Perron-Frobenius Measure of Systemic Risk of Cascading Overload in Complex Clouds: Work in Progress

V. Marbukh
National Institute of Standards and Technology
100 Bureau Drive, Stop 8910
Gaithersburg, MD 20899-8910
E-mail: marbukh@nist.gov

Abstract—**Dynamic resource sharing has both positive and negative effects on the system performance. Current research mostly concentrates on the positive effect resulting from the statistical multiplexing gain. However, the negative effect resulting from the unavoidable overhead, e.g., due to coordination or delays for accessing distant resources, may outweight the positive effect. This possibility has been demonstrated analytically and through simulations as well as has been observed in operational networks with dynamic routing, when a positive effect of bandwidth sharing is offset by using unnecessary long routes. Emergence and proliferation of cloud computing necessitates better understanding of the risk of cascading overload in highly distributed complex resource sharing systems. This paper, which represents work in progress, attempts to make a case for the Perron-Frobenius eigenvalue of the mean-field equations linearized about the "operational" system equilibrium to be viewed as a measure of the systemic risk of cascading overload in complex resource sharing systems. For simplicity of the exposition we consider a loss service model. Future research will investigate practical applicability and a possibility of online estimation of the proposed risk measure.**

Keywords-complex clouds, resource sharing, systemic risk, cascading overload, Perron-Frobenius.

I. INTRODUCTION

Understanding, predicting, and ultimately controlling complex distributed resource sharing systems is a problem of significant interest in a variety of contexts. Recent results indicate benefits of dynamic resource sharing in large-scale clouds due to the statistical multiplexing gain [1]-[3]. However, empirical evidence suggests that dynamic resource sharing may have both positive and negative effects on the overall system performance, where the negative effect is due to the overhead associated with accessing distant resources. Even more troubling is that this negative effect may manifest itself in abrupt transition to unacceptably highly congested system equilibrium through the process of cascading overload.

This paper suggests that an adequate measure of the systemic risk of cascading overload in complex clouds can be obtained by applying Perron-Frobenius theory to the mean-field system performance model. Mean-field approximation has accurately predicted cascading overload in communication networks with dynamic routing [4]-[6]. Unfortunately, despite that mean-field approximation significantly reduces the dimension of the system description, this reduction is not sufficient to allow for a direct solution of the mean-field equations for practical size networks. This paper, which is a work in progress, suggests that the Perron-Frobenius eigenvalue [7] of the mean-field equations linearized about the "operational" system equilibrium quantifies systemic risk of cascading overload in complex resource sharing systems. This approach extends line of research [8]-[9] to a case of discontinues, i.e., first order systemic events.

Our results, which are consistent with [6], indicate that dynamic resource sharing may be beneficial or harmful depending on whether the exogenous load is sufficiently light or heavy. For an intermediate exogenous load, when both "normal" and "congested" system equilibria coexist as metastable, the beneficial effect dominates in the "normal" metastable equilibrium, and the negative effect dominates in the "congested" metastable equilibrium. Assuming that the system remains in the "normal" metastable equilibrium, the optimal system performance is achieved on the stability boundary of this equilibrium, creating an inherent trade-off between equilibrium system performance and systemic risk of cascading overload leading to a highly congested equilibrium.

This paper is organized as follows. Section II introduces a loss resource sharing system. Section III proposes a mean-field approximation for this system performance. Section IV discusses system performance under mean-field approximation in a tractable "homogeneous" case. Section V suggests a combination of mean-field approximation and Perron-Frobenius theory as an adequate and tractable framework for assessing systemic risk of cascading overload in a general case. Finally, Section V briefly summarizes and outlines directions of future research.

II. RESOURCE ALLOCATION SYSTEM

Consider a system with I classes of requests and J classes of servers, where class $j=1,..,J$ includes N_j servers. Requests of class $i=1,..,I$ arrive following a

Poisson process of rate Λ_i, and have exponentially distributed service time with average $1/\mu_{ij}$ on a class j server. The resource allocation strategy is determined by probabilities $Q=(q_{ij})_{i,j=1}^{I,J}$, where $\sum_{j=1}^{J} q_{ij} \le 1$ $q_{ij} \ge 0$. An arriving request of class i is immediately lost with probability $q_{i0} = 1-\sum_{j=1}^{J} q_{ij}$, and occupies an available server of class j selected with probability q_{ij}. If no class j server is available, the request makes another attempt to find an available server with the same probabilities. After K_i unsuccessful attempts the request is lost.

Assuming that an accepted request of class i brings revenue w_i, the resource allocation scheme performance can be characterized by the portion of the lost revenue

$$L=\left(\sum_i w_i\right)^{-1} \sum_i w_i \pi_i \quad (1)$$

where the steady-state probability of a class i request being rejected π_i can in principle be determined as follows. The system evolution is described by a Markov process $X(t)=(x_{ij}(t))_{i,j=1}^{I,J}$, where $x_{ij}(t)$ is the number of a class i requests occupying class j servers. Process $X(t)$ has unique steady-state distribution $P(X)$, which is a solution of the corresponding system of Kolmogorov equations. Once probabilities $P(X)$ are known, rejection probabilities π_i can be determined as follows:

$$\pi_i = q_{i0} + (1-q_{i0})\left(\sum_{j \in J} q_{ij} p_j\right)^{K_i} \quad (2)$$

where probability of all servers of class j being occupied is $p_j = \sum_{x_{1j}+..x_{Ij}=N_j} P(X)$.

Since an astronomically high number of Markov process $X(t)=(x_{ij}(t))_{i,j=1}^{I,J}$ states makes direct solution of the corresponding Kolmogorov system computationally infeasible, Section III proposes a mean-field approximation, which reduces the problem dimension from hyper-exponential to linear in J at the cost of non-linearity of the corresponding mean-field equations. In a "homogeneous" case when $I=J$, $\Lambda_i=\Lambda$, $N_i=N$, $K_i=K$, $\mu_{ii}=\mu$, $\mu_{ij}=\nu<\mu$, $q_{ij}=(1-q_0)/(J-1)$, $i,j=1,..,I$; $i\ne j$, the mean-field system reduces to a single equation considered in Section IV.

III. A Mean-field Approximation

Introduce dynamic variables $\delta_j=1$ if all servers of class j are occupied: $X:\sum_i x_{ij}=N_j$, and $\delta_j=0$ otherwise. In a particular case of static resource allocation $K_i=1$, $i=1,..,I$, it is known [10] that due to Markov process $X(t)$ reversibility, steady-state random variables δ_j are jointly statistically independent for $j=1,..,J$, i.e.,

$$P(\delta_1,..,\delta_J)=\prod_{j=1}^{J}[p_j^{\delta_j}(1-p_j)^{1-\delta_j}], \quad (3)$$

where steady-state probability of all servers of class j being occupied is $p_j = P(\delta_j=1)$.

Generally, in a case of dynamic resource sharing: $K_i \ge 2$, Markov process $X(t)$ is not reversible, and thus (3) does not hold exactly. We view (3) as a simplifying assumption, which originated in statistical physics as "mean-field approximation", and then has been successfully used for performance modeling of various "complex systems" comprised of a large number of components. Accuracy of this approximation has been confirmed by simulations and rigorously proven in an asymptotic sense under much weaker assumptions than reversibility of the underlying Markov process $X(t)$. These results indicate that in a case of dynamic resource sharing scheme considered in this paper, factorization (3) approximately holds for $K_i \ge 2$ when the number of server classes J is large, i.e., $J>>1$, and the aggregate demand on each class of servers is the sum of a large number of "small" streams of different classes ensuring that the aggregate stream is approximately a Poisson process.

The mean-field approximation (3) leads to the following expression for the overflow probabilities $p_j \approx \tilde{p}_j$:

$$\tilde{p}_j = \frac{1}{\tilde{Z}_j} \sum_{n_{1j}+..n_{Ij}=N_j} \prod_{i=1}^{I} \frac{\tilde{\rho}_{ij}^{n_{ij}}}{n_{ij}!}, \quad (4)$$

where $\tilde{Z}_j$ are the normalization constants, the "effective" average loads are

$$\tilde{\rho}_{ij} = \rho_{ij} q_{ij} [1-(\theta_i q_{ij})^{K_i}]/(1-\theta_i q_{ij}) \quad (5)$$

and the probability that a class i request attempts to access a server of a class $k \ne j$ and this attempt is unsuccessful is $\theta_i = \sum_{k \ne i} q_{ik} p_k$. Note that if $K_i \ge 2$ then $\tilde{\rho}_{ij} \ge \rho_{ij}$ due to repeated attempts by requests to access service. Substituting (5) into (4) we obtain a closed system of J non-linear fixed point equations for vector of overflow probabilities $\tilde{p}=(\tilde{p}_j)$:

$$\tilde{p}_j = \varphi_j(\tilde{p}) \quad (6)$$

$j=1,..,J$. After solving system (6), the resource allocation scheme performance can be determined from (1)-(2).

In a homogeneous case, assumption $\tilde{p}_j = \tilde{p}$, $j = 1,..,J$ reduces system (6) to a single fixed-point equation

$$\tilde{p} = \varphi(\tilde{p}) \qquad (7)$$

where function

$$\varphi(\tilde{p}) = \frac{\rho^N}{Z(\tilde{p})} \sum_{n=0}^{N} \frac{1}{n!(N-n)!} \left(\frac{1-\tilde{p}^K}{1-\tilde{p}} \rho\theta \right)^n , \qquad (8)$$

$Z(\tilde{p})$ is the normalization constant, $\rho = \Lambda/\mu$, and parameter $\theta = \mu/\nu \geq 1$ quantifies the overhead associated with resource pooling, e.g., due to the communication delay.

IV. Performance of Homogeneous System

Figure 1 shows solution of fixed-point equation (7)-(8).

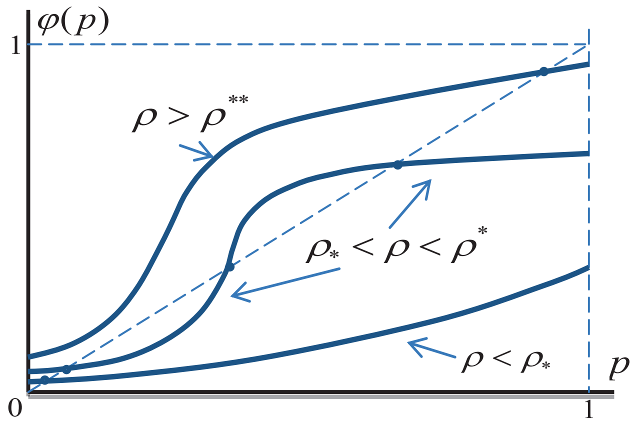

Figure 1. Solution of the fixed-point equation (6)

For sufficiently low and high exogenous loads $\rho < \rho_*$ and $\rho > \rho^*$ equation (7)-(8) has unique globally stable equilibria p_* and p^* respectively. For intermediate load $\rho_* < \rho < \rho^*$, sufficiently large number of servers in each class N, and sufficiently high level of resource sharing characterized by parameter K, equilibria p_* and p^* coexist as locally stable, and are separated by an unstable equilibrium $\tilde{p}$. Critical loads ρ_* and ρ^* depend on the system parameters N, K, θ. In particular, $\rho_*|_{K\to\infty} = 1/\theta$ and $\rho^*|_{K\to\infty} = 1$. Following the accepted practice, e.g., see [4]-[6], we interpret globally stable solutions as describing stable system equilibria, and locally stable solutions as describing metastable system equilibria. Note that the metastability is the result of the positive feedback in the effective load $\tilde{\rho}$ due to repeated attempts by requests to access service for $K \geq 2$.

Figure 2 shows persistent revenue loss (1) as a function of "slowly" changing exogenous load ρ.

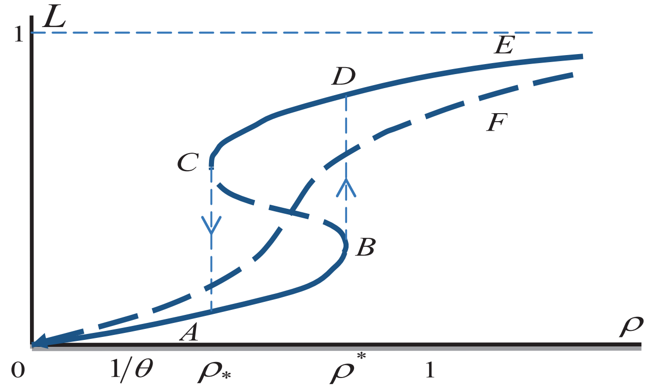

Figure 2. Loss L for "slowly" changing exogenous load ρ.

As load ρ "slowly" increases, the portion of lost revenue L follows curve OABDE. As ρ "slowly" decreases, the loss L follows curve EDCA0. Curves 0A and DE correspond to the globally stable "normal" and "congested" system equilibria in cases of light: $\rho < \rho_*$, and heavy: $\rho > \rho^*$ loads, respectively. Branches AB and CD correspond to the coexisting "normal" and "congested" metastable system equilibria respectively, and branch BC corresponds to the unstable system equilibrium an a case of intermediate load: $\rho_* < \rho < \rho^*$. Note that discontinuities at the critical loads ρ_* and ρ^* as well as hysteresis loop ABDCA indicate discontinuous, i.e., the first order phase transition.

Direct solution of equation (7)-(8) reveals that resource sharing is either beneficial, $K_{opt} = \infty$, or harmful, $K_{opt} = 1$, depending on whether the system is in the "normal" or "congested" equilibrium respectively. Thus, in the case of sufficiently light ($\rho < \rho_*$) or sufficiently heavy ($\rho > \rho^*$) exogenous load, the optimal level of resource sharing is unambiguous. The situation is more complicated in a case of intermediate exogenous load: $\rho_* < \rho < \rho^*$, when "normal" and "congested" system equilibria coexist as metastable. Figure 3 sketches the system phase diagram in the parameter space $(\rho, 1/K)$.

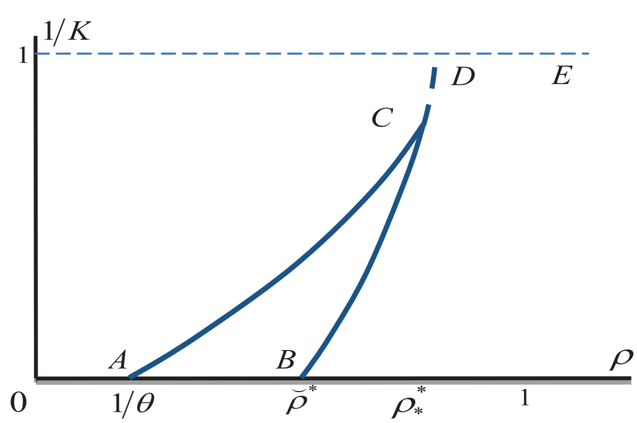

Figure 3. System phase diagram.

Inside region ACBA, the "normal" and "congested" system equilibria coexist as metastable, while outside of this region the system has an unique stable equilibrium. The optimal $K = K_{opt}(\rho)$ follows 0ABCDE assuming that system remains in the "normal" metastable equilibrium.

Since curve BC lies on the boundary of the stability region of the "normal" metastable system equilibrium, the system is likely to transition to the "congested" equilibrium through the process of cascading overload. This instability of the optimal dynamic resource sharing $K = K_{opt}(\rho)$ creates an inherent trade-off between equilibrium system performance and risk of cascading overload. The risk can be reduced by reducing K below $K_{opt}(\rho)$ in the region when the "normal" and "congested" metastable equilibria coexist. This risk reduction and associated increase in the system loss rate L constitute the systemic risk/benefit trade-off.

V. Perron-Frobenius Measure of Systemic Risk

We start by linearizing the mean-field system (6) about the "normal" equilibrium $\tilde{p}_* = (\tilde{p}_{*j})_{j=1}^J$ assuming its existence:

$$\delta\tilde{p}_j = \sum_{i \neq j} a_{ij}(\tilde{p}_*)\delta\tilde{p}_i \qquad (9)$$

where $a_{ij}(\tilde{p}_*) = \left(\partial\varphi_i(p)/\partial p_j\right)_{\tilde{p}=\tilde{p}_*}$ and $\delta\tilde{p}_j = \tilde{p}_j - \tilde{p}_{*j}$. We assume matrix $A(\tilde{p}_*) = (a_{ij}(\tilde{p}_*))_{i,j=1}^J$ to be non-negative: $a_{ij}(\tilde{p}_*) \geq 0$, and irreducible.

According to the Perron-Frobenius theory [7], these assumptions guarantee that matrix $A(\tilde{p}_*)$'s spectral radius γ is an eigenvalue, and the corresponding eigenvector has positive components. It can be shown that the "normal" equilibrium $\tilde{p}_*$ of mean-field system (6) is locally asymptotically stable if $\gamma < 1$, and unstable if $\gamma > 1$. Since equilibrium $\tilde{p}_*$ loses stability as Perron-Frobenius eigenvalue γ crosses point $\gamma = 1$ from below, it is natural to quantify the system stability margin and risk of cascading overload by

$$\Delta = 1 - \gamma \qquad (10)$$

In the rest of this Section we demonstrate that (10) also quantifies risk of dynamic instability due to fluctuating load.

Let $g = (g_j)$ be the normalized right Perron-Frobenius eigenvector of matrix $A(\tilde{p}_*)$, i.e., $\sum_j a_{ij}(\tilde{p}^*)\delta g_j = \gamma g_i$, $\sum_j g_j = 1$. Since linearized mean-field system (9) loses its stability along eigenvector g, it is natural to assume that essential fixed-point dynamics (6) occurs in this one-dimensional subspace. We describe this dynamics by assuming in (6) $\tilde{p} = \tilde{p}_* + \delta\tilde{p}$, where $\delta\tilde{p} = J\alpha g$ and α is the average portion of overloaded server groups:

$$J\alpha g_i = J\sum_j a_{ij}(\tilde{p}_*)\alpha g_j + 0.5 b_j^{(2)}(\tilde{p}_*)(J\alpha)^2 + o(\alpha^3) \qquad (11)$$

as $\alpha \to 0$, where $b_j^{(2)}(p_*) = \partial^2\psi_j(\tilde{p}_* + xg)/\partial x^2 |_{x=0}$ and $\psi_j(p) = \varphi_j(p) - \tilde{p}_{*j} - \sum_i a_{ij}(\tilde{p}_*)\delta p_i$. Adding equations (11) for $j = 1,..,J$ we can solve for α as $\Delta \to 0$:

$$\alpha = \frac{\Delta}{2J}\left(\sum_j b_j^{(2)}(\tilde{p}_*)\right)^{-1} + o(\Delta) \qquad (12)$$

Expression (12) quantifies dynamic stability margin of the "normal" system equilibrium $\tilde{p}_*$: if due to unavoidable fluctuations in the instantaneous load, the portion of the overloaded server group exceed its "normal" average by margin (12), the system goes to the "congested" equilibrium. The critical margin (12) is proportional to the "static risk measure" (10) with proportionality coefficient which can be evaluated numerically under mean-field approximation, or estimated by online measurements.

VI. Conclusion and Future Research

This paper, which represents work in progress, suggests that Perron-Frobenius theory applied to mean-field system model may provide an adequate and tractable framework for assessing and managing risk of cascading overload in complex resource sharing systems. The applicability of the Perron-Frobenius theory is due to non-negativity of the corresponding linearized system, which is a result of the positive congestion feedback: congestion in some system components overflows to other components. Future work should address practicality of the proposed systemic risk measure at the system design and the operational stages. An intriguing possibility is online measurement of the corresponding Perron-Frobenius eigenvalue as system approaches its instability point.

References

[1] J. Tsitsiklis and K. Xu, "On the power of (even a little) centralization in distributed processing," *ACM SIGMETRICS 2011*, San Jose, CA, 2011.

[2] M. Bramson, Y. Li, and B. Prabhakar, "Randomized load balancing with service time distribution," *ACM SIGMETRICS 2010*, 2010.

[3] J. Hu and B. Hajek, The supermaket game, *IEEE ISIT,* 2012

[4] V. Marbukh, "Loss Circuit-Switched Communications Networks: Performance Analysis and Dynamic Routing, Queueing Systems: Theory and Applications, vol. 13, pp. 111-141, 1993.

[5] R. J. Gibbens, P. J. Hunt, and F. P. Kelly, Bistability in communication networks, Disorder in physical systems, Oxford Sci. Publ., Oxford Univ. Press, New York, 1990, pp. 113–127.

[6] V. Marbukh, "Risks/benefits of selfish network operations & management in dynamic environment," *13th IEEE/IFIP Network Operations and Management Symposium* (NOMS 2012), Maui, HI, 2012.

[7] A. Berman, R.J. Plemmons, Nonnegative Matrices in the Mathematical Sciences, 1994, SIAM.

[8] S. Hanly, Congestion measure of congestion in CDMA networks, *IEEE Transactions on Communications*, Vol. 47, Issue 3, 1999, pp. 426-437.

[9] D. O'Neill, D. Julian, and S. Boyd, Seeking Foschini's Genie: Optimal Rates and Powers in Wireless Networks, *IEEE Trans. on Vehicular Technology*, 2004.

[10] F.P. Kelly, Reversibility and Stochastic Networks, Cambridge University Press, 2011.