**Uma Contribuição ao
Estudo de Estabilidade
em Sistemas de Manufatura**

Tito Manlio Brelaz Homem de Mello

DISSERTAÇÃO APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO

PARA OBTENÇÃO DO GRAU DE
MESTRE EM
MATEMÁTICA APLICADA

Área de Concentração: **Ciência da Computação**
Orientador: **Prof. Dr. Carlos Humes Jr.**

–São Paulo, Dezembro de 1992–

Aos meus Amigos

Resumo

Neste trabalho estudamos a classe de modelos associada a Sistemas Flexíveis de Manufatura, apresentada por Perkins e Kumar [16]. Em tais modelos, há um conjunto de M máquinas e P produtos; cada produto tem um caminho de processamento através das máquinas pré-estabelecido, permitindo-se o processamento de várias etapas de produção numa mesma máquina. A concorrência pelo serviço de uma máquina ocasiona a formação de *filas*, sendo necessária a elaboração de *regras* (chamadas *políticas*) que definam qual fila será atendida e quantos elementos desta serão processados. Tais políticas devem ser descentralizadas, ou seja, as decisões devem basear-se apenas em informações locais a cada máquina. Uma característica desejável das políticas utilizadas é que elas garantam a *estabilidade* do sistema, no sentido de que a aplicação contínua das regras não leve a um crescimento ilimitado das filas.

Ao longo do trabalho, estudamos a estabilidade de Sistemas de Manufatura sob dois aspectos: por um lado analisam-se políticas que garantam a estabilidade de qualquer sistema, e por outro lado procuram-se características inerentes a um sistema que assegurem sua estabilidade para uma ampla classe de políticas. Alguns dos principais resultados conhecidos na área são apresentados, e para um caso mais restrito (em que não há tempos de preparação nas máquinas) é proposta uma modelagem através de equações diferenciais no tempo; com esta abordagem, derivamos novas condições suficientes de estabilidade para sistemas com duas máquinas. Finalmente, uma Conjectura de Humes e Kumar sobre condições inerentes de estabilidade é discutida, apresentando-se alguns casos onde ela é verdadeira e exibindo-se um contra-exemplo para o caso geral.

Abstract

In this work we study the class of models associated with Flexible Manufacturing Systems, presented by Perkins e Kumar [16]. In such models, there are a set of M machines and a set of P products; each product has a previously established process path through the machines, with possible re-processing in the same service center. The competition for the services of a machine results in the formation of *queues*, requiring the setting of *policies* in order to define which queue will be served and how many elements from it will be processed. Such policies must be decentralized, that is, the decisions should be based only on local information. A desirable feature of the policies used is that they assure the *stability* of the system, in the sense that their continuous application does not lead to queues of unbounded size.

Throughout this work, we analyze the stability of Manufacturing Systems under two aspects: the study of policies that assure the stability of any system, and the search for inherent characteristics of a system which guarantee its stability when operated under a large class of policies. Some of the main known results are presented, and for a more restricted case (in which there are no set-up times) we propose representing the model using differential equations over time; with this approach, we derive new sufficient stability conditions for 2-machine systems. Finally, a conjecture of Humes and Kumar on a possible set of sufficient stability conditions is discussed, and we present cases where it is valid, as well as a counter-example for the general case.

Agradecimentos

Ao Carlos Humes, por um monte de coisas: por me introduzir ao problema, pela paciência ao longo das inúmeras discussões, pela capacidade de motivação (principalmente nas horas mais difíceis), pela dedicação exemplar para resolver os piores “pepinos” que cruzaram nosso caminho, pela confiança depositada desde a proposta do trabalho e, acima de tudo, pela amizade firmada ao longo desse tempo;

Ao Mané, por agüentar discussões malucas — mas não por isso menos proveitosas —, pela imprescindível colaboração no estudo de Sistemas Dinâmicos e também pelo apoio moral;

À Glaura e ao Odilon, meus pais, pelo apoio fundamental e pela compreensão durante os períodos mais críticos;

Ao Daniel e à Lucia, que mesmo de tão longe colaboraram de forma decisiva na fase final desta tese, “ouvindo” e se dispondo a todo tipo de ajuda, inclusive revisando o texto final;

À Yoshiko, pelo valioso auxílio no tópico referente à Teoria dos Grafos;

Ao Steve, pela tradicional revisão do “abstract”;

Ao pessoal dos Seminários (Sônia, Leo, Celma, etc.) pelas relevantes discussões nas apresentações dos trabalhos relacionados com esta tese;

À “turma do mail”, pela torcida nas horas finais;

E a todos que estavam no mesmo barco!!!

Conteúdo

Introdução	i
1 Conceitos básicos	1
1.1 Definições e notações	1
1.2 Estudo de sistemas com 1 máquina	8
1.2.1 Políticas “clearing”	11
1.2.2 Política “Fotográfica”	13
1.3 Sistemas acíclicos	16
2 Condições e mecanismos de estabilização	19
2.1 Um sistema instável sob CAF	20
2.2 Condições suficientes de estabilidade	27
2.2.1 Condições de estabilidade condicional	27
2.2.2 Enfraquecendo as condições de estabilidade condicional	29
2.2.3 Casos particulares	39

2.3	Mecanismos de estabilização	42
2.3.1	Políticas com “backoff”	43
2.3.2	O Mecanismo Supervisor Estabilizador (USSM)	46
2.3.3	O Mecanismo Regulador	49
2.3.4	Políticas de prioridades de buffers e componentes	58
3	Uma modelagem através de limiares sem tempos de preparação	62
3.1	O Modelo	63
3.2	Condições suficientes de estabilidade	69
3.3	Um caso particular: 2 máquinas	75
4	Analisando a Conjectura de Tempos Decrescentes	85
4.1	Casos válidos	86
4.1.1	O exemplo 1 de Kumar e Seidman	86
4.1.2	O exemplo 2 de Kumar e Seidman	88
4.2	Analisando o diagrama de estados para políticas “clearing”	90
4.3	Um contra-exemplo à Conjectura	100
	Conclusão	111
	Bibliografia	114

Introdução

O problema de programação de uma linha de produção (“scheduling”) há muito tem sido abordado de diversas maneiras, dada a sua importância na realidade de uma indústria ou de um projeto de redes de computadores. A abordagem dominante na literatura, contudo, concentra-se no aspecto *combinatório* do problema, isto é, encontrar a seqüência de processamento que prioriza algum critério visando sua otimização; pode-se objetivar, por exemplo, a minimização de atrasos na entrega, ou a maximização do atendimento aos clientes.

No presente trabalho estudaremos a classe de modelos associada a *Sistemas Flexíveis de Manufatura* apresentada por Perkins e Kumar [16]. Nessa modelagem, não se pretende fazer uma programação global do sistema; ao contrário, o princípio básico é que as decisões devem ser *descentralizadas*, de maneira que sejam independentes do restante do sistema, e portanto permitindo não só uma implementação de decisões em *tempo real*, mas também uma *flexibilidade* do sistema, conforme Graves [3], no sentido de que as máquinas sejam adaptáveis a diferentes tipos de situações.

Mais precisamente, o modelo analisado consiste de várias máquinas e várias diferentes linhas de produção que utilizam essas máquinas. Como cada máquina tem a capacidade de processar somente um tipo de produto de

cada vez, é natural pensar em uma concorrência de vários produtos pelos serviços de uma mesma máquina, o que pode levar à formação de *filas*, que no contexto de uma linha de produção podem ser entendidas como estoques. Note-se ainda que é possível termos diferentes etapas de processamento de um produto na mesma máquina, o que faz com que as filas sejam definidas não só por tipo de produto, mas também pela etapa de processamento de cada um.

A concorrência pelos serviços de uma máquina leva à necessidade de se definirem *regras* que estabeleçam *a priori* como resolver duas questões básicas: *qual* das filas deve ser escolhida para processamento numa dada máquina e *quantos* elementos da fila devem ser processados. No contexto clássico de “scheduling”, Panwalker e Iskander [15], por exemplo, apresentam um sumário com mais de 100 diferentes regras de decisão. No nosso caso, como dissemos acima, interessarão regras que possam se basear apenas em informações locais a cada máquina.

Embora a questão de *quantos* elementos de uma fila devem ser processados aparente ser desnecessária, uma vez que seria equivalente tomar uma nova decisão após o processamento de cada elemento (como na usual disciplina “primeiro a entrar – primeiro a sair”), é importante ressaltar que uma característica do modelo em questão é que considera-se a existência de *tempos de preparação* de uma máquina para o início do processamento de uma nova fila, durante os quais tal máquina nada processa. Dessa forma, é intuitivo pensar que o número de chaveamentos não pode ser muito grande. Por outro lado, uma rotatividade muito pequena entre as filas de uma máquina pode fazer com que algum produto com alta taxa de produção seja atendido de forma insuficiente, ocasionando assim a formação de filas cada vez maiores.

Nosso trabalho se concentrará, assim como as principais referências sobre o assunto ([16, 11, 6, 10]), no estudo de sistemas e de regras (que denominaremos *políticas*) que assegurem uma limitação no tamanho das filas ao longo do tempo — neste caso, diremos que existe *estabilidade* do sistema. No contexto, por exemplo, de redes de computadores, tal propriedade equivale a um dimensionamento do projeto do sistema, à medida em que garante uma limitação máxima no tamanho dos *buffers*. Adotamos a linha, como em [16, 11, 6], de estudar o modelo sob um ponto de vista *contínuo* o que,

embora não corresponda a uma realidade concreta, permite um tratamento matemático mais poderoso, possibilitando apesar disso uma boa aproximação dos casos discretos.

O trabalho está organizado da seguinte maneira: no Capítulo 1 introduzimos os conceitos básicos e a notação que será utilizada ao longo dos outros capítulos; estudamos os casos de sistemas simples, apresentando os principais resultados encontrados na literatura.

No Capítulo 2 passamos à análise de sistemas mais complexos, apresentando novamente alguns resultados conhecidos sob três aspectos diferentes: a demonstração da possível *instabilidade* de um sistema, através de contra-exemplos; a análise da existência de características inerentes a um sistema que garantam sua estabilidade para uma ampla classe de políticas; e, finalmente, o estudo de *mecanismos estabilizadores*, ou seja, políticas que garantam o bom funcionamento do sistema.

No Capítulo 3 procuramos formular o modelo alternativamente através de equações diferenciais no tempo, a fim de estudar a estabilidade de um sistema utilizando algumas ferramentas usuais em Sistemas Dinâmicos. É necessário, contudo, desconsiderar os tempos de preparação nas máquinas, para que o modelo matemático fique tratável. Após apresentarmos as hipóteses básicas, exibimos alguns resultados decorrentes dessa abordagem, ressaltando-se as condições obtidas para sistemas com apenas duas máquinas.

Finalmente, no Capítulo 4, analisamos uma Conjectura, proposta independentemente por C.Humes Jr. e P.R.Kumar, acerca de uma possível condição suficiente de estabilidade sobre um sistema para uma ampla classe de políticas, condição essa que englobaria várias condições fortes conhecidas, como veremos no Capítulo 2. Uma ferramenta alternativa para a análise de estabilidade de sistemas pequenos é discutida, e o trabalho fecha-se com a apresentação de um *contra-exemplo* à Conjectura, bem como a proposta de uma alternativa ao enunciado original.

Capítulo 1

Conceitos básicos

Neste capítulo, introduziremos a notação que será utilizada ao longo do trabalho, bem como algumas definições básicas. Em seguida, estudaremos a aplicação do conceito de *estabilidade* para sistemas com 1 máquina e para sistemas *acíclicos*, relacionando alguns resultados conhecidos na literatura.

1.1 Definições e notações

Os sistemas considerados possuem as seguintes características:

1. Há um conjunto de P produtos $\mathcal{P} = \{1, \dots, P\}$ e um conjunto de M máquinas $\mathcal{M} = \{m_1, \dots, m_M\}$;
2. Cada produto ¹ p segue um caminho de processamento pré-definido $\{\mu_{p,1}, \dots, \mu_{p,n_p}\}$ dentro do sistema, onde:
 - i) $\mu_{p,i} \in \mathcal{M}$ $i = 1, \dots, n_p$;
 - ii) $\mu_{p,1}$ é a máquina onde p é processado inicialmente;

¹Por simplicidade, usaremos o termo “produto” significando os elementos desse produto

- iii) Após ser processado por $\mu_{p,i}$, p é processado por $\mu_{p,i+1}$, $1 \leq i < n_p$;
- iv) Após ser processado por μ_{p,n_p} , o produto deixa o sistema.

Note-se que podemos ter diferentes etapas do processamento de um produto realizadas na mesma máquina, ou seja, $\mu_{p,i} = \mu_{p,j}$ com $i \neq j$.

3. Enquanto aguarda o processamento em $\mu_{p,i}$, um produto p é armazenado em um *buffer*, que denotaremos por (p,i) . Vamos denotar por $B_m := \{(p,i) : \mu_{p,i} = m\}$ o conjunto dos buffers associados à máquina m e por $x_{p,i}(t)$ o nível do buffer (p,i) , ou seja, a quantidade de elementos em (p,i) no instante t . O vetor de todos os $x_{p,i}(t)$ será indicado por $\mathbf{x}(t)$.
4. Cada máquina m processa apenas 1 buffer de cada vez. O processamento de 1 unidade de (p,i) demanda $\tau_{p,i}$ unidades de tempo, e diremos que a taxa *nominal* de processamento de (p,i) , que é a taxa de entrada instantânea máxima em $(p,i+1)$, é $\frac{1}{\tau_{p,i}}$. O chaveamento em m de um buffer b para um buffer b' (com $b, b' \in B_m$) demanda um tempo de *preparação* ("set-up") denotado por $\delta_{b,b'}$, durante o qual nenhum elemento é processado.
5. Estudaremos os sistemas sob a seguinte hipótese, com relação ao fluxo de entrada dos produtos: cada produto tem associado a si uma *taxa de chegada* e um valor d_p de tal modo que, se $e_p(t)$ indica a quantidade do produto p introduzida no sistema até o instante t , então

$$e_p(t) - e_p(s) \leq d_p(t - s) + \xi_p \quad \forall s, t > 0 : s \leq t \quad (1.1)$$

onde ξ_p é uma constante não-negativa. Ao longo deste trabalho, a menos que explicitado o contrário, suporemos a taxa de chegada igual a d_p .

A figura 1.1 ilustra um sistema típico.

A seguinte definição será largamente utilizada em todo o texto:

Definição 1.1 O grau de congestionamento de um buffer (p,i) é definido por $\rho_{p,i} = d_p \tau_{p,i}$. O grau de congestionamento de uma máquina m é definido por $\rho_m = \sum_{(p,i) \in B_m} \rho_{p,i}$.

$$\mu_{1,1} = \mu_{2,3} = m_1$$

$$\mu_{1,2} = \mu_{2,2} = m_2$$

$$\mu_{1,3} = \mu_{1,4} = \mu_{2,1} = m_3$$

$$B_{m_1} = \{(1,1), (2,3)\}$$

$$B_{m_2} = \{(1,2), (2,2)\}$$

$$B_{m_3} = \{(1,3), (1,4), (2,1)\}$$

Figura 1.1: Sistema típico

A utilização de um modelo como o descrito só faz sentido, qualquer que seja a aplicação, se pudermos garantir uma *limitação* no nível dos buffers, de forma a possibilitar um planejamento da capacidade do sistema. Desta forma, estaremos interessados, durante a maior parte deste trabalho, em analisar condições para a *estabilidade* dos sistemas estudados, definida nesse contexto:

Definição 1.2 Um sistema é dito estável se os níveis de todos os buffers permanecem limitados ao longo do tempo, ou seja, se dada uma condição inicial $\mathbf{x}(0)$, existem constantes $\Gamma_{p,i}(\mathbf{x}(0))$ tais que

$$x_{p,i}(t) \leq \Gamma_{p,i}(\mathbf{x}(0)) \quad \forall (p,i) \quad \forall t \geq 0 \quad (1.2)$$

Tal conceito, usualmente, só faz sentido em sistemas onde temos a seguinte *condição de capacidade* satisfeita:

$$\rho_m < 1 \quad \forall m \in \mathcal{M}$$

A razão da afirmação acima deve-se ao fato que $\rho_{p,i}$, com $(p,i) \in B_m$ pode ser visto como a fração de tempo que a máquina m deve dedicar ao buffer (p,i) .

É fácil ver, portanto, que uma máquina com grau de congestionamento acima de 1 terá uma capacidade insuficiente para atender à demanda (os produtos que chegam aos seus buffers), e conseqüentemente haverá formação de filas ao infinito, portanto o sistema não será estável. Dessa maneira, assumiremos daqui para frente que a condição de capacidade está satisfeita, sendo que em alguns casos particulares ela pode ser enfraquecida para $\rho_m \leq 1$. Quando isto ocorrer, destacaremos o fato.

Vamos denotar por $y_p(t)$ a quantidade do produto p processada pelo sistema em $[0, t]$. Para cada buffer (p, i) , vamos denotar por $y_{p,i}(t)$ a quantidade de (p, i) processada em $[0, t]$, e por $e_{p,i}(t)$ a quantidade introduzida em (p, i) nesse período. É claro que $e_p(\cdot) = e_{p,1}(\cdot)$ e $y_p(\cdot) = y_{p,n_p}(\cdot)$. O Lema abaixo caracteriza a estabilidade em função desses valores:

Lema 1.1 *O sistema é estável se e somente se a diferença entre as quantidades processada e introduzida de todo buffer é limitada em qualquer instante de tempo, ou seja,*

$$\text{estabilidade} \iff \exists \gamma(x(0)) : |y_{p,i}(t) - e_{p,i}(t)| < \gamma(x(0)) \quad \forall t \geq 0$$

-Demonstração:

É natural considerarmos a igualdade

$$y_{p,i}(t) - y_{p,i}(s) = e_{p,i}(t) - e_{p,i}(s) + x_{p,i}(s) - x_{p,i}(t) \quad s \leq t$$

ou seja, o que “sai” do buffer em um certo intervalo de tempo é o que “entra” durante esse período, adicionando-se o que havia antes e descontando-se o que há no final do intervalo. Re-escrevendo a expressão acima,

$$(y_{p,i}(t) - e_{p,i}(t)) - (y_{p,i}(s) - e_{p,i}(s)) = x_{p,i}(s) - x_{p,i}(t) \quad s \leq t$$

é fácil ver que existe limitação em $x_{p,i}(t)$ se e só se a diferença $y_{p,i}(t) - e_{p,i}(t)$ é limitada (em módulo) para todo t . ■

No nosso caso, a taxa de chegada satisfaz (1.1), logo vale a seguinte conseqüência de estabilidade:

Lema 1.2 *Se o sistema é estável, então todo buffer (p, i) “replica” a taxa de entrada, isto é,*

$$y_{p,i}(t) - y_{p,i}(s) \leq d_p(t - s) + \xi'_p \quad \forall s, t > 0 : s \leq t$$

onde ξ'_p é uma constante que depende apenas do estado inicial. Em particular, para $i = n_p$ tem-se

$$y_p(t) - y_p(s) \leq d_p(t - s) + \xi'_p \quad \forall s, t > 0 : s \leq t$$

–*Demonstração:*

Como no Lema anterior, consideremos a igualdade

$$y_{p,i}(t) - y_{p,i}(s) = e_{p,i}(t) - e_{p,i}(s) + x_{p,i}(s) - x_{p,i}(t) \quad s \leq t$$

Mas $e_{p,i}(\cdot) = y_{p,i-1}(\cdot)$, $i = 2, \dots, n_p$, logo aplicando essa igualdade sucessivamente obtemos

$$y_{p,i}(t) - y_{p,i}(s) = e_p(t) - e_p(s) + \sum_{j=1}^i x_{p,j}(s) - \sum_{j=1}^i x_{p,j}(t) \quad s \leq t$$

Pela hipótese (1.1) sobre a taxa de chegada, e da definição (1.2) de estabilidade concluímos que

$$y_{p,i}(t) - y_{p,i}(s) \leq d_p(t - s) + \xi_p + \sum_{j=1}^i \Gamma_{p,j}(\mathbf{x}(0))$$

Definindo $\xi'_p = \xi_p + \sum_{i=1}^{n_p} \Gamma_{p,i}(\mathbf{x}(0))$, segue a tese. ■

Algumas vezes nos referiremos à estabilidade em um sub-sistema do sistema analisado, o que deve ser entendido como uma restrição da definição (1.2) de estabilidade (buffers limitados) aos buffers do sub-sistema. Em particular, falaremos em *máquinas estáveis*.

Vimos acima que o grau de congestionamento $\rho_{p,i}$ indica a fração que uma máquina deve dedicar ao buffer (p, i) . A Proposição abaixo e seu corolário formalizam esse conceito, expressando uma consequência de estabilidade em termos do tempo médio dedicado a cada buffer:

Proposição 1.1 *Sejam $A_{p,i}(t)$ o tempo total de processamento do buffer (p,i) em $[0,t]$, e $N_m(t)$ o número de “set-up’s” ocorridos em m durante esse período. Suponhamos ainda todos os tempos de preparação iguais a uma constante δ . Se o sistema é estável, então*

$$\lim_{t \rightarrow \infty} \frac{A_{p,i}(t)}{t} = \lim_{t \rightarrow \infty} \frac{e_{p,i}(t)\tau_{p,i}}{t} = \lim_{t \rightarrow \infty} \frac{e_p(t)\tau_{p,i}}{t} \quad \forall (p,i)$$

$$\lim_{t \rightarrow \infty} \frac{N_m(t)}{t} \leq \frac{1 - \sum_{(p,i) \in B_m} \lim_{t \rightarrow \infty} \frac{A_{p,i}(t)}{t}}{\delta} \quad \forall m$$

–*Demonstração:*

Inicialmente, notemos que $A_{p,i}(t) = y_{p,i}(t)\tau_{p,i}$, ou seja, o tempo total de atendimento a (p,i) é a quantidade processada vezes o tempo de processamento de cada unidade. Se S é estável, então pelo Lema 1.1 temos

$$|y_{p,i}(t) - e_{p,i}(t)| \leq \gamma(x(0))$$

Logo,

$$\left| \frac{y_{p,i}(t)\tau_{p,i}}{t} - \frac{e_{p,i}(t)\tau_{p,i}}{t} \right| \leq \frac{\gamma(x(0))}{t}$$

e portanto

$$\lim_{t \rightarrow \infty} \frac{y_{p,i}(t)\tau_{p,i}}{t} = \lim_{t \rightarrow \infty} \frac{e_{p,i}(t)\tau_{p,i}}{t}$$

Lembrando, como no Lema 1.2, que $e_{p,i}(\cdot) = y_{p,i-1}(\cdot)$, $i = 2, \dots, n_p$, concluímos que

$$\lim_{t \rightarrow \infty} \frac{y_{p,i}(t)\tau_{p,i}}{t} = \lim_{t \rightarrow \infty} \frac{e_p(t)\tau_{p,i}}{t}$$

o que demonstra a primeira parte da Proposição.

Seja m uma máquina qualquer. É claro que, durante o intervalo $[0,t]$, ou m atendeu a algum buffer ou esteve em preparação, logo

$$\sum_{(p,i) \in B_m} A_{p,i}(t) + N_m(t)\delta \leq t$$

e portanto

$$\frac{N_m(t)}{t} \leq \frac{1 - \sum_{(p,i) \in B_m} \frac{A_{p,i}(t)}{t}}{\delta}$$

Tomando o limite para $t \rightarrow \infty$, segue a tese. ■

Corolário 1 Se o sistema é estável e $|e_p(t) - d_p t| \leq \xi_p$, para todo produto p , então

$$\lim_{t \rightarrow \infty} \frac{A_{p,i}(t)}{t} = \rho_{p,i} \quad \forall (p,i)$$

$$\lim_{t \rightarrow \infty} \frac{N_m(t)}{t} \leq \frac{1 - \rho_m}{\delta} \quad \forall m$$

Observe-se que, nesse caso, claramente $\rho_{p,i}$ corresponde ao atendimento médio necessário ao buffer (p, i) .

Finalmente, a natureza seqüencial das máquinas exige que se estabeleçam regras de atendimento ao buffers:

Definição 1.3 Uma política é um conjunto de regras que determina, para cada máquina, qual buffer será atendido e por quanto tempo ele será processado.

Neste trabalho interessarão basicamente as políticas ditas *descentralizadas*, ou seja, políticas em que as regras de atendimento dependem apenas de informações *locais* a cada máquina. Tal restrição baseia-se na premissa de que o custo de gerenciamento global de informação é alto; além disso, a descentralização permite uma *flexibilidade* do sistema, conforme Graves [3], na medida em que possibilita uma fácil adaptação a diferentes tipos de situações, tais como quebras de máquinas, alterações de velocidade, etc. Dessa maneira, torna-se desejável *minimizar* o conteúdo informacional necessário à operação do sistema.

Nosso estudo se concentrará, assim, em dois tópicos principais: a análise de *políticas estáveis para uma classe de sistemas*, ou seja, políticas que definem regras de operação sob as quais qualquer sistema de tal classe é estável, e o estudo de condições gerais que garantam a estabilidade de um sistema para uma larga classe de políticas.

1.2 Estudo de sistemas com 1 máquina

Uma questão natural que surge inicialmente diz respeito à estabilidade de sistemas com uma única máquina. Embora simples, o estudo desses casos já nos dá uma visão de alguns problemas envolvidos na análise de sistemas maiores.

Antes de analisarmos algumas políticas, é útil considerarmos o conceito introduzido por Humes [6]:

Definição 1.4 *Uma política é dita usável se todo sistema com uma única máquina, sujeito somente à “condição de capacidade”, é estável sob essa política.*

A importância de tal conceito deve-se ao fato de que em geral é bastante simples verificar se uma dada política é usável. Dessa maneira, para efeito de análise de estabilidade, podemos nos concentrar somente em políticas que satisfaçam essa condição.

Uma outra importante classe de políticas analisadas serão as políticas *conservantes*, que são definidas por Kleinrock [9] como políticas que *conservam* o trabalho potencial em cada máquina. No nosso contexto, tal classe de políticas pode ser definida da seguinte maneira:

Definição 1.5 *Uma política é dita conservante se não permite que nenhuma máquina fique ociosa (ou sub-alimentada) se algum de seus buffers não está vazio, ou seja, um buffer não vazio não é atendido pela respectiva máquina se e somente se esta estiver processando outro buffer não vazio ou estiver em preparação.*

Consideremos primeiramente sistemas nos quais não há tempo de preparação. O Lema abaixo mostra que, nessa condição, toda política conservante é usável.

Lema 1.3 *Qualquer política conservante estabiliza sistemas com uma única máquina sem tempos de preparação.*

–*Demonstração:*

Inicialmente, devemos lembrar que cada produto pode ter vários processamentos na máquina, como mostra a figura abaixo:

Tomemos a função

$$\beta(t) = \sum_p x_{p,1}(t)(\tau_{p,1} + \dots + \tau_{p,n_p}) + x_{p,2}(t)(\tau_{p,2} + \dots + \tau_{p,n_p}) + \dots + x_{p,n_p}(t)\tau_{p,n_p}$$

Sejam T_0 o instante de início de processamento de um buffer (p^*, i^*) qualquer, e T_1 o instante até o qual este é processado. Em T_1 temos

$$\begin{aligned} x_{p,1}(T_1) &\leq x_{p,1}(T_0) + d_p(T_1 - T_0) + \xi_p & p \neq p^* \text{ ou } i^* \neq 1 \\ x_{p,i}(T_1) &= x_{p,i}(T_0) & i \neq 1, \\ & & (p, i) \neq (p^*, i^*), \\ & & (p, i) \neq (p^*, i^* + 1) \\ x_{p^*, i^*}(T_1) &\leq x_{p^*, i^*}(T_0) + (d_p - \frac{1}{\tau_{p^*, 1}})(T_1 - T_0) + \xi_{p^*} & i^* = 1 \\ x_{p^*, i^*}(T_1) &= x_{p^*, i^*}(T_0) - \frac{1}{\tau_{p^*, i^*}}(T_1 - T_0) & i^* \neq 1 \\ x_{p^*, i^*+1}(T_1) &= x_{p^*, i^*+1}(T_0) + \frac{1}{\tau_{p^*, i^*}}(T_1 - T_0) & \text{se } i^* < n_{p^*} \end{aligned}$$

Logo,

$$\beta(T_1) \leq \beta(T_0) + (T_1 - T_0) \left[\sum_p d_p \sum_{i=1}^{n_p} \tau_{p,i} - \frac{1}{\tau_{p^*, i^*}} \left(\sum_{i=i^*}^{n_{p^*}} \tau_{p^*, i} - \sum_{i=i^*+1}^{n_{p^*}} \tau_{p^*, i} \right) \right] +$$

$$\begin{aligned}
& + \sum_p \sum_{i=1}^{n_p} \xi_p \tau_{p,i} \\
& = \beta(T_0) - (1 - \rho)(T_1 - T_0) + \sum_p \sum_{i=1}^{n_p} \xi_p \tau_{p,i}
\end{aligned}$$

Como, por hipótese, a política é conservante, segue que a máquina sempre estará atendendo a algum buffer, exceto quando todos os buffers estiverem vazios. Consideremos então uma sequência $\{T_1, \dots, T_r\}$ de atendimentos. Ora, a taxa de entrada satisfaz

$$e_p(T_r) - e_p(T_0) \leq d_p(T_r - T_0) + \xi_p$$

logo temos

$$\beta(T_r) \leq \beta(T_0) - (1 - \rho)(T_r - T_0) + \sum_p \sum_{i=1}^{n_p} \xi_p \tau_{p,i}$$

É fácil ver que, como $\rho < 1$, então $\beta(T_r) \rightarrow 0$ quando $r \rightarrow \infty$ e portanto o sistema é estável. Note-se que esse é um caso em que a condição de capacidade pode ser relaxada para $\rho \leq 1$, pois nesse caso concluímos que $\beta(\cdot)$ é menor ou igual a uma constante, o que também implica na estabilidade do sistema. ■

Observemos que o resultado acima não necessariamente vale quando existem tempos de preparação entre os buffers. De fato, o Corolário da Proposição 1.1 mostra que, se tivermos uma política (ainda que conservante, no sentido colocado acima) em que a frequência média de chaveamentos $\frac{N(t)}{t}$ seja maior que $\frac{1-\rho}{\delta}$ para todo t , então a fração de tempo “produtiva” da máquina será menor que ρ , e assim a máquina será incapaz de atender a todos os buffers, resultando na instabilidade.

Considerando então sistemas com tempo de preparação não nulo, analisaremos agora algumas políticas usáveis.

1.2.1 Políticas “clearing”

O raciocínio acima sobre a frequência média de chaveamentos de buffers mostra que, quanto mais congestionado o sistema, ou seja, quanto mais próximo ρ está de 1, menor deve ser o número de “set-up’s”. Isto nos leva a pensar em políticas tipo “clearing”, em que, uma vez escolhido um buffer, ele é processado até que não contenha mais elementos. Assim, seja T_n o instante da n -ésima escolha de buffer, e seja $b^* = (p^*, i^*)$ o buffer escolhido para ser processado. Se $b = (q, j)$ é o buffer que estava sendo processado, então $S_n = T_n + \delta_{b,b^*}$ é o instante a partir do qual b^* será atendido. Como a regra das políticas “clearing” nos diz que b^* será processado até que não contenha mais elementos, podemos então calcular o instante do fim do atendimento, e conseqüentemente nova escolha de buffer:

$$T_{n+1} = \begin{cases} S_n + x_{p^*,i^*}(S_n)\tau_{p^*,i^*} + d_{p^*}(T_{n+1} - S_n)\tau_{p^*,i^*} & \text{se } (p^*, i^*) \text{ tem entrada} \\ & \text{externa} \\ S_n + x_{p^*,i^*}(S_n)\tau_{p^*,i^*} & \text{caso contrário} \end{cases}$$

O termo $x_{p^*,i^*}(S_n)\tau_{p^*,i^*}$ corresponde ao tempo necessário para processar os elementos presentes no buffer no início do atendimento, enquanto que $d_{p^*}(T_{n+1} - S_n)\tau_{p^*,i^*}$ é o tempo que será gasto para processar os elementos que chegam durante o atendimento, caso o buffer em questão tenha alimentação externa.

Da equação acima temos que, no caso de (p^*, i^*) ter entrada externa,

$$T_{n+1} = S_n + \frac{x_{p^*,i^*}(S_n)\tau_{p^*,i^*}}{1 - \rho_{p^*,i^*}} = T_n + \delta_{b,b^*} + \frac{x_{p^*,i^*}(S_n)\tau_{p^*,i^*}}{1 - \rho_{p^*,i^*}}$$

onde $\rho_{p^*,i^*} = d_{p^*}\tau_{p^*,i^*}$, conforme definido anteriormente.

É preciso ainda definir uma regra para a escolha do novo buffer a ser processado nesse instante. Perkins e Kumar [16] consideraram as seguintes variações:

- *Limpar o maior buffer* (“Clear the Largest Buffer” - CLB): Nessa política, escolhe-se para processar o buffer com mais elementos, ou seja,

$$(p^*, i^*) = \underset{(p,i)}{\operatorname{argmax}} x_{p,i}(T_n)$$

- *Limpar uma fração* ("Clear-A-Fraction" - CAF): Generalização da política anterior. Escolhe-se para processar um buffer cujo nível seja no mínimo uma fração fixada do nível total dos buffers. Ou seja, (p^*, i^*) é tal que

$$x_{p^*, i^*}(T_n) \geq \varepsilon \sum_{(p, i)} x_{p, i}(T_n) \quad (1.3)$$

onde ε é a fração fixada. Se mais de um buffer satisfizer o critério, a escolha entre esses é arbitrária, embora isso possa ser um fator determinante de instabilidade em sistemas mais complexos, como veremos no capítulo 4.

- *Limpar o Maior Trabalho* ("Clear the Largest Work" - CLW): Escolhe-se para processar o buffer com maior trabalho contido, ou seja,

$$(p^*, i^*) = \underset{(p, i)}{\operatorname{argmax}} \{ \tau_{p, i} x_{p, i}(T_n) \}$$

É fácil ver que CLB é um caso particular de CAF, tomando-se $\varepsilon = \frac{1}{P-1}$. Perkins e Kumar demonstraram a estabilidade dessas políticas no caso do sistema sem laços mostrando que a função *trabalho*

$$w(t) = \sum_{(p, i)} \tau_{p, i} x_{p, i}(T_n)$$

é limitada superiormente por uma constante dependente apenas do estado inicial. Kumar e Seidman [11] generalizaram o resultado para máquinas com laços.

Note-se que uma política do tipo *Limpar o menor buffer*, em que escolhe-se o buffer com menos elementos, pode ser instável, bastando considerar um sistema com mais do que dois buffers e taxas de entrada apropriadas.

Uma outra política, também descrita em [16] para sistemas sem laços, define o produto a ser escolhido para processamento no instante T_n (como não há laços, cada produto só tem 1 buffer) como aquele que satisfaz

$$p^* = \underset{p}{\operatorname{argmax}} \left\{ \frac{x_p(T_n) + \delta d_p}{d_p \sqrt{\gamma_p^{-1} \rho_p^{-1} (1 - \rho_p)}} \right\}$$

onde os γ_p são constantes positivas. Além de mostrar a estabilidade de tal política, Perkins e Kumar relatam simulações em que o desempenho dessa política, medido em termos de *nível médio de buffer*, dado por

$$\liminf_{t \rightarrow \infty} \frac{1}{t} \int_0^t \left(\sum_p \gamma_p \tau_p x_p(s) \right) ds$$

fica próximo do valor

$$\frac{\delta[\sum_p \sqrt{\gamma_p \rho_p (1 - \rho_p)}]^2}{2(1 - \rho)}$$

que, como é provado também por Perkins e Kumar, é um *limite inferior* para o nível médio de buffer de *qualquer política estável*, comprovando o bom desempenho da política proposta.

Posteriormente, voltaremos ao estudo das políticas CAF, quando da análise de sistemas com mais de 1 máquina.

1.2.2 Política “Fotográfica”

Uma análise cuidadosa de sistemas com mais de 1 máquina operados sob uma política tipo “clearing” mostra — como será discutido em detalhes no Capítulo 2 — que o fato de um buffer ser processado ininterruptamente durante um período muito extenso pode levar o sistema à instabilidade. Seguindo essa linha de raciocínio, podemos considerar então uma política CAF modificada, na qual procura-se evitar “rodadas” de processamento muito longas. Denominaremos tal política como *política fotográfica*, definida formalmente da seguinte maneira:

- *Regra de escolha:* a mesma de CAF, isto é, no instante T_n da escolha toma-se um buffer (p^*, i^*) tal que

$$x_{p^*, i^*}(T_n) \geq \varepsilon \sum_{(p, i)} x_{p, i}(T_n)$$

- *Duração do atendimento:* processam-se apenas os elementos presentes no buffer no instante do início do atendimento. Ou seja, Se $b = (q, j)$

é o buffer que acabou de ser atendido e $b^* = (p^*, i^*)$ é o novo buffer escolhido, então (p^*, i^*) será processado até o instante

$$T_{n+1} = S_n + x_{p^*, i^*}(S_n) \tau_{p^*, i^*}$$

onde $S_n = T_n + \delta_{b, b^*}$ é o instante do início do processamento.

Note-se que a duração do atendimento independe da taxa de entrada, sendo definida apenas em função do nível do buffer no instante do início do processamento — daí o nome “fotográfica”. A política pretende, assim, assegurar uma duração mínima de atendimento, ao mesmo tempo em que procura garantir que o processamento não será longo demais.

Mostraremos agora que a política proposta é *usável*.

Teorema 1.1 *A política “fotográfica” é estável para sistemas com 1 única máquina.*

–*Demonstração:*

Seja

$$\beta(t) = \sum_p x_{p,1}(t)(\tau_{p,1} + \dots + \tau_{p,n_p}) + x_{p,2}(t)(\tau_{p,2} + \dots + \tau_{p,n_p}) + \dots + x_{p,n_p}(t)\tau_{p,n_p}$$

definida como na demonstração do Lema 1.3, e consideremos o instante S_n de início de serviço de (p^*, i^*) . É claro que a política proposta é conservante, logo no intervalo $[S_n, T_{n+1}]$ vale um desenvolvimento análogo ao feito naquele Lema e temos, portanto:

$$\beta(T_{n+1}) \leq \beta(S_n) - (1 - \rho)(T_{n+1} - S_n) + \xi$$

onde $\xi = \sum_p \sum_{i=1}^{n_p} \xi_p \tau_{p,i}$. Lembrando que

$$T_{n+1} - S_n = \beta_{p^*, i^*}(S_n)$$

temos então

$$\beta(T_{n+1}) \leq \beta(S_n) - (1 - \rho)\beta_{p^*, i^*}(S_n) + \xi \quad (1.4)$$

Por outro lado, em $[T_{n+1}, S_{n+1}]$ nenhum buffer é processado, logo

$$\beta(S_{n+1}) \leq \beta(T_{n+1}) + \rho(S_{n+1} - T_{n+1}) + \xi$$

Lembrando que $S_{n+1} - T_{n+1} = \delta_{b^*, b'}$, para algum buffer b' , temos então

$$\beta(S_{n+1}) \leq \beta(T_{n+1}) + \rho\delta_{b^*, b'} + \xi \quad (1.5)$$

De (1.4) e (1.5) concluímos que

$$\beta(S_{n+1}) \leq \beta(S_n) - (1 - \rho)\beta_{p^*, i^*}(S_n) + \rho\delta_{b^*, b'} + 2\xi \quad (1.6)$$

Sejam $\bar{\tau} = \max_{(p,i)} \tau_{p,i}$, $\underline{\tau} = \min_{(p,i)} \tau_{p,i}$, $\bar{\delta} = \max_{b_1, b_2 \in B_m} \delta_{b_1, b_2}$. Definamos uma constante Γ por

$$\Gamma = \frac{\bar{\tau}\bar{\delta}\rho}{\varepsilon\underline{\tau}(1 - \rho)} + \frac{2\xi\bar{\tau}}{\varepsilon(1 - \rho)}$$

Suponhamos que $\beta(S_n) > \Gamma + \rho\bar{\delta} + \xi > \Gamma$. Ora, de (1.5) vem que

$$\beta(S_n) \leq \beta(T_n) + \rho\bar{\delta} + \xi$$

e portanto

$$\beta(T_n) > \Gamma$$

Logo,

$$\sum_{(p,i)} x_{p,i}(T_n) \tau_{p,i} > \frac{\bar{\tau}\bar{\delta}\rho}{\varepsilon\underline{\tau}(1 - \rho)} + \frac{2\xi\bar{\tau}}{\varepsilon(1 - \rho)}$$

então

$$\sum_{(p,i)} x_{p,i}(T_n) > \frac{\bar{\delta}\rho}{\varepsilon\underline{\tau}(1 - \rho)} + \frac{2\xi}{\varepsilon(1 - \rho)}$$

e portanto

$$\varepsilon \sum_{(p,i)} x_{p,i}(T_n) > \frac{\bar{\delta}}{\underline{\tau}} \frac{\rho}{1 - \rho} + \frac{2\xi}{1 - \rho}$$

Mas T_n é o instante em que (p^*, i^*) foi escolhido, logo pela regra da política CAF devemos ter

$$x_{p^*, i^*}(T_n) \geq \varepsilon \sum_{(p, i)} x_{p, i}(T_n) > \frac{\bar{\delta}}{\underline{\tau}} \frac{\rho}{1 - \rho} + \frac{2\xi}{1 - \rho} > \frac{\delta_{b^*, b'}}{\tau_{p^*, i^*}} \frac{\rho}{1 - \rho} + \frac{2\xi}{1 - \rho}$$

Das desigualdades acima e do fato que $x_{p^*, i^*}(S_n) \geq x_{p^*, i^*}(T_n)$ vem que

$$\beta_{p^*, i^*}(S_n) > \frac{2\xi}{1 - \rho} + \delta_{b^*, b'} \frac{\rho}{1 - \rho}$$

e assim

$$2\xi + \rho\delta_{b^*, b'} - (1 - \rho)\beta_{p^*, i^*}(S_n) < 0$$

De (1.6) concluímos portanto que

$$\beta(S_{n+1}) < \beta(S_n)$$

Como $\beta(\cdot)$ é lipschitziana, segue que

$$\beta(t) < \max\{\beta(T_0), \Gamma + \rho\bar{\delta} + \xi\} \quad \forall t > T_0$$

donde $x_{p, i}(t)$ é limitado, $\forall(p, i)\forall t$, e portanto o sistema é estável. ■

1.3 Sistemas acíclicos

Consideremos agora sistemas com várias máquinas, onde os caminhos de processamento dos produtos induzem uma ordenação topológica das máquinas. Ou seja, como em [11], tomemos um grafo dirigido $G_S = (V, A)$ associado ao sistema, definido da seguinte maneira:

$$\begin{aligned} V &= \mathcal{M} = \text{máquinas do sistema} \\ A &= \{(m_i, m_j) : m_i, m_j \in V, m_i \neq m_j \text{ e } \exists p \in \mathcal{P} \mid \mu_{p, k} = m_i \text{ e} \\ &\quad \mu_{p, k+1} = m_j, \text{ para algum } k, 1 \leq k \leq n_p - 1\} \end{aligned}$$

Definição 1.6 Um sistema S é dito acíclico se o grafo G_S não contém ciclos orientados.

Figura 1.2: Sistema acíclico

Figura 1.3: Sistema não-acíclico

As figuras 1.2 e 1.3 ilustram sistemas acíclicos e não-acíclicos.

Humes [6] demonstrou a seguinte propriedade de sistemas acíclicos:

Proposição 1.2 *Qualquer política usável é estável para sistemas acíclicos.*

–*Demonstração:*

Como o grafo G_S é acíclico, podemos considerar uma *ordenação parcial* das máquinas induzida pelos caminhos de processamento dos produtos. Ou seja, podemos dizer que m_i *precede* m_j se e só se existe um caminho dirigido de m_i a m_j em G_S . Provaremos a afirmação proposta por indução nessa ordenação.

Seja m^1 uma máquina *minimal* com relação à ordenação. Como m^1 só tem entradas externas, segue que é estável sob uma política usável. Consideremos agora uma máquina m qualquer. Pela hipótese de indução, todas as máquinas que precedem m na ordenação são estáveis, logo pelo Lema 1.2 temos que as saídas dessas máquinas “replicam” suas entradas e portanto as entradas de m satisfazem a definição (1.1), comportando-se então como entradas externas. Assim, m é estável.

Como m foi escolhida arbitrariamente, segue que toda máquina tem buffers limitados, e portanto o sistema é estável. ■

Um corolário desta proposição é a estabilidade de políticas CAF para sistemas acíclicos, uma vez que, como comentamos anteriormente, Perkins e Kumar [16] mostraram que essa é uma classe de políticas usáveis. Embora em [16] também se prove a estabilidade de CAF para tais sistemas, a demonstração acima é muito mais geral, além de muito mais simples, utilizando fortemente o conceito de *política usável*.

Capítulo 2

Condições e mecanismos de estabilização

Os sistemas acíclicos estudados no capítulo 1 tinham a propriedade de serem estáveis sob qualquer política usável. Isto naturalmente leva ao questionamento acerca da estabilidade de sistemas com topologias mais complexas; em [16], deixa-se em aberto o problema da estabilidade da política CAF em sistemas não-acíclicos. Kumar e Seidman [11], contudo, mostraram que até mesmo sistemas pequenos, com 2 máquinas e 1 produto, podem ser instáveis, quando operados por políticas usáveis.

Neste capítulo, exibiremos os exemplos propostos por Kumar e Seidman, demonstrando sua instabilidade quando operados sob uma política CAF. Em seguida, relacionaremos algumas *condições* que, se satisfeitas por um sistema de qualquer natureza, são *suficientes* para garantir a estabilidade desse sistema sob qualquer política usável. Finalmente, algumas *políticas estabilizadoras* encontradas na literatura são descritas, com particular ênfase nos *mecanismos reguladores* propostos por Humes [6].

2.1 Um sistema instável sob CAF

Consideremos o sistema abaixo, composto de 1 produto e 2 máquinas.

Suponhamos que m_1 e m_2 satisfazem a condição de capacidade, ou seja, $\rho_1 + \rho_4 < 1$ e $\rho_2 + \rho_3 < 1$ (por simplicidade, omitiremos o índice do produto). Como dissemos no capítulo 1, suporemos a taxa de chegada constante igual a d .

Não há tempos de preparação. Seja T_0 um instante tal que

$$\begin{aligned} x_1(T_0) &= x_0 \\ x_2(T_0) &= x_3(T_0) = x_4(T_0) = 0 \end{aligned}$$

O funcionamento do sistema a partir de T_0 vai ocorrer da seguinte maneira:

1. A máquina m_1 começa a processar o buffer 1, alimentando o buffer 2, que passa a ser processado por m_2 .
2. Durante o processamento de (1), ou seja, do buffer 1, a taxa de entrada em (2) é $1/\tau_1$, que por hipótese é maior que sua taxa de processamento ($1/\tau_2$). Assim, o próximo evento será o fim do atendimento de (1), o que ocorre no instante

$$T_1 = T_0 + \frac{x_1(T_0)\tau_1}{1 - \rho_1} = T_0 + \frac{x_0\tau_1}{1 - \rho_1} \quad (2.1)$$

e temos

$$\begin{aligned}x_2(T_1) &= (T_1 - T_0) \left(\frac{1}{\tau_1} - \frac{1}{\tau_2} \right) \\x_3(T_1) &= (T_1 - T_0) \frac{1}{\tau_2}\end{aligned}$$

3. Após T_1 , a máquina m_1 passa a trabalhar *sub-alimentada*, ou seja, passa a processar (1) com uma taxa (d) menor que sua capacidade. Isto ocorre porque os buffers 1 e 4 estão vazios, logo (1) é processado com taxa igual à de chegada, permanecendo portanto com nível zero. Dessa maneira, (2) passa a ter uma taxa de entrada igual a d , logo o próximo evento ocorre quando esse buffer é zerado. Temos então

$$T_2 = T_1 + \frac{x_2(T_1)\tau_2}{1 - \rho_2} = T_1 + (T_1 - T_0) \left[\frac{1}{\tau_1} - \frac{1}{\tau_2} \right] \frac{\tau_2}{1 - \rho_2} \quad (2.2)$$

e também

$$\begin{aligned}x_3(T_2) &= x_3(T_1) + (T_2 - T_1) \frac{1}{\tau_2} = \frac{1}{\tau_2} (T_1 - T_0) + \frac{x_2(T_1)}{1 - \rho_2} = \\&= (T_1 - T_0) \left[\frac{1}{\tau_2} + \left(\frac{1}{\tau_1} - \frac{1}{\tau_2} \right) \frac{1}{1 - \rho_2} \right] = \\&= \frac{T_1 - T_0}{1 - \rho_2} \left[\frac{1}{\tau_1} - \rho_2 \frac{1}{\tau_2} \right] = \\&= (T_1 - T_0) \left[\frac{1}{\tau_1} \frac{1 - \rho_1}{1 - \rho_2} \right] = \\&= \frac{x_0}{1 - \rho_2}\end{aligned}$$

4. Tendo zerado (2), a máquina m_2 passa a processar (3). Conseqüentemente, m_1 passa a processar (4), já que sua taxa de entrada ($\frac{1}{\tau_3}$) é maior que a taxa de processamento ($\frac{1}{\tau_4}$). Como (2) não está sendo processado, a taxa de entrada de (3) é zero e assim o próximo evento será o fim do atendimento de (3), o que ocorre no instante

$$T_3 = T_2 + x_3(T_2)\tau_3 \quad (2.3)$$

e temos

$$\begin{aligned}x_1(T_3) &= d(T_3 - T_2) \\x_4(T_3) &= (T_3 - T_2) \left(\frac{1}{\tau_3} - \frac{1}{\tau_4} \right) = x_3(T_2) \left(1 - \frac{\tau_3}{\tau_4} \right)\end{aligned}$$

5. Após T_3 , a taxa de entrada em (4) passa a ser zero. A máquina m_2 fica ociosa enquanto (4) é processado, o que ocorre até o instante

$$T_4 = T_3 + x_4(T_3)\tau_4 \quad (2.4)$$

e temos

$$x_1(T_4) = x_1(T_3) + d(T_4 - T_3) = d(T_4 - T_2)$$

Ora, das equações (2.3) e (2.4) vem que

$$\begin{aligned}T_4 - T_2 &= x_3(T_2)\tau_3 + x_4(T_3)\tau_4 = \\&= x_3(T_2) \left[\tau_3 + \left(1 - \frac{\tau_3}{\tau_4} \right) \tau_4 \right] = x_3(T_2)\tau_4\end{aligned}$$

Juntando com a equação (2.2) obtemos

$$x_1(T_4) = d\tau_4 \frac{x_0}{1 - \rho_2} = \frac{\rho_4}{1 - \rho_2} x_0$$

Assim, o sistema atinge um *ciclo*, já que temos

$$x(T_4) = \frac{\rho_4}{1 - \rho_2} x(T_0)$$

onde $x(t) = (x_1(t), \dots, x_4(t))$. É fácil ver portanto que

$$x_1(T_{4k}) = \left(\frac{\rho_4}{1 - \rho_2} \right)^k x_0 \quad k = 1, 2, \dots$$

Donde, concluímos que o sistema é instável se $\rho_2 + \rho_4 > 1$, já que nesse caso temos $\lim_{k \rightarrow \infty} x_1(T_{4k}) = +\infty$.

A figura abaixo mostra o comportamento dos buffers em $[T_0, T_4]$ tomando-se $\tau_1 = \tau_3 = 0.2$, $\tau_2 = \tau_4 = 0.6$, $d = 1$.

O estudo do comportamento do sistema apresentado mostra que, entre os instantes T_1 e T_2 , a máquina m_1 fica sub-alimentada, enquanto que entre os instantes T_3 e T_4 , a máquina m_2 fica ociosa. Tais situações, como veremos no capítulo 3, são responsáveis pela instabilidade detectada, sendo resultantes tanto da política utilizada (CAF) quanto da natureza não-acíclica do sistema.

O que surpreende é que, ao contrário do que pode parecer à primeira vista, a instabilidade pode ocorrer mesmo que não haja re-processamentos de produtos, ou seja, mesmo que os caminhos dos produtos sejam acíclicos. De fato, consideremos o sistema abaixo, também apresentado por Kumar e Seidman:

Como no exemplo anterior, suponhamos que não há tempos de preparação, e seja T_0 um instante tal que

$$\begin{aligned} x_{1,1}(T_0) &= x_0 \\ x_{1,2}(T_0) &= x_{2,1}(T_0) = x_{2,2}(T_0) = 0 \end{aligned}$$

A máquina m_1 vai processar $(1,1)$, enquanto que m_2 vai processar $(1,2)$, já que este tem taxa de entrada maior que a taxa de processamento. Assim, o comportamento do sistema é similar ao anterior, isto é:

1. No instante

$$T_1 = T_0 + \frac{x_0 \tau_{1,1}}{1 - \rho_{1,1}}$$

o buffer $(1,1)$ é zerado; a partir daí m_1 fica sub-alimentada e temos

$$\begin{aligned} x_{1,2}(T_1) &= (T_1 - T_0) \left(\frac{1}{\tau_{1,1}} - \frac{1}{\tau_{1,2}} \right) = \frac{x_0}{1 - \rho_{1,1}} \left(1 - \frac{\tau_{1,1}}{\tau_{1,2}} \right) \\ x_{2,1}(T_1) &= d(T_1 - T_0) \end{aligned}$$

2. No instante

$$T_2 = T_1 + \frac{x_{1,2}(T_1) \tau_{1,2}}{1 - \rho_{1,2}} = T_1 + \frac{x_0(\tau_{1,2} - \tau_{1,1})}{(1 - \rho_{1,1})(1 - \rho_{1,2})}$$

o buffer $(1,2)$ é zerado, e temos

$$\begin{aligned} x_{1,2}(T_1) &= (T_1 - T_0) \left(\frac{1}{\tau_{1,1}} - \frac{1}{\tau_{1,2}} \right) = \frac{x_0}{1 - \rho_{1,1}} \left(1 - \frac{\tau_{1,1}}{\tau_{1,2}} \right) \\ x_{2,1}(T_2) &= d(T_2 - T_1) + d(T_1 - T_0) = d(T_2 - T_0) = \\ &= \frac{x_0 \rho_{1,1}}{1 - \rho_{1,1}} + \frac{x_0(\rho_{1,2} - \rho_{1,1})}{(1 - \rho_{1,1})(1 - \rho_{1,2})} = \frac{x_0 \rho_{1,2}}{1 - \rho_{1,2}} \end{aligned}$$

A máquina m_2 passa a processar $(2,1)$ e conseqüentemente m_1 passa a processar $(2,2)$.

3. Notemos que, em T_2 , temos uma situação simétrica à inicial (em T_0), com um buffer não vazio e os demais zerados. Como $\tau_{2,1} < \tau_{2,2}$, o

funcionamento do sistema a partir de T_2 se dará de forma análoga à acima, donde concluimos que, após os instantes T_3 (limpeza de $(2, 1)$) e T_4 (limpeza de $(2, 2)$) temos

$$\begin{aligned} x_{1,1}(T_4) &= \frac{\rho_{2,2}}{1 - \rho_{2,2}} x_{2,1}(T_2) \\ x_{1,2}(T_4) &= x_{2,1}(T_4) = x_{2,2}(T_4) = 0 \end{aligned}$$

Logo,

$$x(T_4) = \frac{\rho_{1,2}}{1 - \rho_{1,2}} \frac{\rho_{2,2}}{1 - \rho_{2,2}} x(T_0)$$

e portanto temos um ciclo.

Como no exemplo anterior, concluimos que, se

$$\frac{\rho_{1,2}}{1 - \rho_{1,2}} \frac{\rho_{2,2}}{1 - \rho_{2,2}} > 1$$

ou seja, se $\rho_{1,2} + \rho_{2,2} > 1$, então o sistema é instável.

É interessante notar também que a instabilidade do sistema depende das condições iniciais. De fato, suponhamos

$$\begin{aligned} \tau_{1,1} &= \tau_{2,1} \\ \tau_{1,2} &= \tau_{2,2} \\ d_1 &= d_2 \end{aligned}$$

satisfazendo a condição de instabilidade detectada acima, ou seja, $\rho_{1,2} = \rho_{2,2} > \frac{1}{2}$. Suponhamos ainda que em T_0 temos

$$\begin{aligned} x_{1,1}(T_0) &= x_{2,1}(T_0) = x_0 \\ x_{1,2}(T_0) &= x_{2,2}(T_0) = 0 \end{aligned}$$

É claro que m_1 vai processar $(1, 1)$ e m_2 vai processar $(2, 1)$, pois os outros buffers estão vazios. O sistema se comporta a partir daí da seguinte maneira:

1. No instante

$$T_1 = T_0 + \frac{x_0 \tau_{1,1}}{1 - \rho_{1,1}}$$

os buffers $(1, 1)$ e $(2, 1)$ são zerados; logo, os buffers $(1, 2)$ e $(2, 2)$ passam a ser processados em m_1 e m_2 , respectivamente, e temos

$$\begin{aligned} x_{1,2}(T_1) &= x_{2,2}(T_1) = (T_1 - T_0) \frac{1}{\tau_{1,1}} = \frac{x_0}{1 - \rho_{1,1}} \\ x_{1,1}(T_1) &= x_{2,1}(T_1) = 0 \end{aligned}$$

2. No instante

$$T_2 = T_1 + x_{1,2}(T_1) \tau_{1,2} = T_1 + \frac{x_0 \tau_{1,2}}{1 - \rho_{1,1}}$$

os buffers $(1, 2)$ e $(2, 2)$ são zerados simultaneamente; logo, os buffers $(1, 1)$ e $(2, 1)$ voltam ser processados em m_1 e m_2 e temos, portanto,

$$\begin{aligned} x_{1,1}(T_2) &= x_{2,1}(T_2) = d_1(T_2 - T_1) = \frac{x_0 \rho_{1,2}}{1 - \rho_{1,1}} \\ x_{1,2}(T_2) &= x_{2,2}(T_2) = 0 \end{aligned}$$

Assim,

$$\mathbf{x}(T_2) = \frac{\rho_{1,2}}{1 - \rho_{1,1}} \mathbf{x}(T_0)$$

e portanto temos um ciclo. Mas $\rho_{1,2} = \rho_{2,2}$ e, como o sistema por hipótese satisfaz a condição de capacidade ($\rho_{1,1} + \rho_{2,2} < 1$) temos

$$\frac{\rho_{1,2}}{1 - \rho_{1,1}} < 1$$

e portanto o sistema é estável. No Capítulo 4 analisaremos esse sistema em detalhes, mostrando condições necessárias e suficientes de estabilidade para políticas CAF.

2.2 Condições suficientes de estabilidade

Os exemplos da seção anterior mostram que a existência de um ciclo simples já pode permitir a instabilidade do sistema, bastando para isso que os graus de congestionamento sejam suficientemente altos. A pergunta que surge em consequência desse fato é: existem condições sobre os graus de congestionamento que garantam a estabilidade do sistema para qualquer política usável? Lembremos que, como vimos no capítulo 1, interessam-nos somente as políticas *descentralizadas*, ou seja, é desejável minimizar a informação global necessária à operação de tal política.

Vamos inicialmente denominar os sistemas que possuem esta característica.

Definição 2.1 *Dizemos que um sistema S é intrinsecamente estável se S é estável quando operado sob qualquer política usável.*

2.2.1 Condições de estabilidade condicional

A fim de analisarmos sistemas mais complexos, temos que considerar, como comentamos acima, os ciclos existentes entre as máquinas. Mais geralmente, devemos analisar conjuntos de máquinas que estão *interligadas*, no sentido de que existem ciclos (não necessariamente do mesmo produto, como vimos acima) contendo qualquer par de máquinas do conjunto.

Será útil, assim, considerar o grafo G_S definido no capítulo 1, isto é, $G_S = (V, A)$, onde

$$\begin{aligned} V &= \mathcal{M} = \text{máquinas do sistema} \\ A &= \{(m_i, m_j) : m_i \neq m_j \text{ e } \exists(p, k) \text{ t.q. } \mu_{p,k} = m_i \text{ e } \mu_{p,k+1} = m_j\} \end{aligned}$$

Para analisarmos os conjuntos de máquinas interligadas, tomemos uma *partição* de G_S , $\{G_1, \dots, G_r\}$, onde cada G_i é uma *componente fortemente conexa*, ou seja,

$m_i, m_j \in G_k \iff$ existem caminhos dirigidos de m_i a m_j e de m_j a m_i em G_S

Sem perda de generalidade, consideremos então que $(G_1, \dots, G_r)$ é uma *ordenação parcial* das componentes de G_S , onde G_i *precede* G_j (o que denotaremos por $G_i \prec G_j$) se e só se existe um caminho dirigido em G_S de G_i a G_j .

Seja $\overline{G}_S$ o grafo obtido a partir de G_S contraindo-se cada componente G_i a um único vértice. Claramente, $\overline{G}_S$ é acíclico. Ora, sabemos que qualquer política usável é estável em sistemas acíclicos, logo isso nos leva a concluir que, se pudermos garantir que cada componente G_i funciona “como uma máquina”, então teremos a estabilidade no sistema original.

Consideremos agora um buffer (p, i) qualquer, e seja G_j a componente fortemente conexa à qual pertence $\mu_{p,i}$. Vamos definir uma *taxa de entrada condicional* $d'_{p,i}$ da seguinte maneira:

$$d'_{p,i} = \begin{cases} d_p & \text{se } i = 1 \text{ ou } \mu_{p,i-1} \in G_k \text{ com } k \neq j \\ \frac{1}{\tau_{p,i-1}} & \text{se } \mu_{p,i-1} \in G_j \text{ e } \mu_{p,i-1} \neq \mu_{p,i} \\ d'_{p,i-1} & \text{se } \mu_{p,i-1} = \mu_{p,i} \end{cases}$$

Definição 2.2 Uma máquina m é chamada *condicionalmente estável* se

$$\rho'_m = \sum_{(p,i) \in B_m} d'_{p,i} \tau_{p,i} < 1$$

A idéia de tal conceito é que uma máquina m *condicionalmente estável* é capaz de “absorver” as taxas de entrada de seus buffers mesmo que estas estejam com seus valores máximos, isto é, $1/\tau_{p,i-1}$ para $(p, i) \in B_m$. Como vimos no Capítulo 1, no caso de re-processamentos consecutivos (laços), podemos considerar a última taxa de entrada *na máquina* antes do buffer em questão.

Com estas definições, Kumar e Seidman [11] demonstraram a seguinte

Proposição 2.1 *Se toda máquina m é condicionalmente estável, então o sistema é estável sob políticas CAF.*

Humes [6] generalizou a Proposição 2.1 enunciando o seguinte

Teorema 2.1 *Se toda máquina m é condicionalmente estável, então o sistema é intrinsecamente estável.*

–*Demonstração:*

Reproduziremos aqui as linhas gerais da demonstração. Tomemos a componente minimal (no sentido de “ $\prec$ ”) G_1 e seja m uma máquina dessa componente. Consideremos o subsistema S_m contendo somente a máquina m . Ora, m é condicionalmente estável, logo temos $\rho'_m = \rho_{S_m} < 1$ e portanto o sistema S_m , tomando como taxas de chegada $d'_{p,i}$, é estável sob qualquer política usável. Como a taxa de entrada de cada $(p, i) \in B_m$ em S é no máximo $d'_{p,i}$, segue que m é estável em S . Portanto, G_1 é estável, logo suas saídas replicam as taxas de chegada no sistema (de acordo com o Lema 1.2) e, aplicando um raciocínio indutivo na ordenação parcial $(G_1, \dots, G_r)$, concluímos que S é estável para qualquer política usável. ■

2.2.2 Enfraquecendo as condições de estabilidade condicional

O Teorema enunciado acima nos permite analisar a estabilidade de um sistema independentemente da política em questão. As condições impostas (estabilidade condicional em cada máquina) são demasiadamente fortes, e assim um esforço tem sido feito no sentido de enfraquecê-las e dessa maneira ampliar a classe dos sistemas intrinsecamente estáveis.

Humes [6] observou que, como máquinas condicionalmente estáveis replicam suas taxas de entrada máximas ($d'_{p,i}$), então *a análise de estabilidade condicional pode ser feita “eliminando-se” essas máquinas do sistema*. Tal processo de eliminação de uma máquina $\overline{m}$, denominado “curto-circuito” por Humes, consiste em considerar o sistema, que denotaremos por $S \setminus \overline{m}$, onde o conjunto de máquinas é $\mathcal{M} \setminus \{\overline{m}\}$ e, para cada buffer (p, i) , aplicar o seguinte algoritmo:

se $(p, i + 1)$ é processado em $\bar{m}$
 então seja (p, j) o primeiro buffer de p após (p, i) que é processado
 fora de $\bar{m}$
 se existe tal (p, j)
 então renumeram-se os buffers $(p, j), (p, j + 1), \dots, (p, n_p)$
 com $(p, i + 1), (p, i + 2), \dots, (p, n_p + i + 1 - j)$
 $n_p \leftarrow n_p + i + 1 - j$
 senão $n_p \leftarrow i$, ou seja, (p, i) é o último processamento de p .
 se $i = 1$ e (p, i) é processado em $\bar{m}$
 então se $n_p > 1$, ou seja, se p é processado por outra máquina
 então renumeram-se os buffers $(p, 2), (p, 3), \dots, (p, n_p)$
 com $(p, 1), (p, 2), \dots, (p, n_p - 1)$
 $n_p \leftarrow n_p - 1$
 senão $\mathcal{P} \leftarrow \mathcal{P} \setminus \{p\}$, ou seja, o produto p é eliminado do
 sistema

A figura 2.1 mostra um exemplo de um sistema S e o correspondente $S \setminus \bar{m}$.

O seguinte Lema caracteriza o grafo $G_{S \setminus \bar{m}}$ em função do sistema original. Mais geralmente, o enunciado vale para o grafo associado ao sistema acrescido dos laços.

Lema 2.1 *Sejam $u, v \in \mathcal{M} \setminus \{\bar{m}\}$, $u \neq v$, e G_S^l o grafo obtido de G_S acrescentando-se os laços correspondentes a buffers consecutivos na mesma máquina. Então, existe o arco (u, v) em $G_{S \setminus \bar{m}}^l$ se e somente se pelo menos uma das afirmações abaixo é válida:*

- i) *Existe o arco (u, v) em G_S^l ;*
- ii) *Existe um buffer (p, i) em S tal que $\mu_{p,i} = \bar{m}$, $\mu_{p,i-1} = u$ e $\mu_{p,j} = v$, onde j é o primeiro buffer de p após (p, i) que é processado fora de $\bar{m}$.*

-Demonstração:

($\Leftarrow$) Se $(u, v) \in G_S^l$, então obviamente $(u, v) \in G_{S \setminus \bar{m}}^l$. Se existe (p, i) em S tal que $\mu_{p,i} = \bar{m}$, $\mu_{p,i-1} = u$ e $\mu_{p,j} = v$, então o primeiro passo do algoritmo

Figura 2.1: Efeito da eliminação de uma máquina

de eliminação descrito anteriormente vai renumerar (p, j) com (p, i) , $(p, j + 1)$ com $(p, i + 1)$ e assim por diante. Logo, $G_{S \setminus \overline{m}}^l$ vai conter o arco (u, v) .

($\Rightarrow$) Pela definição do grafo $G_{S \setminus \overline{m}}^l$, sabemos que existe um buffer (p, i) em $S \setminus \overline{m}$ tal que $\mu_{p,i} = u$ e $\mu_{p,i+1} = v$ em $S \setminus \overline{m}$. Suponhamos que a retirada de $\overline{m}$ tenha causado a eliminação de alguns buffers entre os atuais (p, i) e $(p, i + 1)$. Seja (p, j) a numeração original de $(p, i + 1)$. Pelo algoritmo de eliminação, somente buffers de $\overline{m}$ são retirados, logo $(p, i + 1), \dots, (p, j - 1) \in B_{\overline{m}}$ em S e portanto j é o primeiro buffer de p após $(p, i + 1)$ processado fora de $\overline{m}$.

Por outro lado, se a retirada de $\overline{m}$ não eliminou nenhum buffer entre os atuais (p, i) e $(p, i + 1)$, então em S esses buffers eram numerados com (p, j) e $(p, j + 1)$, respectivamente, para algum j . Logo, existe o arco (u, v) em G_S^l .

■

É interessante notar que a eliminação de uma máquina pode descaracterizar algumas componentes fortemente conexas. Por exemplo, consideremos o sistema abaixo, contendo apenas 1 produto:

A eliminação de m_1 torna o sistema $S \setminus m_1$ acíclico, como mostra a figura abaixo:

A fim de analisarmos a influência da eliminação de uma máquina na composição fortemente conexa, consideremos a coleção $\mathcal{G}_S$ de componentes fortemente conexas de G_S e vamos definir um mapa ponto-a-conjunto

$$K_S : \mathcal{M} \longrightarrow \mathcal{G}_S$$

da seguinte maneira:

$$K_S(m) = G_i \in \mathcal{G}_S \text{ tal que } m \in G_i$$

ou seja, $K_S(m)$ é a componente fortemente conexa à qual pertence m , no sistema original. Lembremos que $\{G_i : G_i \in \mathcal{G}_S\}$ é uma *partição* de G_S .

O Lema abaixo mostra que a eliminação de uma máquina não “concatena” duas componentes fortemente conexas.

Lema 2.2 *Seja $\overline{m}$ a máquina eliminada. Para quaisquer máquinas $m_l, m_j \in \mathcal{M} \setminus \{\overline{m}\}$, vale que*

$$K_{S \setminus \overline{m}}(m_l) = K_{S \setminus \overline{m}}(m_j) \implies K_S(m_l) = K_S(m_j)$$

–Demonstração:

Se m_l e m_j estão na mesma componente em $G_{S \setminus \overline{m}}$, então existe um caminho dirigido de m_l a m_j em $G_{S \setminus \overline{m}}$ e vice-versa. Seja $C = (m_l, v_1, \dots, v_q, m_j)$ um tal caminho de m_l a m_j , e consideremos o arco (m_l, v_1) . Pelo Lema 2.1, sabemos que ou existe o arco (m_l, v_1) em G_S , ou então existe um buffer (p, i) em S tal que $\mu_{p, i-1} = m_l$, $\mu_{p, i} = \overline{m}$ e $\mu_{p, k} = v_1$, onde k é o primeiro buffer de p após (p, i) que é processado fora de $\overline{m}$. Logo, em G_S existe o caminho $(m_l, \overline{m}, v_1)$. Repetindo o mesmo raciocínio para os outros arcos de C , concluímos que existe um caminho dirigido C' em G_S que vai de m_l a m_j , eventualmente passando por $\overline{m}$. Analogamente, mostra-se que também existe um caminho dirigido em G_S de m_j a m_l . Portanto, $K_S(m_l) = K_S(m_j)$. ■

Uma importante observação deve ser feita neste ponto: devemos frisar que, embora máquinas condicionalmente estáveis repliquem suas taxas de entrada, o comportamento do sistema $S \setminus \overline{m}$ é diferente do comportamento de S .

De fato, isso ocorre quando consideramos *atrasos*, que por sua vez podem ser encarados como *tempos de transporte* entre os buffers. Nesses casos, um elemento que sai de um buffer após seu processamento não entra imediatamente no buffer seguinte, mas somente após determinado retardo.

Como observou Humes [6], uma política pode ser usável, no contexto em que estudamos até agora, mas ser instável, mesmo para sistemas com uma máquina isolada, ao considerarmos os atrasos. De fato, tomemos um sistema formado por uma única máquina e um único produto, com três processamentos consecutivos, e tempo de preparação δ entre todos os buffers. Suponhamos que existe um atraso σ entre os buffers (2) e (3), e consideremos uma política em que elementos “atrasados” têm menos prioridade. Note-se que, uma vez que a principal característica das políticas discutidas neste trabalho é a capacidade de permitir decisões *locais*, a definição do que é um elemento “atrasado” deve ter como parâmetro tão somente os buffers da máquina; ou seja, *o atraso só pode ser medido em relação a um processamento anterior na máquina*. Assim, um buffer com entrada externa nunca terá elementos atrasados.

É fácil ver o sistema assim definido é instável para a política proposta, uma vez que claramente o buffer (3) nunca será processado, já que seus elementos sempre terão menos prioridade.

É preciso, portanto, qualificar políticas que tolerem atrasos:

Definição 2.3 *Uma política é dita fortemente usável se todo sistema com uma única máquina é estável sob essa política, mesmo considerando atrasos entre os buffers consecutivos.*

A razão de inserirmos este comentário deve-se ao fato de que o processamento de um elemento, por si só, já é causa de atraso; ou seja, se um produto é processado em m_1 , depois em m_2 e novamente em m_1 , então ao chegar em m_1 na segunda vez o elemento tem um *atraso* em relação ao seu processamento anterior nessa máquina; o valor do atraso é exatamente o tempo gasto pelo elemento para ser processado na outra máquina.

De posse desses resultados, podemos agora mostrar que, sob determinadas condições, a eliminação de uma máquina condicionalmente estável não afeta o estudo da estabilidade do sistema, como afirma o Teorema abaixo, de Humes [6]:

Teorema 2.2 *Seja $\overline{m}$ uma máquina condicionalmente estável em S , e P uma política usável qualquer. Se $S \setminus \overline{m}$ é estável para P , e se vale pelo menos uma das afirmações abaixo:*

- i) Todos os laços de $S \setminus \overline{m}$ estão presentes em S ;*
- ii) P é fortemente usável.*

Então, S também é estável para P .

–Demonstração:

Seja (p, i) um buffer de S tal que $\mu_{p,i} \neq \overline{m}$, e (p, j) a sua numeração em $S \setminus \overline{m}$. Vamos denotar por $d''_{p,j}$ a taxa de entrada condicional de (p, j) em $S \setminus \overline{m}$. Para mostrar que S é estável, basta mostrar que $d''_{p,j}$ é a maior taxa de entrada possível de (p, i) em S , pois a estabilidade de $S \setminus \overline{m}$ garante a estabilidade de $\mu_{p,i}$ em S com as taxas de $S \setminus \overline{m}$ e portanto, nesse caso, teríamos a estabilidade de todas as máquinas de S .

Seja $m = \mu_{p,i}$ em S . Ora, se $m \notin K_S(\overline{m})$, então pelo Lema 2.2 temos que $m \notin K_{S \setminus \overline{m}}(u)$ para todo $u \in K_S(\overline{m})$ e portanto m não é afetada pela eliminação de $\overline{m}$, isto é, $d''_{p,j} = d'_{p,i}$. Logo, pela definição de $d'_{p,i}$, concluímos que $d''_{p,j}$ é a maior taxa de entrada possível de (p, i) em S .

Suponhamos então $m \in K_S(\overline{m})$, e consideremos a máquina $m' = \mu_{p,j-1}$ em $S \setminus \overline{m}$. Temos três possibilidades:

1. *Não existe tal m' :* Neste caso, em S temos ou $(p, i) = (p, 1)$ (ou seja, m é o primeiro processamento de p), ou então $\mu_{p,i-1} = \overline{m}$ e $\overline{m}$ é o primeiro processamento de p . Se vale a primeira situação, então claramente $d''_{p,j} = d'_{p,i} = d_p$. Na segunda situação, a taxa de entrada condicional de

p em $\overline{m}$ é d_p , logo a hipótese de estabilidade condicional de $\overline{m}$ garante que $\overline{m}$ replica essa taxa e portanto $d''_{p,j} = d_p$ é a taxa máxima de (p, i) em S .

2. $m' \in K_{S \setminus \overline{m}}(m)$: Neste caso, se $m' \neq m$ então segue que $d''_{p,j} = \frac{1}{\tau_{p,j-1}}$ e existe o arco (m', m) em $G_{S \setminus \overline{m}}$; Se $m' = m$, então $d''_{p,j} = d''_{p,j-1}$ e temos um novo laço, logo pela hipótese do Teorema a política em questão é *fortemente usável*. Em qualquer caso, podemos concluir pelo Lema 2.1 que em S vale que ou m' de fato é a máquina imediatamente anterior a (p, i) , ou então temos $\mu_{p,l+1} = \mu_{p,l+2} = \dots = \mu_{p,i-1} = \overline{m}$, onde (p, l) é a numeração de $(p, j-1)$ em S . Na última situação, a taxa de entrada condicional de $(p, l+1) \in B_{\overline{m}}$ em S é $d'_{p,l+1} = \frac{1}{\tau_{p,l}}$ donde se conclui que, se $m' \neq m$, então a estabilidade condicional de $\overline{m}$ garante que essa é a taxa máxima de entrada de (p, i) em S ; e se $m' = m$, é fácil ver que a sequência $(m, \overline{m}, m)$ comporta-se como um laço e portanto, como a política é fortemente usável, $d''_{p,j} = d''_{p,j-1}$ (taxa de entrada em m) é a taxa a ser considerada.
3. $m' \notin K_{S \setminus \overline{m}}(m)$: Se também temos $m' \notin K_S(m)$, então claramente $d''_{p,j} = d'_{p,i} = d_p$. Assim, suponhamos $m' \in K_S(m)$. Como $m' \notin K_{S \setminus \overline{m}}(m)$, concluímos que não existe caminho orientado de m a m' em $G_{S \setminus \overline{m}}$. Logo, o Lema 2.1 garante que em S não existe visita posterior do produto p à máquina m' , ou seja, $\mu_{p,k} \neq m', \forall k \geq j$. A mesma relação vale entre m e qualquer máquina $\mu \in K_S(m)$ tal que $K_{S \setminus \overline{m}}(\mu) \not\preceq K_{S \setminus \overline{m}}(m')$, ou seja, μ deve preceder m nos caminhos de todos os produtos que utilizam essas duas máquinas. Tomando então o subsistema S_1 de S , formado pelas máquinas de uma componente minimal (no sentido de $\prec$) de $G_{S \setminus \overline{m}}$, temos que S_1 só tem entradas externas ou através de $\overline{m}$; nesse último caso, uma entrada não externa em um buffer $(q, j) \in S_1$ só pode ocorrer através de uma visita anterior do respectivo produto q a $\overline{m}$ (que por sua vez replica sua taxa de entrada, pois é condicionalmente estável). Como as máquinas de S_1 são estáveis para essas taxas, pois são estáveis em $S \setminus \overline{m}$, concluímos que, em S , o subsistema S_1 replica as taxas externas. Aplicando indutivamente o raciocínio, concluímos que as taxas máximas de entrada nos subsistemas de S formados pelas máquinas das componentes de $G_{S \setminus \overline{m}}$, em particular no buffer (p, i) , são $d''_{p,j} = d_p$.

Em qualquer dos três casos, segue que m é estável, logo S é estável para P .

■

A condição de estabilidade demonstrada acima é, de fato, mais fraca que as condições de estabilidade condicional originais, como mostra o exemplo abaixo:

Note-se que m_1 não é condicionalmente estável, pois $\rho'_{m_1} = 0.6 + \frac{0.3}{0.1} > 1$. Mas m_2 é condicionalmente estável, já que $\rho'_{m_2} = \frac{1}{0.6}(0.4 + 0.1) < 1$. Eliminando m_2 , obtemos o sistema $S \setminus m_2$, que é formado por uma única máquina, logo é estável e portanto, como obtivemos um novo laço, concluímos S é estável para qualquer política *fortemente usável*.

A idéia da eliminação de uma máquina $\bar{m}$ condicionalmente estável nos leva a pensar de modo indutivo. Ou seja, podemos tomar o sistema $S \setminus \bar{m}$ e considerar uma máquina m' condicionalmente estável em $S \setminus \bar{m}$; se $(S \setminus \bar{m}) \setminus m'$, que denotaremos por $S \setminus (m \cup m')$, for estável, então $S \setminus m$ também o será e conseqüentemente S será estável. Se for possível então aplicar seguidamente o mesmo raciocínio, até obtermos um sistema garantidamente estável (por exemplo, com uma única máquina), teremos que cada sistema “reduzido” será estável e portanto o sistema original será estável.

Tais idéias podem ser sintetizadas no seguinte Teorema, que em [6] é apresentado como um algoritmo:

Teorema 2.3 *Seja S um sistema e P uma política usável qualquer. Se existe uma ordenação $(m^1, \dots, m^M)$ das máquinas de S e um índice j , $1 \leq j \leq M - 1$, tais que:*

- i) m^1 é condicionalmente estável em S ;
- ii) m^k é condicionalmente estável em $S \setminus (m^1 \cup \dots \cup m^{k-1})$, $k = 2, \dots, j$;
- iii) $S \setminus (m^1 \cup \dots \cup m^j)$ é estável quando operado sob P ;
- iv) Se P não é fortemente usável, então todos os laços de $S \setminus (m^1 \cup \dots \cup m^k)$ estão presentes em S , $k = 1, \dots, j$;

então, S é estável para a política P proposta.

–*Demonstração:*

Vamos formalizar o raciocínio exposto acima. Seja $(m^1, \dots, m^M)$ uma tal ordenação, e j um índice tal que $S \setminus (m^1 \cup \dots \cup m^j)$ é estável. Suponhamos que $S \setminus (m^1 \cup \dots \cup m^k)$ é estável, $1 \leq k \leq j$. Pela hipótese (ii) do Teorema, m^k é condicionalmente estável em $S \setminus (m^1 \cup \dots \cup m^{k-1})$. Ora, $S \setminus (m^1 \cup \dots \cup m^k) = (S \setminus (m^1 \cup \dots \cup m^{k-1})) \setminus m^k$, logo com a hipótese (iv) o Teorema 2.2 garante que $S \setminus (m^1 \cup \dots \cup m^{k-1})$ é estável, completando a demonstração. ■

Observemos ainda que o Teorema 2.1 é um caso particular do Teorema acima, tomando $j = M - 1$, já que, se S é condicionalmente estável, ou seja, se todas as máquinas são condicionalmente estáveis, então qualquer ordenação das máquinas satisfaz as condições do Teorema, como se pode concluir do seguinte

Fato 2.1 Se m e $\overline{m}$ são condicionalmente estáveis em S , então m é condicionalmente estável em $S \setminus \overline{m}$.

–*Demonstração:*

Basta observar que, na demonstração do Teorema 2.2, dado um buffer $(p, i) \in B_m$, se (p, j) é sua numeração em $S \setminus \overline{m}$, então $d''_{p,j} \leq d'_{p,i}$. ■

Figura 2.2: Um ciclo simples para $n = 6$

2.2.3 Casos particulares

A condição (iii) do enunciado do Teorema 2.3 nos diz que, vista de uma forma algorítmica, a verificação das condições de estabilidade condicional das máquinas em cada sistema reduzido pode ser encerrada quando um sistema intrinsecamente estável for atingido. Por exemplo, se em alguma etapa do processo de eliminação obtém-se um sistema acíclico, então o Teorema garante a estabilidade do sistema original.

Dessa forma, torna-se útil analisar casos particulares de sistemas intrinsecamente estáveis, com o objetivo de obter configurações minimais que possam servir como “base” para o processo de eliminação.

Consideremos um sistema S formado por M máquinas e 1 produto cujo caminho de processamento é um *ciclo simples*, ou seja, $n_p = n = M + 1$ e

$$\mu_i = \begin{cases} m_i & i = 1, \dots, n-1 \\ m_1 & i = n \end{cases}$$

A figura 2.2 mostra um tal sistema

O Teorema 2.3 nos diz que, se m_1 é condicionalmente estável, ou seja, se $d\tau_1 + \frac{1}{\tau_{n-1}}\tau_n < 1$, então S é estável, pois nesse caso a eliminação de m_1 leva ao sistema $S \setminus m_1$, que é acíclico e portanto estável.

No caso particular em que S não tem tempos de preparação, a proposição abaixo garante a estabilidade para uma classe de políticas:

Proposição 2.2 *Seja S um sistema ciclo-simples com M máquinas e 1 produto, sem tempos de preparação. Então, S é estável sob qualquer política conservante.*

-Demonstração:

Sem perda de generalidade, vamos supor que as máquinas $m_i \in \mathcal{M}$ são tais que $\mu_i = m_i$, $i = 1, \dots, n-1$, $\mu_n = m_1$. Conforme vimos acima, se m_1 é condicionalmente estável, então S é estável. Suponhamos que m_1 não é condicionalmente estável, ou seja, $\rho_1 + \frac{1}{\tau_{n-1}}\tau_n \geq 1$. É claro que podemos supor também $\tau_1 \leq \tau_2 \leq \dots \leq \tau_{n-1}$, pois se $\tau_{j-1} > \tau_j$ para algum j tal que $2 \leq j \leq n-1$, então $\frac{1}{\tau_{j-1}}\tau_j < 1$ e portanto μ_j é condicionalmente estável, logo poderia ser eliminada do sistema, de acordo com o Teorema 2.2; assim, podemos considerar o sistema $S \setminus \mu_j$.

Consideremos uma seqüência de instantes T_0, T_1, T_2 , onde:

- T_0 = início do atendimento do buffer 1
- T_1 = fim do atendimento do buffer 1 e início do atendimento do buffer n
- T_2 = fim do atendimento do buffer n e início do atendimento do buffer 1

Consideremos também a função

$$\beta(t) = x_1(t)\tau_1 + [x_1(t) + \dots + x_n(t)]\tau_n$$

É fácil ver que o processamento de uma unidade de um buffer j (com $2 \leq j \leq n-1$) não afeta a função $\beta(\cdot)$, pois essa unidade causa um decréscimo de 1 em (j) e portanto um decréscimo de τ_n na função, mas também causa um acréscimo de 1 em $(j+1)$ e conseqüentemente um acréscimo de τ_n em $\beta(\cdot)$. Já o processamento de uma unidade do buffer 1 causa um decréscimo de τ_1 na função, pois o argumento acima vale para a segunda parcela da expressão de

$\beta(t)$. Finalmente, observemos que o processamento de uma unidade do buffer n causa um decréscimo de τ_n na função, pois n não tem buffer posterior.

Por outro lado, supondo uma taxa de chegada no buffer 1 constante igual a d , temos que, em qualquer intervalo de tempo Δ , a função $\beta(\cdot)$ cresce $d\Delta(\tau_1 + \tau_n) = \Delta(\rho_1 + \rho_n)$.

Em $[T_0, T_1]$ a máquina 1 atende o buffer 1. Como m_1 não é condicionalmente estável, sabemos que o processamento só não ocorre com a taxa nominal $(\frac{1}{\tau_1})$ se (1) e (n) estão vazios, mas neste caso é fácil ver que o sistema todo teria que estar vazio, pois nenhuma máquina é condicionalmente estável, e portanto o sistema seria obviamente estável. Assim, se $x_1(t) + x_n(t) > 0$ durante todo o intervalo, temos que o número de unidades processadas nesse intervalo é $(T_1 - T_0)\frac{1}{\tau_1}$, logo

$$\begin{aligned}\beta(T_1) &= \beta(T_0) + (T_1 - T_0)(\rho_1 + \rho_n) - (T_1 - T_0) \\ &= \beta(T_0) - (T_1 - T_0)(1 - \rho_1 - \rho_n)\end{aligned}$$

Em $[T_1, T_2]$ a máquina 1 atende o buffer n . Mais uma vez, como m_1 não é condicionalmente estável, sabemos que o processamento só não ocorre com a taxa nominal $(\frac{1}{\tau_n})$ se (1) e (n) estão vazios. Supondo então $x_1(t) + x_n(t) > 0$, $\forall t \in [T_1, T_2]$, concluímos que o número de unidades processadas nesse intervalo é $(T_2 - T_1)\frac{1}{\tau_n}$ e portanto

$$\begin{aligned}\beta(T_2) &= \beta(T_1) + (T_2 - T_1)(\rho_1 + \rho_n) - (T_2 - T_1) \\ &= \beta(T_1) - (T_2 - T_1)(1 - \rho_1 - \rho_n)\end{aligned}$$

Juntando as equações acima, e denotando por ρ o valor $\rho_1 + \rho_n$, obtemos

$$\beta(T_2) = \beta(T_0) - (1 - \rho)[T_2 - T_0]$$

Se $\rho \leq 1$, então claramente $\beta(\cdot)$ é não-crescente, logo o sistema é estável. ■

A figura abaixo mostra um exemplo onde a aplicação do Teorema 2.3 e da Proposição anterior garante a estabilidade do sistema:

Suponhamos que não há tempos de preparação. Note-se que m_1 não é condicionalmente estável, pois $\rho'_{m_1} = \rho_1 + \frac{\tau_5}{\tau_4} = 0.6 + \frac{0.3}{0.2} > 1$. Também m_3 não é condicionalmente estável, pois $\rho'_{m_3} = \frac{\tau_3}{\tau_2} = \frac{0.7}{0.3} > 1$. Mas m_2 é condicionalmente estável, pois $\rho'_{m_2} = \frac{\tau_2}{\tau_1} + \frac{\tau_4}{\tau_3} = \frac{0.3}{0.6} + \frac{0.2}{0.7} < 1$. A eliminação de m_2 resulta no sistema

Embora nem m_1 nem m_3 sejam condicionalmente estáveis, o sistema resultante $S \setminus m_2$ é um ciclo simples, logo é estável para qualquer política conservante pela Proposição anterior, e portanto S é estável para essa classe de políticas.

2.3 Mecanismos de estabilização

Vimos na seção anterior condições que, se satisfeitas por um sistema, garantem a estabilidade deste para qualquer política usável. É interessante, contudo, estudar o problema sob um ponto-de-vista mais prático, isto é, responder à pergunta: Dado um sistema, como ele pode ser operado de maneira que sua estabilidade fique assegurada? Ou seja, queremos determinar *políticas estáveis* ou, mais geralmente, *mecanismos de estabilização* que adaptem uma política usável, tornando-a estável.

2.3.1 Políticas com “backoff”

Perkins e Kumar [16] propuseram a implementação da *política CAF com “backoff”*. Tal mecanismo consiste em considerar cada máquina como um sistema *isolado*, isto é, como se a taxa de entrada em cada buffer fosse a taxa de chegada d_p , e aplicar a política CAF nesse sistema “simulado”. Assim, indicando por $z_{p,i}(t)$ o nível do buffer (p, i) no instante t , temos, no sistema original:

- *Regra de escolha*: no instante T_n , escolhe-se um buffer (p^*, i^*) que satisfaça

$$z_{p^*, i^*}(T_n) \geq \varepsilon \sum_{(p, i) \in B_m} z_{p, i}(T_n)$$

- *Duração do atendimento*: processa-se o buffer escolhido (p^*, i^*) como se ele tivesse $z_{p^*, i^*}(T_n)$ elementos, ou seja,

$$T_{n+1} = T_n + \frac{z_{p^*, i^*}(T_n) \tau_{p^*, i^*}}{1 - \rho_{p^*, i^*}}$$

A atualização dos $z_{p,i}(\cdot)$ é feita supondo taxas d_p , ou seja, para (p, i) tal que $\mu_{p,i} = \mu_{p^*, i^*}$ temos:

$$z_{p,i}(T_{n+1}) = \begin{cases} z_{p,i}(T_n) + d_p(T_{n+1} - T_n) & (p, i) \neq (p^*, i^*) \\ z_{p,i}(T_n) + (d_p - \frac{1}{\tau_{p,i}})(T_{n+1} - T_n) & (p, i) = (p^*, i^*) \end{cases}$$

Os valores iniciais são os reais, ou seja, $z_{p,i}(T_0) = x_{p,i}(T_0)$ para todo (p, i) .

Note-se que o mecanismo proposto não é conservante, uma vez que permite que uma máquina fique parada (ou sub-alimentada) mesmo havendo outros buffers a processar, o que ocorrerá se $z_{p,i}(T_n)$ for suficientemente grande para que (p, i) seja escolhido mas tivermos $x_{p,i}(T_n) = 0$. Apesar desta aparente sub-utilização do sistema, em [16] demonstra-se a seguinte

Proposição 2.3 *A política CAF com “backoff” é estável para qualquer sistema.*

Podemos, porém, estender o conceito de Perkins e Kumar, de modo a generalizá-lo para qualquer política usável. Para isso, consideremos também variáveis $z_{p,i}(\cdot)$ representando o nível do buffer (p,i) se a máquina $\mu_{p,i}$ estivesse sendo operada isoladamente sob a política usável em questão. Novamente, a *escolha* do buffer a processar em um instante T_n , bem como o *intervalo de tempo* durante o qual o buffer escolhido será processado são definidos pela aplicação da política com base nos valores $z_{p,i}(T_n)$, aos invés dos valores reais $x_{p,i}(T_n)$. A atualização de $z_{p,i}(\cdot)$ é feita de modo idêntico ao visto no caso da política CAF.

Consideremos um buffer (p,i) e sejam $\{T_{2n}\}_{n \in \mathbb{N}}$ e $\{T_{2n+1}\}_{n \in \mathbb{N}}$ respectivamente as seqüências de instantes de início e fim de processamento de (p,i) . Assim, $\{[T_{2n}, T_{2n+1}]\}_{n \in \mathbb{N}}$ representam os intervalos em que (p,i) é processado, enquanto que $\{[T_{2n+1}, T_{2n+2}]\}_{n \in \mathbb{N}}$ representam aqueles em que (p,i) não é processado, o que inclui o tempo total 2δ (que suporemos positivo) correspondente ao chaveamento *de e para* o buffer (p,i) . Como dissemos acima, tais intervalos são determinados pela política em questão.

O Teorema abaixo estende o resultado de Perkins e Kumar, garantindo a estabilidade de políticas “backoff” em geral.

Teorema 2.4 *Seja P uma política usável. Então, a implementação de P com “backoff” é estável para qualquer sistema.*

–*Demonstração:*

Inicialmente, constatemos que o sistema “simulado” é estável, pois cada máquina é tratada como um sistema isolado e P é usável. Logo, existe $\Gamma > 0$ tal que

$$z_{p,i}(t) \leq \Gamma \quad \forall (p,i) \forall t \quad (2.5)$$

Vamos denotar por $e'_{p,i}(t)$ e $y'_{p,i}(t)$ respectivamente a entrada e saída acumuladas do buffer (p,i) no sistema simulado até o instante t . Pelo Lema 1.1, temos, para todo s, t com $s \leq t$

$$|y'_{p,i}(t) - y'_{p,i}(s) - (e'_{p,i}(t) - e'_{p,i}(s))| \leq |y'_{p,i}(t) - e'_{p,i}(t)| + |e'_{p,i}(s) - y'_{p,i}(s)| \leq 2\gamma$$

onde γ é uma constante não-negativa, dependente do estado inicial.

Como, por definição, $e'_{p,i}(t) = d_p t$, temos então

$$d_p(t-s) - 2\gamma \leq y'_{p,i}(t) - y'_{p,i}(s) \leq d_p(t-s) + 2\gamma \quad \forall(p,i) \forall s, t : s \leq t$$

Mas o processamento ocorre de acordo com o sistema simulado, logo a saída real é no máximo a saída simulada, ou seja,

$$y_{p,i}(t) - y_{p,i}(s) \leq y'_{p,i}(t) - y'_{p,i}(s) \leq d_p(t-s) + 2\gamma \quad \forall(p,i) \forall s, t : s \leq t$$

Por outro lado, é claro que $e_{p,i}(\cdot) = y_{p,i-1}(\cdot)$, $i > 1$, logo da desigualdade acima temos que

$$e_{p,i}(t) - e_{p,i}(s) \leq d_p(t-s) + 2\gamma \quad \forall(p,i) : i > 1 \quad \forall s, t : s \leq t$$

Lembrando da hipótese básica sobre o fluxo de entrada $e_{p,1}(\cdot)$ (ver (1.1)) temos então

$$e_{p,i}(t) - e_{p,i}(s) \leq d_p(t-s) + \gamma' \quad \forall(p,i) \quad \forall s, t : s \leq t \quad (2.6)$$

onde $\gamma' = \max\{2\gamma, \max_p \xi_p\}$.

Sejam (p, i) um buffer, t um instante qualquer. Vamos denotar por $N(t)$ o número de atendimentos de (p, i) em $[0, t]$ com taxa constante, por $\{S_j, T_j\}_{j=1}^{N(t)}$ respectivamente os instantes de início e fim dos atendimentos ($T_{N(t)} = t$ se (p, i) está sendo processado no instante t , $T_0 = 0$), e por ν_j a taxa de processamento em $[S_j, T_j]$.

Assim, temos, para $j = 1, \dots, N(t)$

$$\begin{aligned} z_{p,i}(T_j) &= z_{p,i}(T_{j-1}) + d_p(T_j - T_{j-1}) - \nu_j(T_j - S_j) \\ x_{p,i}(T_j) &= \max\{0, x_{p,i}(T_{j-1}) + e_{p,i}(T_j) - e_{p,i}(T_{j-1}) - \nu_j(T_j - S_j)\} \end{aligned}$$

A expressão de $x_{p,i}(T_j)$ deve-se ao fato que ν_j é determinado de acordo com o valor de $z_{p,i}(S_j)$, ou seja, ν_j pode ser igual à taxa de processamento nominal mas (p, i) pode estar vazio.

Seja r o maior índice j entre 1 e $N(t)$ tal que $x_{p,i}(T_j) = 0$, convencionando que $r = 0$ se $x_{p,i}(T_j) > 0$ para $j = 1, \dots, N(t)$. Temos então

$$\begin{aligned} z_{p,i}(t) &= z_{p,i}(T_r) + d_p(t - T_r) - \sum_{k=r+1}^{N(t)} \nu_k(T_k - S_k) \\ x_{p,i}(t) &= x_{p,i}(T_r) + e_{p,i}(t) - e_{p,i}(T_r) - \sum_{k=r+1}^{N(t)} \nu_k(T_k - S_k) \end{aligned}$$

Por (2.6), e da definição de r segue que

$$\begin{aligned} x_{p,i}(t) &\leq x_{p,i}(0) + d_p(t - T_r) + \gamma' - \sum_{k=r+1}^{N(t)} \nu_k(T_k - S_k) \\ &= x_{p,i}(0) + z_{p,i}(t) - z_{p,i}(T_r) + \gamma' \end{aligned}$$

De (2.5) temos então

$$x_{p,i}(t) \leq x_{p,i}(0) + \Gamma + \gamma'$$

Como t e (p, i) foram escolhidos arbitrariamente, segue que o sistema é estável para a política proposta. ■

2.3.2 O Mecanismo Supervisor Estabilizador (USSM)

Uma primeira análise dos exemplos apresentados na seção 2.1 levou Kumar e Seidman [11] a afirmarem que a instabilidade ali detectada teria como causa o longo período de sub-utilização de uma das máquinas, o que por sua vez é consequência de “rodadas” de processamento muito longas. Dessa maneira, é natural pensarmos em um mecanismo que interrompa, de alguma forma, o processamento de um buffer, quando um valor limite de tempo for atingido.

Baseados nisso, Kumar e Seidman propuseram o *Mecanismo Supervisor Estabilizador* (“Universally Stabilizing Supervisory Mechanism” - USSM), com duas características básicas: a *interrupção* de rodadas de processamento longas e a formação de uma *fila* de atendimento prioritário aos buffers, em cada máquina.

Consideremos, para cada máquina m , uma constante γ_m tal que

$$\gamma_m(1 - \rho_m) > \sum_{b \in B_m} \max_{b' \in B_m} \delta_{b,b'} \quad (2.7)$$

e, para cada buffer (p, i) , uma constante $\xi_{p,i}$ não-negativa. A cada máquina m associa-se uma fila Q_m de buffers, com disciplina Primeiro-a-Entrar-Primeiro-a-Sair.

Seja P uma política qualquer. O mecanismo consiste em adaptar P , aplicando as seguintes regras:

- *Escolha do buffer a processar*: se a fila Q_m está vazia, escolhe-se o buffer de acordo com o critério da política P ; caso contrário, *retira-se* o primeiro buffer da fila.
- *Duração do atendimento*: se o buffer escolhido *não foi* retirado da fila Q_m , então ele é processado durante o mínimo entre o tempo fixado por P e o valor $\gamma_m \rho_{p,i}$; caso contrário, ele é processado durante exatamente $\gamma_m \rho_{p,i}$ unidades de tempo, parando antes somente se o buffer for esvaziado.
- *Entrada na fila*: um buffer (p, i) entra no fim da fila Q_m no instante t se (p, i) não está sendo processado (nem está em "set-up") e $x_{p,i}(t) > \xi_{p,i}$.

Observemos que a desigualdade (2.7) é na verdade uma *condição necessária* à estabilidade da política proposta. De fato, suponhamos por simplicidade que todos os tempos de preparação são iguais a um certo valor δ . Seja m uma máquina qualquer e, para cada buffer $(p, i) \in B_m$, vamos denotar por $N_{p,i}(t)$ o número de atendimentos de m a (p, i) em $[0, t]$. Ora, como vimos acima, cada atendimento tem duração máxima $\gamma_m \rho_{p,i}$. Logo, denotando por $A_{p,i}(t)$ o tempo total de processamento de (p, i) em $[0, t]$, temos:

$$A_{p,i}(t) \leq N_{p,i}(t) \gamma_m \rho_{p,i}$$

e então, dividindo por t e tomando o limite em t , obtemos

$$\lim_{t \rightarrow \infty} \frac{A_{p,i}(t)}{t} \leq \lim_{t \rightarrow \infty} \frac{N_{p,i}(t) \gamma_m \rho_{p,i}}{t} \quad (2.8)$$

Mas pela Proposição 1.1 sabemos que, se o sistema é estável e supondo $|e_p(t) - d_p t| \leq M_p$ com M_p constante positiva, então temos

$$\lim_{t \rightarrow \infty} \frac{A_{p,i}(t)}{t} = \rho_{p,i}$$

De (2.8) concluímos que

$$\lim_{t \rightarrow \infty} \frac{N_{p,i}(t)}{t} \geq \frac{1}{\gamma_m}$$

e então

$$\sum_{(p,i) \in B_m} \lim_{t \rightarrow \infty} \frac{N_{p,i}(t)}{t} \geq \sum_{(p,i) \in B_m} \frac{1}{\gamma_m} \quad (2.9)$$

Por outro lado, notemos que o número de atendimentos aos buffers de m em $[0, t]$, que é o lado esquerdo da desigualdade acima, é pelo menos o número de chaveamentos de m nesse intervalo, valor que é denotado por $N_m(t)$, ou seja,

$$\sum_{(p,i) \in B_m} N_{p,i}(t) \geq N_m(t) \quad (2.10)$$

Mas (2.9) tem que valer sempre, em particular quando tivermos a igualdade em (2.10), ou seja, devemos ter

$$\lim_{t \rightarrow \infty} \frac{N_m(t)}{t} \geq \sum_{(p,i) \in B_m} \frac{1}{\gamma_m} \quad (2.11)$$

Novamente, se o sistema é estável e supondo $|e_p(t) - d_p t| \leq M_p$ temos, pela Proposição 1.1:

$$\lim_{t \rightarrow \infty} \frac{N_m(t)}{t} \leq \frac{1 - \rho_m}{\delta}$$

Juntando a desigualdade acima com (2.11) temos portanto

$$\sum_{(p,i) \in B_m} \frac{1}{\gamma_m} \leq \lim_{t \rightarrow \infty} \frac{N_m(t)}{t} \leq \frac{1 - \rho_m}{\delta}$$

donde concluímos que

$$\gamma_m(1 - \rho_m) \geq \sum_{(p,i) \in B_m} \delta$$

que é a condição (2.7).

Por sua vez, $\xi_{p,i}$ condiciona a aplicação do mecanismo a níveis altos de buffers, já que obviamente não há necessidade de modificar a política original se os buffers permanecem limitados.

Note-se assim que o ajuste dos parâmetros γ_m e $\xi_{p,i}$ possibilita o controle do grau de intervenção do mecanismo na política original, sendo que valores altos

desses parâmetros corresponderão a uma diminuição do número de vezes em que o mecanismo é acionado.

No mesmo trabalho [11], Kumar e Seidman mostraram que o mecanismo proposto é, de fato, estabilizador:

Teorema 2.5 *Qualquer política P , associada ao Mecanismo Supervisor Estabilizador, é estável para todo sistema.*

A demonstração, que não reproduziremos aqui (uma vez que será obtida como corolário da estabilidade de Mecanismos Reguladores, como veremos adiante), consiste em mostrar que: *a)* o tempo decorrido entre a entrada de um buffer na fila Q_m e seu processamento subsequente é limitado; e *b)* a função $\beta_{p,i}(t) = \sum_{j=1}^n \tau_{p,i} x_{p,j}(t)$ é limitada. A partir daí, pode-se concluir (com algum trabalho) que $x_{p,i}(\cdot)$ é limitado, para todo buffer (p, i) .

2.3.3 O Mecanismo Regulador

O estudo das condições suficientes de estabilidade, visto na seção 2.2, em particular o Teorema 2.3, permite que derivemos um mecanismo de estabilização. Humes [6] observou inicialmente que podemos introduzir *máquinas reguladoras* $r_{p,i}$ antes de cada buffer (p, i) , tais que: *a)* $r_{p,i}$ processa somente um buffer, ou seja, $|B_{r_{p,i}}| = 1$; e *b)* a taxa de processamento de $r_{p,i}$, *que passa a ser a taxa de entrada condicional em (p, i)* , é igual a d_p .

É fácil ver que, nesse caso, toda máquina não reguladora passa a ser condicionalmente estável, uma vez que a taxa de entrada máxima em cada buffer é d_p . Eliminando (isto é, efetuando o “curto-circuito”) essas máquinas, obtemos um sistema composto apenas pelas máquinas $r_{p,i}$, que é obviamente acíclico, já que cada uma dessas máquinas processa somente 1 buffer.

O raciocínio acima mostra portanto que, *acrescentando* máquinas apropriadas ao sistema, podemos garantir sua estabilidade, em virtude do Teorema 2.3. Note-se que, uma vez que o objetivo é garantir a estabilidade condi-

cional das máquinas originais, as máquinas reguladoras podem ter taxas de processamento maiores que d_p .

É essencial observar, contudo, que a implementação das máquinas reguladoras pode ser feita de maneira *lógica*, ou seja, pode-se *emular* as máquinas reguladoras através de ponteiros em cada buffer (p, i) que indiquem o nível de (p, i) se este estivesse sendo alimentado pela máquina reguladora $r_{p,i}$, o que faz com que obtenhamos dessa maneira um mecanismo estabilizador semelhante ao mecanismo "backoff" visto na seção 2.3.1. No final desta seção comentaremos as relações entre reguladores, "backoff" e o mecanismo USSM.

Vamos agora formalizar os conceitos, como em [6], começando pela definição de regulador:

Definição 2.4 Um v -regulador é um emulador de uma máquina reguladora, no sentido visto acima, cuja taxa de processamento é v .

A estabilidade do mecanismo regulador é enunciada no teorema abaixo.

Teorema 2.6 Seja S um sistema e P uma política usável. Se a cada buffer (p, i) tal que $i = 1$ ou $\mu_{p,i} \neq \mu_{p,i-1}$ associa-se um $v_{p,i}$ -regulador de tal modo que:

- i) $v_{p,i} \geq d_p \quad \forall (p, i)$
- ii) $\sum_{(p,i) \in B_m} v_{p,i} \tau_{p,i} < 1 \quad \forall m \in \mathcal{M}$

Então, S é estável quando operado sob P com esse mecanismo regulador.

-Demonstração:

Vamos denotar por S_R o sistema obtido de S acrescentando máquinas $v_{p,i}$ -reguladoras que satisfazem as condições (i) e (ii) do Teorema. É claro que sempre podemos tomar tais $v_{p,i}$, pois como S satisfaz a condição de capacidade, temos

$$\sum_{(p,i) \in B_m} d_p \tau_{p,i} < 1 \quad \forall m \in \mathcal{M}$$

logo, $v_{p,i} = d_p$ satisfaz (i) e (ii). Além disso, por (i) concluímos que as máquinas reguladoras satisfazem a condição de capacidade, pois o tempo de processamento de uma máquina $v_{p,i}$ -reguladora é $v_{p,i}^{-1}$ e, como tais máquinas só processam 1 buffer, a condição de capacidade pode ser relaxada para $\rho_m \leq 1$.

Por (ii), temos que toda máquina $m \in \mathcal{M}$ é condicionalmente estável em S_R e portanto pode ser eliminada (no sentido do Teorema 2.3) do sistema. Como o “curto-circuito” de uma máquina não aumenta o número de buffers nas máquinas remanescentes, que só têm um buffer por serem reguladores, temos que $S_R \setminus (m^1 \cup \dots \cup m^M)$ é acíclico, logo é estável e, pelo Teorema 2.3, S_R é estável. ■

Minimizando o número de reguladores

O Teorema acima possibilita a garantia de estabilização de um sistema através da implementação de reguladores. É fácil ver, porém, que não é necessário “regular” todos os buffers; tal é o caso, por exemplo, daquelas que são alimentadas externamente, ou por outra componente fortemente conexa. Do mesmo modo, podemos observar que, se a eliminação de uma máquina “regulada” resultar num sistema acíclico, então os reguladores que garantem a estabilidade condicional dessa máquina são suficientes para estabilizar o sistema todo.

Humes [6] demonstrou o seguinte resultado, que mostra uma relação entre o número de reguladores e a topologia do sistema:

Teorema 2.7 *Seja S um sistema, R um conjunto de d_p -reguladores (ou seja, reguladores com taxa d_p) e S_R o sistema obtido a partir de R acrescentando-se os reguladores de R . Se o grafo $G_{S_R}^R$, que se obtém removendo do grafo G_{S_R} os nós $r \in R$ e os arcos incidentes a esses nós, é um grafo acíclico, então S_R é intrinsecamente estável, ou seja, R é um conjunto suficiente de reguladores para S .*

–*Demonstração:*

Sejam $C_1, \dots, C_k$ as componentes conexas de $G_{S_R}^R$, já que este pode ser não-

conexo. Como $G_{S_R}^R$ é acíclico, então obviamente cada C_i também o é. Como já vimos anteriormente, os arcos em C_i induzem uma ordenação parcial dos nós (máquinas) de C_i . Tomemos então uma máquina m^1 , que seja *minimal* em C_i . É claro que, em S_R , m^1 só tem entradas externas ou através de reguladores, pois do contrário haveria um arco (m, m^1) em $G_{S_R}^R$, para algum $m \in \mathcal{M}$. Logo, m^1 é condicionalmente estável, e portanto replica a taxa d_p .

Aplicando indutivamente o raciocínio em cada componente conexa acíclica C_i , concluímos que as entradas de toda máquina $m \in C_i$ são: ou externas, ou através de reguladores, ou através de máquinas que precedem m na ordenação e portanto replicam d_p . Em qualquer caso, a taxa máxima de entrada de todo buffer de qualquer máquina m é d_p , logo m é estável, e portanto S_R é estável para qualquer política usável. ■

A figura 2.3 mostra um exemplo de aplicação do Teorema acima.

Uma consequência imediata do Teorema é que podemos pensar em *minimizar* o número de reguladores do sistema. Ou seja, queremos descobrir o menor número de buffers que, se regulados, garantam a estabilidade do sistema. É claro que tal valor seria também um limite máximo para o número de reguladores necessários, possibilitando desse modo uma redução na estimativa do custo de implementação.

Formalmente, o problema pode ser posto da seguinte maneira: consideremos o grafo $G_S = (V, A)$ e uma função $\alpha : A \rightarrow \mathbb{N}$ definida como

$$\alpha((u, v)) = |\{(p, i) : \mu_{p,i} = u, \mu_{p,i+1} = v\}|$$

ou seja, $\alpha(a)$ para $a \in A$ é o número de buffers representados em G_S pelo arco a . Queremos encontrar um subconjunto $X \subset A$ tal que

- i) todo ciclo orientado de G_S contenha pelo menos 1 arco de X ;
- ii) $\alpha(X) = \sum_{a \in X} \alpha(a)$ seja mínimo.

É fácil ver, pelo Teorema 2.7 que, se $\bar{X}$ é uma solução do problema acima, então $\alpha(\bar{X})$ é um limitante superior ao número de reguladores necessários à

Figura 2.3: Utilização dos reguladores. Note que $G_{S_R}^R$ é acíclico.

estabilização do sistema original. Isto porque, se colocarmos d_p -reguladores nos buffers associados a cada arco de $\overline{X}$, então claramente a retirada desses reguladores em G_{S_R} tornará esse grafo acíclico.

Infelizmente, porém, o problema é $\mathcal{NP}$ -Completo. De fato, o caso particular em que $\alpha(a) = 1$ para todo arco a é o problema conhecido na literatura como “Feedback Arc Set Problem” (Problema do Conjunto de Arcos de Realimentação), comprovadamente $\mathcal{NP}$ -Completo, conforme Garey e Johnson [2].

Para algumas classes de grafos, existem algoritmos polinomiais: Lucchesi [14] o demonstrou para grafos *planares* e Ramachandran [17] para *grafos de fluxo* (grafos com um nó r tal que existe caminho orientado de r a qualquer outro nó do grafo) que sejam *reduzíveis*. Existem diversas caracterizações para tal classe de grafos, por exemplo as de Hecht e Ullman [4, 5]. Uma delas [5] afirma que um grafo é irreduzível se e somente se seu DAG (subgrafo acíclico maximal) é único. Em [4] mostra-se que um grafo é reduzível se e somente se não contém o grafo abaixo:

onde as linhas onduladas representam caminhos disjuntos nos nós (exceto nas pontas) e os nós r, v_1, v_2, v_3 são dois a dois distintos entre si, sendo a única exceção permitida $r = v_1$.

A figura 2.4 ilustra a minimização de reguladores necessários.

O sistema original

Uma solução

A solução ótima

Figura 2.4: Minimizando o limite superior ao número de reguladores necessários.

Minimizando a informação necessária aos reguladores

Vimos acima que o Teorema 2.7 permite que se obtenha um limite máximo ao número de reguladores necessários à estabilização. É importante observar, contudo, que o Teorema é válido para d_p -reguladores, ou seja, reguladores com taxa de processamento igual à taxa de entrada no sistema. Tal hipótese, utilizada para garantir a estabilidade condicional de toda máquina regulada (inclusive após os “curto-circuitos”) pressupõe porém o gerenciamento ainda de uma informação global, a taxa d_p , o que não é desejável em nossa abordagem, conforme comentamos no capítulo 1.

Uma outra forma de encarar o problema é pensando na minimização da *informação* global necessária aos reguladores. Ou seja, sob esse ponto de vista, queremos estabelecer um mecanismo regulador cujas taxas, para uma máquina m , dependam apenas dos valores de $\tau_{p,i}$, para $(p,i) \in B_m$.

Um exemplo de tal mecanismo seria a implementação de $v_{p,i}$ -reguladores satisfazendo as condições do Teorema 2.6, ou seja, $v_{p,i} \geq d_p$ e $\sum_{(p,i) \in B_m} v_{p,i} \tau_{p,i} < 1$. Como vimos na demonstração daquele Teorema, a existência de tais $v_{p,i}$ é garantida pela condição de capacidade $\rho_m < 1$, e qualquer alteração nas taxas d_p tem que manter essa restrição.

Um mecanismo como esse, entretanto, tem a aparente desvantagem de, num certo sentido, piorar o desempenho do sistema, já que nesse caso os reguladores funcionam como máquinas lentas e portanto “atrasam” a produção até o limite da estabilidade.

A questão da utilização de reguladores mais rápidos continua em aberto; de fato, ela está intimamente relacionada à obtenção de condições suficientes de estabilidade mais fracas do que aquelas encontradas na literatura, e que foram tema da seção 2.2 deste capítulo.

Humes e Kumar propuseram independentemente a seguinte

Conjectura 1 *Um sistema é estável sob qualquer política usável se os tempos de processamento de cada produto são sucessivamente não-crescentes,*

isto é, se

$$\tau_{p,i} \geq \tau_{p,i+1} \quad \forall p, \quad i = 1, \dots, n_p - 1$$

A confirmação da validade da conjectura possibilitaria a obtenção de mecanismos reguladores mais eficientes, já que nesse caso bastaria que a taxa $v_{p,i}$ do regulador satisfizesse a condição $v_{p,i} \geq \tau_{p,i}$. De fato, resultados de nossas simulações apontavam nessa direção; posteriormente, porém, encontramos uma classe de *contra-exemplos* à conjectura, o que será descrito com detalhes no capítulo 4.

Relacionando reguladores com “backoff” e USSM

De certa forma, podemos encarar o Mecanismo Regulador descrito anteriormente como uma *generalização* dos dois outros mecanismos de estabilização vistos neste capítulo, o que permite que se apliquem alguns resultados obtidos para reguladores àquelas políticas, como por exemplo o limite máximo do número de buffers onde deve intervir o mecanismo de estabilização.

É fácil verificar a relação entre o Mecanismo Regulador e as políticas com “backoff”. De fato, uma implementação lógica (e não física) dos reguladores poderia ser feita através da criação de variáveis $z_{p,i}(t)$ indicando, para cada buffer regulado, qual seria o seu nível se esse buffer estivesse sendo alimentado pelo regulador correspondente. Visto dessa forma, fica claro que d_p -reguladores têm uma forte analogia com o mecanismo “backoff”, mas com a vantagem adicional de levar em conta o nível do buffer *real*, ou seja, na verdade toma-se $\min\{z_{p,i}(t), x_{p,i}(t)\}$ e não sempre $z_{p,i}(t)$ como naquele mecanismo. Evitam-se, assim, os períodos de ociosidade forçada que podem ocorrer com políticas com “backoff”, embora ainda permitindo, como naquela política, uma sub-utilização das máquinas, ao supor a taxa de entrada menor que a nominal.

Já a ligação com o Mecanismo Supervisor Estabilizador (USSM) é mais sutil; notando, porém, que o tempo máximo dedicado ao atendimento ininterrupto de um produto é proporcional a d_p , podemos pensar em reguladores alimentando os buffers à taxa d_p (lembremos que a constante γ_m apenas “corrige” o efeito do “set-up”). Os limites $\xi_{p,i}$ fazem o papel de *limiares* para o disparo do mecanismo, já que abaixo de um valor constante a política original pode

funcionar de qualquer maneira, não afetando a estabilidade, como veremos com mais detalhes no Capítulo 3.

2.3.4 Políticas de prioridades de buffers e componentes

Os mecanismos de estabilização vistos nas seções anteriores têm como característica o fato de se basearem nas taxas de chegada (d_p) e/ou nos tempos de processamento ($\tau_{p,i}$). Kumar e Lu [10] estudaram uma outra classe de políticas, baseadas em *prioridades de buffers e de componentes*. Essa abordagem difere fundamentalmente das anteriores em dois aspectos:

1. O modelo é *discreto*, uma vez que os elementos componentes dos produtos são tratados individualmente; assim, “processar um buffer” significa “processar 1 componente desse buffer”; além disso, as políticas são *não-preemptivas*, ou seja, o processamento de um componente nunca é interrompido.
2. A escolha do buffer a processar baseia-se no índice do buffer, em algumas políticas, ou em alguma característica do componente, em outras.
3. No modelo estudado não são considerados os tempos de preparação entre os buffers.

Nesta seção, comentaremos, sem nos aprofundarmos, algumas políticas analisadas em [10]. Dadas a diferenças de abordagem em relação ao contexto desta dissertação, o texto que segue terá caráter apenas informativo, não sendo pré-requisito para a leitura do restante do trabalho.

Como em [10], consideraremos inicialmente sistemas com 1 produto, sendo os buffers do caminho de processamento numerados de 1 a n .

Kumar e Lu analisaram as seguintes políticas:

1. *Atender-Antes-o-Primeiro-Buffer* ("First-Buffer-First-Served" - FBFS): Nessa política, escolhe-se o buffer mais próximo ao *início* do caminho, ou seja, o buffer com menor índice; se este estiver vazio, toma-se o seguinte com menor índice, e assim por diante. Privilegia-se, assim, os componentes menos processados.
2. *Atender-Antes-o-Último-Buffer* ("Last-Buffer-First-Served" - LBFS): Analogamente à anterior, escolhe-se o buffer mais próximo ao *final* do caminho, ou seja, aquele com maior índice. Privilegia-se, assim, os componentes mais processados, portanto mais próximos do acabamento.
3. *Data-de-Entrega-Mais-Cedo* ("Earliest-Due-Date" - EDD): Nessa política, cada componente, digamos π , tem associada a si uma *data de entrega* $\delta(\pi)$. No momento da escolha, toma-se o componente $\bar{\pi}$ cuja data de entrega é mais cedo, ou seja, aquele com menor valor $\delta(\bar{\pi})$. Prioriza-se, assim, o prazo de entrega do componente.
4. *Menor-Folga* ("Least-Slack" - LS): Aqui considera-se, além das datas de entrega $\delta(\pi)$, um valor ξ_i representando a estimativa do tempo de permanência no sistema de um componente do buffer i . Escolhe-se então o componente que, estima-se, estará ao final do processamento total com a menor folga em relação ao prazo de entrega. Ou seja, no momento da escolha, toma-se o componente $\bar{\pi}$ tal que

$$\delta(\bar{\pi}) - \xi_{i(\bar{\pi})}$$

é mínimo, onde $i(\pi)$ indica o buffer onde π está armazenado. Procura-se, dessa forma, atender os prazos de uma forma global, olhando para a possível data de finalização da produção de um componente.

5. *Atender-Antes-quem-Chega-Primeiro* ("First-Come-First-Served" - FCFS): Nessa política, considera-se simplesmente a ordem em que os componentes chegam aos respectivos buffers. Naturalmente, escolhe-se aquele que está há mais tempo esperando para ser processado, priorizando assim a ordem da produção.

Alguns comentários devem ser feitos aqui. Inicialmente, notemos que as duas primeiras políticas (FBFS e LBFS) baseiam-se nos índices dos buffers

candidatos, enquanto que as demais levam em conta alguma característica do componente em si. Outra observação é que, se os produtos chegam ao sistema na ordem em que devem ser entregues, então a política EDD coincide com LBFS. Finalmente, notemos que EDD é um caso particular de LS, onde $\xi_i = \xi$ para todo buffer i .

Kumar e Lu [10] demonstraram a estabilidade de algumas dessas políticas, bem como mostraram, através de um contra-exemplo, que nem toda política de prioridade de buffers é estável. Por simplicidade, enunciaremos a estabilidade em um único Teorema.

Teorema 2.8 *As políticas FBFS, LBFS, EDD e LS são estáveis para todo sistema em que a condição de capacidade está satisfeita.*

Devido à sua complexidade, não reproduziremos aqui a prova. É curioso notar que as demonstrações de FBFS e LBFS são completamente diferentes, a despeito da analogia entre as políticas: na primeira, é feita uma indução nos buffers, de 1 a n , mostrando que cada um replica a taxa de chegada no sistema; na segunda, muito mais complexa, mostra-se que o tempo de permanência de um componente π no sistema pode ser calculado em função do tempo gasto por uma máquina para processar os componentes à frente de π quando de sua chegada ao sistema, e a partir daí constrói-se uma função de Liapunov apropriada. A demonstração de LS é similar à de LBFS, enquanto que a de EDD segue trivialmente da observação acima que essa política é um caso particular de LS.

No caso de sistemas com vários produtos, Kumar e Lu mostraram que, ordenando os buffers do sistema apropriadamente, mantém-se a estabilidade de FBFS e LBFS, nada podendo afirmar-se sobre LS e EDD.

Finalmente, é importante observar que a estabilidade da política FCFS não foi demonstrada nem mesmo para sistemas com 1 só produto. Além de ser uma política bastante conhecida e utilizada na prática, os resultados de simulações relatados em [10] mostram um desempenho, medido em termos de Média + 3 Desvios Padrões do tempo de permanência de um componente no sistema, superior às demais políticas. A prova da estabilidade de FCFS

mostra-se, assim, uma importante questão em aberto.

Note-se, contudo que, se considerarmos *reguladores* no sistema, a estabilidade de FCFS segue trivialmente do Teorema 2.6, uma vez que é uma política usável. Ressalte-se porém que o mesmo princípio não vale para provar a estabilidade de FCFS através de políticas “backoff”, uma vez que haveria uma forma de diferença de sincronismo entre as entradas real e simulada — lembremos que o modelo em questão é discreto. Por outro lado, o mecanismo USSM modifica completamente a política quando os limiares de níveis de buffers são atingidos, descaracterizando a política original. Desta maneira, a única prova conhecida de FCFS adaptada permanece sendo através de reguladores.

Capítulo 3

Uma modelagem através de limiares sem tempos de preparação

A natureza dinâmica dos sistemas estudados neste trabalho nos leva a pensar em uma modelagem do problema através de um conjunto de equações diferenciais no tempo, onde as principais variáveis representam o nível de cada buffer.

Neste capítulo, faremos um estudo da estabilidade de Sistemas de Manufatura com esse enfoque, restringindo-nos a sistemas sem tempos de preparação. Inicialmente, descreveremos o sistema de equações utilizado, bem como as restrições impostas, verificando a compatibilidade de alguns conceitos vistos anteriormente. Para que o modelo fique bem definido, será preciso introduzir o conceito de políticas com *limiares* de atendimento.

Em seguida, utilizando algumas ferramentas da teoria de Sistemas Dinâmicos — basicamente funções de Liapunov — analisaremos algumas condições suficientes para a estabilidade do sistema. Especial atenção será dada ao caso de sistemas com 2 máquinas, para os quais é possível derivar condições de estabilidade diferentes daquelas vistas nos Capítulos 1 e 2.

3.1 O Modelo

Consideremos inicialmente, por simplicidade, sistemas com apenas 1 produto. Como anteriormente, vamos denotar por $x_i(t)$ o nível do buffer i no instante t , por $\mathbf{x}(t)$ a n -upla $(x_1(t), \dots, x_n(t))$ e diremos que o vetor $\mathbf{x}(t)$ indica o *estado* do sistema no instante t . Consideremos também a taxa de entrada instantânea no buffer i no instante t , e vamos denotá-la por $\nu_i(t)$. É claro que $\nu_i : \mathbb{R} \rightarrow \mathbb{R}$, já que as taxas de entrada são finitas.

Como a taxa de entrada instantânea em um buffer é a taxa de saída instantânea do buffer anterior, gostaríamos de escrever

$$\begin{cases} \dot{e}_i &= \nu_i \\ \dot{y}_i &= \nu_{i+1} \\ \dot{x}_i &= \nu_i - \nu_{i+1} \end{cases} \quad (3.1)$$

(onde $e_i(t)$ e $y_i(t)$ denotam, respectivamente, a quantidade introduzida no buffer i durante o intervalo $[0, t]$, e a quantidade produzida do buffer i durante esse intervalo) já que é razoável dizer que a variação instantânea do nível do buffer é a taxa de entrada menos a de saída (e denotamos por $\nu_{n+1}(t)$ a taxa de saída do buffer n). Contudo, como veremos adiante, será preciso uma hipótese adicional para garantir a diferenciabilidade de x_i , e_i e y_i .

Assumiremos a hipótese de que a taxa de saída de um buffer i que está sendo processado e não está vazio é $\frac{1}{\tau_i}$. Porém, se o buffer que está sendo servido estiver vazio, a taxa de saída passa a ser a de entrada, desde que esta não ultrapasse, é claro, $\frac{1}{\tau_i}$.

Assim, temos:

$$\nu_1(t) = d$$

e, para $1 \leq i \leq n$,

$$\nu_{i+1}(t) = \begin{cases} 0 & \text{se o buffer } i \text{ não está sendo processado} \\ \frac{1}{\tau_i} & \text{se } i \text{ está sendo processado e } x_i(t) > 0 \text{ ou } \nu_i(t) \geq \frac{1}{\tau_i} \\ \nu_i(t) & \text{se } i \text{ está sendo processado, } x_i(t) = 0 \text{ e } \nu_i(t) < \frac{1}{\tau_i} \end{cases}$$

Vamos denotar por $a_i(t)$ a função *atividade* do buffer (i) no instante t , isto é,

$$a_i(t) = \begin{cases} 1 & \text{se o buffer } i \text{ está sendo processado no instante } t \\ 0 & \text{caso contrário} \end{cases}$$

Note-se que as variáveis $a_i(\cdot)$ podem ser encaradas como *controles* externos que, na verdade, definem a política utilizada, uma vez que indicam os buffers *ativos* (isto é, que estão sendo processados) em cada instante de tempo.

Para que o modelo fique bem definido, assumiremos sempre a hipótese de que as funções $\nu_i(\cdot)$ e $a_i(\cdot)$, $i = 1, \dots, n$ só têm *descontinuidades de primeira espécie*, o que implica que a quantidade de pontos de descontinuidade dessas funções é enumerável. Ou seja, se $\{t_k\}_{k \in \mathbb{N}}$ é o conjunto de pontos de descontinuidade de $\nu_i(\cdot)$, então existem os limites laterais

$$\nu_i^+(t_k) = \lim_{t \rightarrow t_k^+} \nu_i(t) \quad \text{e} \quad \nu_i^-(t_k) = \lim_{t \rightarrow t_k^-} \nu_i(t)$$

Convencionaremos que $\nu_i(t_k) = \nu_i^+(t_k)$, ou seja, $\nu_i(\cdot)$ é sempre contínua à direita. Definindo analogamente para os $a_i(\cdot)$, a hipótese resume-se então a garantir, para $i = 1, \dots, n$, que

$$\nu_i(t) = \lim_{s \rightarrow t^+} \nu_i(s) \quad (3.2)$$

$$a_i(t) = \lim_{s \rightarrow t^+} a_i(s) \quad (3.3)$$

para todo $t \geq 0$.

Com esta hipótese, podemos então escrever as equações (3.1), com a ressalva de que $\dot{x}_i$, $\dot{e}_i$ e $\dot{y}_i$ representam *derivadas à direita* e assim

$$e_i(t) = \int_0^t \nu_i(s) ds \quad \text{e} \quad y_i(t) = \int_0^t \nu_{i+1}(s) ds$$

portanto

$$x_i(t) = e_i(t) - y_i(t) = \int_0^t (\nu_i(s) - \nu_{i+1}(s)) ds$$

É fácil ver que, uma vez que $\nu_i(\cdot)$ é limitada, então $x_i(\cdot)$, $e_i(\cdot)$ e $y_i(\cdot)$ são lipschitzianas para todo $t \geq 0$.

Durante todo este Capítulo, trabalharemos sempre com *políticas com limiares*. Por *limiares* entendam-se valores ϵ_i estritamente positivos associados a cada buffer (i) indicando o nível a partir do qual o buffer (i), se estiver sendo processado, deve sê-lo à taxa nominal. A razão de impormos tal condição deve-se ao fato de que certas políticas, como por exemplo as políticas CAF vistas no Capítulo 2, não estão definidas na origem. Assim, utilizando as funções definidas acima temos

$$a_i(t) = 1 \text{ e } x_i(t) \geq \epsilon_i \implies \nu_{i+1}(t) = \frac{1}{\tau_i} \quad \forall t \geq 0$$

Outra característica das políticas analisadas é que estas devem ser *conservantes* acima do limiar. Como vimos no Capítulo 1, uma política conservante é aquela que não permite que uma máquina permaneça ociosa ou subalimentada se um dos seus buffers está não-vazio. No nosso caso, imporemos a restrição apenas para a situação em que existe pelo menos um buffer da máquina acima do limiar, isto é, para qualquer máquina m e qualquer instante t vale que

$$\exists i \in B_m \text{ com } x_i(t) \geq \epsilon_i \implies a_{i^*}(t) = 1 \text{ e } \nu_{i^*+1}(t) = \frac{1}{\tau_{i^*}} \text{ para algum } i^* \in B_m \quad (3.4)$$

Observe-se que não é imposto que $x_{i^*}(t) \geq \epsilon_{i^*}$, mas devemos ter $x_{i^*}(t) > 0$ ou $\nu_{i^*}(t) \geq \frac{1}{\tau_{i^*}}$ para que o processamento de i^* possa ser feito à taxa nominal.

Vamos agora re-escrever a condição (3.4) denotando por H_m a região onde ela não é aplicada. Assim, dada uma máquina m , sejam $i_1, \dots, i_{k_m}$ os buffers de B_m , e vamos denotar por $\mathbf{x}_m(t)$ a k_m -upla $(x_{i_1}(t), \dots, x_{i_{k_m}}(t))$. Definindo H_m como o hipercubo (aberto em $\mathbb{R}_+^{k_m}$)

$$H_m = H_{i_1} \times H_{i_2} \times \dots \times H_{i_{k_m}}$$

onde

$$H_i = [0, \epsilon_i[\quad (3.5)$$

a condição (3.4) pode ser então re-escrita como

$$\mathbf{x}_m(t) \notin H_m \implies \begin{cases} a_{i^*}(t) = 1 & \text{para algum } i^* \in B_m \\ a_i(t) = 0 & \text{para } i \in B_m, i \neq i^* \\ \nu_{i+1}(t) = \frac{1}{\tau_i} a_i(t) & \text{para todo } i \in B_m \end{cases} \quad (3.6)$$

Devemos ressaltar aqui a importância de impor a condição (3.6) apenas para a situação em que existe pelo menos um buffer da máquina com nível acima do limiar. Como dentro do hipercubo H_m não é feita a restrição de que o buffer ativo tem que estar sendo processado à taxa nominal, então nessa região existe um grau de liberdade muito maior para a política utilizada; porém, as hipóteses (3.2) e (3.3) têm que valer, seja qual for a política proposta, a fim de que o sistema seja *fisicamente viável*. Note-se que, em particular, a política pode manter dentro do hipercubo o mesmo comportamento (inclusive a propriedade de ser conservante) que tem fora deste, ou seja, a introdução do limiar não significa uma restrição adicional.

Por outro lado, não é preciso, como observamos acima, que o buffer que está acima do limiar seja processado, ainda que este seja o único buffer da máquina nesta situação. Não fosse assim, e não seria possível a modelagem de políticas "clearing", nas quais um buffer é processado até que não contenha mais elementos.

Em vista disso, concluímos, assim, que a classe de políticas cobertas com esta modelagem não é restringida pela condição (3.6), uma vez que engloba todas as políticas conservantes e algumas não-conservantes; tal é o caso, por exemplo, das políticas "idling" introduzidas por Chase e Ramadge [1], nas quais permite-se que uma máquina trabalhe sub-alimentada mesmo tendo buffers positivos, desde que estes não tenham ultrapassado um valor pré-fixado. Claramente, tal valor equivale ao *limiar* considerado em (3.6).

Consideremos novamente a condição (3.6). Para $x(t)$ fora do hipercubo $H = H_1 \times \dots \times H_n$, e definindo $a_0(\cdot) \equiv 1$ e $\tau_0 = d^{-1}$, podemos escrever

$$\dot{e}_i = a_{i-1}(t) \frac{1}{\tau_{i-1}}$$

$$\dot{y}_i = a_i(t) \frac{1}{\tau_i}$$

e

$$\dot{x}_i(t) = a_{i-1}(t) \frac{1}{\tau_{i-1}} - a_i(t) \frac{1}{\tau_i} \quad i = 1, \dots, n$$

que pode ser expressa na forma

$$\dot{x}(t) = Ca(t)$$

onde

$$\dot{\mathbf{x}} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_{n-1} \\ \dot{x}_n \end{bmatrix}, \mathbf{C} = \begin{bmatrix} d & -\frac{1}{\tau_1} & 0 & \dots & \dots & 0 \\ 0 & \frac{1}{\tau_1} & -\frac{1}{\tau_2} & 0 & \dots & 0 \\ & & & \ddots & & \\ 0 & \dots & 0 & \frac{1}{\tau_{n-2}} & -\frac{1}{\tau_{n-1}} & 0 \\ 0 & \dots & \dots & 0 & \frac{1}{\tau_{n-1}} & -\frac{1}{\tau_n} \end{bmatrix}, \mathbf{a} = \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_{n-1} \\ a_n \end{bmatrix}$$

Para completar a descrição do modelo, observemos que é preciso impor a condição de que 2 buffers não podem estar ativos (isto é, em processamento) simultaneamente na mesma máquina, ou seja,

$$\sum_{i \in B_m} a_i(t) \leq 1 \quad \forall m \in \mathcal{M}$$

Re-escrevendo, temos então

$$\begin{cases} \dot{x}_i(t) = \nu_i(t) - \nu_{i+1}(t) \\ \mathbf{a} \in \mathcal{A} = \{\mathbf{a}(\cdot) : \mathbf{a}(t) \in \{0, 1\}^n, \sum_{i \in B_m} a_i(t) \leq 1, \forall m \in \mathcal{M} \quad \forall t \geq 0\} \end{cases}$$

Podemos agora estender a modelagem, de modo a considerar vários produtos. Ora, estamos considerando neste trabalho que os produtos são independentes entre si, no que diz respeito a taxas e buffers, isto é, não há *composição* de novos produtos a partir de produtos originais; logo, concluímos que o sistema de equações considerado vale individualmente para cada produto, e assim temos, para cada $p \in \mathcal{P}$,

$$\begin{cases} \dot{x}_{p,i}(t) = \nu_{p,i}(t) - \nu_{p,i+1}(t) & i = 1, \dots, n_p \\ \sum_{(p,i) \in B_m} a_{p,i}(t) \leq 1 & \forall m \in \mathcal{M} \end{cases} \quad (3.7)$$

com os limiares definidos por constantes $\epsilon_{p,i}$. Analogamente a (3.5) definimos

$$\begin{cases} H_{p,i} = [0, \epsilon_{p,i}[\\ H = H_{1,1} \times \dots \times H_{1,n_1} \times \dots \times H_{P,1} \times \dots \times H_{P,n_P} \end{cases} \quad (3.8)$$

Alguns comentários devem ser feitos neste ponto. Primeiramente, devemos observar a compatibilidade entre os conceitos de estabilidade em Sistemas

de Manufatura e estabilidade em Sistemas Dinâmicos. De fato, para este último existem vários conceitos de estabilidade; por exemplo, um sistema é dito estável *no sentido de Liapunov* em torno da solução $\bar{x} = 0$ se, dada uma vizinhança U_1 de $\bar{x}$, existe uma vizinhança U_2 de $\bar{x}$ tal que, se $x(t_0) \in U_2$, então $x(t) \in U_1$ para todo $t \geq t_0$. Claramente, é equivalente dizer que cada $x_i(t)$ é limitado por uma constante dependente do valor inicial, o que coincide com a definição de estabilidade vista no capítulo 1. Como falamos simplesmente em *sistemas estáveis*, utilizaremos o conceito de estabilidade *no sentido de Lagrange*, que corresponde exatamente àquela definição — as variáveis $x_{p,i}$ limitadas. Assim, utilizaremos o termo *sistema* para nos referirmos tanto a um Sistema de Manufatura quanto ao sistema de equações correspondente. Uma análise de diversos tipos de estabilidade em Sistemas Dinâmicos pode ser encontrada em [12].

Em segundo lugar, notemos que a aparente simplicidade do modelo esconde os fatos de que *a)* a hipótese (3.6) faz com que o conjunto de controles admissíveis dependa do *estado* do sistema; e *b)* a representação de uma política qualquer através das variáveis $a_i(\cdot)$ pode tornar a expressão de tais variáveis bastante complexa. Isto ocorre devido ao caráter *não-autônomo* de grande parte das políticas, no sentido de que a determinação dos buffers ativos num certo instante depende da informação em instantes passados; tal é o caso, por exemplo, das políticas CAF vistas no Capítulo 2, onde podemos dizer que um buffer está ativo se já o estava num instante imediatamente anterior e não está vazio no instante atual, desde que exista outro buffer não-vazio na mesma máquina.

Nesse contexto, uma política com “boa” representação seria, por exemplo, uma *política de prioridade de buffers preemptiva*, já que nesse caso os buffers são atendidos em ordem de prioridade, independentemente de qual buffer estava sendo processado. De fato, consideremos uma máquina m e uma ordenação $j_1, \dots, j_k$ dos buffers de B_m , de tal modo que j_1 é o mais prioritário, depois j_2 , e assim por diante. Definindo

$$u_i(t) = \begin{cases} 1 & \text{se } x_i(t) > 0 \\ 0 & \text{se } x_i(t) = 0 \end{cases}$$

temos, então,

$$\begin{aligned} a_{j_1} &= u_{j_1} \\ a_{j_2} &= u_{j_2}(1 - u_{j_1}) \\ &\vdots \\ a_{j_k} &= u_{j_k}(1 - u_{j_{k-1}}) \dots (1 - u_{j_1}) \end{aligned}$$

ou seja, j_1 é atendido se e só se $x_{j_1}(t) > 0$, j_2 é atendido se e só se $x_{j_2}(t) > 0$ e j_1 não está sendo atendido, etc. O sistema, nesse caso, seria re-escrito como

$$\dot{\mathbf{x}} = \mathbf{Cf}(\mathbf{u}(\mathbf{x}))$$

onde $\mathbf{f} = (f_1(\mathbf{u}), \dots, f_n(\mathbf{u}))$.

Finalmente, devemos lembrar que, uma vez que estamos supondo políticas conservantes acima do limiar, então *não consideraremos tempos de preparação*, já que durante tais períodos a condição (3.6) claramente é violada. Nesse contexto, portanto, toda política conservante é *usável*, de acordo com o Lema 1.3.

3.2 Condições suficientes de estabilidade

A modelagem do sistema através de (3.7), como desenvolvido acima, possibilita o estudo da estabilidade através de *funções de Liapunov*. Tal procedimento, denominado *método direto de Liapunov*, consiste em provar a estabilidade do sistema através de uma função apropriada $V(\mathbf{x}(t))$. Dessa maneira, tentaremos obter condições *suficientes* de estabilidade através da manipulação de tais funções.

No nosso caso, consideraremos funções da forma

$$V_\alpha(\mathbf{x}(\cdot)) = \sum_{p \in \mathcal{P}} \sum_{i=1}^{n_p} \alpha_{p,i} \tau_{p,i} x_{p,i}(\cdot) \quad (3.9)$$

onde os $\alpha_{p,i}$ são constantes estritamente positivas. Como cada $x_{p,i}(\cdot)$ é derivável à direita, então claramente $V_\alpha(\mathbf{x}(\cdot))$ também o é. Denotaremos por

$\dot{V}_\alpha(t)$ a derivada à direita de $V_\alpha(\mathbf{x}(t))$ em relação a t . Assim,

$$\dot{V}_\alpha(\cdot) = \frac{dV_\alpha(\mathbf{x}(\cdot))}{dt} = \nabla_{\mathbf{x}} V_\alpha \cdot \dot{\mathbf{x}}$$

e então

$$\dot{V}_\alpha(\cdot) = \sum_{p \in \mathcal{P}} \sum_{i=1}^{n_p} \alpha_{p,i} \tau_{p,i} \dot{x}_{p,i}(\cdot) \quad (3.10)$$

O Lema abaixo fornece uma condição para o estudo de estabilidade através de $\dot{V}_\alpha$.

Lema 3.1 *Sejam: $V_\alpha(\mathbf{x}(\cdot))$ uma função da classe definida acima, P uma política fisicamente viável (isto é, satisfazendo (3.2) e (3.3)), e H o hipercubo definido em (3.8). Se, para todo $t \geq 0$ vale a condição*

$$\mathbf{x}(t) \notin H \implies \dot{V}_\alpha(t) \leq 0$$

então o sistema é estável para a política P .

-Demonstração:

Vamos denotar por $V_\alpha(\epsilon)$ a expressão $\sum_{i=1}^{n_p} \alpha_{p,i} \tau_{p,i} \epsilon_{p,i}$. Seja Γ o número $\max\{V_\alpha(\mathbf{x}(0)), V_\alpha(\epsilon)\}$. Vamos mostrar que $x_{p,i}(t) \leq \Gamma$ em qualquer instante t , para todo buffer (p, i) .

Observemos inicialmente que, conforme comentamos na seção anterior, $x_{p,i}(\cdot)$ é lipschitziana para todo (p, i) , logo $V_\alpha(\mathbf{x}(\cdot))$ também o é, e portanto obviamente é contínua.

Suponhamos por absurdo que $V_\alpha(\mathbf{x}(\bar{t})) > \Gamma$ para algum $\bar{t}$. Por definição, $V_\alpha(\mathbf{x}(0)) \leq \Gamma$, logo pela continuidade de $V_\alpha(\mathbf{x}(\cdot))$ concluímos que

$$\exists t_0, 0 \leq t_0 < \bar{t} \mid V_\alpha(\mathbf{x}(t_0)) = \Gamma \text{ e } \dot{V}_\alpha(t_0) > 0 \quad (3.11)$$

Mas $\mathbf{x}(t_0) \notin H$, pois do contrário teríamos $V_\alpha(\mathbf{x}(t_0)) < V_\alpha(\epsilon) \leq \Gamma$. Pela hipótese do Lema, temos então $\dot{V}_\alpha(t_0) \leq 0$, contradizendo (3.11).

Concluimos assim que $V_\alpha(\mathbf{x}(t)) \leq \Gamma$ para todo t . Como os coeficientes $\alpha_{p,i}\tau_{p,i}$ de $x_{p,i}(t)$ em $V_\alpha(\mathbf{x}(t))$ são todos positivos, e $x_{p,i}(\cdot) \geq 0$, segue então que $x_{p,i}(\cdot) \leq \Gamma$ para todo buffer (p,i) , o que completa a prova. ■

O Lema acima afirma que, dados os coeficientes $\alpha_{p,i}$, se pudermos garantir que $\dot{V}_\alpha(t) \leq 0$ sempre que $\mathbf{x}(t)$ está fora do hipercubo, então o sistema será estável. De (3.10) temos a expressão de $\dot{V}_\alpha$:

$$\dot{V}_\alpha(t) = \sum_{p \in \mathcal{P}} \sum_{i=1}^{n_p} \alpha_{p,i} \tau_{p,i} (\nu_{p,i}(t) - \nu_{p,i+1}(t))$$

e então, lembrando que $\nu_{p,1}(\cdot) \equiv d_p$,

$$\begin{aligned} \dot{V}_\alpha(t) = & \sum_p \alpha_{p,1} \rho_{p,1} + \sum_p \sum_{i=1}^{n_p-1} (\alpha_{p,i+1} \tau_{p,i+1} - \alpha_{p,i} \tau_{p,i}) \nu_{p,i+1}(t) - \\ & - \sum_p \alpha_{p,n_p} \tau_{p,n_p} \nu_{p,n_p+1}(t) \end{aligned} \quad (3.12)$$

Consideremos inicialmente os coeficientes de $x_{p,i}$ unitários, isto é, $\alpha_{p,i} = 1$. Para esta função, que denotaremos por $V_{\underline{1}}$, a Proposição abaixo nos dá uma condição de estabilidade relacionada à proporção entre os tempos de processamento:

Proposição 3.1 *Se*

$$\max_{\substack{(p,i) \\ i \neq n}} \frac{\tau_{p,i+1}}{\tau_{p,i}} \leq 1 - \sum_p \rho_{p,1}$$

então o sistema é estável quando operado por uma política conservante com limiar.

-Demonstração:

Sejam $\mathbf{x}(\cdot)$ uma solução do sistema (3.7), t_0 um instante tal que $\mathbf{x}(t_0) \notin H_m$. Como a política é conservante com limiar, então pela condição (3.6) sabemos que existe um buffer (p^*, i^*) tal que $a_{p^*, i^*}(t_0) = 1$ e, além disso, $\nu_{p^*, i^*+1}(t_0) = \frac{1}{\tau_{p^*, i^*}} a_{p^*, i^*}(t_0)$ para todo buffer $(p, i) \in B_{m^*}$, onde $m^* = \mu_{p^*, i^*}$ é a máquina onde

(p^*, i^*) é processado. Relembrando de (3.12) a expressão de $\dot{V}_{\underline{1}}$, e tomando $\alpha_{p,i} = 1$ para todo (p, i) , temos:

$$\dot{V}_{\underline{1}}(t_0) = \sum_p \rho_{p,1} + (\tau_{p^*, i^*+1} - \tau_{p^*, i^*}) \frac{1}{\tau_{p^*, i^*}} + \sum_{(p,i) \notin B_{m^*}} (\tau_{p,i+1} - \tau_{p,i}) \nu_{p,i+1}(t_0)$$

onde definimos $\tau_{p, n_p+1} = 0$, para todo p , a fim de simplificar a equação.

Notemos agora que, como consequência da hipótese da Proposição, temos $\tau_{p,i+1} < \tau_{p,i}$ para todo (p, i) , logo podemos escrever

$$\dot{V}_{\underline{1}}(t_0) \leq \frac{\tau_{p^*, i^*+1}}{\tau_{p^*, i^*}} - (1 - \sum_p \rho_{p,1})$$

Novamente, pela hipótese da proposição, o lado direito da desigualdade acima é menor ou igual a zero, logo temos $\dot{V}_{\underline{1}}(t_0) \leq 0$. Logo, pelo Lema 3.1, o sistema é estável para qualquer política conservante com limiar. ■

Vamos tomar agora os coeficientes dos buffers nas funções V_α diferentes entre si. O Teorema abaixo fornece uma condição referente aos $\rho_{p,i}$.

Teorema 3.1 *Seja J um conjunto de buffers tal que:*

$$\rho_J = \sum_{(p,i) \in J} \rho_{p,i} < 1$$

Então, existem constantes $\alpha_{p,i} > 0$ tais que, em qualquer instante $t \geq 0$ vale a condição

$$\left(a_{p,i}(t) = 1 \text{ e } \nu_{p,i+1}(t) = \frac{1}{\tau_{p,i}} \text{ para algum } (p,i) \in J \right) \implies \dot{V}_\alpha(t) \leq 0$$

-Demonstração:

Sejam os $\alpha_{p,i}$ definidos da seguinte forma:

$$\alpha_{p,i} = \frac{1}{\tau_{p,i}} \left(\sum_{k \in J} \tau_{p,k} - \rho_J \sum_{\substack{k \in J \\ k < i}} \tau_{p,k} \right) \quad (3.13)$$

onde por simplicidade escrevemos $\sum_{k \in J}$ ao invés de $\sum_{\substack{(q,k) \in J \\ q=p}}$.

Seja (p, i) um buffer qualquer. Ora, é claro que $\sum_{k \in J} \tau_k \geq \sum_{\substack{k \in J \\ k < i}} \tau_k$. Como $\rho_J < 1$ por hipótese, segue que $\alpha_{p,i} > 0$, satisfazendo a condição imposta no Teorema.

Vamos agora re-escrever a expressão de $\dot{V}_\alpha$ da equação (3.12) para os $\alpha_{p,i}$ definidos acima. Para todo t , temos

$$\begin{aligned} \dot{V}_\alpha(t) = & \sum_p \alpha_{p,1} \rho_{p,1} + \sum_{\substack{(p,i) \in J \\ i < n_p}} (\alpha_{p,i+1} \tau_{p,i+1} - \alpha_{p,i} \tau_{p,i}) \nu_{p,i+1}(t) + \\ & \sum_{\substack{(p,i) \notin J \\ i < n_p}} (\alpha_{p,i+1} \tau_{p,i+1} - \alpha_{p,i} \tau_{p,i}) \nu_{p,i+1}(t) - \sum_p \alpha_{p,n_p} \tau_{p,n_p} \nu_{p,n_p+1}(t) \end{aligned}$$

Mas da definição dos $\alpha_{p,i}$ temos que, para $i < n_p$,

$$\alpha_{p,i+1} \tau_{p,i+1} = \begin{cases} \alpha_{p,i} \tau_{p,i} & \text{se } (p,i) \notin J \\ \alpha_{p,i} \tau_{p,i} - \rho_J \tau_{p,i} & \text{se } (p,i) \in J \end{cases}$$

Além disso,

$$\sum_p \alpha_{p,1} \rho_{p,1} = \sum_p \left(\frac{1}{\tau_{p,1}} \sum_{k \in J} \tau_{p,k} \rho_{p,1} \right) = \sum_p \sum_{k \in J} \rho_{p,k} = \rho_J$$

Logo, a expressão de $\dot{V}_\alpha(t)$ fica

$$\dot{V}_\alpha(t) = \rho_J \left(1 - \sum_{\substack{(p,i) \in J \\ i < n_p}} \tau_{p,i} \nu_{p,i+1}(t) \right) - \sum_p \alpha_{p,n_p} \tau_{p,n_p} \nu_{p,n_p+1}(t)$$

Seja t_0 um instante tal que $a_{p^*,i^*}(t_0) = 1$ e $\nu_{p^*,i^*}(t_0) = \frac{1}{\tau_{p^*,i^*}}$ para algum $(p^*, i^*) \in J$. Vamos mostrar que $\dot{V}_\alpha(t_0) \leq 0$.

Consideremos 2 possibilidades:

1. $i^* < n_{p^*}$

Nesse caso, temos $\tau_{p^*, i^*} \nu_{p^*, i^*}(t_0) = 1$ e então

$$\rho_J \left(1 - \sum_{\substack{(p, i) \in J \\ i < n_p}} \tau_{p, i} \nu_{p, i+1}(t_0) \right) \leq 0$$

Além disso, como $\alpha_{p, n_p} > 0$ e $\nu_{p, n_p+1}(\cdot) \geq 0 \forall p$, então obviamente vale que $\sum_p \alpha_{p, n_p} \tau_{p, n_p} \nu_{p, n_p+1}(t_0) \geq 0$.

2. $i^* = n_{p^*}$

Nesse caso, temos $\tau_{p^*, n_{p^*}} \nu_{p^*, n_{p^*}+1}(t_0) = 1$ e então

$$\rho_J \left(1 - \sum_{\substack{(p, i) \in J \\ i < n_p}} \tau_{p, i} \nu_{p, i+1}(t_0) \right) - \sum_p \alpha_{p, n_p} \tau_{p, n_p} \nu_{p, n_p+1}(t_0) \leq \rho_J - \alpha_{p^*, n_{p^*}} \stackrel{(3.13)}{=} \quad (3.13)$$

$$\begin{aligned} & \stackrel{(3.13)}{=} \rho_J - \frac{1}{\tau_{p^*, n_{p^*}}} \left(\sum_{k \in J} \tau_{p^*, k} - \rho_J \sum_{\substack{k \in J \\ k < n_{p^*}}} \tau_{p^*, k} \right) = \\ & = \frac{1}{\tau_{p^*, n_{p^*}}} \left(\rho_J \tau_{p^*, n_{p^*}} - \sum_{k \in J} \tau_{p^*, k} + \rho_J \sum_{\substack{k \in J \\ k < n_{p^*}}} \tau_{p^*, k} \right) = \\ & = \frac{1}{\tau_{p^*, n_{p^*}}} \sum_{k \in J} \tau_{p^*, k} (\rho_J - 1) < 0 \end{aligned}$$

pois $\rho_J < 1$.

Em qualquer caso, concluímos portanto que

$$\dot{V}_\alpha(t_0) \leq 0$$

■

Corolário 1 *Sejam: H o hipercubo definido em (3.8), J como no Teorema 3.1, P uma política fisicamente viável. Se em todo instante t vale que*

$$x(t) \notin H \implies \left(a_{p,i}(t) = 1 \text{ e } \nu_{p,i+1}(t) = \frac{1}{\tau_{p,i}} \text{ para algum } (p,i) \in J \right)$$

então o sistema é estável para a política P .

–Demonstração:

Imediato, a partir do Lema 3.1. ■

Nos casos particulares em que $J = \{\text{buffers do sistema}\}$ e $J = B_m$ para alguma máquina m , obtemos do Teorema acima as seguintes condições, que são demonstradas diretamente a partir da caracterização (3.6) de políticas conservantes com limiar.

Corolário 2 *Se*

$$\sum_{(p,i)} \rho_{p,i} < 1$$

então o sistema é estável para qualquer política conservante com limiar.

Corolário 3 *Seja P uma política fisicamente viável. Se existe uma máquina m que está sempre ativa trabalhando à taxa nominal, exceto talvez quando todos os buffers do sistema estão abaixo do limiar, então o sistema é estável para a política P .*

Observe-se que a Proposição 2.2, que garante a estabilidade de ciclos simples de 1 produto, é um exemplo de aplicação do Corolário 3.

3.3 Um caso particular: 2 máquinas

Ainda como consequência do Teorema 3.1, mostraremos que, para sistemas com 2 máquinas, é possível obter condições mais fracas de estabilidade.

Consideremos um sistema S com $\mathcal{M} = \{m_1, m_2\}$, sem tempos de preparação. Para $i = 1, 2$, vamos denotar por ρ_{p, m_i} a soma $\sum_{\substack{(q, j) \in B_{m_i} \\ q=p}} \rho_{q, j}$ e por ρ_{m_i} a expressão $\sum_p \rho_{p, m_i}$.

O Lema abaixo mostra condições suficientes de estabilidade para um tal sistema.

Lema 3.2 *Se existem constantes positivas c_1 e c_2 satisfazendo as condições abaixo:*

i) para todo $(p, i) \in B_{m_1}$, tem-se

$$c_1 \left(\sum_{k \in B_{m_1}} \tau_{p, k} - \rho_{m_1} \sum_{\substack{k \in B_{m_1} \\ k < i}} \tau_{p, k} \right) > c_2 \left(\sum_{k \in B_{m_2}} \tau_{p, k} - \rho_{m_2} \sum_{\substack{k \in B_{m_2} \\ k < i}} \tau_{p, k} \right)$$

ii) para todo $(p, i) \in B_{m_2}$, tem-se

$$c_2 \left(\sum_{k \in B_{m_2}} \tau_{p, k} - \rho_{m_2} \sum_{\substack{k \in B_{m_2} \\ k < i}} \tau_{p, k} \right) > c_1 \left(\sum_{k \in B_{m_1}} \tau_{p, k} - \rho_{m_1} \sum_{\substack{k \in B_{m_1} \\ k < i}} \tau_{p, k} \right)$$

Então, o sistema é estável para qualquer política conservante com limiar.

–*Demonstração:*

Consideremos as funções

$$\begin{aligned} V_\alpha(\mathbf{x}(\cdot)) &= \sum_{(p, i)} \alpha_{p, i} \tau_{p, i} x_{p, i}(\cdot) \\ V_\beta(\mathbf{x}(\cdot)) &= \sum_{(p, i)} \beta_{p, i} \tau_{p, i} x_{p, i}(\cdot) \end{aligned}$$

obtidas tomando-se, respectivamente, os conjuntos B_{m_1} e B_{m_2} no Teorema 3.1, e vamos definir uma função V da seguinte maneira:

$$V = \max\{c_1 V_\alpha, c_2 V_\beta\}$$

Sabemos do Teorema 3.1, e porque a política é conservante com limiar, que

$$\begin{aligned} \mathbf{x}_{m_1}(t) \notin H_{m_1} &\implies \dot{V}_\alpha(t) \leq 0 \implies c_1 \dot{V}_\alpha(t) \leq 0 \\ \mathbf{x}_{m_2}(t) \notin H_{m_2} &\implies \dot{V}_\beta(t) \leq 0 \implies c_2 \dot{V}_\beta(t) \leq 0 \end{aligned}$$

É claro que, se $\dot{V}_\alpha(t) \leq 0$ e $\dot{V}_\beta(t) \leq 0$, então $\dot{V}(t) \leq 0$. Para mostrar então que V satisfaz as condições do Lema 3.1, basta provar que existe uma região R em torno do zero fora da qual tem-se $\dot{V}(t) \leq 0$, para todo t .

Relembrando a definição dos $\alpha_{p,i}$ utilizada na demonstração do Teorema 3.1 (ver a equação (3.13)) temos

$$\alpha_{p,i} = \frac{1}{\tau_{p,i}} \left(\sum_{k \in m_1} \tau_{p,k} - \rho_{m_1} \sum_{\substack{k \in B_{m_1} \\ k < i}} \tau_{p,k} \right) \quad 1 \leq i \leq n$$

e então

$$\alpha_{p,i} = \frac{1}{\rho_{p,i}} \left(\rho_{p,m_1} - \rho_{m_1} \sum_{\substack{k \in B_{m_1} \\ k < i}} \rho_{p,k} \right) \quad 1 \leq i \leq n$$

A definição dos $\beta_{p,i}$ é análoga, trocando-se m_1 por m_2 , logo

$$\beta_{p,i} = \frac{1}{\rho_{p,i}} \left(\rho_{p,m_2} - \rho_{m_2} \sum_{\substack{k \in B_{m_2} \\ k < i}} \rho_{p,k} \right) \quad 1 \leq i \leq n$$

Notemos que

$$\begin{aligned} &(c_1 \alpha_{p,i} - c_2 \beta_{p,i}) \tau_{p,i} = \\ &= \frac{c_1}{\rho_{p,i}} \left(\rho_{p,m_1} - \rho_{m_1} \sum_{\substack{k \in B_{m_1} \\ k < i}} \rho_{p,k} \right) \tau_{p,i} - \frac{c_2}{\rho_{p,i}} \left(\rho_{p,m_2} - \rho_{m_2} \sum_{\substack{k \in B_{m_2} \\ k < i}} \rho_{p,k} \right) \tau_{p,i} \\ &= c_1 \left(\tau_{p,m_1} - \rho_{m_1} \sum_{\substack{k \in B_{m_1} \\ k < i}} \tau_{p,k} \right) - c_2 \left(\tau_{p,m_2} - \rho_{m_2} \sum_{\substack{k \in B_{m_2} \\ k < i}} \tau_{p,k} \right) \end{aligned}$$

onde τ_{p,m_i} denota, naturalmente, $\frac{\rho_{p,m_i}}{d_p}$.

Assim, pelas hipóteses do Lema temos

$$c_1 \alpha_{p,i} > c_2 \beta_{p,i} \quad (p,i) \in B_{m_1} \quad (3.14)$$

$$c_2 \beta_{p,i} > c_1 \alpha_{p,i} \quad (p,i) \in B_{m_2} \quad (3.15)$$

Vamos escrever

$$\begin{aligned} V_\alpha &= \sum_{(p,i) \in B_{m_1}} \alpha_{p,i} \tau_{p,i} x_{p,i} + \sum_{(p,i) \in B_{m_2}} \alpha_{p,i} \tau_{p,i} x_{p,i} \\ V_\beta &= \sum_{(p,i) \in B_{m_1}} \beta_{p,i} \tau_{p,i} x_{p,i} + \sum_{(p,i) \in B_{m_2}} \beta_{p,i} \tau_{p,i} x_{p,i} \end{aligned}$$

e consideremos então os dois casos:

1. $x_{m_1}(t) \notin H_{m_1}$ e $x_{m_2}(t) \in H_{m_2}$

Vamos mostrar que existe uma constante positiva Γ_1 tal que

$$c_2 V_\beta(x(t)) \leq \max\{c_1 V_\alpha(x(t)), \Gamma_1\} \quad (3.16)$$

Suponhamos que $c_2 V_\beta(x(t)) > \Gamma_1$. Temos então

$$c_2 \sum_{(p,i) \in B_{m_1}} \beta_{p,i} \tau_{p,i} x_{p,i} + c_2 \sum_{(p,i) \in B_{m_2}} \beta_{p,i} \tau_{p,i} x_{p,i} > \Gamma_1$$

Como $x_{m_2}(t) \in H_{m_2}$, isto é, $x_{p,i}(t) < \epsilon_{p,i}$ para todo $(p,i) \in B_{m_2}$, segue da desigualdade acima que

$$c_2 \sum_{(p,i) \in B_{m_1}} \beta_{p,i} \tau_{p,i} x_{p,i} > \Gamma_1 - c_2 V_\beta(\epsilon_{m_2})$$

onde $V_\beta(\epsilon_{m_2})$ denota a soma $\sum_{(p,i) \in B_{m_2}} \beta_{p,i} \tau_{p,i} \epsilon_{p,i}$. Definindo

$$\begin{aligned} \bar{\beta} &= \max_{(p,i) \in B_{m_1}} \beta_{p,i} \\ \bar{\tau} &= \max_{(p,i)} \tau_{p,i} \end{aligned}$$

temos então

$$\sum_{(p,i) \in B_{m_1}} x_{p,i}(t) > \frac{\Gamma_1 - c_2 V_\beta(\epsilon_{m_2})}{c_2 \bar{\beta} \bar{\tau}} = \Gamma'_1 \quad (3.17)$$

Tomemos

$$\Gamma'_1 = \frac{\sum_{(p,i) \in B_{m_2}} (c_2 \beta_{p,i} - c_1 \alpha_{p,i}) \tau_{p,i} \epsilon_{p,i}}{\min_{(p,i) \in B_{m_1}} (c_1 \alpha_{p,i} - c_2 \beta_{p,i}) \tau_{p,i}}$$

Note-se que, de acordo com (3.14) e (3.15), Γ'_1 tem o numerador e o denominador positivos. Da equação acima vem que

$$\min_{(p,i) \in B_{m_1}} (c_1 \alpha_{p,i} - c_2 \beta_{p,i}) \tau_{p,i} \Gamma'_1 = \sum_{(p,i) \in B_{m_2}} (c_2 \beta_{p,i} - c_1 \alpha_{p,i}) \tau_{p,i} \epsilon_{p,i} \quad (3.18)$$

Lembrando que

$$\begin{aligned} \sum_{(p,i) \in B_{m_1}} (c_1 \alpha_{p,i} - c_2 \beta_{p,i}) \tau_{p,i} x_{p,i}(t) &\geq \min_{(p,i) \in B_{m_1}} (c_1 \alpha_{p,i} - c_2 \beta_{p,i}) \tau_{p,i} \sum_{(p,i) \in B_{m_1}} x_{p,i}(t) \\ &\stackrel{(3.17)}{>} \min_{(p,i) \in B_{m_1}} (c_1 \alpha_{p,i} - c_2 \beta_{p,i}) \tau_{p,i} \Gamma'_1 \end{aligned}$$

concluimos então, de (3.18) que

$$\begin{aligned} \sum_{(p,i) \in B_{m_1}} (c_1 \alpha_{p,i} - c_2 \beta_{p,i}) \tau_{p,i} x_{p,i}(t) &> \sum_{(p,i) \in B_{m_2}} (c_2 \beta_{p,i} - c_1 \alpha_{p,i}) \tau_{p,i} \epsilon_{p,i} \\ &> \sum_{(p,i) \in B_{m_2}} (c_2 \beta_{p,i} - c_1 \alpha_{p,i}) \tau_{p,i} x_{p,i}(t) \end{aligned}$$

ou seja,

$$c_1 V_\alpha(\mathbf{x}(t)) > c_2 V_\beta(\mathbf{x}(t))$$

e portanto vale (3.16).

2. $\mathbf{x}_{m_1}(t) \in H_{m_1}$ e $\mathbf{x}_{m_2}(t) \notin H_{m_2}$

Analogamente ao caso anterior, concluimos que

$$c_1 V_\alpha(\mathbf{x}(t)) \leq \max\{c_2 V_\beta(\mathbf{x}(t)), \Gamma_2\}$$

onde

$$\frac{\Gamma_2 - c_1 V_\alpha(\epsilon_{m_1})}{c_1 \bar{\alpha} \bar{\tau}} = \Gamma'_2 = \frac{\sum_{(p,i) \in B_{m_1}} (c_1 \alpha_{p,i} - c_2 \beta_{p,i}) \tau_{p,i} \epsilon_{p,i}}{\min_{(p,i) \in B_{m_2}} (c_2 \beta_{p,i} - c_1 \alpha_{p,i}) \tau_{p,i}}$$

Resumindo os casos 1 e 2 acima, temos então que, em todo instante t valem

$$\begin{aligned} (\mathbf{x}_{m_1}(t) \notin H_{m_1} \text{ e } \mathbf{x}_{m_2}(t) \in H_{m_2}) &\implies V(\mathbf{x}(t)) \leq \max\{c_1 V_\alpha(\mathbf{x}(t)), \Gamma_1\} \\ (\mathbf{x}_{m_1}(t) \in H_{m_1} \text{ e } \mathbf{x}_{m_2}(t) \notin H_{m_2}) &\implies V(\mathbf{x}(t)) \leq \max\{c_2 V_\beta(\mathbf{x}(t)), \Gamma_2\} \end{aligned}$$

Definindo portanto $\Gamma = \max\{\Gamma_1, \Gamma_2\}$ e a região $R = ([0, \Gamma])^{P_{n_P}}$, concluímos que V tem derivada não-positiva fora de R , logo pelo Lema 3.1 o sistema é estável. ■

Uma aplicação do Lema 3.2 pode ser percebida se considerarmos um sistema S com $\mathcal{M} = \{m_1, m_2\}$ e apenas 1 produto (omitiremos assim o índice do produto, para simplificar a notação). Sem perda de generalidade, vamos supor que o buffer (1) está em B_{m_1} , e denotaremos por $k_1, \dots, k_r$ os buffers onde há mudança de máquina, convencionando que $k_r = n$; imporemos também a restrição de que $k_1 = 1$. A figura abaixo mostra um sistema típico.

Assim, $\mu_{k_j} \neq \mu_{k_{j+1}}$, $j = 1, \dots, r-1$. Note-se que o valor de r é arbitrário, exceto é claro pelo fato de que $1 \leq r \leq n$. Sem perda de generalidade, suporemos $r \geq 3$, uma vez que se $r = 1$ ou $r = 2$ então o sistema é acíclico e, como vimos no Capítulo 1, é estável para qualquer política usável. Ou seja, o sistema considerado é um sistema qualquer com 2 máquinas e 1 produto (sem tempos de preparação), exceto pela restrição de que $k_1 = 1$. Note-se, contudo, que tal restrição não é forte, uma vez que a estabilidade desse sistema implica na estabilidade de qualquer outro obtido a partir deste — modificando-se apenas o valor de k_1 —, bastando para isso impor que os buffers $1, \dots, k_1$ sejam sempre processados consecutivamente.

A Proposição abaixo mostra uma condição suficiente de estabilidade para tais sistemas.

Proposição 3.2 *Se, num sistema com as características acima, os tempos de processamento são tais que:*

$$i) \tau_{k_j} > \tau_{k_{j+1}} + \dots + \tau_{k_{j+1}-1} \quad j = 1, \dots, r-1$$

$$ii) \tau_{k_j} < \tau_{k_{j+1}} + \dots + \tau_{k_{j+1}-1} + \tau_{k_{j+1}} \quad j = 1, \dots, r-2$$

então, o sistema é estável para qualquer política conservante com limiar.

-Demonstração:

Vamos mostrar que $c_1 = \rho_{m_2}$ e $c_2 = \rho_{m_1}$ satisfazem as condições do Lema 3.2. Como só há 1 produto, então

$$\rho_{m_i} = \sum_{k \in B_{m_i}} \rho_k = d \sum_{k \in B_{m_i}} \tau_k$$

Logo, as condições do Lema podem ser re-escritas como

$$i) \rho_{m_2} \frac{\rho_{m_1}}{d} \left(1 - \sum_{\substack{k \in B_{m_1} \\ k < i}} \rho_k \right) > \rho_{m_1} \frac{\rho_{m_2}}{d} \left(1 - \sum_{\substack{k \in B_{m_2} \\ k < i}} \rho_k \right) \quad \forall i \in B_{m_1}$$

que é equivalente a

$$\sum_{\substack{k \in B_{m_1} \\ k < i}} \rho_k < \sum_{\substack{k \in B_{m_2} \\ k < i}} \rho_k \quad \forall i \in B_{m_1} \quad (3.19)$$

$$ii) \rho_{m_1} \frac{\rho_{m_2}}{d} \left(1 - \sum_{\substack{k \in B_{m_2} \\ k < i}} \rho_k \right) > \rho_{m_2} \frac{\rho_{m_1}}{d} \left(1 - \sum_{\substack{k \in B_{m_1} \\ k < i}} \rho_k \right) \quad \forall i \in B_{m_2}$$

que é equivalente a

$$\sum_{\substack{k \in B_{m_2} \\ k < i}} \rho_k < \sum_{\substack{k \in B_{m_1} \\ k < i}} \rho_k \quad \forall i \in B_{m_2} \quad (3.20)$$

Vamos mostrar inicialmente, por indução em i , que a expressão

$$f_i = \sum_{\substack{k \in B_{m_2} \\ k < i}} \tau_k - \sum_{\substack{k \in B_{m_1} \\ k < i}} \tau_k$$

é positiva, para todo $i \in B_{m_1}$.

Para $i = k_2 + 1$, que é o primeiro buffer que "volta" a m_1 , temos $f_{k_2+1} = \tau_2 + \dots + \tau_{k_2} - \tau_{k_1}$, logo pela condição (ii) do Teorema segue que $f_{k_2+1} > 0$.

Seja $j \in B_{m_1}$, $j < n$, e suponhamos que a afirmação vale para todo $i \in B_{m_1}$, $i \leq j$. Seja k_l o último buffer de m_2 antes de j (existe tal k_l , pois $k_1 = 1$ e $n > 1$). Temos duas situações:

1. $j + 1 \in B_{m_1}$: Nesse caso, temos $k_l < j < k_{l+1}$. Como $j \in B_{m_1}$, o primeiro termo da expressão de f_j é igual ao de f_{j+1} , logo

$$\begin{aligned} f_{j+1} &= \sum_{\substack{k \in B_{m_2} \\ k < j}} \tau_k - \sum_{\substack{k \in B_{m_1} \\ k \leq j}} \tau_k = \sum_{\substack{k \in B_{m_2} \\ k < k_l}} \tau_k + \tau_{k_l} - \left(\sum_{\substack{k \in B_{m_1} \\ k < k_l}} \tau_k + \sum_{\substack{k \in B_{m_1} \\ k_l < k \leq j}} \tau_k \right) \\ &= f_{k_l} + \tau_{k_l} - (\tau_{k_l+1} + \dots + \tau_j) \end{aligned}$$

Ora, $f_{k_l} > 0$, pela hipótese de indução. Por outro lado, pela hipótese (i) da Proposição temos

$$\tau_{k_l} > \tau_{k_l+1} + \dots + \tau_{k_{l+1}-1} \geq \tau_{k_l+1} + \dots + \tau_j$$

e portanto $f_{j+1} > 0$.

2. $j + 1 \notin B_{m_1}$: Nesse caso, temos $j = k_{l+1}$, logo $k_{l+1} + 1, \dots, k_{l+2} \in B_{m_2}$ e portanto temos que mostrar que $f_{k_{l+2}+1} > 0$ (se $k_{l+2} = n$, isto é, se $l + 2 = r$, então j é o último buffer de B_{m_1} , logo não há o que provar). Mas

$$f_{k_{l+2}+1} = \sum_{\substack{k \in B_{m_2} \\ k \leq k_{l+1}}} \tau_k + \sum_{k_{l+1} < k \leq k_{l+2}} \tau_k - \sum_{\substack{k \in B_{m_1} \\ k \leq k_{l+1}}} \tau_k$$

$$\begin{aligned}
&= \sum_{\substack{k \in B_{m_2} \\ k < k_{l+1}}} \tau_k + \sum_{k_{l+1} < k \leq k_{l+2}} \tau_k - \left(\sum_{\substack{k \in B_{m_1} \\ k < k_{l+1}}} \tau_k + \tau_{k_{l+1}} \right) \\
&= f_{k_{l+1}} + (\tau_{k_{l+1}+1} + \dots + \tau_{k_{l+2}}) - \tau_{k_{l+1}}
\end{aligned}$$

Como $f_{k_{l+1}} = f_j > 0$ pela hipótese de indução, e o segundo termo é positivo pela hipótese (ii) da Proposição, segue que $f_{k_{l+1}+2} > 0$.

De 1 e 2 acima concluímos que $\sum_{\substack{k \in B_{m_2} \\ k < i}} \tau_k - \sum_{\substack{k \in B_{m_1} \\ k < i}} \tau_k > 0 \forall i \in B_{m_1}$, certificando portanto a condição (3.19).

Para mostrar a condição (3.20), vamos definir, analogamente ao caso acima,

$$f_i = \sum_{\substack{k \in B_{m_1} \\ k < i}} \tau_k - \sum_{\substack{k \in B_{m_2} \\ k < i}} \tau_k$$

e vamos mostrar que $f_i > 0$ para todo $i \in B_{m_2}$.

Para $i = 2$, temos $f_2 = \tau_1 > 0$.

Seja $j \in B_{m_2}$, e suponhamos que a afirmação vale para todo $i \in B_{m_2}$, $i \leq j$. Seja k_l o último buffer de m_1 antes de j . Ora, com um raciocínio idêntico ao desenvolvido no caso 1 concluímos que

1. $j < k_{l+1} \implies f_{j+1} > 0$
2. $j = k_{l+1} \implies j+1, \dots, k_{l+2} \in B_{m_1}$ e $f_{k_{l+2}+1} > 0$

Logo, temos $\sum_{\substack{k \in B_{m_1} \\ k < i}} \tau_k - \sum_{\substack{k \in B_{m_2} \\ k < i}} \tau_k > 0 \forall i \in B_{m_2}$, portanto vale (3.20).

Aplicando então o Lema 3.2, concluímos que o sistema é estável para qualquer política conservante com limiar. ■

A figura abaixo mostra um exemplo de aplicação das condições obtidas. Note-se que o sistema não satisfaz nenhuma das condições anteriormente estudadas. Por exemplo, a condição de *estabilidade condicional* vista no Capítulo 2 não é satisfeita para nenhuma das duas máquinas, uma vez que (supondo $d=1$)

$$\begin{aligned}\rho'_{m_1} &= 0.5 + \frac{1}{0.3}(0.2 + 0.2) > 1 \\ \rho'_{m_2} &= \frac{1}{0.5}(0.3 + 0.1 + 0.3) > 1\end{aligned}$$

Observemos que neste caso temos $k_1 = 1$, $k_2 = 4$ e $k_3 = 6$.

Por outro lado, se as condições de estabilidade condicional forem satisfeitas, então certamente a condição (ii) da Proposição 3.2 não valerá, uma vez que nesse caso teremos $\frac{1}{\tau_{k_j}}(\tau_{k_j+1} + \dots + \tau_{k_{j+1}-1} + \tau_{k_{j+1}}) < 1$, $j = 1, \dots, r-2$. Concluimos assim que as condições obtidas e as condições de estabilidade condicional são disjuntas.

Outra aplicação interessante da Proposição 3.2 é para o caso em que não há laços em nenhuma das duas máquinas. É fácil ver que, neste caso, temos $k_j = j$, $j = 1, \dots, n$, e portanto a condição suficiente de estabilidade decorrente daquela Proposição é escrita como

$$\tau_1 < \tau_2 < \dots < \tau_{n-1}$$

Capítulo 4

Analizando a Conjectura de Tempos Decrescentes

Vimos, ao longo deste trabalho, diversas condições, encontradas na literatura, que garantem a estabilidade de um sistema onde certos requisitos são satisfeitos. Conforme comentamos no Capítulo 2, tais condições são contudo muito restritas, e a questão de obter condições suficientes mais fracas permanece em aberto.

Também vimos no Capítulo 2 que, com esse objetivo, Humes e Kumar conjecturaram (independentemente) que um sistema seria *intrinsecamente estável* se os tempos de processamento fossem sucessivamente não-crescentes. Além de ampliar consideravelmente a classe dos sistemas intrinsecamente estáveis, a comprovação da validade da Conjectura teria grande impacto sobre os Mecanismos Reguladores estudados no Capítulo 2 já que, como discutimos ali, obteríamos um mecanismo de estabilização informacionalmente eficiente, uma vez que não dependeria das taxas d_p de entrada no sistema. Efetivamente, simulações realizadas principalmente com os dois contra-exemplos de estabilidade para políticas CAF (ver a seção 2.1) corroboraram tal hipótese.

Neste capítulo, estudaremos a Conjectura sob dois aspectos: inicialmente, analisaremos os dois sistemas apresentados por Kumar e Seidman [11] como

contra-exemplos de instabilidade, mostrando que, para esses casos, a Conjectura é verdadeira, o que explica os resultados das simulações; em seguida, exibiremos uma topologia de sistema para a qual existem condições adicionais sobre os tempos de processamento que, se satisfeitas, levam o sistema à instabilidade, quando operado sob uma determinada política CAF, demonstrando assim a falsidade da afirmação proposta. Alternativamente, exibiremos na seção 4.2 uma análise de um sistema específico através do seu *diagrama de estados*, mostrando que tal técnica, embora de aplicação bastante restrita, possibilita a obtenção de condições *necessárias* e *suficientes* de estabilidade.

A importância dedicada ao estudo da Conjectura deve-se não só à sua relevância no contexto de estabilidade mas principalmente por ser a motivação principal deste trabalho. Para facilitar a referência ao longo do capítulo, vamos enunciá-la novamente:

Conjectura 1 *Se os tempos de processamento de um sistema são tais que*

$$\tau_{p,i} \geq \tau_{p,i+1} \quad \forall p, i = 1, \dots, n_p$$

então o sistema é intrinsecamente estável.

4.1 Casos válidos

4.1.1 O exemplo 1 de Kumar e Seidman

Consideremos o sistema

sem tempos de preparação. No Capítulo 2 (seção 2.1) vimos que, obedecidas as condições de capacidade $\rho_1 + \rho_4 < 1$ e $\rho_2 + \rho_3 < 1$ e *partindo da configuração*

$$\begin{aligned} x_1(t_0) &= x_0 > 0 \\ x_2(t_0) &= x_3(t_0) = x_4(t_0) = 0 \end{aligned}$$

podemos concluir que, sob uma política CAF, o sistema é

$$\begin{aligned} &\text{estável, se } \rho_2 + \rho_4 \leq 1 \\ &\text{instável, se } \rho_2 + \rho_4 > 1 \end{aligned}$$

A proposição abaixo mostra uma condição suficiente de estabilidade, para uma configuração inicial qualquer e outras políticas:

Proposição 4.1 *O sistema acima é estável para qualquer política conservante fortemente usável se $\tau_1 > \tau_2$.*

–*Demonstração:*

Vamos aplicar a Proposição 3.2. Os buffers onde há mudança de máquina são 1,3 e 4; logo, de acordo com aquela Proposição, uma condição suficiente de estabilidade é que os τ_i satisfaçam

$$\begin{aligned} \tau_1 &> \tau_2 \\ \tau_1 &< \tau_2 + \tau_3 \end{aligned}$$

Suponhamos então $\tau_1 > \tau_2$. Se $\tau_1 < \tau_2 + \tau_3$, então as condições acima estão satisfeitas. Por outro lado, se $\tau_1 \geq \tau_2 + \tau_3$, então a máquina m_2 é *condicionalmente estável* (conforme o conceito visto no Capítulo 2), pois nesse caso o “grau de congestionamento máximo” de m_2 é

$$\frac{1}{\tau_1}(\tau_2 + \tau_3) \leq 1$$

e portanto, uma vez que não há tempos de preparação, podemos eliminar m_2 do sistema, como no Teorema 2.3, se a política é fortemente usável. Obtemos assim um sistema de apenas 1 máquina, que é estável, portanto de acordo com

aquele Teorema o sistema é estável para toda política usável, em particular para uma conservante.

Portanto, $\tau_1 > \tau_2$ é condição suficiente de estabilidade para o sistema em questão. ■

Note-se ainda que

$$\tau_2 < \tau_1 \implies \rho_2 < \rho_1 < 1 - \rho_4$$

ou seja,

$$\tau_2 < \tau_1 \implies \rho_2 + \rho_4 < 1$$

4.1.2 O exemplo 2 de Kumar e Seidman

Consideremos agora o sistema

sem tempos de preparação. Como no caso anterior, vimos no Capítulo 2 que, partindo da configuração

$$\begin{aligned} x_{11}(t_0) &= x_0 > 0 \\ x_{12}(t_0) &= x_{21}(t_0) = x_{22}(t_0) = 0 \end{aligned}$$

(onde x_{11} denota $x_{1,1}$, e assim por diante) podemos concluir que, sob uma política CAF, o sistema é

$$\begin{aligned} &\text{estável, se } \rho_{12} + \rho_{22} \leq 1 \\ &\text{instável, se } \rho_{12} + \rho_{22} > 1 \end{aligned}$$

A proposição abaixo mostra que a Conjectura de tempos não-crescentes é válida para esse sistema:

Proposição 4.2 *O sistema acima é estável para qualquer política conservante se $\tau_{11} > \tau_{12}$ e $\tau_{21} > \tau_{22}$.*

–*Demonstração:*

Sejam $c_1 = \tau_{12}\tau_{21}$, $c_2 = \tau_{11}\tau_{22}$. Vamos mostrar que c_1 e c_2 satisfazem as condições do Lema 3.2. Ora, $J_1 = \{(1, 1), (2, 2)\}$ e $J_2 = \{(1, 2), (2, 1)\}$, logo as condições do Lema são re-escritas como:

$$\text{i) para } (p, i) = (1, 1): \tau_{12}\tau_{21} \tau_{11} > \tau_{11}\tau_{22} \tau_{12}, \text{ ou seja,}$$

$$\tau_{21} > \tau_{22}$$

$$\text{ii) para } (p, i) = (2, 2): \tau_{12}\tau_{21} \tau_{22} > \tau_{11}\tau_{22}(\tau_{21} - \rho_{m_2}\tau_{21}), \text{ ou seja,}$$

$$\tau_{12} > \tau_{11}(1 - \rho_{m_2})$$

$$\text{iii) para } (p, i) = (1, 2): \tau_{11}\tau_{22} \tau_{12} > \tau_{12}\tau_{21}(\tau_{11} - \rho_{m_1}\tau_{11}), \text{ ou seja,}$$

$$\tau_{22} > \tau_{21}(1 - \rho_{m_1})$$

$$\text{iv) para } (p, i) = (2, 1): \tau_{11}\tau_{22} \tau_{21} > \tau_{12}\tau_{21} \tau_{22}, \text{ ou seja,}$$

$$\tau_{11} > \tau_{12}$$

As condições (i) e (iv) seguem diretamente da hipótese de tempos não-crescentes. Consideremos agora as taxas máximas de entrada nos buffers de m_1 . Temos $d'_{11} = d_1$, $d'_{22} = \frac{1}{\tau_{21}}$. Logo, se

$$\rho'_{m_1} = \rho_{11} + \frac{1}{\tau_{21}}\tau_{22} \leq 1$$

então m_1 é condicionalmente estável (já que não há tempos de preparação, e portanto a condição pode ser relaxada), e assim podemos efetuar um “curto-circuito” de m_1 , logo o sistema é estável. Por outro lado, se

$$\rho_{11} + \frac{1}{\tau_{21}}\tau_{22} > 1$$

então temos

$$\tau_{22} > \tau_{21}(1 - \rho_{11}) > \tau_{21}(1 - \rho_{11} - \rho_{22})$$

satisfazendo portanto a condição (iii).

Analogamente, concluímos que ou m_2 é condicionalmente estável (e portanto o sistema é estável) ou vale a condição (ii). Portanto, pelo Lema 3.2, o sistema é estável para qualquer política conservante. ■

4.2 Analisando o diagrama de estados para políticas “clearing”

Uma abordagem alternativa no estudo de estabilidade pode ser conseguida através da análise de um *diagrama de estados* do sistema. Para ilustrar tal técnica, tomemos novamente o Exemplo 2 de Kumar e Seidman, e vamos estudar o comportamento do sistema sob uma política CAF a partir de uma configuração inicial qualquer. O diagrama será representado por um grafo, definido da seguinte maneira:

- cada nó do grafo representa uma possível combinação dos *sinais* das variáveis $x_{p,i}$, juntamente com a indicação dos *buffers ativos* naquele instante;
- dois nós e_1 e e_2 estão ligados por um arco quando existe uma *transição* do estado e_1 para o estado e_2 . Uma vez que no máximo 2 buffers podem estar ativos simultaneamente, e como o sistema é operado sob uma política “clearing”, então as transições sempre ocorrem quando um buffer é “zerado”. Assim, dado um estado e_j , denotaremos por $T_{p,i}^j$ o tempo necessário para esvaziar o buffer (p, i) a partir do seu valor no estado e_j .

Em cada nó, ao lado do sinal de $x_{p,i}$ indicaremos também o sinal da *derivada à direita* $\dot{x}_{p,i}$, e colocaremos o rótulo “SUB” nos estados onde ocorre subalimentação de alguma máquina.

A figura abaixo mostra, como exemplo, um subgrafo do grafo de estados.

A transição (e_1, e_2) ocorre se o buffer $(1, 1)$ é esvaziado antes de $(2, 1)$. Neste caso, o buffer $(1, 2)$, que não estava sendo processado, agora tem nível positivo, mas tem que esperar o fim do processamento de $(2, 1)$. Por outro lado, $(2, 2)$ torna-se ativo em m_2 no estado e_2 , uma vez que não há tempos de preparação. No caso de $(1, 1)$ e $(2, 1)$ serem zerados simultaneamente, estão imediatamente $(1, 2)$ e $(2, 2)$ tornam-se ativos.

Assumiremos que os tempos de processamento satisfazem as relações

$$\tau_{11} < \tau_{12}$$

$$\tau_{21} < \tau_{22}$$

Note-se que isso implica em que, se $(1, 1)$ e $(1, 2)$ estão ativos num estado e_j , com $x_{11} > 0$ e $x_{12} > 0$, então a única transição possível é a dada por $T_{11}^j < T_{12}^j$.

Finalmente, suporemos as taxas de entrada externas em $(1, 1)$ e $(2, 1)$ constantes, e assumiremos que no instante inicial o sistema não está completa-

mente vazio, uma vez que a política CAF não é definida na origem. É fácil constatar que, dado um intervalo de tempo *finito*, o sistema nunca atinge a origem novamente durante esse intervalo.

A figura 4.1 mostra o diagrama.

Uma observação deve ser feita, em relação aos estados $e_1, \dots, e_4$: as variáveis ali indicadas com $x_{p,i} \geq 0+$ na realidade representam uma “concatenação” dos estados em que $x_{p,i} = 0+$ e $x_{p,i} > 0+$. Tal procedimento, utilizado para simplificar o diagrama, não implica em perda de informação, uma vez que nesses estados o buffer correspondente (p, i) não está ativo e, se $x_{p,i}(t) = 0$ com $\dot{x}_{p,i}(t) > 0$, então num instante $t+h$ arbitrariamente próximo de t tem-se $x_{p,i}(t+h) > 0$ e $\dot{x}_{p,i}(t+h) > 0$.

Uma análise cuidadosa do grafo de estados mostra que o conjunto $X = \{e_7, e_8, e_9, e_{10}\}$ é tal que *não existem arcos saindo de X*, isto é, dado um arco (a, b) , se $a \in X$ então $b \in X$. Em outras palavras, o conjunto de estados X é um *invariante*, ou seja, uma vez que o sistema atinge um desses estados permanece para sempre neles. Veremos adiante que, dependendo das condições iniciais, podem existir dois invariantes.

Nossa tarefa se dividirá em duas partes: analisar as condições em que o invariante X não é atingido, bem como o conseqüente comportamento do sistema, e estudar o caso em que o invariante X é alcançado. A partir daí, obteremos condições necessárias e suficientes de estabilidade para o sistema em questão.

Condição de existência de dois invariantes

Observando o diagrama, vemos que todos os arcos, exceto o arco (e_4, e_1) , são orientados no sentido do invariante X ; ou seja, o sistema só não atinge X se, toda vez em que e_4 for atingido, valer a transição $T_{12}^4 = T_{22}^4$; nesse caso, o sistema passa por e_1 com $x_{12} = x_{22} = 0$.

Suponhamos então o sistema no estado e_1 num instante t_0 . Temos

$$x_{11}(t_0) = A > 0$$

Figura 4.1: Grafo de estados do sistema

$$\begin{aligned}
x_{12}(t_0) &= 0 \\
x_{21}(t_0) &= C > 0 \\
x_{22}(t_0) &= 0
\end{aligned}$$

sendo que (1, 1) e (2, 1) estão ativos. Podemos calcular T_{11}^1 e T_{21}^1 :

$$\begin{aligned}
T_{11}^1 &= \frac{x_{11}(t_0)\tau_{11}}{1 - \rho_{11}} = \frac{A\tau_{11}}{1 - \rho_{11}} \\
T_{21}^1 &= \frac{x_{21}(t_0)\tau_{21}}{1 - \rho_{21}} = \frac{C\tau_{21}}{1 - \rho_{21}}
\end{aligned}$$

Consideremos as 3 possibilidades:

1. $T_{11}^1 < T_{21}^1$: Neste caso, em $t_1 = t_0 + T_{11}^1$ o sistema está em e_2 :

$$\begin{aligned}
x_{11}(t_1) &= 0 \\
x_{12}(t_1) &= \frac{1}{\tau_{11}}T_{11}^1 = \frac{A}{1 - \rho_{11}} \\
x_{21}(t_1) &= x_{21}(t_0) + \left(d_2 - \frac{1}{\tau_{21}}\right)T_{11}^1 = C - \frac{(1 - \rho_{21})A\tau_{11}}{\tau_{21}(1 - \rho_{11})} \\
x_{22}(t_1) &= \frac{1}{\tau_{21}}T_{11}^1 = \frac{A\tau_{11}}{\tau_{21}(1 - \rho_{11})}
\end{aligned}$$

A próxima transição é obrigatória para e_4 e ocorre no instante $t_2 = t_1 + T_{21}^2$. Como

$$T_{21}^2 = \frac{x_{21}(t_1)\tau_{21}}{1 - \rho_{21}} = \frac{C\tau_{21}}{1 - \rho_{21}} - \frac{A\tau_{11}}{1 - \rho_{11}}$$

temos então

$$\begin{aligned}
x_{11}(t_2) &= d_1T_{21}^2 = d_1 \left(\frac{C\tau_{21}}{1 - \rho_{21}} - \frac{A\tau_{11}}{1 - \rho_{11}} \right) \\
x_{12}(t_2) &= x_{12}(t_1) = \frac{A}{1 - \rho_{11}} \\
x_{21}(t_2) &= 0
\end{aligned}$$

$$\begin{aligned}
x_{22}(t_2) &= x_{22}(t_1) + \frac{1}{\tau_{21}}T_{21}^2 - \frac{1}{\tau_{22}}T_{21}^2 = \\
&= \frac{A\tau_{11}}{\tau_{21}(1-\rho_{11})} + \left(\frac{1}{\tau_{21}} - \frac{1}{\tau_{22}}\right) \left(\frac{C\tau_{21}}{1-\rho_{21}} - \frac{A\tau_{11}}{1-\rho_{11}}\right) \\
&= \frac{C}{1-\rho_{21}} - \frac{1}{\tau_{22}} \left(\frac{C\tau_{21}}{1-\rho_{21}} - \frac{A\tau_{11}}{1-\rho_{11}}\right)
\end{aligned}$$

Para que ocorra a transição de volta para e_1 , é preciso que $T_{12}^4 = T_{12}^4$, ou seja, $x_{12}(t_2)\tau_{12} = x_{22}(t_2)\tau_{22}$. Logo, devemos ter

$$\frac{A\tau_{12}}{1-\rho_{11}} = \frac{C\tau_{22}}{1-\rho_{21}} - \frac{C\tau_{21}}{1-\rho_{21}} + \frac{A\tau_{11}}{1-\rho_{11}} \quad (4.1)$$

e aí em $t_3 = t_2 + T_{12}^4 = t_2 + T_{22}^4$ temos

$$\begin{aligned}
x_{11}(t_3) &= x_{11}(t_2) + d_1T_{22}^4 = \\
&= d_1 \left(\frac{C\tau_{21}}{1-\rho_{21}} - \frac{A\tau_{11}}{1-\rho_{11}}\right) + d_1 \left(\frac{C\tau_{22}}{1-\rho_{21}} - \frac{C\tau_{21}}{1-\rho_{21}} + \frac{A\tau_{11}}{1-\rho_{11}}\right) \\
&= \frac{d_1C\tau_{22}}{1-\rho_{21}} = A' \\
x_{12}(t_3) &= 0 \\
x_{21}(t_3) &= d_2T_{12}^4 = \frac{d_2A\tau_{12}}{1-\rho_{11}} = C' \\
x_{22}(t_3) &= 0
\end{aligned}$$

2. $T_{11}^1 > T_{21}^1$: Neste caso, o desenvolvimento é análogo ao caso 1, pois o sistema é simétrico, logo a condição $T_{12}^4 = T_{12}^4$ fica

$$\frac{C\tau_{22}}{1-\rho_{21}} = \frac{A\tau_{12}}{1-\rho_{11}} - \frac{A\tau_{11}}{1-\rho_{11}} + \frac{C\tau_{21}}{1-\rho_{21}} \quad (4.2)$$

que é a mesma equação de (4.1), e temos em e_1

$$\begin{aligned}
x_{11} &= \frac{d_1C\tau_{22}}{1-\rho_{21}} = A' \\
x_{21} &= \frac{d_2A\tau_{12}}{1-\rho_{11}} = C' \\
x_{12} &= x_{22} = 0
\end{aligned}$$

3. Finalmente, se $T_{11}^1 = T_{21}^1$, ou seja, se

$$\frac{A\tau_{11}}{1 - \rho_{11}} = \frac{C\tau_{21}}{1 - \rho_{21}} \quad (4.3)$$

então temos, em e_4 ,

$$\begin{aligned} x_{11} &= x_{21} = 0 \\ x_{12} &= \frac{A}{1 - \rho_{11}} \\ x_{22} &= \frac{C}{1 - \rho_{21}} \end{aligned}$$

e a condição $T_{12}^4 = T_{12}^4$ fica

$$\frac{A\tau_{12}}{1 - \rho_{11}} = \frac{C\tau_{22}}{1 - \rho_{21}}$$

que, juntamente com (4.3), implica na condição (4.1).

Assim, a condição necessária e suficiente para que ocorra a volta ao estado e_1 (e a conseqüente fuga do invariante) é a equação (4.1), que pode ser re-escrita como

$$\frac{A}{C} = \frac{1 - \rho_{11}}{1 - \rho_{21}} \frac{\tau_{22} - \tau_{21}}{\tau_{12} - \tau_{11}} \quad (4.4)$$

e, na volta a e_1 , tem-se

$$\begin{aligned} x_{11} &= \frac{d_1 C \tau_{22}}{1 - \rho_{21}} = A' \\ x_{21} &= \frac{d_2 A \tau_{12}}{1 - \rho_{11}} = C' \\ x_{12} &= x_{22} = 0 \end{aligned}$$

Repetindo o raciocínio para os valores A' e C' , concluímos que, se o sistema volta novamente para e_1 , isto é, se vale a condição

$$\frac{A'}{C'} = \frac{1 - \rho_{11}}{1 - \rho_{21}} \frac{\tau_{22} - \tau_{21}}{\tau_{12} - \tau_{11}} \quad (4.5)$$

então na próxima volta a e_1 teremos

$$\begin{aligned}
 x_{11}(t_8) &= \frac{d_1 C' \tau_{22}}{1 - \rho_{21}} = \frac{A d_1 d_2 \tau_{12} \tau_{22}}{(1 - \rho_{11})(1 - \rho_{21})} \\
 &= A \frac{\rho_{12}}{1 - \rho_{21}} \frac{\rho_{22}}{1 - \rho_{11}} \\
 x_{21}(t_8) &= \frac{d_2 A' \tau_{12}}{1 - \rho_{11}} = \frac{C d_2 d_1 \tau_{22} \tau_{12}}{(1 - \rho_{21})(1 - \rho_{11})} \\
 &= C \frac{\rho_{12}}{1 - \rho_{21}} \frac{\rho_{22}}{1 - \rho_{11}} \\
 x_{12}(t_8) &= x_{22}(t_8) = 0
 \end{aligned}$$

Substituindo os valores de A' e C' em (4.5), concluímos que a condição vale se e só se

$$\frac{C}{A} \frac{d_1 \tau_{22}(1 - \rho_{11})}{d_2 \tau_{12}(1 - \rho_{21})} = \frac{1 - \rho_{11}}{1 - \rho_{21}} \frac{\tau_{22} - \tau_{21}}{\tau_{12} - \tau_{11}}$$

Trazendo agora de (4.4) o valor de $\frac{C}{A}$, a segunda condição fica

$$\frac{1 - \rho_{21}}{1 - \rho_{11}} \frac{\tau_{12} - \tau_{11}}{\tau_{22} - \tau_{21}} \frac{d_1 \tau_{22}(1 - \rho_{11})}{d_2 \tau_{12}(1 - \rho_{21})} = \frac{1 - \rho_{11}}{1 - \rho_{21}} \frac{\tau_{22} - \tau_{21}}{\tau_{12} - \tau_{11}}$$

ou seja,

$$\frac{d_1 \tau_{22}(1 - \rho_{21})}{d_2 \tau_{12}(1 - \rho_{11})} = \left(\frac{\tau_{22} - \tau_{21}}{\tau_{12} - \tau_{11}} \right)^2 \quad (4.6)$$

Notemos agora que, na segunda volta a e_1 , temos

$$\mathbf{x}(t_8) = \frac{\rho_{12}}{1 - \rho_{21}} \frac{\rho_{22}}{1 - \rho_{11}} \mathbf{x}(t_0)$$

Ora, se o sistema satisfaz a condição de capacidade

$$\begin{aligned}
 \rho_{12} + \rho_{21} &< 1 \\
 \rho_{22} + \rho_{11} &< 1
 \end{aligned}$$

então temos $\mathbf{x}(t_8) = \varphi \mathbf{x}(t_0)$ com $\varphi < 1$, e portanto o sistema é estável.

Assim, as condições (4.4) e (4.6) são necessárias e suficientes para que o sistema não atinja o invariante X , e conseqüentemente seja estável.

Atingindo o invariante X

Vamos analisar agora o caso em que o invariante X é atingido, o que ocorre, como vimos, se não vale (4.4) ou (4.6).

Suponhamos que no instante t_0 o sistema está no estado e_9 :

$$\begin{aligned}x_{11}(t_0) &= A > 0 \\x_{12}(t_0) &= 0 \\x_{21}(t_0) &= 0 \\x_{22}(t_0) &= 0\end{aligned}$$

Em $t_1 = t_0 + T_{11}^9 = t_0 + \frac{A\tau_{11}}{1-\rho_{11}}$ o sistema está em e_8 , e temos

$$\begin{aligned}x_{11}(t_1) &= 0 \\x_{12}(t_1) &= \frac{1}{\tau_{11}}T_{11}^9 - \frac{1}{\tau_{12}}T_{11}^9 = \frac{A(\tau_{12} - \tau_{11})}{\tau_{12}(1 - \rho_{11})} \\x_{21}(t_1) &= d_2T_{11}^9 = \frac{d_2A\tau_{11}}{1 - \rho_{11}} \\x_{22}(t_1) &= 0\end{aligned}$$

A próxima transição é obrigatória para e_{10} e ocorre no instante $t_2 = t_1 + T_{12}^8$. Como m_1 está sub-alimentada, então

$$T_{12}^8 = \frac{x_{12}(t_1)\tau_{12}}{1 - \rho_{12}} = \frac{A(\tau_{12} - \tau_{11})}{(1 - \rho_{11})(1 - \rho_{12})}$$

Logo,

$$\begin{aligned}x_{11}(t_2) &= 0 \\x_{12}(t_2) &= 0 \\x_{21}(t_2) &= x_{21}(t_1) + d_2T_{12}^8 = \frac{d_2A\tau_{11}}{1 - \rho_{11}} + \frac{d_2A(\tau_{12} - \tau_{11})}{(1 - \rho_{11})(1 - \rho_{12})} \\&= \frac{d_2A\tau_{12}}{1 - \rho_{12}} = C \\x_{22}(t_2) &= 0\end{aligned}$$

O comportamento do sistema de e_{10} para e_7 e de e_7 para e_9 é análogo ao calculado acima, logo em t_3 o estado e_9 é novamente atingido, e temos

$$\begin{aligned}x_{11}(t_3) &= \frac{d_1 C \tau_{22}}{1 - \rho_{22}} = \frac{d_1 d_2 A \tau_{12} \tau_{22}}{(1 - \rho_{12})(1 - \rho_{22})} \\&= A \frac{\rho_{22}}{1 - \rho_{12}} \frac{\rho_{12}}{1 - \rho_{22}} \\x_{12}(t_3) &= 0 \\x_{21}(t_3) &= 0 \\x_{22}(t_3) &= 0\end{aligned}$$

Portanto,

$$\mathbf{x}(t_4) = \varphi \mathbf{x}(t_0), \quad \text{onde } \varphi = \frac{\rho_{22}}{1 - \rho_{12}} \frac{\rho_{12}}{1 - \rho_{22}}$$

É fácil ver que

$$\varphi \leq 1 \iff \rho_{12} + \rho_{22} \leq 1$$

Logo,

$$\rho_{12} + \rho_{22} \leq 1 \tag{4.7}$$

é uma condição necessária e suficiente para que o sistema, desde que atinja o invariante, seja estável.

Vamos resumir os resultados acima na seguinte Proposição:

Proposição 4.3 *Para o sistema em questão, consideremos as condições:*

i) *Para todo instante t_k tal que $x_{11}(t_k) > 0$, $x_{21}(t_k) > 0$ e $x_{12}(t_k) = x_{22}(t_k) = 0$, tem-se*

$$\frac{x_{11}(t_k)}{x_{21}(t_k)} = \frac{1 - \rho_{11}}{1 - \rho_{21}} \frac{\tau_{22} - \tau_{21}}{\tau_{12} - \tau_{11}}$$

$$\text{ii) } \frac{d_1 \tau_{22}(1 - \rho_{21})}{d_2 \tau_{12}(1 - \rho_{11})} = \left(\frac{\tau_{22} - \tau_{21}}{\tau_{12} - \tau_{11}} \right)^2$$

$$\text{iii) } \rho_{12} + \rho_{22} \leq 1$$

Então, o sistema é estável para qualquer política “clearing” e sob qualquer condição inicial se e somente se valem (i) e (ii) ou vale (iii).

É interessante notar que a Proposição acima atrela a estabilidade do sistema às condições iniciais, ou seja, um sistema tipicamente instável (em que $\rho_{12} + \rho_{22} > 1$) pode, sob determinada condição inicial, funcionar de modo estável.

Um exemplo disso pode ser obtido tomando-se $\tau_{11} = \tau_{21}$, $\tau_{12} = \tau_{22}$ e $d_1 = d_2$. Se $\rho_{12} + \rho_{22} > 1$, ou seja, se $\rho_{12} = \rho_{22} > \frac{1}{2}$, então dada uma condição inicial $\mathbf{x}(t_0)$ tal que $x_{12}(t_0) = x_{22}(t_0) = 0$, o sistema é

$$\begin{array}{ll} \text{estável,} & \text{se } x_{11}(t_0) = x_{21}(t_0) \\ \text{instável,} & \text{se } x_{11}(t_0) \neq x_{21}(t_0) \end{array}$$

4.3 Um contra-exemplo à Conjectura

Consideremos o sistema abaixo, formado por duas máquinas e 2 produtos:

com as seguintes características:

- O sistema é simétrico, isto é, $\tau_{1,i} = \tau_{2,i}$, $i = 1, 2, 3$ e $d_1 = d_2$. Omitiremos assim assim o índice do produto, ou seja, escreveremos apenas τ_i e d .
- Existe um tempo de preparação em m_1 correspondente ao chaveamento de (1, 1) para (2, 2), que denotaremos por $\bar{\delta}$. Analogamente, o mesmo

$\bar{\delta}$ denota também o tempo de preparação em m_2 correspondente ao chaveamento de $(2, 1)$ para $(1, 2)$; os demais tempos de preparação serão considerados nulos.

- Valem as seguintes relações:

$$\rho_1 + \rho_2 + \rho_3 < 1 \quad (4.8)$$

$$\tau_1 > \tau_2 \quad (4.9)$$

$$\tau_1 < \tau_2 + \tau_3 \quad (4.10)$$

A relação (4.8) corresponde à *condição de capacidade* do sistema, enquanto que (4.9) é a hipótese básica de tempos decrescentes (note-se que não há restrição quanto à relação entre τ_2 e τ_3). A desigualdade (4.10) assegura a formação de filas em m_2 , o que terá papel fundamental nesta demonstração de instabilidade.

O sistema será operado sob uma política tipo CAF. Relembremos, do Capítulo 1, a definição de tais políticas:

Políticas CAF — Seja T_k o instante da escolha de buffer na máquina m . Escolhe-se para processar um buffer (p^*, i^*) tal que

$$x_{p^*, i^*}(T_k) \geq \varepsilon \sum_{(p, i) \in B_m} x_{p, i}(T_k)$$

onde ε é uma fração fixada. O buffer é processado até que não contenha mais elementos. Se mais de um buffer satisfizer o critério, *escolheremos aquele que está mais próximo do fim do caminho de processamento do produto*, ou seja, escolheremos o buffer (p, i) com maior índice i . A ênfase no critério de desempate deve-se ao fato de que tal regra será essencial para esta prova.

A fim de simplificar a notação, denotaremos por x_i ao invés de $x_{1, i}$ o nível do buffer $(1, i)$, e denotaremos por z_i ao invés de $x_{2, i}$ o nível do buffer $(2, i)$.

Enunciaremos a seguir alguns Lemas, que serão utilizados na demonstração de instabilidade do sistema em questão.

Lema 4.1 *Seja $[t_1, t_2]$ um intervalo tal que $x_3(t_1) = 0$. Se m_2 processa somente os buffers $(1, 2)$ e $(1, 3)$ durante todo o intervalo, sem ficar subalimentada ou ociosa, então para todo t em $[t_1, t_2]$ vale que*

$$x_1(t) + x_2(t) + x_3(t) + w(t_1, t) = x_1(t_1) + x_2(t_1) + x_3(t)\rho_2 + w(t_1, t)(\rho_2 + \rho_3)$$

onde $w(s, t) = y_1(t) - y_1(s)$ = quantidade do produto 1 que deixou o sistema em $[s, t]$.

–Demonstração:

Basta observar que o tempo decorrido entre t_1 e t pode ser expresso em termos da quantidade do produto 1 processada nesse intervalo. Como m_2 trabalhou com taxa nominal durante esse tempo, e dado que $(1, 3)$ estava vazio no início do intervalo, então $w(t_1, t)$ indica a quantidade de elementos que passaram por $(1, 2)$ e $(1, 3)$ durante esse período, consumindo $w(t_1, t)(\tau_2 + \tau_3)$ unidades de tempo. Da mesma forma, $x_3(t)$ indica a quantidade de elementos que passaram por $(1, 2)$ somente, gastando $x_3(t)\tau_2$ unidades de tempo. Durante $[t_1, t]$, a taxa de entrada no sistema é constante igual a d , logo o número de elementos do produto 1 que chegaram ao sistema durante esse período pode ser expresso como

$$e_1(t) - e_1(t_1) = d [w(t_1, t)(\tau_2 + \tau_3) + x_3(t)\tau_2]$$

Ora, como vimos no Capítulo 1, é claro que vale a igualdade

$$\sum_{i=1}^3 x_i(t) = \sum_{i=1}^3 x_i(t_1) + (e_1(t) - e_1(t_1)) - w(t_1, t)$$

Como $x_3(t_1) = 0$, então temos

$$x_1(t) + x_2(t) + x_3(t) + w(t_1, t) = x_1(t_1) + x_2(t_1) + x_3(t)\rho_2 + w(t_1, t)(\rho_2 + \rho_3)$$

■

Lema 4.2 *Com as hipóteses do Lema 4.1, e se a máquina m_1 está subalimentada processando o buffer $(1, 1)$, então a escolha de $(2, 1)$ em m_2 sempre ocorre num instante tal que $(1, 3)$ acabou de ser esvaziado, ou seja, ela nunca ocorre logo após o processamento de $(1, 2)$.*

–*Demonstração:*

Seja T_{2k} um instante tal que $x_2(T_{2k}) > 0$, $x_3(T_{2k}) = 0$. Suponhamos que em T_{2k} temos

$$x_2(T_{2k}) \geq \varepsilon(x_2(T_{2k}) + z_1(T_{2k})) \quad (4.11)$$

logo, pela regra da CAF, $(1, 2)$ é escolhido para o processamento. Consideremos agora o instante T_{2k+1} em que $(1, 2)$ é “zerado”. Denotando por Δ a diferença $T_{2k+1} - T_{2k}$ temos

$$\begin{aligned} x_3(T_{2k+1}) &= x_2(T_{2k}) + d\Delta \\ z_1(T_{2k+1}) &= z_1(T_{2k}) + d\Delta \end{aligned}$$

e assim

$$\begin{aligned} \varepsilon(x_3(T_{2k+1}) + z_1(T_{2k+1})) &= \varepsilon(x_2(T_{2k}) + z_1(T_{2k}) + 2d\Delta) \\ &\stackrel{(4.11)}{\leq} x_2(T_{2k}) + 2\varepsilon d\Delta \\ &\stackrel{\varepsilon < 1/2}{\leq} x_2(T_{2k}) + d\Delta \\ &= x_3(T_{2k+1}) \end{aligned}$$

Logo, $(1, 3)$ é escolhido. ■

Lema 4.3 *Consideremos as hipóteses do Lema 4.1, e suponhamos que o buffer $(1, 2)$ está sendo alimentado por uma taxa constante λ tal que $\lambda\tau_2 < 1$. Vamos denotar por $\{T_{2n}\}_{n \in \mathbb{N}}$ os instantes de início de processamento de $(1, 2)$ e por $\{T_{2n+1}\}_{n \in \mathbb{N}}$ os instantes de início de processamento de $(1, 3)$. Então,*

$$\begin{aligned} x_2(T_{2n+1}) &= 0 & x_3(T_{2n+1}) &= x_2(T_{2n}) \frac{1}{1-\lambda\tau_2} \\ x_2(T_{2n+2}) &= x_2(T_{2n}) \frac{\lambda\tau_3}{1-\lambda\tau_2} & x_3(T_{2n+2}) &= 0 \end{aligned}$$

–*Demonstração:*

Basta observar que $T_{2n+1} - T_{2n}$ é o tempo necessário para esvaziar $(1, 2)$, logo

$$T_{2n+1} - T_{2n} = \frac{x_2(T_{2n})\tau_2}{1 - \lambda\tau_2}$$

e temos $x_3(T_{2n+1}) = \frac{1}{\tau_2}(T_{2n+1} - T_{2n})$. Por outro lado, $T_{2n+2} - T_{2n+1}$ é o tempo necessário para esvaziar $(1, 3)$, logo

$$T_{2n+2} - T_{2n+1} = x_3(T_{2n+1})\tau_3$$

e temos $x_2(T_{2n+2}) = \lambda(T_{2n+2} - T_{2n+1})$. ■

Para completar os lemas preparatórios, vamos enunciar o seguinte

Lema 4.4 *Se, iniciando com $x_1(0) = x_0 > 0$ e os demais buffers vazios, o sistema atinge em algum instante T a configuração $z_1(T) = cx_1(0)$ e os demais buffers vazios, então para estas condições iniciais $(x(0), z(0))$ o sistema tem*

$$\begin{array}{ll} \text{buffers limitados} & \text{se } c \leq 1 \\ \text{buffers ilimitados} & \text{se } c > 1 \end{array}$$

–*Demonstração:*

Trivial, dada a simetria do sistema. ■

Utilizando os lemas acima, mostraremos então que, com uma restrição razoável nos valores de ρ_i , o sistema torna-se instável para a política CAF considerada, com uma escolha conveniente do valor de ε e as condições iniciais do Lema 4.4.

Antes de formalizarmos a prova de instabilidade, alguns comentários devem ser feitos, a fim de possibilitar uma melhor compreensão daquela demonstração. Inicialmente, consideremos a seguinte situação inicial: $x_1(0) = x_0 + \delta > 0$ (onde $\delta = \frac{1}{\tau_1}\bar{\delta}$) e os demais buffers vazios, com m_2 preparada para o processamento de $(2, 1)$. Seja $T_0 = \bar{\delta}$ o instante após o chaveamento de $(2, 1)$ para $(1, 2)$. Forçando a taxa de entrada no sistema para zero durante esse período (o que mantém a restrição sobre taxas de entrada $e_p(t) - e_p(s) \leq d(t - s)$) obtemos

$$x_1(T_0) = x_0 \quad x_2(T_0) = \delta \quad x_3(T_0) = z_i(T_0) = 0$$

Figura 4.2: Comportamento dos buffers (1,2) e (1,3). Note-se a mudança na taxa de entrada de (1,2) no instante T_k .

A partir de T_0 , a taxa de entrada instantânea em (1,2) passa a ser $d' = \frac{1}{\tau_1}$. Ora, a hipótese (4.9) de tempos não-crescentes garante que $d'\tau_2 < 1$, logo o comportamento dos buffers (1,2) e (1,3) é ditado pelo Lema 4.3, ou seja, temos:

$$\begin{aligned} x_2(T_{2n}) &= \delta \left(\frac{d'\tau_3}{1-d'\tau_2} \right)^n & x_3(T_{2n}) &= 0 \\ x_2(T_{2n+1}) &= 0 & x_3(T_{2n+1}) &= \delta \left(\frac{d'\tau_3}{1-d'\tau_2} \right)^n \frac{1}{1-d'\tau_2} \end{aligned}$$

Mas $d'(\tau_2 + \tau_3) > 1$ de acordo com a hipótese (4.10), logo os valores de *pico* de $x_2(\cdot)$ (que ocorrem nos instantes $\{T_{2n}\}_{n \in \mathbb{N}}$) e de $x_3(\cdot)$ (que ocorrem nos instantes $\{T_{2n+1}\}_{n \in \mathbb{N}}$) crescem, até o instante T_k em que (1,1) é esvaziado. A partir daí, a máquina m_1 fica sub-alimentada processando o buffer (1,1), ou seja, a taxa de entrada instantânea em (1,2) passa a ser d . Agora, $d(\tau_2 + \tau_3) < 1$, portanto os valores de pico de x_2 e x_3 decrescem, até o instante T_l em que (2,1) é escolhido pela regra da política CAF. A figura 4.2 mostra o comportamento desses buffers.

Assim, em $[T_0, T_l]$ podemos aplicar o Lema 4.1, e obtemos

$$x_1(t) + x_2(t) + x_3(t) + w(T_0, t) = x_0 + \delta + x_3(t)\rho_2 + w(T_0, t)(\rho_2 + \rho_3)$$

onde $w(T_0, t)$ é definido como no Lema 4.1. A partir de T_k temos $x_1(t) = 0$, logo da equação acima concluímos que

$$w(T_0, t) = \frac{x_0 + \delta - x_2(t) - x_3(t)(1 - \rho_2)}{1 - \rho_2 - \rho_3} \quad \forall t \in [T_k, T_l] \quad (4.12)$$

Por outro lado, como $(2, 1)$ não é processado em $[T_0, T_l]$, então $z_1(t) = d(t - T_0)$, ou seja, usando novamente o Lema 4.1 temos

$$z_1(t) = x_3(t)\rho_2 + w(T_0, t)(\rho_2 + \rho_3)$$

e então, utilizando (4.12),

$$z_1(t) = \frac{(x_0 + \delta)(\rho_2 + \rho_3)}{1 - \rho_2 - \rho_3} - \frac{x_2(t)(\rho_2 + \rho_3)}{1 - \rho_2 - \rho_3} - \frac{x_3(t)\rho_3}{1 - \rho_2 - \rho_3} \quad \forall t \in [T_k, T_l] \quad (4.13)$$

Além disso, pelo Lema 4.2 podemos concluir que o instante T_l é tal que $x_2(T_l) > 0$, $x_3(T_l) = 0$. Denotando por Γ o primeiro termo do lado direito de (4.13), e considerando que

$$x_2(T_l) < \varepsilon(x_2(T_l) + z_1(T_l))$$

podemos escrever então

$$x_2(T_l) < \varepsilon \left[x_2(T_l) + \Gamma - \frac{x_2(T_l)(\rho_2 + \rho_3)}{1 - \rho_2 - \rho_3} \right]$$

Resumindo,

$$x_2(T_l) < \varepsilon(x_2(T_l) + z_1(T_l)) \iff x_2(T_l) < \frac{\varepsilon\Gamma}{(1 - \varepsilon) + (1 - 2\varepsilon)(\rho_2 + \rho_3)} \quad (4.14)$$

Substituindo (4.14) em (4.13) obtemos então

$$z_1(T_l) > \Gamma - \frac{\varepsilon\Gamma}{(1 - \varepsilon) + (1 - 2\varepsilon)(\rho_2 + \rho_3)} \frac{\rho_2 + \rho_3}{1 - \rho_2 - \rho_3}$$

Podemos agora enunciar a instabilidade do sistema:

Teorema 4.1 *O sistema considerado tem buffers ilimitados se valem as condições abaixo:*

- i) *As condições iniciais são: $x_1(0) = x_0 + \delta > 0$ e os demais buffers vazios;*
- ii) *A política CAF é aplicada com uma fração ε' , onde*

$$\varepsilon' \leq \frac{\varepsilon \rho_3 (1 + \rho_2 + \rho_3)}{(1 - \rho_2)(1 + \rho_2 + \rho_3) - \varepsilon \rho_3 (1 + 2\rho_2 + 2\rho_3)(1 - \rho_2 - \rho_3)}$$

$$\varepsilon \leq \min \left\{ \frac{1}{2}, \frac{[(1 + c)\rho - c][1 - \rho^2]}{\rho^2 + (1 + \rho - 2\rho^2)[(1 + c)\rho - c]} \right\}$$

(onde $c > 1$ e $\rho = \rho_2 + \rho_3$);

- iii) *Os eventuais empates na escolha do buffer a processar são resolvidos tomando-se o buffer com maior índice i ;*
- iv) *$2(\rho_2 + \rho_3) > 1$.*

–Demonstração:

Inicialmente, vamos mostrar que $z_1(T_l) \geq cx_1(0)$ para algum $c > 1$, onde T_l é, como definido acima, o instante em que $(2, 1)$ é escolhido em m_2 . Notemos que

$$\begin{aligned} z_1(T_l) &\geq cx_1(0) \stackrel{(4.13)}{\iff} \\ \stackrel{(4.13)}{\iff} (x_0 + \delta - x_2(T_l)) \frac{\rho_2 + \rho_3}{1 - \rho_2 - \rho_3} &\geq c(x_0 + \delta) \iff \\ \iff x_2(T_l) &\leq (x_0 + \delta) \left[1 - \frac{c(1 - \rho_2 - \rho_3)}{\rho_2 + \rho_3} \right] \end{aligned}$$

Ou seja,

$$z_1(T_l) \geq cx_1(0) \iff x_2(T_l) \leq (x_0 + \delta) \frac{(1 + c)(\rho_2 + \rho_3) - c}{\rho_2 + \rho_3}$$

Como (4.14) vale, é suficiente impor que

$$\frac{\varepsilon}{(1 - \varepsilon) + (1 - 2\varepsilon)(\rho_2 + \rho_3)} \frac{(x_0 + \delta)(\rho_2 + \rho_3)}{1 - \rho_2 - \rho_3} \leq (x_0 + \delta) \frac{(1 + c)(\rho_2 + \rho_3) - c}{\rho_2 + \rho_3}$$

Para simplificar a notação, vamos denotar por ρ a expressão $\rho_2 + \rho_3$. Assim, a desigualdade acima vale se e somente se

$$\begin{aligned} \varepsilon \rho^2 &\leq [(1+c)\rho - c][(1-\varepsilon) + (1-2\varepsilon)\rho][1-\rho] \iff \\ \iff \varepsilon \rho^2 &\leq [(1+c)\rho - c][(1+\rho) - \varepsilon(1+2\rho)][1-\rho] \iff \\ \iff \varepsilon [\rho^2 + (1+2\rho)(1-\rho)((1+c)\rho - c)] &\leq [(1+c)\rho - c][1-\rho^2] \end{aligned}$$

Se $(1+c)\rho - c > 0$, isto é, se

$$c < \frac{\rho}{1-\rho}$$

(sempre existe tal c , pois pela hipótese (iv) do Teorema temos $2(\rho_2 + \rho_3) > 1$ e isto implica que $\frac{\rho}{1-\rho} > 1$), então todos os termos da desigualdade acima são positivos. Logo, tomando

$$\varepsilon \leq \frac{[(1+c)\rho - c][1-\rho^2]}{\rho^2 + (1+\rho-2\rho^2)((1+c)\rho - c)} \quad (4.15)$$

temos $z_1(T_l) \geq cx_1(0)$.

Observemos agora que, se $(1, 2)$ fosse processado em T_l ao invés de $(2, 1)$, e se forçássemos a taxa de entrada no sistema para zero de T_l até o instante T_m em que $x_2(T_m) = x_3(T_m) = 0$, então teríamos um “semi-ciclo” de instabilidade, como no Lema 4.4, uma vez que, nesse caso,

$$\begin{aligned} z_1(T_m) &= z_1(T_l) \geq cx_1(0) \quad \text{com } c > 1 \\ z_i(T_m) &= 0 \quad i = 1, 2 \\ x_i(T_m) &= 0 \quad i = 1, 2, 3 \end{aligned} \quad (4.16)$$

Portanto, para completar a prova precisamos apenas mostrar que existe um $\varepsilon' < \varepsilon$ tal que $(2, 1)$ não é escolhido pela regra da CAF em T_l , ou seja,

$$\exists 0 < \varphi < 1 : \quad x_2(T_l) \geq \varphi \varepsilon (x_2(T_l) + z_1(T_l)) \quad (4.17)$$

Consideremos o instante T_{l-2} em que $x_2(T_{l-2}) > 0$, $x_3(T_{l-2}) = 0$ imediatamente antes de T_l , isto é,

$$T_l = T_{l-2} + \frac{x_2(T_{l-2})}{1-\rho_2}(\tau_2 + \tau_3)$$

pelo Lema (4.3) temos então

$$x_2(T_l) = x_2(T_{l-2}) \frac{\rho_3}{1 - \rho_2} \quad (4.18)$$

Pela definição de T_l temos

$$x_2(T_{l-2}) \geq \varepsilon(x_2(T_{l-2}) + z_1(T_{l-2}))$$

e então de (4.14) vem que

$$x_2(T_{l-2}) \geq \frac{\varepsilon\Gamma}{(1 - \varepsilon) + (1 - 2\varepsilon)(\rho_2 + \rho_3)} \quad (4.19)$$

A fim de assegurar que (4.17) vale, vamos impor de (4.18) e (4.19) que

$$\frac{\varepsilon\Gamma}{(1 - \varepsilon) + (1 - 2\varepsilon)(\rho_2 + \rho_3)} \frac{\rho_3}{1 - \rho_2} \geq \frac{\varphi\varepsilon\Gamma}{(1 - \varphi\varepsilon) + (1 - 2\varphi\varepsilon)(\rho_2 + \rho_3)} \quad (4.20)$$

Mas essa desigualdade vale se e só se tivermos

$$\begin{aligned} [(1 - \varphi\varepsilon) + (1 - 2\varphi\varepsilon)\rho]\rho_3 &\geq \varphi(1 - \rho_2)[(1 - \varepsilon) + (1 - 2\varepsilon)\rho] \iff \\ \iff \rho_3 - \varphi\varepsilon\rho_3 + \rho_3\rho - 2\varphi\varepsilon\rho_3\rho &\geq \varphi(1 - \rho_2)[(1 - \varepsilon) + (1 - 2\varepsilon)\rho] \iff \\ \iff \rho_3(1 + \rho_2 + \rho_3) &\geq \varphi[(1 - \rho_2)(1 - \varepsilon) + (1 - \rho_2)(1 - 2\varepsilon)\rho + \varepsilon\rho_3 + 2\varepsilon\rho_3\rho] \end{aligned}$$

Como todos os termos da desigualdade acima são positivos, então concluímos que (4.20) vale se e só se

$$\varphi \leq \frac{\rho_3(1 + \rho_2 + \rho_3)}{(1 - \rho_2)(1 + \rho_2 + \rho_3) - \varepsilon\rho_3(1 + 2\rho_2 + 2\rho_3)(1 - \rho_2 - \rho_3)} \quad (4.21)$$

Mas

$$\varepsilon \leq \frac{1}{2} < \frac{1 + \rho_2 + \rho_3}{1 + 2(\rho_2 + \rho_3)} < \frac{1 + \rho_2 + \rho_3}{1 + 2(\rho_2 + \rho_3)} \frac{1}{\rho_3}$$

logo

$$\varepsilon \rho_3(1 + 2\rho_2 + 2\rho_3)(1 - \rho_2 - \rho_3) < (1 + \rho_2 + \rho_3)(1 - \rho_2 - \rho_3)$$

e então

$$\rho_3(1 + \rho_2 + \rho_3) < (1 - \rho_2)(1 + \rho_2 + \rho_3) - \varepsilon \rho_3(1 + 2\rho_2 + 2\rho_3)(1 - \rho_2 - \rho_3)$$

Portanto, o lado direito de (4.21) é menor que 1, e conseqüentemente (4.17) vale. Assim, a condição (4.16) é satisfeita e, aplicando o Lema 4.4, concluimos que o sistema tem buffers ilimitados se operado da maneira descrita, e portanto é instável, o que completa a prova. ■

O Teorema acima mostra que a Conjectura dos tempos não-crescentes é, portanto, falsa, no caso de tempos estritamente decrescentes. Porém, com a mesma topologia de sistema é possível mostrar também que a instabilidade pode ser obtida com $\tau_1 = \tau_2$, o que não faremos aqui para não sermos repetitivos.

Permanece em aberto, assim, uma reformulação da Conjectura que tenha as mesmas propriedades daquela, ou seja, que exiba condições suficientes de estabilidade para qualquer política usável baseadas apenas nos tempos de processamento dos produtos, sem utilizar a informação global d_p . Uma alternativa possível, que podemos intuir a partir dos exemplos e do contra-exemplo apresentados, seria considerar a relação de ordem não entre um buffer e seu sucessor, mas sim entre um buffer e os eventuais laços consecutivos do seu sucessor. Formalmente, teríamos:

Conjectura 2 *Seja $\mathcal{B} = \{(p, i) : (p, i + 1) \notin B_{\mu_{p,i}}\}$. Para todo $(p, i) \in \mathcal{B}$, seja $k(i)$ o menor índice tal que $k(i) > i$ e $(p, k(i)) \in \mathcal{B}$. Então, se*

$$\tau_{p,i} \geq \tau_{p,i+1} + \dots + \tau_{p,k(i)} \quad \forall (p, i) \in \mathcal{B}$$

o sistema é estável para qualquer política usável.

Conclusão

Estudamos neste trabalho o problema da estabilidade em Sistemas Flexíveis de Manufatura sob diversos aspectos. Vimos algumas condições sobre a topologia e sobre os tempos de processamento que, se satisfeitas, asseguram a estabilidade de um sistema para qualquer política *usável* ou *fortemente usável*. Por outro lado, analisamos alguns mecanismos que, adaptados a políticas, garantem a estabilidade de sistemas de qualquer natureza.

Quanto ao primeiro aspecto (condições de estabilidade), devemos ressaltar as condições enfraquecidas de estabilidade condicional, apresentadas por Humes [6], que estudamos no Capítulo 2. Tais propriedades, como vimos, constituem-se nas mais fracas condições *intrínsecas* conhecidas para sistemas gerais. No caso particular de sistemas com duas máquinas, sem tempos de preparação, é possível encontrar, como o fizemos no Capítulo 3, um conjunto de condições suficientes de estabilidade disjunto do de condições derivadas da estabilidade condicional, o que se constitui numa ferramenta adicional quando da análise de sistemas desse tipo.

Ainda com relação ao estudo de condições de estabilidade, analisamos também *topologias* estáveis, ou seja, topologias em que, quaisquer que sejam os tempos de processamento (sujeitos somente à condição de capacidade, é claro), a estabilidade é assegurada, para uma classe de políticas. Como

exemplo temos a topologia acíclica vista no Capítulo 1 e a topologia “ciclo-simples”, que apresentamos e demonstramos no Capítulo 2.

Com relação ao segundo aspecto (mecanismos estabilizadores), destaque-se o Mecanismo Regulador proposto por Humes [6] que, como comentamos no Capítulo 2, generaliza as políticas “backoff” [16] e USSM [11] estudadas anteriormente. Vimos, contudo, que o problema de encontrar um limitante superior para o número de reguladores necessários é $\mathcal{NP}$ -Completo; de qualquer modo, uma implementação com poucos reguladores aumenta a dependência das taxas globais d_p , o que pode não ser desejável. Nesse sentido, a abordagem de Kumar e Lu [10], embora num contexto sem tempos de preparação, é ideal, uma vez que as políticas ali apresentadas dependem apenas da prioridade do buffer ou de cada componente do produto em si.

Com isso em mente, concluímos que a união dos dois aspectos mencionados acima ocorre quando pensamos em mecanismos reguladores independentes das taxas globais, uma vez que, como discutimos, tal meta está estreitamente relacionada à obtenção de condições suficientes de estabilidade. Por essa razão, seria de fundamental importância a comprovação da Conjectura de Humes e Kumar sobre tempos não-crescentes. Nosso contra-exemplo apresentado no Capítulo 4, no entanto, demonstrou a falsidade da Conjectura, embora permitindo uma reformulação do enunciado original, como propusemos no final daquele capítulo, questão esta que permanece em aberto.

O assunto estudado é relativamente novo, e dessa forma muitas questões permanecem irrespondidas. A obtenção de condições suficientes de estabilidade mais fracas, por exemplo, é um desses pontos, bem como a demonstração da estabilidade de outras políticas; em [8], por exemplo, mostra-se a estabilidade de políticas tipo “open-loop” para 1 máquina isolada. A prova da estabilidade da política FCFS (primeiro-a-entrar-primeiro-a-sair) também permanece em aberto, mesmo para sistemas sem tempos de preparação.

Outro ponto importante é a derivação de condições também *necessárias* de estabilidade. Exibimos no Capítulo 4 um exemplo para um caso extremamente particular, mas o fizemos através da análise do *grafo de estados* do sistema que, como comentamos, torna-se inviável para sistemas gerais. Nesse sentido, uma abordagem que nos parece promissora fundamenta-se nos

resultados da Teoria de Controle e Teoria de Sistemas Lineares com retardo, seguindo a linha de modelar o problema através de equações diferenciais, como desenvolvemos no Capítulo 3.

Finalmente, podemos pensar em *extensões* do modelo proposto por Perkins e Kumar, como por exemplo a utilização de *tecnologias alternativas* de produção, ou seja, caminhos de processamento opcionais; também se poderia pensar em introduzir uma abordagem estocástica ao problema, considerando taxas não mais fixas, mas regidas por uma distribuição de probabilidades, ou ainda supor máquinas sujeitas a falhas.

Bibliografia

- [1] C.Chase e P.J.Ramadge
On Real-Time Scheduling Policies for Flexible Manufacturing Systems
IEEE Transactions on Automatic Control, AC-37 (4):491-496, Apr. 1992

- [2] M.R.Garey e D.S.Johnson
Computers and Intractability: a Guide to the Theory of NP-Completeness
San Francisco, Freeman, 1981.

- [3] S.C.Graves
A Review of Production Scheduling
Operations Research, Vol. 29, 646-675, Jul.-Aug. 1981

- [4] M.S.Hecht e J.D.Ullman
Flow Graph Reducibility
SIAM J. Computing 1,2 (June 1972), 188-202

- [5] M.S.Hecht e J.D.Ullman
Characterization of Reducible Flow Graphs
J. Assoc. Comput. Mach., 21, No. 3 (1974) 167-175

- [6] C.Humes Jr.
A Regulator Stabilization Technique: Kumar-Seidman Revisited
 Aceito para publicação no IEEE Transactions on Automatic Control

- [7] C.Humes Jr.
Some Results in the Perkins-Kumar Framework for Distributed Activity Scheduling
 In Pirooz Vakili, editor, NSF - Workshop on Hierarchical Control for Real-Time Scheduling of Manufacturing Systems, pag. A.C.1-8, Lincoln, New Hampshire, outubro 1992

- [8] C.Humes Jr. e A.F.P.C.Humes
On the Stability of Clearing Open-Loop Policies
 RT-MAC 9203, Instituto de Matemática e Estatística, USP, 1992

- [9] L.Kleinrock
Queueing Systems, vol. 2
 John Wiley and Sons, 1976

- [10] P.R.Kumar e C.H.Lu
Distributed scheduling based on due dates and buffer prioritization
 A ser publicado no IEEE Transactions on Automatic Control

- [11] P.R.Kumar e T.I.Seidman
Dynamic instabilities and stabilization methods in distributed real-time scheduling of manufacturing systems
 IEEE Transactions on Automatic Control, AC-35 (3):289-298, Mar. 1990

- [12] N.Rouche, P.Habets e M.Laloy
Stability Theory by Liapunov's Direct Method
 Applied Mathematical Sciences 22
 Springer-Verlag, 1977

- [13] D.G.Luenberger
Introduction to Dynamic Systems
John Wiley and Sons, 1979

- [14] C.L.Lucchesi e D.H.Younger
A minimax theorem for directed graphs
Journal of the London Mathematical Society (2) 17:369-374 (1978)

- [15] S.S.Panwalkar e W.Iskander
A Survey of Scheduling Rules
Operations Research, Vol. 25, 45-61, Jan.-Feb. 1977

- [16] J.R. Perkins e P.R. Kumar
Stable, distributed, real-time scheduling of flexible manufacturing/assembly/disassembly systems
IEEE Transactions on Automatic Control, AC-34 (2):139-148, Feb. 1989

- [17] V.Ramachandran
Finding a Minimum Feedback Arc Set in Reducible Flow Graphs
Journal of Algorithms 9, 299-313 (1988)