

CONTRATAÇÃO DE SERVIÇOS DE DESENVOLVIMENTO DE *SOFTWARE* NA ADMINISTRAÇÃO PÚBLICA FEDERAL POR MEIO DE UMA MÉTRICA ORIENTADA AO COMPORTAMENTO DE SISTEMA (REGRAS DE NEGÓCIO)

Por Claudson dos Santos Melo

Servidor do Tribunal Superior do Trabalho, graduado em Análise de Sistemas de Informação pela União de Ensino Superior de Brasília, especialista em Engenharia de Software pela Universidade Católica de Brasília.

RESUMO

No presente trabalho propõe-se uma nova métrica de *software* de tamanho funcional derivada da NESMA (Advinda da Análise de Pontos de Função), porém sem algumas de suas deficiências conhecidas. A nova métrica, denominada Ponto de Especificação por Exemplo – PEEX possui maior correlação com o esforço despendido no desenvolvimento de software do que a métrica de pontos de função. A métrica PEEX está alinhada diretamente ao resultado da prestação do serviço pois mensura o esforço despendido dos desenvolvedores de *software* para construção dos objetivos de negócio dos usuários demandantes de solução de *software*, ademais, a métrica auxilia no aumento da qualidade do software entregue uma vez que toda a especificação do *software* construída no Tribunal Superior do Trabalho - TST é automatizada gerando benefícios de manutenibilidade e sustentação posterior após entrega do produto.

Palavras-chave: Engenharia de *Software*; Métrica de *Software*; Pontos de Função, NESMA, BDD, Especificação por Exemplo.

Introdução

O Decreto nº 2.271, de 7 de julho de 1997, sinaliza que a contratação de serviços de desenvolvimento e manutenção de sistemas deve ser realizada, preferencialmente, de forma terceirizada, com a contratação de empresas especializadas.

Com a ampla contratação de serviços de desenvolvimento de *software* pela Administração Pública Federal – APF e, após análise do TCU sobre diversos desses contratos, o Tribunal de Contas emitiu vários acórdãos relativos às contratações de serviços de desenvolvimento de *software* e sobre o uso de métricas nesse tipo de instrumento, segue alguns exemplos: acórdãos nº 2.094/2004, nº 2.138/2005, nº 2.658/2007, nº 1.782/2007, nº 1.910/2007, nº 2.172/2008, nº 2.471/2008, nº 1.125/2009, nº 2.348/2009, nº 1.274/2010 e Nota Técnica 6/2010-Sefti/TCU, nº 1.647/2010; nº 2.314/2013 e Nota Técnica 505/2015 (TC 002.116/2015-4).

Abaixo segue trecho do relatório do TCU sobre seis auditorias de gestão de contratos de TI na APF pelo TCU¹:

Tecnologia da Informação (TI) tem-se tornado, cada vez mais, elemento indispensável ao alcance dos objetivos organizacionais. No âmbito público federal, a legislação prioriza a execução indireta de tarefas executivas por meio da contratação de empresas para a prestação de serviços, de modo que ganham especial relevância as contratações de serviços de TI efetuadas pelos órgãos e entidades federais.

(...)

Considerando que o atual modelo de contratação de soluções de TI é calcado no planejamento da contratação, na divisão do objeto, no pagamento por resultados, na avaliação da qualidade, no controle efetivo da execução contratual e na existência de recursos humanos capacitados, observa-se, ante as fiscalizações realizadas por esta Corte no ano de 2008, evolução no panorama das contratações de serviços de TI adotado pelos entes públicos, com maior

¹ Avaliação da gestão de contratos de TI em seis auditorias, e acórdão 916/2015-TCU-Plenário de 22 de abril de 2015. <https://portal.tcu.gov.br/biblioteca-digital/avaliacao-da-gestao-de-contratos-de-ti-em-seis-auditorias.htm>, acessado em 24/8/2017

disseminação do planejamento da contratação, avanços no pagamento por resultados e na avaliação da qualidade dos serviços.

Considerando as diretrizes do atual modelo de contratação de soluções de TI, a presente avaliação focou a mensuração e o pagamento dos serviços por resultados, a avaliação sistemática de qualidade dos serviços prestados e o efetivo controle da execução contratual, além de alguns pontos a respeito do planejamento da contratação. (grifo meu)

Quanto à mensuração dos resultados, o TCU determinou a restrição do uso da métrica homem-hora, conforme Súmula nº 269 do Tribunal de Contas da União (TCU): a remuneração dos serviços deve ser vinculada a entrega de resultados e restringem a utilização da métrica homem-hora;

A Instrução Normativa nº. 4 MP/SLTI, 12 de novembro de 2010 (versão inicial) e sua edição de mesmo número IN 4 MP/SLTI Nº 4, 11 de setembro de 2014 também vetou o uso da métrica homem-hora conforme inciso VIII, Art. 7º:

Art. 7º. É vetado:

(...)

VIII - adotar a métrica homem-hora ou equivalente para aferição de esforço, salvo mediante justificativa e sempre vinculada à entrega de produtos de acordo com prazos e qualidade previamente definidos;

O Tribunal de Contas recomendou aos Órgãos ditos como reguladores das diversas esferas Federais que elaborassem normativo sobre as fases da contratação, a saber, planejamento, acompanhamento e controle. Resulta-se dessa recomendação a Instrução Normativa - SLTI/MP 4/2010 (principalmente o art. 11, inciso II) e a Resolução CNJ 182/2013.

Assim, para viabilizar contratação de serviço de desenvolvimento de *software* é imprescindível que, na fase de planejamento, seja definida de forma clara e objetiva a métrica a ser utilizada para mensuração do serviço prestado e consequente pagamento baseado na métrica.

O presente artigo apresentará métrica criada no âmbito do TST que está alinhada ao resultado da prestação do serviço, aos objetivos de negócio dos usuários demandantes de solução de *software* e que visa auxiliar no aumento da qualidade do *software* entregue.

Uso de métrica para aferição de resultado para serviços de desenvolvimento de *software*

O pagamento dos serviços prestados deve ter correlação com o resultado obtido e, para aferição do resultado é necessário se ter uma métrica associada. Para medição de *software* tem-se no mercado atual algumas métricas que tratam de diferentes abordagens, como: tamanho funcional de *software*; tamanho não-funcional do *software*; artefatos gerados nas diversas fases de construção de *software*; unidade de esforço despendido para construção de artefatos de *software*; dentre outras. Algumas dessas métricas podem ser complementares umas às outras em seu uso.

São exemplos de métricas relacionadas ao tamanho funcional do *software* e normatizadas pela ISO 14143: FISMA (IEC, 2010), COSMIC-FFP (IEC, 2011b), MKII FPA (IEC, 2002b), NESMA (IEC, 2005) e Análise de Pontos de Função.

Por outro lado, há também métricas definidas pelos Contratantes que estão mais relacionadas ao esforço, como exemplo, Horas de Serviço Técnico, Unidade de Serviço Técnico, Profissional Ideal, dentre outras. Essas métricas utilizam catálogos ou repertório de serviços onde há lista de artefatos de *software* e associados aos profissionais de TI responsáveis por confeccionar tais artefatos. A quantidade de unidade de medida atribuída refere-se ao esforço despendido para produção de cada artefato.

Há também métricas em desuso e que não são adequadas para contratação de serviço de desenvolvimento de *software*, e.g. LOC (linhas de código) e, conforme já dito, homem-hora.

Assim, o Decreto n.º 2.271/97, o qual dispõe, no § 1º do seu art. 3º, que sempre que a prestação de serviço, objeto da contratação, puder ser avaliada por determinada unidade quantitativa de serviço prestado, esta será utilizada como um dos parâmetros de aferição do resultado.

A métrica de Pontos por Função

No mercado brasileiro de construção de *software*, a métrica de Ponto de Função é amplamente utilizada nas contratações de serviço que envolvem desenvolvimento de novos sistemas, manutenção evolutiva e, em menor escala, migração de sistemas.

A métrica Pontos de Função é uma medida de tamanho funcional de projetos de *software*, considerando as funcionalidades implementadas, sob o ponto de vista do usuário. Tamanho funcional é definido como “tamanho do *software* derivado pela quantificação dos requisitos funcionais do usuário” (Dekkers, 2003).

A métrica foi definida por Allan Albrecht em 1979, tem sido utilizada de forma crescente pela indústria de *software*. O IFPUG (*International Function Point Users Group*), criado em 1986, é responsável pela atualização das regras de Contagem de Pontos de Função, descritas no CPM (*Counting Practices Manual*), que se encontra na versão 4.2.1, publicada em 2005 no IFPUG. O IFPUG também é responsável pelo exame de certificação de especialistas em contagem de Pontos de Função, denominada CFPS (*Certified Function Point Specialist*).

A métrica de Ponto de Função está compreendida basicamente em dois tipos de função e suas respectivas complexidades.

Funções de dados

Funções de transação

Cláudia Hazan² recomenda o uso da métrica Pontos de Função Não Ajustados como unidade de medida nos contratos de fábrica de *software*.

Contudo, é importante ressaltar que o detalhamento abordado na métrica de ponto de função não leva em consideração os detalhes do comportamento com suas respectivas lógicas de negócio.

A métrica de Ponto de Função no TST

Desde 2009, o TST utiliza a métrica de pontos de função para contratação de Fábrica de *software*. Os preços das contratações/atas foram de R\$ 269,89 (PE 124/2009), R\$ 445,00 (PE 146/2012) e R\$ 449,00 (PE 009/2016).

Para apoiar na realização da mensuração dos produtos entregues pelas Fábricas de *Software* o TST fez uso de contratações de empresas especializadas para realização de contagem em pontos de função.

Dificuldades no uso da métrica de Ponto de Função

Com as diversas contagens realizadas no TST utilizando-se da métrica de Ponto de Função foram vivenciados os seguintes problemas:

- Divergências recorrentes entre contratante e contratada na identificação das funções de dados por ser a métrica subjetiva e inadequada para a produção de *software* da atualidade e das necessidades dos usuários que são diferentes das necessidades da década de 80;
- Não consideração do esforço empreendido para se construir sistemas complexos, em contrapartida, a métrica contabiliza muitos pontos de função em atividades simples como elaboração de entidades caracterizadas por feituas de *CRUD* (*create, retrieve, update e delete*);

² Hazan, Cláudia. Como evitar armadilhas nos contratos de Fábrica de Software. <<http://revista.tcu.gov.br/ojs/index.php/RTCU/article/download/279/292>>. Acessado em 24/08/17.

- Inviabilidade econômica para sustentação de *software*, pois na mensuração por pontos de função há dificuldade em remunerar evoluções em sistemas já existentes, pois quando as funções de dados e as entidades de dados já estão criadas, a contratada somente receberá 50% das funções de transação, ainda que, a alteração resulte em inserção de grande quantidade de regras de negócios que redundam em alta complexidade e/ou que demande grande esforço de codificação de testes automatizados, a exemplo do sistema TST-Saúde. Isso pode gerar mensuração em PF consideravelmente inferior quando comparado ao desenvolvimento de novos sistemas.
- Gestores – geralmente, fiscais demandantes –, têm dificuldade em administrar saldo de PF diante das necessidades de suas respectivas áreas devido ao desconhecimento da métrica;
- A contagem de pontos de função por ser complexa, gera muito desperdício de tempo dos envolvidos e, ainda mais, quando há necessidade de arbitragem de conflitos;
- Necessidade de capacitação e desejável especialização na métrica de Pontos de Função, além de prática constante em contagens;
- Custo com contratação de empresa especializada para aferir *software* e esforço posterior para gerir os contratos firmados; e
- Necessidade de utilizar outra métrica, além do Ponto de Função para contemplar a sustentação de sistemas.

Outras dificuldades encontradas no uso da métrica Ponto de Função

Vários autores têm feito críticas à validade e à aplicabilidade da Análise de Pontos de Função. Isso põe em xeque a correção do seu uso em contratações e fragilizam a confiabilidade de sua aplicação como ferramenta para a gestão e para a governança da TI (ABRAN; ROBILLARD, 1994; FENTON; PFLEEGER, 1998; KITCHENHAM, 1997; KITCHENHAM; KÄNSÄLÄ, 1993; KITCHENHAM et al., 1995; KRALJ et al., 2005; PFLEEGER et al., 1997; TURETKEN et al., 2008; XIA et al., 2009), conforme citam CASTRO e HERNANDES³.

Relatório operacional do TCU – uso de métricas

O Relatório da Auditoria Operacional realizada mediante autorização do Tribunal exarada no Acórdão nº 505/2015 – TCU–Plenário, prolatado no bojo do processo TC-002.402/2015-7, sessão de 11/3/2015 (peça 1), apontou a insatisfação do Banco do Brasil no uso da Análise de Pontos de Função e, como solução, o desenvolvimento de métrica própria, como se segue:

76. ... a organização optou por desenvolver sua própria métrica devido ao fato de não estar completamente satisfeita com a Análise de Pontos de Função em contratos anteriores, já que, na visão daquela organização: (i) a correlação entre o valor expresso em pontos de função (PF) e o custo da prestação do serviço nem sempre se mostrava adequada em situações de software de elevada complexidade; (ii) havia dificuldade de uso de PF para estimativa e planejamento do serviço a ser executado, uma vez que o demandante não conhecia a técnica; e (iii) esses fatores resultavam em maior ônus para medição do volume de serviços prestados.

Semelhantemente ao Banco do Brasil, os técnicos do TST que mensuram as entregas utilizando-se da métrica Ponto de Função também encontraram dificuldades com a métrica de PF devido aos diversos problemas citados na seção anterior.

Nesse mesmo relatório, o TCU ressaltou a importância sobre o grau de maturidade do desenvolvimento de *software* daquele órgão que favoreceram a adoção de métrica própria:

77. Além da insatisfação com a métrica, gestores dessa organização afirmaram desconhecer métrica de mercado que se mostre mais adequada que a Análise de Pontos de Função. Por fim, o Banco possuía uma série de condições para especificar sua métrica própria, como, por exemplo: pessoal qualificado disponível, histórico de casos anteriores e processo de

³ CASTRO, Marcus Vinícius Borela de; HERNANDES, Carlos Alberto Mamede. Uma métrica de tamanho de software como ferramenta para a governança de TI. Revista do TCU. Jan/abril 2016. Pág. 61.

desenvolvimento de software estabelecido, o Processo de Desenvolvimento de Aplicativos do Banco do Brasil (PDABB).

78. A USTIBB é composta por um catálogo de serviços que podem ser demandados pelas áreas de negócio. Esses serviços, que estão divididos por tipo de plataforma computacional, complexidade, entre outras, são frequentemente atualizados pela entidade, a fim de atender a novas tecnologias ou melhorias do processo de desenvolvimento de software. Além disso, as demandas são controladas por meio de sistema informatizado, o que permite acompanhar, em tempo real, o processo de desenvolvimento e os custos a ele associados. Não se pode deixar de mencionar que parte da viabilidade de tal modelo está diretamente relacionada à organização que o implementou (i) possuir histórico de projetos, (ii) reter conhecimento adquirido com sucessos e fracassos anteriores, (iii) conhecer profundamente a disciplina de desenvolvimento de software, (iv) buscar experiências de outras instituições e (v) ser capaz de prevenir e mitigar riscos que outras instituições menos estruturadas não conseguem.

79. Verifica-se, com base nas entrevistas com os gestores de TI da entidade, que a USTIBB, em si, não está em desacordo com a legislação ou a jurisprudência do TCU, visto que a métrica visa à garantia de pagamento vinculado a resultados, para todo o ciclo de desenvolvimento de software. A dificuldade de expansão para outras organizações reside no fato de nem todas possuírem as condições que esta entidade possui, exemplificados no parágrafo 77. Isto porque a USTIBB é uma métrica aderente à realidade do Banco, o que significa que eventual utilização em outra organização somente deveria ocorrer com prévia adaptação, a qual não prescinde das condições previamente lá existentes. (grifos nossos)

Para viabilizar o uso de nova métrica em contratações de fábrica de software, o TCU apontou outros aspectos que favorecem o uso de tal inovação, tais como: (i) pessoal qualificado disponível; ii) histórico de casos anteriores; (iii) processo de desenvolvimento de software estabelecido; (iv) possuir histórico de projetos; (v) reter conhecimento adquirido com sucessos e fracassos anteriores; (vi) conhecer profundamente a disciplina de desenvolvimento de software; (vi) buscar experiências de outras instituições; e (viii) ser capaz de prevenir e mitigar riscos que outras instituições menos estruturadas não conseguem.

Situação semelhante está sendo vivenciada pelo TST, ademais, pode-se ainda associar alguns outros aspectos relevantes para a criação da métrica PEEX:

1. Requisito de software sendo dada a devida importância (com inserção de testes automatizados nos cenários dos requisitos apontados) com o objetivo de mitigar risco de fracasso nos projetos por falha nessa fase. Como bem observa o PMI: *A gestão de requisitos é uma competência fundamental na gestão de projetos, mas também destaca que é a causa de 47% dos projetos que fracassam.*
2. Empresa com acesso ao ambiente do TST (integração contínua, banco de homologação, dentre outros);
3. Histórico da métrica computada nos últimos projetos com comparações realizadas para mitigar risco de licitação ser mais onerosa para o erário;

Quanto a não obrigatoriedade do uso da Análise de pontos de função, o relatório aponta:

4.1.3 Uso de Análise de Pontos de Função não é obrigatório

80. A jurisprudência do TCU é de que os pagamentos por serviços de TI devem ser efetuados por resultados, nos termos da Súmula-TCU 269, não havendo obrigatoriedade de métrica específica que deva ser utilizada. Ou seja, a escolha da métrica fica a cargo dos gestores, devendo ela importar pagamentos por resultados. Ainda assim, durante a auditoria foi feita uma pesquisa ampla na jurisprudência, objetivando identificar situações em que poderia estar havendo uma determinação direta para que os gestores adotassem a Análise de Pontos de Função como métrica para remuneração de desenvolvimento de software, cujo resultado resume-se nos próximos parágrafos.

81. Um julgado do Tribunal que menciona especificamente o uso da Análise de Pontos de Função é, por exemplo, o Acórdão 2.024/2007-TCU-Plenário. Note-se que o uso da referida métrica é apresentado de forma exemplificativa, sendo impositiva apenas a mensuração por resultados:

9.2. fixar, com fulcro no inc. IX do art. 71 da Constituição Federal c/c o art. 45 da Lei 8.443/1992 e com o art. 251 do Regimento Interno desta Corte, o prazo de sessenta dias para que o Departamento de Logística do Exército Brasileiro/Ministério da Defesa:

9.2.2. atente, caso ainda haja interesse na contratação referida, para que, na elaboração da peça editalícia correspondente, tanto na hipótese de instauração de novo procedimento licitatório quanto na de retomada do certame já iniciado, os seguintes aspectos sejam observados, de modo a evitar-se irregularidades identificadas no Edital de Tomada de Preços 001/2007-D LOG:

9.2.2.2. prever metodologias de mensuração de serviços prestados que privilegiem a remuneração da contratada mediante a mensuração de resultados, a exemplo da análise por Pontos de Função (método padronizado largamente utilizado no mercado nos dias de hoje para a mensuração de serviços de desenvolvimento e manutenção de sistemas, considerando as funcionalidades implementadas, sob o ponto de vista do usuário), buscando eliminar a possibilidade de remunerar a contratada com base na quantidade de horas trabalhadas ou nos postos de trabalho disponibilizados ou, caso tal caminho não se mostre comprovadamente viável, restando como única opção a remuneração de serviços por horas trabalhadas, cuidar para que sejam previamente definidos e especificados os serviços a serem executados e estabelecidos, também de antemão, os valores máximos de horas aceitáveis para cada um desses serviços, assim como explicitada a metodologia a ser utilizada para a identificação desse quantitativo de horas; (grifou-se)

82. Já o Acórdão 2.836/2008-TCU-Plenário determinou expressamente que as contratações da Fundação Nacional de Saúde fossem feitas com uso de pontos de função:

9.6. determinar à Fundação Nacional de Saúde - FUNASA que:

9.6.4. observe a determinação exarada no item 9.3.3 do Acórdão 667/2005 – Plenário e, partir de 2/1/2009, as orientações expedidas pelo Ministério do Planejamento na IN-SLTI/MP 4/2008, promovendo o treinamento do seu quadro técnico para que, nas futuras licitações na área tecnologia da informação, as contratações sejam realizadas por serviços e por pontos de função; (grifou-se)

83. Referido acórdão menciona o item 9.3.3 do Acórdão 667/2005-TCU-Plenário (transcrito a seguir) como fundamentação do comando. Entretanto, este não menciona a métrica, conforme será mostrado a seguir. Ocorre que na época em que esses acórdãos foram prolatados, a jurisprudência a respeito de métricas para contratação de desenvolvimento de software estava em fase de construção, o que explica, naquele momento, entendimentos aparentemente divergentes.

9.3. determinar à SPOA/MDIC que, quando da abertura dos novos procedimentos licitatórios em substituição à Concorrência 01/2005, observe as determinações expedidas no item 9.3 do Acórdão 1.094/2004-Plenário, bem como os seguintes preceitos na elaboração dos editais:

9.3.3. adote metodologias de mensuração de serviços prestados que privilegiem a remuneração das contratadas mediante a mensuração de resultados e que eliminem a possibilidade de remunerar as empresas com base na quantidade de horas trabalhadas ou nos postos de trabalho;

84. Já o Acórdão 1.153/2013-TCU-2ª Câmara foi prolatado com o seguinte teor a respeito do uso de pontos de função:

1.8. Recomendar ao Centro Tecnológico de Informática do Ministério da Saúde que:

1.8.4. utilize a métrica pontos de função para a remuneração de serviços de desenvolvimento de sistemas, em substituição à métrica de homens-hora, consoante especificações da Nota Técnica 6/2010 - Sefti/TCU.

85. Por ter sido citada no referido acórdão, cabe transcrever a parte da Nota Técnica 6/2010-Sefti/TCU (cuja publicação foi autorizada pelo Acórdão 1.233/2013-TCU-Plenário) a respeito do uso de pontos de função. Percebe-se que o comando expedido pela unidade especializada menciona, de forma exemplificativa, a Análise de Pontos de Função.

12. Nessa forma de contratação, outro aspecto que deve ser destacado é a mensurabilidade dos resultados. Os serviços de TI comumente contratados pela APF podem ser adequadamente mensurados, seja por unidades de medida de tamanho (e.g. pontos de função em desenvolvimento de software) ou por indicadores de nível de serviço (mais utilizados para serviços de suporte a banco de dados e rede de computadores, entre outros). (grifou-se)

86. Como último exemplo de comando para uso de pontos de função, foi identificado o Acórdão 2.393/2013-TCU-Plenário. Neste há determinação de que se use pontos de função, mas a determinação decorre do fato de o órgão possuir norma interna por meio da qual se obriga ao uso de tal métrica, no caso a Portaria-MF 47/2011.

9.1. determinar à Secretaria da Receita Federal do Brasil que:

9.1.2. preveja que o método para mensuração e pagamento dos serviços de Desenvolvimento e Manutenção de Sistemas seja o de análise de ponto de função (APF) não ajustado, em atendimento ao disposto na Portaria – MF 47/2011, art. 2º, § 5º, c/c o art. 54, § 1º da Lei 8.666/1993;

87. Concluindo, neste achado foram apresentados os acórdãos que poderiam dar margem a interpretação no sentido da obrigatoriedade da Análise de Pontos de Função. Também foi demonstrado que esse não é o entendimento que coaduna com a jurisprudência majoritária do TCU, consolidada na Súmula-TCU 269. Também foram citados casos em que organizações estão fazendo uso de outras formas de medição como UST e USTIBB, sem que esse fato seja contrário à jurisprudência do TCU. Tais fatos permitem concluir que a obrigação é de que sejam usados critérios objetivos e baseados em resultados, não exclusivamente a Análise de Pontos de Função.

88. O entendimento do TCU, apresentado de maneira clara como a que aqui se propõe, serve como importante balizador para os gestores públicos que dependem de conhecimento adequado da jurisprudência para as constantes tomadas de decisão sobre contratação de serviços de desenvolvimento de software.

89. As notas técnicas são instrumento utilizado para orientar os auditores do TCU e os gestores públicos acerca de entendimentos do TCU, tendo a sua jurisprudência como principal fonte. A Nota Técnica 6/2010-Sefti/TCU foi elaborada com objetivo de avaliar a aderência de determinada forma de contratação ao conceito de pagamento por resultados.

90. Em vista disso, propor-se-á ao Tribunal que autorize a Sefti a proceder aos devidos ajustes na Nota Técnica 6/2010-Sefti/TCU, de forma a considerar as conclusões do item 4.1 deste relatório.

A inviabilidade do uso da métrica de pontos de função para remuneração de sistemas complexos é fato reconhecido pelo TCU, conforme trecho abaixo extraído do acórdão de seu Plenário:

46. Diante deste cenário, algumas organizações têm buscado a contratação de serviço de desenvolvimento de sistemas com escopo mais amplo, fazendo uso de empresas de desenvolvimento de software generalistas, o que traz como benefício a flexibilidade de atendimento de várias demandas no âmbito de um mesmo contrato, mas que também traz riscos, como os de menor especialização da empresa em determinados tipos de projeto, área de negócio ou plataforma tecnológica e maior imprecisão na formação de preços.

47. *Os riscos tendem a aumentar porque a formação de preços, no momento da licitação torna-se imprecisa: os preços devem variar segundo a tecnologia utilizada e também de acordo com a área de conhecimento dos sistemas a desenvolver. A imprecisão de preços foi enfrentada pelo TCU por meio do Acórdão 161/2012-TCU-Plenário, nos seguintes termos:*

9.2. determinar ao Conselho Nacional de Justiça, com base no art. 251, caput, do Regimento Interno do Tribunal que:

9.2.2. nas próximas licitações para contratação de empresa para prestação de serviços técnicos de fábrica de software:

9.2.2.1. aponte a proporção de cada linguagem operacional/plataforma tecnológica a ser utilizada no total da quantidade de pontos de função necessários, assim como a linguagem que será utilizada para desenvolver cada sistema, quando for o caso, demonstrando analiticamente a metodologia de cálculo usada para chegar ao quantitativo de ponto de função estabelecido para cada sistema;

48. *Hipoteticamente falando, para melhor explicar o risco envolvido, o preço a ser exigido por um fornecedor para um sistema focado em cálculos atuariais, por exemplo, deveria ser maior que aquele exigido para um sistema de controle de protocolo de documentos. Isso tendo em vista a especialização das equipes.*

49. *Logo, se o fornecedor não tiver conhecimento, de antemão, de que tipos de sistemas poderão ser demandados, a tendência é que ele seja conservador e majore os preços, a fim de evitar prejuízos. Neste caso, se forem demandados apenas sistemas de baixa complexidade, a APF estaria pagando valores acima do daqueles realmente praticados no mercado para o que fora demandado.*

50. *Entretanto, com a dinâmica de preços mais comum em certames licitatórios, poderá ocorrer situação oposta, na qual fornecedores mais arrojados, contando com a possibilidade de que a maioria dos sistemas demandados sejam de baixa complexidade, reduzam o preço a ponto de a execução contratual se tornar inexecutável (ver item 4.2.1) para sistemas complexos. Se estes últimos vierem a ser demandados pela APF, pode correr a materialização do risco de não entrega ou entrega com qualidade inadequada do objeto. Em que pese haver possibilidade de glosas e sanções, o interesse público, de ter os sistemas implementados e em funcionamento, não terá sido atendido.*

Em suma:

- A métrica de Ponto de Função, além de não ser obrigatória, há cenários em que ela não é aplicável;
- Há iniciativas na Administração Pública Federal, desde 2010, para confecção de novas métricas que melhor atendam as expectativas dos Órgãos;
- A métrica de Ponto de Função não considera de forma adequada a complexidades dos *softwares*; e, por fim,
- As condições de vivência do TST em contratações e pelas condições de evolução tecnológicas (com muita automação) colaboram para o desenvolvimento de métrica associada ao produto e ao processo de desenvolvimento de *software*, contribuindo para a eliminação de desperdício.

Crítica à Análise de Pontos de Função que motivaram a criação da nova métrica

Inviabilidade de execução de serviço de manutenção evolutiva em projetos de média/alta complexidade com a métrica atualmente vigente (Análise de Pontos de Função).

O serviço de manutenção evolutiva enseja em que ele já tenha sido parcialmente construído e esteja em funcionamento em ambiente de produção e carece de evolução de requisitos de negócios.

Tal situação foi encontrada no sistema de Autogestão de Plano de Saúde do TST. Esse sistema é um dos mais complexos do TST devido ao seu grande volume de regras de negócios e carecendo de respectivas validações dessas regras para um perfeito funcionamento.

Em avaliação prévia desse Sistema e contrato firmado (PE 009/2016-D) para sua evolução, a empresa contratada apresentou carta informando inviabilidade econômica de execução do contrato devido a dois fatores: a) projeto de alta complexidade e b) baixa contagem de Pontos de função.

Tal fato foi o principal motivador para elaboração de uma métrica que pudesse abranger aspectos do comportamento do *software* com suas respectivas regras de negócio.

Metodologias Ágeis de Desenvolvimento de Sistemas e seu uso no TST

Por muitos anos a Administração Pública Federal teve insucessos em suas contratações. Uma das causas apontadas por diversos Órgãos foi o uso da metodologia RUP (*Rational Unified Process*). Com o surgimento de metodologias ágeis, alguns Órgãos passaram a usar essa metodologia para contratações de desenvolvimento de software e os resultados apresentados, principalmente, em eventos sobre agilidade no Governo, têm se mostrado positivos.

Esse novo cenário, ou seja, de uso de metodologias ágeis, está descrito no Ac. 2314/2013- TCU- Plenário em auditoria de visita realizada pelo TCU em alguns Órgãos da Administração Pública Federal.

A métrica é fruto do uso de prática ágil de desenvolvimento de sistemas, a saber, especificação por exemplo, advinda do *Behaviour Driven Development* – BDD.

Algumas das práticas ágeis de desenvolvimento de *software*, preconizam que os testes sejam pensados desde o início do processo de desenvolvimento, ou seja, desde o detalhamento dos requisitos. Assim, há várias práticas ágeis que contribuem para isso, alguns exemplos são: *Behaviour Driven Development* – BDD, *Acceptance Test Driven Development* – ATDD, *Design Driven Development* – DDD, dentre outros que podem ser associados, como por exemplo: *Test Driven Development* – TDD.

O BDD (*Behaviour Driven Development*), ou Desenvolvimento orientado por comportamento, visa auxiliar a atuação conjunta entre o time de desenvolvimento e o usuário demandante para a construção de *software* de qualidade por meio de especificação do *software* por exemplo (Especificação por Exemplo), como detalhado a seguir.

Exemplo: Sacar dinheiro de Caixa eletrônico

História: Sacar dinheiro de um caixa eletrônico

Para movimentar minha conta bancária (funcionalidade)

Como um cliente de conta bancária (papel)

Desejo sacar dinheiro por meio de um caixa eletrônico (objetivo)

Possíveis, porém não exaustivos, cenários para essa história:

Cenário: Sacar dinheiro com saldo suficiente

Cenário: Sacar dinheiro com saldo insuficiente

Cenário: Tentativa de sacar dinheiro usando um cartão inválido para a conta.

Exemplo do cenário detalhado com seu comportamento (Especificação por exemplo)

Cenário: Sacar dinheiro com saldo suficiente

Dado que eu tenha R\$ 110 de saldo na minha conta

Quando eu solicitar R\$ 50,00

Então o sistema atualiza o saldo da conta para R\$ 60

A descrição do comportamento do Sistema visa também a observância de um princípio ágil que é entregar certo da primeira vez.

No TST, principalmente para contratações de fábrica de *software*, utiliza-se a abordagem de especificação por exemplo desde o ano de 2014.

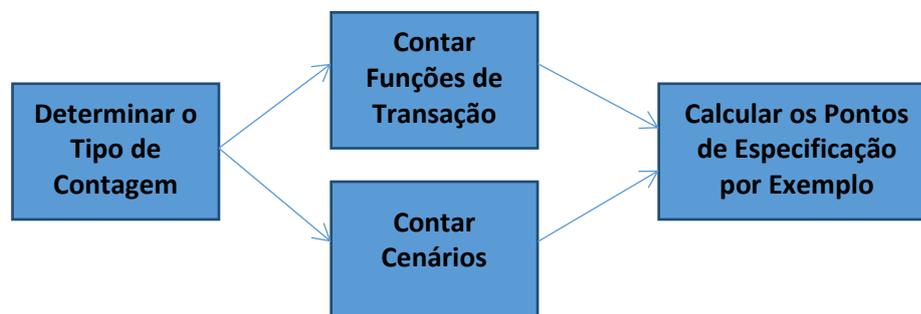
A métrica Ponto de Especificação por Exemplo

A métrica proposta pelo TST, chama-se Ponto de Especificação por Exemplo – PEEEX.

Ela é uma métrica não direta, pois depende do conjunto: tamanho funcional e dos cenários de especificação por exemplo.

Procedimento de Contagem de PEEEX

Esta Seção apresenta o procedimento para contagem de pontos de especificação por exemplo.



A Fórmula PEEEX

A mensuração do Ponto de Especificação por Exemplo – PEEEX é dada pela soma das funções de transações (TR - aferidas pela contagem estimada – NESMA) mais o quantitativo de cenários de testes contidos nas histórias de usuários (CT).

Assim, na contagem estimada NESMA, Entrada Externa – EE e Consulta Externa – CE valem 4 pontos e Saída Externa – SE vale 5 pontos.

Cada cenário de história equivale a 1 (um) ponto na métrica PEEEX.

$$PEEX = \square TR + \square CT$$

Onde:

TR = pontuação das transações NESMA (EE, CE e SE)

CT = cenários de histórias

Regras de Ponto de Especificação por Exemplo

A seguir são descritas as regras para consideração de cenários de testes.

Os cenários de teste para serem considerados na contagem devem:

- Ser cenários sob a ótica do usuário;
- Registrar as lógicas de negócio do sistema;
- Ser escrito utilizando especificação por exemplo; e,

Para quando exigido:

- Possuir teste automatizado codificado associado;

Pontos que devem ser observados

- O conceito de fronteira da aplicação será sempre a do sistema em desenvolvimento/manutenção evolutiva.

- Os cenários das histórias devem conter exemplos contemplando um ou mais caminhos positivos e um ou mais caminhos negativos.
- A especificação por exemplo pressupõe escrita de cenários que sejam compreensíveis por todos os envolvidos, principalmente pelos demandantes da solução, e deve conter:
 - Pré-condições;
 - Entradas; e;
 - Saídas associadas.
- Validações de formatos de arquivos XML com suas respectivas regras estipuladas em arquivos (XSD, DTD...) não são considerados como cenários para pontuação de cenários de teste no PEEEX.
- Deve haver apenas um cenário para verificação de campos obrigatórios. Esse deve ser considerado no mesmo cenário que realiza o cadastro de uma entidade.

Por exemplo:

Cenário: Cadastro de Estagiário

!-- O cpf tem de ser válido

!-- A data tem de ser no futuro

Dado que o nome do estagiário é <nome>

Dado que a instituição de ensino é <instituição>

Dado que o cpf do estagiário é <cpf>

Dado que a data de encerramento do estágio é <data>

Quando for solicitado o cadastro

Então é informado que a situação do cadastro é <situação>

Exemplos:

nome	instituição	cpf	data	situação
James T. Kirk	Academia da Federação	12345678901	20/02/2200	SUCESSO
James T. Kirk	Academia da Federação	12345678901	----	SUCESSO
James T. Kirk	-----	12345678901	20/02/2200	FALHA
James T. Kirk	Academia da Federação	-----	20/02/2200	FALHA
----	Academia da Federação	12345678901	20/02/2200	FALHA

- Os cenários têm por objetivo a especificação das lógicas de negócio, escritas sob a ótica do usuário, como requisitos a serem desenvolvidos.

Por este motivo, validações naturais como validação de CPF, CNPJ, datas, número de telefone, CEP, formatos de arquivos, etc., além de campos que tenham validação firmada por restrição de banco de dados, como tipos de sessão (ordinária/extraordinária), unidade administrativas ativas, códigos de servidores, não devem ser considerados como cenários para a contagem de PEEEX.

Por exemplo, na história a seguir:

Para faturar um lote de guias de consultas médicas

Como um prestador

Desejo uma funcionalidade de envio de arquivo XML padrão TISS 3

São exemplos de cenários válidos para computo de PEEEX para histórias semelhantes a essa:

1. Validação se o beneficiário estava ativo no momento da consulta médica;
2. Validação se o prestador estava ativo no momento da consulta médica; e
3. Validação se o valor cobrado é igual ou inferior ao valor de tabela.

Não são exemplos de cenários válidos para computo de PEEEX para histórias semelhantes a essa:

1. Validação CNPJ/CPF do prestador;
 2. Validação se matrícula do beneficiário é válida;
 3. Validação se tipo de consulta é válida;
 4. Validação se formato do XML é válido;
 5. Validação se código da consulta é válido;
 6. Validação se valor total cobrado corresponde à soma das guias (caso não exista regra especial de arredondamento).
- Questões de implementação de tela devem estar especificadas, quando necessárias, em protótipo anexado à história de usuário.
 - Serão considerados com um único cenário:
 - Múltiplas entradas de dados (múltiplos exemplos) com suas respectivas saídas; e;
 - Cenários acessados por diferentes papéis.

Por exemplo, os cenários 1 e 2 descritos abaixo, por serem de múltiplas entradas de dados são considerados como um único cenário para computo de PEEEX, assim, devem ser escritos em um único cenário com múltiplos exemplos, conforme descrito no “Cenário Unificado”:

Cenário 1: operador da bolsa não deve ser alertado caso a valor esteja abaixo do limite.

Dado que a ação é STK1 e o limite é 10,00

Quando a ação atingir o valor de 5,00

Então o alerta deve ter situação de INATIVO

Cenário 2: operador da bolsa deve ser alertado caso a valor esteja acima do limite.

Dados que a ação é STK1 e o limite é 10,00

Quando a ação atingir 11,00

Então o alerta deve ter situação de ATIVO

Os dois cenários acima refletem a mesma regra de negócio e poderiam ser escritos em apenas um cenário contendo múltiplos exemplos:

Cenário Unificado: operador da bolsa deve ser alertador caso a valor esteja acima do limite.

Dado que a ação é <símbolo> e o limite é <limite>

Quando a ação atingir o valor de <valor>

Então o alerta deve ter situação de <situação>

Exemplos:

símbolo	limite	valor	situação
STK1	10.0	5.0	INATIVO
STK1	10.0	11.0	ATIVO
STK1	10.0	10.0	INATIVO
STK1	10.0	10.01	ATIVO

Benefícios do uso da métrica PEEX

- Esforço mensurado refletindo em pagamento por resultado que agrega valor à qualidade do produto;
- Viabilidade de utilização de contratação para sistemas em sustentação;
- Esforço adequado dos profissionais de TI do TST na especificação de cenários de teste;
- Simplicidade e clareza na contagem dos produtos entregues;
- Agilidade na contagem;
- Diminuição de subjetividade nas contagens;
- Capacitação simplificada de profissionais para realização da contagem;
- Redução de custos com controle;
- Redução de custo com empresa contratada para realizar contagem de pontos de função;
- Maior facilidade no gerenciamento dos serviços;
- Maior transparência de custos, uma vez que os produtos e serviços contratados estarão alinhados aos objetivos corporativos e que a remuneração por resultados cria uma relação convergente entre a remuneração do fornecedor e os objetivos institucionais.
- Manutenibilidade dos sistemas pós-entrega.
- Maior qualidade dos sistemas.

Conclusão

A métrica PEEX está alinhada diretamente ao resultado da prestação do serviço pois especifica e mensura esforço despendido pelos profissionais na construção do software alinhado com os objetivos de negócio dos usuários demandantes de solução de *software*, ademais, a métrica auxilia no aumento da qualidade do *software* entregue uma vez que toda a especificação da aplicação construída no TST é automatizada gerando benefícios de manutenibilidade e sustentação posterior após entrega do produto.

A Coordenadoria de Desenvolvimento de Sistemas do TST realizou estudo de viabilidade para utilização da métrica PEEX. Esse estudo consta no Estudo Técnico Preliminar do Processo Administrativo 503.735/2017. A métrica foi apresentada para os seguintes atores:

- Coordenadoria de Licitações e Contratos do TST – CLCON. Em 19/05/2017. Participaram da reunião os integrantes administrativos da equipe de planejamento da contratação;
- Coordenadoria de Auditoria de Gestão Administrativa – CAUGE. Em 22/5/2017.
- Secretário de Tecnologia da Informação – SETIN. Em 22/5/2017;
- Tribunal de Contas da União – TCU. Em 31/5/2017. Participaram da reunião cinco auditores desse Órgão de Controle;
- Representante do Gartner Group. Em 6/6/2017;
- Audiência Pública (10 empresas convidadas). Em 14/6/2017. Participaram desse evento 16 profissionais representantes das seguintes empresas: Capgemini, Indra, GlobalWeb, CTIS, Avanade, Cast e Polisy;
- Consulta Pública. Ampla participação da comunidade. No período de 12 a 26 de julho de 2017;
- Agilistas e servidores públicos no Congresso Agile Trends Gov 2017 – Brasília. Em 16/08/2017;

Vários desses atores responderam positivamente à proposição da métrica, inclusive muitas empresas, tanto as participantes da Audiência Pública quanto do Congresso Agile Trends Gov, deram retorno positivo quanto à tratativa de remuneração mais justa, ou seja, diretamente proporcional ao esforço despendido para a construção de *software*.