



Provider Data Architecture Blueprint (Part 1 - Data Aggregation)

Enterprise Architecture
April 29, 2015

Acknowledgements

Enterprise Architecture Team	Architects
Information Architecture	Sumit Bhindwale, Kevin Gareau, Sunil Pimenta , Sanjay Seth, Gene Vilain
Application Architecture	Vaibhav Bhomkar, Swarada Deshpande, Harika Durga, Jeff Eng, Alex Finogenov, Sharad Gupta, Ian Lopez, Michael Reynolds, Satya Tirupachur
Technology Architecture	Tim Madams, Jay Shen, Dean Watson
Security Architecture	Wayne Harriman, Tom Lunzer, Richard Marsh, Joan Rose, John Walker, Mike Wolfe
EA Leadership	Steve Kong, Alvin Loving, Alan Pearson, Marc Wojtowicz, Karen Xie

Stakeholders	Contributors
IT Infrastructure Services	Fernando Acosta, Govinda Chinthada, Stefan Dalthorp, Silvester Divas, Bryan Green, Ravi Kalaga, Shelby Kocherry, Bhaskar Lakshmikanthan, Bags Perumal, Michelle Shih, Chandhra Vadlamudi, Mike Windbigler
Application Services	Sergey Chudnovsky, Chendra Grandhi, Supal Patel
IT Program Delivery	Sadhana Chandramouli, Angela McArthur
IT Shared Services	Alicia Domecus
Data Liquidity Team	Gautam Karmakar, Jo Monterola
Other	Enterprise Data Governance Steering Committee

Table of Contents

- 1. Executive Summary**
- 2. Concept Definition and Model Framework**
- 3. Environmental Analysis**
- 4. Current State Architecture Models**
- 5. Future State Architecture Models**
- 6. Architecture Analysis**
 - **Gap Analysis**
 - **Impact Analysis**
 - **Technology Solutions Analysis**
 - **Risk Analysis**
- 7. Roadmap**
- 8. Recommendation**
- 9. Appendices**

2015 IT Strategic Initiatives

#	Initiative Name
1	Infrastructure Stabilization & Sustainability
2	Advanced Infrastructure Capabilities
3	Best-in-Class Infrastructure Capabilities
4	Application Delivery Excellence
5	Service Business Management
6	Enterprise Demand and Sourcing Management
7	Compliance & Risk Management
8	Technology Architecture & Simplification
9	Middleware Integration & Interoperability
10	Training, Talent, and Communications
11	Enterprise Architecture Enablement
12	Data Liquidity
13	Portfolio Management
14	Shield Advance
15	Application Lifecycle Manage
16	Application Retirement
17	Risk-based IT Security Capabilities

Table of Contents

- 1. Executive Summary**
2. Concept Definition and Model Framework
3. Environmental Analysis
4. Current State Architecture Models
5. Future State Architecture Models
6. Architecture Analysis
 - Gap Analysis
 - Impact Analysis
 - Technology Solutions Analysis
 - Risk Analysis
7. Roadmap
8. Recommendation
9. Appendices

Executive Summary

Ask	Develop the enterprise architecture blueprint for provider data which will guide the implementation of provider data initiatives including Provider Book of Record (BoR) and Provider data quality. (Part 1 of 2: Data Aggregation; Part 2 of 2: Data Consumption)	
Current State Highlights	<ul style="list-style-type: none"> • Provider BoR – Provider data is fragmented with no single authoritative source for core provider information • Provider Index (MDM) – Provider “golden record” does not exist, with duplicate provider records within & across provider systems • Provider Data Quality – Inaccurate provider information in directories and portals have resulted in failed DMHC audits • Data Integration – lack of a data replication tool for change data capture and real-time data delivery • Technology Solutions: <ul style="list-style-type: none"> ◦ Master Data Mgmt – IBM Infosphere MDM SE; Data Replication - gap; Data Balancing – Infogix Assure; Job Orchestration (Nettezza) – ABC framework; Data Archiving – Optim Archive; Analytics – Business Objects; Data Masking – Optim Data Privacy; Database Activity Monitoring – Imperva; PDQ (bulk updates) – gap; Data Profiling/Cleansing – Informatica IDQ; BoR Data Platform – Nettezza; BoR Disaster Recovery – Nettezza PTS; File Transfer – SFTP (ext.), FTP(int.); Code Version Control – Unfuddle; Scheduling – Tidal; Database Backup – NetBackup; Staging Database – Oracle; ETL/ELT – Informatica, Nettezza ELT scripts 	
High-Level B.I.A.T.S Requirements	<p>Business – Implement a single source of truth for Provider data (Provider BoR) and Provider Index “golden record”</p> <p>Information – Provider BoR consists of 8 data topics including provider demographics, credentialing, contracts, ratings, networks, services, review and relations; Provider Index creates “golden record” for practitioner, provider organization, business entity and corresponding locations</p>	<p>Application – Implement a suite of data management services including master data management, data replication, data quality and data integration (ELT)</p> <p>Technology – Leverage and build infrastructure for hosting Provider BoR, Provider Index and operational data with HA/DR capabilities</p> <p>Security – Comply with BSC security policies and implement security controls</p>
Future State and Roadmap Recommendations	<p>Future State:</p> <ul style="list-style-type: none"> • Provider BoR – Provider BoR deployed on the Nettezza platform based on Provider Enterprise Information Model (EIM) • Provider Index (MDM) – Provider Index deployed with “golden record” for 6 provider entities using Infosphere MDM • Provider Data Quality – Provider Data Quality (PDQ) custom tool deployed for bulk data updates to PIMS • Data Integration – GoldenGate data replication tool deployed for enabling real-time operational reporting and data services • Technology Solutions: <ul style="list-style-type: none"> ◦ Master Data Mgmt – IBM Infosphere MDM SE; Data Replication – Oracle GoldenGate (new); Data Balancing – Infogix Assure; Job Orchestration (Nettezza) – Custom built (new); Data Archiving – Optim Archive; Analytics – Business Objects, SAS, Tableau; Data Masking – Optim Data Privacy; Database Activity Monitoring – Imperva; PDQ (bulk updates) – Custom built (new); Data Profiling/Cleansing – Informatica IDQ; BoR Data Platform – Nettezza; BoR Disaster Recovery – Nettezza PTS; File Transfer – IBM Sterling File Gateway (external), MQ MFT (internal); Code Version Control – Apache Subversion; Scheduling – Tidal; Database Backup – NetBackup; Staging Database – Oracle; ETL/ELT – Informatica, Nettezza ELT scripts <p>Roadmap - Key Next Steps:</p> <ol style="list-style-type: none"> 1. Implement a provider data quality program that includes governance, auditing and monitoring – Owner: Provider Data Owner & Stewards 2. Develop and implement the data integration ELT framework (guidelines, best practices) – Owner: Provider BoR project team, Data Liquidity team 3. Implement the rules for mastering the 6 entities in Provider Index and generating golden records – Owner: Provider BoR project team, Data Liquidity team 4. Complete the Provider BoR implementation, including re-mediation of identified gaps – Owner: Provider BoR project team, Data Liquidity team 5. Purchase and deploy recommended data replication tool (GoldenGate) – Owner: Infrastructure Services 6. Implement the Provider Data Quality workflow tool, including Enclarity integration and address standardization – Owner: PDQ project team 7. Implement other data mgmt services: job orchestration, data archiving, activity monitoring, data balancing, DR – Owner: Infra. Services, Data Liquidity team 8. Expand Provider data architecture blueprint to include data consumption (Part 2 of 2) – Owner: Enterprise Architecture 9. Initiate Rosetta retirement planning – Owner: Actuarial & Analytics 	

Table of Contents

1. Executive Summary
- 2. Concept Definition and Model Framework**
3. Environmental Analysis
4. Current State Architecture Models
5. Future State Architecture Models
6. Architecture Analysis
 - Gap Analysis
 - Impact Analysis
 - Technology Solutions Analysis
 - Risk Analysis
7. Roadmap
8. Recommendation
9. Appendices

Provider Data: Key Definitions

Provider Book of Record (BOR)	Provider BOR is the authoritative data source for Provider data that is derived from integrating multiple data sources (internal and external). The data model is based on the Enterprise Information Model for Provider subject area and technology components are aligned with the Data Liquidity blueprint.
Data Liquidity	Data Liquidity is concerned with implementing data governance, supporting capabilities and the foundational technology components necessary for providing the right data to the right person at the right time.
Enterprise Information Model (EIM)	EIM is the model that provides an integrated view of the data produced and consumed across the enterprise. It incorporates an appropriate industry perspective. It represents a single integrated definition of data, unbiased of any system or application and is aligned with the Business Glossary.
Business Glossary	Business Glossary is the definitive dictionary of business terms used across an enterprise. The definitions are designed to engender a common understanding of what is meant by the term for all employees and key business partners regardless of business function.
Master Data	Master Data is the consistent , unique and uniform set of identifiers and extended attributes that describes the core entities of the enterprise, including providers. (Source: Gartner)
Provider Index (MDM)	MDM is a technology-enabled discipline in which business and IT work together to ensure the uniformity, accuracy, stewardship, governance, semantic consistency and accountability of the enterprise's master data assets, in this case - Provider. (Source: Gartner)
Source System of Record	Source System of Record for a business data element is the system or application where the data is entered, or originates for the first time.
Data Quality	Data Quality is the state of completeness, validity, consistency, integrity, timeliness and accuracy that makes data appropriate for a specific use. (Source: TDWI)
Provider Golden Record	Provider Golden Record is the provider record representing the best set of data attributes across all provider data sources as determined by the Provider Index (MDM) system.

Provider Data: Conceptual Model (Level 1)



Table of Contents

1. Executive Summary
2. Concept Definition and Model Framework
- 3. Environmental Analysis**
4. Current State Architecture Models
5. Future State Architecture Models
6. Architecture Analysis
 - Gap Analysis
 - Impact Analysis
 - Technology Solutions Analysis
 - Risk Analysis
7. Roadmap
8. Recommendation
9. Appendices

External Research: Healthcare Industry Trends

Healthcare Consumerism (B2C model)

- Consumers choose own products from range of preselected offerings on public/private exchanges
- Rise in premium sharing, deductibles, co-insurance places decision-making on consumers
- Move toward a consumer-centric marketplace, with more options for consumers

Transparency and Self-Service

- Consumers expect provider quality scores (institution/physician) for procedures
- Physician finder tools expected to list office locations, operation hours, training etc
- Consumers expect transparency around service costs, provider's experience in delivering procedure, including outcomes

Narrow Provider Networks

- Networks being narrowed by limiting providers, driving utilization to favored providers or contracts with integrated providers (ACO's)
- Recognition that providers are the supply chain for payers and networks need active management

Partnerships and Acquisitions

- Digital & health partnerships outside traditional marketplace offer improved device integration, care management
- Increased adoption of vertically integrated care and shared risk models

External Research: Data Management Trends

Data Access	<ul style="list-style-type: none">• Shift to real-time information that is accessible anywhere, anytime on multiple form factors including mobile devices
Data Integration	<ul style="list-style-type: none">• Intelligent combination of data from warehouses, marts, and the cloud using multiple modes including batch, messaging, replication and data virtualization
Business Analytics	<ul style="list-style-type: none">• Instant, real-time analytics with in-memory databases• Predictive and prescriptive analytics using big data sources (internal, external)
Information Governance and MDM	<ul style="list-style-type: none">• Information governance and MDM initiative address all enterprise data (including external over which there is less control)
Database Scalability and Operations	<ul style="list-style-type: none">• Big data, cloud-scale with exabytes of data, with self-healing infrastructure, databases, systems and applications

Source : Gartner's 2015 Planning Guide for Data Management

External/Internal Research: Provider Data Quality Challenges

Based on research from Enclarity, typically, 30-40% of a payer's provider records contain errors or missing information

Statistics from Enclarity's database (6.5 mill. provider profiles)

- 2 - 2.5% of provider demographic data changes every month (50% degradation over 18 months if not corrected), resulting in:

In a Typical Provider File	Average
Duplicate provider records	28%
Providers with inaccurate or missing NPI's	12%
Wrong or missing phone numbers	15%
Wrong or missing addresses	12%
Providers with sanctions	1.3%
Deceased providers	0.2%

Statistics from assessment of BSC Provider File using Enclarity (2014)

- 10% of provider data changes every month

In BSC Provider File	%
Duplicate facility providers with same address/TIN	27.67%
Duplicate individual provider with same address/TIN	11.58%
Individual providers with inaccurate or missing NPI's	8.49%
Facility providers with inaccurate or missing NPI's	2.87%
Wrong or missing practice phone numbers	16.9%
Wrong or missing practice addresses	23.6%
Individual providers with sanctions	0.71%
Deceased individual providers	0.07%

- BSC has ~ 60,000 network providers
- In 2014, ~304,000 provider records updated
- End 2014, ~31,000 in pending inventory

Sources: (a) Enclarity Whitepaper: A_Business_Case_for_Fixing_Provider_Data_Issues (b) IDV_CIC_Report.pdf, FAC_CIC_Report.pdf, January 2014 (BSC), (c) BSC Provider Information & Enrollment (PIE) team

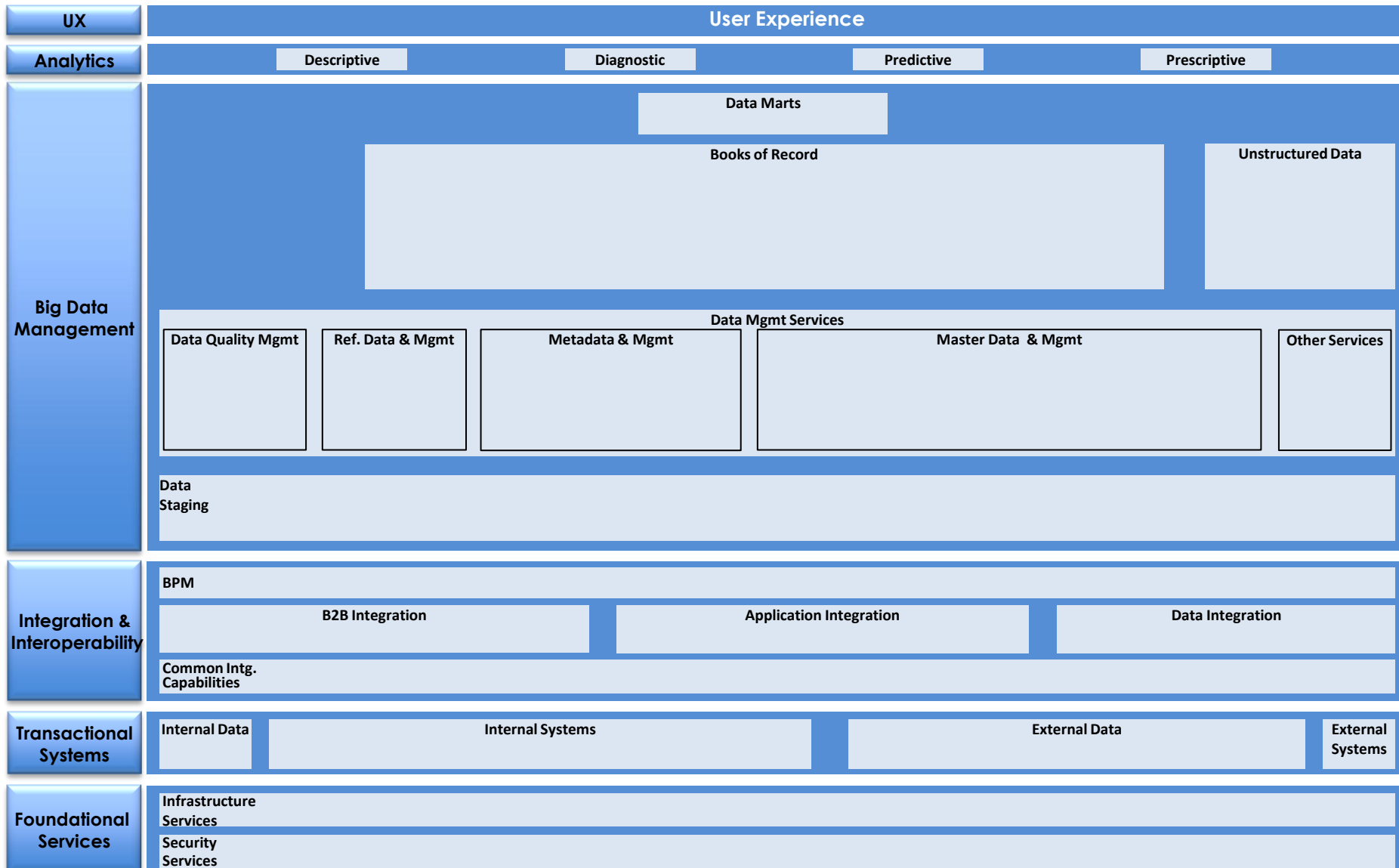
Provider Data: High-Level Requirements

Business Requirements	<ul style="list-style-type: none">• Implement a single authoritative source for Provider data - Provider Book of Record.• Implement Provider Index master data management tool for probabilistic matching, merging and de-duplication of provider records from multiple sources and generation of a unique “golden record” for provider entities.• Implement a workflow tool for ingesting provider data updates from provider groups, regulatory agencies, external data vendors, portals and applying bulk updates to PIMS after manual review of changes or automated based on business rules.
Information Requirements	<ul style="list-style-type: none">• The Provider Book of Record will be based on the Provider Enterprise Information Model and consist of data topics including provider demographics, credentialing, contracts, ratings, networks, provider services, provider review and provider relations• The Provider Index system will match/merge provider records from internal and external sources and create golden records for 6 provider entities, including practitioner, provider organization, business entity and their corresponding service locations• Leverage data quality tools for provider data quality measurement and data cleansing using address standardization modules.
Application Requirements	<ul style="list-style-type: none">• Implement a tool based match, merge and de-dupe functionality configured using business rules and UI for managing linkages and duplicates, while keeping customization to a minimum.• Implement a custom workflow tool for applying bulk update of provider information to PIMS in either real-time or batch mode, after either a manual review or automated based on configurable business rules.
Technology Requirements	<ul style="list-style-type: none">• Deploy a real-time log based data replication tool for logical data replication from PIMS to PROPS and other databases.• Leverage secure file gateways for exchange of files between BSC and external vendors/partners.• Ensure technology versions must be either current (N) or one behind current (N-1) release for effective vendor support.• Maintain appropriate backup/recovery and disaster recovery (DR) capabilities for the data and application platforms to ensure business continuity.
Security Requirements	<ul style="list-style-type: none">• Implement BSC security requirements for user authentication, password policy, role based access control (authorization), separation of duties, application session and timeouts, encryption, logging and monitoring of security related events and ensure security risk assessment is completed.• PHI data will be masked in the non-production environments, code development needs to follow scripting security requirements, change management and use of test data, access to provider data services will be protected by DataPower.

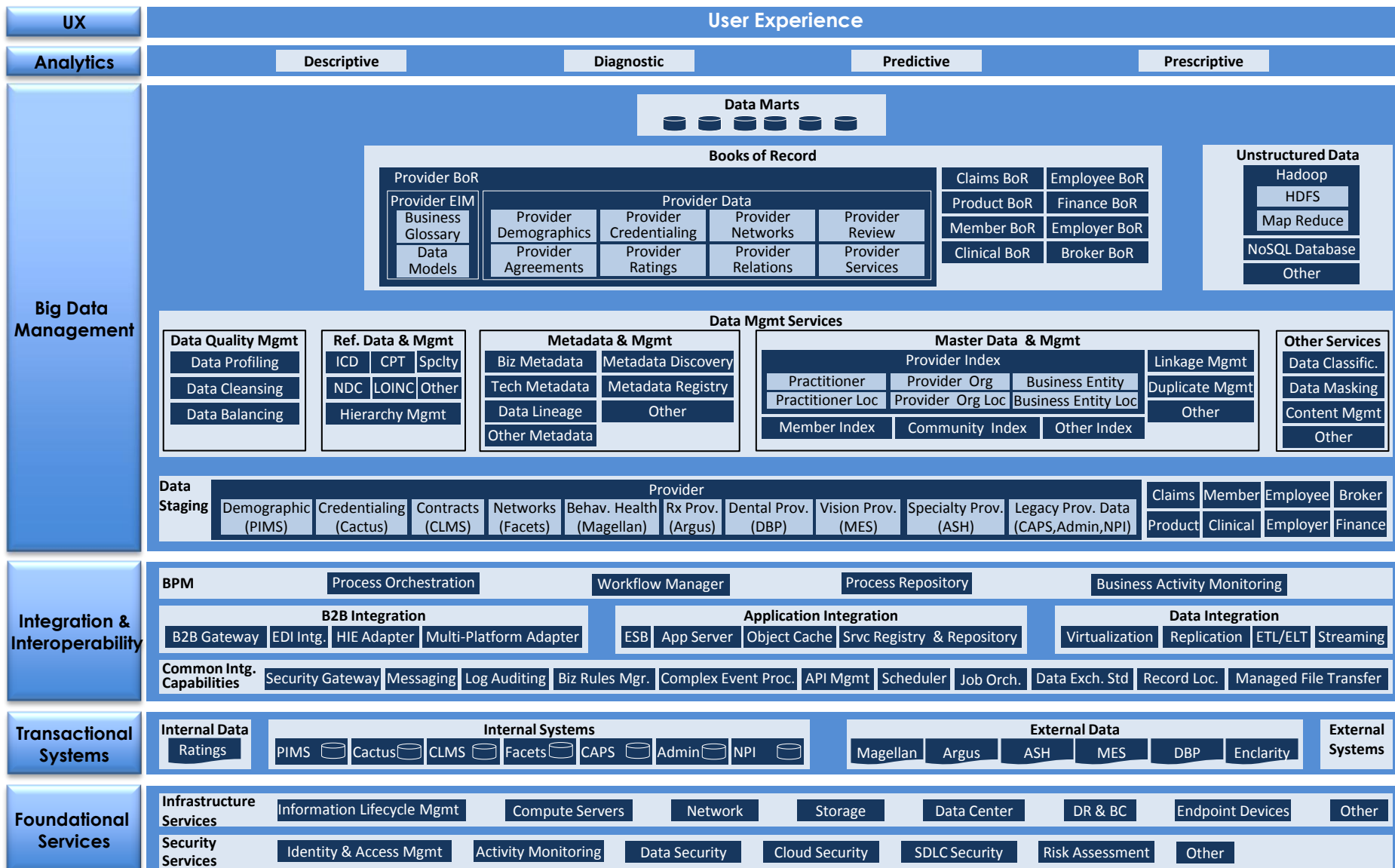
Table of Contents

1. Executive Summary
2. Concept Definition and Model Framework
3. Environmental Analysis
- 4. Current State Architecture Models**
- 5. Future State Architecture Models**
6. Architecture Analysis
 - Gap Analysis
 - Impact Analysis
 - Technology Solutions Analysis
 - Risk Analysis
7. Roadmap
8. Recommendation
9. Appendices

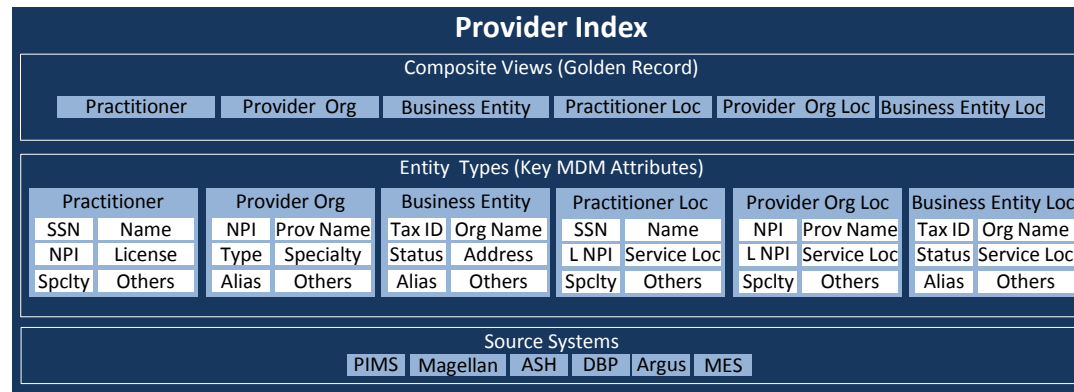
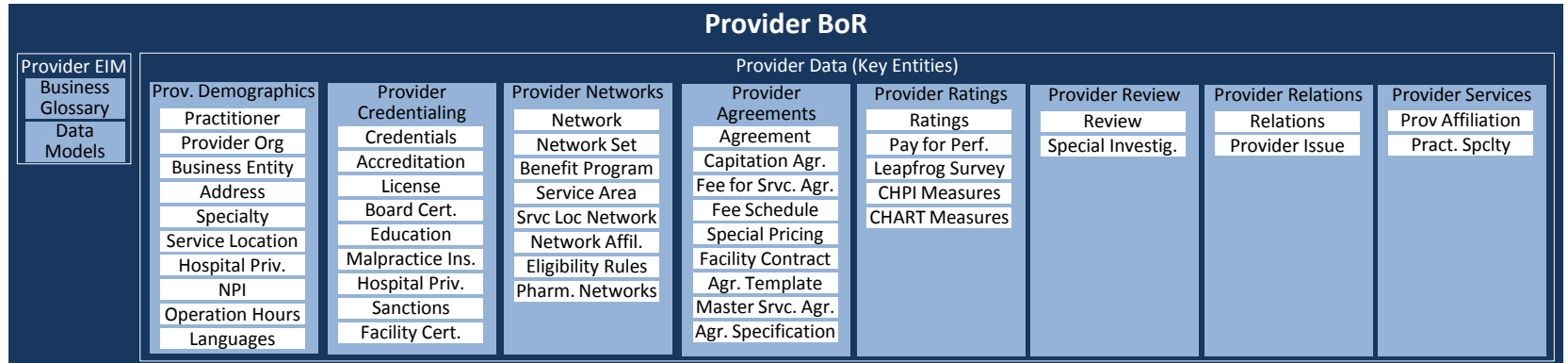
Provider Data: Conceptual Model (Level 2)



Provider Data: Conceptual Model (Level 3)



Provider Data: Conceptual Model (Level 4)



Provider Data: Conceptual Model (Level 5)

#	DATA TOPIC	ENTITY	ATTRIBUTE
1	Provider Agreements	Ancillary-Alternate Care Service Providers	Description Of Pricing
2		Ancillary-Alternate Care Service Providers	Description Of Services
3		Ancillary-Alternate Care Service Providers	UC Care Tier
4		Benefit Program	benefit program type
5		Benefit Program	effective date
6		Benefit Program	termination date
7		Capitation Agreement	Aggregate Stop Loss Amount
8		Capitation Agreement	Capitation Base Amount
9		Capitation Agreement	Capitation Payment Frequency
10		Capitation Agreement	Capitation Rate
11		Capitation Agreement	Capitation Rate Adjustment Factor
12		Capitation Agreement	Capitation Rate Adjustment Reason
13		Capitation Agreement	Capitation Rate Basis
14		Capitation Agreement	IPA Claim Submission Source
15		Capitation Agreement	IPA Specialty Pharmacy
16		Capitation Agreement	Out Of Network Referral Assessment Program
17		Capitation Agreement	Percentage of Revenue
18		Capitation Agreement	Pprevnar and new Pediatric vaccines approved after 01/01/01 Contracted Rate
19		Capitation Agreement	Pprevnar and new Pediatric vaccines approved after 01/01/01 Risk Allocation
20		Capitation Agreement	Richmam Hill List of Injectable Drug Exceptions Contracted Rate
21		Capitation Agreement	Richmam Hill List of Injectable Drug Exceptions Risk Allocation
22		Capitation Agreement	Specific Stop Loss Amount
23		Capitation Agreement	Withhold Percentage
24		Commercial-Medicare Hospital Transplant	Number of Facilities
25		Commercial-Medicare Hospital Transplant	Transplant Type Applied
26		Exception	type of network
27		Fee For Service Agreement	Allowed Amount
28		Fee For Service Agreement	Arbitration Months
29		Fee For Service Agreement	Deadline For Claims Submission
30		Fee For Service Agreement	Final Appeal Days
31		Fee For Service Agreement	Initial Appeal Days
32		Fee For Service Agreement	Negotiation Contact Name
33		Fee For Service Agreement	Negotiation Contact Phone Number
34		Fee For Service Agreement	Negotiation Contact Title
35		Fee For Service Agreement	Provider Fee Code
36		Healthcare Provider	Medicare Provider Number
37		Healthcare Provider	Primary Provider Identifier
38		Healthcare Provider	Provider Phone Number
39		Healthcare Provider	Type of Provider
40		Healthcare Provider	effective date

Provider Agreements

#	DATA TOPIC	ENTITY	ATTRIBUTE
41	Provider Agreements	Healthcare Provider	initial start date
42		Healthcare Provider	termination date
43		Healthcare Provider	termination reason
44		Provider Agreement	Agreement Effective End Date
45		Provider Agreement	Agreement Effective Start Date
46		Provider Agreement	Agreement Identifier
47		Provider Agreement	Agreement Version
48		Provider Agreement	Ammdment Number
49		Provider Agreement	Contract Term
50		Provider Agreement	Discount Amount
51		Provider Agreement	Highlights of Agreement
52		Provider Agreement	Incentive Program
53		Provider Agreement	Interest Exception Reason
54		Provider Agreement	Late Payment Interest Percentage
55		Provider Agreement	Network Manager Name
56		Provider Agreement	Network Manager Phone Number
57		Provider Agreement	Prompt Payment Discount Percentage
58		Provider Agreement	Stop Loss Attachment Points Percent Of Increase
59		Provider Agreement	Type of Agreement
60		Provider Agreement	Year Of Base Agreement
61		Provider Agreement Language	Language Type
62		Provider Agreement Narrow Network Coding Instruction	Network Effective Date
63		Provider Agreement Narrow Network Coding Instruction	Network Exception Reason
64		Provider Agreement Narrow Network Coding Instruction	Network Rate
65		Provider Agreement Narrow Network Coding Instruction	Network Type
66		Provider Agreement Narrow Network Decrement	Network Decrement
67		Provider Agreement Narrow Network Decrement	Notes
68		Provider Agreement Product Type	Product Type
69		Provider Agreement Renewal Date	Renewal Date
70		Provider Agreement Template	Template Version
71		Provider Agreement Template	Type of Agreement
72		Provider Fee Schedule	DRG code
73		Provider Fee Schedule	POS Code
74		Provider Fee Schedule	Procedure Code
75		Provider Fee Schedule	Procedure Code Set
76		Provider Fee Schedule	Procedure Modifier Code
77		Provider Fee Schedule	Service Fee Amount
78		Provider Fee Schedule	Service Fee Effective Date
79		Provider Fee Schedule	Service Fee Expiration Date
80		Provider Fee Schedule	Service Fee Units

Provider Agreements

Provider Data: Conceptual Model (Level 5)

#	DATA TOPIC	ENTITY	ATTRIBUTE	#	DATA TOPIC	ENTITY	ATTRIBUTE
81	Provider Agreements	Provider Fee Schedule	Standard Service Fee Amount	124	Provider Credentialing	Healthcare Provider License	license number
82		Provider Fee Schedule	Standard Service Fee Amount Qualifier	125		Healthcare Provider License	license restriction
83		Provider Fee Schedule	fee schedule effective date	126		Healthcare Provider License	license type
84		Provider Fee Schedule	fee schedule expiration date	127		Healthcare Provider License	licensure
85		Provider Fee Schedule	major surgery during inpatient stay	128		Healthcare Provider License	persistent verification indicator
86		Rate Increase	Percent Of Increase	129		Healthcare Provider License	state
87		Rate Increase	Percent Of Increase Type	130		Healthcare Provider License	status
88		Skilled Nursing Facility	Commercial And Medicare Share Rates	131		Healthcare Provider License	termination date
89		Skilled Nursing Facility	Other Payer Surplus Applies	132		Healthcare Provider License	termination reason
90	Provider Credentialing	Accreditation	accreditation body name	133		Initial Credentialing	effective start date
91		Accreditation	accreditation decision description	134		Initial Credentialing	original date
92		Accreditation	effective date	135		Institution	Institution name
93		Accreditation	end date	136		Malpractice Insurance Coverage	carrier name
94		Accreditation	notes	137		Malpractice Insurance Coverage	claim limit amount
95		Accreditation	survey date	138		Malpractice Insurance Coverage	coverage type
96		Background Check	background check indicator	139		Malpractice Insurance Coverage	effective date
97		Background Check	begin practice date	140		Malpractice Insurance Coverage	expiration date
98		Board Certification	board status	141		Malpractice Insurance Coverage	institution code
99		Board Certification	board type	142		Malpractice Insurance Coverage	policy number
100		Board Certification	certification date	143		Malpractice Insurance Coverage	specialty type
101		Board Certification	expiration date	144		Malpractice Insurance Coverage	total claim limit amount
102		Board Certification	persistent verification indicator	145		Provider Credential	credential type
103		Credentialing Application	application received date	146		Provider Credential	credential type description
104		Credentialing Application	application sent date	147		Provider Credential	fax number
105		Credentialing Application	credential type	148		Provider Credential	phone number
106		Credentialing Committee	credentialing group type	149		Provider Credentialing	category code
107		Education	degree type	150		Provider Credentialing	credentialing complete indicator
108		Education	education program type	151		Provider Credentialing	credentialing group type
109		Education	finish date	152		Provider Credentialing	effective date
110		Education	graduate complete indicator	153		Provider Credentialing	effective end date
111		Education	specialty type	154		Provider Credentialing	entity name
112		Education	start date	155		Provider Credentialing	status
113		Education Verification	source of education	156		Provider Credentialing	verification date
114		Healthcare Provider	Medicare Provider Number	157		Provider Credentialing	verification indicator
115		Healthcare Provider	Primary Provider Identifier	158		Provider Credentialing	verification type
116		Healthcare Provider	Provider Phone Number	159		Provider Service Organization	Provider Services Organization name
117		Healthcare Provider	Type of Provider	160		Provider Service Organization	Provider Services Organization name usage
118		Healthcare Provider	effective date	161		Provider Service Organization	Provider Services Organization type
119		Healthcare Provider	initial start date	162		Provider Service Organization	directory print suppress indicator
120		Healthcare Provider	termination date	163		Provider Service Organization	medicare number
121		Healthcare Provider	termination reason	164		Provider Service Organization	primary location indicator
122		Healthcare Provider License	awarded date	165		Provider Service Organization	urgent care indicator
123		Healthcare Provider License	expiration date	166		Provider Service Organization	urgent care service comments

Provider Data: Conceptual Model (Level 5)

#	DATA TOPIC	ENTITY	ATTRIBUTE
167	Provider	Provider Service Organization	walk in allowed indicator
168	Credentialing	Re-credentialing	effective start date
169	Provider Networks	Benefit Program	benefit program type
170		Benefit Program	effective date
171		Benefit Program	termination date
172		Healthcare Provider	Medicare Provider Number
173		Healthcare Provider	Primary Provider Identifier
174		Healthcare Provider	Provider Phone Number
175		Healthcare Provider	Type of Provider
176		Healthcare Provider	effective date
177		Healthcare Provider	initial start date
178		Healthcare Provider	termination date
179		Healthcare Provider	termination reason
180		Inherited Practitioner Location Network	age restriction
181		Inherited Practitioner Location Network	directory print indicator
182		Inherited Practitioner Location Network	effective date
183		Inherited Practitioner Location Network	gender limitation
184		Inherited Practitioner Location Network	highest valid age
185		Inherited Practitioner Location Network	lowest valid age
186		Inherited Practitioner Location Network	panel limit
187		Inherited Practitioner Location Network	panel status
188		Inherited Practitioner Location Network	specialty role
189		Inherited Practitioner Location Network	termination date
190		Inherited Practitioner Location Network	termination reason
191		Pharmacy Network	Number of Pharmacies
192		Practitioner	Blue Shield Practitioner Identifier Number
193		Practitioner	SSN
194		Practitioner	date of birth
195		Practitioner	date of death
196		Practitioner	first name
197		Practitioner	full name
198		Practitioner	gender
199		Practitioner	last name
200		Practitioner	medicaid number
201		Practitioner	medicare number
202		Practitioner	middle initial
203		Practitioner	noncertified provider
204		Practitioner	noncertified provider start date
205		Practitioner	noncertified provider termination date
206		Practitioner	practitioner type
207		Practitioner	prefix name
208		Practitioner	salutation name
209		Practitioner	suffix name
210		Practitioner	title name

#	DATA TOPIC	ENTITY	ATTRIBUTE
211	Provider Networks	Practitioner Location Network	age restriction
212		Practitioner Location Network	directory print indicator
213		Practitioner Location Network	effective date
214		Practitioner Location Network	gender limitation
215		Practitioner Location Network	highest valid age
216		Practitioner Location Network	lowest valid age
217		Practitioner Location Network	panel limit
218		Practitioner Location Network	panel status
219		Practitioner Location Network	specialty role
220		Practitioner Location Network	termination date
221		Practitioner Location Network	termination reason
222		Practitioner Service Location	directory print indicator
223		Practitioner Service Location	effective date
224		Practitioner Service Location	primary location indicator
225		Practitioner Service Location	termination date
226		Practitioner Service Location	termination reason
227		Practitioner Service Location	urgent care indicator
228		Practitioner Service Location	urgent care service comments
229		Practitioner Service Location	walk in allowed indicator
230		Product Offering	effective end date
231		Product Offering	effective start date
232		Product Offering	name
233		Product Offering	product identifier
234		Provider Network	network allowance
235		Provider Network	network code
236		Provider Network	network identifier
237		Provider Network	network name
238		Provider Network	ownership status
239		Provider Network	type of network
240		Provider Service Organization	Provider Services Organization name
241		Provider Service Organization	Provider Services Organization name usage
242		Provider Service Organization	Provider Services Organization type
243		Provider Service Organization	directory print suppress indicator
244		Provider Service Organization	medicare number
245		Provider Service Organization	primary location indicator
246		Provider Service Organization	urgent care indicator
247		Provider Service Organization	urgent care service comments
248		Provider Service Organization	walk in allowed indicator
249		Provider Service Organization Location	address type
250		Provider Service Organization Location	effective date
251		Provider Service Organization Location	termination date
252		Provider Service Organization Location	termination reason
253		Specialty	ancillary specialty indicator

Provider Data: Conceptual Model (Level 5)

#	DATA TOPIC	ENTITY	ATTRIBUTE	#	DATA TOPIC	ENTITY	ATTRIBUTE
254	Provider Networks	Specialty	specialty Identifier	298	Provider Demographics, Services	Healthcare Provider	initial start date
255		Specialty	specialty name	299		Healthcare Provider	termination date
256		Specialty	specialty type	300		Healthcare Provider	termination reason
257		Specialty	specialty usage	301		Healthcare Provider Address	address type
258		Specialty	taxonomy code	302		Healthcare Provider Address	effective date
259	Provider Demographics, Services	Address	accessible by train	303		Healthcare Provider Address	termination date
260		Address	accessing by bus	304		Healthcare Provider Address	termination reason
261		Address	address line 1	305		Healthcare Provider License	awarded date
262		Address	address line 2	306		Healthcare Provider License	expiration date
263		Address	city name	307		Healthcare Provider License	license number
264		Address	country name	308		Healthcare Provider License	license restriction
265		Address	county name	309		Healthcare Provider License	license type
266		Address	handicap accessibility	310		Healthcare Provider License	licensure
267		Address	latitude number	311		Healthcare Provider License	persistent verification indicator
268		Address	longitude number	312		Healthcare Provider License	state
269		Address	public transit route	313		Healthcare Provider License	status
270		Address	state code	314		Healthcare Provider License	termination date
271		Address	zip code	315		Healthcare Provider License	termination reason
272		Address	zip plus4 code	316		Inherited Practitioner Location Network	age restriction
273		Alias	alias full name	317		Inherited Practitioner Location Network	directory print indicator
274		Alias	alias name	318		Inherited Practitioner Location Network	effective date
275		Alias	alias type	319		Inherited Practitioner Location Network	gender limitation
276		Alias	effective date	320		Inherited Practitioner Location Network	highest valid age
277		Alias	first name	321		Inherited Practitioner Location Network	lowest valid age
278		Alias	last name	322		Inherited Practitioner Location Network	panel limit
279		Alias	middle initial	323		Inherited Practitioner Location Network	panel status
280		Alias	rank	324		Inherited Practitioner Location Network	specialty role
281		Alias	salutation name	325		Inherited Practitioner Location Network	termination date
282		Alias	suffix name	326		Inherited Practitioner Location Network	termination reason
283		Alias	termination date	327		Language	language name
284		Alias	termination reason	328		National Provider Identifier	NPI number
285		Authorized Signator	signer name	329		National Provider Identifier	NPI type
286		Authorized Signator	signer role	330		National Provider Identifier	description
287		Claim Healthcare Provider Payment	Provider Claim Payment Exception	331		National Provider Identifier	effective date
288		Claim Healthcare Provider Routing	Provider Claim Routing Description	332		National Provider Identifier	termination date
289		Corporate Organization / System	corporate name	333		National Provider Identifier	termination reason
290		Corporate Organization / System	corporate organization identifier	334		Practitioner	Blue Shield Practitioner Identifier Number
291		Corporate Organization / System	corporate type	335		Practitioner	SSN
292		Corporate Organization / System	tax identifier	336		Practitioner	date of birth
293		Healthcare Provider	Medicare Provider Number	337		Practitioner	date of death
294		Healthcare Provider	Primary Provider Identifier	338		Practitioner	first name
295		Healthcare Provider	Provider Phone Number	339		Practitioner	full name
296		Healthcare Provider	Type of Provider	340		Practitioner	gender
297		Healthcare Provider	effective date				

Provider Data: Conceptual Model (Level 5)

#	DATA TOPIC	ENTITY	ATTRIBUTE
341	Provider Demographics, Services	Practitioner	last name
342		Practitioner	medicaid number
343		Practitioner	medicare number
344		Practitioner	middle initial
345		Practitioner	noncertified provider
346		Practitioner	noncertified provider start date
347		Practitioner	noncertified provider termination date
348		Practitioner	practitioner type
349		Practitioner	prefix name
350		Practitioner	salutation name
351		Practitioner	suffix name
352		Practitioner	title name
353		Practitioner Hospital Privilege	effective date
354		Practitioner Hospital Privilege	hospital privilege type
355		Practitioner Hospital Privilege	primary hospital indicator
356		Practitioner Hospital Privilege	rank
357		Practitioner Hospital Privilege	restriction indicator
358		Practitioner Hospital Privilege	termination date
359		Practitioner Hospital Privilege	termination reason
360		Practitioner Language	Proficiency
361		Practitioner Location Network	age restriction
362		Practitioner Location Network	directory print indicator
363		Practitioner Location Network	effective date
364		Practitioner Location Network	gender limitation
365		Practitioner Location Network	highest valid age
366		Practitioner Location Network	lowest valid age
367		Practitioner Location Network	panel limit
368		Practitioner Location Network	panel status
369		Practitioner Location Network	specialty role
370		Practitioner Location Network	termination date
371		Practitioner Location Network	termination reason
372		Practitioner Service Location	directory print indicator
373		Practitioner Service Location	effective date
374		Practitioner Service Location	primary location indicator
375		Practitioner Service Location	termination date
376		Practitioner Service Location	termination reason
377		Practitioner Service Location	urgent care indicator
378		Practitioner Service Location	urgent care service comments
379		Practitioner Service Location	walk in allowed indicator
380		Practitioner Specialty	primary specialty indicator
381		Practitioner Specialty	specialty role
382		Provider Affiliation	affiliation type
383		Provider Affiliation	effective date
384		Provider Affiliation	rank

#	DATA TOPIC	ENTITY	ATTRIBUTE
385	Provider Demographics, Services	Provider Affiliation	restriction indicator
386		Provider Affiliation	termination date
387		Provider Affiliation	termination reason
388		Provider Business Organization	Provider Business Organization class
389		Provider Business Organization	Provider Business Organization identifier
390		Provider Business Organization	Provider Business Organization name
391		Provider Business Organization	tax identifier
392		Provider Business Organization Location	address type
393		Provider Business Organization Location	effective date
394		Provider Business Organization Location	termination date
395		Provider Business Organization Location	termination reason
396		Provider Network	network allowance
397		Provider Network	network code
398		Provider Network	network identifier
399		Provider Network	network name
400		Provider Network	ownership status
401		Provider Network	type of network
402		Provider Service Organization	Provider Services Organization name
403		Provider Service Organization	Provider Services Organization name usage
404		Provider Service Organization	Provider Services Organization type
405		Provider Service Organization	directory print suppress indicator
406		Provider Service Organization	medicare number
407		Provider Service Organization	primary location indicator
408		Provider Service Organization	urgent care indicator
409		Provider Service Organization	urgent care service comments
410		Provider Service Organization	walk in allowed indicator
411		Provider Service Organization Location	address type
412		Provider Service Organization Location	effective date
413		Provider Service Organization Location	termination date
414		Provider Service Organization Location	termination reason
415		Service Location Contact	contact method type
416		Service Location Contact	contact person full name
417		Service Location Contact	contact text
418		Service Location Contact	effective date
419		Service Location Contact	gender
420		Service Location Contact	termination date
421		Service Location Contact	title name
422		Service Location Hours of Operation	close time
423		Service Location Hours of Operation	open close
424		Service Location Hours of Operation	open time
425		Service Location Hours of Operation	week day name
426		Specialty	ancillary specialty indicator
427		Specialty	specialty Identifier

Provider Data: Conceptual Model (Level 5)

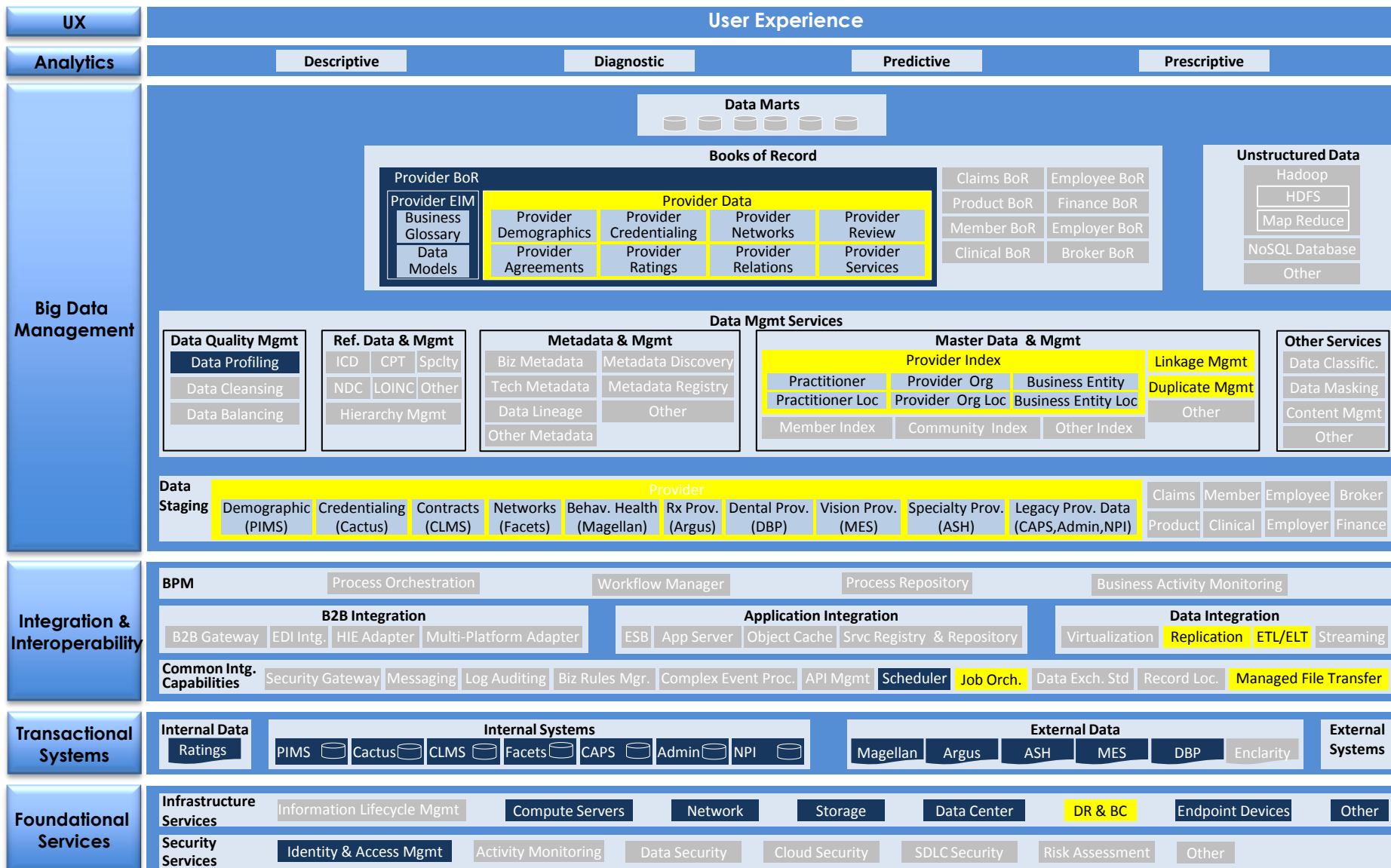
#	DATA TOPIC	ENTITY	ATTRIBUTE	#	DATA TOPIC	ENTITY	ATTRIBUTE
428	Provider Demographics, Services	Specialty	specialty name	470	California Health Performance Information (CHPI)	Clinical Quality Measures Behavioral Health	
429		Specialty	specialty type	471		Clinical Quality Measures Cardiovascular Conditions	
430		Specialty	specialty usage			Clinical Quality Measures Diabetes	
431		Specialty	taxonomy code			Clinical Quality Measures Medication Management	
432		Urgent Care Service	available service name			Clinical Quality Measures Musculoskeletal Conditions	
433	Provider Relations	Healthcare Provider	Medicare Provider Number			472	Clinical Quality Measures Prevention and Screening
434		Healthcare Provider	Primary Provider Identifier	473		Clinical Quality Measures Respiratory Conditions	
435		Healthcare Provider	Provider Phone Number	474	CHART name		
436		Healthcare Provider	Type of Provider	475	CHART participant indicator		
437		Healthcare Provider	effective date		CHART top performer indicator		
438		Healthcare Provider	initial start date	476	Cleanliness of Hospital Environment		
439		Healthcare Provider	termination date	477	Communication About Medicines		
440		Healthcare Provider	termination reason	478	Communication with Doctors		
441		Provider Agreement	Agreement Effective End Date	479	Communication with Nurses		
442		Provider Agreement	Agreement Effective Start Date	480	Discharge Information		
443		Provider Agreement	Agreement Identifier	481	Likelihood to Recommend		
444		Provider Agreement	Agreement Version	482	Overall Hospital Rating		
445		Provider Agreement	Amendment Number	483	Pain Control		
446		Provider Agreement	Contract Term	484	Quietness of Hospital Environment		
447		Provider Agreement	Discount Amount	485	Responsiveness of Hospital Staff		
448		Provider Agreement	Highlights of Agreement	486	Medicare Provider Number		
449		Provider Agreement	Incentive Program	487	Primary Provider Identifier		
450		Provider Agreement	Interest Exception Reason	488	Provider Phone Number		
451		Provider Agreement	Late Payment Interest Percentage	489	Type of Provider		
452		Provider Agreement	Network Manager Name	490	effective date		
453		Provider Agreement	Network Manager Phone Number	491	initial start date		
454		Provider Agreement	Prompt Payment Discount Percentage	492	termination date		
455		Provider Agreement	Stop Loss Attachment Points Percent Of Increase	493	termination reason		
456		Provider Agreement	Type of Agreement	494	Hospital Accreditation Indicator		
457		Provider Agreement	Year Of Base Agreement	495	Computerized Drug Orders Measure Result		
458	Provider Ratings	Blue Distinction Centers Designation	Bariatric Surgery Designation Indicator	500	High Risk Treatment Measure Result		
459		Blue Distinction Centers Designation	Cardiac Care Designation Indicator	501	ICU Physician Staffing Measure Result		
460		Blue Distinction Centers Designation	Knee and Hip Replacement Surgery Designation Indicator	502	Preventing Serious Errors Measure Result		
461		Blue Distinction Centers Designation	Spine Surgery Designation Indicator	503	Quality Index Measure Result		
462		CG CAHPS CHPI Clinical Group Performance Results	Coordination of Care Measure	504	appropriate resource use measures		
463		CG CAHPS CHPI Clinical Group Performance Results	Health Promotion Measure	505	meaningful use of IT measures		
464		CG CAHPS CHPI Clinical Group Performance Results	Office Staff Measure				
465		CG CAHPS CHPI Clinical Group Performance Results	Overall PCP Rating Measure				
466		CG CAHPS CHPI Clinical Group Performance Results	Overall Ratings of Care Measure				
467		CG CAHPS CHPI Clinical Group Performance Results	Overall Specialist Rating Measure				
468		CG CAHPS CHPI Clinical Group Performance Results	Patient Access Measure				
469		CG CAHPS CHPI Clinical Group Performance Results	Patient Doctor Interaction Measure				

Provider Data: Conceptual Model (Level 5)

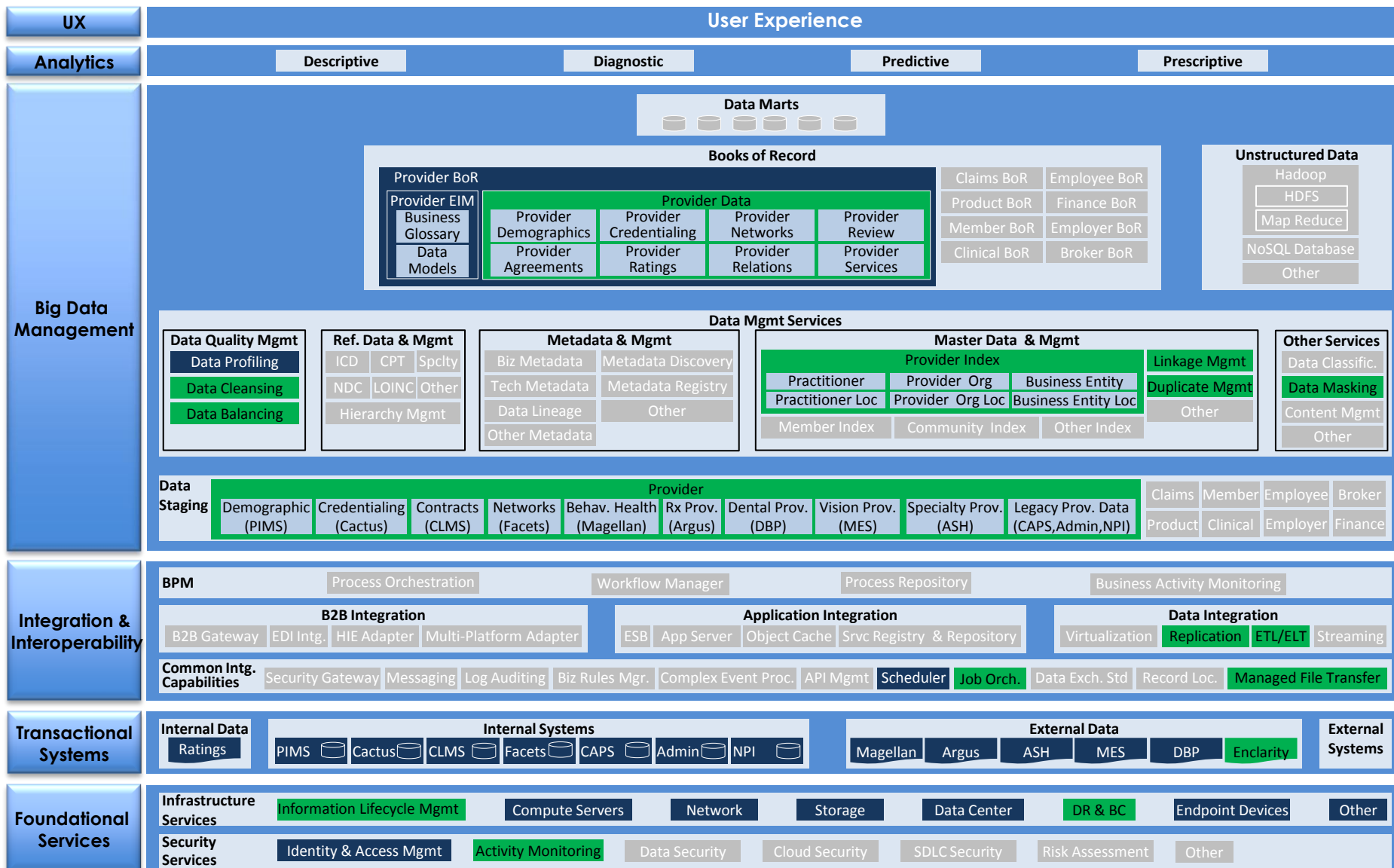
#	DATA TOPIC	ENTITY	ATTRIBUTE
506	Provider Ratings	Pay For Performance	overall patient satisfaction score
507		Pay For Performance	overall quality score
508		Pay For Performance	patient satisfaction measures
509		Pay For Performance	total cost of care trend
510	Provider Review	Healthcare Provider	Medicare Provider Number
511		Healthcare Provider	Primary Provider Identifier
512		Healthcare Provider	Provider Phone Number
513		Healthcare Provider	Type of Provider
514		Healthcare Provider	effective date
515		Healthcare Provider	initial start date
516		Healthcare Provider	termination date
517		Healthcare Provider	termination reason
518		Provider Agreement	Agreement Effective End Date
519		Provider Agreement	Agreement Effective Start Date
520		Provider Agreement	Agreement Identifier
521		Provider Agreement	Agreement Version
522		Provider Agreement	Ammendment Number
523		Provider Agreement	Contract Term
524		Provider Agreement	Discount Amount
525		Provider Agreement	Highlights of Agreement
526		Provider Agreement	Incentive Program
527		Provider Agreement	Interest Exception Reason
528		Provider Agreement	Late Payment Interest Percentage
529		Provider Agreement	Network Manager Name
530		Provider Agreement	Network Manager Phone Number
531		Provider Agreement	Prompt Payment Discount Percentage
532		Provider Agreement	Stop Loss Attachment Points Percent Of Increase
533		Provider Agreement	Type of Agreement
534		Provider Agreement	Year Of Base Agreement
535		Provider Credentialing	category code
536		Provider Credentialing	credentialing complete indicator
537		Provider Credentialing	credentialing group type
538		Provider Credentialing	effective date
539		Provider Credentialing	effective end date
540		Provider Credentialing	entity name
541		Provider Credentialing	status
542		Provider Credentialing	verification date
543		Provider Credentialing	verification indicator
544		Provider Credentialing	verification type
545	Provider Services	Healthcare Provider	Medicare Provider Number
546		Healthcare Provider	Primary Provider Identifier
547		Healthcare Provider	Provider Phone Number

#	DATA TOPIC	ENTITY	ATTRIBUTE
548	Provider Services	Healthcare Provider	Type of Provider
549		Healthcare Provider	effective date
550		Healthcare Provider	initial start date
551		Healthcare Provider	termination date
552		Healthcare Provider	termination reason
553		Practitioner	Blue Shield Practitioner Identifier Number
554		Practitioner	SSN
555		Practitioner	date of birth
556		Practitioner	date of death
557		Practitioner	first name
558		Practitioner	full name
559		Practitioner	gender
560		Practitioner	last name
561		Practitioner	medicaid number
562		Practitioner	medicare number
563		Practitioner	middle initial
564		Practitioner	noncertified provider
565		Practitioner	noncertified provider start date
566		Practitioner	noncertified provider termination date
567		Practitioner	practitioner type
568		Practitioner	prefix name
569		Practitioner	salutation name
570		Practitioner	suffix name
571		Practitioner	title name
572		Practitioner Specialty	primary specialty indicator
573		Practitioner Specialty	specialty role
574		Provider Affiliation	affiliation type
575		Provider Affiliation	effective date
576		Provider Affiliation	rank
577		Provider Affiliation	restriction indicator
578		Provider Affiliation	termination date
579		Provider Affiliation	termination reason

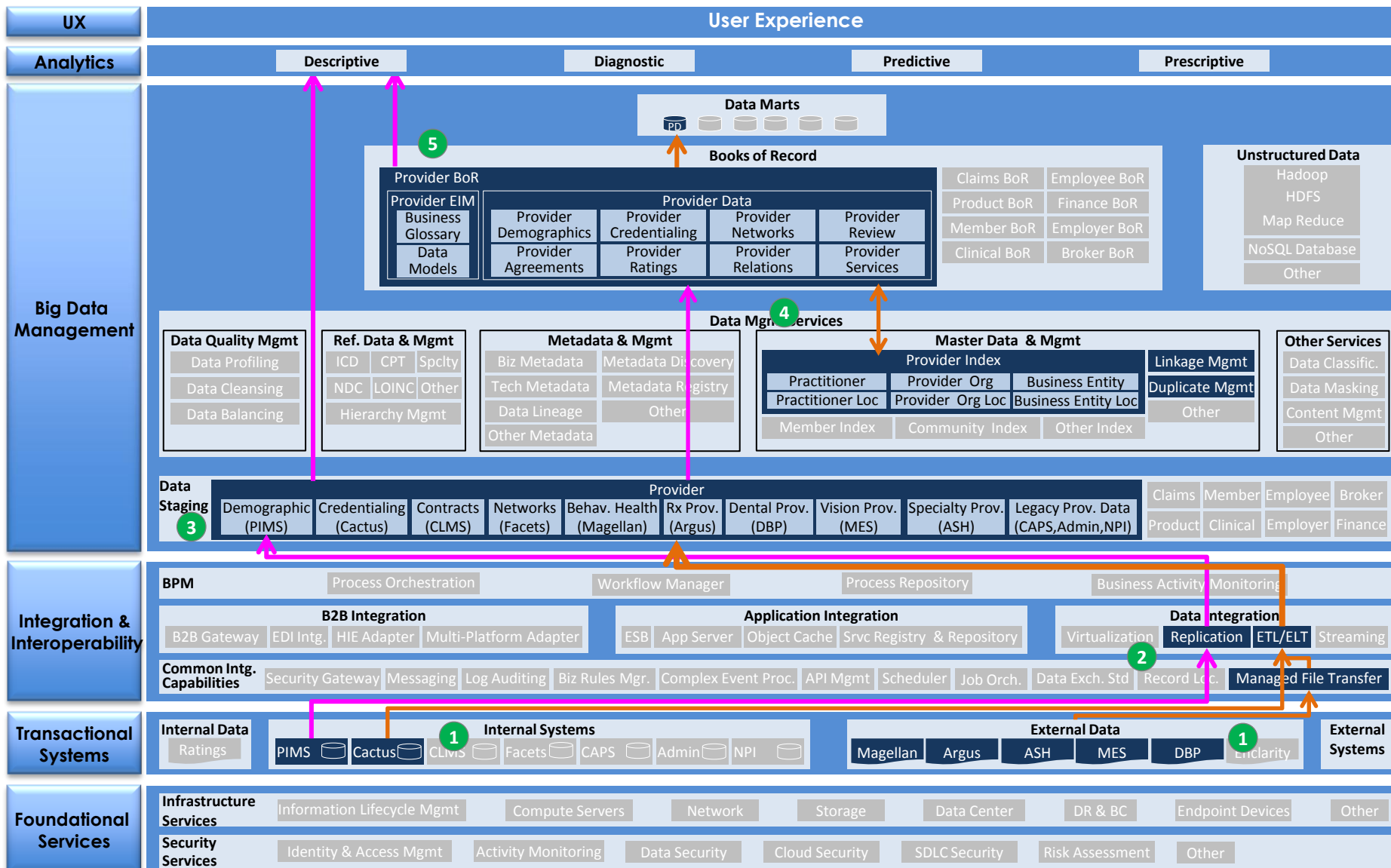
Provider Data: Current State



Provider Data: Future State



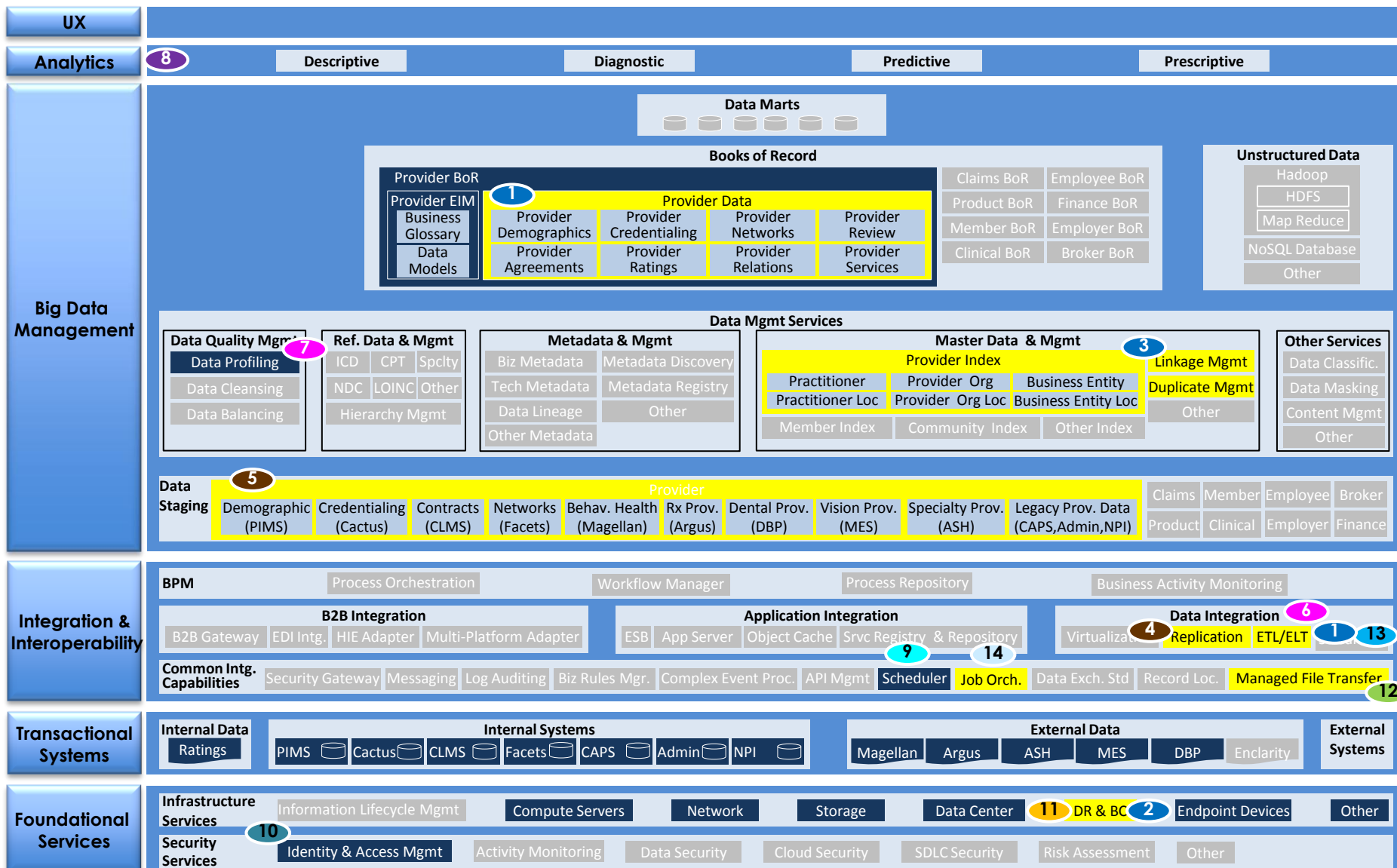
Provider Data: Provider BoR Use Case



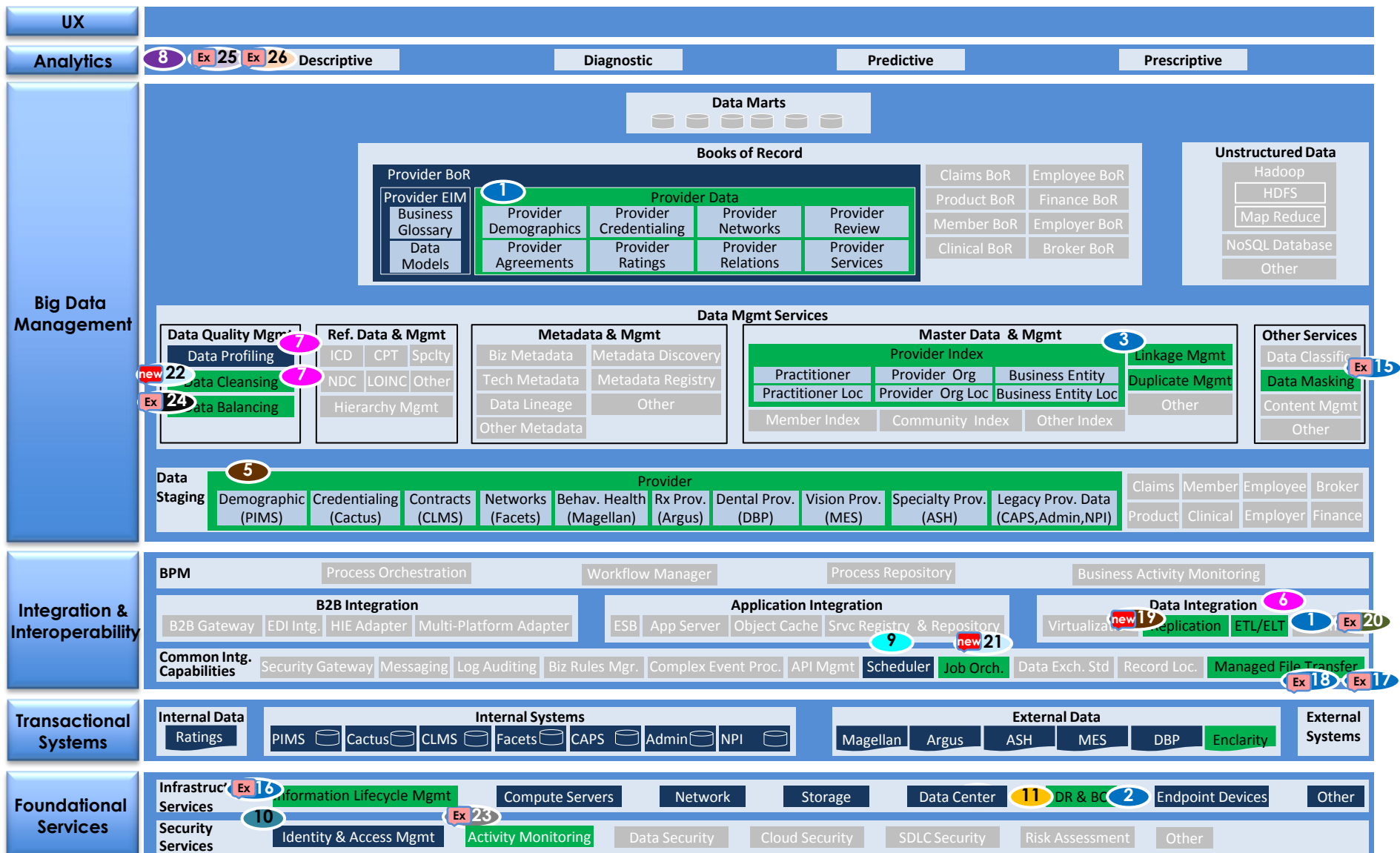
Provider Data: Current Technology Footprint

Technology	Vendor	Product	Capability
1	IBM	PureData System for Analytics (Netezza)	Data warehouse/analytics appliance platform for data storage and data integration in extract, load, transfer (ELT) pattern
2	IBM	Netezza Replication Services (PTS)	Asynchronous data replication for disaster recovery
3	IBM	Infosphere MDM Standard Edition (SE)	Master data management in registry (virtual) style
4	Oracle	Data Guard	Data replication (physical) for disaster protection and high availability
5	Oracle	Oracle Relational Database	Relational database management
6	Informatica	Informatica PowerCenter	Data integration in extract, transform, load (ETL) pattern
7	Informatica	Informatica Data Quality	Data quality - data profiling, data cleansing and address standardization
8	SAP	Business Objects	Reporting tool for reporting and ad hoc analysis using a semantic layer
9	Cisco	Tidal Scheduler	Batch job automation and scheduling
10	Microsoft	Active Directory	Directory service for user authentication
11	Symantec	NetBackup	Database backup and recovery
12	Red Hat	SFTP	Secure file transfer
13	Unfuddle	Unfuddle	Software versioning and revision control system
14	BSC IT	ABC Framework	ETL/ELT Job orchestration, audit and control

Provider Data: Current State Technology



Provider Data: Future State Technology



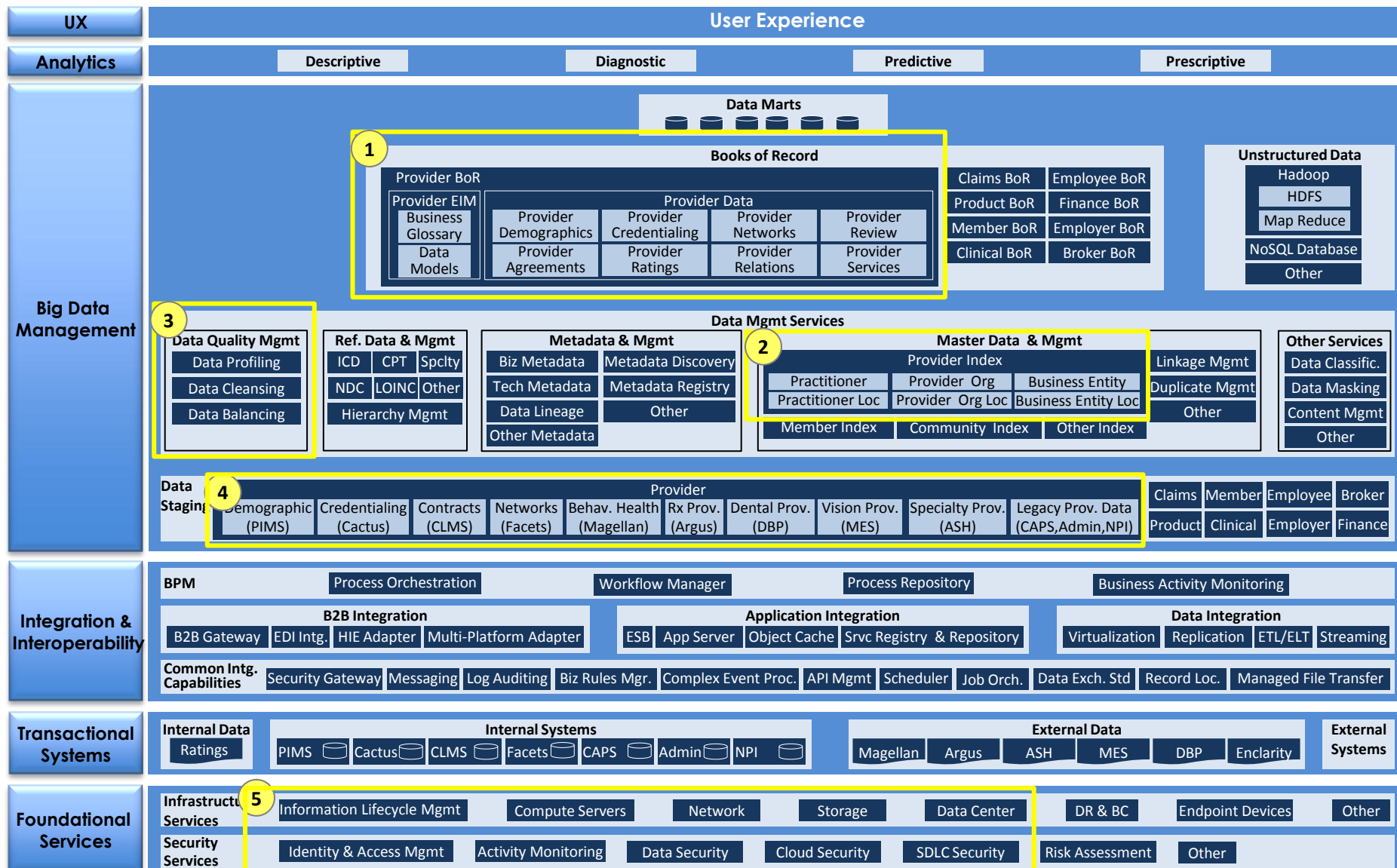
Provider Data: Future State Technology Footprint

Technology	Vendor	Product	Capability
1	IBM	PureData System for Analytics (Netezza)	Data warehouse/analytics appliance platform for data storage and data integration in extract, load, transfer (ELT) pattern
2	IBM	Netezza Replication Services (PTS)	Asynchronous data replication for disaster recovery
3	IBM	Infosphere MDM Standard Edition (SE)	Master data management in registry (virtual) style
Ex 15	IBM	Optim Data Privacy	De-identification of confidential data to protect data privacy
Ex 16	IBM	Optim Archive	Archive historical data for reduced storage costs and compliance
Ex 17	IBM	Sterling File Gateway	External file transfer gateway for security exchange of file-based data
Ex 18	IBM	MQ Managed File Transfer	Real-time transfer of files within the MQ network
new 19	Oracle	Golden Gate	Data integration using real-time log based data replication
5	Oracle	Oracle Relational Database	Relational database management
6	Informatica	Informatica PowerCenter	Data integration in extract, transform, load (ETL) pattern
7	Informatica	Informatica Data Quality	Data quality - data profiling, data cleansing and address standardization
8	SAP	Business Objects	Reporting tool for reporting and ad hoc analysis using a semantic layer
9	Cisco	Tidal Scheduler	Batch job automation and scheduling
10	Microsoft	Active Directory	Directory service for user authentication
11	Symantec	NetBackup	Database backup and recovery
Ex 20	Apache	Subversion	Software versioning and revision control system
new 21	BSC IT	Netezza Job Orchestration	ELT Job orchestration, audit and control
new 22	BSC IT	Provider Data Quality (PDQ)	Workflow tool for reviewing data updates and updating PIMS
Ex 23	Imperva	SecureSphere Database Activity Monitor	Full auditing and visibility into database data usage
Ex 24	Infogix	Infogix Assure	Data balancing for ensuring data integrity between source and BoR
Ex 25	SAS	SAS	Tool for advanced analytics including data mining and predictive modeling
Ex 26	Tableau	Tableau	Tool for data visualization, dashboards and interactive reporting

Table of Contents

1. Executive Summary
2. Concept Definition and Model Framework
3. Environmental Analysis
4. Current State Architecture Models
5. Future State Architecture Models
- 6. Architecture Analysis**
 - Gap Analysis
 - Impact Analysis
 - Technology Solutions Analysis
 - Risk Analysis
7. Roadmap
8. Recommendation
9. Appendices

Provider Data: Areas for Gap Analysis



Provider Data: Gap Analysis

#	Technical Capability Set	Capability Description	Current State	Technology Gap / Opportunities
1A	Provider BoR – Data Loading	<ul style="list-style-type: none"> Load data from sources using truncate and load method 	<ul style="list-style-type: none"> Custom Java script called Data Pump 	<ul style="list-style-type: none"> Instead of custom Java scripts, opportunity to use standard data movement pattern using Informatica PowerCenter (ETL standard)
1B	Provider BoR - Data Model	<ul style="list-style-type: none"> BoR data model based on the EIM approved by Enterprise Data Governance 	<ul style="list-style-type: none"> data model based on P360 project requirement 	<ul style="list-style-type: none"> Opportunity to base BoR data model on the Enterprise Information Model for provider approved by Data Governance
1C	Provider BoR - Data Content	<ul style="list-style-type: none"> Single authoritative source of truth for Provider data 	<ul style="list-style-type: none"> Gap 	<ul style="list-style-type: none"> Implement Provider BoR including presentation layer
1D	Provider BoR - ELT Error Handling and Recovery	<ul style="list-style-type: none"> Gracefully handle errors during ELT processes, including logging, recovery to ensure system resiliency 	<ul style="list-style-type: none"> Rudimentary error handling and recovery processes 	<ul style="list-style-type: none"> Define and implement a robust error handling and recovery strategy in Netezza ELT processes
1E	Provider BoR - Intraday Batches	<ul style="list-style-type: none"> Run ELT batches multiple times a day 	<ul style="list-style-type: none"> Ability exists; need of minor updates to ELT scripts 	<ul style="list-style-type: none"> Update Netezza ELT scripts for intra-day batch capability
1F	Provider BoR - Pre-processing Data Stores	<ul style="list-style-type: none"> Provide persistence during pre-processing stages, integration with Provider Index before loading to BoR target tables 	<ul style="list-style-type: none"> Multiple Netezza databases (Intake, Stage, Work and Xref) 	<ul style="list-style-type: none"> Consolidate Intake, Stage, Work and Xref databases to reduce the number of databases for backend pre-processing
1G	Provider BoR - Code Version Control	<ul style="list-style-type: none"> Manage changes to source code, including a repository for storing versions of code 	<ul style="list-style-type: none"> Unfuddle tool (cloud open source) 	<ul style="list-style-type: none"> Leverage BSC standard tool for code version control – Apache Subversion
1H	Provider BoR – ELT	<ul style="list-style-type: none"> ELT pattern for MPP performance but governance framework needed to ensure code consistency and design for supportability 	<ul style="list-style-type: none"> Netezza ELT scripts 	<ul style="list-style-type: none"> Opportunity to define and implement ELT framework for consistent code development and enforce it using code reviews and governance

Provider Data: Gap Analysis (contd.)

#	Technical Capability Set	Capability Description	Current State	Technology Gap / Opportunities
1G	Provider BoR – Roles & Responsibility	<ul style="list-style-type: none"> • Delivery of the solution design for BoR project 	<ul style="list-style-type: none"> • App Services is unclear about role and responsibility 	<ul style="list-style-type: none"> • Opportunity for Solution Design team to partner with Data Liquidity team for detailed design of BoR initiatives
2A	Provider Index – Match/merge and Golden Record	<ul style="list-style-type: none"> • Probabilistic matching to link business entities across systems of record and merging of attributes based on source priority and trust rules to create “golden record” 	<ul style="list-style-type: none"> • Provider Index using Infosphere MDM tool (design completed for Provider Org entity) 	<ul style="list-style-type: none"> • Implement Provider Index with 6 provider entities (Practitioner, Provider Org, Business Entity, Practitioner Location, Provider Org Location, Business Entity Location) using IBM Infosphere MDM tool
2B	Provider Index – Data Integration with BoR	<ul style="list-style-type: none"> • Batch internal transfer of Provider Index inbound and outbound files from Provider BoR pre-processing data stores 	<ul style="list-style-type: none"> • SFTP 	<ul style="list-style-type: none"> • Opportunity to leverage BSC standard tool MQ MFT for internal file transfers
3A	Provider Data Quality – Workflow Automation for Bulk Update of PIMS	<ul style="list-style-type: none"> • Ingest provider data updates from Provider IPA/Groups, regulatory agencies, external data vendors and apply bulk updates to PIMS either after review or fully automated 	<ul style="list-style-type: none"> • Gap 	<ul style="list-style-type: none"> • Implement custom Provider Data Quality (PDQ) workflow tool to review and make bulk updates to PIMS, leveraging data updates from external and internal sources
3B	Provider Data Quality –Verification/ Augmentation using Vendor feed	<ul style="list-style-type: none"> • Improve provider demographic data accuracy by verification and augmentation using vendor referential data feed 	<ul style="list-style-type: none"> • Gap 	<ul style="list-style-type: none"> • Update provider data in PIMS leveraging Enclarity verification/augmentation data feeds, either manually or using PDQ workflow tool
3C	Provider Data Quality - Data Cleansing and Standardization	<ul style="list-style-type: none"> • Standardize and cleanse data based on business requirements or leveraging standardization tools, preferably at source or close to source 	<ul style="list-style-type: none"> • Gap 	<ul style="list-style-type: none"> • Opportunity to standardize provider addresses, cleanse data based on business rules in PIMS and Provider BoR, leveraging Informatica IDQ and PDQ tool
3D	Provider Data Quality - Data Profiling	<ul style="list-style-type: none"> • Profile data based on business rules to create scorecards and metrics to monitor data quality 	<ul style="list-style-type: none"> • Informatica IDQ used – but limited use 	<ul style="list-style-type: none"> • Opportunity to implement business defined data profiling scorecards to measure data quality at PROPS and BoR touch points

Provider Data: Gap Analysis (contd.)

#	Technical Capability Set	Capability Description	Current State	Technology Gap / Opportunities
3E	Provider Data Quality - Data Balancing & Reconciliation	<ul style="list-style-type: none"> Balance and reconcile data between source systems and Provider BoR based on business defined controls 	<ul style="list-style-type: none"> Gap 	<ul style="list-style-type: none"> Leverage Infogix to implement out of band balancing process using business defined control metrics between sources and BoR for providing assurance, compliance and audit ability
3F	Provider Data Quality – People	<ul style="list-style-type: none"> Assign people to create data quality plan, identify data quality metrics, perform initial assessment, articulate rules for cleansing data and set up monitoring systems to maintain adequate levels of data quality 	<ul style="list-style-type: none"> Gap 	<ul style="list-style-type: none"> Opportunity for the provider data owner and data stewards to launch a data quality program that addresses the missing capabilities from a people perspective
3G	Provider Data Quality – Process	<ul style="list-style-type: none"> Define and operate a governance process for managing the data quality plan, data quality metrics, data quality assessment, data cleansing rules and monitoring of data quality 	<ul style="list-style-type: none"> Gap 	<ul style="list-style-type: none"> Opportunity for the provider data owner and data stewards to launch a data quality program that addresses the missing capabilities from a process perspective
4A	Provider Data Staging- Data Storage	<ul style="list-style-type: none"> Collect data from multiple provider data sources (internal and external) without transformation and support real-time operational reporting, data quality measurement, auditing and real-time data services 	<ul style="list-style-type: none"> PROPS database for non-PIMS sources; PIMS standby database for PIMS PIMS as operational reporting data source 	<ul style="list-style-type: none"> Inability to link PIMS with non-PIMS provider data; opportunity to co-locate all sources in schemas in a single physical database Business Objects operational reports need to leverage PROPS instead of PIMS database
4B	Provider Data Staging – External Batch Data Transfer	<ul style="list-style-type: none"> Exchange data files with external partners/vendors in a secure manner 	<ul style="list-style-type: none"> SFTP server 	<ul style="list-style-type: none"> SFTP is not the standard for B2B data transfers; opportunity to leverage Sterling File Gateway (BSC Standard)

Provider Data: Gap Analysis (contd.)

#	Technical Capability Set	Capability Description	Current State	Technology Gap / Opportunities
4C	Provider Data Staging - Real-time Data Replication	<ul style="list-style-type: none"> Capture, routing, transformation and delivery of operational data between heterogeneous databases in real time with minimal overhead 	<ul style="list-style-type: none"> Oracle Data Guard – a disaster protection and high availability tool used for PIMS standby database 	<ul style="list-style-type: none"> Tight coupling between source and target,, latency issues etc; opportunity to use GoldenGate – for logical data replication in real-time, with support for filtering, heterogeneous databases, change data capture
5A	Common Integration Capabilities - Job Orchestration	<ul style="list-style-type: none"> Orchestrate ELT jobs - define jobs, cycles, parameters, dependencies; error logging and notification; restart ability, visualization, job metadata, audit, balance and control 	<ul style="list-style-type: none"> ABC framework (dependent on Informatica) 	<ul style="list-style-type: none"> Opportunity to implement Netezza based job orchestration framework for ELT and remove dependency on Informatica
5B	Disaster Recovery & Business Continuity (DR&BC) – Netezza DR	<ul style="list-style-type: none"> Ability for recovery and continuity of technology infrastructure and systems following a natural or human-induced disaster 	<ul style="list-style-type: none"> IBM PTS Replication tool installed but not operational Data transfers over unsecure channels 	<ul style="list-style-type: none"> Complete implementation of PTS Replication; Develop and test Netezza DR strategy Implement link level encryption between data centers for secure data transfers between data centers
5C	Information Lifecycle Management (ILM) - Data Archiving	<ul style="list-style-type: none"> Archive provider data for 10 years with about 3 years to be maintained online to comply with BSC Records Retention policy 	<ul style="list-style-type: none"> Gap 	<ul style="list-style-type: none"> Define and implement data archiving strategy for BoR data to ensure compliance with BSC data retention policy, leveraging IBM Optim Archive tool
5D	Data Management Services - Data Masking	<ul style="list-style-type: none"> Mask PHI (if any) and IIPI (SSN, Tax Id, License number) in non-production environments and if required create integrated sub-sets of data 	<ul style="list-style-type: none"> Gap 	<ul style="list-style-type: none"> While copying production data to non-production environments, leverage Optim Data Masking tool for data masking and creating integrated sub sets.
5E	Security Services – Database Activity Monitoring	<ul style="list-style-type: none"> Ability to monitor and audit activities by users using privileged accounts or pooled connections from other applications 	<ul style="list-style-type: none"> Gap 	<ul style="list-style-type: none"> Opportunity to monitor and audit user access using Imperva database activity monitoring tool

Provider Data: Impact Analysis

Capability Set	People	Process	Technology
1. Provider BoR	<ul style="list-style-type: none"> Need for maintenance and governance of Provider EIM by Enterprise Data Governance data stewards and IT Need to train business users of Provider BoR in data content and usage 	<ul style="list-style-type: none"> Need to create or leverage process for managing data issues in Provider BoR including issue identification, tracking, resolution, communication and escalation Need to set up access and security processes for managing access requests to Provider BOR 	<ul style="list-style-type: none"> Need for remediation to extracts and reporting/analytics using provider data to now source from Provider BoR Need to retire use of Rosetta marts as source of truth for provider information
2. Provider Index (MDM)	<ul style="list-style-type: none"> Need to train IT development teams in customization and configuring of the Infosphere MDM tool to avoid reliance on vendor professional services Need to train business data stewards in linkage management, duplicate management and other administrative tasks using the Infosphere MDM tool, evolving into a MDM competency center over time 	<ul style="list-style-type: none"> Need to create a governance framework for managing provider entities, with defined metrics for measuring progress and communicating success to stakeholders – and align it with Enterprise Data Governance 	<ul style="list-style-type: none"> Implement Provider Index (Infosphere MDM) with 6 provider entities Upgrade the other instances of Infosphere MDM SE tool (Provider Match and EMPI) to current version 11 and assess feasibility of merging Provider Match instance with Provider Index
3. Provider Data Quality	<ul style="list-style-type: none"> Need for training the business team managing provider information in effective use of the PDQ workflow tool and Informatica data profiling tool Using PDQ workflow tool will optimize staffing and reduce reliance on SME's for managing basic provider data 	<ul style="list-style-type: none"> Need for creating a business driven process for ongoing provider data quality measurement and identification of data cleansing rules for improving data quality in PIMS and Provider BoR 	<ul style="list-style-type: none"> Deploy the custom built Provider Data Quality (PDQ) workflow tool Implement data exchange with Enclarity

Provider Data: Impact Analysis (contd.)

Capability Set	People	Process	Technology
4. Provider Data Staging	<ul style="list-style-type: none"> Need for GoldenGate training to IT development and database support teams Need for awareness in IT solution architecture and design teams for applying data replication design pattern to appropriate use cases 	<ul style="list-style-type: none"> Need to implement governance process to ensure PROPS is only used for operational reporting, real-time data source and data quality and auditing 	<ul style="list-style-type: none"> Deploy GoldenGate as enterprise data replication tool
5. Other Foundational Services	<ul style="list-style-type: none"> Need for training IT development and support teams in custom built Netezza Orchestration tool and Infogix data balancing tool 		<ul style="list-style-type: none"> Deploy the Netezza Job Orchestration tool Deploy Netezza PTS Replication for DR Deploy the Imperva DAM agent on the Netezza database

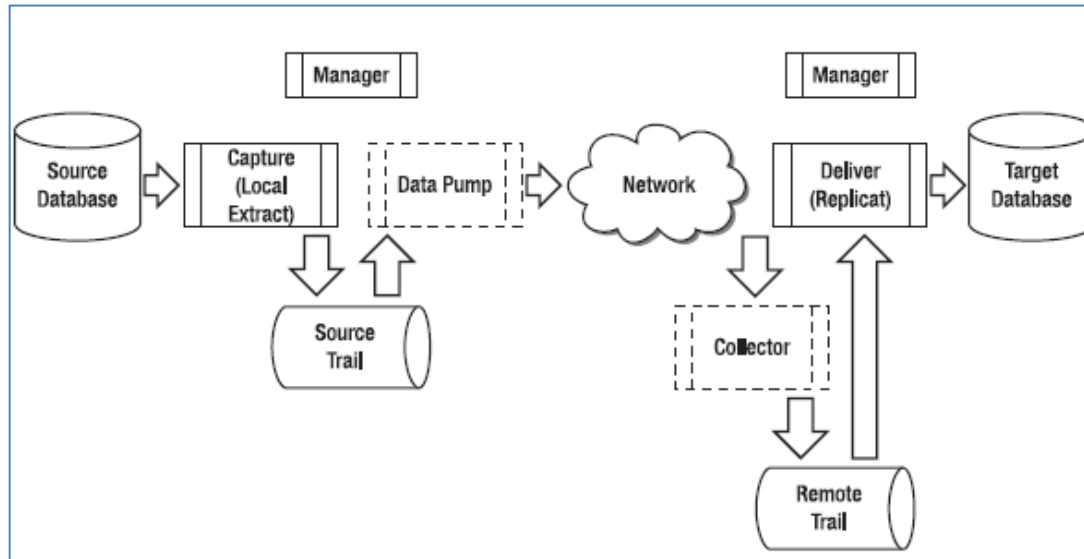
Impact Analysis: Business Component/Skill Matrix

Business Component	Skills
Data Governance (Data Stewards)	<ul style="list-style-type: none">• Strong communication skills• Detail oriented with strong organizational skills• Analytical and problem solving skills• Understanding of data concepts (entities)
Provider Index	<ul style="list-style-type: none">• Understanding of data concepts (entities)• Business rules• Understanding of the concept of golden record• SQL / other query skills
Data Quality	<ul style="list-style-type: none">• Business rules• Understanding of measuring data quality• Understanding of business processes• Detail oriented with strong organizational skills• Root cause analysis
Books of Record	<ul style="list-style-type: none">• SQL / other query skills• Understanding of the concept of golden record• Understanding of business processes

Technology Solution Analysis: Data Replication Tool Architecture

Key Features & Benefits

- **Low impact data replication** - Moves thousands of transactions per second with negligible impact on source and target systems
- **Integration** - Integrates with ETL tools as well as has adapters for JMS and flat files
- **Heterogeneous support**- Supports heterogeneous databases and platforms, Hadoop to increase IT flexibility
- **Real-time data** - Immediately captures, routes, transforms, and delivers transactional data to other systems with sub-second latency
- **Transaction integrity** - Maintains transaction commit boundaries and atomicity, consistency, isolation, and durability (ACID) properties as transactions are moved between source and target systems, ensuring data consistency and referential integrity
- **Reliability** - Delivers all committed records to the target, even in the event of network outages. Moves data without requiring system interruption or outage windows
- **Flexible topology support** - Moves data in one-source-to-one-target, one-to-many, many-to-one, many-to-many, cascading, and bidirectional configurations



How Data Replication Works

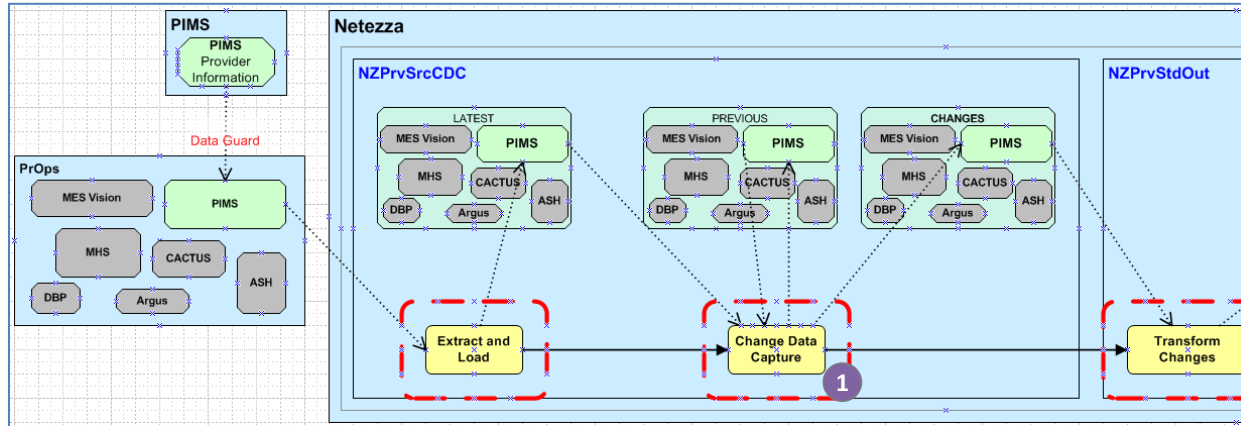
1. New and changed database data is **captured** from the **source database**.
2. The captured data is written to a file called the **source trail**.
3. The trail is then read by the **data pump**, sent across the **network** to the destination
4. At the destination server, it is written to a **remote trail** file by the **collector** process
5. The **delivery** function reads the remote trail and updates the **target database**
6. Each of the components is managed by the **manager** process

Data Replication Tool: Usage Scenarios

	Use Case Description
Use Case 1: Changed Data Capture (CDC)	Some source systems (i.e. PIMS, Facets) do not support audit columns that help change data capture (CDC) using traditional ETL tools. Consequently, CDC has to be done using brute force full table compare in the target system staging area. A data replication tool can provide the change timestamps to enable CDC.
Use Case 2: Data Provisioning for BI/Analytics	HPXR and RDI use Informatica PowerCenter (ETL) to copy data from Facets for BI data marts, including full table scans and joins during the daily load cycle. Using a low-impact data replication tool will reduce performance impact to Facets and also simplify data acquisition for BI and analytics through leverage of CDC and table replication with filters. For, e.g. RDI uses 80+ Informatica mappings to copy Facets tables over. A data replication tool can enable access to transactional source data without bogging down production systems.
Use Case 3: Operational Reporting & Dashboards	Operational Reporting and Dashboards need near real time operational data for workforce management, inventory management, command center, etc. Currently we use logical standby databases which create constraints for data management and reporting. A data replication tool can provision selective real-time data for operational BI applications.
Data Services	Data Services for exposing real-time information via web services and data virtualization, would need to go against the transactional systems adding to the workload. Replicating source data to a operational/staging data stores using a data replication tool will enable efficient operational data integration use cases.
Zero-downtime Operations	Enable uninterrupted business operations during system update, migration and maintenance activities. Also, useful for high availability to prevent data loss and downtime prevention.

Use Case 1: Provider BoR Change Data Capture

Current State

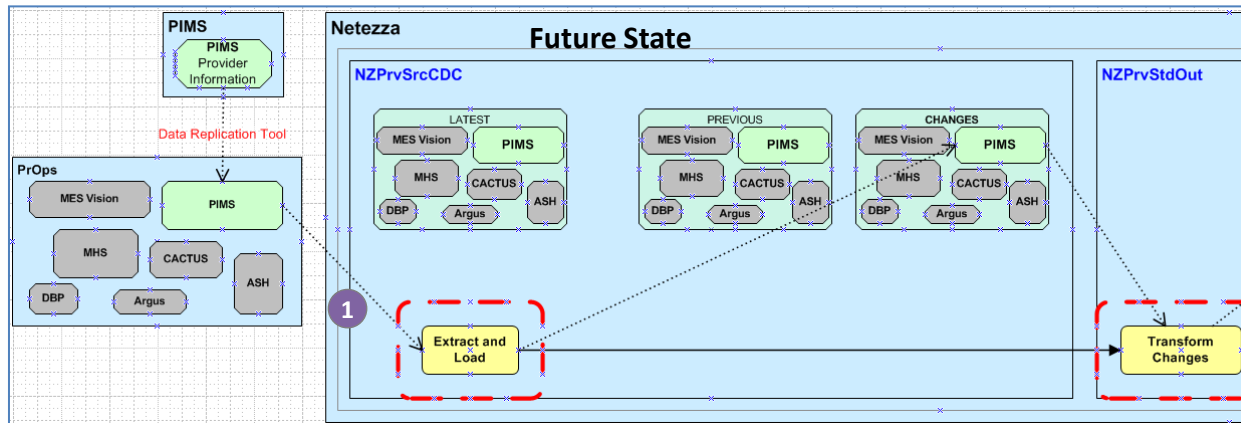


challenges

1 Change Data Capture

- PIMS is a COTS application and database tables do not have audit timestamps for CDC via ETL tool
- CDC done using brute force, full table compare in Netezza for about **140+ tables** – wasting storage and compute cycles

Future State

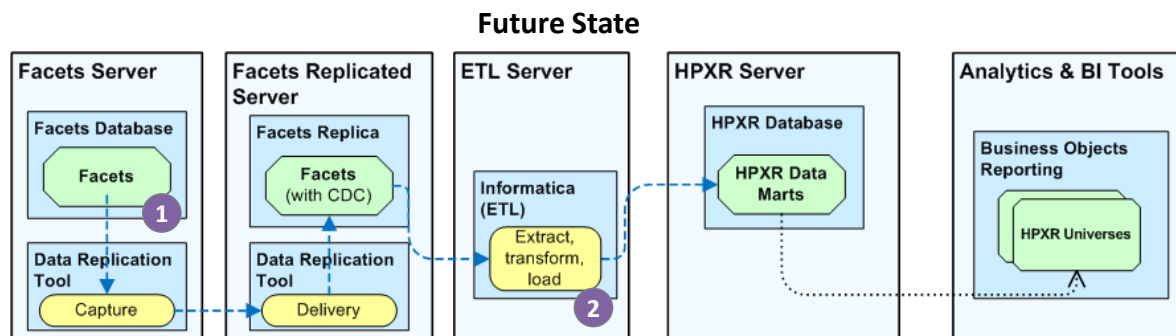
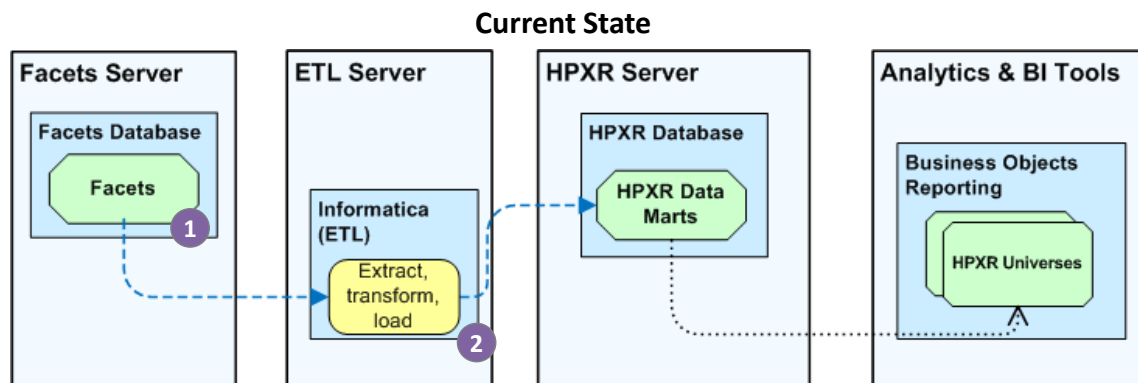


future state recommendation

1 Change Data Capture

- Using Data Replication tool, the change timestamps can be recorded in the PROPS database and the Netezza extract process can leverage it for incremental loads

Use Case 2: Facets/HPXR ETL Refresh Improvement



challenges

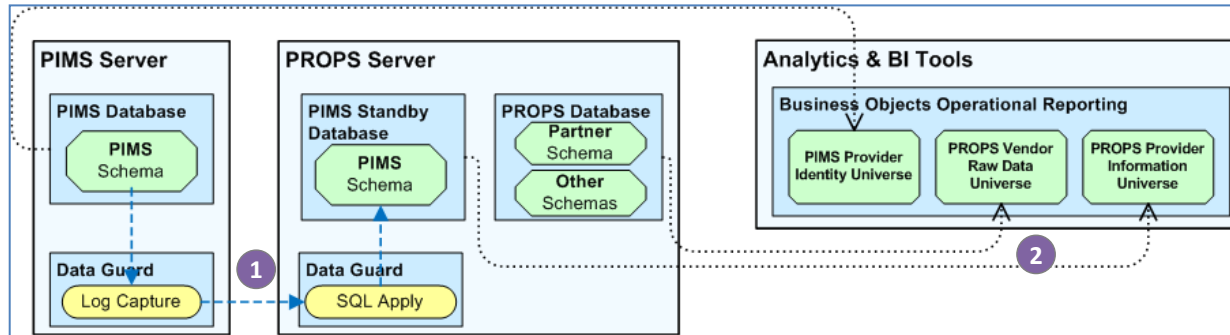
- 1 **Impact to Facets Transactional System**
 - HPXR daily refresh using Informatica ETL run against Facets transactional database during business hours
 - Lack of audit timestamps on Facets tables requires joins on entire tables during extraction (**about 300 tables**)
 - HPXR refresh **imposes average 10% CPU utilization** on Facets DB server
- 2 **HPXR Refresh Cycle exceeds SLA**
 - Inability for incremental loads, resulted in refresh time to gradually **increase from 2 hrs in Jan 2015 to 9-10 hrs in May (SLA is 5 hrs)**

future state recommendation

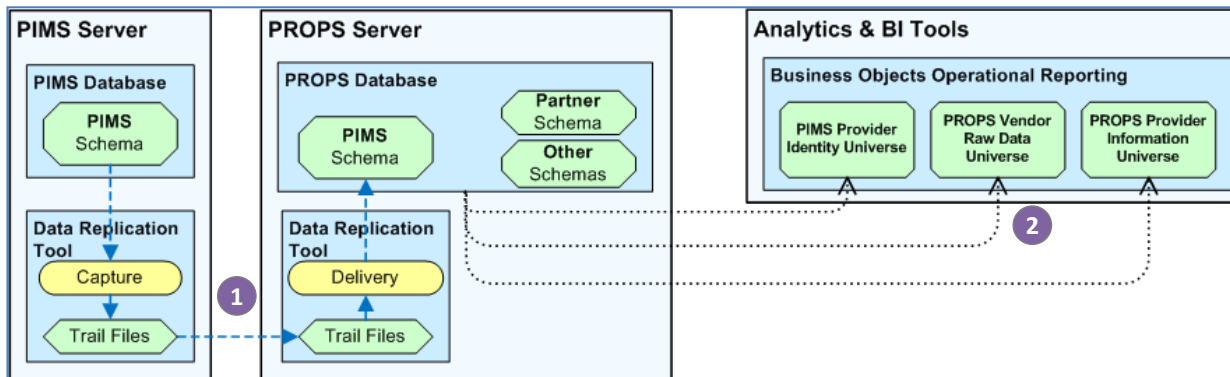
- 1 **Impact to Facets Transactional System**
 - Data replication to Facets Replicated database will result in negligible impact and overhead on Facets transactional system
- 2 **HPXR Refresh Cycle exceeds SLA**
 - With Data Replication tool based CDC, the Informatica mappings can be re-mediated to only pick incremental changes and drastically reduce the processing and cycle time for HPXR refresh
 - Facets is unlocked for any other reporting/analytics – for the entire enterprise

Use Case 3: PIMS to PROPS Replication

Current State



Future State



challenges

- 1 **PIMS to PROPS Replication**
 - PIMS and non-PIMS data forced into separate databases on PROPS, not allowing joins between PIMS and other sources
 - Latency issues due to source load or network issues – requires rebuild of standby database
 - Cost issues as target database needs to be same size as source
 - No ability for filtering, transformation during data transfer – all 1340 tables are replicated
- 2 **Operational Reporting (BO)**
 - Latency issues and two physical databases require BO universes to go against transactional system and multiple universes

future state recommendation

- 1 **PIMS to PROPS Replication**
 - Data Replication tool allows replication to single target database, ability to select tables, transformation, filter (columns, rows) and heterogeneous target databases
- 2 **Operational Reporting (BO)**
 - Replication to single target at low latency allows for efficient operational reporting

Technology Solutions Analysis: Data Replication

Assessment Category	Assessment Factor	Option 1 Oracle GoldenGate	Options 2 IBM Infosphere Data Replication	Option 3 Informatica Data Replication
Functionality	Support for foundational functionality (unobtrusive log capture, real-time processing, heterogeneous databases)	5	5	5
	Support for differentiating functionality (guaranteed delivery, transactional integrity, light transformations, bidirectional, encryption)	5	4	4
Technology Maturity	Technology maturity for mission critical use (scalability, performance, security, reliability)	5	5	4
	Technology adoption in the healthcare industry. Ability to support future BSC needs	5	5	5
Agility and Usability	Usability of tools and support for agile development	5	5	5
Skill-set	Availability of skill-set – commodity skill-set versus specialized vendor-specific skills	5	4	4
Vendor Supportability	Vendor and product (roadmap) viability	5	5	4
	Vendor support & SLA	5	5	5
Cost	Software acquisition	3	4	4
	Implementation and ongoing maintenance	5	4	4
Current BSC Supportability	Technology support structure and integration with existing tools. In-house skills	5	4	4
	Synergies with existing investments (transactional databases, data integration tools)	5	3	3
		58	53	51

Table 2. Product/Service Rating on Critical Capabilities

Product or Service Ratings	Action	Adeptia	Cisco	IBM	Informatica	Information Builders	Microsoft	Oracle	SAP	SAS	Syncsort	Talend
Bulk/batch data movement	5.0	4.3	2.0	5.0	5.0	4.2	4.4	4.5	4.6	4.2	5.0	4.3
Data federation/virtualization	1.0	2.0	4.8	4.0	4.0	4.0	2.3	2.9	3.6	3.9	1.4	1.2
Message-oriented movement	3.0	3.5	1.0	4.5	3.8	3.5	3.6	3.9	3.5	2.3	1.1	3.0
Data replication & synchronization	2.8	1.1	1.0	4.7	4.5	3.0	2.9	5.0	4.1	3.0	2.7	2.8

Source: Critical Capabilities for Data Integration Tools, Gartner, Nov 2014

Source: EA Assessment

Rating Legend

Best Fit 5 4 3 2 1 Poor Fit

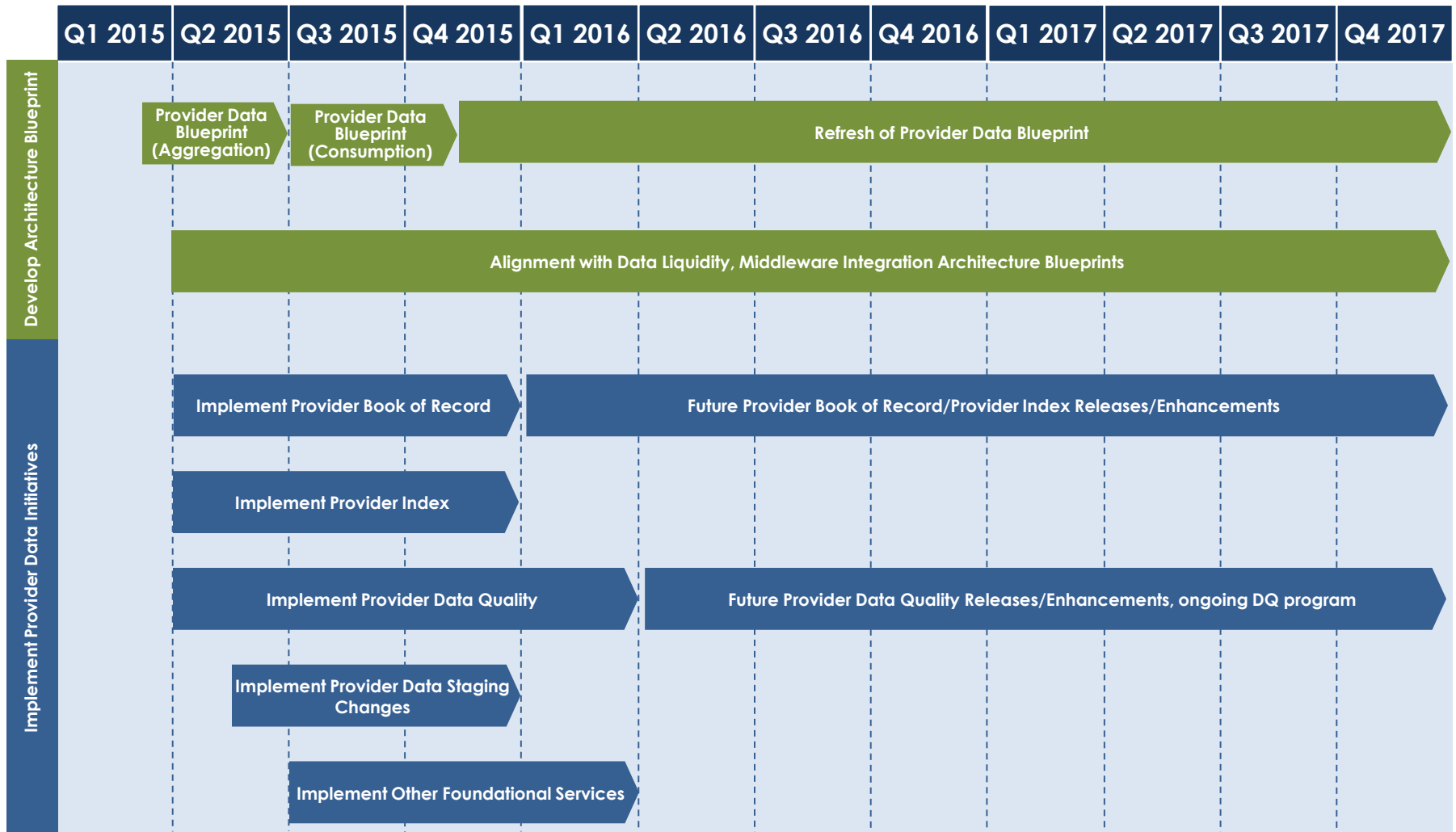
Provider Data: Risk Analysis

	Description
Provider BoR	The Infosphere MDM tool provides composite views for golden records, and relies on Provider BoR to persist the golden record, maintain attribute change history, traceability – adding to complexity of physical data model and provisioning of a presentation layer will help ease of user navigation.
Provider Index	The Provider Index data model with 6 entities and relationships is complex and will need clear business rules for match/merge, thresholds etc., and keep customizations low. Also, the hardware sizing may need to be re-validated to ensure SLA's can be met.
Network Infrastructure	Although, sufficient network bandwidth and SSL encrypted channels exist between COLO1 and COLO2, traffic gets routed through EDH and Lodi legacy data centers (retiring end 2015) and until then there is potential risk of spikes during large bulk loads (Netezza DR replication) or additional load due to Imperva DAM monitoring.
Skill-Set and BSC Supportability	Introduction of new technologies like GoldenGate to BSC as well as will require training of solution delivery and support resources. Also, new investments may be required in administration and monitoring tools and creation of operational processes.
Funding & Resources	Availability of funding and resources (external vendor and internal) for the development and deployment of the Provider BoR, Provider Data Quality and related activities in a timely manner will impact the delivery schedule and execution of roadmap initiatives.

Table of Contents

1. Executive Summary
2. Concept Definition and Model Framework
3. Environmental Analysis
4. Current State Architecture Models
5. Future State Architecture Models
6. Architecture Analysis
 - Gap Analysis
 - Impact Analysis
 - Technology Solutions Analysis
 - Risk Analysis
- 7. Roadmap**
8. Recommendation
9. Appendices

High-Level Roadmap (Level 1)



High-Level Roadmap (Level 2)

id	Initiative	2015				2016				2017			
		q1	q2	q3	q4	q1	q2	q3	q4	q1	q2	q3	q4
1	Provider Book of Record												
1.1	Develop and implement the data integration ELT framework (guidelines, best practices)												
1.2	Complete the implementation of Provider BOR addressing gaps including data loading, data model, error handling and recovery, intra-day batches, data masking, source version control and database consolidation		✓										
2	Provider Index (MDM)												
2.1	Implement the 6 provider entities using Infosphere MDM SE for match/merge and generating provider “golden record”		✓										
2.2	Implement MQ MFT based processes for batch data transfers between Provider Index and Provider BoR												
3	Provider Data Quality												
3.1	Implement the PDQ workflow automation tool for bulk updates to PIMS, for use cases including IDQ enabled address standardization and Enclarity feeds		✓										
3.2	Implement a provider data quality program that includes governance, auditing and monitoring and leverages IDQ												
4	Provider Data Staging												
4.1	Deploy data replication tool (GoldenGate) for PIMS replication and re-mediate gaps including database topology, operational reporting data source etc												
4.2	Re-mediate external data batch processes to use Sterling												
5	Other												
5.1	Implement the foundational data management services that apply to all BoR's including Netezza DR, data archiving, job orchestration, database activity monitoring and data balancing												
5.2	Expand blueprint to include Provider data consumption												
5.3	Initiate Rosetta retirement planning												

Roadmap Key Next Steps

	Initiative	Owner	Participants	Target	Key Requirements	Expected Outcomes
1	Implement a provider data quality program that includes governance, auditing and monitoring	• Provider Data Owner	• Infrastructure Services	Q3, 2015 – Q1, 2016, ongoing	• Data quality charter and sponsorship • Data quality metrics	• Defined data quality policies, procedures, roles, responsibilities and governance process • Informatica IDQ based scorecards for measuring and monitoring provider data quality including support model
2	Develop and implement the data integration ELT framework (guidelines, best practices)	• BoR Project Team, Data Liquidity team	• Enterprise Architecture	Q2, 2015	• ELT best practices	• Coding standards, naming conventions, guidelines and best practices, design patterns etc for ELT development and as guidance for code reviews; to ensure consistency and design for supportability
3	Implement the rules for mastering the 6 entities in Provider Index and generating golden records	• BoR Project Team, Data Liquidity team	• Enterprise Architecture • Application Services • Infrastructure Services	Q2-Q4, 2015	• Business requirements for mastering 6 provider entities in Provider Index	• Provider Index application with 6 provider entities deployed to Production, including data integration with Provider BoR • Business MDM data stewards actively managing tasks and administering the MDM tool
4	Complete the Provider BoR implementation, including re-mediation of identified gaps	• BoR Project Team, Data Liquidity team	• Infrastructure Services • Application Services • Enterprise Architecture	Q2-Q4, 2015	• Business requirements for first release of Provider BoR • Enterprise Information Model (EIM) for Provider	• Provider BoR deployed to Production with initial population and on-going incremental loads • Robust error handling and recovery processes implemented to ensure resiliency of ELT load process • Data masking process operational to ensure de-identification of sensitive data in non-production environments
5	Purchase recommended data replication tool (GoldenGate) and deploy for replicating PIMS data into PROPS	• Infrastructure Services	• IT Shared Services • Application Services	Q2-Q4, 2015	• Deploy data replication tool • Re-mediate gaps: database topology, operational reporting, B2B file transfer	• Data Replication tool deployed for PIMS to PROPS replication • Business Objects universes use PROPS as data source instead of PIMS • B2B file transfers via Sterling File Gateway
6	Implement the Provider Data Quality workflow tool, including Enclarity integration and address standardization	• PDQ Project Team	• Infrastructure Services • Application Services • Enterprise Architecture	Q2, 2015 – Q1, 2016	• Business requirements for Provider Data Quality application • Integration with Enclarity • Address standardization	• PDQ application deployed to Production, supporting 3 identified use cases Integration with Enclarity data and updates to PIMS via PDQ • Address standardization in PIMS leveraging Informatica IDQ
7	Implement other data mgmt services including job orchestration, data archiving, database activity monitoring, data balancing, disaster recovery	• Data Liquidity team, Infrastructure Services	• IT Shared Services • Application Services	Q3, 2015 – Q1, 2016	• Netezza job orchestration • data archiving for BoR's • DR, Database activity monitoring for Netezza • Data balancing for BoR's	• Netezza DR and database activity monitoring operational • Data archiving strategy for BoR defined and implemented • Netezza job orchestration tool deployed and used for BoR development • Infogix data balancing implemented between data sources and BoR for audit and assurance
8	Expand Provider Data Architecture Blueprint (Part-2: Data Consumption)	• Enterprise Architecture	• Application Services • IT Shared Services	Q3, 2015	• Data consumption by Portals, Provider Directories, External extracts, Internal applications	• Identification of technology solutions for future state • Roadmap for addressing gaps identified in the blueprint
9	Initiate Rosetta retirement planning	• Actuarial & Analytics	• Enterprise Architecture	Q4, 2015	• Provider BoR	• Approach and plan for retiring use of Rosetta as a source of Provider information

Table of Contents

1. Executive Summary
2. Concept Definition and Model Framework
3. Environmental Analysis
4. Current State Architecture Models
5. Future State Architecture Models
6. Architecture Analysis
 - Gap Analysis
 - Impact Analysis
 - Technology Solutions Analysis
 - Risk Analysis
7. Roadmap
- 8. Recommendation**
9. Appendices

Future State Recommendation

- **Provider BoR** – Provider BoR deployed on the Netezza platform based on Provider Enterprise Information Model (EIM)
- **Provider Index (MDM)** – Provider Index deployed with “golden record” for 6 provider entities using Infosphere MDM
- **Provider Data Quality** – Provider Data Quality (PDQ) custom tool deployed for bulk data updates to PIMS
- **Data Integration** – GoldenGate data replication tool deployed
- **Technology Standards:**

Core Capability	Sub Core Capability	Component	Current Technologies	Future Technology Solution Standards
Big Data Management	Master Data Mgmt	Master Data Management	IBM Infosphere MDM SE	
Big Data Management	Books of Record	BoR Data Platform	IBM Netezza	
Big Data Management	Data Staging	Staging Data Platform	Oracle RDBMS	
Big Data Management	Data Quality Mgmt	Data Balancing	Gap	Infogix Assure Ex
Big Data Management	Data Quality Mgmt	Data Profiling	Informatica IDQ / Informatica Analyst	
Big Data Management	Data Quality Mgmt	Data Cleansing	Gap	Informatica IDQ/ Address Validator Ex
Big Data Management	Data Quality Mgmt	Workflow tool bulk update (PDQ)	Gap	Custom PDQ Tool new
Big Data Management	Other Services	Data Masking	Gap	IBM Optim Data Privacy Ex
Integration & Interoperability	Data Integration	Data Replication	Oracle Data Guard	Oracle GoldenGate new
Integration & Interoperability	Data Integration	ETL	Informatica PowerCenter	
Integration & Interoperability	Data Integration	ELT	Netezza scripts	
Integration & Interoperability	Data Integration	Code Version Control	Unfuddle	Apache Subversion Ex
Integration & Interoperability	Common Intg. Capabilities	ELT Job Orchestration		Custom Netezza Orchestration Tool new
Integration & Interoperability	Common Intg. Capabilities	Managed File Transfer (external)	SFTP	IBM Sterling File Gateway Ex
Integration & Interoperability	Common Intg. Capabilities	Managed File Transfer (internal)	FTP	IBM MQ MFT Ex
Integration & Interoperability	Common Intg. Capabilities	Scheduling	Cisco Tidal	
Integration & Interoperability	Infrastructure Services	Data Archiving (ILM)	Gap	IBM Optim Archive Ex
Integration & Interoperability	Infrastructure Services	BoR Disaster Recovery	IBM Netezza Replication Services (PTS)	
Integration & Interoperability	Infrastructure Services	Database Backup	Symantec NetBackup	
Foundational Services	Security Services	Access Management	Microsoft Active Directory	
Foundational Services	Security Services	Database Activity Monitoring	Gap	Imperva SecureSphere Ex
Analytics	Analytics	Analytics	SAP Business Objects	SAP Business Objects, SAS, Tableau Ex

Recommendation

Roadmap – Key Next Steps:

1. Implement a provider data quality program that includes governance, auditing and monitoring and leverages Informatica IDQ tool – **Owner:** Provider Data Owner & Stewards
2. Develop and implement the data integration ELT framework (guidelines, best practices) – **Owner:** Provider BoR project team, Data Liquidity team
3. Implement the rules for mastering the 6 entities in Provider Index and generating golden records – **Owner:** Provider BoR project team, Data Liquidity team
4. Complete the Provider BoR implementation, including re-mediation of identified gaps – **Owner:** Provider BoR project team, Data Liquidity team
5. Purchase and deploy recommended data replication tool (GoldenGate) – **Owner:** Infrastructure Services
6. Implement the Provider Data Quality workflow tool, including Enclarity* integration and address standardization – **Owner:** PDQ project team
7. Implement other data mgmt services including job orchestration, data archiving, database activity monitoring, data balancing, disaster recovery – **Owner:** Data Liquidity team, Infrastructure Services
8. Expand Provider data architecture blueprint to include data consumption(Part 2 of 2) – **Owner:** Enterprise Architecture
9. Initiate Rosetta retirement planning – **Owner:** Actuarial & Analytics

ARB Decision (4/29/15):

1. Directionally approved the recommended future state architecture model (Levels 1-5)
2. Approved the recommended technology solution standards, with the exception of the data replication tool which will be further reviewed with a updated business case to be reviewed at a later date
3. Approved the roadmap key next steps with recommendation to delineate owner and supporting team roles

Ask ARB (5/20/15):

1. Decision on recommended technology solution standard for data replication tool (GoldenGate)

Table of Contents

1. Executive Summary
2. Concept Definition and Model Framework
3. Environmental Analysis
4. Current State Architecture Models
5. Future State Architecture Models
6. Architecture Analysis
 - Gap Analysis
 - Impact Analysis
 - Technology Solutions Analysis
 - Risk Analysis
7. Roadmap
8. Recommendation
- 9. Appendices**

Appendix 1: Glossary

Capability	Description
Descriptive Analytics	Descriptive Analytics <i>provides information about the current (and past) state of the business</i> and its environment. The descriptive realm uses regular reports for events that already happened, and ad hoc reports to help examine facts about what happened where, how often.
Diagnostic Analytics	Diagnostic analytics provides information on why certain business events occurred in the past. Diagnostic analytics is essential to improve business operations and processes by determining what is broken with the existing business processes
Predictive Analytics	Predictive Analytics provides information what could happen in the future. The predictive realm looks at real-time events and alerts to suggest actions, and it uses simulation models to suggest what can happen. The predictive analytics provides for these approaches: forecasting, which looks at the probability for these trends to continue; predictive modeling, which allows you to look into what-if situations.
Prescriptive Analytics	Prescriptive Analytics provides optimization techniques to examine how you can achieve the best outcome for a particular situation. It provides for optimization on how to mitigate or even avoid uncertain risks. The most common examples are optimization methods such as linear and non-linear programming, decision analysis methods such as decision trees.
Books of Record	Provider BOR is the authoritative data source for Provider data that is derived from integrating multiple data sources (internal and external). The data model is based on the Enterprise Information Model for Provider subject area and technology components are aligned with the Data Liquidity blueprint.
Data Marts	Data Marts provide persistence for data transformed from Books of Record or other data stores into de-normalized, physical views and data formats to support specific reporting requirements, user facing applications, and external partner/vendor extracts.
Unstructured Data	Unstructured data provide for persistence, management of unstructured and semi-structured data, distributed processing of unstructured data via Map Reduce, development frameworks such as Pig and Hive, management and integration frameworks
Data Staging	Data Staging provides persistence for raw data from disparate sources both internal and external, for leverage in supporting operational reporting, real-time data services, data quality monitoring, auditing etc. The data formats reflect the source data (as-is) and they are typically not transformed or integrated. The data is refreshed at a low latency, even near real-time in some cases and the data is highly volatile (data is overlaid in-place as it changes on source system).
Business Process Management (BPM)	Business process management is about managing business processes for improving business performance outcomes and operational agility by linking people, information flows, systems and other assets to create and deliver value to customers and constituents.
B2B Integration	B2B Integration is comprised of software integration middleware technology capabilities that are used to facilitate integration and interoperability of internal BSC application and processes with external entities, such trading partners and SaaS vendors.
Application Integration	Application integration is comprised of software integration middleware technology capabilities that are used to facilitate integration and interoperability of internal BSC applications, SOA services, and processes
Data Integration	Data integration is comprised of software integration middleware technology capabilities that are used to facilitate data integration across BSC data sources
Common Integration Capabilities	Common integration capabilities is comprised of software integration middleware technology capabilities that are foundational and can be used across many different types of integrations.

Appendix 1: Glossary – contd.

Capability	Description
ESB Service Gateway	Single point of entry for service consumers (service virtualization). Mediates the service request for routing, protocol conversion, transformation, lightweight stateless orchestration, auditing, logging, SLA management.
ESB Complex Integration Flows	Multi-step integration flows involving multiple systems and service providers. Integration flows requiring maintaining session state between service invocations. E.g. composite services that encapsulate channel specific integration that requires applying channel specific rules (e.g. EDI)
Application Server	Application platform that acts as host (or container) for web services, web components, business logic, and other application components . Supports standard java programming models, such as Servlet, JSP, JSF, JMS, JAX RS, JAX WS, SCA, JPA, etc.
Object Cache	Object caching in the services layer to reduce database I/O, including in-memory, distributed object caching, elastic scalability, web session caching, and replication.
Service Registry & Repository	Repository to maintain services metadata, including service contracts, QoS & security policies, versioning, and governance artifacts. Supports the runtime binding and service virtualization needs of service gateway.
Data Virtualization	Provides for real-time, on-demand retrieval and manipulation of data while abstracting the technical details about the source, location or format of the data. Resolves differences between source and consumer formats and semantics, transforming as needed. Depending on the form it takes might include <i>data federation</i> which combines results sets from multiple source systems and provides a common logical data access point by publishing result sets as views or data services.
ETL/ELT	A batch process for extracting data from data sources, transforming and integrating the data into the designed structure and format, sometimes in complex, multi-step workflows, and loading the result into the target data stores. The intermediate steps can include data validation and data cleansing. With ETL, the intermediate, transformation processing takes place in either the data sources or the ETL servers. With ELT, the intermediate transformation processing takes place in the target data stores, usually highly performing data warehouses.
Data Replication	For a relational database source, the transaction log is continuously scanned and the log records are transported and applied to a target database, usually in real-time but possibly in micro-batches. This method preserves the logical integrity of the source database and guarantees transactional consistency between the source and target databases.
Streaming	Transferring data so that it can be processed as a steady and continuous stream. The data may be structured or unstructured. Usually continual analysis is performed on the data stream while it is in-memory, typically using time windows of data and special functions that operate on time windows.
Data Masking	Data Masking is also known as data obfuscation or data de-sensitization, the treatment of sensitive production data, such as personal identification or financial information, for safe and compliant use in non-production environments for development, testing or user training.
Data Quality Management	Business processes, practices and tools for measuring, monitoring and improving data quality as set by data governance policy and business rules for data quality.
Information Lifecycle Management	Technical processes, practices and systems for managing data throughout its lifecycle: creation, on-boarding, development, testing ,production usage, archival and retirement.
Reference Data Management	The tools and processes for managing data that is used to categorize other data within enterprise applications and databases. While reference data is commonly stored in the form of code tables, this capability includes its centralized management, stewardship, and distribution. (Source: IBM)
Data Classification	Data Classification is the categorization of data to enable/help organization to effectively answer following questions: What data types are available? Where are certain data located? What access levels are implemented? What protection level is implemented and does it adhere to compliance regulations?

Appendix 1: Glossary – contd.

Capability	Description
Identity and Access Management	Identity and access management (IAM) is the security discipline that enables the right individuals to access the right resources at the right times for the right reasons.
Cloud Security	Cloud computing security or, more simply, cloud security is an evolving sub-domain of computer security, network security, and, more broadly, information security. It refers to a broad set of policies, technologies, and controls deployed to protect data, applications, and the associated infrastructure of cloud computing.
Data Security	Data security refers to protective digital privacy measures that are applied to prevent unauthorized access to computers, databases and websites. Data security also protects data from corruption. Data security is the main priority for organizations of every size and genre.
SDLC Security	These are the security controls applied to the application to ensure the proper protections to confidentiality, integrity, and availability as the development team moves through the phases of the SDLC
Risk Assessment	A risk assessment helps the organization ensure it is compliant with HIPAA's administrative, physical, and technical safeguards. A risk assessment also helps reveal areas where the organization's protected health information (PHI) could be at risk
Activity Monitoring	Database activity monitoring is a database security technology for monitoring and analyzing database activity that operates independently of the database management system (DBMS) and does not rely on any form of native (DBMS-resident) auditing or native logs such as trace or transaction logs. DAM is typically performed continuously and in real-time.
DR & BC	Disaster recovery (DR) involves a set of policies and procedures to enable the recovery or continuation of vital technology infrastructure and systems following a natural or human-induced disaster. Business continuity (BC) encompasses a loosely defined set of planning, preparatory and related activities which are intended to ensure that an organization's critical business functions will either continue to operate despite serious incidents or disasters that might otherwise have interrupted them, or will be recovered to an operational state within a reasonably short period
Endpoint Devices	Endpoint devices can include desktop or laptop computers, as well as portable devices like tablets and smart phones. Other types of hardware installations, like retail kiosks, also may fall under the category of endpoint devices.

Appendix 1: Glossary – contd. (MDM terms)

Capability	Description
Entity	An entity is the logical relationship between two or more records. Entities are represented in the MDM engine as records that share an Enterprise ID. There can be an unlimited number of records in an entity.
Enterprise ID	The ID used to represent an entity (same object across multiple sources or within the same source). It is assigned by the engine and can be shared by more than one member record. This identifier is displayed in MDM user applications. It is initially assigned by the MDM engine, but can be changed by users of client applications like Inspector
Member	A member is defined as a set of demographic information that represents one individual (for example, a person), a group of individuals (organization), or object (for example, a car part). A member object is the representation of what a single source system asserts to be true about an individual or thing.
Relationships	A relationship in the MDM engine is a type of association that can exist between two entities (they can be the same or different entity types). For example, a person can manage another person, or an organization can legally own another organization.
Attributes	Information that describes the member object. For example, name, date of birth, address, Social Security or other identifier number, part number, company location and business type, loan or account numbers.
Entity Types	An entity type provides a distinction between the way members are viewed and linked within the MDM operational server. Each entity type has a specific algorithm configuration.
MDM Engine	The MDM engine contains the logic that is at the heart of a virtual MDM implementation. Its configuration is a collection of data attributes and a custom algorithm specific to the organizational domain. The algorithms and rule definitions in your configuration are used by the engine to form entities and relationships between records.
Linkages	Two or more records that are grouped by a common Enterprise ID. A linkage set can have an unlimited number of records. There are two types of linkages. 1) Auto-link, records that are automatically linked and assigned common Enterprise IDs by the MDM engine. 2) Manual-link, records that are assigned a common Enterprise ID by user action

Appendix 2: Provider Golden Record

The provider Golden Record is the record that represents the best set of data attributes across all provider sources, as determined by source priority and trust rules configured in the Provider Index MDM engine.

Illustrative Example:

MDM Attributes:	FULL NAME	SSN	ADDRESS				PHONE
Source 1	Kate Lamb	SSN 555-55-1234					
Source 2	K. Jones		1000 Main St.	San Francisco	CA	94105-1524	(415) 123-4567
Source 3	Katherine J. Jones	SSN 555-55-1234		San Francisco	CA	94105-2539	
Source 4	Catherine Lamb						(415) 123-4567
	⋮	⋮	⋮	⋮	⋮	⋮	
GOLDEN RECORD	Katherine J. Jones	SSN 555-55-1234	1000 Main St.	San Francisco	CA	94105-1524	(415) 123-4567

Creating Golden Record:

1. Using probabilistic algorithm on the attributes (e.g., Full Name, SSN, Address) of an entity, identify the “matching” records from multiple data sources. Assign an Enterprise Identifier (EID) to the “matching” records.
2. Create a composite view (golden record). A composite view of an entity is the combination of the “best” set of attribute values for the “matching” records with the same EID. The “best” attribute set is determined by the business specified source trust and priority rules.

Appendix 3: Provider Business Glossary and EIM

What/When

The Provider Business Glossary and Enterprise Information Model (EIM) for all data topics were developed during working sessions starting January 6, 2015 and completed and approved by March 17, 2015.

Why

Among the many reasons to build the Business Glossary and the EIM are:

- Engenders trustworthy, timely & relevant information in support of informed business decisions
- Enable business users to access technical information using business terms
- Business and IT have a common vocabulary to communicate more effectively
- Accountability for management of company data assets
- Develop a continuously improving process

Who

Book of Record (BoR)

Domain

Provider

Network
Credentialing
Contract (Agreement)
Provider Review
Provider Rating
Provider Demographics
Provider Relations
Provider Services

EDG Steering Committee member

Marcus Thygeson

Domain Owner

Kenny Deng

Data Stewards

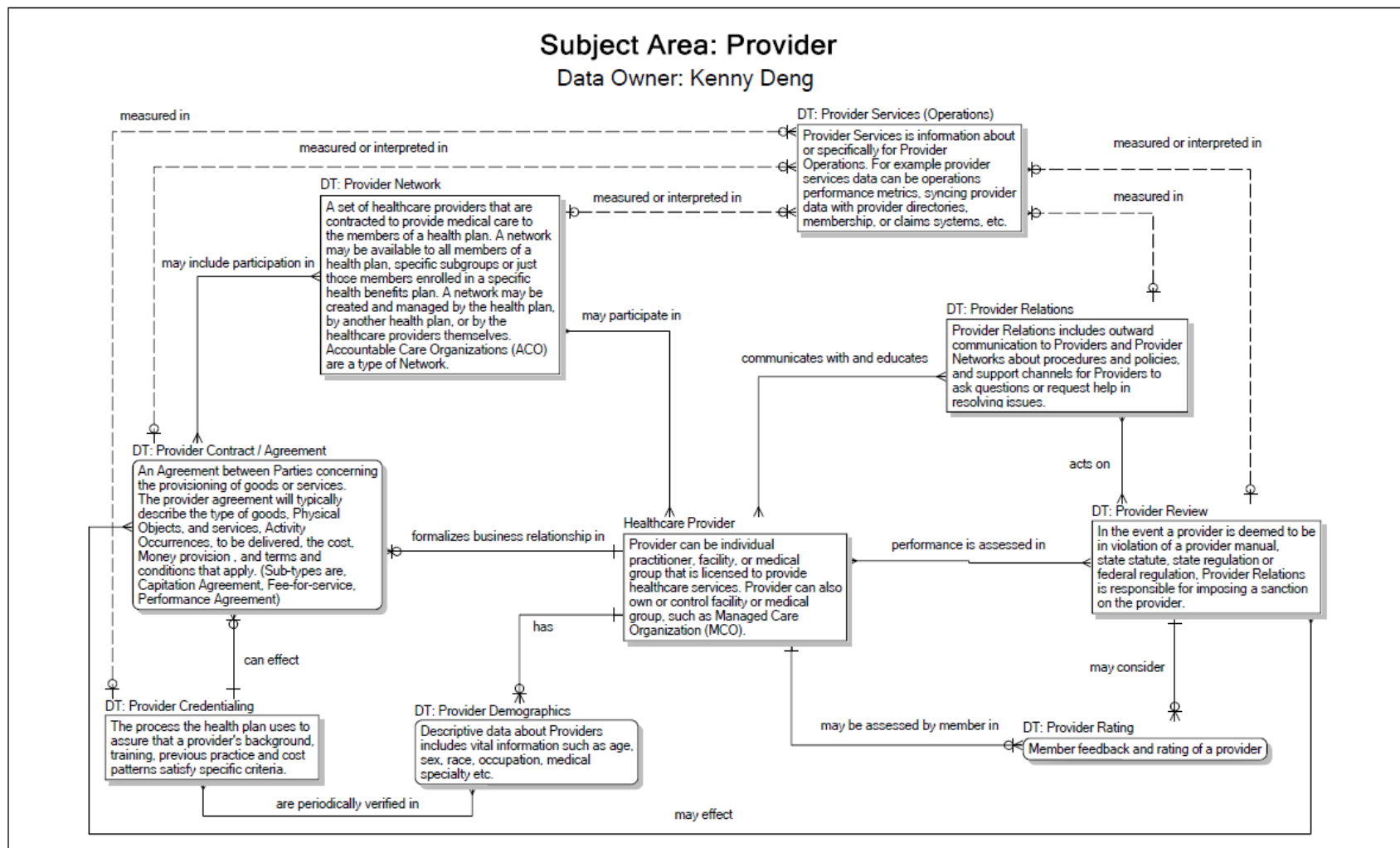
Armine Papouchian/Hugo Florez
Rosemary/Richard Rice
Armine Papouchian/Tina Shabanian
Hugo Florez/Jamie Ostroff
Amy Lung
Edie Parker/Jason Hagen
Wendy Bavan/Angela Schoenfeld
Edie Parker/Jason Hagen

Outcome

Deliverables completed and approved:

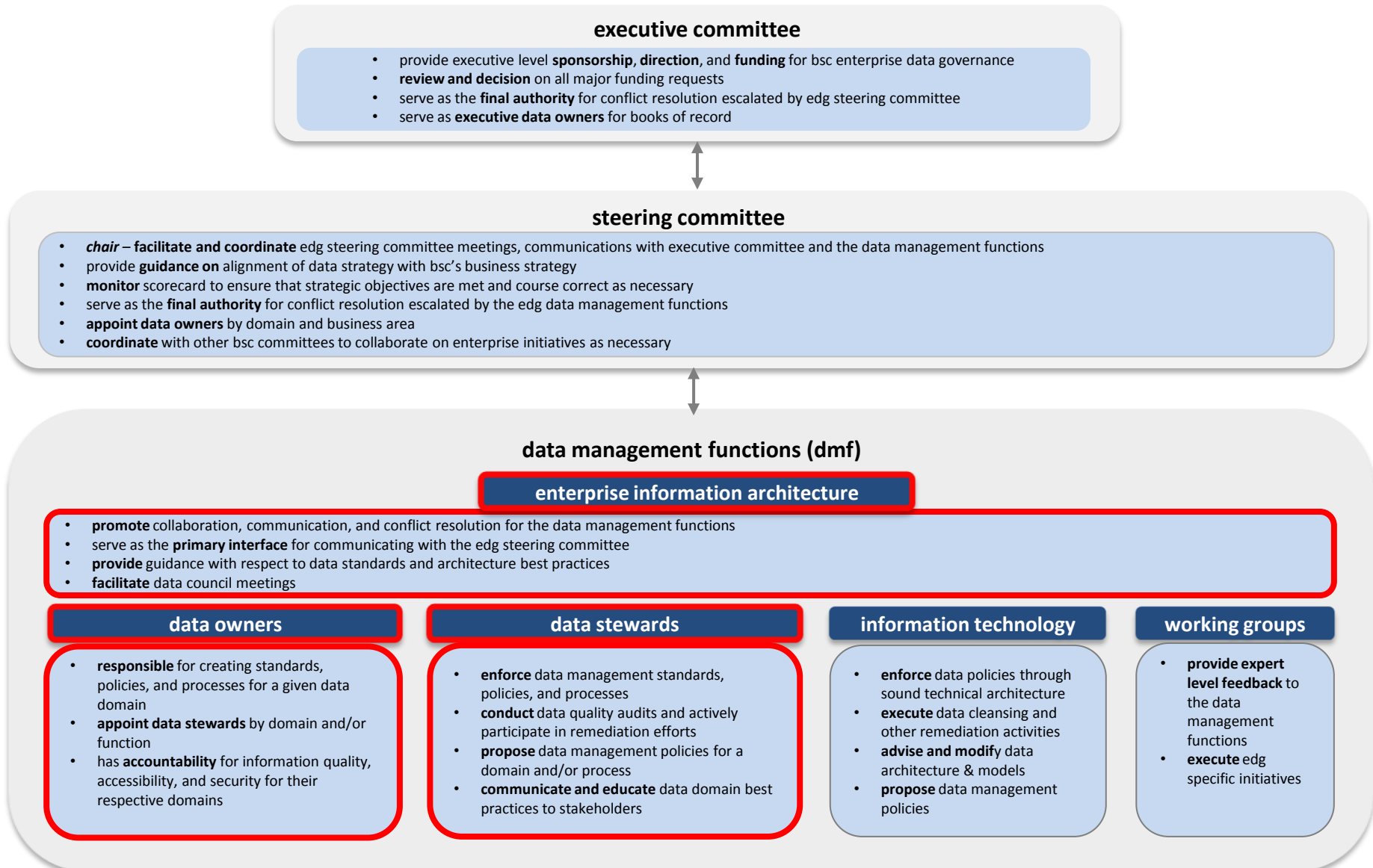
- Provider Subject Area Business Glossary
- Enterprise Information Model (EIM) aligned with the Business Glossary

Appendix 4: Provider Data Topics



Links: [Approved Provider Business Glossary](#) [Approved Provider EIM Data Models](#)

Appendix 5: Enterprise Data Governance Model



Appendix 6: Data Replication Tool Comparison

Assessment Category	Assessment Factor	Oracle GoldenGate	IBM Infosphere Data Replication	Informatica Data Replication
Functionality	Support for foundational functionality (unobtrusive log capture, real-time processing, heterogeneous databases)	Supports all foundational functionality for data replication use cases in a single product, and manages all scenarios in a unified fashion , including unobtrusive log capture, CDC, full HA/DR, real-time replication, heterogeneous databases	Supports all foundational functionality for data replication use cases in a single product, including unobtrusive log capture, CDC, full HA/DR, real-time replication, heterogeneous databases	Supports all foundational functionality for data replication use cases in a single product, including unobtrusive log capture, CDC, full HA/DR, real-time replication, heterogeneous databases
	Support for differentiating functionality (guaranteed delivery, transactional integrity, light transformations, bidirectional, encryption)	Supports guaranteed delivery, transactional integrity, transformations, encryption, high availability, recoverability, lowest impact on source, event based processing, adapters for Java and flat files, auto-restart	Supports guaranteed delivery, transactional integrity, transformations high availability, recoverability, low impact on source, event based processing. Does not support encryption during transfer.	Supports guaranteed delivery, transactional integrity, transformations high availability, encryption, recoverability, low impact on source. Does not support target side filtering, SQL Server database, UDT columns, conflict resolution
Technology Maturity	Technology maturity for mission critical use (scalability, performance, security, reliability)	Mature and best-in-class technology. Proven scalability, reliability, and performance. Vendor placed in the Gartner magic quadrant leader category	Mature technology. Proven scalability, reliability, and performance. Placed in the Gartner magic quadrant leader category	Mature and best-in-class technology. . Placed in the Gartner magic quadrant leader category
	Technology adoption in the industry. Ability to support future BSC needs	More than 4000 customers, including the majority of Fortune 500	Significant customer base. Until recently, CDC and Replication Server were separate products.	Limited customer base. Technology acquired and not integrated into core Informatica product; overlaps with PowerExchange CDC product
Agility and Usability	Usability of tools and support for agile development and deployment	High usability through use of GUI console for design, deployment and monitoring. Agility in terms of development for replication jobs	High usability through use of GUI console for design, deployment and monitoring. Agility in terms of development for replication jobs	High usability through use of GUI console for design, deployment and monitoring. Agility in terms of development for replication jobs
Skill-set	Availability of skill-set – commodity skill-set versus specialized vendor-specific skills	GoldenGate is fully integrated with Oracle database, including end-to-end monitoring, integration with RMAN, support for RAC, ASM, TDE; no programming required for creating replication jobs	Will need training for database team to pick up additional vendor-specific skills for deployment, management and support and integration with Oracle database products	Will need training for database team to pick up additional vendor-specific skills for deployment, management and support and integration with Oracle database products
Vendor Supportability	Vendor and product (roadmap) viability	Established vendor. Low vendor risk.	Established vendor. Low vendor risk.	Vendor was recently acquired by private equity. Significant changes to product architecture at current time and future direction not clear
	Vendor support & SLA	High (traditional software support model)	High (traditional software support model)	High (traditional software support model)
Cost	Estimated software acquisition costs	First year (acquisition): \$10,500/core*	First year (acquisition): \$4,500/core**	vendor did not provide
	Implementation and ongoing maintenance (complexity, maturity, skills)	Low to medium. GUI console for designing replication mappings; no programming required for creating replication jobs	Low to medium. GUI console for designing replication mappings; no programming required for creating replication jobs	Low to medium. GUI console for designing replication mappings; no programming required for creating replication jobs
Current BSC Supportability	Technology support structure and integration with existing tools. In-house skills	Existing infra support structure. Integration with existing development and management & monitoring tools In-house skill-set	Requires new infra support structure. New tools and integrations for management & monitoring.	Requires new infra support structure. New tools and integrations for management & monitoring.
	Synergies with existing investments (transactional databases, data integration tools)	Strong synergies with existing transactional Oracle databases. Performance advantage as GoldenGate can read logs directly from memory.	No synergies with existing transactional Oracle databases or data integration tools.	No synergies with existing transactional Oracle databases. Also, IDR is a standalone product with no synergy with PowerCenter

Appendix 6: HPXR Performance Issues

HPXR Performance Dashboard for week ending- 05/01

Average batch run time (excluding tidal delays) for the week ending 05/01 is around 09 hour.

Batches of 04/28 and 04/29 were not executed completely and have therefore not been considered in this Dashboard.

Workflows wf_DIM_ELG_PCP_CLM_BLG_P4A and wf_FACT_APP_AUTH_CUST_P5F have been the major long running workflows leading to delays in batch completions.

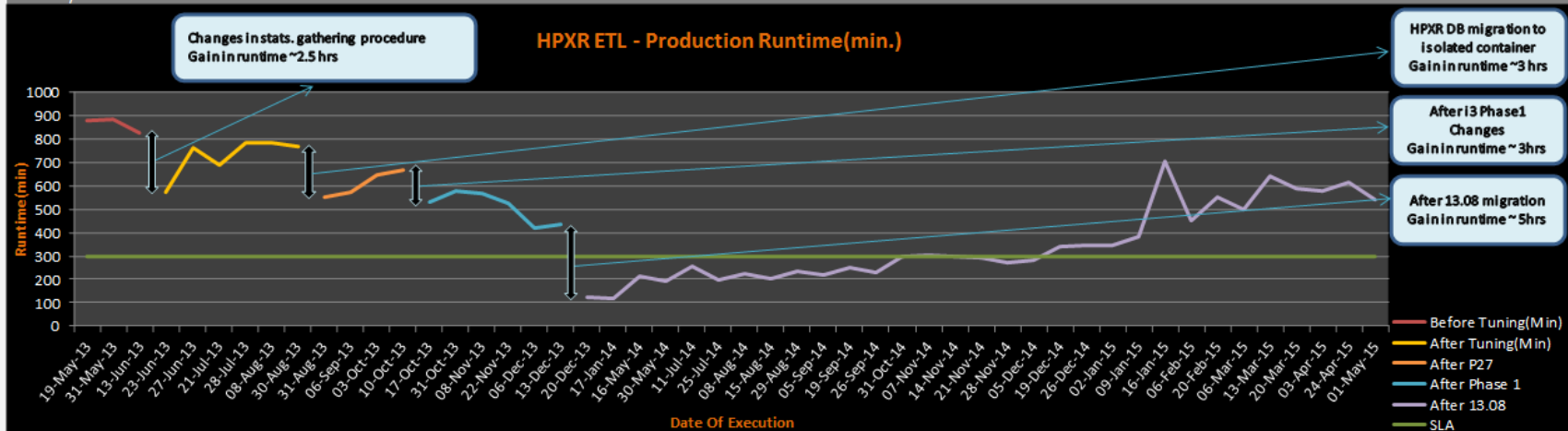
Sessions s_m_CLAIM_DIAGNOSIS, s_m_CLAIM_LINE_PRCNG_BRIDGE and s_m_CUSTSERV_FACT of these workflows have been the major longer running sessions delaying workflows' completion. Longer time taken by Transformation threads caused delays in various sessions' completions in general.

On 04/28, session s_m_CUSTSERV_FACT ran for around 4 hour 30 minute, with cache creation for Lookup transformation LKP_CLAIM_DIM taking an exceptional time of over 2 hour.

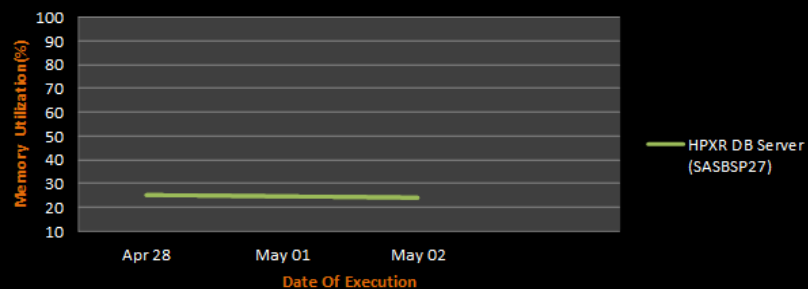
CPU and Memory utilization metrics have been collected from OEM and these metrics are not available for INFA App. Server (SASBSP24) for the week reported in this Dashboard.

CPU and Memory utilization metrics for INFA App. Server (SASBSP24) are available in OEM only till 12/28/2014 5 PM PST.

Memory utilization metrics for Facets DB Servers are not available in OEM.



Memory Utilization- Production HPXR ETL



CPU Utilization- Production HPXR ETL

