NASCIO
Representing Chief Information
Officers of the States

NASCIO
Cloud
Computing
Series

Capitals in the Clouds

The Case for Cloud Computing in State Government Part II:
Challenges and Opportunities to Get Your Data Right

NASCIO Staff Contact:
Eric Sweden
Program Director,
Enterprise Architecture &
Governance
NASCIO

NASCIO represents state chief information officers and information technology executives and managers from state governments across the United States. For more information visit www.nascio.org.

201 East Main Street, Suite 1405
Lexington, KY 40507
Phone: (859) 514-9153
Fax: (859) 514-9166
NASCIO@AMRms.com
www.NASCIO.org

Introduction

As discussed in NASCIO's first issue in this series, cloud computing as arrived and will gain emphasis as an enabler of efficiencies and innovation. If there was one foundational theme for describing cloud computing as an *enabler* of business and information management strategy in state government, it should be the concept of thinking "enterprise." However, an unmanaged cloud environment will create just the opposite situation. If departments, agencies, or divisions are able to acquire cloud services, typically public cloud services, independent of an *enterprise governance structure*, particularly data governance, state government will be creating the next generation legacy system *silos* that are not integrated or interoperable. This might be termed *cloud legacy*, or *legacy cloud system silos*.

In August, 2010 NASCIO, TechAmerica, and Grant Thornton LLP published the results of a survey of state chief information officers - *2010 State CIO Survey – Perspectives and trends from state government IT leaders*.[1] Fifty-percent of these respondents stated that they are investigating cloud computing. One-third reported that they are running active or pilot cloud projects. The word "investigating" is an important one. This can be interpreted to mean state CIOs are being somewhat, and justifiably, cautious as they evaluate alternatives for delivering IT services. Cloud computing is one of many approaches to sharing resources. And, cloud computing itself offers multiple alternative approaches for deployment and service delivery.

As stated in NASCIO's Part I in this series on cloud computing, the predominant stress that is driving state and local governments toward *serious consideration and adoption* of cloud computing are the continued budget woes and demands for operational savings. These pressures must be managed intelligently to avoid pushing government into a future situation that could constitute greater

What is Cloud Computing?

A model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.

National Institute of Science and Technology (NIST)

cost, and more difficulty in achieving interoperability of government lines of business and government jurisdictions.

Is Cloud Computing Really Something New?

Cloud computing presents a technological innovation, or as some say, a rebirth of an old idea. There are other flavors of resource sharing that have appeared in past, and are still employed today.

- Timesharing
- Client/Server
- Utility Computing
- Grid Computing
- Application Service Provider
- On-demand Dynamic Provisioning
- Thin Client
- Net-Centric Computing
- Software-as-a-Service

Anyone "selling" cloud services is probably aware of this fact. Cloud computing is not as much of an innovation as it is a rebranding and a rescaling of old ideas. Technology advances such as ubiquitous high-speed internet connectivity, virtualization and inexpensive mass storage have made cloud services a viable option for replacing the traditional computing paradigm. However, the prolific messaging and marketing of cloud together with the current budget crises have opened the door so that business and IT are in conversation about something that is very strategic – *removing unnecessary redundant IT investment*. The term used in one state is "optimization." Optimization and reduction in redundant investment can certainly be justified as a concept and a business goal. There are necessary steps in the process of "evaluating" cloud computing, *as well as other shared resource approaches*, as an approach for achieving this optimization. Several significant considerations that must be addressed are *data issues, jurisdictional issues, and human resources issues*. These are not simple matters. They are complex, having evolved over a period of two centuries since this nation was founded. Further, these issues have not been faced at the scale now being considered. Over time there will be some very hard lessons learned, unintended consequences, surprises, precedence set, and the emergence of best practices. The focus of this issue brief on cloud computing is *data issues*.

Cloud Computing: One Idea in a Portfolio of Solutions

Before moving onto an exploration of data issues, it must be emphasized that cloud computing is not the only approach for sharing resources. It is also not the only approach that requires this process called *data harmonization*. Data warehouse processes include what is termed Extract, Transform, and Load (ETL). If the data from the various partners is to be brought together into a shared data warehouse, then that data must be examined and evaluated for the inherent semantic business rules that reside in the data. Those rules, or *semantics*, must then be reconciled, or *harmonized*, into a single set of seman-

tics. This step is avoided if data is simply extracted and loaded in a shared *storage* environment. This latter scenario is typical of open data initiatives that are simply *publishing* data sets. However, if the intent is to bring together an *enterprise view* of data involving two or more government lines of business, then the exercise of evaluating these inherent semantics becomes relevant and potentially one of the most challenging activities in the initiative.

Again, cloud computing is not the only approach for sharing resources. There are multiple technologies and best practices that are also gaining greater adoption as state and local government look for ways to bring enterprise thinking and efficiencies into government. These include but are not limited to the following general concepts:

- IT-Enabled Innovation and Improvement

- Effective Governance of IT

- Shared Services

- Managed Services

- Business Intelligence and Business Analytics

- Infrastructure Consolidation and Optimization

- Virtual Computing

Disciplines, frameworks and methodologies employed across government IT include:

- ITIL / ITSM
- COBIT - Control OBjectives for Information and related Technology
- Project Management
- Portfolio Management
- Agile Programming
- Component Development
- Service Oriented Architecture

Examples of these and other approaches can be found in a number of resources including:

- NASCIO's awards database at www.nascio.org/awards.

- TechAmerica's companion report to the 2010 CIO survey – *Why Think Enterprise? – State and Local Government Options for Reducing Costs and Improving Services.*

- Gartner Hype Cycles for Emerging Technologies

State government is already employing many of these approaches and adoption will continue to grow. Cloud providers, whether they are internal government cloud providers, or external commercial providers, can be expected to also be

employing these kinds of disciplines in delivering shared services, or cloud services.

Cloud computing should be considered as one contributing element to a portfolio of ideas and approaches for delivering government services more efficiently, more economically, and, in many cases, more effectively. Some organizations may move completely into a cloud computing environment. Gartner has predicted that by 2012, 20% of *businesses* will have no investment in hardware having embraced a level of cloud computing that no longer requires it.[2]

Life in the Cloud - Data Issues Are Significant

Data is the lifeblood of state government operations and critical for service delivery. The concerns with data are issues that are present in any scenario. The only difference may be the scale or magnitude of this issue with the advent of an enterprise cloud strategy. There are multiple terms that are used to describe the bringing together of two or more organizations, agencies, functions, departments, or jurisdictions. These include but are not limited to:

- shared services,
- consolidation,
- harmonization,
- unification,
- optimization,
- simplification,
- integration, and
- interoperability.

The scope of such efforts can involve two organizations, or fifty. For example, optimization initiatives can vary in scope.

1. Agency only
2. Multi-Agency
3. Multi-jurisdiction within a single state that include counties and cities
4. Multi-jurisdiction across states

We'll use the term *data harmonization* to describe the rationalizing of the data and the inherent business rules imbedded in the data from the partner organizations that are now coming together to share IT services and potentially business processes. In any of these scenarios, *data harmonization* will be significant. In fact, if there is one exercise in past data warehouse projects that has been the primary cause of project overruns it is the process of bringing together the data from the partnering organizations into a *single enterprise view*.[3] Data warehouse initiatives have been facing this issue for some time. This process of *data harmonization* is extremely difficult.

Opportunity to "Harmonize" Data and Business Processes

Cloud computing brings with it *opportunities, issues* and *risks*. The *opportunities* include the possibility for transformation of business processes into an enterprise integration and coordination of business processes. There is also the

possibility of more immediate exploitation of advances in technology. The latter depends on the cloud provider's capabilities for continual innovation. Cloud computing initiatives also bring the opportunity for creating *consistencies in business processes and information management*. The latter occurs because shared services or cloud services are "common IT services." Common services will require agreement on business processes and business rules, and the understanding and structure of business information. That *agreement* will be reflected in a common process architecture and a common information architecture, and achieves *consistency* in government operations. That *consistency* will bring with it a new capability for government lines of business to work together as a single enterprise in delivering government services.

Enterprise business process and information management capabilities should be employed and exploited by policy makers to greatly improve government's ability to *deliver the government intended* as put forth by state constitutions, laws and regulations.

This means, or could mean, more effective capabilities for:

- governors to govern;
- legislatures to deliberate and craft the right laws;
- executive branch agencies to execute the delivery of regulations and services *perfectly aligned* with the intent of legislation;
- executive branch agencies working together in an orchestrated manner;
- courts to adjudicate laws and regulations with a clear understanding of the intent of law and the intent of regulations that support law;
- procurement to purchase products and services with economic and functional efficiency;
- use of analytics to manage government services and programs from an enterprise perspective;

Essentially, delivery efficient and effective government. However, to achieve this will require effective *data harmonization or data unification*.

Challenges facing Optimization and Shared Services – Business Rules / Semantics

Consistency in business rules or semantics is a necessary ingredient in the recipe for success in bringing two or more agencies together into a single enterprise. That consistency is not always easily achieved particularly when the conceptual and logical understanding, or the definition of something, is different from one agency to another. Keep in mind, the existing inconsistencies in data definitions and organization have been created over a significant period of agency and government line of business *independence*.

This disparity in data definitions and attributes doesn't preclude government from creating and employing cloud solutions or other shared resource approaches. It does present the necessity for proper evaluation, reconciliation, and agreement on data definitions, data relationships, and business rules. Data reconciliation is a necessary phase in the path toward cloud computing as well
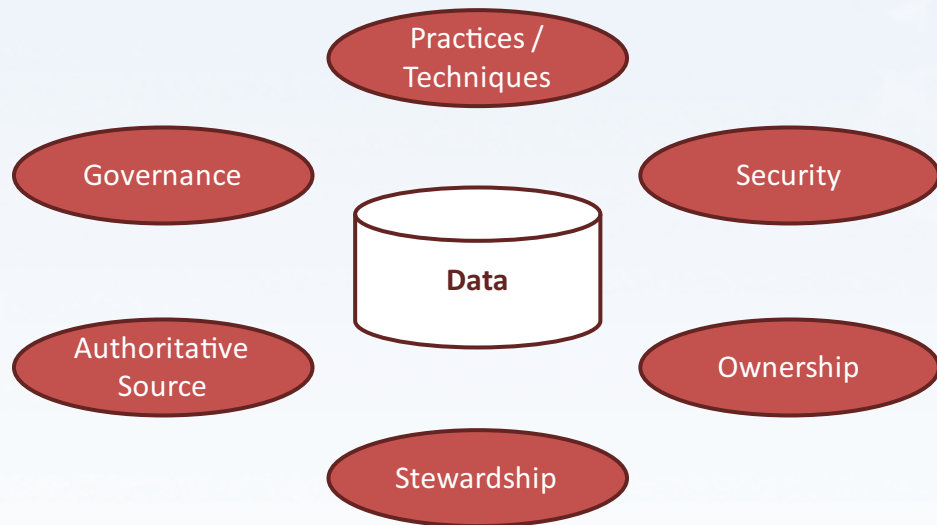
Customer "Lock-In"

Customer "lock-in" is a very common commercial practice and frequently an integral part of pricing strategy. Standards are the antidote to lock-in, and this applies to data just as it does to plugs and sockets.

Data Portability – The Antidote to Data "Lock-In"[4]

as any shared service scenario. Additionally, data quality, data valuation and data security become points of contention when there is a lack of consistency across the organizations that are now coming together to share IT resources.[5]

Agencies may differ in their renderings of subject areas, entities, associations, attributes, integrity constraints, or business rules. Agencies may also differ in how they define and apply aspects of data and data management.

Data Consistency Issues Across Partner Organizations

Issues of potential contention across agencies include:

- data ownership;
- jurisdictional issues;
- governance issues;
- knowledge and information management discipline:
- valuation and security discipline:
- procurement discipline:
- supplier management;
- project management;
- required authorization and approvals

These issues must be evaluated relative to the service provider(s), or cloud computing provider(s) under consideration. Some aspects of data management must not be handed over to an outside service provider. One example is *data ownership* and *data control*. State governments are advised not to sign over ownership of its information or data assets to an outside vendor or hosted provider. Allowing a vendor to assume control of government data, especially mission critical information for public safety, may jeopardize the state's ability to protect its citizens.

This has happened in past, and has created regrettable circumstances for government. One example highlights not only data control and data access, but also the necessity of maintaining an "enterprise perspective" of state data.

State agency data must be seen as a shared resource and therefore appropriately shared across the state government enterprise. In this example, a new state driver license data base project was outsourced to a vendor. Not all stakeholders, or users of the data, were considered or consulted. Public safety related access was not included in the scope of the project. Once the project was completed, the state public safety agencies found that they were required to request permission from the vendor for driver license data through very limiting and time consuming procedures.

Other instances have occurred where small state government agencies or local governments with very limited budgets have sought to save money by moving records management to an outside vendor. Given their limited budget and shortage of other resources, these agencies did not apply an *enterprise perspective*, or a long term view, in project scoping, contracting, terms of service, system design, and delivery of the solution. Although these agencies saw cost savings in the beginning, the full impact and cost, was not encountered until later years. The vendor's solution and underlying technology eventually became obsolete and could not meet new requirements from the agencies. The agencies moved to a new solution and *expected* to extract historical records from the first vendor's system. However, the fees charged by the first vendor for the agencies to "re-obtain" historical government data were substantial. The agencies could not afford to pay such fees. The result: the data was not recovered by the agencies. That data, which is government data, remains with the first vendor.

Risks related to cloud computing include new possibilities for loss of data, lack of adequate internal controls and security, and vendor and technology lock in. These risks increase when moving from highly controlled private clouds – to highly controlled community clouds – to highly controlled public clouds - to loosely contracted public clouds. The risks don't preclude government from going down the path of cloud computing, nor are they unique to cloud computing. However, these inherent risks must be recognized and properly addressed to ensure state government information and data are secure and available; state government does not incur undo risks; and all parties perform per expectations. The issue must be addressed within governance, operating procedures, contracting, and vendor management. Similarly, contentions in data management discipline do not preclude agencies from forming a community of interest that share IT services. As stated in NASCIO's publication, *Transforming Government Through Change Management – The Role of the State CIO*[6] - consideration must be given to "inter-enterprise architecture" as well as recognizing potential collisions up front – admit they exist, then deal with them in a constructive way.

As stated, cloud computing initiatives bring the opportunity to *examine the underlying business processes and business information management*. This exercise will achieve consistency across the parties that are now sharing a service. It also affords the opportunity to determine if previous understanding of the actual data rules was correct. It may be determined that some representations of data relationships were only *approximations* that were either convenient, or adequate for certain government services.

When a variety of agencies, or government business services, now begin to share data and computing services, there must be a common agreement on

data definitions, entities, attributes, relationships, cardinalities, integrity con-straints, naming and aliases, and other components of conceptual and logical data models, and modeling. The most robust rules may have to prevail if com-puting services are to be shared. That can in fact create significant change for those agencies that have previously been using more simple approximations of the actual data relationships. There will also be times when the complexity of existing business rules should be challenged. It may be discovered that a level of complexity is not necessary or useful. On the other hand, if agencies have been using mere approximations of actual business rules, that will have to change as well. Over simplification and over complexity are both to be avoided. What is required is a *single, accurate, version of the truth*.

What does this "enterprise" term mean?

The term *enterprise* is a generic term that can refer to a state government, a state agency, a county, a city, a for-profit company, a non-profit organization, an academic institution, or an individual. As *enterprises* come together to share resources, and potentially use a cloud deployment model to do so, they must evaluate and reconcile their business rules to create a new single enter-prise set of business rules. In the process of evaluating and reconciling busi-ness rules, the partner enterprises may discover that each has been dealing with different business scenarios – both of which represent valid business rules. Bringing these enterprises together into a new single enterprise now intro-duces, or uncovers the fact that different business rules are applied based on certain conditions. The two separate enterprises may have a business context that only dealt with a limited set of conditions. Bringing the two enterprise sets of rules together may uncover a more full set of conditions rather than a conflict. This assumes they are going to share common IT resources such as software, data storage, or business process. If they are sharing infrastructure or platforms, then they must evaluate business rules regarding the delivery of that infrastructure or platform. The level of review and reconciliation depends on *what* is being shared. If they are going to share applications, business processes, or business information, then they must evaluate their business processes and business information to determine the level of "fit." It is almost guaranteed that the *business architectures* developed independently for two separate enterprises will differ to at least some degree.

How Different Could Business Rules Really Be?

To use a very simple yet realistic example, two enterprises that keep informa-tion on individuals may have modeled the concept of *person* quite differently and subsequently designed data base structures and the applications that ac-cess these structures quite differently. The two enterprises may vary quite dif-ferently in the level of sophistication in their knowledge, skills, experience and tools for developing data models. Further, a provider of cloud services may have modeled the concept with yet another version of understanding which may present an international view of *person* that accounts for many different structures for geographical and political regions and subdivisions.

Master Data Management

Reference and Master Data Management is the ongoing reconciliation and maintenance of reference data and master data.

DAMA Guide to the Data Management Body of Knowledge[7]

The example here is a very simple illustration of the differences in approaches for modeling the same concept. Logical data models are intended to describe WHAT information the enterprise cares about. There is no "flow" implied as in a process model. Rather, a data model presents how concepts are related to one another. These models will include entities like person, address, and service. Entities are usually depicted using a box. An entity is a person, place, thing, event, or concept. Entities have attributes and relationships. Attributes are characteristics of the entity and in this case include a person's *name*. Relationships describe how entities are related. Relationships are sometimes termed associations by some data modelers. Relationships describe how entities are related and present the business rules of the enterprise that describe *how many*. Relationships can become entities, termed associative entities, if the relationships have attributes that the enterprise cares about. The models presented are *entity relationship diagrams,* or *ERDs*. Enterprises may use different approaches for data modeling such as Unified Modeling Language (UML).

Data models have different purposes so there are different types of data models. Some are used to describe the business information, and the business rules. These are sometimes termed *business information models*. *Conceptual models* are very high level and are intended to explore and discover subject areas. *Logical data models* are sometimes referred to as *fully attributed logical data models*. They should include all the attributes, relationships, cardinalities and *sometimes* detailed integrity constraints. Some data models are used to plan and construct electronic data bases. These are often termed *physical* data models as opposed to *logical* data models. There are many volumes written on data modeling. And, more often than not, these terms are defined differently. In fact, there is such a disparity on terminology, that definition is required when using any of these terms.

"Data models" are often created by people who are only interested in creating a physical database. In many cases, the people creating these models have a single purpose in mind and therefore create models that are not really *logical data models* or *business information models* devoid of physical data base characteristics. However, they may use those terms to describe what they created. For example, XML schemas are sometimes called logical data models when in fact they are physical models that present data in a hierarchical structure. It is not possible to always *accurately* reverse engineer business rules from an XML schema. XML schemas are created to implement information exchanges. They are not appropriate for creating conceptual or logical data models. Bottom line - be careful with *terminology*.

In this example, two separate enterprises have made some decisions about how they will do business and what business rules they adopt. Those decisions are either on purpose and deliberate, or they were never really explored so the data modeler, and data base administrator, or programmer, made some decisions regarding how to render the business rules that subsequently affect how the database was structured. The database structure and the applications that access that structure essentially *impose* strict rules on business operations. The situation here shows a fairly dramatic difference in business rules. In reality, the difference in business rules or semantics may be much more complex.

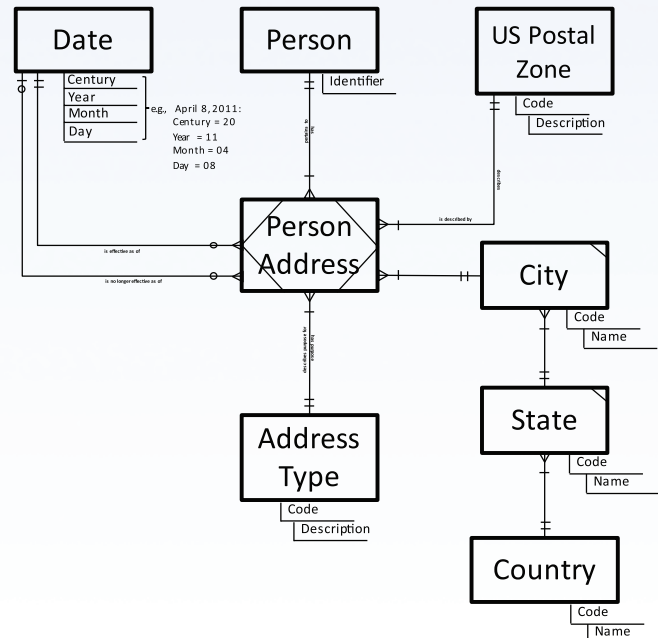## An Illustration of Differing Business Rules / Semantics

The "State of Dismay" is an enterprise that has decided – again, either on purpose, or by default – that it would allow two addresses for a *person*. It has also assumed that these addresses would be in the United States, so there are attributes for the *state* and the *city*. They may also have added *county*.

If the customer, citizen, or taxpayer calls this enterprise to add an address they will enter into a dialogue as follows with either a customer service representative, a call center employee/contractor, or an automated system such as an interactive voice response (IVR). The IVR provides the greatest opportunity to experience the *new dependence on systems and imposed business rules*.
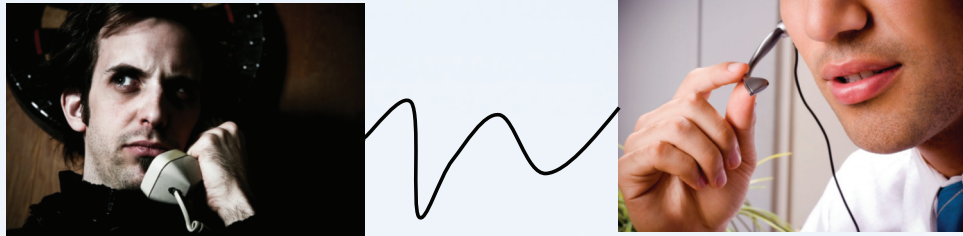

State of Dismay
"Logical Data Model"

State of Enlightment
"*More* Logical Data Model"

Two Very Different Data Models for PERSON

An Operational Scenario with the *State of Dismay*

Citizen call to Call Center – Scenario – State of Dismay

Citizen: "I need to add another address."

Call Center: "So, you mean you want to change one of your two addresses?"

Citizen: "No, I want to add a third address."

Call Center: "Well, that's not possible. You're only allowed two addresses. Which one do you want to change?"

Citizen: "Like I said, I need to add a third address. The other two you have are still in effect."

Call Center: "Like I said, you can only have two addresses. Why would anyone have three addresses?"
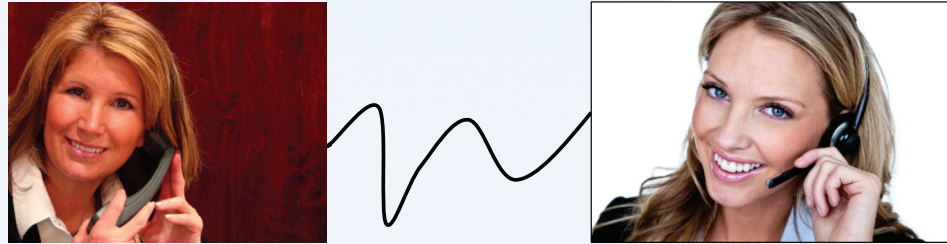
Citizen: "Why can't you add a third address?"

Call Center: "The <u>system</u> only allows two addresses. Now which one do you want to change?"

Citizen: dial tone . . . . .

Call Center: "Hello. Are you still there?"

In contrast, if the customer, citizen, or taxpayer calls the State of Enlightenment to add an address, they will have a completely different experience. The State of Enlightenment has a far more robust data base that was designed from a far more robust set of business rules that were explicitly modeled in a far more robust logical data model.

An Operational Scenario with *State of Enlightenment*

Citizen call to Call Center – State of Enlightenment

Citizen: "I need to add another address."

Call Center: "OK. Let me help you with that. First what type of address is this? Is it for business, is it another residence, is it temporary, or none of those?"

Citizen: "Actually, it's a temporary vacation address."

Call Center: "What dates do you expect this address will be effective?"

Citizen: "My lease begins on April 8, 2011. I can receive mail from you at that address starting then. My lease is up on July 30th, 2011, so make the end date July 20th, 2011."

Call Center: "OK, that address has been added to our records. We'll use it as a mailing address for our purposes during that period?

Citizen: "Thank you!"

These conversations and the related data models are intended to present the reality that different enterprises, governments, organizations, agencies, or departments will have different business rules imbedded into their systems. If the intention is to bring these multiple organizations into a cloud environment, or any type of *shared resources environment*, where they share common systems or resources, then they must also share common business rules. Or, the business rules implemented must accommodate the complete portfolio of scenarios these organizations encounter.

This situation is also evident in a Software as a Service (SaaS) scenario. The provider will have imbedded business rules in the portfolio of applications they offer. There *may be* some flexibility in modifying those business rules. This is the case when business rules are not hard coded into the application(s). Instead, business rules are stored in tables which can be modified by the customer of the provider. Nevertheless, a cloud provider or service provider will have a data base schema which describes the physical organization of the provider's data base. If the provider only has the physical schema, then that

schema will have to suffice as input from the provider in evaluating data harmonization and "fit." Expect data base schemas to be more complicated to interpret for business rules than logical data models. This is due to the fact that a physical data base schema will most likely not compare well with the customer's, *any customer's,* logical data model.
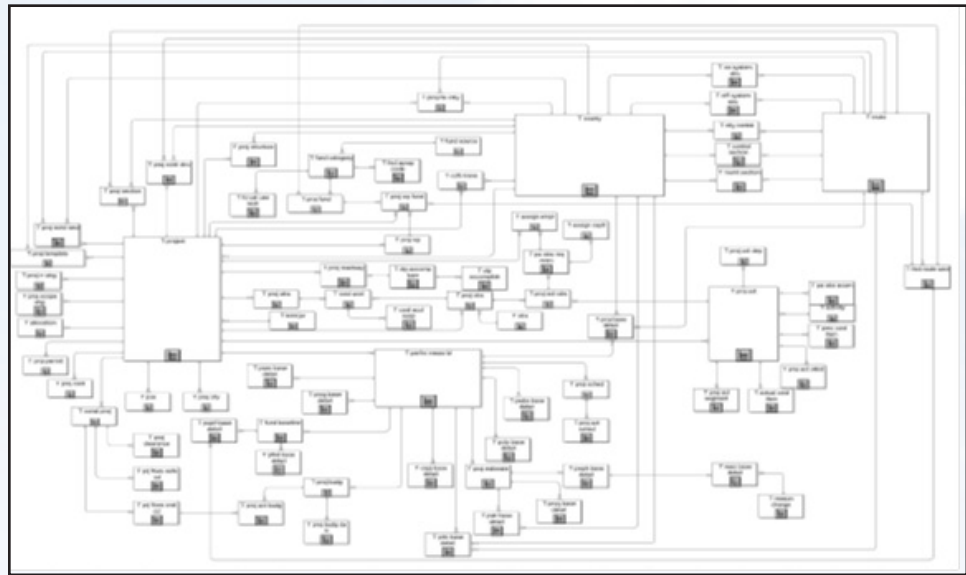
Most application service providers have created data base schemas to accommodate performance by reducing the number of inputs and outputs (I/Os) required to: retrieve and present data, add data, delete data, and to update data. Some application service providers *do not* have a logical data model partly because they don't have skilled logical data modelers on staff with the expertise to create one. However, they may have on staff people skilled at creating data bases that may or may not be derived from logical data models. Possibly the best advice is, "buyer beware." Know what you are buying. Once the data is migrated into a provider's data base, it may not be so easy to get it back out again if the customer decides to "switch" providers, or bring the service back in house. Up-to-date and accurate logical data models are necessary references when considering first time service contracts as well as subsequent consideration of moving to a new service provider. Further, when the provider's database does not perfectly map to the customer logical data model, the gaps must be documented and kept up to date. Those "compromises" would potentially be "backed out" when the customer either moves to a new provider with a better "fit", or the customer moves back in house with an application that may have a more "perfect fit."

Does this mean cloud computing is a bad choice? No, but it does create *bounding* on what cloud services are truly viable, particularly for government or any other enterprise that has established business rules, and the anticipation of long term enterprise viability of those rules.

Multiple images of an application will be required if there are multiple customers with differing business rules that can not be harmonized, or where partnering enterprises are unwilling to adopt another's business rules. Or, the partnering enterprises can not agree on common business rules. This will of course be more expensive since the service provider is not able to achieve some economies of scale that are used to determine pricing. If government is trying to achieve the savings associated with economies of scale that a cloud provider can offer, then there will need to be a rationalization / optimization of business rules across the cohort of organizations that are intending to share a resource, assuming that resource is not simply storage, or infrastructure.
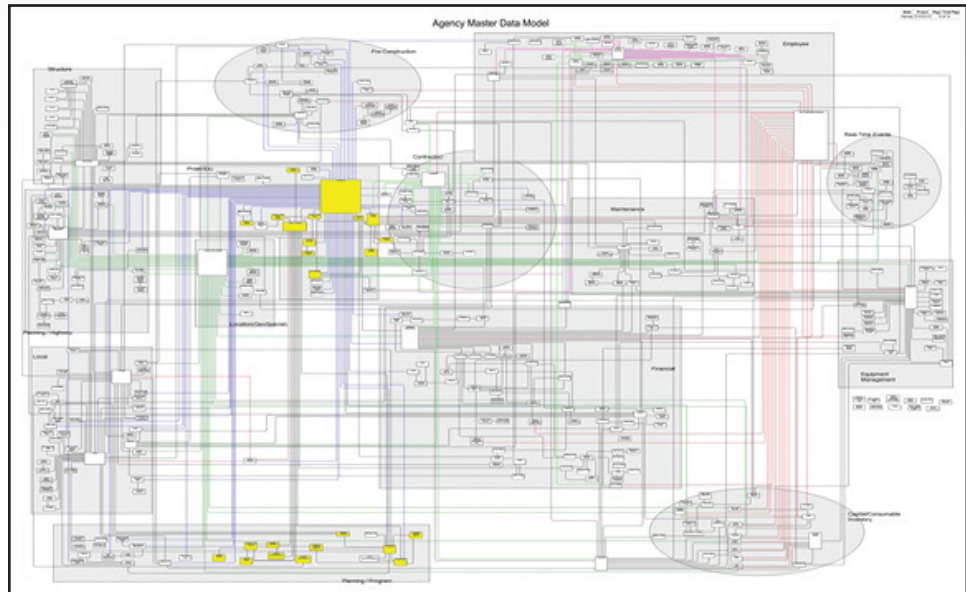
As depicted in the example above. This is not a straight forward exercise. The data models presented in the example are relatively simple with many business rules left out. In reality, business information models, or logical data models, are far more complex. An agency logical data model may contain hundreds of entities, and thousands of attributes and relationships.

This is a more realistic logical data model from one state's department of transportation. Simply note the number of relationships (lines) and the number entities (boxes) in this model.

Example 1 of an actual Logical Data Model

Here is another example that is even more complex. The shaded circles are intended to depict major subject areas.

Example 2 of an actual Logical Data Model

If state government, or an agency, or other enterprise *does not* have a logical data model, then one will have to be created, or adopted. The cloud provider, or software application provider should have something in hand to describe the rules inherent to *their* system. Again, oftentimes, what is presented by a

provider is a *physical data model* that presents the data base structures. The customer will need to understand the inherent business rules that are imbedded in the provider's system since they will be essentially adopting those rules if they engage the provider.

If state government, or an agency, or other enterprise *does* have a logical data model – similar to Example 1 or Example 2 – that is truly a logical data model - they will use that model to evaluate IT service providers, or cloud providers, in order to judge the level of "fit". That "fit" will have to be judged as "close enough" before the customer enterprise engages the service provider. Further, state government, or an agency, or other enterprise will need to understand the adaptability of that provider's system if the customer's business rules change.

If state government, or an agency, or other enterprise *does* have a logical data model, and is in the process of joining a *community of interest* within state government to design, build and maintain a shared IT service, then the community of interest will use that data model, and the data models from other partners, to essentially *negotiate* a community of interest data model which will be used to:

- rationalize business rules – potentially making business process improvements, or business process enhancements
- create a new common understanding / adoption of definitions, relationships and attributes – potentially for the first time creating a common understanding of the *single, accurate version of the truth* regarding how government should operate
- evaluate potential service providers outside of government for "fit"
- evaluate joining other larger communities of interest

As with any technology employment strategy, cloud computing initiatives must look at the underlying business of government that can benefit from its use. Cloud deployment models and service models must be evaluated for appropriateness for the "enterprise in focus" or "business area in focus." Some cloud deployment models are simply not appropriate for certain government business processes, or government information.

Finally, there must be a business goal driving adoption of cloud computing. The business goal will have associated economic and social benefits that justify the cost of human and financial capital that will be expended to plan, evaluate and potentially deliver the cloud service, or the shared resource, including the process of *data harmonization*.

Calls to Action for the State CIO

- Not surprisingly, anticipate high interest in cloud computing. Be proactive in establishing an enterprise cloud strategy that takes full account of the data management issues.

- Understand that cloud computing does not preclude the need for enterprise architecture – particularly process and data architecture. These architectures will be necessary disciplines for understanding how the shared service fits into the overall state government enterprise architecture. Possibly, the importance of these disciplines increases as state government evaluates shared services including cloud computing services.

- Anticipate significant data harmonization issues when creating communities of interest and evaluating shared service models such as cloud computing.

- Begin planning for creating appropriate data governance as agencies are coming together to form communities of interest. Be the advocate and promoter of these communities. Emphasize the need for data reconciliation. This activity should be guided by an experienced data architect.

- Engage data management staff in evaluating shared service offerings to evaluate the impact of employing such a service.

- Examine the logical data model and the physical data model of the service provider. If they don't have either of these, keep looking for a provider that does. It is absolutely essential that state government understand the "fit" of a particular IT service. Fit analysis includes reconciliation of process and data.

- Employ the DAMA Body of Knowledge as an important reference in exploring data management issues.

- Anticipate that you will still need qualified enterprise architects with skills in logical data modeling, master data management, and process modeling if you engage service providers or cloud providers.

- Never turn ownership of business rules, data models, or data, over to an outside service provider. The consequences of an outside entity having legal ownership of data, and business rules can be serious – particularly in public safety, health, environmental, and security lines of business.

- Stay tuned to NASCIO and the community of state CIOs for more on Cloud Computing.

Contributors

David Behen, Chief Information Officer, State of Michigan

Patricia Cummens, Account Executive, ESRI Corporation

Albert Decker, First Data – Technical Manager, Office of Temporary and Disability Assistance, State of New York

Paul Warren Douglas, Enterprise Architect, State of Washington

Scot Ellsworth, Chief Enterprise Architect and Director of the Office of Enterprise Architecture, State of Michigan

James Earl, J.D., Executive Director, Technological Crime Advisory Board, State of Nevada

Lauren Farese, Public Sector Senior Director, Oracle Corporation

Mike Fenton, Director of Enterprise Architecture, State of North Carolina

Jeremy Foreman, Public Sector Enterprise Architecture Program Director, Oracle

Sam L. Hearn, Jr., Graphic Designer, AMR Management Services

Sam Holcman, President of the Zachman Institute for Framework Advancement (ZIFA), the Chairman of the Pinnacle Business Group, Inc., and the Managing Director of the Enterprise Architecture Center of Excellence (EACOE)

Christopher G. Ipsen, CISSP-ISSAP, CISM, Chief Information Security Officer, State of Nevada, Department of Information Technology

Dr. Dave McClure, Associate Administrator for Citizen Services and Innovative Technology (OCSIT), Government Services Administration Office of Citizen Services and Innovative Technologies

Bob McDonough, Chief Cloud Architect, State of Michigan

Darlene Meskell, Director, Global Government Innovation Networks, Government Services Administration Office of Citizen Services and Innovative Technologies

Maury Mitchell, Director State of Alabama Criminal Justice Information Center

Andris Ozols, Chief Policy Advisor, State of Michigan, Department Technology, Management and Budget

Daniel J. Paolini, CBIP, CDMP, CDPS, Deputy Chief Technology Officer, Data Management, Information Architecture and Administrative Services Affinity Group, State of New Jersey Office of Information Technology

Carolyn Parnell, Chief Information Officer, State of Minnesota

Doug Robinson, Executive Director, NASCIO

Bill Roth, Chief Information Technology Architect, State of Kansas

Gary Therrien, Manager, Advantech Group LLC

Glenn J. Thomas, PMP, CDMP, CPM, Director, IT Governance, Commonwealth of Kentucky, Office of Technology

Shawn K. Vaughn, Membership & Communications Coordinator, NASCIO

Tom Walters, Enterprise Business Architect, Commonwealth of Kentucky, Office of Technology, Office of Enterprise Technology, Division of IT Governance

Appendix A – Cloud Computing References

The Australian Government Cloud Computing Strategic Direction Paper
http://www.finance.gov.au/e-government/strategy-and-governance/docs/final_cloud_computing_strategy_version_1.pdf

The Department of Finance and Deregulation, through the Australian Government Information Management Office, has consulted with government agencies, industry and the public to develop an Australian Government Computing Strategic Direction paper to explore the opportunities and impacts of cloud computing.

Cloud Computing Use Cases Group (Google group)
http://groups.google.com/group/cloud-computing-use-cases

This group is devoted to defining common use cases for cloud computing.

Computer Crime & Intellectual Property Section, United States Department of Justice
http://www.justice.gov/criminal/cybercrime/ssmanual/

The purpose of this publication is to provide Federal law enforcement agents and prosecutors with systematic guidance that can help them understand the legal issues that arise when they seek electronic evidence in criminal investigations. Chapter 3 of this publication presents the Stored Communications Act (SCA). The significance of the SCA is that it imposes restrictions on voluntary disclosures by providers of services to the public, but it also includes exceptions to those restrictions. The SCA governs how investigators can obtain stored account records and contents from network service providers, including Internet service providers ("ISPs"), telephone companies, and cell phone service providers.

Cloud Customers' Bill of Rights
Information Law Group LLP – www.infolawgroup.com

The InfoLawGroup has issued a "Cloud Customers' Bill of Rights" to serve as the foundation of a cloud relationship, allow for more transparency and enable a better understanding of potential legal risks associated with the cloud.

Detailed description of the Cloud Customers' Bill of Rights:
http://www.infolawgroup.com/2010/10/articles/cloud-computing-1/cloud-computing-customers-bill-of-rights/

The Cloud Security Alliance (CSA)
https://cloudsecurityalliance.org/about/

The Cloud Security Alliance (CSA) is a not-for-profit organization with a mission to promote the use of best practices for providing security assurance within Cloud Computing, and to provide education on the uses of Cloud Computing to help secure all other forms of computing. The Cloud Security Alliance is led by a broad coalition of industry practitioners, corporations, associations and other key stakeholders.

Federal Cloud Computing Strategy
http://www.cio.gov/documents/Federal-Cloud-Computing-Strategy.pdf

This Federal Cloud Computing Strategy is designed to:
Articulate the benefits, considerations, and trade-offs of cloud computing
Provide a decision framework and case examples to support agencies in migrating towards cloud computing
Highlight cloud computing implementation resources
Identify Federal Government activities and roles and responsibilities for catalyzing cloud adoption

The Jericho Forum (The Open Group)
http://www.opengroup.org/jericho/

Jericho Forum is the leading international IT security thought-leadership association dedicated to advancing secure business in a global open-network environment. Members include top IT security officers from multi-national Fortune 500s & entrepreneurial user companies, major security vendors, government, & academics. Working together, members drive approaches and standards for a secure, collaborative online business world.

National Institute of Standards and Technology Cloud Computing Program
http://www.nist.gov/itl/cloud/index.cfm

The long term goal of this program is to provide thought leadership and guidance around the cloud computing paradigm to catalyze its use within industry and government. NIST aims to shorten the adoption cycle, which will enable near-term cost savings and increased ability to quickly create and deploy enterprise applications. NIST aims to foster cloud computing systems and practices that support interoperability, portability, and security requirements that are appropriate and achievable for important usage scenarios.

The Open Cloud Manifesto
http://www.opencloudmanifesto.org/

Dedicated to the belief that the cloud should be open. This effort intends to initiate a conversation that will bring together the emerging cloud computing community (both cloud users and cloud providers) around a core set of principles. We believe that these core principles are rooted in the belief that cloud computing should be as open as all other IT technologies.

Appendix B – Data Management References

NASCIO on Data Governance – www.nascio.org/publications

Data Governance - Managing Information As An Enterprise Asset:
Part I - An Introduction
April 2008

Data governance entails a universe of concepts, principles, and tools intended to enable appropriate management and use of the state's investment in information. Part I on data governance presents an introduction that describes the basic concepts. Governance, and particularly data governance, is an evolutionary process. It begins with an understanding of the current investment and then manages that investment toward greater value for the state.

Data Governance Part II: Maturity Models – A Path to Progress
March 2009

Data governance maturity models provide a foundational reference for understanding data governance and for understanding the journey that must be anticipated and planned for achieving effective governance of data, information and knowledge assets. This report continues to build on the concepts presented in Data Governance Part I. It presents a portfolio of data governance maturity models.

Data Governance Part III: Frameworks – Structure for Organizing Complexity
May 2009

This issue brief presents the concept of frameworks that describes what constitutes a data governance program, with a focus on frameworks from the Data Management Association (DAMA), the Data Governance Institute (DGI), and IBM. Use of frameworks can assist state government in planning and executing on an effective data governance initiative. They assist in achieving completeness in a program. In any subject or discipline frameworks and maturity models assist in describing the scope – both breadth and depth – of an initiative. This holds true as well for data, information and knowledge

DAMA Data Management Body of Knowledge – DMBOK
http://www.dama.org/i4a/pages/index.cfm?pageid=1

DAMA International is a non-profit, vendor-independent, global association of technical and business professionals dedicated to advancing the concepts and practices of information and data management. ISBN 9780977140084. See "recommended reading" at the end of each section.

The Data Administration Newsletter (TDAN)
www.tdan.com

TDAN.com is an industry leading publication for people interested in learning about data administration and data management disciplines & best practices. Each monthly issue addresses the most challenging issues of the day.

---

[1] *2010 State CIO Survey – Perspectives and trends from state government IT leaders is available at www.nascio.org/publications and www.techamerica.org/2010-state-cio-survey.*

[2] *Gartner Highlights Key Predictions for IT Organizations and Users in 2010 and Beyond.* Retrieved on August 18, 2011, from www.gartner.com/it/page.jsp?id=1278413.

[3] O'Neil, B., "Introducing Data Unification and Harmonization: Critical Success Factor For Data Warehousing," *The Data Administration Newsletter (TDAN),* July 1, 2006. Retrieved on September 3, 2011, from http://www.tdan.com/view-articles/5005.

[4] Benson, P., *Data Portability – The Antidote to Data "Lock-In,"* The Data Administration Newsletter – TDAN.com. Retrieved on September 13, 2011, from http://www.tdan.com/view-articles/11187.

[5] Kruschwitz, N., Shockley, R, *MIT Sloan Management Review,* "First Look: The Second Annual New Intelligent Enterprise Survey", summer 2011. Reprint Number 52413.

[6] Transforming Government Through Change Management: The Role of the State CIO, NASCIO, April, 2007. Available at www.nascio.org/publications.

[7] The DAMA Guide to the Data Management Body of Knowledge, 1st Edition, 2009. ISBN 9780977140084. p. 171