Ministério da Justiça

UnB

Centro de Apoio ao Desenvolvimento Tecnológico

latitude
Laboratório de tecnologias da tomada de decisão

Termo de Cooperação/Projeto:

Acordo de Cooperação Técnica FUB/CDT e MJ/SE Registro de Identidade Civil – Replanejamento e Novo Projeto Piloto

Documento:

RT Preliminar dos Inventários de Serviços de Identificação

Data de Emissão:

20/09/2014

Elaborado por:

Universidade de Brasília – UnB
Centro de Apoio ao Desenvolvimento Tecnológico – CDT
Laboratório de Tecnologias da Tomada de Decisão – LATITUDE.UnB

MINISTÉRIO DA JUSTIÇA

José Eduardo Cardozo
Ministro

Marivaldo de Castro Pereira
Secretário Executivo

Helvio Pereira Peixoto
Coordenador Suplente do Comitê Gestor do SINRIC

UNIVERSIDADE DE BRASÍLIA

Ivan Marques Toledo Camargo
Reitor

Paulo Anselmo Ziani Suarez
Diretor do Centro de Apoio ao Desenvolvimento Tecnológico – CDT

Rafael Timóteo de Sousa Júnior
Coordenador do Laboratório de Tecnologias da Tomada de Decisão – LATITUDE

EQUIPE TÉCNICA

Ana Maria da Consolação Gomes Lindgren
Alexandre Cardoso de Barros
Andréa Benoliel de Lima
Beatriz Merguiso Garrido
Celso Pereira Salgado
Delluiz Simões de Brito
Domingos Soares dos Santos
Elaine Fabiano Tocantins
Felipe Bragança Itaborahy
Fernando Saliba
Fernando Teodoro Filho
Guilherme Braz Carneiro
Jhon Kennedy Férrer Lima
José Alberto Sousa Torres
Joaquim de Oliveira Machado
Marcelo Martins Villar
Narumi Pereira Lima
Paulo Cesar Vieira dos Santos
Raphael Fernandes de Magalhães Pimenta
Rodrigo Borges Nogueira
Rodrigo Gurgel Fernandes Távora
Sara Lais Rahal Lenharo

EQUIPE TÉCNICA

Flávio Elias Gomes de Deus (Pesquisador Sênior)
William Ferreira Giozza (Pesquisador Sênior)
Ademir Agustinho de Rezende Lourenço
Adriana Nunes Pinheiro
Alessandro Zimmer
Alysson Fernandes de Chantal
Amanda Almeida Paiva
Andréia Campos Santana
Andreia Guedes Oliveira
Antônio Claudio Pimenta Ribeiro
Carolinne Januária de Souza Martins
Caio Rondon Botelo de Carvalho
Cristiane Faiad de Moura
Daniela Carina Pena Pascual
Danielle Ramos da Silva
Eduarda Simões Veloso Freire
Fábio Lúcio Lopes Mendonça
Fábio Mesquita Buiati
Glaidson Menegazzo Verzeletti
João Luiz Xavier M. de Negreiros
Johnatan Santos de Oliveira
José Carneiro da Cunha Oliveira Neto
José Elenilson Cruz
Kelly Santos de Oliveira Bezerra
Luciano Pereira dos Anjos
Luciene Pereira de Cerqueira Kaipper
Luiz Antônio de Souto Evaristo
Luiz Claudio Ferreira
Marcos Vinicius Vieira da Silva
Marco Schaffer
Mirele Maria Cavalcante Rocha
Pedro Augusto Oliveira de Paula
Renata Elisa Medeiros Jordão
Roberto Mariano de Oliveira Soares
Sandro Augusto Pavlik Haddad
Sergio Luiz Teixeira Camargo
Soleni Guimarães Alves
Suzane Lais De Freitas
Valério Aymoré Martins
Vinicius de Morais Alves
Wladimir Rodrigues da Fonseca

HISTÓRICO DE REVISÕES

Data	Versão	Descrição
16/05/2014	0.1	Versão inicial
18/07/2014	0.2	Versão modificada por definição do Grupo de Trabalho
31/07/2014	0.3	Versão ampliada dos serviços com maiores detalhes de serviços específicos
20/08/2014	0.4	Versão ampliada com os principais projetos de uso pelo Ecossistema e CANRIC
30/08/2014	0.5	Versão ampliada com os principais elementos propostos
02/09/2014	0.6	Alteração do título e adição do conteúdo sobre gestão de identidade
20/09/2014	0.7	Versão Preliminar Final

latitude
Laboratório de tecnologias da tomada de decisão

Universidade de Brasília – UnB
Campus Universitário Darcy Ribeiro - FT – ENE – Latitude
CEP 70.910-900 – Brasília-DF
Tel.: +55 61 3107-5598 – Fax: +55 61 3107-5590

SUMÁRIO

1 INTRODUÇÃO

A Secretaria Executiva (SE/MJ), vinculada ao Ministério da Justiça (MJ), é responsável por viabilizar o desenvolvimento e a implantação do Registro de Identidade Civil, instituído pela Lei nº 9.454, de sete de abril de 1997, regulamentado pelo Decreto nº 7.166, de 5 de maio de 2010.

Atualmente, a República Federativa do Brasil conta com sistema de identificação de seus cidadãos amparado pela Lei nº 7.116, de 29 de agosto de 1983. Essa lei assegura validade nacional às Carteiras de Identidade, ou Cédulas de Identidade; confere também autonomia gerencial às Unidades Federativas no que concerne à expedição e controle dos números de registros gerais emitidos para cada documento. Essa condição de autonomia, ao contrário do que pode parecer, fragiliza o sistema de identificação, já que dá condições ao cidadão de requerer legalmente até 27 (vinte e sete) cédulas de identidades diferentes. Com essa facilidade legal, inúmeras possibilidades fraudulentas se apresentam de maneira silenciosa, pois, na grande maioria dos casos, os Institutos de Identificação das Unidades Federativas não dispõem de protocolos e aparato tecnológico para identificar as duplicações de registro vindas de outros estados, ou até mesmo do seu próprio arquivo datiloscópico. Consoante aos fatos, os Institutos de Identificação não trabalham interativamente para que haja trocas de informações de dados e geração de conhecimento para manuseio inteligente e seguro para individualização do cidadão em prol da sociedade.

Com foco na busca de soluções para tais problemas, o Projeto RIC prevê a administração central dos dados biográficos e biométricos dos cidadãos no Cadastro Nacional de Registro de Identificação Civil (CANRIC) e ABIS (do inglês *Automated Biometric Identification System*), respectivamente. A previsão desse novo modelo sustenta a não duplicação de registros e a consequente identificação unívoca dos cidadãos brasileiros natos e naturalizados. O Projeto RIC, portanto, visa aperfeiçoar o sistema de identificação e individualização do cidadão brasileiro nato e naturalizado com vistas a um perfeito funcionamento da gestão de dados da sociedade, os quais agregam valor à cidadania, à gestão administrativa, à simplificação do acesso aos serviços disponíveis ao cidadão e à segurança pública do país.

Nesse contexto, o termo de cooperação entre MJ/SE e FUB/CDT define um projeto que

objetiva identificar, mapear e desenvolver parte dos processos e da infraestrutura tecnológica necessária para viabilizar a implantação do número único de Registro de Identidade Civil – RIC no Brasil.

Resultante de um subconjunto das atividades previstas para inicialização da cooperação MJ/SE e FUB/CDT, o presente documento contempla uma primeira visão de infraestrutura que poderá ser posta em teste na forma de Prova de Conceito visando obter alguns indicativos de desempenho e modelos que podem ser utilizados.

2 ARQUITETURA ORIENTADA A SERVIÇOS

Atualmente há uma clara tendência das organizações em buscar tecnologias para instrumentalizar a sua gestão de processos por meio do uso de ferramentas que permitam entender e melhorar seus processos de negócio a partir dos princípios de uma administração científica (ROSEMANN, 2006). Esse processo resulta no aumento da procura pelos conceitos e práticas de gestão de processos, advindos de diversas fontes: abertura de capital, adaptabilidade dos serviços frente à demanda evolutiva e oscilante, organização dos serviços de negócios compatíveis com políticas internas e regulamentações externas, diversidade dos modelos de negócios da nova economia e possibilidade de integração da visão de negócios com a visão de tecnologia de informação.

Este movimento atual busca uma gestão de processos suportada por tecnologia. Nesta gestão, tecnologias e ferramentas são apresentadas para suportar cada tarefa individualmente e são indicadas como soluções de integração entre os vários sistemas de informação que suportam a gestão e a execução dos processos organizacionais.

2.1 Conceituação e História

O uso da Arquitetura Orientada a Serviços (SOA) vem crescendo ao longo dos anos. SOA surgiu como uma forma de integrar negócio e sistemas, facilitando a interação entre sistemas (primordialmente a comunicação interna entre eles) e promovendo principalmente a reusabilidade tecnológica.

Como o principal elemento tecnológico da arquitetura baseada em serviços, as plataformas baseadas em SOA permitem que um mesmo serviço possa ser utilizado por aplicações, sistemas e/ou módulos diferentes, visto que encapsula uma lógica de negócio

comum a elas. Assim, os sistemas baseados nessas plataformas provêm visões com conteúdo diferentes, mas que possuem o mesmo propósito e a mesma lógica de negócio, pois são visões do modelo organizacional com mesma lógica de negócio em um único serviço centralizado.

Thomas Erl (Erl, Service-Oriented Architecture: Concepts, Technology & Design, 2005) explica que, na medida em que estão centrados em torno de serviços, os modelos organizacionais baseados em SOA associam as funcionalidades tecnológicas diretamente aos objetivos de negócios, em um encadeamento de processos integrados.

SOA é um paradigma de projeto, e não um paradigma tecnológico, influenciado por diversas plataformas e inovações tecnológicas, como o BPM, EAI, AOP, *Web Services*, Orientação a Objetos, dentre outras. A arquitetura viabiliza a criação, padronização, administração e documentação de serviços ou funções únicas para um mesmo propósito, sem que necessariamente sejam reescritos cada vez que usadas por outras aplicações.

2.2 Benefícios Práticos das Arquiteturas Orientadas a Serviços

Entre os principais impactos tecnológicos que justificam e tornam viável o uso de SOA estão a reusabilidade, o baixo acoplamento, a interoperabilidade, a integração, a manutenibilidade, a agilidade, a qualidade das soluções, o monitoramento dos processos de negócio, a capacidade de tratar a heterogeneidade de fontes de informação e o desenvolvimento integrado em uma linguagem "universal" dos processos de negócio. Estas características são explicadas a seguir.

- Reusabilidade: em SOA, um dos focos da reusabilidade é a possibilidade da criação de coleções que abrigam diferentes serviços que podem ser reutilizados por um grande número de aplicações. Dentre os principais benefícios do reuso (OLIVEIRA, 2006) podem-se citar os seguintes.
    - Redução de custos e prazos de novos projetos por meio do aumento de produtividade proporcionado pelo reaproveitamento de componentes/serviços.
    - Padronização da arquitetura dos produtos desenvolvidos pela organização.
    - Oportunidade de reutilização de serviços legados por meio da exposição de suas interfaces como novos serviços.
    - Agilidade na manutenção das aplicações devido à padronização de sua

arquitetura.

- Aumento na qualidade dos produtos gerados por meio da reutilização de partes prontas e testadas.

- Baixo acoplamento: é a característica definida pela independência dos serviços e pela necessidade desses interagirem entre si, apenas por meio de interfaces bem definidas. É através dessa característica que o impacto de uma mudança no sistema é minimizado. Como a dependência entre os serviços é mínima e a mudança estaria concentrada em algum serviço, bastando apenas alterá-lo, não se faz necessário alteração em todos os serviços, tão pouco em todo o sistema. Plataformas SOA buscam um uso de uma arquitetura técnica permitindo esta comunicação entre serviços e sua alta disponibilidade através de seu canal de descoberta de serviços e mediação da comunicação, denominado de Barramento de Serviços Corporativos (ESB - *Enterprise Service Bus*).
- Interoperabilidade: é a coexistência e comunicação flexíveis entre sistemas independentemente de fabricantes ou tecnologias. Padrões de mercado denominados de WS-*, principalmente de protocolos de troca de informação e definição estruturais, são ponto comum da arquitetura (SOA). Porém, além dos padrões de mercado, a interoperabilidade deve ser garantida pela aplicação de políticas, padrões e outros critérios de projeto durante o ciclo de vida do serviço, principalmente durante a identificação de serviço.
- Integração: tem como principal objetivo a obtenção de sistemas que facilitem o acesso a dados e procedimentos sem qualquer barreira funcional. Em consequência, as aplicações resultantes podem corresponder a combinações de componentes de diferentes áreas tecnológicas. Este ponto se destaca entre os principais benefícios da implementação de uma estrutura de TI integrada, baseada em SOA.
- Manutenibilidade: a união das características de baixo acoplamento e reusabilidade refletem diretamente na redução do tempo e do custo de manutenção. O fato dos serviços serem independentes entre si, facilita na hora da manutenção corretiva ou evolutiva de um serviço específico, já que uma alteração qualquer será realizada apenas localmente no serviço.
- Agilidade: a agilidade organizacional é aumentada pela representação funcional do

negócio, pela abstração dos serviços correspondentes e pelo baixo acoplamento entre a lógica de negócios e a lógica da aplicação, a partir da utilização da camada de serviço (Erl, Service-Oriented Architecture: Concepts, Technology & Design, 2005). Existe ainda um fator de agilidade na manutenção das aplicações devido à padronização de sua arquitetura. Neste sentido, as plataformas de SOA devem oferecer recursos para inventariança dos seus ativos, que em geral é disponibilizada sobre a denominação de Inventário ou Registro de Governança.

- Monitoramento dos processos de negócio: plataformas SOA, em geral, provêm a gestão dos serviços e de suas utilizações (monitoramento) na geração de métricas para a avaliação do seu desempenho, que varia em função dos indicativos de negócios e da sua aderência às expectativas, pois a avaliação dos serviços expõe as oportunidades de aperfeiçoamento do modelo, completando um ciclo de alinhamento e interlocução que se autoalimenta. Este princípio é em geral disponibilizado nos módulos de Monitoramento das Atividades de Negócio (BAM - *Business Activity Monitoring*).
- Observação dos fluxos de negócio em linguagem padrão: de seus princípios tecnológicos as plataformas SOA buscam disponibilizar, em elementos tecnológicos da plataforma SOA conhecidos como Servidores de Processo de Negócio (BPS), a capacidade de execução de fluxos de processo de negócio. Essas plataformas de BPS são mais transparentes ainda quando permitem a automatização da transformação dos processos de negócios definidos por meio da notação BPMN (*Business Process Modeling Notation*) em fluxos de linguagem automatizadas BPEL (*Business Process Execution Language*).
- Acesso à informações de fontes heterogêneas por meio de uma interface única (os serviços): como a área de tecnologia da informação tem se deparado com várias plataformas e uma diversidade de sistemas heterogêneos, esse contexto geralmente ocasiona problemas de falta de integração entre os sistemas e, consequentemente, dificulta o compartilhamento das informações. Plataformas SOA integram o acesso à informação proporcionando economia e eficiência de recursos, racionalizando os processos de negócio e garantindo integridade às informações extraídas dos sistemas, abstraindo o acesso às fontes heterogêneas de dados por meio de camadas de serviços compostos.

2.3 Governança SOA

Além dos benefícios elencados acima, devem-se ainda considerar aspectos relativos à Governança SOA. De acordo com o Gartner (Hotle, 2010), o maior desafio para expansão de SOA em uma corporação, no curto prazo, é operar e gerenciar processos e estruturas organizacionais, e não o uso das tecnologias de suporte. Decisões de governança devem ser pensadas desde o início dos esforços empregados na construção de sistemas baseados em SOA. A sofisticação dos processos de governança aumenta com o crescimento das iniciativas SOA.

Diferentemente dos modelos de governança usual, empregados em muitos ambientes de TI, em SOA é de fundamental importância a identificação da autoridade capaz de tomadas de decisão, de modo que a governança não se restrinja a seguir regras, mas também medir o uso e o cumprimento para promover a adoção de tais regras, bem como de políticas, metodologias, responsabilidades e procedimentos. A autoridade, em geral, se responsabiliza e deve ser capaz, também, de participar das definições de processos de negócio suportados pelas técnicas SOA, além de projetar e gerenciar serviços, identificando níveis de serviço, direitos de acesso e responsáveis pelos serviços.

As decisões iniciais quando da implantação de um ambiente baseado em SOA são fundamentais para se atingir os benefícios oferecidos. Que serviços serão implementados primeiro? Como determinar quão reutilizável é um serviço? Quem ou que setor da organização será responsável financeiro pelo desenvolvimento e manutenção de um serviço?. A implantação de processos de governança SOA precisa ser adiada em decorrência dos riscos envolvidos ao se tentar uma profunda alteração na governança envolvendo áreas chave de uma organização. Estes riscos podem ser resultado de problemas como: os mecanismos de governança atuais são insuficientes para suportar mesmo a implantação de processos mínimos para a governança SOA; a implantação desses processos atravessa fronteiras na cadeia de decisão da qual faz parte o(s) patrocinador(es) do projeto SOA; políticas, regulações, legislação e outros fatores podem restringir mudanças na governança em determinadas organizações, especialmente no Governo.

Além das respostas a essas perguntas é necessário o entendimento de que a construção inicial de SOA pode ser mais cara e demandar mais tempo que em abordagens

tradicionais com as tecnologias existentes na corporação. Os benefícios, entretanto, desde que sob eficaz governança, são percebidos no médio e longo prazo, pelos quais a construção de novas aplicações e mudanças realizadas sobre a infraestrutura SOA existente tem grande potencial de significar menos custos.

Destaca-se que as plataformas que favorecem a adoção da Arquitetura Orientada a Serviços - SOA realizam a gestão dos serviços e de suas utilizações (monitoramento) pela disponibilização de módulos voltados ao Monitoramento das Atividades de Negócio - BAM (*Business Activity Monitoring*). Esses módulos atuam na geração de métricas para a avaliação do desempenho dessas plataformas, a qual varia em função dos indicativos de negócios e da sua aderência às expectativas. Isto permite que a avaliação do cumprimento e do desempenho dos serviços exponha as oportunidades de aperfeiçoamento do modelo, completando um ciclo de alinhamento e interlocução que se autoalimenta. Maiores detalhes do BAM podem ser analisados junto ao "RT de Soluções de *Message Queue* e Barramentos de Serviço Corporativos".

2.4 Requisitos dos Serviços em SOA

Técnica e metodologicamente, o que foi abordado até o momento impõe que SOA e suas plataformas atendam a requisitos mínimos essenciais, a saber.

- Serviços SOA possuem a capacidade de serem localizados: serviços devem ser bem projetados e seus contratos devem ser publicados e estarem visíveis para os possíveis clientes poderem acessá-los. Serviços podem ser publicados através de: registros de serviços, repositórios de metadados, subdiretórios, ou uma localização qualquer conhecida. Após serem publicados, os serviços devem ser notificados aos usuários potenciais (propaganda dos serviços). Eles devem ter clientes (ou consumidores) identificados e padrões de reuso identificados antes de serem criados ou expostos. Um cliente para qualquer dispositivo, usando qualquer sistema operacional, em qualquer linguagem de programação, pode acessar um serviço SOA a fim de criar um novo processo do negócio.
- Serviços SOA têm que ter durabilidade: os serviços devem existir ao longo do tempo. Os serviços podem ser alterados, mas o negócio ou tema do processo que o serviço implementa deve persistir. Neste sentido, os serviços de negócio para um processo de negócio pode sofrer alterações, mas o processo em si sempre deve existir.

- Serviços SOA devem atender a requisitos técnicos de composição: serviços devem ser projetados a fim de permitir que possam ser utilizados por outros serviços. Eles devem permitir seu acoplamento a fluxos de processos orquestrados, não devem manter estados (*stateless*) e devem ser atômicos (outras estruturas devem ser utilizadas para manutenção de estado e de contexto). Novas funcionalidades ou novos padrões podem ser adicionados aos serviços sem quebrar o contrato com os consumidores atuais do serviço, mantendo interoperabilidade e as funcionalidades existentes anteriormente.
- Deve haver uma meta de atingimento de maturidade corporativa em SOA: é muito importante que uma arquitetura bem modelada orientada a serviços produza soluções para processos de negócios sem haver uma característica dominante perante a infraestrutura utilizada, pois a mesma solução poderá (e deve) ser utilizada para outros tipos de aplicações, gerando assim a agilidade nos processos. Neste sentido SOA, é um conceito de arquitetura corporativo, o qual permite criar, padronizar, documentar serviços genéricos, únicos e interoperáveis, para que possam de maneira fácil ser reutilizados por diversas aplicações diferentes, sem a necessidade de ser desenvolvido novamente, tornando o processo de desenvolvimento mais ágil.
- Arquiteturas SOA devem prever e prover soluções de integração: serviços em arquitetura SOA devem ser planejados permitindo o seu compartilhamento e a sua reutilização em ambientes distribuídos. O resultado desse planejamento, o qual alia tecnologia e negócio, é um conjunto de serviços interligados que perpassam a transferência de dados e a coordenação de atividades. Neste sentido, aplicações SOA grandes e complexas devem ser evitadas e substituídas por um conjunto de aplicações pequenas e simples, ou seja, uma aplicação passa a ser fisicamente composta por vários e pequenos módulos especializados, distribuídos, acessados remotamente, interoperáveis e reutilizáveis de *software* que são unidos graças a padronizações adotadas, podendo ainda ser rapidamente recomposta para o processo desejado (Erl, Service Oriented Architecture - SOA: Principles Of Service Design, 2007).
- As bases dos serviços SOA devem ser analisadas visando um alto reuso: serviços devem ser implementados com reuso claro em processos de negócio. Deve-se estar

ciente de que o reuso de um serviço pode levar à utilização do serviço por um novo cliente ou consumidor não determinado previamente. Deve-se considerar a infraestrutura (tamanho apropriado de *hardware* e de largura da banda de rede) a fim de garantir desempenho.

- Serviços SOA devem atender a padrões de interoperabilidade: a área de tecnologia da informação tem se deparado com várias plataformas e com uma diversidade de sistemas heterogêneos, o que geralmente ocasiona problemas de falta de integração entre os sistemas e, consequentemente, dificulta o compartilhamento das informações. A integração é um processo vital para manter a competitividade, pois proporciona economia e eficiência de recursos, racionaliza os processos de negócio e garante integridade às informações extraídas dos sistemas. A transformação que SOA provoca nos processos empresariais afeta todas as camadas da arquitetura corporativa, desde o direcionamento estratégico até a plataforma computacional e rede física na qual a empresa opera.
- SOA deve ser sustentado por altos padrões de política e governança: a Governança SOA é um processo contínuo de alinhamento dos objetivos estratégicos, novas oportunidades táticas e uso de experiência adquirida na implantação de iniciativas SOA. Essa governança é essencialmente diferente daquela utilizada em TI, porque necessita do envolvimento de gestão de processos e de pessoas ligadas ao negócio da organização (Schepers, Iacob, & Van Eck, 2008). A governança em SOA é essencial para a implantação de SOA, pois quanto maior o universo SOA, maior a necessidade de governança e mais complexos devem ser os papéis e mecanismos relacionados. A Governança SOA leva tempo para ser projetada e instalada e não é fácil de ser garantida, mas sem ela todo projeto estará em risco (Malinvero, 2006). Surge então a necessidade de altos padrões de administração de serviços, por meio de uma função responsável por desenvolver e administrar de forma centralizada as estratégias, políticas, procedimentos e práticas para o processo de gerência dos serviços de informações, incluindo planos para sua definição, padronização, organização, proteção, localização, divulgação e utilização.

O resultado da aplicação de SOA é um alto desempenho, agilidade e a diminuição de custos para as organizações que utilizam essa abordagem, alçando-as a um nível superior

na adequação do alinhamento entre o negócio e a Tecnologia da Informação. Porém, mesmo com toda conceituação e os elementos tecnológicos existentes, restam porém algumas questões que estão pendentes no entendimento dos gestores e interessados nas técnicas e tecnologias envolvidas, a saber.

- O que ainda falta para SOA provar seu valor? SOA terá que se tornar um modelo tecnologicamente automatizado de integração, de modelagem de processos de negócio e de governança em cada área de negócio.
- Por que "arriscar" em SOA como tecnologia em maturação? SOA pode ser necessária nestes tempos de alta competitividade e de necessidades de rápidas mudanças nos processos organizacionais. O orçamento de TI deve ser direcionado a novas implementações, em vez de reinventar, reintegrar e reconstruir os mesmos serviços, mesmo que a energia a ser gasta em SOA ainda seja muito alta devido a aspectos de maturidade tecnológica.

Para suporte à viabilidade econômica, quanto a aplicação de SOA pode representar em Retorno do Investimento (ROI)? SOA e seus fundamentos de Governança devem ser adotadas para direcionar e controlar os investimentos realizados e para assegurar o ROI dos projetos envolvidos. Quando se opta por desenvolver uma solução de TI em formato de serviço existe um investimento a mais, e se estes ativos criados não forem reutilizados, a organização simplesmente gasta este investimento a mais e não tem nenhum retorno por isso. SOA ainda pode dar destaque ao Retorno do Investimento (ROI) por dar visibilidade aos sistemas legados por meio de interfaces ou extrair processos e serviços em domínios existentes, possibilitando a mudança de processos sem custos elevados e com respostas rápidas.

Quando há o Retorno do Investimento em SOA? Para isto, é fundamental que se pense em longo prazo, pois SOA é uma metodologia de TI de longo prazo, a qual não deve ser suscetível a flutuações estratégicas ou econômicas de curto prazo.

Então a aplicação de SOA traz benefícios? Depende. Os benefícios advêm da aplicação correta dos controles de Governança SOA, os quais permitem a maximização da reutilização de serviços e componentes e o aumento do ROI em projetos de TI. É pela aplicação rígida da Governança SOA que é possível medir e gerenciar todos os processos organizacionais e identificar pontos falhos que inibem a maximização do reuso e a obtenção do ROI.

Os benefícios podem ser classificados da seguinte maneira.

- Benefícios Operacionais: redução de custo de administrar, executar e manter a infraestrutura de TI, melhora na produtividade devida à otimização de processos completamente automatizados e integrados; melhoria no planejamento da cadeia de suprimentos, visto que as cadeias interna e externa estão integradas.
- Benefícios Gerenciais: o processo de reengenharia que ocorre pela aplicação de SOA resulta em processos de negócio mais organizados e melhorados, resultando em melhora no desempenho, bem como na qualidade dos dados, o que por sua vez melhora o desempenho e a gestão por dar suporte ao processo de tomada de decisão em toda a infraestrutura integrada.
- Benefícios Estratégicos: aumento do retorno sobre investimento, atingimento da satisfação do consumidor e melhoria na colaboração entre parceiros.
- Benefícios Técnicos: infraestrutura de TI flexível, gerenciável e passível de manutenção; redução de redundância de dados e sistemas; implementação mais rápida e barata do que soluções individualizadas.
- Benefícios Organizacionais: maior efetividade nos negócios, devido à organização, automação e integração dos processos; redução das tarefas manuais e eliminação de tarefas desnecessárias e/ou redundantes.

Assim, quando tudo é feito de maneira correta, com governança, a aplicação de SOA oferece efetividade e eficiência no processo de implantação de serviços sistêmicos alinhados ao negócio ou à integração de sistemas legados, permitindo às organizações governar seus processos, acompanhar o crescimento do reuso de serviços e evidenciar o retorno dos investimentos aplicados.

2.5 Modelos de Serviços em SOA

Um serviço SOA é a unidade lógica projetada segundo os princípios da orientação a serviços. São estruturas sistêmicas que possuem um contrato formal com declaração de suas entradas, saídas, possíveis exceções, funcionamento, versão, etc. Um serviço representa uma funcionalidade do negócio, o qual tem por objetivo a troca de dados entre fornecedor e consumidor; estes podem ser outros blocos sistêmicos ou interfaces como o usuário.

- Reutilizável.
- Possui contrato formal.
- Prevê um baixo acoplamento.
- Abstrai a lógica.
- É capaz de ser composto para construção de outro serviço.
- É autônomo e único.
- Evita alocação de recursos por longos períodos.
- Apresenta facilidades na sua descoberta.

Modelos de serviços são classificações estabelecidas para criar camadas lógicas de abstração que organizam um inventário de serviços e introduzem domínios de governança. Modelos de serviço provêm meios de atribuir uma etiqueta permanente a serviços que representam seu propósito e função geral. Um modelo de serviço, essencialmente, estabelece um contexto funcional base para um serviço que irá então ser refinado durante o processo de modelagem de serviço.

Modelos de serviço são estabelecidos a partir das seguintes distinções de lógica.

- Lógica de negócio vs. Lógica utilitária.
- Lógica agnóstica vs. Lógica não agnóstica.

A lógica é classificada como *lógica de negócio* quando ela é derivada de modelos de análise e especificações de negócio. Exemplos desse tipo de documentos incluem fluxo de trabalho (*workflow*) ou definições de processos de negócio, especificações BPM, ontologias, taxonomias, modelos lógicos de dados, diagramas de entidades de negócio e uma variedade de outros documentos relacionados com a análise negocial, análise de dados e arquitetura da informação.

Partes da lógica de processamento que não são relacionadas ou derivadas da lógica de negócio são classificadas como *lógica utilitária*. Esse tipo de lógica está relacionado a diversas funções de processamento disponíveis como recursos tecnológicos que apoiam a construção de soluções tecnológicas automatizadas.

Logica agnóstica é relacionada ao desenvolvimento e execução de serviços compartilhados, os quais podem ser utilizados por diferentes consumidores estabelecidos em diferentes partes da corporação. Um contexto de serviço agnóstico não tem propósito

específico e, portanto, é potencialmente útil para muitos, constituindo uma lógica suficientemente genérica de modo a não ter conhecimento sobre uma tarefa pai em particular. Como o conhecimento específico a uma tarefa de propósito único é intencionalmente omitido, uma lógica agnóstica é considerada de multi-propósito e reutilizável.

Por outro lado, uma lógica que é específica (tem conhecimento sobre) a uma tarefa de propósito único é classificada como *lógica não-agnóstica*.

Nesta AR-SOA, são considerados cinco modelos comuns de serviço, conforme ilustrado na **Figura 2.1**.

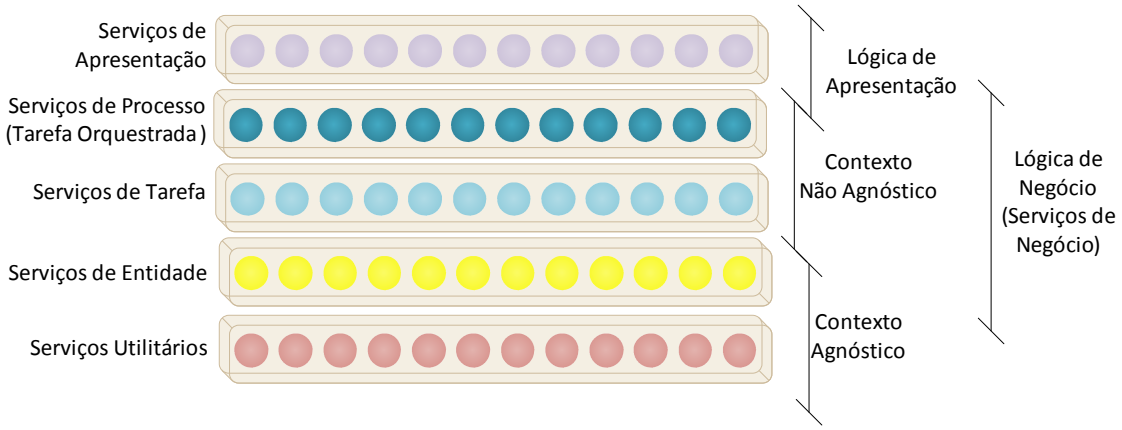


Figura 2.1: Modelos de Serviços

- Serviços de Apresentação: são *software* que realizam interfaces de usuário compartilháveis e extensíveis que podem ser invocados por outros elementos (i.e. serviços ou plataforma de SOA) de uma solução SOA. Estes serviços estão associados com a camada de apresentação do modelo lógico desta arquitetura de referência. Uma lógica residindo na camada de apresentação pode não precisar, de fato, ser encapsulada por um serviço.
- Serviços de Processos (Tarefas Orquestradas): são serviços que encapsulam lógica de processos de negócio pai. Estes serviços estão associados com a camada de apresentação do modelo lógico desta arquitetura de referência. Uma lógica residindo na camada de processo pode não precisar, de fato, ser encapsulada por um serviço. Representam *workflows* ou processos de negócio a longo prazo, e envolvem um estágio de conhecimento que abrange o *SOA habilitado para processos*. Este tipo

de serviço, em geral, tem um “estado” que se mantém estável durante múltiplas chamadas (*stateful*).

- Serviços de Tarefa: são serviços que encapsulam lógica de negócio específica de uma tarefa de negócio, frequentemente relacionada a atividades de processo (i.e. produz um resultado/saída). De natureza intencionalmente não agnóstica, possuem contexto e propósito específico e, portanto, podem ter potencial de reuso limitado ou inexistente. Usualmente atuam como controladores em composições complexas, compondo serviços de entidade e utilitários, além de outros serviços de tarefa. Frequentemente, assumem configuração na qual sua execução torna-se dependente de estado (*statefull*). Normalmente, expõem um número reduzido de capacidades.

- Serviços de Entidade: são serviços cujo contexto é derivado de uma entidade de negócio específica, por exemplo, usuário, profissional, estabelecimento de saúde ou de um grupo de entidades de negócio relacionadas. Capacidades de serviços de entidade provêm funcionalidades centradas em torno do processamento do conjunto de informações associado com a entidade de negócio. Isso geralmente resulta na definição de capacidades de CRUD. Serviços de entidade se situam na parte de baixo da hierarquia de composições, realizando composição apenas de serviços utilitários e de outros serviços de entidade com granularidade funcional mais refinada.

- Serviços Utilitários: provêm funcionalidade que endereça interesses múltiplos, enquanto mantêm um escopo funcional não negocial. Frequentemente, encapsulam recursos corporativos, tais como sistemas legados e bases de dados e os expõem em contextos únicos. Residem na parte de baixo de hierarquias de composição, nas quais são compostos por serviços de tarefa e de entidades. Alguns serviços utilitários podem, até mesmo, ser compostos por outros serviços utilitários.

Em algumas arquiteturas é possível ainda definir serviços de dados, os quais escrevem dados de/para um sistema *backend*. Podem prover algumas funcionalidades mínimas do negócio. É ideal que seja autônomo e podem ser atômicos (garantem a leitura ou escrita) ou consistentes (garantem a consistência).

A Figura 2.2 ilustra a relação entre os modelos de serviço definidos na arquitetura e os tipos de lógica que compõem cada contexto respectivo.

	Lógica de Apresentação	Lógica Negocial	Lógica Utilitária	Lógica Agnóstica	Lógica Não-Agnóstica
Serviços de Apresentação					
Serviços de Processo					
Serviços de Tarefa					
Serviços de Entidade					
Serviços Utilitários					

Figura 2.2: Modelos de Serviços *vs* Tipo de Lógica

Outros modelos especializados de serviço que podem ser definidos, conforme dado a seguir.

- Serviços de Persistência e Adaptadores de Banco de Dados (Utilitários): provêm acesso direto a recursos de persistência de dados, usando estruturas de dados nativas.
- Serviços de Sistemas Legados (Utilitários): encapsulam lógica de sistemas legados, expondo-a como serviço.
- Serviços de ETL (Utilitários): provêm funções de processamento de múltiplas transações/registros, em lote (*batch*).
- Serviços de Replicação de Dados (Utilitários): provêm lógica de replicação de dados entre aplicações diferentes. Podem operar tanto em modo síncrono quanto em modo assíncrono, dependendo da arquitetura de replicação adotada.
- Serviços de Gerenciamento do Modelo Mestre de Dados (MDM) (Utilitários): provêm acesso a lógica de MDM corporativa.
- Serviços de Convergência de Interatividade (Utilitário): provêm acesso a recursos de comunicação e interatividade, tais como telefonia, e-mail, chat, videoconferência, entre outros.
- Serviços Canônicos (Entidade): provêm acesso a bases de dados corporativas normalizadas, por meio de modelos de representação de informações e seus relacionamentos consistentes e padronizados.
- Serviços de Regras de Negócio (Tarefa): provêm centralização de regras de negócio.

- Serviços de Monitoramento de Negócio (Tarefa): provêm lógica de monitoramento de processos, normalmente realizada em tempo real (BAM).
- Serviços de Apoio a Decisão (Tarefa): provêm acesso a dados analíticos (lógica de Business Intelligence – BI).

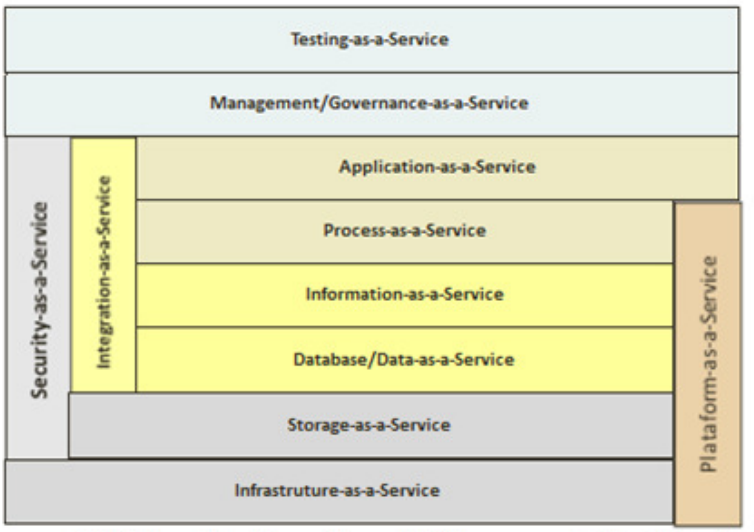
3 ARQUITETURA DA NUVEM DE SERVIÇOS PARA O RIC

Para o estabelecimento de uma infraestrutura aderente a Arquitetura Orientada a Serviços (SOA) e com condições de suportar o esperado volume de acesso massivo, contínuo e distribuído baseado em serviços web como idealizado para o RIC, há de se conhecer ao mínimo o escopo do RIC, os arquétipos necessários e, principalmente, os serviços essenciais que devam ser providos pela infraestrutura do mesmo. Soluções baseadas em SOA são tipicamente sistemas cooperativos abertos, nas quais novas entidades podem dinamicamente passar a compor o sistema, evoluírem, ou mesmo, deixarem de existir.

Além disto, em virtude da distribuição e do uso comum de *middlewares* orientado a mensageria interoperável, existe a possibilidade de que sistemas baseados em SOA sejam suportados por ambientes heterogêneos distribuídos geograficamente, isto é, os componentes dos sistemas (serviços) podem estar sendo executados em máquinas distintas, fisicamente distantes e heterogêneas, precisando apenas se comunicar por troca de mensagens. O grande diferencial dessa arquitetura é o ideal perseguido de aumentar a capacidade de compor ou recompor conjunto de serviços na resolução de contextos funcionais superiores.

Assim, como no desenvolvimento de qualquer modelo baseado em serviços web e SOA, é prevista a aplicação de diversos níveis de serviços (as-a-Service) de forma a contribuir com a possível implantação do modelo em infraestruturas funcionais e escaláveis aos moldes de nuvens computacionais.

A Figura 1 a seguir ilustra os diversos níveis de serviços conforme apontado por Linthicum (2009).



D. Linthicum. Cloud Computing and SOA Convergence in Your Enterprise A Step-by-Step Guide. Addison-Wesley. 2009. pg 199

Figura 1 - Camada de serviços (**-as-a-Service*)

Reconhecida essas possibilidades no RIC, e visto o grande volume de acesso e o acesso diferenciado com fortes preocupações com condições de privacidade, será necessária, então, a aplicação de padrões básicos de projeto de serviços estabelecidos junto a Arquitetura Orientada a Serviços, tais como de Contratos Concorrentes (*Concurrent Contracts*) e de Implementação Redundante (*Redundant Implementation*). Tais abstrações são documentadas e detalhadas no "RT de Soluções de *Message Queue* e Barramentos de Serviço Corporativos".

Ainda dos cenários discutidos no "RT Infraestrutura Tecnológica: Avaliação na Necessidade de Contratação de Múltiplos Parceiros" v3.1 do Programa RIC, no que se refere a capacidade de distribuição do modelo, têm-se diversos modelos arquiteturais propostos com diversos cenários possíveis, como ilustrado na Figura 2, e descritos a seguir.

- Cenário 1 - Uma única empresa é definida como parceiro tecnológico do RIC e é responsável por hospedar a solução e prover todos os serviços. Esta empresa deve possuir dois ou três (conforme será definido no futuro) sites espelhados permitindo o balanceamento de carga das transações efetuadas entre os sites e mantendo a disponibilidade dos serviços em pelo menos um site.
- Cenário 2 - Duas empresas são definidas como parceiros tecnológicos do RIC,

sendo que as duas empresas ficam responsável por hospedar todos os serviços do RIC, atuando uma como site espelho da outra. Neste modelo, deve haver balanceamento das transações entre os sites das duas empresas, ou seja, as requisições serão distribuídas entre os sites de modo proporcional aos tempos de resposta, custos, etc. Neste modelo, cada uma das empresas não precisa ter um espelho próprio, já que a redundância dos dados é provida pela outra.

- Cenário 3 - Duas empresas são definidas como parceiros tecnológicos do RIC, sendo que cada uma das empresas fica responsável por prover parte dos serviços do RIC. Por exemplo, uma empresa hospeda o serviço de cadastramento e a outra o serviço de autenticação. Cada uma das empresas deve manter um espelhamento dos seus serviços em um site de backup, que só será acionado em caso de queda do site principal ou funcionarão com balanceamento de carga.
- Cenário 4 - Duas ou mais empresas são definidas como parceiros tecnológicos do RIC, sendo que cada uma das empresas fica responsável por prover uma camada diferenciada no modelo tecnológico do RIC. Por exemplo, uma empresa provê o acesso ao processamento do RIC e outra provê o acesso aos dados básicos. Cada uma das empresas deve manter um espelhamento dos seus serviços em um site de backup, que só será acionado em caso de queda do site principal ou funcionarão com balanceamento de carga.

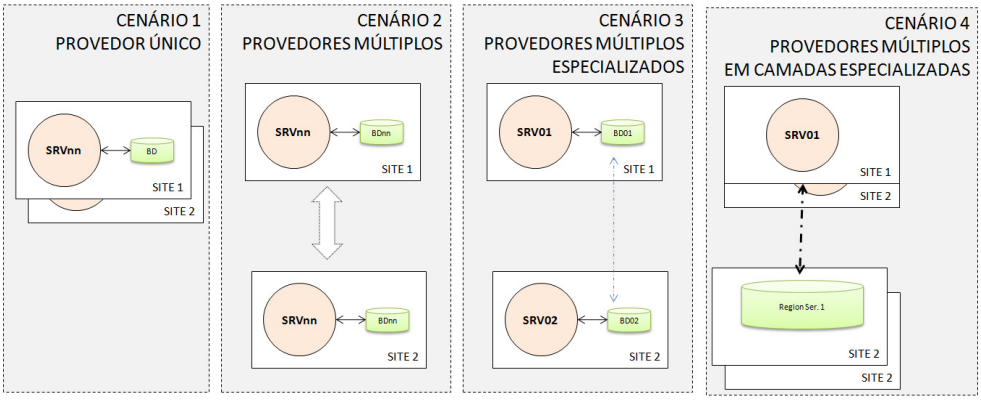


Figura 2 - Cenários Avaliados

Para entendimento desses cenários, cabe-nos apontar nossa visão da infraestrutura para atendimento de serviços necessários ao RIC. Neste sentido entendemos haver serviços de processamento e serviços de armazenamento de dados. Fica óbvia que, no

caso de sua separação física, é necessária a elaboração de um forte mecanismo de garantia de confiança entre as partes, de velocidade de interligação e de acesso irrestrito do serviço de processamento ao serviço de armazenamento (Figura 3).

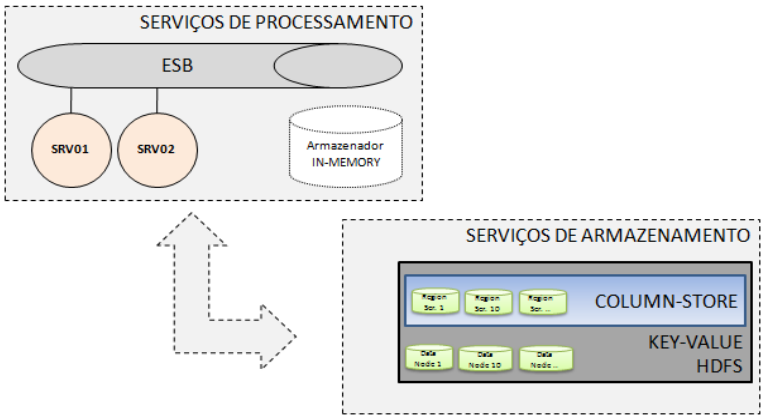


Figura 3 - Serviços de processamento e serviços de armazenamento de dados.

Nos serviços de processamento, todo processo que tenha incluídas lógicas de negócio devem residir neste ambiente, o qual tem por fim prover suporte aos diversos modelos de processamento necessários ao RIC. Por meio desse ambiente deve-se buscar eliminar a possibilidade de acesso direto aos dados armazenados, bem como permitir a criação de mecanismos que impeçam ao fiel depositário dos dados (serviço de armazenamento) o conhecimento dos dados por ele armazenado. Este conjunto tem similaridades com o que se conhece como Barramento de Serviços Corporativos (ESB - *Enterprise Service Bus*), sendo necessária a existência de elementos comuns aos mesmos tais como suporte a fila de mensagens (MQ - *Messaqe Queue*), mensageria confiável e tolerância a falhas. Estes serviços devem ter em sua essência uma baixíssima latência, pois o tempo necessário de processamento estaria acesso aos serviços de armazenamento de dados.

Nos serviços de armazenamento, teríamos a capacidade não só de recuperação, mas todo conjunto de operações de persistência de dados. Esses serviços dividem-se em serviços de acesso a dados básicos (Dados como Serviço - DaaS) e serviços de informações (Informações como Serviço – IaaS, necessários para entrega para ambientes de BI). Assim, de nossas primeiras discussões, algumas observações são tratadas:

- os serviços de processamento, e o correspondente serviço de acesso a dados, referentes a autenticação e autorização devem se dar ao mínimo possível de escalabilidade geográfica, necessitando que dados mínimos residam junto aos serviços de processamento de forma a garantir um conjunto com mais baixa

latência possível;

- aos serviços de processamento estão associados serviços de tarefa básicos que são usados na formação / execução dos fluxos de processos de negócio necessários (tarefas complexas);
- o acesso a dados referentes a autenticação e autorização deve utilizar de modelos muito rápidos de recuperação de dados armazenados, sendo possível o uso de SGBD's distribuídos e/ou modelos de armazenadores NoSQL chave-valor ou HDFS básico, com possível uso de bancos em memória;
- em uma primeira análise do proposto, os serviços de armazenamento são fortemente agnósticos e são a base para uso dos serviços de tarefas e de tarefas orquestradas;
- os serviços que atuam no âmbito da manutenção de dados devem tratar tanto o acesso aos dados biométricos – ABIS, e a correspondente manutenção das bases de dados de autenticação e autorização, quanto o cadastramento de outros dados que podem ocorrer utilizando-se um modelo de SGBD's distribuídos /ou modelos de armazenadores orientado a colunas.

Para entendimento dessas camadas e suas relações com Arquiteturas Orientadas a Serviços, o modelo sugerido se reporta aos princípios de camadas da arquitetura como pode ser observado na Figura 4 nos elementos (a) e (b) como proposto por Josuttis (2008) e Erl (2005), respectivamente. Mais detalhes sobre esses elementos e sua relação com o Barramento de Serviços Corporativos são discutidos no documento "RT de Soluções de *Message Queue* e Barramentos de Serviço Corporativos"

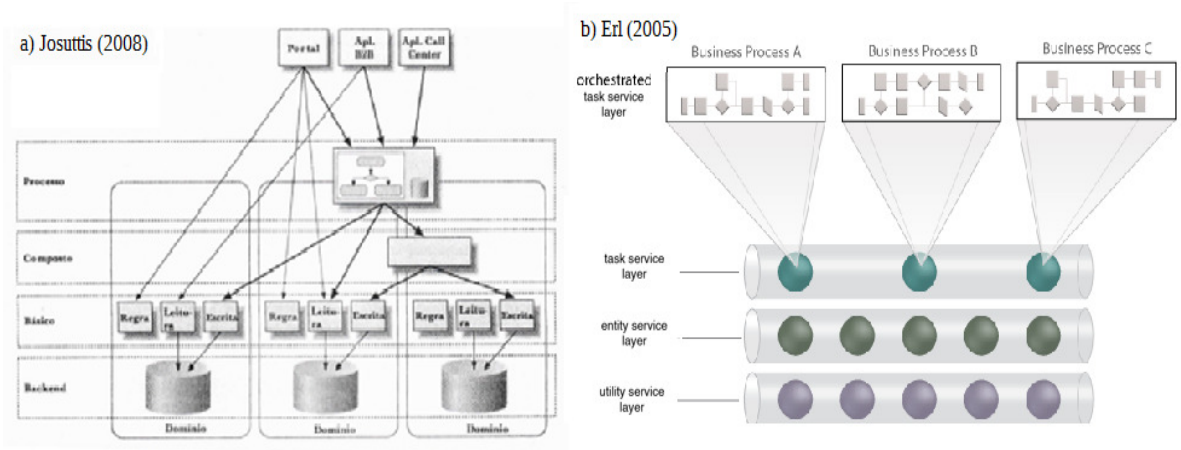


Figura 4 – Camadas Tradicionais da Arquitetura Orientada a Serviços

A seguir ilustramos na Figura 5 nosso entendimento da arquitetura pretendida.

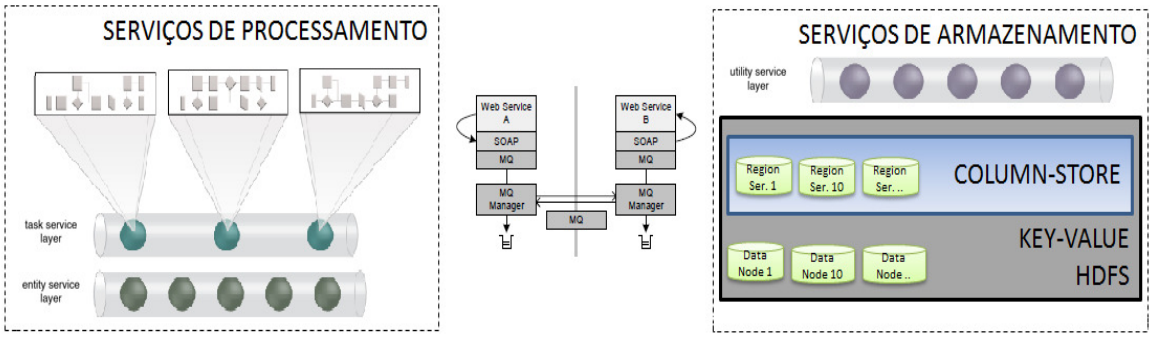


Figura 5 – Arquitetura sugerida com as camadas comunicando-se através do MQ

Em continuidade, em uma análise mais profunda dos serviços que serão providos pela camada de armazenamento e pela camada de processamento, em conformidade com as características apresentadas pelo projeto RIC, pode-se idealizar inicialmente grandes grupos de serviços básicos - que ora são caracterizados tão somente como serviços, mas da forma como apresentados representam estruturas maiores que agrupam os prováveis candidatos de serviços, que terão sua definição final na fase de Análise e Modelagem de Serviços - segundo seu padrão de uso e comportamento perante o ecossistema, e que poderão ser usados em composições em aplicações distintas. Esses grandes grupos são os seguintes.

- **Serviços de Autenticação (Validação e Identificação):** acesso com mais baixa latência, pouco recurso de composição de serviços, alta escalabilidade (incluindo escalabilidade geográfica) e bases de dados de alto desempenho, se possível,

usando SGBD em memória (*in memory*).

- **Serviços de Federação (Identidade Federada): a**cesso através de chaves que garantam a portabilidade de credenciais na execução de processos dentro do ecossistema RIC.
- **Serviços de Registro (Criação da chave primária):** composição de serviços entre os ABIS e a criação da chave única do RIC com suporte a redistribuição nas bases de dados localizadas.
- **Serviços de Cadastramento (Cadastro de dados básico e documentos):** modelo de média latência porém síncrono.
- **Serviços de Informação (Requisição-Resposta a solicitações de terceiros):** serviços síncronos de informação.
- **Serviços de Integração (Requisição-Resposta a requisições do RIC):** modelo síncrono de solicitações do RIC para serviços de terceiros no modelo requisição-resposta; com fortes requisitos de tolerância às falhas.
- **Serviços de Comunicação (Mensageria Direcionado por Eventos):** modelo assíncrono baseado em princípios de padrões de mensageria *publish-subscribe*.
- **Serviços de Atendimento (Processo de Agendamento):** operações de CRUD tanto para o processo de agendamento quanto para outros atendimentos aos cidadãos.
- **Serviços de Operacionalização (Processo de Emissão de Documentos RIC):** operações de CRUD para informação aos prestadores de serviços de emissão do documento referente ao RIC.
- **Serviços de Auditoria (Notificar Log, Auditoria e "Problemas Graves"):** mensageria assíncrona – notificação para bases de dados de auditoria do registro do RIC bem como de acesso ao RIC.
- **Serviços de Sincronismo (Sincronizar Bases Geodistribuídas):** modelo assíncrono de garantia de sincronia entre a base centralizada e os elementos distribuídos do modelo.
- **Serviços Analíticos (*Business Intelligence*):** utiliza-se de relatórios pré-formatados e cubos multidimensionais para consultas ad-hoc e análise

multidimensional.

Todos esses serviços se caracterizam pelo uso de modelos de mensageria diferenciados no barramento, de comunicação síncrona ou não, de características de entrada e saída (*inbound* ou *outbound*), e de necessidades diferenciadas de latência no processamento de dados. Isto reforça a capacidade de execução segundo os requisitos do projeto para os usuários, associando esses requisitos a modelos tecnológicos diferenciados na infraestrutura arquitetural e com a forte possibilidade de aplicação de modelos distribuídos no atendimento de cada grupo de serviços que representam os requisitos iniciais de tarefas do RIC.

Estende-se que esses grupos de serviços representam as entidades de tarefas a serem realizadas, e que irão realizar composições de serviços nas camadas de armazenamento (denominada também como serviços utilitário de dados, que compõe a camada persistência de arquiteturas orientada a serviços). Ainda, para mais detalhes dos componentes que irão prover esses padrões de mensageria, sugerimos a leitura do documento “"RT de Soluções de Message Queue e Barramentos de Serviço Corporativos".

O Quadro 1 ilustra uma representação visual que posiciona cada um desses grupos de serviços com os requisitos que são cumpridos pelos mesmos.

Autenticar	• Validação e Identificação de cidadãos
Federar	• Autorização de ações no sistema
Registrar	• Criação da chave primária
Cadastrar	• Entrada de dados básicos e cadastramento de documentos
Informar	• Requisição-Resposta a solicitações de terceiros
Integrar	• Requisição-Resposta a requisições do RIC
Comunicar	• Mensageria Direcionado por Eventos
Atender	• CRUD da agenda e ações de atendimento ao cidadãos
Operacionalizar	• Ações sistêmicas na emissão de documentos do RIC
Auditar	• Notificar Log, Auditoria e “Problemas Graves”
Sincronizar	• Sincronizar Bases Geodistribuídas
Analisar	• Business Intelligence

Quadro 1- Representação Visual dos Grupos de Serviços e os Requisitos Cumpridos

Essa divisão também oferece a possibilidade de análise de perfis e papéis de acesso segundo características de uso e de forma de acesso (somente leitura ou leitura-escrita).

Na subseção seguinte iremos detalhar essas as sugestões iniciais de grandes grupos de serviços, suas capacidades e os padrões de mensageria sugeridos. Nesta análise já são observados dados iniciais tais como a nomeação de algumas das capacidades esperadas, sua funcionalidade (contexto funcional a ser atendido), suas entradas e saídas como padrões de mensagens, suas características e de acesso (com detalhe para a expressão "*too-many*" que indica um acesso massivo ao serviço) e volume nominal estimado, e o apontamento do tipo de entidade ou entidade direcionada a ser atendida por tais capacidades de serviços.

Considera-se ainda que esta documentação minimamente aponta sugestões, já que envolve responsabilidades afetas ao Grupo de Ecossistema no que se referem as atividades de: Aplicações Governamentais e Privadas, Definição do Visionamento das Aplicações, Cooperação Internacional.

Ainda, no que se refere a sugestões nos modelos canônicos e estrutura física de armazenamento do RIC, o mesmo também envolve responsabilidades afetas ao mesmo Grupo de Ecossistema no que se referem às atividades de definição das informações gerenciadas pelo projeto RIC.

4 PERFIL DOS GRUPOS DE SERVIÇOS

4.1 Serviços de Autenticação

Referem-se fundamentalmente ao atendimento dos princípios de autenticação por meio de recursos de identificação, pesquisa e validação do acesso a partir de chaves do RIC.

Capacidade:	**Validar**
Contexto Funcional:	Identifica única e inequivocamente um registro no RIC.A mensageria de retorno informa apenas esta validade (S/N).
Input:	dadoValidacaoRIC
Output:	mensagemRIC
Característica de Acesso:	*Read-too-many*

Clientes Potenciais:	Na camada de armazenamento, é utilizado por composição por diversos serviços de tarefas.

Capacidade:	**Identificar**
Contexto Funcional:	Lista os prováveis elementos identificados em uma pesquisa nos registros do RIC. A mensageria de retorno é a lista de dados básicos de pesquisa no RIC. Os dados básicos são limitados pela característica de acesso público.
Input:	dadosIdentificacaoRIC
Output:	dadoPesquisaRIC [unbouded]
Característica de Acesso:	*Read-many*
Clientes Potenciais:	Na camada de armazenamento, é utilizado por composição por diversos serviços de tarefas.

4.2 Serviços de Federação

Referem-se fundamentalmente ao acesso através de chaves que garantam a federação na execução de processos dentro do ecossistema RIC. Envolve conceitos de Identidade Federada e uma implementação em uma arquitetura suplementar adjacente (logo não é classificado como um serviço interno ao modelo de serviços do RIC). Este modelo suplementar ora se utiliza da capacidade de Validação dos grupos de serviços de autenticação.

4.3 Serviços de Registro

Referem-se ao processo geral de criação, remoção e expiração da chave primária do RIC. Assim, o processo de cadastramento de uma nova identificação é utilizado em uma composição com esse serviço para realizar todo o cadastramento.

Capacidade:	**Criar**
Contexto Funcional:	Cria uma chave única e inequívoca no registro no RIC.
Input:	dadoIdentificacaoRIC
Output:	mensagemRIC
Característica de Acesso:	*Write-too-many*
Clientes Potenciais:	Sistemas adjacentes ao projeto RIC.

Capacidade:	**Apagar**
Contexto Funcional:	Apaga uma chave única no registro no RIC tratando sua existência e realizando um processo assíncrono de informação de apagamento de todas as referências a mesma.
Input:	dadoIdentificacaoRIC
Output:	mensagemRIC
Característica de Acesso:	*Write-once*
Clientes Potenciais:	Sistemas adjacentes ao projeto RIC.

Capacidade:	**Expirar**
Contexto Funcional:	Marca como inativa uma chave única e inequívoca no registro no RIC.
Input:	dadoValidacaoRIC
Output:	mensagemRIC
Característica de Acesso:	*Write-once*
Clientes Potenciais:	Sistemas adjacentes ao projeto RIC.

4.4 Serviços de Cadastramento

Referem-se fundamentalmente aos serviços que atendem a manutenção da base do RIC e a respectiva entrega de informações como serviço (dados como serviço – DaaS). Junto aos serviços de cadastramento e informação encontra-se capacidade (ou serviço específico) para atender ao processo de agendamento de cadastramento que sugerimos ser suportada por base de dados distinta, mas integrada ao RIC.

Capacidade:	**Cadastrar**
Contexto Funcional:	Permite o cadastramento (inclusão) na base de dados do RIC.
Input:	dadosIdentificacaoRIC
Output:	mensagemRIC
Característica de Acesso:	*Read-Write-too-many*
Clientes Potenciais:	Sistemas adjacentes ao projeto RIC.

Capacidade:	**Informar**
Contexto Funcional:	Solicita permissão de acesso em nível de confidencialidade específico ao RIC.
Input:	dadoValidacaoRIC dadoAgendaCadastramentoRIC

Output:	dadosRIC
Característica de Acesso:	*Read-too-many*
Clientes Potenciais:	Sistemas adjacentes ao projeto RIC.

Capacidade:	**Atualizar**
Contexto Funcional:	Permite a atualização na base de dados do RIC.
Input:	dadosIdentificacaoRIC
Output:	mensagemPesquisaRIC
Característica de Acesso:	*Write-many*
Clientes Potenciais:	Sistemas adjacentes ao projeto RIC.

Capacidade:	**Expirar**
Contexto Funcional:	Permite o cancelamento lógico no RIC, o que invalida as ações nos serviços de autenticação no RIC. Esta exclusão é lógica, não envolvendo a deleção de dados no RIC. A mensageria de retorno informa apenas a validade dessa ação (S/N).
Input:	dadoValidacaoRIC
Output:	dadosRIC
Característica de Acesso:	*Write-once*
Clientes Potenciais:	Sistemas adjacentes ao projeto RIC.

Capacidade:	**Cancelar**
Contexto Funcional:	Realiza o cancelamento físico no RIC, eliminando a existência da instância associada a chave do RIC. A mensageria de retorno informa apenas a validade dessa ação (S/N).
Input:	dadoValidacaoRIC
Output:	dadosRIC
Característica de Acesso:	*Write-once*
Clientes Potenciais:	Sistemas adjacentes ao projeto RIC.

4.5 Serviços de Informação

Referem-se fundamentalmente aos serviços que suportam o atendimento de terceiros no acesso a dados e situações no RIC. Como ainda está em desenvolvimento, esse mapeamento de relações, o desenvolvimento da expectativa de capacidades futuras bem como a apropriação dos aspectos tecnológicos necessários ainda não é possível.

4.6 Serviços de Integração

Referem-se a serviços que serão utilizados por soluções do RIC para solicitação de informações externas ao ambiente do mesmo. Envolve padrões de mensageria do tipo requisição-resposta oriundas do ambiente do RIC.

Capacidade:	**Verificar**
Contexto Funcional:	Atua verificando informações a partir de fontes confiáveis adjacentes.
Input:	dadoPesquisaRIC
Output:	dadosRIC
Característica de Acesso:	*Read-many*
Clientes Potenciais:	Cartórios, etc.

Capacidade:	**Homologar**
Contexto Funcional:	Suporta o registro de informações oriundas de fontes confiáveis junto ao RIC.
Input:	dadoPesquisaRIC
Output:	dadosRIC
Característica de Acesso:	*Write-once*
Clientes Potenciais:	Cartórios, etc.

4.7 Serviços de Comunicação

Referem-se a serviços de informação baseados em eventos oriundos do modelo proposto para informação intempestiva a sistemas suportados pelo RIC. Baseia-se fundamentalmente em processos de mensageria direcionada por eventos (*publish-subscribe*) e pré-estabelece a existência de qualquer modelo de MQ ou ESB junto ao RIC.

Capacidade:	**Centralizar**
Contexto Funcional:	Trazer automaticamente informações distribuídas ou descentralizadas para o RIC.
Input:	dadosIdentificacaoRIC dadosRIC
Output:	mensagemRIC
Característica de Acesso:	*publish-subscribe*
Clientes Potenciais:	Bases regionalizadas

Capacidade:	**Sincronizar**
Contexto Funcional:	Repassar automaticamente informações distribuídas ou descentralizadas do RIC para outras bases de dados associadas ao projeto.
Input:	dadosIdentificacaoRIC dadosRIC
Output:	mensagemRIC
Característica de Acesso:	*publish-subscribe*
Clientes Potenciais:	Bases regionalizadas

4.8 Serviços de Atendimento

Referem-se fundamentalmente a serviços em todos os níveis que são utilizados no desenvolvimento de soluções baseadas na orquestração de processos para a atividade de agendamento no cadastramento do RIC.

- Agenda (Solicitar, Reagendar, Confirmar, Realizar)
- Cadastramento e Atualização (conjunto de serviços de tarefa que se utilizam do motor de orquestração do ESB/SOA para implementar uma solução que preveja os macroprocessos do Sistema RIC.

Para implementação deste modelo sugere-se o uso de serviços básicos de gerenciamento e execução de processos de negócio baseados em tarefas orquestradas.

4.9 Serviços de Operacionalização

Referem-se fundamentalmente aos serviços em todos os níveis que são utilizados no desenvolvimento de soluções baseadas na orquestração de processos para a atividade de operacionalização junto a terceiros da emissão de documento do RIC. Dada sua característica de uso por entidades externas ao ecossistema do Sistema RIC - como no caso as emissoras de documento do RIC - esses foram isolados do grupo de serviços de Atendimento para tratamento dos aspectos específicos de segurança.

- Emissão do Documento do RIC (Geração de Lotes, Confecção, Entrega, Cancelamento)

4.10 Serviços de Auditoria

Referem-se fundamentalmente a serviços assíncronos que suportam todos mecanismos de registro de log e auditoria de atividades gerais do RIC.

- Log (Geração do RIC)
- Auditoria (Uso do RIC)
- “Problemas Graves”

Esses serviços podem ser supridos pelo suporte de infraestrutura subjacente de controle do Barramento Corporativo de Serviços - ESB.

4.11 Serviços de Sincronismo

Estabelecem um grupo de serviços que permite um possível sincronismo de dados entre unidades regionalizadas do RIC. Esta extensão é uma situação provável, não estando estabelecidas ainda diretrizes que permitam a definição dessas atividades sistêmicas.

- (DE/IN) Bases Geodistribuídas
- (PARA/OUT) Bases Geodistribuídas.

No conjunto de soluções de serviços de sincronismo, dependendo do tipo de escalabilidade e distribuição previsto para implementação do RIC, pode ser substituído por modelos de sincronismo, replicação ou técnicas modernas de distribuição junto ao sistema gerenciador de armazenamento de dados.

Diversas soluções de armazenamento massivo de dados já preveem recursos de sincronismo e/ou distribuição para atender as questões de consistência eventual (*Eventual Consistency*) de suas arquiteturas de armazenamento (detalhes junto ao "RT de Infraestrutura Tecnológica do RIC: Armazenadores NoSQL”).

4.12 Serviços Gerenciais

Referem-se a disponibilização de dados como serviço na forma de fonte de dados utilizáveis por modelos analíticos para suporte às soluções OLAP.

Capacidade:	**ConsultarOLAP**
Contexto Funcional:	Dispõem dados multidimensionais como serviço.
Input:	<XML-MDX><ODATA Request>
Output:	<XML/A><ODATA Request>
Característica de Acesso:	*Read-only*
Clientes Potenciais:	Atores gerenciais associados a execução do RIC

5 ESQUEMA PARA ATENDER OS SERVIÇOS

Para dar entendimento aos perfis de grupos de serviços apresentados, a Figura 10 ilustra os prováveis modelos de mensageria previstos para o RIC (à exceção do processo de deduplicação), ao qual iremos usar como base para tratar a associação do tipo de mensagem e os grupos de serviços.

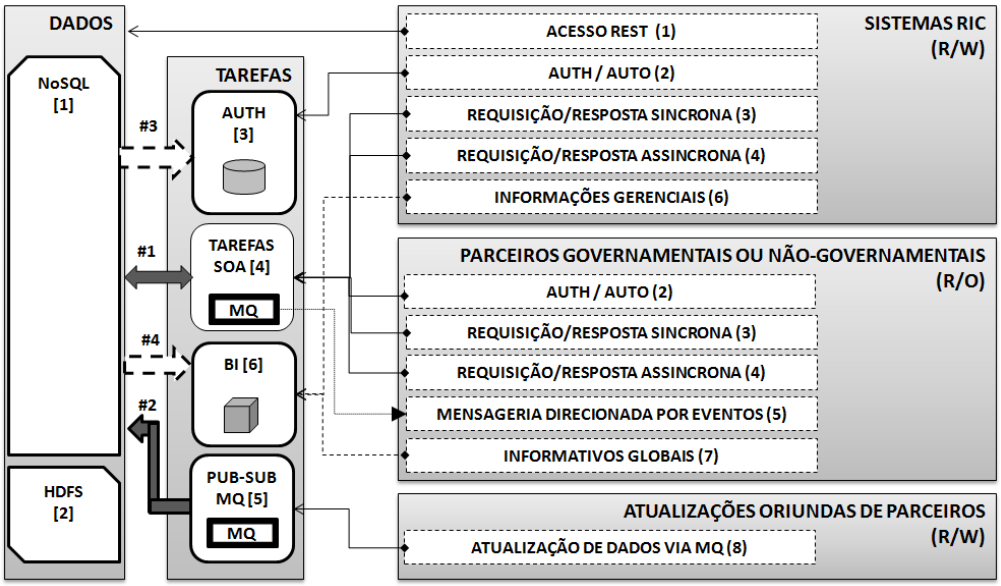


Figura 10 - Esquema Geral Proposto

5.1 Dos Elementos Sugeridos (notação: [ ])

[1] Uma arquitetura de armazenamento massivo de dados, a qual utilize técnicas de consistência eventual e que permita a implementação de bases de dados distribuídas escalável com alta performance e baixa latência.

[2] Uma arquitetura de armazenamento massivo de arquivos, a qual utilize técnicas de consistência eventual e redundância, que permita a implementação de processos de mapeamento e redução na distribuição no armazenamento dos originais de digitais, faces, íris e assinatura [3]

[4] Uma arquitetura de Barramento Corporativo de Serviços com um *Message Queue* (MQ), a qual aplique padrões de mensageria confiável, de enfileiramento assíncrono, de implementações redundantes de serviços e de contratos SOA concorrentes, e que possibilite a implantação de mensageria baseada em eventos (pub/sub).

[5] Uma arquitetura independente com um *Message Queue* (MQ), a qual aplique padrões de mensageria confiável, de enfileiramento assíncrono, de implementações redundantes de serviços e de contratos SOA concorrentes, e que possibilite a implantação de mensageria baseada em eventos (pub/sub), com requisitos restritos de segurança e confiança computacional.

5.2 Dos Modelos de Mensageria Permitidos (notação: ( ))

(1) Acesso REST: acesso exclusivo e restrito a funcionalidades de CRUD diretamente sobre o modelo de dados do RIC. Este acesso será utilizado somente por funcionalidades específicas e restritas dentro do Sistema RIC.

(2) AUTH: modelo de comunicação direta dentro do ESB aos dados básicos válidos ao perfil de serviços de Autenticação.

(3) Modelo de Requisição-Resposta Síncrono: modelo de mensageria no qual a aplicação requisitante encaminha uma mensagem de requisição (*request*) e imediatamente suspende o seu processamento no aguardo uma mensagem de resposta (*response*). Mensagens de falhas (*message faults*) são previstas para indicar situações imprevistas ou de indisponibilidade dos serviços.

(4) Modelo de Requisição-Resposta Assíncrona (com suporte de MQ): modelo de mensageria no qual a aplicação requisitante encaminha uma mensagem de requisição (*request*) e não suspende o seu processamento. Essas requisições são empilhadas em filas para processamento (*messages queues*) que terá como função adicional reencaminhar uma mensagem de resposta. Assim, a grande vantagem que MQ de mensagem assíncrona pode delegar um pedido com nós diferentes, escaláveis e balanceados. Ele pode redirecionar os pedidos entre computadores segundo uma análise de carga a fim de suportar o processamento massivo de mensagens. Quando o pedido é concluído, o próprio computador que processou a mensagem pode retornar a resposta diretamente para o terminal de origem da mensagem. Mensagens de falhas (*message faults*) são previstas para indicar situações imprevistas ou de indisponibilidade completa dos serviços.

(5) Modelo de mensageria baseada em eventos (com suporte do MQ): modelo de

pub/sub, utilizando o MQ presente no ESB.

(6) Acesso a relatórios gerenciais pré-definidos.

(7) Acesso a serviços de BI com recursos de disponibilização de relatórios pré-definidos, consultas ad-hoc e suporte ao acesso a análise multidimensional.

(8) Modelo de MQ independente com suporte a mensageria baseada em eventos com fortes requisitos de segurança.

5.3 Dos Fluxos de Dados Entre Camadas (notação: setas largas escuras)

- Fluxo #1 de acesso a camada de dados: acesso das tarefas síncronas e assíncronas em (3), (4) e (5), atualizando o modelo de dado. Esse fluxo se refere às atualizações oriundas do Sistema RIC, ou de consulta tanto do Sistema RIC quando de requisições (R/O) dos sistemas de parceiros privados (não-governamentais).
- Fluxo #2 de atualizações restritas: acesso das requisições de atualização dos modelos de dados com tratamento intensivo de validade, prioridade e auditoria por se tratar de atualizações oriundas de requisições (R/O) dos sistemas de parceiros privados (não-governamentais).

5.4 Dos Fluxos de Consistência Eventual (notação: setas largas tracejadas)

- Fluxo #3 de consistência eventual ao armazenamento chave-valor para suporte aos processos de Validação.
- Fluxo #4 de consistência eventual dos Cubos Multidimensionais (ETL).

6 PERFIS E PROVÁVEIS SERVIÇOS NO ECOSSISTEMA

Das definições de perfis no ecossistema do RIC, têm-se: o uso dos serviços por sistemas elaborados em suporte às atividades do RIC (Sistema RIC), o uso por prováveis parceiros governamentais, o uso por parceiros privados (não governamentais) e o uso direto pelos cidadãos (limitado as suas informações junto ao modelo de armazenamento do RIC). Desses perfis é possível detalhar o que será descrito a seguir.

6.1 Sistema RIC

Entende-se o Sistema RIC como um sistema que trata das atividades (com tarefas sistêmicas ou não) que envolvem os seguintes aspectos, segundo o perfil básico de Grupos de Registro.

- Processo de emissão de 1ª via do Documento RIC (Serviços de Registro + Serviços de Cadastramento)
- Processo de emissão de 2ª via do Documento RIC (Serviços de Cadastramento)
- Cancelamento da Solicitação (Serviços de Registro)
- Cancelamento do Documento RIC após a entrega (Serviços de Registro + Serviços de Cadastramento)
- Verificação de Histórico dos Dados Biográficos e Biométricos (Serviços de Informação)
- Tratamento de Erro de Fabricação ou Perda do Documento RIC (Serviços de Operacionalização)
- Tratamento de Erro de Ofício do Documento RIC e Emissão de Certificado Digital (desconhecido)
- Encaminhamento de Relatórios Pré-formatados (Serviços Analíticos)
- Análise de Informações Gerenciais (Serviços Analíticos)

No item "RELAÇÃO ENTRE ESQUEMAS E SERVIÇOS" deste documento será tratado pontualmente cada item sugerido durante o levantamento original do sistema, e seu posicionamento junto aos serviços.

6.2 Parceiros Governamentais

Da análise do documento "Levantamento de potenciais aplicações governamentais" do Ecossistema do RIC, é apontada uma matriz multiatributos que quantifica um modelo de viabilidade institucional para parceiros governamentais notórios e comuns para os serviços ofertados pelo sistema RIC. Ao Quadro 2, o qual reproduz essa matriz quantificada, foi incorporado o entendimento dos perfis e papéis de serviços adequados às aplicações geridas pelos mesmos.

Órgãos	INDICADOR QUALITATIVO	TIPO DE SERVIÇOS	SISTEMA
MINISTÉRIO DA PREVIDÊNCIA SOCIAL	401.739.779.957	SINCRONISMO	Cadastro Nacional de Informações Sociais (CNIS)
MINISTÉRIO DA SAÚDE	106.019.264.465	SINCRONISMO	Cartão Nacional de Saúde (CNS)
		VALIDAÇÃO	Farmácia Popular
		VALIDAÇÃO	Saúde Suplementar (ANS)
MINISTÉRIO DA EDUCAÇÃO	94.490.611.520	VALIDAÇÃO E INFORMAÇÃO / INTEGRAÇÃO (RELACIONAMENTOS)	Bolsa Família X Acompanhamento da Frequência Escolar:
		VALIDAÇÃO E INFORMAÇÃO / INTEGRAÇÃO (RELACIONAMENTOS)	Exames, Programas e Fundos de Financiamento para Alunos: ENEM - Exame Nacional do Ensino Médio, ENADE - Exame Nacional de Desempenho de Estudantes, PROUNI – Programa Universidade para Todos, FIES – Fundo de Financiamento Estudantil
		VALIDAÇÃO	SISU – Sistema de Seleção Unificada
MINISTÉRIO DA DEFESA	74.017.108.772	SINCRONISMO	Identidade Militar
MINISTÉRIO DO DESENVOLVIMENTO SOCIAL E COMBATE À FOME	68.607.635.321	SINCRONISMO	Cadastro Único para Programas Sociais (Cadúnico) e outros programas
MINISTÉRIO DO TRABALHO E EMPREGO	50.098.129.556	VALIDAÇÃO E INFORMAÇÃO / INTEGRAÇÃO (RELACIONAMENTOS)	Cadastro Geral de Empregados e Desempregados - CAGED, Sistema Nacional de Emprego - SINE
		INFORMAÇÃO/INTEGRAÇÃO	Conselho Nacional de Imigração - CNIg
MINISTÉRIO DAS CIDADES	26.706.655.445	VALIDAÇÃO	Carteira Nacional de Habilitação (CNH)
MINISTÉRIO DA FAZENDA	25.998.883.674	SINCRONISMO	CPF - Cadastro de Pessoa Física
		VALIDAÇÃO (SóCIOS)	CNPJ - Cadastro Nacional da Pessoa Jurídica
		VALIDAÇÃO	CAFIR - Cadastro de Imóveis Rurais
		SINCRONISMO	CEI - Castrado Específico do INSS
MINISTÉRIO DOS TRANSPORTES	21.068.400.360	VALIDAÇÃO E INFORMAÇÃO / INTEGRAÇÃO (RELACIONAMENTOS)	Passe Livre
MINISTÉRIO DO PLANEJAMENTO, ORÇAMENTO E GESTÃO	19.154.319.533	VALIDAÇÃO	Serviço de Informação ao Cidadão (SIC)
MINISTÉRIO DA JUSTIÇA	11.962.736.692	VALIDAÇÃO	CNRF - Cadastro Nacional de Reclamações Fundamentadas
		INFORMAÇÃO/INTEGRAÇÃO	MJ/DPF - Passaporte, Controle migratório internacional, Registro e expedição de cédula de identidade de estrangeiro (CIE)
		VALIDAÇÃO	MJ/DPF - Certidão de Antecedentes Criminais
		VALIDAÇÃO	RENACH - Registro Nacional de Condutores Habilitados
		VALIDAÇÃO	Cadastro Nacional de Crianças e Adolescentes Desaparecidos
TRIBUNAL SUPERIOR ELEITORAL	6.077.120.836	SINCRONISMO	SECAD - Cadastro de Eleitores - Secad
MINISTÉRIO DO DESENVOLVIMENTO AGRÁRIO	4.897.205.500	-	-
MINISTÉRIO DAS RELAÇÕES EXTERIORES	2.345.081.277	INFORMAÇÃO/INTEGRAÇÃO	Solicitação de passaporte (comum, oficial ou diplomático) no exterior

		SINCRONISMO	Legalização de documentos emitidos no exterior e no Brasil
		VALIDAÇÃO E INFORMAÇÃO / INTEGRAÇÃO (DEMOGRÁFICOS)	Emissão de demais documentos consulares, tais como atestados de residência
		VALIDAÇÃO E INFORMAÇÃO / INTEGRAÇÃO (RELACIONAMENTOS)	Autorização de viagem para menores
SECRETARIA DE DIREITOS HUMANOS	317.527.886	VALIDAÇÃO E INFORMAÇÃO / INTEGRAÇÃO (DEMOGRÁFICOS)	Defesa de Direitos de Cidadãos
CONSELHO NACIONAL DE JUSTIÇA (CNJ)	219.262.114	SINCRONISMO	SIRC, INFOSEG, INFOJUD
INSTITUTO NACIONAL DE TECNOLOGIA DA INFORMAÇÃO	11.855.963	SINCRONISMO	AR Biométrica

Quadro 2 – Prováveis parceiros governamentais e usos segundo o perfil de serviços

6.3 Parceiros Privados (Não-Governamentais)

Da análise do documento "Levantamento de potenciais aplicações privadas" do Ecossistema do RIC, é apontada uma oferta viável de parceiros privados para os serviços ofertados pelo sistema RIC. Ao Quadro 2, o qual reproduz essa matriz quantificada, foi incorporado o entendimento dos perfis e papéis de serviços adequados às aplicações geridas pelos mesmos.

Órgãos	TIPO DE SERVIÇOS	SISTEMA
FEBRABAN	VALIDAÇÂO e SINCRONISMO	Fraudes eletrônicas CNF - Confederação Nacional das Instituições Financeiras CCS - Cadastro de Clientes do Sistema Financeiro Nacional Cadastro positivo
TELECOMUNICAÇÕES	VALIDAÇÂO	Telefonia Fixa e Móvel, e Comunicação Multimídia
SERVIÇO DE PROTEÇÃO AO CRÉDITO – SPC	VALIDAÇÂO	Cartão de Crédito, Cadastro Positivo, SPC
SETOR DE DISTRIBUIÇÃO DE ENERGIA ELÉTRICA	VALIDAÇÂO	Tarifas - Consumidores Finais
SETOR DE ÁGUA E COLETA DE ESGOTO	VALIDAÇÂO e SINCRONISMO	CNARH - Cadastro Nacional de Usuários de Recursos Hídricos
PLANOS DE SAÚDE	VALIDAÇÂO	Padrão TISS (Troca de Informações na Saúde Suplementar)
ASSOCIAÇÃO BRASILEIRA DAS EMPRESAS DE CARTÃO DE CRÉDITO E SERVIÇOS – ABECS	VALIDAÇÂO e SINCRONISMO	Mercado de Adquirentes, Empresas / Bandeiras.

Quadro 2 – Prováveis parceiros privados e usos segundo o perfil de serviços

6.4 Uso Previsto: Acesso Direto pelo Cidadão

Inicialmente está previsto ao cidadão:

- a entrega de serviços seguros e validados das suas informações cadastradas

na base do RIC (INFORMAÇÃO);

- o acesso seguro, validado e multimodo para atualização de seus dados demográficos (CADASTRAMENTO - Atualização Oriunda de Parceiros);
- o conhecimento de todo registro dos Serviços de Auditoria no que se refere a parte ou todo das informações de entidades ou organizações que acessaram suas informações (AUDITORIA).

7 PROCESSO DE CADASTRO E DEDUPLICAÇÃO

7.1 Introdução

O conceito de deduplicação refere-se ao processo de examinar, durante o procedimento de cadastramento, se a amostra biométrica que está sendo cadastrada possui alguma correspondente em todo o banco de dados já existente. Desta forma, a amostra é comparada as "N" amostras já cadastradas, uma a uma. Se existe alguma amostra correspondente, o indivíduo não é cadastrado, e, portanto, não recebe uma nova identidade a fim de evitar uma entrada duplicada. Caso não exista amostra correspondente, o usuário é cadastrado de forma correta e um número único é associado a amostra apresentada pelo individuo (Decann, 2013).

O processo de deduplicação é necessário para garantir que todos os indivíduos da população tenham apenas um único número no banco de dados. Por exemplo, uma entrada duplicada pode ser criada intencionalmente por um impostor para futuramente fraudar o sistema e obter algum benefício utilizando a identidade de outro indivíduo.

O processo de deduplicação vem ganhando notoriedade e atenção nos últimos anos, devido particularmente a programas de identificação civil de grande escala como o caso da Índia, México e Indonésia, em que enormes bases de dados precisam garantir a unicidade dos indivíduos cadastrados.

A deduplicação também pode ser considerada um processo de identificação, no qual é feito uma comparação 1:N. Entretanto, para cada indivíduo que está sendo cadastrado é realizada uma comparação, o que exige alto poder de processamento do sistema. Como exemplo da complexidade dessa operação, a Índia cadastra, diariamente, 1 milhão de indivíduos, os quais são comparados com os 650 milhões já cadastrados, gerando mais de

500 trilhões de comparações por dia.

7.2 Cadastro e Deduplicação no Sistema RIC

No que se refere ao Serviço de Cadastramento do RIC, um fluxo assíncrono, suportado por mensageria confiável, remete as informações cadastrais às bases do modelo de armazenamento do RIC e a entidade deduplicadora à geração e solicitação de incorporação das bases do RIC de um novo registro (com a intempestiva geração do número do RIC).

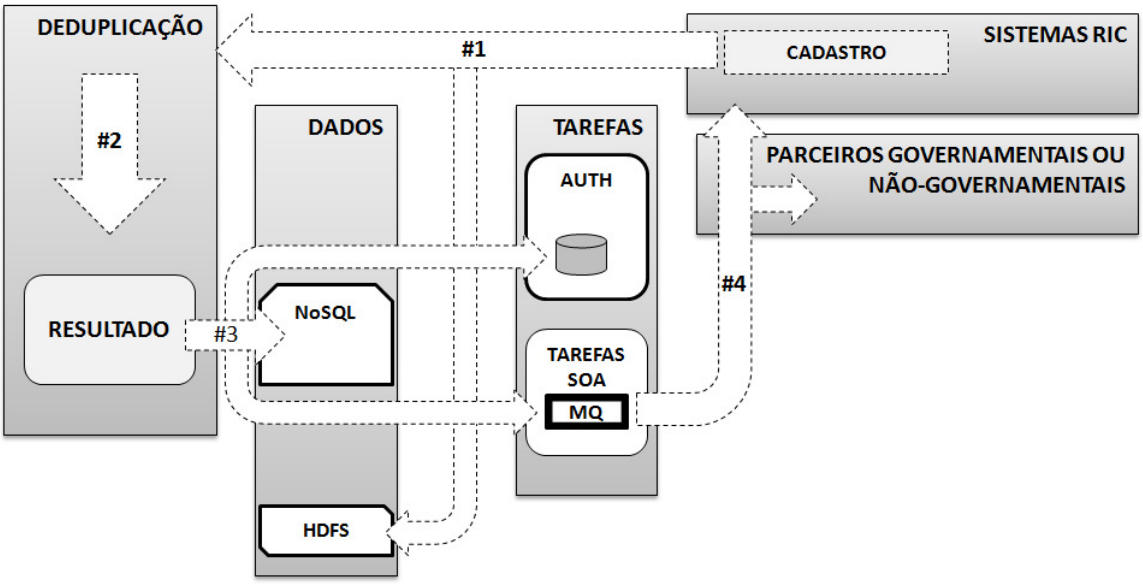


Figura 9 - Fluxo de Dados e Mensagens dos serviços de Cadastro / Atualização

Do modelo ilustrado na Figura 9 têm-se os seguintes fluxos previstos.

- **Fluxo #1**: repasse de informações massivas (imagens) ao modelo de armazenamento de dados do RIC em conjunto com o repasse das informações biométricas a entidade deduplicadora. Essas informações (biométricas e de cadastramento) são encaminhadas associadas com respectivas informações de trilha de auditoria (quem solicita - autorização, de onde - endereço MAC, com que valor de vetor de hora - *vector timestamp*) e, fundamentalmente, com o número de protocolo gerado especificamente com um código *hash* dos dados de auditoria para associação futura com os dados encaminhados às bases de

armazenamento do RIC.

- **Fluxo #2**: realização da deduplicação das imagens em *templates*, suportada por enfileiramento confiável e assíncrono das solicitações.
- **Fluxo #3**: após o exaustivo processo de deduplicação (Fluxo #2), é encaminhado associado com o código *hash* do protocolo (gerado para o Fluxo #1) para registro dos *templates* nas bases de dados do RIC. Somente esse encaminhamento poderá registrar os *templates* nessa base.
- **Fluxo #4**: ao final do processo, isto é, depois de dada uma garantia de seu encerramento efetivo, é atualizada a base de chaves e valores dos Serviços de Autenticação (armazenadores *in-memory*) que está posicionada junto aos componentes do Barramento de Serviços Corporativos e é gerado o evento de GERAÇÃO DO NÚMERO DE RIC junto aos serviços de pub/sub existentes no barramento para geração de mensagem de notificação aos Sistemas de Cadastramento do RIC e aos parceiros governamentais e privados assinantes do serviço de pub/sub.

8 RELAÇÃO ENTRE O ESQUEMA E OS SERVIÇOS

8.1 Autorização

Como é premissa que serviços SOA atendam à requisitos técnicos de composição, todos os serviços projetados para atendimento do RIC devem estabelecer um acoplamento inicial ao processo de autorização (especificamente à uma capacidade de validação).

Esse processo e o esquema para ele idealizado prevê que esses serviços não mantenham estados (*stateless*), sejam extremamente atômicos (outras estruturas devem ser utilizadas para manutenção de estado e de contexto) e seja suportado por modelos de armazenamento de baixíssima latência (no caso, *in-memory* dos pares de chaves / número do RIC e valores / *templates* biométricos). Além disso, para garantir a disponibilidade efetiva desse serviço, o modelo de armazenamento a ser idealizado deverá estar presente junto ao Barramento de Serviços Corporativo, para acesso direto pelo mesmo.

Isto posto, percebe-se que haverá uma Consistência Eventual entre o modelo de armazenamento de dados do RIC e esse conjunto de soluções para alta disponibilidade dos serviços, sendo que a oportunidade dessa eventualidade - já que somente é afetada

por um cadastro ou atualização de dados biométricos - já é ofertada pelo longo tempo de processamento para deduplicação das mesmas.

8.2 Federação (Autenticação)

Tem por objetivo congregar os serviços relacionados à autenticação do usuário, por meio da adoção de um dos tradicionais modelos de gestão de identidade.

JøSANG E POPE (2005) classificam em quatro os modelos de gerenciamento de identidade: o modelo isolado, ou em silo; o federado; o centralizado; e o centrado no usuário. Segundo Torres (2014), o modelo em silo é o mais utilizado, o qual é caracterizado pelo fato de que cada provedor de serviço faz a gerência de sua própria base de usuários, atuando como o seu próprio provedor de identidade. A vantagem deste modelo é a facilidade de implementação por parte dos provedores de serviço e o fato de que o usuário tem a sensação de que os dados pessoais fornecidos estarão visíveis apenas para aquele provedor e para uma finalidade conhecida (Figura 10).

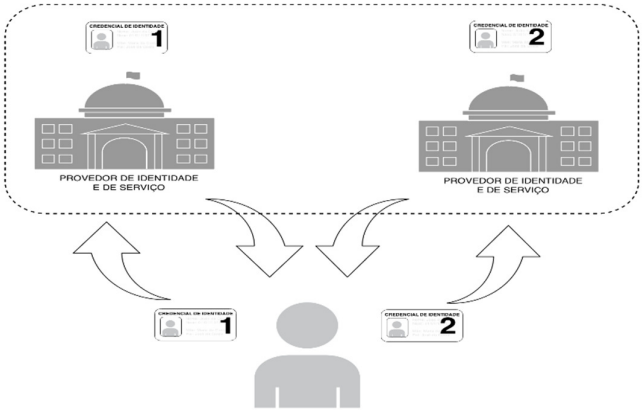


Figura 10. Modelo Isolado ou em Silo (Torres, 2014)

A principal desvantagem vem do fato de que o usuário precisa de um identificador e uma senha própria para cada um dos provedores de identidade/serviço em que efetua transações. Outro ponto negativo é a grande quantidade de senhas a serem memorizadas pelos usuários, levando ao esquecimento das senhas e aumento da complexidade na utilização dos serviços.

Já no modelo centralizado, uma autoridade central atua como único provedor de

identidade do sistema, sendo responsável por gerenciar de forma centralizada os dados dos usuários, identificadores e *tokens* de autenticação. Neste modelo, todos os provedores de serviço irão realizar a identificação do indivíduo validando remotamente as credenciais junto ao IdP, conforme exibido na Figura 11.

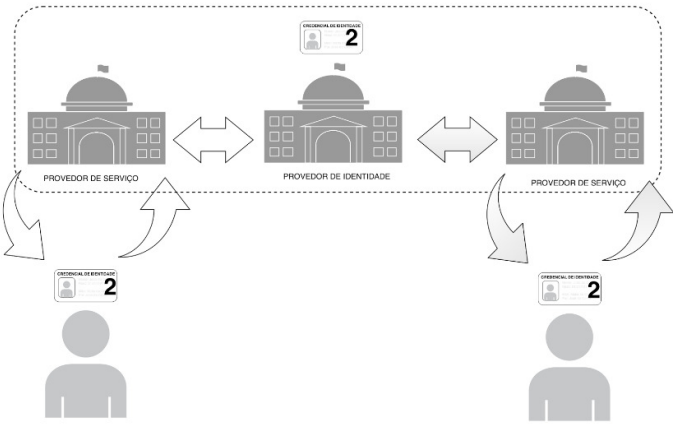


Figura 11. Modelo centralizado (Torres, 2014)

É um modelo de fácil gerenciamento, tanto pela perspectiva do provedor de serviço quanto do usuário, já que o controle dos dados do usuário é centralizado em uma única entidade. A principal desvantagem do modelo é o fato de que os Provedores de Serviço perdem parte do domínio de conhecimento sobre os seus usuários, já que o processo de identificação é realizado por outra entidade e de forma remota.

O terceiro modelo, o Federado, se apresenta como uma alternativa entre os modelos isolado e centralizado. Identidade federada pode ser definida como uma série de acordos, padrões e tecnologias que permitem a um grupo de provedores de serviço reconhecer identificadores e direitos de usuários oriundos de outros provedores de serviço de um domínio federado. Um mapeamento é estabelecido entre diferentes identificadores pertencentes a um mesmo usuário mas em diferentes domínios, de forma a interligar as identidades associadas resultando em um único domínio virtual de identidade (JøSANG E POPE, 2005).

A Figura 12 permite visualizar o conceito do modelo federado. É possível perceber que o comportamento do provedor de serviço consegue permitir o acesso a determinado usuário a partir de uma credencial emitida por um outro provedor de serviço pertencente ao mesmo

domínio de identidade.

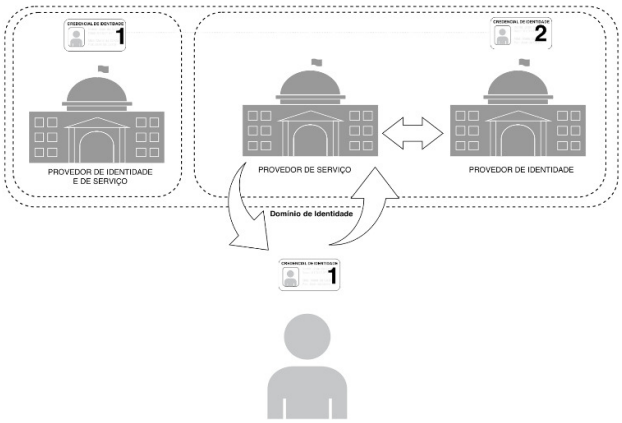


Figura 12. Modelo Federado (Torres, 2014)

O que se observa no modelo federado é que o usuário tem a ilusão de que o processo de autenticação é centralizado, ao tempo que cada provedor de serviço pode ter a sua própria base de usuários, levando por terra as principais desvantagens tanto do modelo isolado quanto do centralizado.

O último modelo apresentado é o centrado no usuário, que tem por principal característica o fato de que o usuário passa a ter maior controle dos seus dados de atributo. BHARGAV-SPANTZEL (2007) diferencia este paradigma em duas diferentes abordagens, uma com foco no relacionamento, no qual o usuário mantém relacionamento apenas com os IdPs e cada transação de transmissão de informações de identidade para um provedor de serviço sempre envolve o IdP apropriado, e uma outra com foco na credencial, no qual credenciais de longo prazo são obtidas do IdP e armazenadas localmente, sendo utilizadas para prover informação de identidade aos provedores de serviço sem, no entanto, precisar envolver os IdPs.

Em uma das propostas de JøSANG E POPE (2005), exibida na Figura 13, o usuário utiliza um dispositivo físico de sua propriedade para armazenar e relacionar as diferentes credenciais para acesso a serviços oriundos de diferentes provedores. É possível notar que tanto o relacionamento entre os diferentes provedores quanto o controle dos dados que estão sendo compartilhados estão sob controle total do usuário.

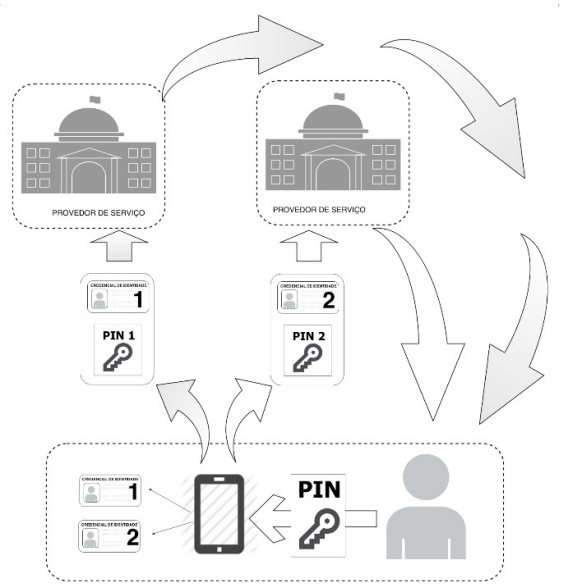


Figura 13. Modelo Centrado no usuário (Torres, 2014)

Mais detalhes sobre o processo de gestão de identidade e sua aplicação no Registro de Identidade Civil serão tratados em Relatório Específico de Levantamento de Diretrizes em Gestão de Identidade.

8.3 Registro e Cadastro

Estes grupos de serviços, além de tratar do armazenamento dos dados biográficos e biométricos impõem outras características e tarefas agregadas, a saber:

- tratamento do tempo de resposta devido à transferência massiva de dados oriunda dos aplicativos clientes;
- encaminhamento coordenado de dados junto a infraestrutura de dados do RIC e a infraestrutura de deduplicação, garantida por uma chave temporária de processo de RIC (status de "em geração de número RIC");
- garantia de acesso restrito oferecida pela infraestrutura subjacente de gestão de identidades federadas.

Do levantamento original do sistema, foram observadas como funcionalidades básicas do CANRIC e apontadas para serem oferecidas por esse grupo de serviços, a saber.

- Coleta de dados biográficos e biométricos do cidadão *off-line*.

- Processamento de dados biográficos e biométricos do cidadão *off-line*.
- Verificação dos dados biográficos e biométricos coletados.
- Registro de informações (número RIC, minúcias, dados biográficos).
- Requisição da individualização biométrica do cidadão AFIS/DPF (encaminhamento coordenado de dados junto a infraestrutura de dados do RIC e a infraestrutura de deduplicação) - que se dá de forma automática pela aplicação cliente.
- Cancelamento da solicitação de emissão do Documento RIC.
- Rotina para expirar o Documento RIC.

8.4 Informação e Integração

Esses dois modelos baseiam-se no princípio básico de mensageria requisição-resposta, considerada suas duas origens: a infraestrutura de serviços RIC solicitando informações de entidades parceiras, ou o acesso por organizações governamentais e não-governamentais, apenas para leitura de informações existentes no armazenamento do RIC.

As solicitações oriundas do RIC são empilhadas em uma fila de requisições, suportada por dinâmica de mensageria confiável, o que imputa que os parceiros do modelo implementem algum tipo de mecanismo de consulta a esses dados solicitados como serviço. Não haverá nenhuma cobrança de garantia de entrega de informações pelos parceiros, sendo imposto a fila o suporte a um mecanismo de tentativas tempestivas o caso de falha nos parceiros.

As solicitações aos serviços do RIC se dão de forma síncrona ou assíncrona, diretamente aos serviços do RIC, estando o solicitante devidamente registrado no serviço de identidade federada para tal fim.

Do levantamento original do sistema, foram observadas as seguintes funcionalidades básicas do CANRIC e apontadas para serem oferecidas por esse grupo de serviços.

- Disponibilidade de serviços para interoperação com os sistemas de Órgãos de Identificação (requisição do Sistema RIC)
- Disponibilidade de serviços para interoperação com os sistemas de Órgãos de Identificação (requisição de um Parceiro)

- Histórico de dados biográficos e biométricos do cidadão

8.5 Comunicação

Modelo de processos de mensageria direcionada por eventos (*publish-subscribe*) gerenciados pelo ambiente RIC e acionado por ocorrências específicas nos seus dados armazenados (ex.: extinção de RIC por falecimento). Esse modelo será suportado por arranjos tecnológicos entre os gestores do RIC e definidos parceiros governamentais ou não governamentais, para encaminhamento de uma mensagem de notificação para esses com garantia de entrega (algum suporte de mensageria confiável). É importante que esse modelo registre consistentemente a informação que a mensagem foi entregue corretamente e em tempo aos parceiros assinantes (*subscribers*).

Do levantamento original do sistema, foram observadas as seguintes funcionalidades básicas do CANRIC e apontadas para serem oferecidas por esse grupo de serviços.

- Notificação de problemas nos dados de solicitação de emissão do RIC

8.6 Atendimento

Para implementação deste modelo sugere-se o uso de serviços básicos de gerenciamento e execução de processos de negócio baseados em tarefas orquestradas junto a motores de execução BPM/BPEL providas por plataformas profissionais de SOA.

Do levantamento original do sistema, foram observadas como funcionalidades básicas do CANRIC e apontadas para serem oferecidas por esse grupo de serviços, a saber.

- Criação e manutenção da agenda de atendimentos dos Postos de Identificação
- Agendamento da solicitação de emissão do RIC
- Cancelamento do agendamento da solicitação de emissão do RIC

8.7 Operacionalização

Para implementação deste modelo sugere-se o uso de serviços básicos de gerenciamento e execução de processos de negócio baseados em tarefas orquestradas

junto a motores de execução BPM/BPEL providas por plataformas profissionais de SOA.

Especificamente para os serviços de operacionalização, é observado que diversas de suas capacidades devem ser implementadas pela execução, ao nível do sistema operacional, por meio de rotinas em "*batch*".

Do levantamento original do sistema, foram observadas as seguintes funcionalidades básicas do CANRIC e apontadas para serem oferecidas por esse grupo de serviços.

- Registro de informações de controle de solicitações de emissão do RIC e do Documento RIC.
- Registro de envio/recebimento do arquivo ao fornecedor de impressão / personalização do Documento RIC.
- Geração do lote de números tipográficos do Documento RIC (número de suporte)
- Registro de recebimento das remessas contendo os Documentos RIC confeccionados.
- Registro de informações do certificado digital emitido pela AC
- Requisição de emissão e/ou revogação do certificado digital a AC
- Entrega do Documento RIC (*on-line* e *off-line*)

8.8 Auditoria

Como um elemento presente nos Barramentos de Serviços Corporativos, pressupõem-se o uso do processo de Monitoramento de Atividades do Negócio - BAM, em conjunto com os padrões de mensageria unidirecionais BAM, na automatização das capacidades de registros de trações de log e na formação de saídas para controle da capacidade e desempenho do sistema.

Assim, esses serviços podem ser supridos pelo suporte de infraestrutura subjacente (Barramento Corporativo de Serviços - ESB) em modelos de filas unidirecionais no *middleware* (*Notify Services by Unidirectional Queuing*). Neste caso, mensagens unidirecionais serão geradas de forma a possibilitar aos modelos de serviços dos RIC (como um todo) a capacidades de arquivamento e auditoria. Adicionalmente, *triggers* e/ou serviços *publish-subscribe* devem ser implementados visando atuar após o registro desses

eventos em situações ocasionais de identificação - por máquina - da ocorrência de atividades não autorizadas (observada durante o processo de registro de auditoria).

Do levantamento original do sistema, foram observadas como funcionalidades básicas do CANRIC e apontadas para serem oferecidas por esse grupo de serviços, a saber.

- Registro do login de operações realizadas pelo usuário do sistema

8.9 Sincronismo

Estabelece dois grupos de serviços para sincronismos de dados (de e para) de forma a atender quaisquer solicitações de sincronia a determinados dados ou recursos digitais (imagens de digitais, faces, íris e assinaturas) entre o RIC e seus parceiros regionais. Ainda não existem maiores detalhes que possam ser tratados nssas realizações.

8.10 Gerencial

Refere-se às atividades voltadas a disponibilização de serviços de acesso a cubos multidimensionais, tratando dos assuntos:

- dos banco de dados (em memória) dos cubos multidimensionais;
- da consistência eventual desses cubos multidimensionais por processo;
- da permissão de acesso a dados globalizados para parceiros e outros;
- da permissão de acesso ao Sistema RIC-BI que prevê com recursos/saídas, a saber.
  - Relatórios Formatados: de acesso periódico ou intempestivo, podendo ser acessado pelo usuário ou encaminhado diretamente (anexos de *e-mail*).
  - Consultas *Ad-hoc*: criadas sob demanda especificamente com conteúdo, *layout* ou cálculo necessário e que podem agilizar ou facilitar uma tomada de decisão pontual.
  - Análise Multidimensional: suportado diretamente pelo uso de ferramentas OLAP, permite ao usuário interagir com o resultado por meio de processos de análise utilizando-se de operações multidimensionais (*drills*, *pivot*, etc.). Em geral, o analista opera buscando situações de:
    - analisar dados contextualizados;

- dar suporte ao trabalho com hipóteses; e/ou
- procurar relações de causa e efeito.

Do levantamento original do sistema, foram observadas as seguintes funcionalidades básicas do CANRIC e apontadas para serem oferecidas por esse grupo de serviços.

- Disponibilização de consultas e relatórios gerenciais na Base Nacional do Registro de Identificação Civil

9 DOCUMENTO BÁSICO DE MODELAGEM DE PERFIL DE SERVIÇOS

A evolução deste relatório irá apontar as necessidades tecnológicas e documentais, as quais servirão de base para a construção dos documentos que definem os elementos tecnológicos necessários para o barramento corporativo de serviços (RT de Arquitetura Orientada a Serviços) e para suporte inicial a análise e modelagem de serviços. Durante a fase de análise e modelagem é necessária a formação de documentação referente ao perfil de serviços necessários similar ao modelo apresentado a seguir.

9.1 Perfil de Serviço

Nome do Serviço	Declaração do nome do serviço no formato “<nome>Service”.
Inventário/Domínio	Sigla/Nome do domínio de inventário
Modelo do Serviço	Modelo do serviço, conforme definido no “"RT de Soluções de *Message Queue* e Barramentos de Serviço Corporativos".
Descrição	Breve descrição do serviço e suas funcionalidades.
Propósito	Declaração detalhada do propósito do serviço.
Palavras-chaves	Lista de palavras-chaves associadas aos serviços, separadas por “;”.
Versão	Versão atual do serviço no formato x.y
***Status* do Serviço**	Situação do serviço (Candidato, Projeto, Implementação, Homologação, Produção, Aposentado).
Data de Entrega	No caso de serviços no *status* <Produção> assume o valor de entrada em produção. Em outros casos assume a data provável de sua entrega.

Desenvolvedores	Nome, telefone e *e-mail* do analista, do projetista e do desenvolvedor principal do serviço.
Guardião	Nome, telefone, *e-mail* e organização do guardião do serviço.
Contrato do Serviço (Interface Técnica)	Link para o artefato de interface técnica do contrato de serviço (e.g. WSDL).
Esquemas de Dados (XSD)	Link para os artefatos de especificação de esquemas de dados utilizados no serviço.
Políticas (WS-*Policy*)	Link para os artefatos de especificação de políticas de serviço.
Mensagens	Nomes das mensagens, separadas por ";", com seus respectivos links para perfis de mensagens.
Capacidades	Nomes das capacidades, separadas por ";", com seus respectivos links para perfis de capacidades.
***Bindings* do Serviço**	Define como o serviço pode ser acessado (*endpoints*). No caso de múltiplas instâncias, separá-las por ";".
Design de Lógica de Serviço	Link para o artefato *Design de Lógica de Serviço*.
SLA	Link para o artefato *Especificações Adicionais e SLA.*
Manual de Uso	Link para o artefato *Manual de Uso*.
Composição (diretas)	Lista de serviços que o serviço compõe diretamente.
Dependências	Listas de sistemas dos quais o serviço depende ou consome funcionalidades (bancos de dados, sistemas legados, outros serviços).
Consumidores conhecidos	Lista de sistemas consumidores do serviço.
Histórico de Versões	Descritivo do histórico de versões com links para os respectivos artefatos de *Perfil de Serviço*.

9.2 Perfil de Capacidade de Serviço

Nome da Capacidade	Declaração do nome da capacidade do serviço.
Situação de Terminação	Identifica com "C" se o serviço realiza alguma chamada de composição ou mediação, e com "F" se é capacidade que apenas executa lógica (autocontida).
Descrição	Breve descrição da capacidade e suas funcionalidades.
Propósito	Declaração detalhada do propósito da capacidade.
***Status* da Capacidade**	Situação atual da capacidade
Data de Entrega	No caso de capacidades de serviços no *status* <Produção> assume o valor de entrada em produção. Em outros casos assume a data provável de entrega.
Contrato do Serviço	Link para o artefato de interface técnica do contrato de serviço (e.g. WSDL).
Endereço da Capacidade (*Address*)	Endereço de como a capacidade deve ser chamada.
Composição (diretas)	Lista de capacidade / serviço que o serviço compõe diretamente.

Mensagem de Entrada	Nome da mensagem, com respectivo link para os *namespaces* / esquemas de dados / tipo de dados (complexos ou não) utilizados pela mensagem como "*input*".
Mensagem de Saída	Nome da mensagem, com respectivo link para os *namespaces* / esquemas de dados / tipo de dados (complexos ou não) utilizados pela mensagem como "*output*".
Mensagem de Falha	Nome da mensagem, com respectivo link para os *namespaces* / esquemas de dados / tipo de dados (complexos ou não) utilizados pela mensagem de falha "*messagefault*".

9.3 Perfil de Mensagem (Opcional)

Nome da Mensagem		Nome do Pacote de Mensagem			
Uso Primário		[REQUEST][RESPONSE]			
Descrição (*Annotation*)		Definição do Pacote			
XSD		Nome com a localização do arquivo XSD			
Default		[qualified]			
Target Namespace		"targetnamespace"			
Versão		Versão do Esquema de Mensagem no formato x.y			
XMLNS		Lista separada por "," dos XML *namespaces*			
Import		Lista dos *imports* realizados separados por ","			
Usado nos serviços		Lista dos XML *namespaces* separados por ","			
Tabelas Relacionadas		Lista de tabelas com seus BD separados por ","			
Tipo		[sequence]			
Nome	**Supertipo**	**Ocorrências**	**Valores**	**Formato**	**Descrição**
Nome da Variável	Link para Supertipo	Limites da Instância no formato {0,1,m:0,1,n}	Limitação de valores em lista separados por [;]	<Máscara>	<Descrição do metadado>

9.4 Perfil de Transformação (Opcional)

Nome da Mensagem		Nome do Pacote de Transformação
XSD associado		Nome do Pacote de Mensagem (modelo canônico associado) de entrada e o de saída
Descrição (*Annotation*)		Definição do Pacote
XSLT		Nome com a localização do arquivo XSLT
Versão		Versão do Esquema de Transformação no formato x.y
NOME-ENTRADA	**NOME-SAIDA**	**TRANSFORMAÇÃO**

Nome e Tipo da Variável de Entrada	Nome e Tipo da Variável de Saída	Transformação realizada.

10 MODELO DE ARMAZENAMENTO e MODELO CANÔNICO DE DADOS

Recentemente, a imprensa especializada e a academia vêm discutindo os novos paradigmas associados à computação distribuída, principalmente no que se refere ao armazenamento e processamento distribuído e paralelo para tratar pesquisa em dados massivos. Este paradigma implica em aproveitamento de grande número de recursos computacionais que trabalham em paralelo para resolver essa problemática.

Essa discussão vem sendo construída a partir do surgimento das técnicas de *MapReduce* e os armazenadores NoSQL que revêm o conceito ACID (principalmente de integridade por uma consistência eventual). A técnica de *MapReduce* é atraente porque oferece uma modelo simples por meio do qual os usuários podem expressar suas lógicas de armazenamento e processamento distribuído de maneira sofisticada, o que vem levando a um interesse significativo no educativo da comunidade.

Dado este interesse em *MapReduce*, é natural que se pergunte: "Por que não usar um SGBD paralelo em vez disso?”

Sistemas de banco de dados paralelos (que todos compartilham um projeto arquitetônico comum) estão comercialmente disponíveis há quase duas décadas, e agora existem cerca de uma dúzia de mercado. Eles são robustas plataformas de computação de alto desempenho.

Assim, com foco a atender o problema, as duas abordagens devem ser discutida, lembrando que SGBD relacionais requerem que o armazenamento estejam em conformidade com um esquema bem definido, ao passo que modelos de armazenamento NoSQL podem até ter um formato arbitrário ou sem esquema - o que não é nossa necessidade. Outras diferenças incluem também em como cada sistema fornece indexação, otimizações por compressão, modelos programacionais de acesso ao dados e, principalmente, a maneira pela qual os dados são distribuídos e as estratégias de execução de consultas.

Neste foco, será iniciado por um tópico correspondente sugerindo o uso de SGBDs

relacionais distribuídos, e outro tratando sobre armazenadores NoSQL. Nessas discussões serão apresentadas as necessidades do modelo arquitetural proposto, porém a decisão desse uso passa pelo estudo direto com comparativo dos SGBDs/Armazenadores NoSQL a ser detalhado em outro Relatório Técnico.

10.1 Modelo de Banco de Dados Relacionais Paralelo e Distribuído para o RIC

Sistemas de banco de dados capaz de rodar em *clusters* já existem desde o final de 1980. Estes sistemas usam o modelo relacional padrão de tabelas com suporte de SQL, e assim, o fato de que os dados são armazenados em várias máquinas é transparente para o usuário final. Os dois aspectos fundamentais que permitem a execução paralela são de que (1) a maioria (ou mesmo todas) das tabelas são divididas ao longo dos nós em um *cluster* e que (2) o sistema usa um otimizador que traduz comandos SQL em um plano de consulta cuja execução é dividida entre vários nós. Assim, o modelo de consultas é especificado em uma linguagem de alto nível, fazendo que não sejam sobrecarregados com os detalhes de armazenamento subjacentes, tais como opções de indexação.

Desse modo, SGBDs paralelos usam o conhecimento da distribuição dos dados por meio de suas localizações como vantagem. Isto se dá na existência de um otimizador de consulta paralela que se esforça para equilibrar as cargas de trabalho computacional, minimizando a quantidade de dados transmitidos pela rede e que conecta os nós do *cluster*.

Isto permite junções e tratamento de dados particionados, podendo ser executados em paralelo (Osthoff et al., 2000) em todos os nós. Após o tratamento de cada nó, existe um modelo de agregação a fim de realizar a resposta para a associação. Assim, um passo final de "*roll-up*" é necessário para calcular a resposta final a partir destes agregados parciais (DeWitt and Gerber., 1985).

Neste sentido, SGBDs distrbuídos proveem algum mecanismo de particionamento horizontal, em geral na forma de contextos (context, namespaces and range partitioned base table), nos quais a raiz de distribuição é um campo, ou um conjunto agregado de campos, de cada SGBDs.

Para esse paralelismo, o projeto para processamento de regras em um ambiente de banco de dados distribuído envolve, basicamente, três questões:

- como decompor uma condição de regra;
- como distribuir os dados e uma condição de regra entre os diferentes sites;

- como avaliar uma condição de regra num ambiente distribuído e garantir sua consistência e exatidão numa avaliação distribuída.

Com base na distribuição das tabelas do banco de dados, transações do usuário são distribuídas para executarem em múltiplos sites ao mesmo tempo. Como premissa básica, o efeito da execução da transação distribuída deve ser equivalente ao efeito da transação se a mesma estivesse sendo executada em um banco de dados centralizado, com as mesmas características do banco de dados distribuído.

Adicionalmente, como vantagem desses modelos, todos esses usam modelos de *hash* ou árvore-B para formação de índices para acelerar o acesso aos dados. Se alguém está procurando um subconjunto de registros será um índice adequado que reduz o escopo da pesquisa de forma dramática. A maioria dos sistemas de banco de dados também suportam vários índices por tabela. Assim, o otimizador de consulta pode decidir qual o índice a ser usado para cada consulta ou simplesmente executar uma força bruta e busca sequencial.

Índices com particionamento local têm sido adicionados como suporte para o banco de dados. Um índice particionado local é aquele particionado em uma tabela particionada com um mapeamento um-para-um de partições de tabela. É como dividir a tabela maior em tabelas menores e seus índices, do subconjunto dessa tabela.

10.2 Modelo de Armazenamento de Dados NoSQL para o RIC

Portanto, sugere-se que o projeto RIC realize a implementação de um modelo que preveja um particionamento horizontal com distribuição baseada nos princípios de escalabilidade geográfica e computação distribuída com sincronismo de dados por arquitetura orientada a eventos (*Staged Event Driven Architecture*). No uso de particionamento horizontal, sugere-se a aplicação de princípios de *Map/Reduce* visando o suporte adequado de desempenho no modelo de armazenadores de dados distribuídos.

Sugere-se também que o projeto realize a implementação de um modelo que preveja um particionamento vertical, garantindo acesso privilegiado fisicamente isolado no contexto de dados. Isto se dá prevendo o uso de modelos de armazenamento baseados em colunas (*Column-Store*). Em modelos de armazenamento em colunas usuais, existem elementos agrupadores de colunas definidos como famílias de colunas (*Column Families*). A análise de base dessa sugestão encontra-se detalhada no documento "RT de Infraestrutura Tecnológica: Armazenadores NoSQL".

A Figura 14 ilustra a estrutura interna de uma base de dados colunar, no caso o HBase, em que a família da coluna é simplesmente uma unidade de organização em que grupos relacionados de colunas são armazenados em conjunto, isto é, em estreita proximidade no disco. Essas famílias de colunas podem ser dados dinâmicos ou identificadores estáticos. Na aplicação sugerida, essa situação deve ser apontada como identificadores estáticos a fim de permitir modelar esses dados em um formato relacional. Como exemplo, no armezenador HBase os nomes de família de coluna são sempre estáticos identificadores, tal qual os nomes das colunas de uma tabela relacional. Já o armazenador Cassandra permite esses dois tipos diferentes de família de colunas. A Figura 10 ilustra um modelo de armazenamento com famílias de colunas.

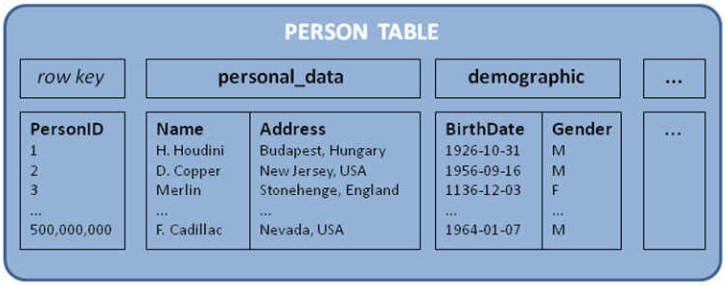


Figura 14 – Modelo Lógico de um Armazenamento Orientado a Colunas

Obs.: *Porém, esse uso do princípio de particionamento vertical deve ser considerado com cautela, pois pode comprometer o desempenho se as partições forem muito grandes.*

Neste sentido, outro modelo possível é o uso de soluções modernas de SGBD's distribuídos, tal como modelo denominados NewSQL. NewSQL é uma classe de sistemas de gerenciamento de banco de dados relacionais modernas, que buscam oferecer o mesmo desempenho escalável de sistemas NoSQL para processamento de transações *online* (OLTP) de leitura e escrita de cargas de trabalho, mantendo as garantias ACID de um sistema de banco de dados tradicional (Aslett, 2011; Hoff, 2012; Lloyd, 2012).

A) DadosRIC

- **(Família de Colunas) Determinante**
    - numeroRIC: varchar(26)
    - BiometricaPrimáriaRIC
    - BiometriasSecundáriasRIC [unbounded]

- **(Família de Colunas) Biográficos**
    - DadosRIC: dadosRIC
    - IdentificacaoRIC: identificacaoRIC [unbounded]
- **(Família de Colunas) Relacionamentos**
    - FamiliaresRIC: relativoRIC [unbounded]
- **(Família de Colunas) Demográficos**
    - ProfissionalRIC: demograficosRIC [unbounded]

B) AuditoriaRIC

- AuditoriaRIC: auditoriaRIC [unbounded]

10.3 Modelo Estrutural Básico:

Figura 15 – Famílias de Colunas para o RIC e suas escalas de acesso

a) **Armazenamento de Dados do RIC**

- Determinante
- Biográficos
- Relacionamentos
- Demográficos

Obs: É provável uma implementação de acesso ao dados Demográficos pelo próprio cidadão na atualização de dados dessa família de colunas.

b) **Armazenamento de Dados para suporte ao RIC**

- Auditoria

c) **Armazenamento de Dados das MatrizesDigitais do RIC**

- Imagens

10.4 Modelagem Canônica de Dados

A troca de mensagens, seu modo de transmissão e a semântica são partes fundamentais em sistemas distribuídos, pois são a única maneira que os componentes de uma arquitetura orientada a serviços possuem de se comunicar e sincronizar suas ações.

Mesmo reconhecendo que os esquemas canônicos representem uma visão *front-end* do modelo de dados - pois eles existem para definir conceitos e agir como mediadores ideais - o que se idealiza é que o arquétipo de armazenagem de dados seja par-a-par com a Modelagem Canônica da Mensageria que será usada como base dos modelos de mensagens utilizados na troca interoperável de dados entre o RIC e os demais sistemas ou consultas existentes pelos sistemas externos por meio de terminais oficiais (*endpoints*). Isto facilitaria sobremaneira a recuperação e encaminhamento dos resultados, nos quais os terminais dos serviços (expressos pelas suas capacidades) não implementem nenhuma lógica que agregue latência aos resultados de consultas ao RIC.

Porém, mesmo sugerindo esse acoplamento entre modelo de dados e mensageria, este trabalho desacopla o desenvolvimento da semântica do modelo de mensageria do projeto de serviços de forma a garantir uma independência entre dados e capacidade dos serviços a serem projetados durante a etapa de análise, impondo um baixo acoplamento entre o fluxo de dados e os requisitos sistêmicos no projeto das capacidades candidatas de serviços.Do processo de modelagem canônica das mensagens (como estabelecido na literatura) e seus resultados finais, têm-se os seguintes padrões de projeto usuais.

- *Russian Doll*: é baseado na redução do tamanho do esquema da mensagem pela colocação de um esquema dentro do outro. O padrão de *Russian Doll* define todos os seus subelementos localmente, assim cada elemento e o seu tipo são encapsulados por elementos-pais. São considerados altamente dissociados (os elementos não são globalmente dependente em outros elementos) e coesos (os elementos são agrupados em um único elemento-pai independente). Este padrão sintetiza esquemas destinados a ter pouca interação com outros sistemas e sem reutilização dos seus componentes.
- *Salami Slice*: esse padrão já expõe o conteúdo dos modelos, pois são movidos todos os seus elementos definidos localmente em definições globais.
- *Venetian Blind*: esse estilo utiliza definições de tipo global para aumentar a

reutilização das capacidades dos esquemas, porque todos os subelementos são localizados, ele vem com vantagem adicional de ser capaz de esconder os elementos dentro do espaço de nomes (*namespaces*). Esta abordagem permite que se exponha as definições de estrutura para reutilização usando-se do *elementFormDefault,* funcionando como um interruptor para ocultar ou expor todos os elementos no espaço de nomes (*namespaces*).

- *Garden of Eden*: assim como a abordagem SalamiSlice esse modelo é modular e estruturado por meio da definição de todos os elementos a nível global. Como a abordagem *Venetian Blind*, nesse modelo todas as definições de tipo também são declaradas globalmente. Desse modo, cada elemento é definido globalmente como um filho imediato do nó (“<schema>”) e os tipo de elementos (atributos) podem ser definido como um dos tipos complexos nomeados globalmente.

No projeto efetivo dos serviços é de nossa sugestão o uso da abordagem de *Garden of Eden* que implementa características da abordagem *Venetian Blind* com a *Salami Slice*. A vantagem dessa abordagem é que os esquemas são reutilizáveis, pois ambos, os elementos e os tipos, são definidos globalmente, e assim ambos estão disponíveis para reutilização. Neste sentido, esta abordagem oferece o máximo de conteúdo reutilizável.

10.5 Mensageria Canônica Sugerida para a Prova de Conceito do RIC

mensagemRIC:

- codigoConfirmacao
- codigoProcesso
- codigoErro
- mensagemErro

biometricaRIC:

- tipoBiometria
- dadoBiometriaBruta
- Template

dadosRIC:

- nome

- dataNascimento
- nomeMae
- ufNascimento
- situacao (ativo, extinto, canceladom, etc)

identificacaoRIC:

- cpf
- cnh
- ctps
- identidade
- nis
- passaporte
- pispasep
- tituloEleitor
- certidao
- certificadoNaturalizacao

relativoRIC:

- TipoFamiliar
- chaveRICFamiliar

demograficosRIC:

- endereco
- celular
- email

AuditoriaRIC

- numeroRIC
- timestamp
- origem
- serviço

11 SEGURANÇA: CONFIDENCIALIDADE E PRIVACIDADE

Pela implementação do modelo do RIC em armazenadores em coluna e com o uso de família de colunas é possível ainda o uso de esquemas de criptografia em camadas dependendo da família de colunas (*onion model*) ou em lógicas baseadas em atributos (*attribute-based encryption* - ABE) por família de colunas. O uso de padrões de

comunicação criptografados (SS) e a aplicação de criptografia também no armazenamento colunar pode garantir a segurança no acesso a dados por perfis.

Além disso, pode ser implementado recursos modernos de *proxy* de criptografia, o que pode permitir que os dados do RIC sejam alocados em terceiros que não terão acesso aos dados persistidos em nenhum momento do processo, garantindo uma forte política de privacidade. A Figura a seguir ilustra o modelo de exposição dos serviços do RIC com segurança e privacidade.

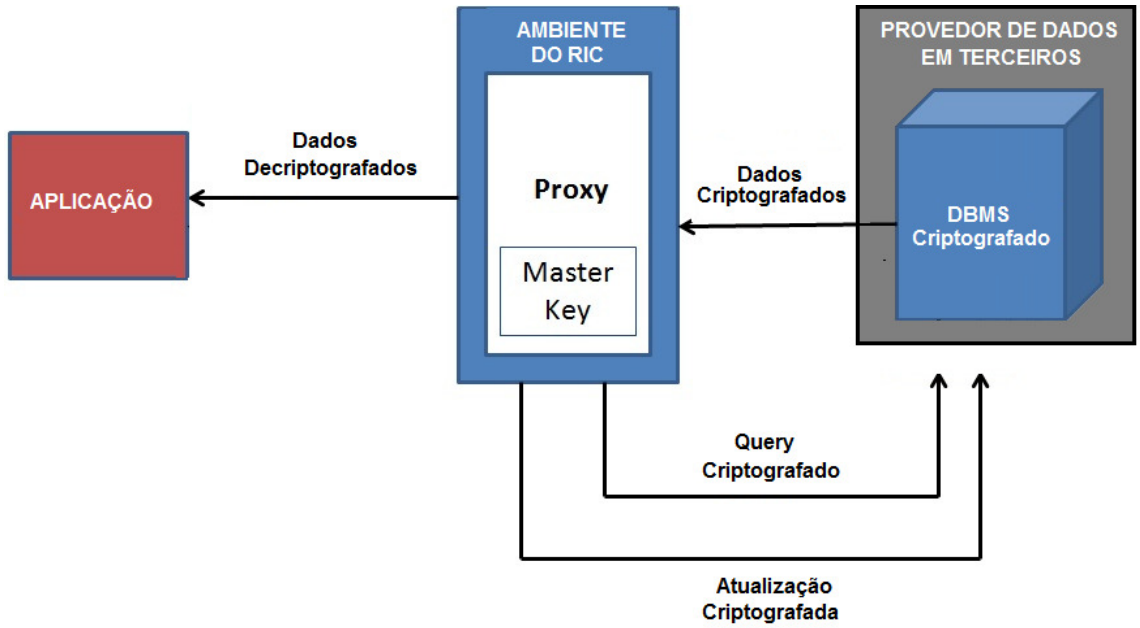


Figura 16 – Segurança através de Proxies de Criptografia

11.1 Criptografia Baseada em Atributos

De Silva, Araujo e Cerqueira (2014) afirmam:

(...) "Por causa das limitações desses métodos tradicionais, Sahai e Waters (2005) criaram a primeira versão ou forma do ABE como objetivo de oferecer segurança via criptografia, além de controle de acesso. Esse recurso criptográfico permite que dados sejam cifrados com base em uma lista de atributos. Múltiplos usuários podem decifrar os dados contanto que suas chaves (que são distintas) estejam associadas a um número mínimo, pré-estabelecido, de atributos que também foram utilizados no momento da cifragem. Isso permite um compartilhamento mais intuitivo e sem as dificuldades mencionadas anteriormente em métodos tradicionais".

Os atributos utilizados no ABE podem representar características reais de usuários, como papeis, perfis, identificadores, profissão, etc. A disponibilização desses atributos, para que possam ser utilizados em qualquer operação de cifragem de dados, é de responsabilidade de uma Autoridade de Atributos (AA). Essa AA também é responsável pela criação de chaves. Assim, ela requer confiança.

Posteriormente, Bethencourt et al. (2007) propuseram uma adaptação para o ABE, chamada de CP-ABE (*Ciphertext-Policy Attribute-Based Encryption*). No CPABE, ao invés de requerer uma lista de atributos, é necessário uma política de acesso baseada em atributos no momento da cifragem. Nessa proposta, a chave continua sendo associada a uma lista de atributos. A Figura a seguir ilustra um modelo baseado em atributos:

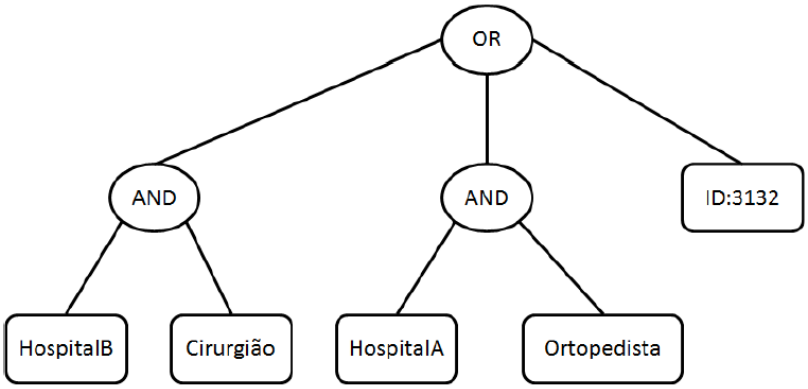


Figura 17 – Modelo de Árvore de Atributos

11.2 Criptografia no Armazenamento de Dados

Com o intuito de impedir a invasão e descoberta de dados sigilosos em sistemas de identificação, uma nova abordagem é apresentada para proteger a confidencialidade dos dados mesmo quando os atacantes têm acesso aos dados do servidor, utilizando sistemas que realizam consultas em dados criptografados sem acesso a chave de decodificação. O modelo proposto é aplicado em um cenário real: um sistema distribuído hospedado pelo Google (*Encrypted BigQuery*), o qual envolve o uso de criptografia em dados que são armazenados em bancos de dados não-relacionais (NoSQL) desenvolvidos para o tratamento massivo de dados. Os resultados mostram que suportam uma variedade de aplicações com baixo custo operacional.

Entende-se que a segurança desses dados armazenados e das informações transmitidas no seu acesso por meio de qualquer sistema é extremamente importante, pois é ela que garante que o sistema conseguirá fazer aquilo pelo qual foi projetado, de maneira correta, e disponibilizada apenas para usuários autorizados. Muitas informações mantidas são sensíveis e sigilosas, portanto essa proteção é fundamental e consiste em atender aspectos de confiabilidade, integridade, disponibilidade, autenticidade e irretratabilidade. Na transmissão de informações, a criptografia de canal vem sendo o mecanismo mais utilizado e, dependendo do sigilo imposto às chaves utilizadas, é eficiente e eficaz.

Assim, a preocupação com a criação e manutenção de ambientes seguros se tornou a ocupação principal de administradores de redes, de sistemas operacionais e de bancos de dados. Esta preocupação é reforçada, pois estudos comprovam que a maioria dos ataques relatados, de roubos de informações e acessos não autorizados, é feita diretamente na origem dos dados (Unisys, 2014), isto é, junto aos seus bancos de dados, ou por pessoas pertencentes a organização ou pela falta de mecanismos que evitem a leitura direta das bases de dados remotamente por terceiros. Por esse motivo, as organizações se esforçam para criar e usar artifícios com a finalidade de eliminar os acessos não-autorizados ou diminuir as chances de sucesso das tentativas de invasão (internas ou externas).

Dos trabalhos relacionados, Song et al.(2000) descrevem esquemas criptográficos para efetuar uma busca por palavras-chave em dados cifrados usando um servidor inseguro. Estes esquemas são usados na composição deste trabalho. O esquema é simples e rápido, para um documento de tamanho n, os algoritmos de criptografia e busca precisam de apenas O(n) operações de cifra de fluxo e de bloco. Um método para buscas exatas que não requer a leitura de toda a base de dados é apresentado em Amanatidis et al. (2007). Bao et al. (2008) aprimora as técnicas de busca de Amanatidis em dados criptografados para o caso de multiusuários. Boneh & Waters (2007) apresentam um modelo com esquemas de chaves públicas para comparação, checagem de subconjuntos e consultas de conjunção em dados cifrados, porém esses esquemas têm cifras de tamanho exponencial em relação ao texto em claro, limitando a sua aplicação prática. Em 2009, Boldyreva et al. apresentam um modelo de criptografia simétrica com preservação de ordem, equivalente a um mapeamento aleatório que preserva a ordem.

No que se refere as implementações existentes, o CryptDB (Popa et al., 2011) é um

sistema que provê confidencialidade prática e comprovada para aplicações baseadas em bancos de dados SQL. Evita a criptografia totalmente homomórfica, no entanto, processa as consultas usando uma coleção de esquemas criptográficos eficientes admitidos pela configuração SQL (Popa et al., 2011). CryptDB direciona as ameaças passivas para o SGBD e as ameaças ativas para o Proxy, que atua como um agente mediador entre o cliente e o SGBD. As consultas vindas do cliente são interceptadas pelo *Proxy* que as reescreve de forma que possam ser executadas em dados cifrados. O servidor efetua as consultas e retorna os textos cifrados correspondentes para o *Proxy*. O *Proxy* ,então, decifra as cifras recebidas e envia os resultados para o cliente.

Ainda, em 2013, Arasu et al. apresentam outra implementação, o *Cipherbase*, um sistema de banco de dados SQL completo que permite as organizações utilizarem as vantagens da computação em nuvem e ao mesmo tempo manterem a confidencialidade dos dados sensíveis. *Cipherbase* é baseado em uma arquitetura de coprocessadores com o objetivo de decompor o processamento entre o *hardware* tradicional (inseguro) e o confiável. Utiliza a combinação de criptografia de dado estático (*encryption at rest*), servidores seguros e o uso de criptografia parcialmente homomórfica para atingir segurança ortogonal, isto é, permitir que as organizações desenvolvam suas aplicações e configurem suas políticas de nível de segurança independentemente de qualquer desempenho, escalabilidade ou custo.

Porém, como citado, é requisito deste trabalho a implementação de segurança em modelos de armazenamento orientado a colunas, do Google *BigQuery*, sendo que a Google provê este mecanismos através de uma interface de acesso criptografado, o *Encrypted BigQuery* (EBQ).

O EBQ é uma interface que permite a criptografia dos dados por parte do cliente seguido do carregamento (*upload*) para o *BigQuery* (Tigani & Naidu, 2014). A fonte não cifrada e a chave usada na criptografia não são enviadas pela rede, assim, não há a possibilidade de qualquer administrador do *BigQuery* obter acesso aos dados.

O funcionamento do EBQ se inicia com a construção do arquivo esquema, que indica qual será a estrutura dos campos presentes da tabela a ser carregada. EBQ inclui um campo extra (*encrypt*), o qual determina os tipos de esquemas criptográficos que serão usados. Feito isso, os dados são carregados de maneira semelhante ao *BigQuery*, especificando o arquivo da chave usada e seguindo os esquemas criptográficos apontados

no campo *encrypt*. O dado é então armazenado na nuvem do Google, mais precisamente no segmento Google *Cloud Storage*.

As consultas a serem feitas são reescritas antes de serem enviadas para que possam fazer sentido a base de dados do *BigQuery*. Deste modo, as consultas feitas usando EBQ passam por uma espécie de interpretador, que adapta com criptografia os parâmetros a serem buscados de forma que, a base de dados do *BigQuery* não obtém conhecimento acerca dos resultados retornados. Semelhantemente, os resultados obtidos são recebidos cifrados são decodificados com a chave usada pelo EBQ armazenada no cliente e apresentados a aplicação, como ilustrado na Figura 1. Todo o acesso ao dado armazenado se dá de forma distribuídas pela integração com o *framework Hadoop* no Google *BigQuery* junto a plataforma de armazenamento.

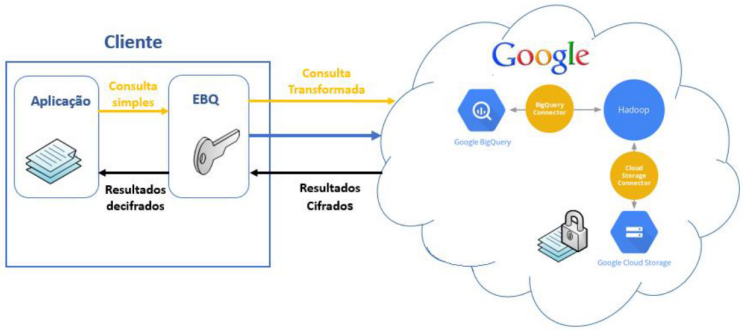


Figura 18 - Arquitetura EBQ.

Assim, a interface EBQ se mostra versátil, pois trata dois aspectos atuais: o processamento de grande volume de dados, auxiliado por um ambiente distribuído e a consulta a dados cifrados.

No que se refere aos requisitos apresentados – o uso junto ao EBQ - para a consulta de dados criptografados é necessário um conjunto de criptossistemas que trate esses dados adequadamente, sem a perda de confidencialidade. Dessa forma, os dados brutos são cifrados de maneiras diferentes antes de serem inseridos em um banco de dados. Dependendo do tipo de consulta a ser realizada num determinado campo, uma técnica distinta será usada. São elas descritas a seguir.

11.2.1 Algoritmo Probabilístico

A cifra é dita probabilística quando, para valores diferentes existirão cifras diferentes com grande probabilidade (Kaufman et al., 2002). Para uma construção eficiente é usada uma cifra de bloco, como AES no modo CBC juntamente com um vetor de inicialização (VI) gerado aleatoriamente. Provê máxima segurança devido a propriedade de indistinção sobre um Ataque de Texto em Claro Escolhido (IND-CPA), um adversário será incapaz de distinguir pares de textos cifrados baseados em mensagens cifradas por ele. Apesar de garantir confidencialidade e integridade dos dados em um banco de dados, não é possível realizar manipulações com os dados cifrados, exceto o comando *SELECT*.

11.2.2 Algoritmo Determinístico

O modelo de criptografia é dito determinístico quando é gerada a mesma cifra para a mesma mensagem em claro. O esquema é feito com a cifra de bloco AES no modo CBC com um VI de zeros. Assim, o esquema pseudônimo realiza uma permutação pseudo-aleatória (Goldreich, 2009) por meio do uso de cifras de bloco, o que representa uma diminuição no nível de segurança: adversários podem ter o conhecimento de quais valores cifrados correspondem ao mesmo valor em claro. A escolha desse esquema permite a realização de consultas de igualdade, isto é, pode realizar o comando *SELECT* com predicados de igualdade, junções de igualdade, *COUNT*. (Popa et al., 2011).

11.2.3 Busca por Palavras

Este esquema não é necessariamente um esquema de criptografia. É calculada a função *hash*, normalmente SHA-1, de todas as sequências possíveis de palavras. Em seguida, os *hashes* são mantidos em um campo e separados por espaços. Caso haja algum caractere especial inserido no texto, este deve ser informado no arquivo *schema*.

Representa o nível mais baixo de segurança implementado, uma vez que *hash* é uma função pública e facilmente reproduzida. Um ataque de dicionário conseguiria relacionar a mensagem em claro com o seu *hash* correspondente. Portanto, é utilizado na criptografia de campos que necessitem da funcionalidade de busca por palavras, entretanto não é necessário um nível de segurança alto. Podem ser usadas como atributo para cláusula *WHERE* com checagem de conteúdo (usando *CONTAINS*) com palavras-chaves inteiras, mas não aparecem em consultas *SELECT* simples.

11.2.4 Busca Probabilística por Palavras

A busca probabilística consiste em buscar uma palavra W e retornar todas as posições onde W aparece no texto em claro (Song et. al, 2000). Erros quanto ao aparecimento de posições erradas podem ocorrer, porém podem ser controlados ajustando o tamanho do texto cifrado gerado. Sendo □□={0,1}□□ o conjunto de textos cifrados, cada posição errada será retornada com probabilidade 12□□⁄(Song et. al, 2000).

Primeiramente, o texto é dividido em palavras-chave usando delimitadores padrão de uma palavra em português. As repetições são eliminadas, as posições aleatoriamente permutadas para garantir mais segurança e é feito o preenchimento (*padding*) para que cada palavra tenha o mesmo tamanho. A ideia principal do algoritmo proposto por Song et. al (2000) é cifrar um texto realizando a operação XOR bit a bit entre o texto em claro e uma sequência pseudo-aleatória de bits com uma estrutura específica. Essa configuração permite a busca de dados sem revelar nada sobre o conteúdo do texto em claro.

Este modelo criptográfico oferece um nível maior de segurança que o modelo de Busca por Palavra, uma vez que provê consultas isoladas para as buscas, ou seja, o servidor inseguro não pode aprender nada do texto em claro além do resultado da busca. Fornece uma busca controlada, para que o servidor não possa buscar uma palavra qualquer sem a autorização do usuário e, além disso, provê a consulta oculta de maneira que o usuário solicita uma busca de uma palavra secreta ao servidor sem revelar a palavra ao mesmo. Permite consultas *SELECT* para checagem de conteúdo, com o uso da cláusula *WHERE* e atributo *CONTAINS* ou *LIKE*.

11.2.5 Criptografia Homomórfica

Criptografia homomórfica é uma forma de criptografia que permite lidar com tipos específicos de cálculos em cifras e gerar um resultado também cifrado que, quando decifrado, corresponde ao resultado de operações realizadas. Ou seja, com as cifras C(x) e C(y), é possível obter C(x # y) sem decifrar x e y, para (Gentry, 2009).

Esse modelo criptográfico permite quatro operações matemáticas básicas, nas quais um dos termos é um número inserido pelo usuário, podem ser feitas com todos os campos numéricos, uma vez que, a resposta da consulta é decifrada primeiro e, em seguida, calculada a expressão antes de exibir os resultados.

O Quadro 4 apresenta um resumo destes modelos, o qual aponta que o criptossistema

que possui maior nível de segurança é o esquema Probabilístico, seguido da criptografia Homomórfica, Busca Probabilística por Palavras, Busca por Palavras e Determinístico.

Esquema	Construção	Funções	Sintaxe SQL
Probabilístico	AES em CBC	Dado estático	SELECT, UPDATE, DELETE
Homomórfico	Pailier (1999)	Adição	SUM, +
Busca por palavras	Hash SHA-1 das palavras	Busca de palavras	SELECT – WHERE, CONTAINS
Busca Prob. por palavras	Song *et. al* (2000)	Busca de palavra	SELECT – WHERE, CONTAINS, ILIKE
Determinístico	AES em CBC com VI nulo	Igualdade	=, !=, IN, COUNT

Quadro 3 - Resumo dos Esquemas Criptográficos Existentes no EBQ.

12 CONCLUSÃO

Por meio de um trabalho coordenado e interdependente entre as equipes da MJ/SEe da Universidade de Brasília, as atividades de elaboração deste RT foram planejadas, discutidas, executadas e documentadas.

Como modelo para avaliação e sugestões, este produto não representa um produto finalístico, sendo provável sua revisão após ciclos de análise e reavaliação e, principalmente, quando do atingimento de resultados parciais na realização das atividades envolvidas no subprojeto de Ecossistema do RIC e das avaliações obtidas por meio de Provas de Conceito.

Porém, o resultado de uma primeira avaliação deste documento pode permitir o direcionamento inicial de uma Prova de Conceito (PoC) mínima visando nortear a avaliação de quesitos específicos de infraestrutura, tal como os produtos e ferramentas que suportam os modelos e arquétipos sugeridos, questões de desempenho e curvas de ruptura na entrega de serviços, identificação de requisitos mínimos de *hardware* de infraestrutura de armazenamento, entre outros. É fator fundamental dessa Prova de Conceito também comprovar a eficácia e eficiência do modelo dada a capacidade de resposta de um número de acesso previsto para o registro da totalidade de cidadãos.

Neste sentido, este primeiro desenvolvimento deste Produto é importante pois, minimamente, elementos gerados podem dar suporte em todos os demais RTs do projeto.

As atividades envolvidas nesta etapa observaram formalmente a execução dos passos da metodologia elencada para gestão do projeto, PMI/PMBoK.

A equipe da UnB considera que teve acesso a todas as informações necessárias à boa condução dos trabalhos e que a disponibilização dessas informações pela equipe da MJ/SE/RIC, assim como as atividades conjuntas de análise e discussão, leva a etapa do projeto a bom termo.

13 REFERÊNCIAS

BHARGAV-SPANTZEL, A., CAMENISCH, J., GROSS, T., E SOMMER, D. User centricity: a taxonomy and open issues. Journal of Computer Security, 15(5), 493-527. 2007

DECANN, Brian; ROSS, Arun. De-duplication errors in a biometric system: An investigative study. In: Information Forensics and Security (WIFS), 2013 IEEE International Workshop on. IEEE, 2013. p. 43-48.ERL, T. Service-Oriented Architecture: Concepts, Technology, And Design. Editora Pearson Education, 2005. ISBN 813171490X, 9788131714904

JøSANG, A. e POPE, S. (2005). User centric identity management. In AusCERT Asia Pacific Information Technology Security Conference 2005.

JOSUTTIS, N. M. SOA na Prática. Tradução de Ivan Bosnic. 1. ed. Rio de Janeiro: Alta Books, 2008. 265 p. ISBN 978-85-7608-184-5.

LINTHICUM, D. Cloud Computing and SOA Convergence in Your Enterprise: A Step-by-Step Guide. Addison-Wesley, 2009. pg. 199

PROGRAMA RIC no RT de Soluções de Message Queue e Barramentos de Serviço Corporativos.

PROGRAMA RIC, RT Infraestrutura Tecnológica: Avaliação na Necessidade de Contratação de Múltiplos Parceiros.

TORRES, J.A.S., Diagnóstico Comparado do Governo Eletrônico Brasileiro - Uma Análise com Base na Estratégia de Gerenciamento de Identidades. Monografia – Especialização em Segurança da Informação e Comunicações. Instituto de Ciências Exatas - Departamento de Ciência da Computação – Universidade de Brasília, 2014.

Universidade de Brasília – UnB

Centro de Apoio ao Desenvolvimento Tecnológico – CDT

Laboratório de Tecnologias da Tomada de Decisão – LATITUDE

www.unb.br– www.cdt.unb.br – www.latitude.eng.br