

Os ilusórios 20% de tempo

Esse é o texto original da barra lateral do capítulo 2, na utilização da capacidade, de Alexei Zheglov. Tivemos que editá-lo por razões de espaço para o livro.

Os 20% de tempo tem se tornado popular na indústria de software nos últimos anos. Mesmo que a maioria dos engenheiros de software não trabalhem em empresas que tenham os 20% de tempo, a maioria tem ouvido ou conhece alguém que trabalha em um lugar como o Google, onde os funcionários gastam 80% de suas horas trabalhando no que a empresa exige que eles façam e 20% em seus próprios projetos. Ou é o que nos tem sido dito.

A loja do outro lado da cidade está fazendo isso e agora queremos fazer isso também. Muitos tentaram introduzir os 20% de tempo em seus locais de trabalho e isso tem se provado bem difícil. Então, como podemos fazer isso? Quais são os prós e os contras? Há alguma teoria por trás dessa prática?

A principal razão para os 20% do tempo é manter a utilização média da capacidade em 80% em vez de 100%.

Podemos pensar em empresas de desenvolvimento de software como um sistema que transforma pedidos de funcionalidades em funcionalidades desenvolvidas. Podemos então, modelar seu comportamento utilizando a teoria das filas.

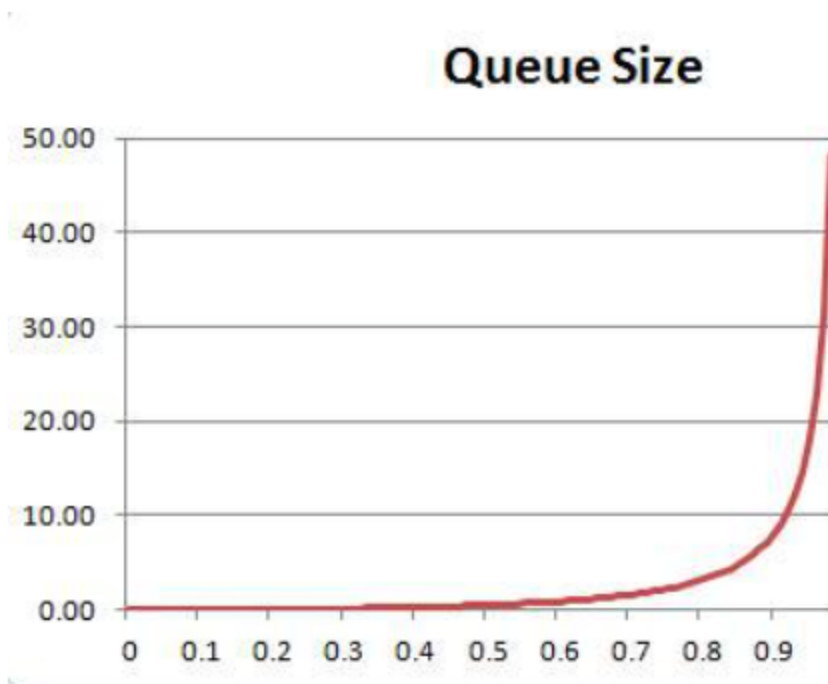
Teoria

Se os pedidos chegam mais rápido do que o sistema pode atendê-los, eles entram na fila. Quando as chegadas são mais lentas, o tamanho da fila diminui. Como as chegadas e os processos de atendimento são aleatórios, o tamanho da fila muda aleatoriamente com o tempo. Os inclinados matematicamente podem se perguntar sobre esta aleatoriedade: deve haver alguma distribuição de probabilidade, então, qual será o tamanho da fila, em média? A Matemática (teoria das filas) tem uma resposta para isso:

$$N = \frac{\rho}{1-\rho}$$

(Aqui a letra grega rho denota o coeficiente de aproveitamento igual a razão das taxas de atendimento e de chegada). Esta fórmula simples se deve ao pressuposto de que as chegadas e os processos de atendimento são os chamados processos de Markov. É uma aproximação das chegadas

no mundo real e dos processos de atendimento. A matemática é mais complicada para os processos não-Markov, mas leva a mesma conclusão.)



Traçando essa função podemos ver que o tamanho médio da fila permanece baixo enquanto a utilização é de 0,8, em seguida, sobe de forma acentuada e vai para o infinito. Podemos entender isto intuitivamente pensando sobre o CPU do nosso computador: quando sua utilização é baixa, ele responde instantaneamente aos nossos insumos, mas quando uma tarefa em segundo plano empurra sua utilização próximo de 100%, o computador torna-se frustrantemente lento a cada clique.

Prática

A economia de desenvolvimento de software é tal que as empresas de software incorrem em custos muito elevados de atraso quando suas filas estão em estados de alta fila. Isso inclui perda de oportunidades de mercado, produtos obsoletos, projetos atrasados, e gastos causados pela construção de funcionalidades em antecipação à demanda. Os 20% de tempo são, portanto, a resposta científica ao problema de otimização de resultados econômicos: evitar os estados de alta fila, evitando a alta utilização que os causa. É essencialmente a folga que mantém o sistema responsivo.

Várias conclusões práticas sobre **o que não fazer** imediatamente a seguir:

- contabilidade de custos (custo do tempo dos engenheiros X o que a empresa pode/não pode pagar). O benefício econômico vem da redução dos custos de atraso.
- a criação de um sistema de submissão-revisão-aprovação do projeto de 20% do tempo
- acompanhar os 20% do tempo através do preenchimento de planilhas de horas

- utilizar a inovação como um motivador para os 20% do tempo. Enquanto novos produtos vêm de projetos de 20% do tempo, eles não são o ponto. Se sua empresa não pode inovar durante suas horas de atividade principal, isto é um problema!
- contar com 20% de tempo para estimular a criatividade. Dizer que você vai soltar sua criatividade com 20% do tempo levanta a questão de por que você ainda não é criativo o suficiente durante suas horas de trabalho de atividade principal.
- alocar os 20% de tempo à sexta-feira todas as semanas...

Essas são todas as questões a não fazer, onde está o que fazer?

Ok, e que tal fazer isso certo? Vamos responder com a melhor pergunta que ouvimos ao discutir este assunto com profissionais: "Se 20% de sua capacidade é delegada a ser preenchida com itens não enfileirados, então você somente encolhe sua capacidade de 80%, e os 80% são seu novo 100%. Certo?"

Sim, "80 é o novo 100" □ destaca o principal problema com as tentativa de ordenar os 20% de tempo sem entender a teoria. Você quer escapar da armadilha de utilização, não continuar na armadilha e alocar o tempo de forma diferente.

Lembre-se que a utilização depende de dois processos: processos de chegada (demanda) e processos de serviço (capacidade). Você realmente não pode escolher sua utilização. Ele é o que é porque os processos são o que são. Você pode, no entanto, trabalhar nos processos: ao melhorar a capacidade de entrega de software e moldar a demanda. Ao fazer progresso, a folga vai surgir.

Escape da armadilha da utilização!