Phase Plane Analysis of Congestion Control in Data Center Ethernet Networks

Fengyuan Ren and Wanchun Jiang

Dept. of Computer Science and Technology, Tsinghua University, Beijing, 100084, China
Email: renfy@tsinghua.edu.cn

***Abstract*—Ethernet has some attractive properties for network consolidation in the data center, but needs further enhancement to satisfy the additional requirements of unified network fabrics. Congestion management is introduced in Ethernet networks to avoid dropping packets due to congestion. The BCN (Backward Congestion Notification) mechanism is a basic element of several standard drafts, and its stability underlies normal network operations. Because the linear stability analysis method is incapable of handling the nonlinearity of the variable structure control employed by the BCN mechanism, some particular phenomena are unexposed and the insights are insufficient. In this paper, we propose the concept of strong stability of the queuing system to satisfy the requirements of no dropped packets in the data center, and build a fluid-flow analytical model for the BCN congestion control system. Considering the nonlinearity involved in the rate regulation laws, we classify the system into different categories according to the shapes of phase trajectories, and conduct a nonlinear stability analysis using phase plane analysis techniques on a case by case basis. The analysis details can provide a comprehensive understanding about the behaviors of the overall congestion control system. Finally, we also deduce an explicit stability criterion presenting the parameters constraints for the strongly stable BCN system, which can provide straightforward guidelines for proper parameter settings.**

***Index Terms*—Data Center Ethernet, Backward Congestion Notification, Strong stability and Phase Trajectory.**

I. Introduction

In today's data centers, it is common to deploy an Ethernet network for IP traffic, one or two storage area networks (SANs) for block mode Fibre Channel traffic, and an InfiniBand network for interprocess communication (IPC) traffic. As data centers grow in size and complexity, the effort to manage different interconnect technologies for traffic from multiple applications is becoming resource and cost intensive. Data center networking will become a new paradigm for providing high-bandwidth and low-latency interconnection fabrics to computing and storage applications. The deployment of unified network fabrics in data centers is expected to be a methodical process. It offers many research challenges in the areas of computing, storage, virtualization and networking. It also causes major shifts in the industry, and witnesses the convergence of the different industries of computing, storage and networking.

Ethernet is the most widespread wired local area network technology because of its low cost, wide availability, simplicity, and broad compatibility. However, in the past its bandwidth limitations kept it from being a candidate for unified fabric in data centers. With recent advances in Ethernet speeds and the development of 40 and 100 Gbps, Ethernet has become an attractive choice for unified network fabrics in data centers [1]. However, Ethernet networks being a best effort datagram service still need further enhancement to satisfy additional requirements, such as low latency, high reliability, and no packets loss. In the IEEE standard body, several working groups are addressing these issues to ensure that Ethernet will be equipped to meet data center's requirements. This enhanced Ethernet is called Data Center Ethernet (DCE).

DCE is an architectural collection of Ethernet extensions designed to improve Ethernet networking and management in the data center, which has been well thought out to provide the next-generation infrastructure for data center networks by taking advantage of classical Ethernet's strengths and adding several crucial extensions, such as priority-flow control [**?**], enhanced transmission selection [**?**], shortest path bridging [**?**] and end-to-end congestion management [3].

Avoiding frame drops is mandatory for carrying native storage traffic, since storage traffic does not tolerate frame drops. SCSI is designed with the assumption of running over a reliable transport in which failures are rather rare. Fibre Channel is the primary protocol used to carry storage traffic, and it avoids frame drops through a link flow control mechanism based on credits called buffer-to-buffer flow control. Historically Ethernet has been a lossy network, since Ethernet switches do not use any mechanism to signal to the sender that they are out of buffers [2]. A few years ago, IEEE 802.3 added a PAUSE mechanism to Ethernet, which can be used to stop the sender for a period of time. However it cannot properly alleviate congestion and maintain high throughput because the congestion can roll back from switch to switch, affecting flows that do not contribute to the congestion, but happen to share a link with flows that do. To deal with this problem and be competent in unified network fabrics in data centers, the link flow control is introduced in Ethernet networks to avoid packets dropping due to transient congestions. The IEEE 802.1 standard committee has launched a new task group to develop the end-to-end congestion control mechanism for switched Ethernet networks [3], and four proposals are currently discussed.

As a general congestion management scheme, the BCN mechanism is adopted by several proposals. Its stability is crucial for normal network operation. The theoretical stability

1063-6927/10 $26.00 © 2010 IEEE
DOI 10.1109/ICDCS.2010.33

analysis of BCN will definitely provide essential insights, which are beneficial for confirming a proper congestion control algorithm and accelerating the finalization of the specification. In [4], the inventors of BCN applied the stability criterion in the classical linear control theory to analyze the BCN mechanism and provided the stable sufficient condition satisfied by parameters. The limitation of their work is that the overall congestion control system governed by the BCN mechanism is intentionally divided into two linear subsystems to analyze their stability independently, and the transient behavior of the switching process between two subsystems and its impact on the system stability are neglected, thus some important phenomena and properties are likely unexplored. In this paper, we model BCN behavior as an autonomous second-order system described by a set of nonlinear ordinary differential equations using fluid-flow approximation technique. Remaining the nonlinearity involved in the variable structure rate control of the BCN mechanism, and taking the switching process into account, we conduct a nonlinear stability analysis applying the phase plane method to provide a comprehensive understanding about the behaviors of the overall congestion control system, and deduce an explicit stability criterion.

The remainder of this paper is organized as follows. The background of congestion control in DCE networks and the basic BCN mechanism are introduced in Section II. Subsequently, the analytical model of the BCN congestion control system is built, and the concept of strong stability is introduced. In Section IV, based on various phase trajectories of the BCN system, the stability is analyzed case by case and a stability criterion is deduced. Finally, the conclusions are drawn in Section V.

II. Background

A. Congestion Control in Ethernet

While numerous investigations have been made on congestion control in TCP/IP networks, Ethernet networks, even after 30 years of their invention, run without congestion control in the link layer. To be competent for data center applications, it is crucial to introduce congestion management in the link layer for Ethernet networks. The main tasks include: (1) Providing congestion detection and information feedback to push congestion from the core towards the edge of the network; (2) Supporting rate control at the edge switches to shape flows causing congestion. In addition, an appropriate congestion management method for Ethernet networks should be compatible with the current IEEE802.1/802.3 standards and also work in harmony with the upper layer congestion control scheme, such as TCP flow control.

The IEEE 802.1Qau working group [3] is developing a new specification for congestion management in Ethernet networks. Until now, the specification has not been finalized. Four representative proposals, however, have been discussed and investigated. The general Ethernet Congestion Management (ECM) framework employing the Backward Congestion Notification (BCN) mechanism was proposed by D. Bergamasco [5]. The basic mechanism of BCN was initially used in frame

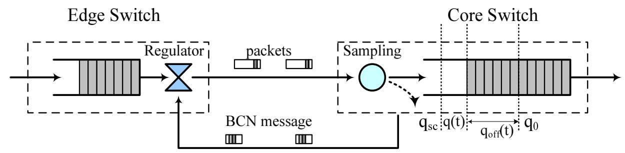

Fig. 1. Framework of BCN

relay networks called Backward Explicit Congestion Notification (BECN) [6]. Compared to FECN (Forward Explicit Congestion Notification) [6], the control messages in BCN are sent directly from switches back to sources when the congestion happens. Due to early congestion indication, BCN is more responsive to congestion, which is attractive for super high-speed Ethernet networks, but needs switches to generate extra backward packets to carry control messages. Unlike traditional BECN whose control message only carries binary congestion indication, the BCN mechanism also sends back the current queue status, including queue length and variation, to guide the adjustment of sending rate (described in detail later). Raj Jain argued for the explicit feedback of allowed rates and proposed the Forward Explicit Rate Advertising (FERA) mechanism [7], which is a variant of ERICA developed for explicit flow control in ATM networks [8]. The researchers at IBM Zurich lab combined some ideas of BCN and FERA to develop a new scheme called Explicit Ethernet Congestion Management (E2CM) [9]. The fourth proposal is quantized congestion notification (QCN) [10] in which the ECM queue feedback is quantized to a few bits and the network only provides negative feedback. Except for FERA, the other three proposals, including ECM, E2CM and QCN, follow the paradigm of BCN and send the queue dynamics back to the sources to instruct rate adjustment. In other words, the BCN mechanism is a cornerstone for most proposals of congestion control in DCE networks. For the convenience of modeling and analysis, we firstly describe the main components of the BCN mechanism in the next subsection.

B. BCN mechanism

We shall only introduce those parts of the BCN mechanism which are relevant for our stability analysis, and omit concrete implementation. The whole description can be seen in [5].

BCN is a rate-based closed-loop feedback control mechanism. Its framework and main components are shown in Fig.1. The technical goal of BCN is to hold the queue length of the core switch at equilibrium point q_0 so that the buffer is neither over-utilized nor under-utilized. The core switches monitor the instantaneous queue length $q(t)$. If $q(t)$ exceeds a high threshold $q_{sc} > q_0$, then the network is regarded as severely congested and the core switch sends out the PAUSE frames defined in IEEE 802.3x to require all its uplink neighbors to stop forwarding packets. In addition, the core switches sample the incoming packets with a deterministic probability p_m, and also simply count the number of arrival and departure packets to obtain the difference of instantaneous queue length

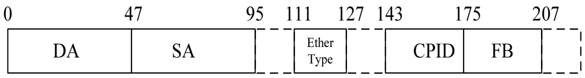

Fig. 2. Format of BCN message

$\Delta q(t)$ in a sampling interval. When a packet is sampled, the core switches determine the congestion level on the link by computing the key variable σ, which consists of a weighted sum of the instantaneous queue length offset and $\Delta q(t)$ over the last sampling interval:

$$\sigma(t) = [q_0 - q(t)] - w\Delta q(t) \tag{1}$$

where w is a weight parameter. The core switches may send a BCN message to the source of the sampled packet if necessary. The BCN messages follow 802.1Q VLAN tag format to ensure the coexistence and interoperability between BCN-aware and BCN-unaware switches. The key fields in BCN message are shown in Fig.2. Here, the 48-bits *DA* and *SA* fields contain the source address of the sampled packet and the address of the switch, respectively. *EtherType* tells the switches and end stations whether it is a BCN message. The *CPID* field is the ID of the congestion point and its purpose is to univocally identify a congestion entity, which should at least include the MAC address of the switch interface. The FB field carries the key measure $\sigma(t)$. If $\sigma(t) > 0$, BCN message is positive, else negative. When a normal BCN message reaches the congestion reaction point, which refers to a rate regulator associated with the source receiving the BCN message and is usually located in network interface card in the edge switch, the reaction point adjusts its sending rate using a modified Additive Increase and Multiplicative Decrease (AIMD) algorithm since it has been proven to be stable, convergent and fair under common network environments [11]. AIMD is defined in BCN as follows:

$$r(t) \leftarrow \begin{cases} r(t) + G_i R_u \sigma(t) & \text{if} \quad \sigma(t) > 0 \\ r(t)[1 + G_d \sigma(t)] & \text{if} \quad \sigma(t) < 0 \end{cases} \tag{2}$$

where $r(t)$ is the rate of sources, R_u is the increase rate unit parameter, G_i and G_d are the rate additive increase and multiplicative decrease gains, respectively. They need to be appropriately set by the network manager. A congestion reaction point receiving a negative BCN message will associate itself with the congestion point identified by the corresponding CPID. Its subsequent packets will contain a rate regulator tag (RRT) to carry this CPID. If these packets are sampled by the core switch, the value of CPID matches with the switch's CPID, a possible positive BCN message is generated and sent back the source of the sampled packets when $q(t) < q_0$, or else no BCN message is sent.

III. Analytical Model of BCN

A. Assumption

In this section, we firstly build an analytical model of the BCN mechanism. Considering regular and symmetrical network topologies in data centers, such as Monsoon [12], Fat-Tree [13] and D-Cell [14], and special traffic patterns driven by the parallel reads/writes in cluster file systems, such as Lustre [15] and Panasas [16], we assume that the sources are homogeneous, namely they have the same characteristics, follow the same routes, and experience the same round-trip propagation delays. In addition, the propagation delay in DCE networks is normally within the order of a few microseconds since the diameter of the network is only a few hundreds meters. Compared with the queuing delay in the order of several tens or hundreds microseconds, the propagation delay could be negligible. Furthermore, since links are assumed to be of high capacity in DCE networks, the number of bit stream in the links is so large that it appears like a continuous flow fluid, the fluid-flow approximation extensively used in network modeling work (in [17] and [4] etc.) is assumed to be practicable in our investigation.

B. Modeling

Considering the queue associated with the bottleneck link, and assuming that the queue length $q(t)$ is continuous and differentiable and the propagation delay is negligible, we have:

$$\frac{dq(t)}{dt} = \sum_{i=1}^{N} r_i(t) - C \tag{3}$$

where $r_i(t)$ is the rate of the i-th source in the link, $q(t)$ is the queue length, N is the number of active flows and C denotes the capacity of the bottleneck link. Because of the homogeneity of sources in DCE networks, (3) can be written as

$$\frac{dq(t)}{dt} = N\left[r_i(t) - \frac{C}{N}\right] \tag{4}$$

The difference of the queue length $\Delta q(t)$ in a sampling interval is

$$\Delta q(t) = \Delta t \frac{dq(t)}{dt} = \frac{1}{p_m C}\frac{dq(t)}{dt} \tag{5}$$

where Δt is the average sampling period. Combining (1), (4) and (5), the feedback variable $\sigma(t)$ in the BCN message can also be rewritten as

$$\sigma(t) = -\left\{[q(t) - q_0] + \frac{wN}{p_m C}\left[r_i(t) - \frac{C}{N}\right]\right\} \tag{6}$$

Referring to (2), we readily obtain

$$\frac{dr_i(t)}{dt} = \begin{cases} G_i R_u \sigma(t) & \sigma(t) > 0 \\ G_d \sigma(t) r_i(t) & \sigma(t) < 0 \end{cases} \tag{7}$$

So far, the dynamic of BCN congestion control system in DCE networks can be described by a set of first order ordinary nonlinear differential equations (4) and (7), which consist of a typical autonomous system. It is noted that equation (7) is characteristic of segmented nonlinearity. The switching function $\sigma(t) = 0$ divides the whole state space into two parts. The behaviors in the different state spaces are determined by the different differential equations.

C. Understanding of BCN Stability

The BCN mechanism aims to limit the queue length within buffer size. Since the empty queue wastes the link resource and the buffer overflow results in dropping packets, the rational solution is to regulate the queue length to the reference. Therefore, the stability of queue can be generally regarded as the stability of the whole system. In some investigations of Internet congestion control, the stability criterions in the classical linear control theory are straightforwardly used to evaluate system stability, such as [18] and [19]. The stability analysis of the BCN system in [4] follows the same paradigm. After the BCN system is approximated as two linear isolated subsystems, the conditions for the stability of each of the two subsystems are derived separately using Nyquist stability criterion. The sufficient condition for the stability of the overall BCN system is obtained through combining the conditions for stable subsystems stiffly. The conclusion can provide some insights, but its limitations are obvious. Because the nonlinearity of the variable structure control employed by the BCN mechanism is missed, some particular phenomena are not opened out. For example, the stability criterion given in [4] cannot definitely explain why queue oscillations occur in BCN system. Moreover, the conclusion in [4] can only tell the truth of steady behaviors in BCN system. When taking the physical constraint of buffer into account, packet droppings due to the transient overshoot of queue system cannot be properly described by the analysis in [4]. Therefore, in this work, we will use the phase plane analysis method in nonlinear control theory to conduct a comprehensive and in-depth investigation on transient and steady behaviors in BCN system where the nonlinearity of the variable structure control and the physical constraint of buffer are held.

The phase plane analysis is a graphic method to analysis the transient phenomena of second order ordinary differential equation. It can not only analyze the stability and oscillations but also give an explicit trajectory of the dynamics of the system, much better than displaying variables against time. The differential equations (4) and (7) tell that the queue controlled by the BCN mechanism is a second order nonlinear system. Assuming the phase trajectory of the BCN queue system is the curve $l\{\dot{q}(t), q(t)\}$ where $\dot{q}(t) = \frac{dq(t)}{dt}$, we illustrate the relationship between phase trajectories and queue stability. Some possible shapes of phase trajectories are shown in Fig. 3, where the queue motion is restricted to the strip shaded area on the phase plane because of the physical limitations of buffer (B denotes the buffer size). The system equilibrium state lies at point $(q_0, \frac{C}{N})$, and the switching line divides the shaded area into two parts, corresponding to rate decrease and increase, respectively. Apparently, the queue motions described by the phase curves l_1 and l_2 are unstable since they diverge from the equilibrium point. It is not difficult to obtain the similar conclusion using the stability criterions in control theory, just like in [4]. Naturally, these stability criterions also suggest that the queue motions described by the other phase curves is stable. Seemingly, the queue motions described by the phase curves

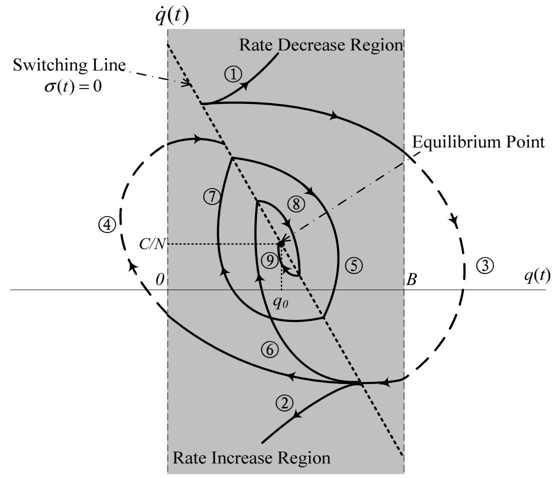

Fig. 3. Phase trajectories

l_3 and l_4 are stable since both of them eventually approach the equilibrium point. However, the movements along the dashed lines would never occur due to the physical limitations of the buffer. The queue stays either full or empty, which implies that the incoming packets are dropped or the link is wasted. The linear stability analysis, which does not take the nonlinearity of physical limitation of buffer into account, is incapable of opening up this phenomena. To overcome it, we introduce a more meaningful definition of the stability with respect to queue analysis in congestion control in DCE networks.

Definition 1: $\exists t_0 > 0$*, when* $t > t_0$*, if and only if* $0 < q(t) < B$*, the queue system is stable.*

By this definition, the queue described by the phase curves l_3 or l_4 are unstable since the transient overshoot empties buffer or causes overflow, which conflicts with the design goal of the BCN mechanism, namely improving throughput and avoiding packet droppings. To distinguish this definition from common definitions of stability, which are strictly defined by Lyapunov in control theory [20], we call it strong stability. According to this definition, the BCN system described by the phase curves l_5, l_6, l_7, l_8 and l_9 (exclusive of l_1, l_2, l_3 and l_4) is of strong stability.

In addition, there is a peculiar branch of phase trajectories, consisting of the curve l_5 in the rate decrease region and the curve l_7 in the rate increase region in Fig.3. It is a closed trajectory in the phase plane, which implies that the queue motion appears to be periodic oscillations. This phenomenon is called limit cycle, appearing in nonlinear systems frequently [21]. The linear stability analysis is incapable of revealing it. Even if the limit cycle could be stable, it would impose a negative impact on the fairness of the congestion control mechanism since the system can hardly converge towards the equilibrium point eventually. Therefore, it should be avoided through design of a proper rate regulation law or by fixing the appropriate parameter settings.

IV. Stability Analysis

Considering the limitations in linear stability analysis, we use the phase plane techniques in nonlinear control theory to analyze the stability of the BCN system. As explained in the previous subsection, the main advantage of this approach is particularly suitable for the analysis of the system with segmented nonlinear characteristics as the phase plane can be divided into regions corresponding to operation on a particular linear segment of the nonlinearity. The BCN congestion control system just possesses these properties. In addition, the transient behaviors of switching process between different regions can be completely captured.

A. General property of BCN system

Let $a = R_u G_i N, b = G_d$ and $k = \frac{w}{p_m C}$, and define the variable $x(t) = q(t) - q_0$ and $y(t) = Nr_i(t) - C$, then the equilibrium point becomes the origin $(0, 0)$. The switching function becomes $\sigma(t) = -[x(t) + ky(t)]$, and equations (4) and (7) are transformed into:

$$\begin{cases} \frac{dx(t)}{dt} = y(t) \\ \frac{dy(t)}{dt} = \begin{cases} -a[x(t) + ky(t)] & \sigma(t) > 0 \\ -b[y(t) + C][x(t) + ky(t)] & \sigma(t) < 0 \end{cases} \end{cases} \tag{8}$$

Define functions $f_1(t) = y(t), f_2(t) = -a[x(t) + ky(t)]$, and $f_3(t) = -b[y(t) + C][x(t) + ky(t)]$. Since for $i = 1, 2, 3$ and any $\vec{z} = (z_1, z_2) = (x(t), y(t))$

$$\| f_i(t, \vec{z}_1) - f_i(t, \vec{z}_2) \| \leq L \| \vec{z}_1 - \vec{z}_2 \|$$

Namely *Lipschitz* condition is satisfied, the set of nonlinear differential equations (8) have a solution uniquely determined by the initial value (x_0, y_0) [22]. Also since $x = 0, y = 0$ is unique solution of a set of equations

$$\begin{cases} f_1(t) = 0 \\ f_2(t) = 0 \end{cases} \quad \text{or} \quad \begin{cases} f_1(t) = 0 \\ f_3(t) = 0 \end{cases}$$

the origin is unique singular point of the nonlinear system (8). Lyapunov has shown that the stability of and the behavior of the trajectories in the neighborhood of a singular point can be found from a linearized version of nonlinear differential equations about the singular point [23]. Expanding the equations (8) in a Taylor series at the singular point, i.e. the equilibrium point $(0, 0)$, the equations of the first approximation are

$$\begin{cases} \frac{dx(t)}{dt} = y(t) \\ \frac{dy(t)}{dt} = \begin{cases} -ax(t) - aky(t) & \sigma(t) > 0 \\ -bCx(t) - \frac{bw}{p_m} y(t) & \sigma(t) < 0 \end{cases} \end{cases} \tag{9}$$

The eigenvalues of system (9) are roots of its characteristic equation:

$$\lambda^2 + m_i \lambda + n_i = 0 \qquad i \in \{1, 2\} \tag{10}$$

where $m_1 = ak$ and $n_1 = a$ as $\sigma(t) > 0$, or $m_2 = \frac{bw}{p_m}$ and $n_2 = bC$ as $\sigma(t) < 0$.

Considering the physical meaning of the network parameters, the coefficients m_i and n_i must be positive. According to Routh-Hurwitz stability criterion [24], we can straightforwardly get the following proposition.

Proposition 1: Both increasing and decreasing rate subsystems are stable in the context of pure linear control theory analysis.

Remarks: Without considering the nonlinear factors in the BCN congestion management system, such as buffer limitation and variable structure control, the conclusion drawn from separate analysis to subsystems is impractical because the transient behavior of switching between two subsystems is missed.

B. Typical phase trajectories

If the physical limitations of the buffer and the transient behavior of switching procedures are taken into account, we need to carefully check the phase trajectories to judge whether the BCN system is strongly stable. The phase trajectories of system (9) depend on its eigenvalues. Obviously,

$$\lambda_{1,2} = \frac{-m_i \pm \sqrt{m_i^2 - 4n_i}}{2} \tag{11}$$

Case1: $m_i^2 - 4n_i < 0$

In this case, the eigenvalues are complex conjugates $\lambda_{1,2} = \alpha \pm j\beta$, here $\alpha = -m_i/2$ and $\beta = \sqrt{4n_i - m_i^2}/2$. The solutions of a set of differential equations (9) are:

$$\begin{cases} x(t) = Ae^{\alpha t} \cos(\beta t + \varphi) \\ y(t) = A\alpha e^{\alpha t} \cos(\beta t + \varphi) - A\beta e^{\alpha t} \sin(\beta t + \varphi) \end{cases} \tag{12}$$

here $A = \sqrt{(\alpha^2 + \beta^2)[x(0)]^2 - 2\alpha x(0)y(0) + [y(0)]^2}/\beta$, $\varphi = -\arctan \frac{y(0) - \alpha x(0)}{\beta x(0)}$, where $x(0)$ and $y(0)$ are the initial values. A compact form of (12) is defined as $\{x(t), y(t)\} = \mathcal{H}\{t | x(0), y(0)\}$. Accordingly, $t = \mathcal{H}^{-1}\{x(t), y(t) | x(0), y(0)\}$ denotes solving t from equations (12) when $\{x(t), y(t)\}$ is known. The curves $\mathscr{H}\{x(t), y(t)\}$ are phase trajectories in this case .

Multiplying both sides of the first equation in (12) by $-\alpha$, then adding both sides of the second equation in (12), we have

$$y(t) - \alpha x(t) = -A\beta e^{\alpha t} \sin(\beta t + \varphi) \tag{13}$$

Combining (13) and the first equation in (12), we can yield

$$[y(t) - \alpha x(t)]^2 + [\beta x(t)]^2 = (A\beta)^2 e^{2\alpha t} \tag{14}$$

$$\tan(\beta t + \varphi) = -\frac{y(t) - \alpha x(t)}{\beta x(t)} \tag{15}$$

Solving t from (15), then substituting it into (14), we have

$$[y(t) - \alpha x(t)]^2 + [\beta x(t)]^2 = c_1 \exp\left\{-\frac{2\alpha}{\beta} \arctan \frac{y(t) - \alpha x(t)}{\beta x(t)}\right\} \tag{16}$$

here $c_1 = (A\beta)^2 \exp(-\frac{2\alpha\varphi}{\beta})$. Let $r(t)\cos\theta(t) = \beta x(t)$ and $r(t)\sin\theta(t) = \alpha x(t) - y(t)$. Since $\tan\theta(t) = -\frac{y(t) - \alpha x(t)}{\beta x(t)} = \tan(\beta t + \varphi)$, the form of (16) in polar coordinates should be

$$\begin{cases} r(t) = \sqrt{c_1} \exp\left\{\frac{\alpha}{\beta}\theta(t)\right\} \\ \theta(t) = \beta t + \varphi \end{cases} \tag{17}$$

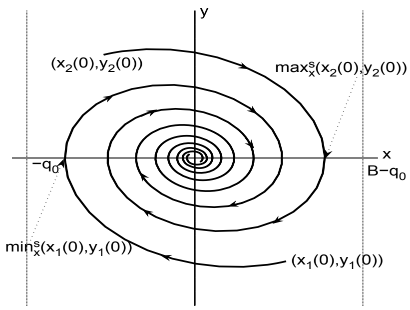

Fig. 4. Phase trajectories as $m_i^2 - 4n_i < 0$

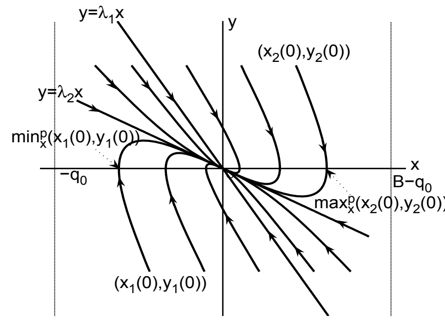

Fig. 5. Phase trajectories as $m_i^2 - 4n_i > 0$

Thus, the phase trajectory $\mathscr{H}\{x(t), y(t) \mid (x(0), y(0))\}$ beginning at point $(x(0), y(0))$ actually is a logarithmic spiral. Since $\alpha < 0$ and $\beta > 0$, r decreases but θ increases as time t increases, which implies that the motion along the phase trajectory eventually approaches the origin. Therefore, the singular point of the system is a stable focus. Fig.4 illustrates two branches of these phase trajectories originating from different initial points. Since $y(t) = \frac{dx(t)}{dt}$, $x(t)$ reaches its local extrema when $y(t) = 0$, namely $x(t)$ crosses the horizontal axis with its local extremum. Substituting $y(t) = 0$ into (12), we can obtain $\tan(\beta t + \varphi) = \frac{\alpha}{\beta}$. Since the inverse tangent is a multivalued function, $x(t)$ has infinite local extrema as shown in Fig.4. As $y(0) > 0$, the local extremum closest to the initial point $(x(0), y(0))$ is the local maximum denoted by $\max_x^s\{x(0), y(0)\}$, while $y(0) < 0$, it is the local minimum denoted by $\min_x^s\{x(0), y(0)\}$. Let t^* denote the time at which the system reaches the closest extremum. Since $t^* \geq 0$, we have

$$t^* = \begin{cases} \frac{1}{\beta}[\tan^{-1}\frac{\alpha}{\beta} + \tan^{-1}\frac{y(0)-\alpha x(0)}{\beta x(0)}] & x(0)y(0) \geq 0 \\ \frac{1}{\beta}[\pi + \tan^{-1}\frac{\alpha}{\beta} + \tan^{-1}\frac{y(0)-\alpha x(0)}{\beta x(0)}] & x(0)y(0) < 0 \end{cases} \tag{18}$$

where $\tan^{-1}(z)$ denote the principal value of the inverse tangent. Substituting (18) and $\cos(\beta t + \varphi) = \pm\beta/\sqrt{\alpha^2 + \beta^2}$ into (12), we readily obtain

$$\max_x^s\{x(0), y(0)\} = \frac{A\beta}{\sqrt{\alpha^2 + \beta^2}} \exp(\alpha t^*) \tag{19}$$

$$\min_x^s\{x(0), y(0)\} = -\frac{A\beta}{\sqrt{\alpha^2 + \beta^2}} \exp(\alpha t^*) \tag{20}$$

Case2: $m_i^2 - 4n_i > 0$
In this case, there are two different negative real eigenvalues. Assuming $\lambda_1 < \lambda_2 < 0$, the solutions of differential equations (9) are:

$$\begin{cases} x(t) = A_1 e^{\lambda_1 t} + A_2 e^{\lambda_2 t} \\ y(t) = A_1 \lambda_1 e^{\lambda_1 t} + A_2 \lambda_2 e^{\lambda_2 t} \end{cases} \tag{21}$$

where

$$A_1 = \frac{\lambda_2 x(0) - y(0)}{\lambda_2 - \lambda_1}; \quad A_2 = \frac{\lambda_1 x(0) - y(0)}{\lambda_1 - \lambda_2}$$

A compact expression of (21) is defined as $\{x(t), y(t)\} = \mathcal{F}\{t|x(0), y(0)\}$. The curves $\mathscr{F}\{x(t), y(t)\}$ are phase trajectories in this case.

Multiplying both sides of the first equation in (21) by λ_1 and λ_2, respectively, then subtracting the results from both sides of the second equation in (21), we have:

$$y(t) - \lambda_1 x(t) = (\lambda_2 - \lambda_1) A_2 e^{\lambda_2 t} \tag{22}$$

$$y(t) - \lambda_2 x(t) = (\lambda_1 - \lambda_2) A_1 e^{\lambda_1 t} \tag{23}$$

If $(x(0), y(0))$ satisfies with $y(0) - \lambda_1 x(0) = 0$, then $A_2 = 0$. From (22), we have

$$y(t) - \lambda_1 x(t) = 0 \tag{24}$$

If $y(0) - \lambda_2 x(0) = 0$, in the same way, we also have

$$y(t) - \lambda_2 x(t) = 0 \tag{25}$$

The phase trajectories defined (24) and (25) are particular. Both of them are straight lines. As $A_1 \neq 0$ and $A_2 \neq 0$, Combining (22) with (23), we make some algebraic operations, and then yield

$$[y(t) - \lambda_2 x(t)]^{\lambda_2} = c_2 [y(t) - \lambda_1 x(t)]^{\lambda_1} \tag{26}$$

where

$$c_2 = \frac{[(\lambda_1 - \lambda_2)A_1]^{\lambda_2}}{[(\lambda_2 - \lambda_1)A_2]^{\lambda_1}} = \frac{[y(0) - \lambda_2 x(0)]^{\lambda_2}}{[y(0) - \lambda_1 x(0)]^{\lambda_1}}$$

Let $u(t) = y(t) - \lambda_1 x(t)$ and $v(t) = y(t) - \lambda_2 x(t)$, equation (26) is transformed into

$$v(t) = \sqrt[\lambda_2]{c_2}\, u(t)^{\frac{\lambda_1}{\lambda_2}} \tag{27}$$

Hence, the phase trajectory $\mathscr{F}\{x(t), y(t) \mid (x(0), y(0))\}$ originating from the initial point $(x(0), y(0))$, here $y(0) \neq \lambda_{1,2} x(0)$, is analogous to a parabola. Some phase trajectories originating from various initial points are depicted in Fig.5. Since the motion along the phase trajectory eventually approaches the origin, the singular point is called a stable node.

As $y(t) = 0$, $x(t)$ arrives its global extremum denoted by $\text{mum}_x^p\{x(0), y(0)\}$:

$$\text{mum}_x^p\{x(0), y(0)\} = -\left\{\frac{(-\lambda_1)^{\lambda_1}[y(0) - \lambda_2 x(0)]^{\lambda_2}}{(-\lambda_2)^{\lambda_2}[y(0) - \lambda_1 x(0)]^{\lambda_1}}\right\}^{\frac{1}{\lambda_2 - \lambda_1}} \tag{28}$$

$\text{mum}_x^p\{x(0), y(0)\}$ is the maximum as $y(0) > 0$ or the minimum while $y(0) < 0$

Case3: $m_i^2 - 4n_i = 0$

In this case, two real eigenvalues are identical, i.e. $\lambda_{1,2} = \lambda = -\frac{m_i}{2}$.The solutions of equations (9)become:

$$\begin{cases} x(t) = (A_3 + A_4 t)e^{\lambda t} \\ y(t) = (A_3\lambda + A_4 + A_4\lambda t)e^{\lambda t} \end{cases} \tag{29}$$

where $A_3 = x(0)$ and $A_4 = y(0) - \lambda x(0)$. A compact form of (29) is defined as $\{x(t), y(t)\} = \mathcal{L}\{t|x(0), y(0)\}$, and the curves $\mathscr{L}\{x(t), y(t)\}$ describe the phase trajectories in this case.

From (29), we can get

$$y(t) - \lambda x(t) = A_4 e^{\lambda t} \tag{30}$$

If $(x(0), y(0))$ satisfies with $y(0) - \lambda x(0) = 0$, then $A_4 = 0$. Thus, the phase trajectory originating from $(x(0), y(0))$ is a straight line:

$$y(t) - \lambda x(t) = 0 \tag{31}$$

From (29), we can also have

$$y(t)(A_3 + A_4 t) = x(t)(\lambda A_3 + A_4 + A_4\lambda t) \tag{32}$$

Solving t from (32), and substituting it into (30)

$$y(t) - \lambda x(t) = A_4 \exp\left\{\lambda \frac{A_4 x(t) - A_3[y(t) - \lambda x(t)]}{A_4[y(t) - \lambda x(t)]}\right\} \tag{33}$$

Since $\lambda < 0$, $\lim_{t\to\infty} te^{\lambda t} = 0$. Therefore, the movement along the phase trajectory eventually approaches the origin. The singular point is a stable node. Although the solution form of equations (9) in this case is different from that of the previous case, the phase trajectories $\mathscr{L}\{x(t), y(t)\}$ are similar to $\mathscr{F}\{x(t), y(t)\}$. The only difference is that the former contain one straight line, but the latter include two straight lines. Because of limited space, we omit the illustration of the phase trajectory in this case. As $y(t) = 0$, $x(t)$ reaches its unique extremum denoted by $\text{mum}_x^u\{x(0), y(0)\}$:

$$\text{mum}_x^u\{x(0), y(0)\} = -\frac{A_4}{\lambda} e^{-\frac{\lambda A_3 + A_4}{\lambda A_4}} \tag{34}$$

$\text{mum}_x^u\{x(0), y(0)\}$ is the maximum as $y(0) > 0$ or the minimum while $y(0) < 0$.

C. Stability analysis of BCN system

Next, according to the definition of strong stability, we use the phase plane techniques to analyze the stability of the BCN system described by equations (9) with linear segmented nonlinear properties.

Generally, the initial state of the BCN system is $q(0) = 0$ and $r_i(0) = \mu$, here μ is the initial rate of sources. This initial state corresponds to the point $(-q_0, N\mu - C)$ in the phase plane. Whatever rate control laws are employed, the system will move from the point $(-q_0, N\mu - C)$ to the point $(-q_0, 0)$ along the straight line $x = -q_0$. In other words, the point $(-q_0, 0)$ is reachable at all time. Because the queue is always empty in this accelerating stage, the core switch can hardly sense the queue variance, and $y(t)$ complies with the regulation law $\frac{dy(t)}{dt} = aq_0$. Thus, this acceleration will last $T_0 = \frac{C - N\mu}{aq_0}$. Since the phase trajectory of the warm-up process is identical for any regulation laws and the point $(-q_0, 0)$ is always reachable, in our phase plane analysis, assume the initial point is $(-q_0, 0)$, namely the queue associated with the bottleneck link is empty and the aggregation rate of N sources equal to the capacity of link C. Referring to Fig.4 and Fig.5 and considering the switching line $\sigma(t) = x(t) + ky(t) = 0$, the phase trajectories of the dynamic system defined by the nonlinear differential equations (9) would present rather various shapes if the parameters are unconstrained. Taking the physical limitations of parameters and their relationship into account, there are only six basic types of phase trajectories. Since $m_2 = \frac{bw}{p_m} = kbC$, equation (10) can be rewritten in the following form

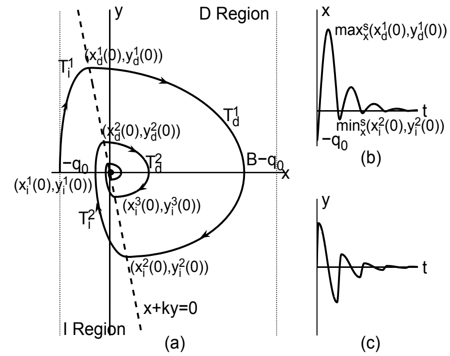

Fig. 6. Phase trajectory and dynamic behaviors as $a < \frac{4p_m^2 C^2}{w^2}$ and $b < \frac{4p_m^2 C}{w^2}$

$$\lambda^2 + kn_i\lambda + n_i = 0 \qquad i \in \{1, 2\} \tag{35}$$

where $n_1 = aN$ as $\sigma(t) > 0$, or $n_2 = bC$ as $\sigma(t) < 0$. As the characteristic equation (35) has two different negative real roots $\lambda_{1,2} = \frac{-kn_i \mp \sqrt{(kn_i)^2 - 4n_i}}{2}$, and $\lambda_2 + \frac{1}{k} = \frac{\sqrt{(k^2 n_i - 2)^2 - 4} - (k^2 n_i - 2)}{2k} < 0$, we have $-\frac{1}{k} > \lambda_2 > \lambda_1$. Thus, some theoretically feasible phase trajectories impossibly appear in the actual BCN system. Subsequently, we discuss the stability of the BCN queue moving along six basic phase trajectories case by case.

Case1: $a < \frac{4p_m^2 C^2}{w^2}$ **and** $b < \frac{4p_m^2 C}{w^2}$

The queue moves along $\mathscr{H}$-type phase trajectories (i.e. logarithmic spiral) in rate increase and decrease regions. A trajectory is shown in Fig.6 (a), where the kth round of rate increasing and decreasing starts from $(x_i^k(0), y_i^k(0))$[1] and

[1]To avoid confusion, the k-th power of $x_i^k(0)$ is denoted by $[x_i^k(0)]^k$

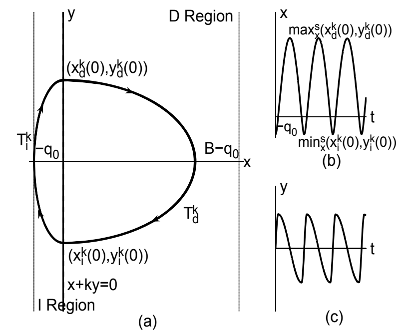

Fig. 7. Limit cycle motion

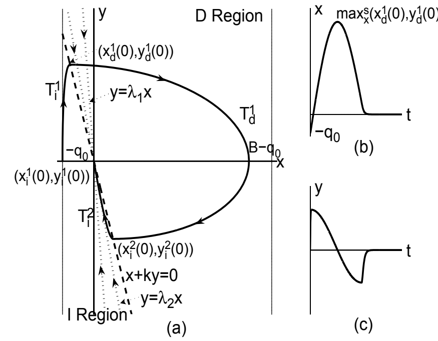

Fig. 8. Phase trajectory and dynamic behaviors as $a > \frac{4p_m^2C^2}{w^2}$ and $b < \frac{4p_m^2C}{w^2}$

$(x_d^k(0), y_d^k(0))$ and lasts T_i^k and T_d^k, respectively. Fig.6 (b) and Fig.6 (c) describe the evolution of queue length and its variance, respectively. Let A_i^k, φ_i^k and A_d^k, φ_d^k denote the coefficients of (12) in the the kth round of rate increasing and decreasing, respectively. Since $(x_i^1(0), y_i^1(0)) = (-q_0, 0)$, and by the expressions of coefficients in (12), we have

$$A_i^1 = \frac{2q_0\sqrt{a}}{\sqrt{4a-a^2k^2}}; \quad \varphi_i^1 = -\tan^{-1}(\frac{ak}{\sqrt{4a-a^2k^2}})$$

Let $x(t) = x_d^1(0), y(t) = y_d^1(0)$ in (12), and since $x_d^1(0) = -ky_d^1(0)$, we can obtain

$$\begin{aligned} T_i^1 &= \mathcal{H}^{-1}\{x_d^1(0), y_d^1(0) \mid -q_0, 0\} \\ &= \frac{2}{\sqrt{4a-a^2k^2}}[\tan^{-1}(\frac{2-ak^2}{k\sqrt{4a-a^2k^2}}) - \varphi_i^1] \end{aligned}$$

and

$$x_d^1(0) = -\frac{kA_i^1\sqrt{4a-a^2k^2}}{2}e^{-\frac{ak}{2}T_i^1}$$

As both $x_d^1(0)$ and $y_d^1(0)$ are determined, we readily get

$$A_d^1 = \frac{2|y_d^1(0)|}{\sqrt{4bC-(kbC)^2}}; \quad \varphi_d^1 = \tan^{-1}(\frac{2-bk^2C}{k\sqrt{4bC-(kbC)^2}})$$

Since $x_d^1(0)y_d^1(0) < 0$, we can get the maximum of $x(t)$ by Equation (19)

$$\begin{aligned} \max\nolimits_1\{x(t)\} &= \max\nolimits_x^s\{x_d^1(0), y_d^1(0)\} \\ &= \tfrac{|x_d^1(0)|}{k\sqrt{bC}}\exp\{\tfrac{\alpha_d}{\beta_d}(\pi + \tan^{-1}\tfrac{\alpha_d}{\beta_d} - \varphi_d^1)\} \end{aligned} \tag{36}$$

where $\frac{\alpha_d}{\beta_d} = \frac{-bkC}{\sqrt{4bC-(kbC)^2}}$. In the same way, we can also get

$$T_d^1 = \frac{2\pi}{\sqrt{4bC-(kbC)^2}},$$

$$x_i^2(0) = -A_d^1\frac{k\sqrt{4bC-(kbC)^2}}{2}e^{-\frac{bkC}{2}T_d^1},$$

$$A_i^2 = \frac{2|x_i^2(0)|\sqrt{a}}{k\sqrt{4a-a^2k^2}} \quad \text{and} \quad \varphi_i^2 = \tan^{-1}(\frac{2-ak^2}{k\sqrt{4a-a^2k^2}})$$

Substituting them into (20), yield

$$\begin{aligned} \min\nolimits_1\{x(t)\} &= \min\nolimits_x^s\{x_i^2(0), y_i^2(0)\} \\ &= -\tfrac{|x_i^2(0)|}{k\sqrt{a}}\exp\{\tfrac{\alpha_i}{\beta_i}[\pi + \tan^{-1}\tfrac{\alpha_i}{\beta_i} - \varphi_i^2]\} \end{aligned} \tag{37}$$

where $\frac{\alpha_i}{\beta_i} = \frac{-ak}{\sqrt{4a-a^2k^2}}$. Therefore, we have the following proposition.

Proposition 2: As $a < \frac{4p_m^2C^2}{w^2}$ and $b < \frac{4p_m^2C}{w^2}$, if $\max_1\{x(t)\} < B - q_0$ and $\min_1\{x(t)\} > -q_0$, here $\max_1\{x(t)\}$ and $\min_1\{x(t)\}$ are defined by (36) and (37) respectively, then the BCN system is strongly stable.

In this case, there is a special queue motion. when $x_i^k(0) = x_i^{k+1}(0)$ or $x_d^k(0) = x_d^{k+1}(0)$, the limit cycle appears. Fig.7 illustrates a typical limit cycle motion, where the queue of switches and the rate of sources oscillate with an identical amplitude, and the system cannot eventually approach the equilibrium point. This phenomena has been observed in some experiments of [4], but the analysis based on linear control theory cannot provide reasonable explanations.

Case2: $a > \frac{4p_m^2C^2}{w^2}$ and $b < \frac{4p_m^2C}{w^2}$

The phase trajectory of the queue is a parabola in the rate increasing region, but a spiral in the rate decreasing region. The system motion in this case is shown in Fig.8. As in the aforementioned discussion, $-\frac{1}{k} > \lambda_{1,2}$, thus the queue phase trajectory must traverse the switching line $x + ky = 0$ in the second quadrant. After traversing the switching line for the second time in the fourth quadrant, it gradually approaches the equilibrium point and the straight line $y = \lambda_2 x$, but never traverse the latter because it is an asymptote of the phase trajectory. Substituting $x_i^1(0) = -q_0, y_i^1(0) = 0$ and $y_d^1(0) = -kx_d^1(0)$ into (26), we can get

$$y_d^1(0) = q_0\left[\frac{(k+1/\lambda_1)^{\lambda_1}}{(k+1/\lambda_2)^{\lambda_2}}\right]^{\frac{1}{\lambda_2-\lambda_1}}$$

Following the same method in case 1, we also have

$$\begin{aligned} \max\nolimits_2\{x(t)\} = \tfrac{q_0}{\sqrt{bC}}\left[\tfrac{(k+1/\lambda_1)^{\lambda_1}}{(k+1/\lambda_2)^{\lambda_2}}\right]^{\frac{1}{\lambda_2-\lambda_1}}\exp\Big\{\tfrac{\alpha_d}{\beta_d}[\pi \\ + \tan^{-1}\tfrac{\alpha_d}{\beta_d} - \varphi_d^1]\Big\} \end{aligned} \tag{38}$$

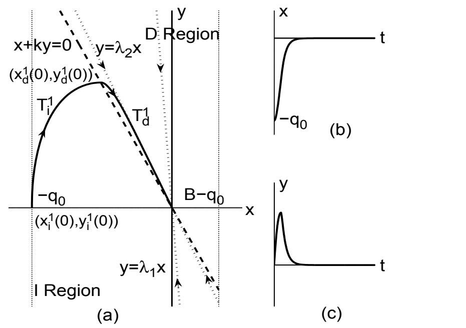

Fig. 9. Phase trajectory and dynamic behaviors as $a < \frac{4p_m^2C^2}{w^2}$ and $b > \frac{4p_m^2C}{w^2}$

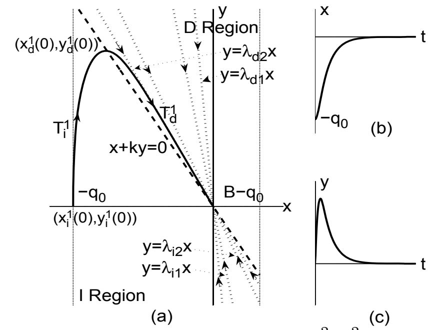

Fig. 10. Phase trajectory and dynamic behaviors $a > \frac{4p_m^2C^2}{w^2}$ and $b > \frac{4p_m^2C}{w^2}$

where $\frac{\alpha_d}{\beta_d} = \frac{-bkC}{\sqrt{4bC-(kbC)^2}}$, and $\lambda_{1,2} = \frac{-ka\mp\sqrt{a^2k^2-4a}}{2}$. Thus, we have the proposition.

Proposition 3: As $a < \frac{4p_m^2C^2}{w^2}$ and $b < \frac{4p_m^2C}{w^2}$, if $\max_2\{x(t)\} < B - q_0$, here $\max_2\{x(t)\}$ is determined by (38), then the BCN system is strongly stable.

Case3: $a < \frac{4p_m^2C^2}{w^2}$ **and** $b > \frac{4p_m^2C}{w^2}$

Contrary to case 2, the phase trajectory is a logarithmic spiral in the rate increasing region, but a parabola in the decreasing region. Fig.9 illustrates the representative behaviors of the BCN system in this case. After crossing the switching line, the phase trajectory moves directly to the equilibrium point. The queue motion in phase plane is limited in the second quadrant because the line $y = \lambda_2 x$ is a asymptote for the phase trajectory originating from the initial point $(x_d^1(0), y_d^1(0))$ as $y_d^1(0) \neq \lambda_2 x_d^1(0)$, which implies that the queue length does not overshoot the reference value q_0 as shown in Fig.9 (b). Therefore, the BCN system is strongly stable forever in this case.

Case4: $a > \frac{4p_m^2C^2}{w^2}$ **and** $b > \frac{4p_m^2C}{w^2}$

In this case, the phase trajectory is a parabola in both rate increasing and decreasing regions. The dynamic evolution of system is shown in Fig.10. The strong stability can be always guaranteed based on the same reasons in case 3.

Case5: $a = \frac{4p_m^2C^2}{w^2}$ **or** $b = \frac{4p_m^2C}{w^2}$

In this case, the switching line $x + ky = 0$ is a special phase trajectory due to $\lambda_{1,2} = -\frac{1}{k}$. Therefore, the phase trajectory only appears in the rate increasing region as $a = \frac{4p_m^2C^2}{w^2}$, or approaches the equilibrium point along the switching line as $b = \frac{4p_m^2C}{w^2}$. Naturally, the system is strongly stable.

The analysis in these three cases can be summarized as the following proposition.

Proposition 4: As $b \geq \frac{4p_m^2C}{w^2}$ or $a = \frac{4p_m^2C^2}{w^2}$, the BCN system is strongly stable.

Summing up *Proposition 2 to 4*, we can draw a stability criterion for the BCN congestion control system in DCE networks.

Theorem 1: The sufficient condition for the strongly stable BCN system is $\left\{1 + \sqrt{\frac{R_uG_iN}{G_dC}}\right\} q_0 < B$

Proof: Since $\frac{\alpha_d}{\beta_d} < 0$ and $(\pi + \tan^{-1}\frac{\alpha_d}{\beta_d} - \varphi_d^1) > 0$, from (36), we have

$$\max_1\{x(t)\} < \frac{|x_d^1(0)|}{k\sqrt{bC}}$$

Also $x_d^1(0) = -\frac{kA_i^1\sqrt{4a-a^2k^2}}{2}e^{-\frac{ak}{2}T_i^1}$ and $A_i^1 = \frac{2q_0\sqrt{a}}{\sqrt{4a-a^2k^2}}$, thus $\max_1\{x(t)\} < \sqrt{\frac{a}{bC}}q_0$ In the same way, from (37), we have

$$\min_1\{x(t)\} \geq -\frac{|x_i^2(0)|}{k\sqrt{a}}$$

Referring to the variables $x_i^2(0)$ and A_d^1 defined in case 1, and considering $y_d^1(0) = -\frac{x_d^1(0)}{k}$, we can easily find $\min_1\{x(t)\} > -q_0$. In other words, as $\left\{1 + \sqrt{\frac{a}{bC}}\right\} q_0 < B$, the conditions required by *Proposition 2* can be satisfied.

In case 2, $\lambda_{1,2} = \frac{-ka\mp\sqrt{a^2k^2-4a}}{2} < 0$. Since $\lambda_1 < \lambda_2 < -\frac{1}{k} < 0$, we have $\frac{k+1/\lambda_1}{k+1/\lambda_2} < 1$, thus $\left[\frac{k+1/\lambda_1}{k+1/\lambda_2}\right]^{\lambda_1} < 1$. Also

$$\frac{(k+1/\lambda_1)^{\lambda_1}}{(k+1/\lambda_2)^{\lambda_2}} = \frac{(k+1/\lambda_1)^{\lambda_1}}{(k+1/\lambda_2)^{\lambda_1}(k+1/\lambda_2)^{\lambda_2-\lambda_1}}$$

we can get the inequality

$$\left[\frac{(k+1/\lambda_1)^{\lambda_1}}{(k+1/\lambda_2)^{\lambda_2}}\right]^{\frac{1}{\lambda_2-\lambda_1}} < \frac{1}{k+1/\lambda_2} = \frac{2a}{ka-\sqrt{a^2k^2-4a}}$$
$$= \frac{2/k}{1-\sqrt{1-[\frac{2}{k\sqrt{a}}]^2}} < \sqrt{a}$$

Thus, (38) can be written as

$$\max_2\{x(t)\} < \sqrt{\frac{a}{bC}}q_0$$

which means that the condition required by *Proposition 3* can be satisfied as $\left\{1 + \sqrt{\frac{a}{bC}}\right\} q_0 < B$. Since $a = R_uG_iN$ and $b = G_d$, $\left\{1 + \sqrt{\frac{R_uG_iN}{G_dC}}\right\} q_0 < B$ is a sufficient condition that the BCN system is strongly stable ■

Remarks: Since the concept of strong stability is introduced, our stability criterion, which is relative to buffer size B, can impose constraints on both steady and transient behaviors. The stability criterion in [4] is irrelative to the parameter B

because it does not reflect the impact of transient behaviors on buffer overflow. *Theorem1* tells that the maximum of the queue length $\max q(t) \propto \sqrt{\frac{R_u G_i N}{G_d C}} q_0$, which increases with $\sqrt{\frac{N}{C}}$ and is irrespective to the other parameters, such as the slope of the switching line. In other words, the control parameters w and p_m do not affect the stability of system, and merely impose the impacts on transient performance, such as the speed of convergence and the existence of limit cycles. Since $\max q(t) \propto q_0$, a small reference value of queue length is in favor of strong stability, but prolongs the start-up process since $T_0 = \frac{C - N\mu}{N R_u G_i q_0}$. A reasonable trade-off has to be made in parameter configuration. In addition, *Theorem1* also implies that the classical rule-of-thumb for buffer dimensioning, i.e. the bandwidth delay product rule, is becoming unsustainable if packets cannot be dropped due to congestion. Assume that $N = 50, C = 10$Gbps, and the propagation delay of the link is $0.5us$ (i.e. the length is $100m$). According to the bandwidth delay product rule, the buffer size should be 5 Mbits. Let $q_0 = 2.5$ Mbits, and the other parameters are fixed at the values recommended by standard draft [5], i.e. $G_i = 4, G_d = \frac{1}{128}$ and $R_u = 8$ Mbits, the strongly stable BCN queue system requires the buffer size of 13.75Mbits according to the estimation of *Theorem1*, which is nearly three times as large as the bandwidth delay product. We can simply decrease the control parameter G_i or increase G_d to shrink the required buffer size, but it is likely to deteriorate the transient performance. For example, the convergence becomes sluggish. The proper parameter settings are prone to optimize the comprehensive performance, and the stability is only one of the basic metrics.

V. Conclusion

Data center networks present some opportunities and challenges for developing new networking technology. DCE is a potential candidate for unified network fabrics in data centers. To avoid dropping packets due to congestion, the congestion management schemes are introduced into DCE networks. The BCN mechanism is a basic element of several proposals of standard drafts. The stability of the BCN mechanism underlies normal network operations. In this paper, we introduce the concept of strong stability to fit for the requirements of no dropped packets in DCE networks, and build a fluid-flow model for BCN congestion control systems. Considering the nonlinearity of the variable structure control employed by the BCN mechanism, we investigate the strong stability of BCN system using phase plane analysis techniques on a case by case basis, and synthesize the fragmentary results into a stability criterion which explicitly presents the constraints complied with by the control parameters. It can provide straightforward guidelines for network management in practice. In our future work, we will follow the same analysis method to investigate the transient behaviors of BCN system and evaluate the impact of parameters on the transient performance.

Acknowledgment

This work is supported by National Natural Science Foundation of China (No.60971102 and 60773138), National Grand Fundamental Research 973 Program of China (No.2009CB320504 and 2006CB303000), and National Science and Technology Major Project (NSTMP) under Grant No. 2009ZX03006-001

References

[1] "Data center ethernet: Cisco's innovation for data center networks," White Paper, Cisco. [Online]. Available: http://www.cisco.com
[2] C. DeSanti and S. Gai, *I/O Consolidation in the Data Center*. Cisco Press, 2009.
[3] *End-to-end congestion management*, IEEE Working Draft 802.1Qau. [Online]. Available: http://www.ieee802.org/1/pages/802.1au.html
[4] Y. Lu, R. Pan, B. Prabhakar, D. Bergamasco, V. Alaria, and A. Baldini, "Congestion Control in Networks with No Congestion Drops," in *Proc. the 44nd Allerton Annual Conference on Communication, Control and Computing*, Spe. 2006.
[5] D. Bergamasco, "Data Center Ethernet Congestion Management: Backward Congestion Notification." [Online]. Available: http://www.ieee802.org/1/files/public/docs2007/au-bergamasco-ecm-v0.1.pdf
[6] J. T. Buckwalter, *Frame Relay: Technology and Practice* . Addison-Wesley Professional, 1999.
[7] J. Jiang, R. Jain, and C. So-In, "An Explicit Rate Control Framework for Lossless Ethernet Operation," in *Proc. IEEE ICC 2008*, May 2008.
[8] S. Kalyanaraman, R. Jain, S. Fahmy, R. Goyal, and B. Vandalore, "The ERICA switch algorithm for ABR traffic management in ATM networks," *IEEE/ACM Trans. on Networking*, vol. 8, no. 1, pp. 87–98, Feb. 2000.
[9] M. Gusat, "Extended Ethernet Congestion Management(E2CM)," IEEE 802.1 Meeting. [Online]. Available: http://www.ieee802.org/1/files/public/docs2007/
[10] R. Pan, B. Prabhakar, and A. Laxmikantha, "QCN: Quantized Congestion Notification," IEEE 802.1 Meeting. [Online]. Available: http://www.ieee802.org/1/files/public/docs2008/
[11] D. M. Chiu and R. Jain, "Analysis of Increase and Decrease Algorithms for Congestion Avoidance in Computer Networks," *Computer Networks and ISDN Systems*, vol. 17, no. 1, pp. 1–14, Jan. 1989.
[12] A. Greenberg, P. Lahiri, D. A. Maltz, P. Patel, and S. Sengupta, "Towards a Next Generation Data Center Architecture: Scalability and Commoditization," in *Proc. of the ACM workshop on Programmable routers for extensible services of tomorrow (2008)*, Aug. 2008, pp. 57–62.
[13] M. Al-Fares, A. Loukissas, and A. Vahdat, "A Scalable, Commodity Data Center Network Architecture," in *Proc. of the ACM SIGCOMM 2008*, Aug. 2008.
[14] C. Guo, H. Wu, K. Tan, L. Shi, Y. Zhang, and S. Luz, "DCell: A Scalable and Fault-Tolerant Network Structure for Data Centers," in *Proc. of the ACM SIGCOMM 2008*, Aug. 2008.
[15] P. J. Braam, "File systems for clusters from a protocol perspective." [Online]. Available: http://www.lustre.org
[16] B. Welch, M. Unangst, Z. Abbasi, G. Gibson, B. Mueller, J. Zelenka, and B. Zhou, "Scalable Performance of the Panasas Parallel File System," in *Proc. of USENIX Conference on File and Storage Technologies*, Feb. 2008.
[17] V. MISRA, W. GONG, and D. TOWSLEY, "Fluid-based Analysis of a Network of AQM Routers Supporting TCP Flows with an Application to RED," in *Proc. of the ACM SIGCOMM 2000*, Aug. 2000.
[18] R. Srikant, *The Mathematics of Internet Congestion Control*. Springer Verlag, 2004.
[19] C. Hollot, V. Misra, D. Towsley, and W.-B. Gong, "A Control Theoretic Analysis of RED," in *Proc. of IEEE INFOCOM 2001*, April 2001, pp. 1510–1519.
[20] A. Lyapunov, *Stability of Motion*. Academic Press, 1966.
[21] D. Atherton, *Nonlinear Control Engineering*. Van Nostrand Reinhold, 1982.
[22] V. Arnol'd, *Ordinary Differential Equations*. Springer-Verlag, 1992.
[23] E. Coddington and N. Levionson, *Theroy of Ordinary Differential Equations*. McGraw Hill, 1955.
[24] E. J. Routh, *Stability of Motion*. Halsted Press, 1975.