that is in place within the OS and architecture stack is the main differentiator.

> *"We have been concentrating on beefing up our [x86] security over the last year, trying to make sure that we do not have any further loss of our IP. The new measures are very strong and we have all been pleased with them. So we were totally blindsided by a large and blatant attack that successfully navigated our protections and destroyed large amounts of data. We think, but cannot be sure, that additional IP loss also occurred.*
>
> *The threat came through our big data analytics. Apparently there is some type of hidden threat that rides in on emails or images. At this point, we have had to put our big data stuff on hold while we develop a response. That delay was extremely unpopular, to say the least. The reliance on the big data insights is growing and this totally derails that until we can figure out how to block this type of threat.*
>
> *You asked us about the costs of the attacks. With the loss of the IP, the destruction of over 100 GB of important data and all of the other mess, I would put the loss at around $100 million, and the total is still growing.*
>
> Pharmaceutical CIO

SIL has seen an extremely rapid rise in this type of incursion over the last 14 months. Costs of this incursion type are very large, since both the actual data and the derived results are vulnerable. Detailed tracking has shown that even though IBM Power platforms have received tainted data in documents and emails, that the embedded attacks have not been successfully launched, due to the tightly integrated security stacks in the Power platform, the OS and Power/VM virtualization. The accompanying table shows the platform security breaches that were reported in the study organizations for their big data and business analytics production environment.

The only reported security incursions on Power platforms for the BDA deployments were tied to misuse of password access. No destruction of data was reported for Power-servers. However, the number of other platform security incursion reports is on a radical rise, with totally new incursion types that exploit the basic structure of the BDA implementations.

In each of these cases, the threat has entered the customer environment in email, pictures and other unstructured data. With the spotty governance of this data, less rigorous protection is employed to prevent the threat from breaching the perimeter. When the BDA processes incorporate the infected data containers, damage to data and IP result. A short summary of the reported incursion counts, damages and tracked action notes are documented below. Further studies will focus more on this area in detail.

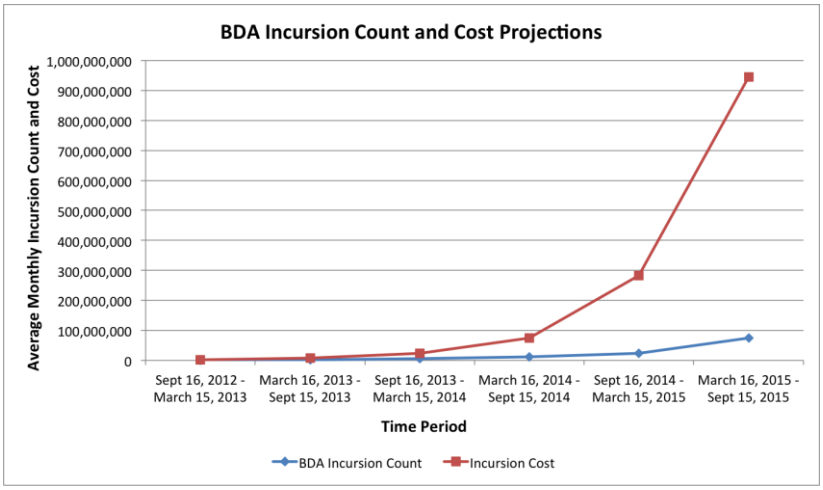
Platform	Attempted Incursion Count	Entry Incursion Count	Total Incursion Count	Estimated Damage (USD)	Notes
IBM Power	215,524	0	0	0	No successful incursions
X86	1,406,584	831,224	14,594,261	361.4 M	Only initial damages, follow-on damages expected
Competitive UNIX	904,268	205,485	6,953,227	96.1 M	Only initial damages, follow-on damages expected
Other	113,980	48,638	109,253	84.5 M	Only initial damages, follow-on damages expected

SIL maintains a Security Watch oversight, with organizations worldwide reporting on incursions, costs and effects. When that larger group is examined for BDA security effects, the numbers are similar, providing a slightly expanded view of the overall threat level. Since no platform is secure from internal password misuse and other internal

violations, the percentage of the incursions that stem from this type of threat are annotated.

Platform	Attempted Incursion Count	Entry Incursion Count	Total Incursion Count	Estimated Damage (USD)	% Internal Process Violations
IBM Power	635,267	0	68	$0.003M	100%
X86	11,268,083	7,919,752	44,594,261	$1,384.5 M	0.14%
Competitive UNIX	8,904,873	4,805,233	16,953,227	$687.1 M	5.29%
Other	289,305	132,580	209,253	$395.3 M	13.27%

Over the last 6 months, the frequency and impact of BDA-leveraged incursions have increased substantially. This increase can be seen in both the number of attacks and the financial cost of those incursions. The growth of both of these metrics is shown on the accompanying chart. Other components in the chart show a projection of future levels of these two metrics based on monthly acceleration over the next two-year period. The figures are expressed as a monthly average.



The costs and counts are expressed as averages, which can be somewhat misleading in that it minimizes the actual perceived fiscal exposure. While the possible financial risk is high based on the average figures, the actual exposure is much higher. It is critical to balance the short-term acquisition costs of the platform with the proven cost and security exposure of the entire BDA operational environment.

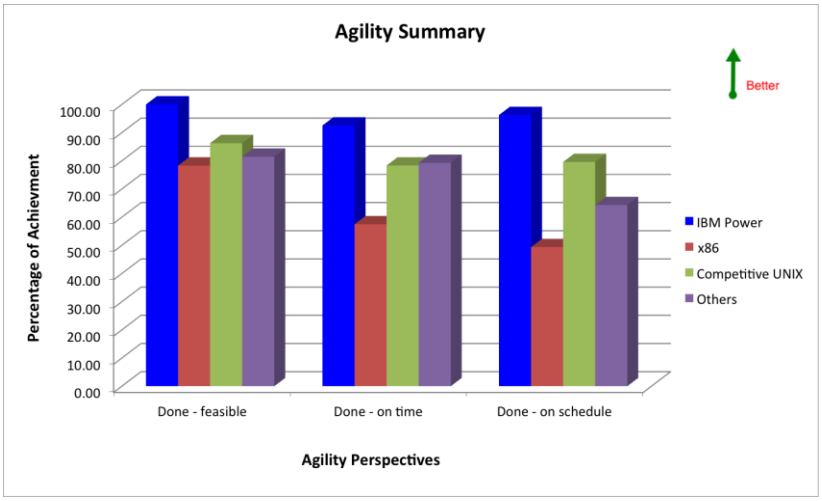
It remains to be seen if the BDA incursions will start to show on Power servers but up to this point, no successfully breach of the OS or Power/VM virtualization for BDA has been reported. SIL will continue to monitor this carefully. This strength makes a singular advantage in this growing threat area.

Agility

The views of agility differ among BDA users, as previously discussed. For three of the perspectives on agility it is possible to build a metric for validation. These three are:

- Execution completion (feasibility)
- Execution time compliance (speed)
- Execution schedule compliance (dependability)

The fourth view of agility involves the value of the insight that results from the BDA itself. Since that is an organization-specific evaluation and does not translate across organizational lines well, it has been left out of the comparison for architectures.



A comparison of the three dimensions of agility that remain, highlights several areas of differentiation in the architectures. It also provides a summary perspective on the success of the big data and business analytics system implementation. The ability to execute quickly and complete the analysis is high for several of the architectures, primarily the Power and competitive UNIX platforms. This feasibility evaluation shows that the variance between the highest rated architecture (Power) and the lowest rated one (x86) represents a substantial 21.68% advantage for Power.

The advantages for the Power architecture increase as the facets of on-time execution and on-schedule execution are added to the metric. This is where the resiliency of the architectures starts to show a significant impact on the repeated performance of BDA processes. When the simpler view of execution timing is examined, the advantage that the resiliency brings to the Power BDA deployments shows in a 37.87% advantage over the competitive field. When more complication is added by including the dimensions of dependability and consistency, that advantage increases to 48.72%.

The results in the analysis of agility for BDA reveal a facet of platform selection that is somewhat difficult to translate between the technical aspects of the architecture as a tool for business and the business results themselves. In this one area of focus, the affect of the resiliency and throughput of the entire platform can be correlated to the end results of agility, with its aspects of feasibility, speed and dependability.

"...the Oracle systems that we have running on Power platforms are very stable. I was concerned about implementing this "big data" analytics because we cannot afford to mess up our operation. However, the effect has been pretty transparent to everyone, except operations itself. Our jobs are still coming out on time. And the new information is pretty interesting."

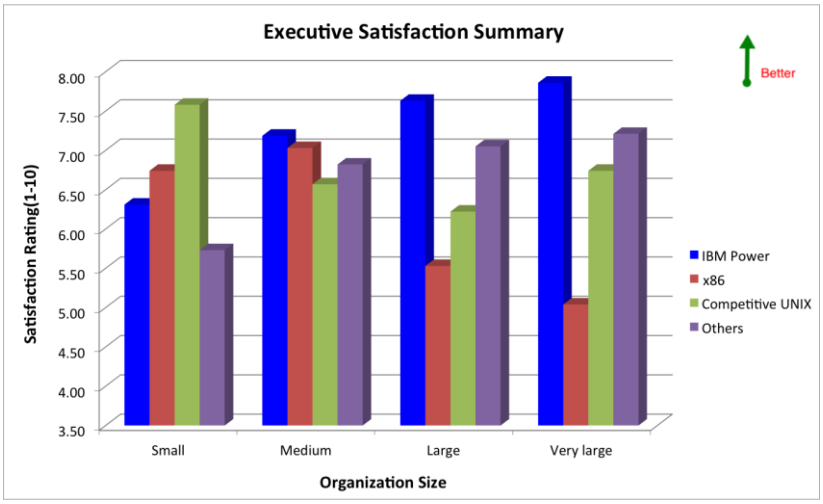
Manufacturing CIO

"We got pretty ambitious on the Big Data front. It was obvious to us that it would have value, as long as we could figure out how to incorporate the findings in a strategic manner. So we slammed a lot of different data together to start looking – and I have to say our IBM (Power) boxes have been great – no problems, no crashes, etc. The combination of SAP and all of the other stuff we are looking at was an area of concern, but so far, only good things are coming out of it."

Retail CEO

Customer Satisfaction

The ultimate metric on a successful implementation is *customer satisfaction*. SIL tracks this split out between the executive management and the operational input from a customer, since the perspective of the customer may radically differ between those two groups. The satisfaction of the customer executive management about their IT systems tends to focus on the application, rather than the deployment platform, although no application can work as well with a poorly configured or fragile big data and business analytics deployment. That being said, the satisfaction with IT implementation and operation provides the most general metric for evaluation. This satisfaction rating was obtained from a large group of customers and provides a singular perspective on the overall success of the big data and business analytics system implementation. While this is a subjective rating provided by high-level organizational management, it does provide the business' actual perception of success.

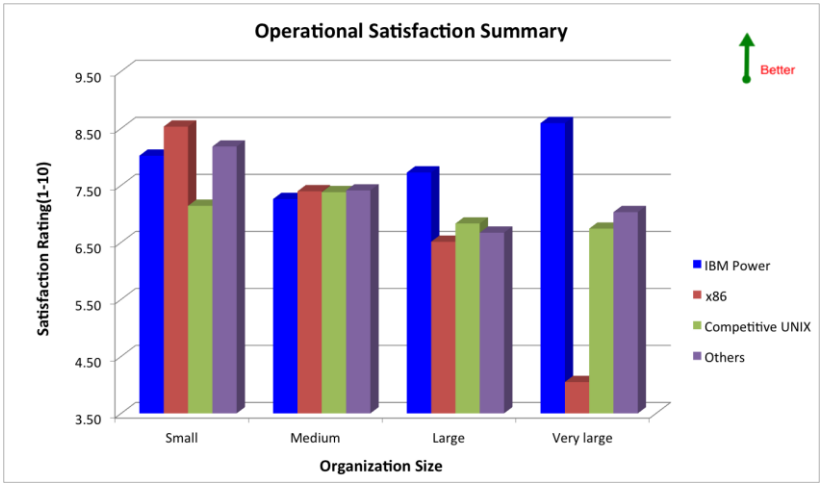


The executive satisfaction is better by as much as 55.95% for IBM Power, with the difference in satisfaction increasing as the size of the organization, and the size of the BDA effort, grows. The advantages seen by the reporting clients show increasing satisfaction in the big data and business analytics deployments run on Power, although there are some further areas of interest in the satisfaction ratings from this group.

While the specific satisfaction rating can be affected by management techniques, application design and other factors, the relative comparison is a legitimate indicator of how well the operating system supports the processing at the organization. The top reasons cited by reporting customers for satisfaction ratings on Power were:

- Smooth running operation with little downtime and complaints
- Adherence to planned budget levels
- Transparency on performance to operational systems

During the study data collection, many of the executives commented on their operation, satisfaction and other areas of the big data and business analytics implementation. Two of the typical comments are encapsulated in the quotes above. One of the quotes is pertinent to BDA running against an Oracle database and over 2.3 TB of unstructured data on Power platforms, while the other addresses BDA that includes SAP operations running with an Oracle database, two additional DB2 databases and 4 TB of unstructured data running on Power.



The operational perception of the customer, based on a variety of component metrics (e.g. support levels, communication, price, etc.), demonstrates satisfaction and success at the most generic level. This satisfaction metric is different from the overall satisfaction metric described earlier, in that the previous metric was gathered from the executive management level, while this metric examines the feedback from the operational side of the organization. This specific metric comes from information reported both by the IT departments and the line-of-business (LOB) groups. Since both areas are directly involved with the execution and use side of the deployment, their satisfaction ratings come from similar perspectives and are grouped together.

The variations in operational satisfaction do not favor the Power platform at the smaller organization level in general. However, there is a significant difference in satisfaction as large amounts of data are included in the BDA efforts. In the larger deployments there is as much as a 112.10% higher operational satisfaction with the Power platform.

The satisfaction of the IT operational staff and the LOB with the Power big data and business analytics deployments reflect the reliability and resiliency of the platform as a deployment choice. The most highly cited reasons for the satisfaction were:

- Ability to shift resources as needed
- Smooth running operation with little downtime and complaints
- Consistent runtimes and load

More than 98% of the reporting customers cited one or more of these reasons for their satisfaction.

"We are running our big data analysis on IBM Power platforms. We have had no crashes at all, and have managed to expand our big data efforts by ten-fold in last year without having to really do much but add some additional storage space."

Finance company director

"Big data is the new catch word, so of course we need big data. Sorry about the sarcasm, but the reality is that we are doing more and more wide-scale analysis and it is useful. Our Power box is what we run all of our big-view stuff on. That was deliberate since it is the only one that we knew would not crash or mess things up. It really helps that we can move resources where we need them. Saves me LOTS of money."

Telecomm CIO

Conclusion

The Solitaire Interglobal Ltd. analysis of big data and business analytics deployments shows substantial advantages to incorporating IBM Power offerings within an IT architecture. The overall findings can be summarized as shown in the following table.

The leader in each category is marked with a green highlight. When separate rankings were discussed in the study analysis for different organizational size categories, those categories are presented in discrete ratings. If the summation is presented in an overall form, the ranking is identified only in the "general" column.

Metric Scorecard - Summary

	IBM Power					x86					Competitive UNIX					Others				
Org Size	Sm	Med	Lrg	X-Lrg	Gen	Sm	Med	Lrg	X-Lrg	Gen	Sm	Med	Lrg	X-Lrg	Gen	Sm	Med	Lrg	X-Lrg	Gen
Cost - TCO	2	4	4	4		4	2	2	1		3	3	3	2		1	1	1	3	
Cost - TCI	3	4	4	4		4	2	1	1		2	3	2	2		1	1	3	3	
Staff	2	4	4	4		3	2	1	1		1	1	2	2		4	3	3	3	
Risk - Resiliency					4					1					2					3
Risk - Scaling					4					1					2					3
Risk - Security Incursions	4	4	4	4		1	1	1	1		2	2	3	3		3	3	3	2	
Agility - Feasibility					4					1					3					2
Agility - Speed					4					1					2					3
Agility - Dependability					4					1					3					2
Customer satisfaction - executive	2	4	4	4		3	3	1	1		4	1	2	2		1	2	3	3	
Customer satisfaction - operational	2	4	4	4		4	3	1	1		1	1	3	2		3	2	2	3	

The leader in each ranking is assigned a weight of 4. The runner-up in that category is assigned a weighting of 3, and so on. The bottom ranked platform has only 1 point for that metric.

The ranking scorecard shows the significant advantage that the IBM Power platform has in many areas and in most organizational sizes. There are some notable holes, however, in the general leadership position that Power has. These are primarily for organizations that are smaller. In these deployments, the complexity, amount of data and resource requirements that Power provides often do not make the best general business case. In any of the averaged situations there are some exceptions, however. Each organization should be conscious of both the opportunity and the challenges that come with BDA and factor in its specific requirements, situation and goals.

The purpose of this analysis was to examine the real-world impact on businesses that deploy big data analytics on a variety of platform, including IBM's Power platforms. The comparison group includes Power, x86, competitive UNIX and a group of miscellaneous platforms. The metrics used to analyze the differences in platforms were both objective and subjective. The objective metrics include reported data points on costs, run times, resource usages, security and so on. The subjective metrics include responses on various levels and sources of customer satisfaction. While overall customer satisfaction uses a variety of qualitative and quantitative measures, it still provides an end-result measurement of deployment success for the customer. Since BDA adoption is highly correlated to the success that a customer has seen in the initial stages of deployment, all of the metrics, i.e., cost, risk and satisfaction, are based on a perception of initial success.

A few of the highlighted findings can be seen in the quick summary below.

Quick Summary

Category	Commentary	Quick Byte
Risk – Resiliency	Resiliency is higher for Power than for the other architectures by as much as 89.33%. This means that Power resiliency is as much as *9.37 times* as the competition.	Resilient platforms like IBM Power mean that overall cost and disruption is less for BDA deployments.
Risk - Scaling	The critical scaling within a platform that is needed for BDA deployments is a massive advantage for IBM Power, with scaling that is as much as 42.76 times higher than the competitive platforms.	If you want massive options for scaling an explosive BDA deployment, choose Power.
Risk - Agility	Whether the view of BDA agility is for simple feasibility, speed of execution or dependability, Power provides some significant advantages. With the agility advantages to Power running as high as 48.72%, it becomes a major consideration when looking at BDA and agility.	For dependable, consistent and speedy BDA, Power provides the best solution.
Risk - Security	No destruction of data was reported for Power servers, while competitive platforms report rising damages.	The safest place for BDA is Power.
Total Cost of Information	The architectural characteristics of Power deliver cost effective workloads and execution for BDA, with cost savings as much as 49.88% for TCO, 74.65% for TCI.	To make the most use out of the overall IT spend, put BDA on Power.

These key findings are all substantial reasons to consider Power platforms for an organization's big data and business analytics deployment. If the study organizations are indicators of the growth in BDA IT resource demand, companies that fully deploy BDA should expect overall IT workloads and storage to expand significantly. With processing, memory, I/O and other resource requirements increasing to 16 times the base operational load, the choice of platform is critical to an organization's market position and profitability. Irrespective of the high customer satisfaction, the more objective metrics evaluated in this study would indicate a significant correlation between the choice of platform and the ongoing success, cost, agility and security of the BDA deployment.

Research Note There are significant variations in satisfaction based on how well an organization utilizes the results of BDA. The main factors that affect the underlying satisfaction are: - Percentage of data covered in BDA - Integration of the BDA with the organizational decision-making - Maturity of the BDA practice Each of these causes significant variation in the underlying satisfaction with BDA overall, and influences the return that organizations see from their BDA practice. Additional SIL research will be available in the forthcoming months in this area.

The economic benefits of selecting IBM Power platform architecture as the big data and business analytics deployment choice are also apparent in overall expense. The IBM Power platform provides a strong and stable base for the critical big data and business analytics processes, irrespective of deployment size. These advantages increase the effectiveness of big data and business analytics deployment and translate to real-world positive results experienced and reported by the businesses in this study.

The strong virtualization functions included in the Power offering also make a measurable difference. These functions provide the ability to sweep capacity resources to targeted processes, and result in the need for fewer overall system cycles. Coupled with the allocation automation, personnel time, hardware, software, and personnel costs can all be minimized. This produces efficient application deployment and cost-effective expenditures, while displaying a risk profile that is substantially lower than the other solutions examined in this study.

The rising security threat within BDA has pointed to another significant area of concern for an organizational BDA deployment. An organization is opened to considerable risk with the possibility of embedded land mines of viral code in the unstructured components of BDA. With its tightly integrated security stacks, the Power platform demonstrates a very secure architecture, which limits losses due to incursions.

Overall, critical effects on staffing, security, cost and satisfaction makes the IBM Power platform a strong contender for an organization's big data and business analytics deployment choice.

About Solitaire Interglobal Ltd.

Solitaire Interglobal Ltd. (SIL) is an expert services provider that specializes in applied predictive performance modeling. Established in 1978, SIL leverages extensive AI technology and proprietary chaos mathematics to analyze prophetic or forensic scenarios. SIL analysis provides over 4,100 customers worldwide with ongoing risk profiling, performance root cause analysis, environmental impact, capacity management, market trending, defect analysis, application Fourdham efficiency analysis, organizational dynamic leverage identification, as well as cost and expense dissection. SIL also provides RFP certification for vendor responses to government organizations around the world and many commercial firms.

A wide range of commercial and governmental hardware and software providers work with SIL to obtain certification for the performance capabilities and limitations of their offerings. SIL also works with these vendors to improve throughput and scalability for customer deployments and to provide risk profiles and other risk mitigation strategies. SIL has been involved deeply in the establishment of industrial standards and performance certification for the last several decades and has been conducting active information gathering for the Operational Characterization Master Study (OPMS) – chartered to develop better understanding of IT-centric organizational costs and behavioral characteristics. The OPMS has continued to build SIL's heuristic database, currently exceeding 110 PB of information. The increased statistical base has continued to improve SIL accuracy and analytical turnaround to unmatched levels in the industry. Overall, SIL runs over 52,000 models per year in support of both ongoing subscription customers and ad hoc inquiries.

Methodology Notes

In order to understand the impact of IBM Power platforms as a key part of an organization's big data and business analytics deployment, a significant number of deployments were examined. The relative degree of difference in operating behavior for each factor, i.e., total number of outages, etc., was then compared to understand the net effect of the respective combinations. The effects were observed in general performance and capacity consumption, as well as other business metrics.

The approach taken by SIL uses a compilation and correlation of operational production behavior, using real systems and real business activities. For the purposes of this investigation, 31,000+ environments were observed, recorded and analyzed to substantiate the findings. Using a large mass of customer and industry experiential data, a more accurate understanding of real-world behavior can be achieved. The data from these systems was used to construct a meaningful perspective on current operational challenges and benefits. The reported behavior of the systems was analyzed to isolate characteristics of the architecture from both a raw performance and a net business effect perspective. All input was restricted to those organizations using systems in versions that were current in calendar years 2011-2013. Since many of the components in this environment have releases at staggered points in time, only those components that were either the current version or a -1 version based on those calendar restrictions were included in the study. Additional information on the methodology and study diversity can be found in additional methodology notes at the end of this document.

In a situation such as that presented by this study, SIL uses a methodology that incorporates the acquisition of operational data, including system activity information at a very detailed level. It should be noted that customers, running on their production platforms, provided all of the information. It is essential to understand that none of the data was captured from artificial benchmarks or constructed tests, since the value in this study comes from the understanding of the actual operational process within an organization, rather than the current perception of what is being done. Therefore, these sites have tuning that is representative of real-life situations, rather than an artificial benchmark configuration. Since the focus of this analysis was not to tightly define the differences among different minor variations of operating system or hardware, the various releases were combined to show overall architectural differences. This provides a more general view of architectural strategy.

In order to support the comprehensive nature of this analysis, information from diverse deployments, industries, geographies, and vendors were obtained. In any collection of this type, there is some overlap that occurs, such as when multiple vendors are present at an organization. In such cases, the total of the discrete percentages may exceed 100%. Those organizations with a multi-layered deployment, such as multiple geographical locations or industrial classifications, have been analyzed with discrete breakouts of

their feedback for all metrics. Additional filtering was performed to eliminate those implementations that substantially failed to meet best practices. Since the failure rates, poor performance and high costs that appear in a large number of those implementations have little to do with the actual hardware and software choices, these projects were removed from the analytical base of this study.

The industry representation covers manufacturing (31.80%), distribution (21.01%), healthcare (5.87%), retail (8.32%), financial (6.62%), public sector (4.92%), communications (21.34%) and a miscellaneous group (0.12%).

The geographies are also well represented with North, South and Central America providing 45.79% of the reporting organizations, Europe 31.91%, Pacific Rim and Asia 17.83%, Africa 2.40%, and those organizations that do not fit into those geographic divisions reporting 0.06% of the information.

Since strategies and benefits tend to vary by organization size, SIL further groups the organizations by the categories of small, medium, large and extra large. These categories combine the number of employees and the gross annual revenue of the organization. This staff count multiplied by gross revenue creates a metric for definition that is used throughout the analysis. In this definition, a small organization could be expected to have fewer than 100 employees and gross less than $20 million, or a value of 2,000, e.g., 100 (employees) X 20 (million dollars of gross revenue). An organization with 50 employees and gross revenue of $40 million would have the same size rating, and would be grouped in the analysis with the first company. The classifications used by SIL use thresholds of 2,000 (small), 10,000 (medium), 100,000 (large) and 1,000,000 (extra large).

The information in this study has been gathered as part of the ongoing data collection and system support in which SIL has been involved since 1978. Customer personnel executed all tests at SIL customer sites. The results of the tests were posted to SIL via the normal, secured data collection points that have been used by those customers since their SIL support relationship was initiated. As information was received at the secure data point, the standard SIL AI processing prepared the data in a standard format, removing all detailed customer references. This scrubbed data was then input to the analysis and findings.

Attributions and Disclaimers

