

4: SBRC

RECIFE - 24 A 26 DE MARÇO 86

Apropriação de uso dos recursos computacionais em uma rede de computação com vistas a contabilização de preço.

Paulo Guilherme Teixeira Freire

SERPRO (Serviço Federal de Processamento de Dados)

DIOPE/DTE/GPTEL - Brasília - DF

Sumário:

Inicialmente é apresentada a dificuldade que existe em se realizar a apropriação do uso dos recursos computacionais em um rede de computação. Paralelamente também é apresentada uma deficiência existente na forma mais tradicional de se realizar o faturamento de usuarios de aplicações multiusuarios "on-line".

Em seguida são propostas algumas soluções para os problemas apresentados, e analisadas as vantagens e desvantagens de cada uma delas.

Finalmente é apresentado um estudo que mostra a viabilidade da solução considerada mais eficiente.

1. INTRODUÇÃO

O SERPRO (Serviço Federal de Processamento de Dados) possui atualmente, assim como inumeras outras empresas, a sua propria rede de computação. Ainda igualmente a tantas outras empresas, a sua rede de computação se baseia no modelo popularmente conhecido pela denominação "estrela".

A utilização do modelo em "estrela" no que diz respeito à apropriação dos recursos computacionais utilizados por serviços "on-line" é vantajosa, visto que todas as informações

necessárias se acham concentradas em uma só máquina. Assim o processo de coleta de dados é relativamente simples.

No entanto uma rede desse tipo possui algumas desvantagens em outros aspectos, desvantagens essas que tendem a se agravar com o crescimento da rede.

Devido a isso o SERPRO deu início ao desenvolvimento de um projeto (denominado Projeto Teleinformática) cujo objetivo é o de promover o desenvolvimento de uma rede de computação capaz de minimizar, ou até mesmo eliminar, as desvantagens existentes no modelo atual.

O modelo de rede preconizado pelo Projeto Teleinformática prega, em resumo, a distribuição de atividades entre vários processadores. Esse modelo, que resolve vários dos problemas existentes no modelo anterior, possui no entanto uma desvantagem no que se refere a apropriação da utilização dos recursos computacionais. Isso porque com esse modelo deixa de existir a figura do processador que detem todos os dados necessários para a apropriação de utilização dos recursos computacionais, tal como no modelo "estrela".

Independente e paralelamente a isso também foi percebido, durante o estudo em questão, uma deficiência nos modelos tradicionais de faturamento de usuário com relação aos serviços "on-line". Essa deficiência, em resumo, resulta do fato de que muitas vezes são apresentadas faturas com base em medidas que tornam difícil, se não mesmo impossível, ao usuário realizar qualquer tipo de aferição. Exemplos de tais medidas utilizadas são o tempo total de CPU utilizada, a quantidade de bits transmitidos, etc.

Este artigo se propõe a mostrar esses dois problemas bem como propor soluções para os mesmos.

Este artigo foi dividido em seis partes, sendo a primeira esta introdução. O segundo tópico é dedicado a uma breve explicação sobre o modelo de rede proposto pela projeto Teleinformática. O terceiro tópico visa a explicar melhor ambos os problemas existentes. No quarto tópico são apresentadas as alternativas visualizadas, bem como vantagens e desvantagens de cada uma delas. O quinto tópico é dedicado ao estudo que prova a viabilidade de uma dessas soluções, e finalmente no sexto tópico é apresentada a conclusão.

2. O MODELO DA NOVA REDE SERPRO

O Projeto Teleinformatica, que se encontra atualmente em fase de implantação de uma rede-piloto, preve para a nova rede SERPRO uma arquitetura baseada em Processadores de Comunicações (PC's). Os Processadores de Comunicações são mini-computadores destinados a controlar toda a rede.

Assim a realização de qualquer transação na rede sera controlada pelos PC's. Todas as mensagens transmitidas e recebidas por um usuario da rede passam necessariamente pelos PC's. A fim de facilitar a visualização disso e' apresentada a figura 1.

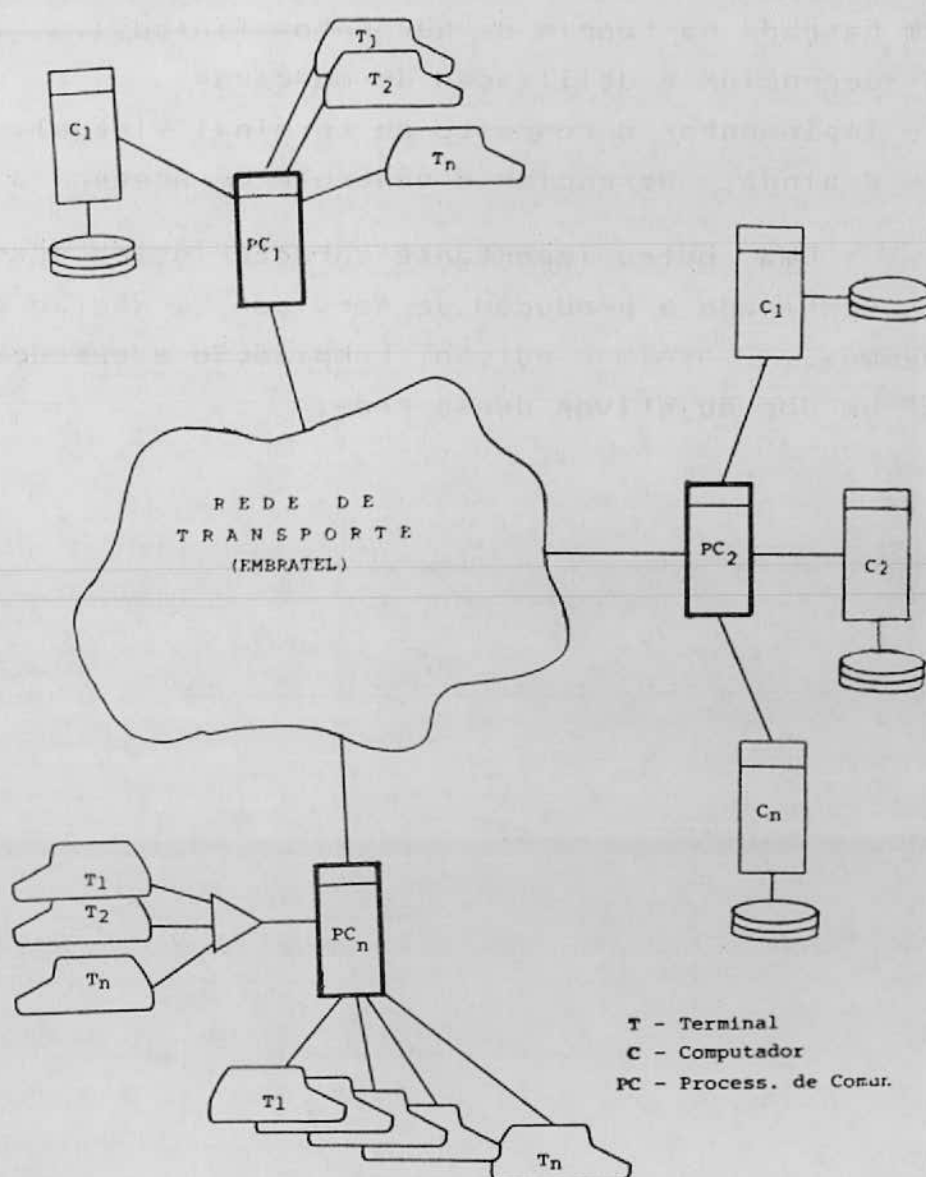


Figura 1. Esquema basico da Rede Serpro de Tratamento de Dados.

A título de curiosidade a rede-piloto, ora em desenvolvimento, preve a instalação de 4 (quatro) processadores de comunicações (PC's), sendo dois em Brasília, um no Rio de Janeiro e um em São Paulo. Os mini-computadores a serem utilizados na rede piloto serão todos eles do modelo COBRA 480.

Entre outras tarefas esses mini-computadores irão:

- "Frontalizar" todos os processadores ("hosts's") hospedeiros de aplicações.
- Concentrar todos os terminais conectados à rede.
- Gerenciar a porção de controle das aplicações (para tanto a porção de controle das aplicações serão descritas em uma linguagem baseada na teoria de automatos finitos).
- Gerenciar a utilização de mascaras.
- Implementar o conceito de terminal virtual.
- e ainda, Gerenciar o controle de acesso à rede .

Uma outra importante característica dessa rede é' que ela é' destinada a produção de serviços, e não ao desenvolvimento das mesmas. Assim a edição, compilação e execução de programas não é' um dos objetivos dessa rede.

3. OS DOIS PROBLEMAS

3.1 O PROBLEMA DA APROPRIAÇÃO DO USO DE RECURSOS COMPUTACIONAIS EM UMA REDE DE COMPUTAÇÃO NA ACEPÇÃO RESTRITA DA PALAVRA

A apropriação de recursos computacionais com vistas à contabilização de preços exige duas informações básicas: a quantidade de recursos consumidas em um determinado período (medida, por exemplo, em tempo de CPU, tempo de IO, integral de memória, etc.) e a identificação do usuário que fez uso desses recursos nesse mesmo período.

Em um modelo de rede em "estrela" existe um processador central que detém todas as informações necessárias para a apropriação do uso dos recursos computacionais. Esse processador pode registrar em seu "log" tanto a quantidade de recursos consumidos quanto a identificação dos usuários.

Foi visto anteriormente que o novo modelo de rede do SERPRO será baseado na figura dos processadores de comunicações, e que uma das principais tarefas assumidas pelos PC's é o controle de acesso à rede.

Dessa forma os PC's possuirão a informação de quem fez uso dos recursos computacionais da rede, mas somente os próprios processadores hospedeiros de aplicações é que possuirão as informações acerca da quantidade de recursos computacionais consumidos. Assim sendo, as informações que antes se encontravam concentradas em um único processador passam a estar divididas em dois processadores.

E mais, com as facilidades que a nova rede irá oferecer pode-se esperar no futuro que :

- i) Sejam construídas aplicações residentes em mais de um processador hospedeiro de aplicações ;
- ii) Que os usuários realizem acessos a rede via processadores de comunicações (PC's) distintos ;

Tudo isso contribui para dificultar ainda mais a realização da apropriação de recursos computacionais discriminado por usuário.

3.2 O PROBLEMA DO MODELO TRADICIONAL DE FATURAMENTO

Provavelmente por razões históricas ocorre hoje em dia o fato anteriormente citado de se faturar um usuário com base em informações do tipo "quantidade de tempo de CPU consumido". A fim de facilitar a compreensão deste problema será realizada a seguinte analogia.

Por exemplo, as companhias telefônicas costumam tarifar as ligações interurbanas por tempo (e também por "degrau tarifário")!

Uma razão para tanto é a de permitir ao usuário a confirmação de sua utilização, ou seja, o usuário dispõe de meios concretos que lhe permitam aferir uma fatura.

Uma segunda razão é que isso permite ao usuário saber, a priori, qual será exatamente o custo (para si) de uma ligação de "x" minutos. Assim sendo, não importa que a ligação interurbana solicitada faça uso de duas, três ou quatro sub-estações, ou ainda que faça uso de satélites, que o preço a ser cobrado do usuário será exatamente o mesmo (respeitado também o esquema de "degrau tarifário").

Em suma, deve-se fornecer ao usuário (de sistemas multi-usuários) um meio do mesmo aferir a sua fatura, além de um meio do mesmo poder estimar o custo (para si) de uma consulta.

Uma aplicação "on-line" multi-usuário que seja faturada de acordo com os tempos de CPU contabilizados no LOG sofre da deficiência de não poder fornecer a priori ao usuário um "orçamento" para uma consulta, além de não poder fornecer ao usuário um meio do mesmo aferir a fatura apresentada. Como saber, antes da consulta ser efetuada, quantos segundos de CPU serão gastos? É mais importante, como confirmar, após a realização da consulta, a quantidade CPU utilizada?

4. AS SOLUÇÕES, SUAS VANTAGENS E DESVANTAGENS

A fim de resolver o primeiro problema apontado foram vislumbradas as seguintes alternativas :

a) Não realizar apropriação de recursos computacionais discriminado por usuário.

b) Cada uma das partes (o PC e o processador hospedeiro das aplicações) deve apropriar as informações que possuir disponíveis. Após um determinado período os arquivos "log's" de ambas as partes devem ser confrontados. Como informação chave para ser realizado o confronto de registros poderia ser utilizado um número sequencial associado à mensagem, ou ainda a hora da mensagem.

c) Toda a apropriação deve ser realizada nos PC's. Para tanto, sempre que um consulta (ou uma sessão) for efetuada, o computador hospedeiro deve fornecer a quantidade de recursos computacionais gastos naquela consulta (ou sessão).

d) Toda a apropriação deve ser realizada nos computadores hospedeiros das aplicações. Para tanto cada mensagem enviada do PC ao "host" deve ser acompanhada de uma identificação do usuário (esse esquema também pode ser imaginado para funcionar a nível de sessão).

Além dessas foi ainda visualizada mais uma alternativa, a saber:

e) Toda a apropriação deve ser realizada nos PC's, porém sem a necessidade do computador hospedeiro enviar em cada mensagem, ou a cada sessão, a quantidade de recursos computacionais consumidos. A fim de suprir tal informação deve ser montada (pelos administradores da rede) uma "tabela" de estimativa de consumo de recursos computacionais por aplicação. Esta "tabela" deve possuir como entrada alguma medida passível de apropriação pelo PC e como saída a estimativa de consumo de recursos computacionais.

Com os olhos voltados para a contabilização de preços pode-se dizer que a primeira alternativa propõe simplesmente que todos os usuarios da rede sejam cobrados identicamente, por um valor padrão, independentemente da quantidade e da qualidade dos serviços realmente utilizados.

Essa solução, por motivos obvios, foi descartada.

A segunda alternativa também foi descartada pelos seguintes motivos :

A utilização da hora como informação para realizar o fechamento dos "log's" não é confiável, devido aos problemas existentes para a realização de uma perfeita sincronização entre a máquina PC e o processador hospedeiro da aplicação. Dessa forma se impõe a utilização de um numero sequencial de mensagem (ou sessão). Em consequencia tem-se a desvantagem da necessidade de espaço gerada por essa informação, que devera' circular agregada a cada mensagem (ou sessão) a transitar pela rede.

Porém o que inviabiliza esta alternativa é a desvantagem que o fechamento de "log's" devera' ser realizado não só entre todos os Processadores de Comunicação, mas como também entre todos os processadores hospedeiros de aplicação.

A terceira alternativa requer que a cada mensagem (ou sessão) a transitar pela rede sejam informados todos os dados relativos ao consumo de recursos computacionais. Essa alternativa possui como forte desvantagem a quantidade de espaço necessária para conter todas as informações relativas ao consumo de recursos computacionais (tais como por exemplo, consumo de CPU, consumo de IO, consumo de memoria, etc).

Alem disso esta alternativa exigiria modificações severas nos monitores de Teleprocessamento residentes no processadores de aplicações, a fim de que os mesmos pudessem captar as informações necessarias.

Devido a isso, tal alternativa também foi descartada.

A quarta alternativa também requer que a cada mensagem (ou sessão) a transitar pela rede seja acoplada uma informação, no caso, a identificação do usuário. Possui portanto a desvantagem da quantidade de espaço necessária para conter as informações relativas ao usuário, e que devem estar juntas a cada mensagem (ou sessão) que transitar pela rede.

Essa alternativa também exige que sejam realizadas algumas modificações nos monitores de TP dos processadores hospedeiros das aplicações.

Porem de todas as alternativas que preservam o que se pode denominar por custeio histórico (ou real) (alternativas "b", "c" e "d") essa é a que apresenta menos desvantagens, devendo por isso ser preservada como uma boa opção (A necessidade de se possuir uma alternativa capaz de realizar o custeio histórico é explicada principalmente pela necessidade de se possuir um bom meio de se realizar a apropriação de recursos computacionais voltada para a contabilização de custos).

A quinta e última alternativa é, a princípio, vantajosa sobre as demais, visto que não possui a desvantagem de requerer espaço para a agregação de informações em cada mensagem que transita pela rede, nem de requerer modificações nos monitores de TP dos processadores hospedeiros de aplicação.

Porem essa alternativa implica na adoção de um esquema de "preço-padrão" por aplicação (ou por subrotinas das aplicações). Assim sendo, a viabilidade desse esquema esta sujeita a resposta da seguinte questão :

"Dois usuários distintos, com atividades distintas, mas que façam uso de uma mesma aplicação, terão uma mesma razão de utilização dos recursos computacionais ?".

Razão de recursos computacionais, no caso, significa a relação de utilização dos recursos computacionais (por exemplo, o tempo de CPU, a quantidade de memória, etc.) comparado a uma medida previamente especificada (por exemplo, a quantidade de transações, a quantidade de bits transmitidos, etc.).

No momento e' conveniente deixar essa questao em aberto, e passar a analise da relacao das solucoes apresentadas com o segundo problema mencionado. Notar que todas as alternativas foram apresentadas com os olhos voltados unicamente para a solucao do primeiro problema (o problema apresentado em 3.1).

Inicialmente deve ser observado que e' inerente a resolucao do segundo problema uma substituicao da medida empregada.

Deve-se substituir as medidas mais tradicionais, tal como por exemplo o tempo de CPU consumido, por outras medidas passiveis de afericao pelo usuario.

Como solucao foram imaginadas a utilizacao das seguintes medidas :

- i) Tempo-de-sala
- ii) Quantidade de bits transmitidos
- iii) Quantidade de transacoes.

A utilizacao do tempo-de-sala resolve o ultimo problema apresentado, mas em contrapartida sofre da desvantagem da correlacao do tempo-de-sala com o consumo real de recursos computacionais nao ser muito boa, alem de outras deficiencias tal como por exemplo, a dificuldade que existe para o computador hospedeiro da aplicacao ou para o proprio PC perceber a "queda" de um terminal assincrono (nesses casos, se nao for adotado um esquema alternativo ocorreria um superfaturamento sobre o usuario). Deve por isso ser descartada.

A quantidade de bits transmitidos foi incluída ai' devido ao argumento de que se o usuario for auxiliado pelo proprio computador que controla o terminal, o usuario conseguira' realizar uma afericao adequada. Assim a utilizacao da quantidade de bits transmitidos e' uma alternativa. Apesar da mesma nao fornecer um meio do usuario saber a priori o valor de uma consulta, o usuario pode estima-lo com alguma seguranga. O SERPRO hoje faz uso de tal medida para o faturamento de alguns de seus usuarios.

A utilizacao da quantidade de transacoes e' uma alternativa que supera as desvantagens citadas anteriormente. E' uma medida que pode ser facilmente aferida pelo usuario, alem de permitir ao mesmo saber a priori qual sera' o custo (para si) de

uma determinada consulta. Esta alternativa pode ser comparada com alguns sistemas tradicionais, tal como por exemplo sistemas de folha de pagamento, que mesmo sendo executados em "batch" são cobrados por funcionario processado.

Visto isso pode-se optar pela utilização ou da medida quantidade de bits transmitidos ou pela medida quantidade de transações, a fim de solucionar o segundo problema (problema 3.2).

Observar agora que a ambas as medidas selecionadas podem ser apropriadas tão somente pelos Processadores de Comunicações. Devido a isso (e também pela substituição das medidas tradicionais) a alternativa "d" perde a sua força.

Em consequencia tem-se que a unica alternativa capaz de resolver ambos os problemas e' a alternativa "e".

Dado que essa alternativa e' a mais indicada resta agora resolver a questão que ficou em aberto, ou seja, verificar a razão da utilização dos recursos computacionais com relação as medidas selecionadas (quantidade de transações e quantidade de bits transmitidos).

Para resolver tal questão foi realizado um estudo que se encontra descrito no topico seguinte.

5. O ESTUDO DA CORRELAÇÃO DO CONSUMO DE RECURSOS COMPUTACIONAIS COM A QUANTIDADE DE TRANSAÇÕES E COM A QUANTIDADE DE BITS TRANSMITIDOS

5.1 INTRODUÇÃO DO ESTUDO

No topico anterior foi sugerido a utilização da apropriação de recursos computacionais (com vistas à contabilização de preços) baseada unicamente em informações disponiveis nos Processadores de Comunicações. As medidas eleitas foram a quantidade de transações e a quantidade de bits transmitidos.

No entanto qual sera' a relação entre essas medidas e o consumo real de recursos computacionais? Em outras palavras, essas duas formas de medição são a principio boa para o usuario, mas serão tambem boa para o administrador dos recursos computacionais?

Curiosamente as intuições pessoais de cada um a respeito são tão diversas quanto eram as intuições pessoais de cada um anos atras quando se colocava em questão a produtividade de um programador. Naquela epoca a afirmação de que um programador tipico produzia "tão somente" 10 linhas-fonte por dia era considerada por muitos um absurdo.

Assim sendo, a fim de dirimir duvidas e basear as discussões futuras em fatos e não em sentimentos, foi realizado um estudo a respeito. Neste topico foi descrito apenas um resumo do estudo realizado. Para maiores detalhes aconselha-se a leitura da referencia /FREIRE 85/.

5.2 OS DADOS ANALISADOS

Para que o estudo fosse efetivado era necessario obter dados reais de sistemas multi-usuarios. Alem disso era necessario que os dados a serem obtidos possuíssem ao menos as tres variaveis basicas estabelecidas, a saber: uma medida acerca dos recursos computacionais utilizados, uma medida acerca da quantidade de transações e uma medida acerca da quantidade de bits transmitidos.

Dando inicio à fase de levantamento de dados foi verificado que no SERPRO ja' existia, ha' muito, um bom sistema de apropriação de recursos (SAR), que sistematicamente coleta dados a respeito das aplicações processadas na empresa. Dessa

forma a fase de levantamento de dados foi em muito facilitada.

Entre as varias medidas coletadas e disponiveis tem-se:

- a) A quantidade de UPP's consumidas (UPP significa Unidade Padrão de Processamento).
- b) A quantidade de transações
- c) e A quantidade de bit's transmitidos

Dado que este estudo se propoz a analisar os recursos consumidos pela maquina se utilizou a medida UPP que em resumo e' uma medida que pondera os tempos de CPU consumidos com a maquina real utilizada. Adicionalmente leva em consideração tambem a quantidade de operações de entrada e saída.

As medidas quantidade de transações e quantidade de bits transmitidos são as duas outras medidas basicas necessarias para a realização do estudo, e que se encontram disponiveis exatamente na forma originalmente idealizada.

Os dados analisados foram apropriados pelo SAR, e dizem respeito a todos os sistemas de teleprocessamento processados na I.URO (Brasilia), no mes de maio de 1985. Alem disso foram executadas algumas outras analises sobre dados semelhantes obtidos a partir dos sistemas de teleprocessamento processados nos meses de fevereiro, março e abril de 1985.

A titulo de curiosidade, durante o mes de maio de 1985, nos sistemas de teleprocessamento foram consumidas cerca de 34,5 UPP's e foram transmitidos cerca de 1,1 bilhão de bits e foram executadas cerca de 200.000 transações. Isso resulta em uma media de aproximadamente 1,6 UPP's/dia util, 50 milhões de bits/dia util e aproximadamente 9.000 transações/dia util.

Durante esse mesmo periodo (maio de 1985) 40 usuarios fizeram uso de 17 sistemas.

Dos sistemas quatro foram excluidos da analise por terem sidos utilizados por poucos usuarios distintos (tres ou menos). Alem deles um outro "sistema" foi excluido da analise devido ao fato de que a maioria dos usuarios contabilizados (11 de um total de 13) terem realizados menos de 10 transações durante todo o mes.

5.3 AS ANÁLISES ESTATÍSTICAS REALIZADAS

Para cada um dos 12 sistemas disponíveis foram realizadas análises de regressão linear sobre as variáveis UPP, quantidade de transações e quantidade de bits transmitidos.

A utilização de regressão linear múltipla entre as três variáveis foi tentada, porém, conforme será visto logo adiante, a sua utilização se mostrou tecnicamente contraproducente devido à ocorrência de um fenômeno denominado por multicolinearidade. Além disso também não deve ser esquecido um dos objetivos primordiais do trabalho que é o de fornecer um meio alternativo mais simples para a apropriação de preço (e/ou utilização dos recursos computacionais).

Em cada análise de regressão linear realizada o primeiro dado a ser analisado é o valor de "F" calculado (ou "F_c" ou ainda simplesmente "F"). Este valor é assim denominado devido a sua distribuição obedecer à uma distribuição F de Snedecor. O valor de F calculado tem por objetivo dizer se existe ou não regressão.

No caso de se chegar a conclusão de que realmente existe regressão, passa-se a análise do valor de R^2 , a saber o coeficiente de determinação. O R^2 é uma das medidas mais importantes de um modelo de regressão linear e tem por finalidade indicar qual o grau de explicação do modelo, ou seja, em que proporção as variáveis independentes do modelo explicam a variável dependente (Obs: em uma equação do tipo $Y = f(X)$ diz-se que X é a variável independente e que Y é a variável dependente).

O valor do coeficiente de determinação pode variar de 0 a 1 (ou de outra forma, de 0% a 100%). Por exemplo, quando duas variáveis são perfeitamente correlacionadas, isto é uma e é exatamente uma função matemática da outra, o valor de R^2 é 100%. De modo inverso, quando duas variáveis não possuem nenhuma correlação linear o valor de R^2 é de 0% (Na realidade basta o valor de R^2 ser próximo de 0% para se concluir pela inexistência de correlação linear entre as duas variáveis).

A fim de se ter uma melhor ideia sobre o coeficiente de determinação, para fins de comparação, foi montada a seguinte tabela :

Valor do coeficiente de determinação	Força da correlação
Menor que 50 %	Correlação muito fraca ou inexistente.
50 % < R ² < 70 %	Correlação fraca.
70 % < R ² < 80 %	Correlação regular.
80 % < R ² < 90 %	Correlação forte.
Maior que 90 %	Correlação muito forte.

Tabela 1. Classificação dos valores de R².

Apos a análise do valor de R² pode-se passar à análise acerca da hipótese de "nulidade" dos coeficientes na equação de regressão.

Para tanto, para cada coeficiente da equação (e entre eles inclui-se a constante da equação) calcula-se um valor (denominado de "t") que obedece a uma distribuição t de Student. Analisando-se este valor pode-se concluir pela nulidade ou não desse coeficiente na equação.

Também não deve ser esquecido a importante contribuição dos gráficos para uma análise de regressão. Através de um gráfico se torna possível analisar os valores reais observados sob um outro ângulo de visão. Um valor real observado que não se enquadre na regressão possivelmente significa alguma situação anômala para a qual deve ser procurada uma explicação particular.

A verificação de valores anômalos pode ser feita visualmente com a ajuda de gráficos, bem como por meio de testes de hipóteses. Para tanto pode-se construir para o modelo um intervalo de confiança, dentro do qual se pode aceitar a hipótese de normalidade dos valores.

Para finalizar deve ser dito que, neste trabalho em particular não é de muita importância o conhecimento puro e simples dos coeficientes da equação, nem a realização de um intervalo de confiança para a reta de regressão. Isso porque neste trabalho está-se interessado em simplesmente provar ou não a existência de correlação entre as variáveis.

5.4 RESULTADOS OBTIDOS

Para cada um dos doze sistemas multi-usuários disponíveis foram realizadas análises de regressão linear simples entre as variáveis UPP total e quantidade de transações, e UPP total e quantidade de bits transmitidos. Adicionalmente foram desenhados os gráficos de todas as análises de regressão realizadas, alguns dos quais poderão ser visto logo adiante.

Também foram realizados cálculos para verificar a possibilidade de se realizar uma análise de regressão linear múltipla entre a variável UPP e as variáveis qtde. de bits transmitidos e quantidade de transações. Para tanto foi calculado, para cada caso, o coeficiente de correlação linear (de Pearson) entre as duas variáveis independentes. É importante lembrar que quando esse coeficiente é alto uma das duas variáveis deve ser retirada do modelo de regressão, sob o risco de estar-se recaído no problema da multicolinearidade (este problema, em suma, causa uma elevação artificial tanto no valor do coeficiente de determinação quanto nas variâncias dos coeficientes da equação).

Em todas as análises realizadas o valor do coeficiente de correlação linear entre aquelas duas variáveis se mostrou muito elevado (sempre maior que 0,7) o que desaconselhou a análise de regressão linear múltipla.

Foi realizada no estudo a análise de regressão linear simples de todos os sistemas, com os dados exatamente na forma como foram fornecidos pelo SAR, para ambos os casos (UPP x qtde.de.transações e UPP x qtde.de.bits.transmitidos).

Em todas as análises realizadas provou-se haver de fato correlação positiva entre as variáveis. Ou estatisticamente falando, os valores de F calculado sempre foram superiores (e na verdade muito superiores) aos valores tabelados de F ao nível de 0.5% de significância.

Além disso o coeficiente de determinação na maioria absoluta dos casos se mostrou muito forte. Veja os valores obtidos na tabela 2.

SISTEMA	Qtde. de dife- rentes usuarios	Coeficiente de Determinação entre	
		UPP total e Qtde. de trans.	UPP total e Qtde. de bits tr.
1	6	99.9 %	99.8 %
2	9	99 %	99 %
3	15	99 %	98 %
4	18	97 %	99 %
5	15	97 %	98 %
6	19	97 %	92 %
7	17	96 %	92 %
8	17	94 %	96 %
9	22	91 %	45 %
10	20	88 %	93 %
11	37	84 %	87 %
12	16	76 %	69 %

Tabela 2. Coeficientes de determinação observados.

Feito isso passou-se a análise dos coeficientes da equação, mais exatamente a análise de significância da constante da equação (teste da hipótese de nulidade da constante).

Em uma regressão linear simples a expressão linear pode ser escrita da seguinte maneira: $Y = a.X + b$. É relativamente óbvio que, na análise em curso, a constante "b" tenha o valor zero. Isso partindo do princípio de que se o usuário não realizar nenhuma transação ele não terá gasto nenhum recurso computacional. O mesmo pensamento se aplica em relação à variável bits transmitidos.

Assim sendo foram realizados testes de hipótese para verificar se, em cada caso, a constante "b" pode ser igual a zero.

Os resultados encontrados, em todas as análises, mostraram que tal hipótese é verdadeira, ou seja, em todas as equações de regressão a constante verdadeira (porém desconhecida) pode ser zero.

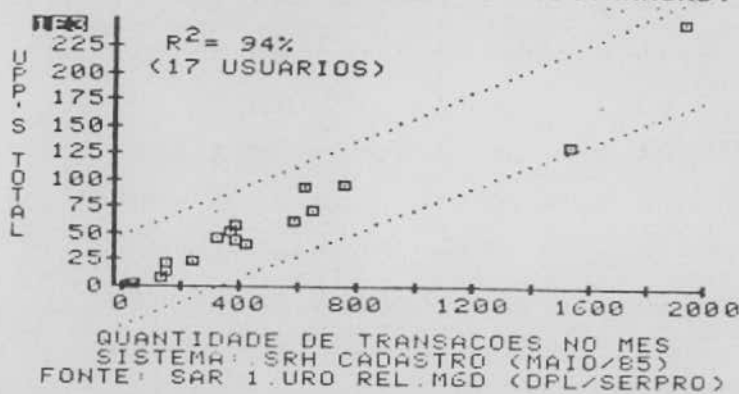
Dado que essa hipótese é verdadeira poderia-se realizar novas análises de regressão, a fim de se recalcular o valor do coeficiente da variável independente (o multiplicador "a" na

equação de regressão), partindo do princípio que o valor da constante ("b") é zero.

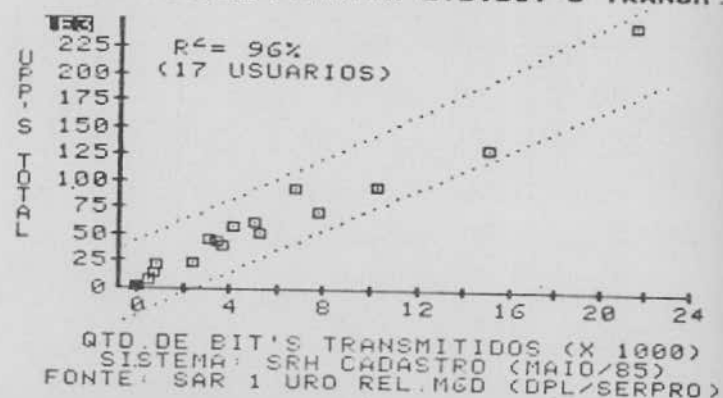
Tal procedimento não foi aqui adotado porque este trabalho não tem por objetivo encontrar a melhor equação ajustante, ou seja, os valores dos coeficientes da equação per si são irrelevantes para este estudo.

Para finalizar são apresentados tres dos graficos realizados onde se pode notar os pontos observados hem como os intervalos de confiança traçados.

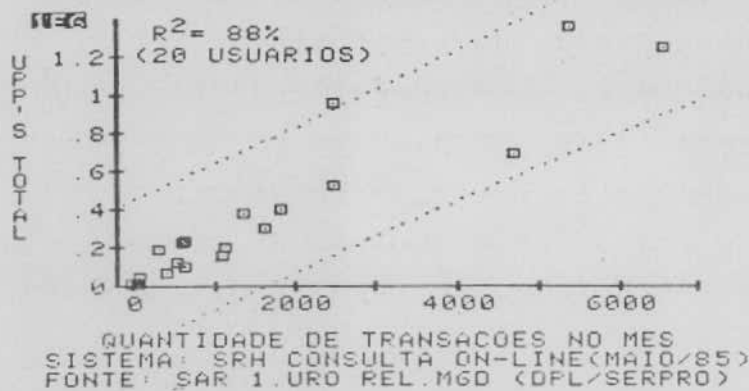
UPP'S TOTAL EM FUNCÃO DA QTD. TRANSAC.



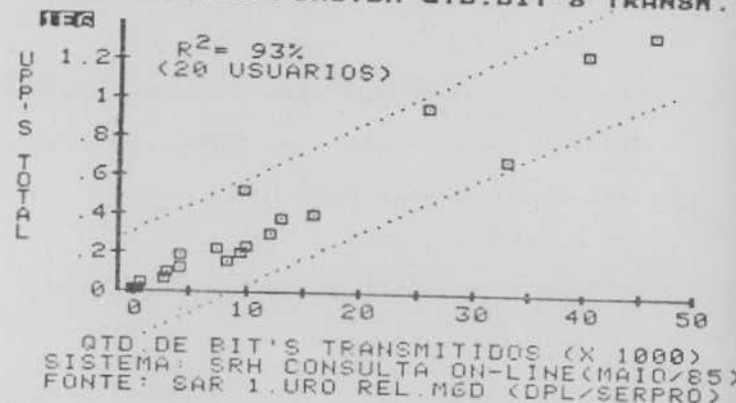
UPPS TOT.EM FUNC.DA QTD.BIT'S TRANSM.



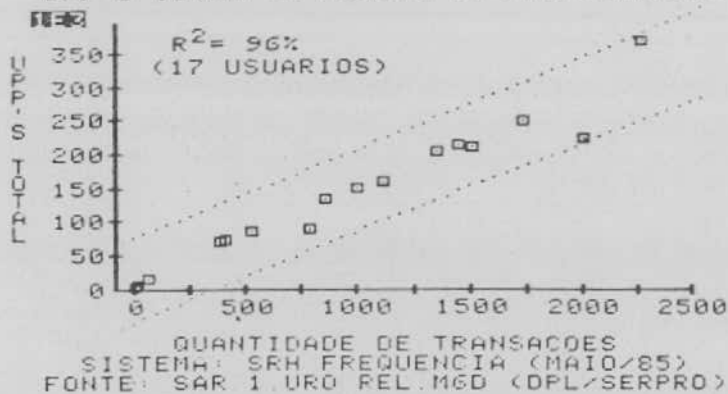
UPP'S TOTAL EM FUNCÃO DA QTD. TRANSAC.



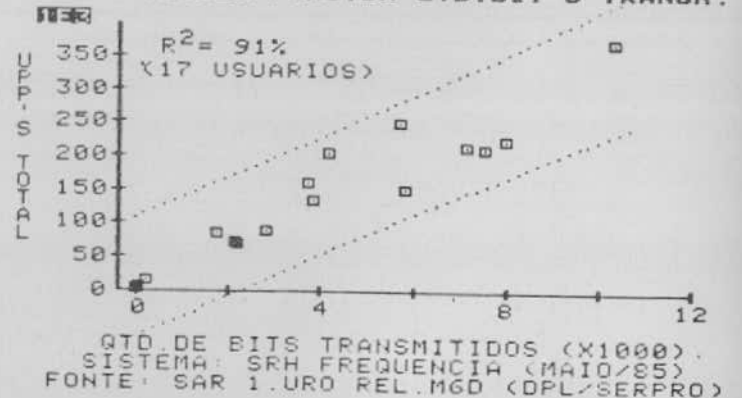
UPPS TOT.EM FUNC.DA QTD.BIT'S TRANSM.



UPP'S TOTAL EM FUNCÃO DA QTD. TRANSAC.



UPPS TOT.EM FUNC.DA QTD.BIT'S TRANSM.



5.5 CONCLUSÃO DO ESTUDO

A principal conclusão a que se chega é a de que independentemente do usuário que utiliza uma determinada aplicação existe uma forte correlação linear entre a quantidade de transações e a quantidade de recursos computacionais consumidos por essa mesma aplicação.

A mesma conclusão também é válida para a correlação existente entre a medida quantidade de bits transmitidos e a quantidade de recursos computacionais consumidos.

Assim sendo, os métodos inicialmente sugeridos, de se realizar a apropriação de preços de aplicações computacionais pelo volume de transações ou ainda pelo volume de bits transmitidos, são perfeitamente válidos.

Também é interessante observar que os dados disponíveis indicam que o grau de correlação existente na regressão UPP x Qtde.de.transações é ligeiramente superior ao grau de correlação da outra regressão. No entanto, a diferença existente entre ambos os graus de correlação é tão pequena que se torna improvável que a realização de um teste de hipóteses indique haver alguma diferença entre ambos. Independente disso não deve ser esquecido que, no aspecto considerado, a medida "qtde. de transações" só possui vantagens sobre a medida quantidade de bits transmitidos, e portanto, se em alguma ocasião for possível optar por uma das duas medidas deve-se ficar com a primeira.

6. CONCLUSÃO

A utilização da apropriação do uso de recursos computacionais exclusivamente nos Processadores de Comunicações (PC's), baseada na medida "quantidade de transações" é uma forma perfeitamente viável, e que quando voltada para a utilização na contabilização de preços elimina as desvantagens inerentes as medidas mais tradicionais.

É importante observar que isso não exclui a necessidade de se continuar realizando apropriações sobre o uso de recursos computacionais nos próprios computadores hospedeiros de aplicações, embora nesses casos tais apropriações sejam feitas a nível global de aplicação, isto é sem a discriminação por usuário. Tais apropriações são necessárias a fim de serem realizadas aferições periódicas acerca da "tabela" de estimativas utilizada.

Alem disso também seria interessante que, esporadicamente, fosse utilizado o método "d" anteriormente citado, com a finalidade de serem realizadas aferições mais precisas na "tabela" de estimativas.

Finalmente deve ser dito que apesar da viabilidade técnica desse método de apropriação, a implementação do mesmo esta sujeita a flexibilidade do sistema de custos e do sistema de faturamento existente na empresa.

BIBLIOGRAFIA

- /CARVALHO 84/ CARVALHO, Marco A. & Freire, P.G.T
Aspectos sobre a administração da nova rede
Relatorio Interno, SERPRO/DIRIN/CRT/GPTEL, 1984
- /CARVALHO 85/ CARVALHO, Marco A.
A influencia da nova arquitetura de Rede de Tele-
procesamento no ambiente de desenvolvimento do
SERPRO
Anais do 3.º Simposio Brasileiro sobre redes de
computadores (SBRC), Rio de Janeiro, 1-3 abril 85.
- /COSTA 77/ COSTA NETO, Pedro L. de O.
Estatística
Ed. Edgard Blucher Ltda, São Paulo, 1977.
- /DCC 83/ DCC
Manual de Custos
Manual Técnico, SERPRO/DCC, Julho de 1983
- /DeROSE 85/ De ROSE, Luiz A. C.
Configuração atual da rede SERPRO
Relatorio interno, SERPRO/DIOPE/DTE/GPTEL, fev.85,
37 paginas
- /FREIRE 85/ FREIRE, Paulo G. T.
Um estudo da correlação do consumo de CPU a quan-
tidade de transações e com a quantidade de bits
transmitidos
Relatorio interno, SERPRO/DIOPE/DTE/GPTEL, nov/85.
- /LEONE 77/ LEONE, George
Custos - um enfoque administrativo
FGV, Rio de Janeiro, 1977, 3.ed.rev.
- /LOPES 80/ LOPES, Roberto L. & Choi, T.J.
Determinação da Unidade Padrão de Processamento do
CTI-SP
Comunicação técnica do SERPRO/CTI, N. 80/01,
São Paulo, 22/JAN/80.

/MALVAR 78/ MALVAR, Henrique T.
Modelos Econometricos
Programa do curso de Formação de assessores em
metodos Quantitativos, IE-EST/UnB (Universidade de
Brasilia), 1978.

AGRADECIMENTOS

Sem a colaboração de inumeras pessoas e orgãos da empresa a realização deste trabalho certamente não seria possível. Mesmo correndo o risco de ficar em falta com alguém gostaria de agradecer nominalmente : A DPL na figura do Ibrahim Armele, e aos idealizadores e implementadores do sistema SAR, pelo fornecimento dos dados analisados e das explicações fornecidas; A DCC na figura do Wagner e Carlos pelas conversações mantidas sobre o sistema de custos da empresa; O CTI de uma forma geral, pela troca de ideias que se tornaram mesmo em um novo ponto de partida deste trabalho; A DIREC na figura de Ricardo Ribeiro pelo fornecimento do pacote estatístico utilizado; E finalmente ao GPTEL e ao grupo de desenvolvimento do novo sistema Senha pela oportunidade proporcionada e pela convivencia durante o periodo de desenvolvimento do trabalho.