
Enterprise Data Center Globality Measurement

Ruairí de Fréin[†] and Joel Pfaff^{††} and Thomas Paré^{††}

[†]KTH - Royal Institute of Technology, Stockholm,
Sweden

^{††}Amadeus SAS, Sofia Antipolis,
Nice, France

web: <https://robustandscalable.wordpress.com>

in: 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing, to appear. See also `BIBTEX` entry below.

BIB_TE_X:

```
@article{rdefrein15DASCGlobality,  
author={Ruairí de Fréin† and  
Joel Pfaff†† and  
Thomas Paré††},  
journal={2015 IEEE International Conference on Computer and Information Technology;  
Ubiquitous Computing and Communications;  
Dependable, Autonomic and Secure Computing;  
Pervasive Intelligence and Computing, to appear},  
title={Enterprise Data Center Globality Measurement},  
year={2015},  
pages={9},  
month={Oct.}, }
```

© 2015 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.



Enterprise Data Center Globality Measurement

Ruairí de Fréin

KTH, Royal Institute of Technology, Sweden
Email: rdefrein@gmail.com

Joel Pfaff

Amadeus SAS
France

Thomas Paré

Amadeus SAS
France

Abstract—Measurement of the globality of On-Line Transaction Processing (OLTP) workloads in Enterprise Data Centers is considered. Providing OLTP workload isolation (application, services and databases) for performance-sensitive enterprise workloads, so that activity in one workload cannot interfere with another, remains a challenge. We demonstrate that traditional aggregate OLTP workload globality measurement frameworks can generate mis-leading globality measures. We propose a higher-order globality measurement framework which addresses this problem. We derive two high dimensional structured measurement matrices, namely a *template* and *measurand* matrix, with special spectral properties, which account for globality measurement 1) boundedness; 2) programmability; 3) multiplicity; 4) relativity; 5) spatial correlation and 6) the appropriate sensitivity of the measure to changes in the distribution of the workload. We demonstrate that these properties are exhibited by the new measure by ordering OLTP workloads by their globality measure. We evaluate the measure using a stochastic layered block model for data center topology and OLTP workload generation and demonstrate that it is consistent.

I. INTRODUCTION

Cloud networking promises cost-efficient and reliable service delivery across data communications networks [1]. The location of services and the potential of virtualization of hardware [2] and software is stressing communication network and protocols [3]. The computing aspects of cloud technologies are steadily advancing; however, until recently lower attention has been paid to the networking aspects of this problem, e.g. monitoring [4] and service level prediction [5]. A crucial design goal is to localize or isolate workloads in modern data-centers. We address the problem of measuring the globality, the extent of localization, of On-Line Transaction Processing (OLTP) workloads in Enterprise Data-Centers (EDC).

Real-world scenario: The increasing emphasis on elasticity and flexibility, driven in part by affordability constraints, is causing EDCs to evolve. However, due to legacy design issues, messaging between different layers of an EDC is unavoidable in many current EDC deployments [6]. The advent of Software Defined Networking (SDN), exemplified by Meridian [7], promises gains in the programmability of EDC Networking, which may give rise to successful strategies for service and database-entry migration. But a major stumbling block is to quantify “how much” the workload globality has been improved. Consider the case of Amadeus SAS who are the leading service provider to the travel industry. Amadeus operates all software out of its fully-instrumented (which makes

off-line reconstruction of OLTP traces and this study possible, cf. Section II) EDC as a *community solution*, sharing hardware and software for all clients using a given product. Examples of OLTP workloads include the ordered sequence/tree of potentially hundreds of messages invoking data-base queries and other subsystems, residing on hundreds of machines in different domains in a large EDC, involved in flight bookings, payment and management etc. One of the many challenges in this scenario is to provide very high numbers of transactional flows, without scheduled downtime, for critical applications (such as airport operations). Globally the Amadeus EDC is hit by $> 10^5$ transactions-per-second (tps), running many complex stateful transactions. Resource management and careful design of system topologies is of the highest importance in order to build commercially viable, highly available systems. SDN technologies offer one solution. However, frameworks for measuring the level of improvement that may be garnered by SDN do not exist. The case for SDN would be considerably enhanced if a demonstrable reduction in the globality of workload in an operational EDC could be measured, and comparison of many measurements be performed in a consistent way. The globality measure of reconstructed OLTP traces is a key indicator of EDC performance and may be computed using the traces, sequence of machines, services and switches visited.

At the heart of globality measurement lies the notion that EDCs and the OLTP workload they support should submit to an ordering (from low globality to high globality), and that future SDN enabled solutions should aim to achieve a low globality. This ordering should induce a globality measure which provides a consistent mapping from the high dimensional space of EDC and OLTP statistics (arising from $> 10^5$ tps) to the real number line, between 0 and 1, where 0 indicates locality and 1 indicates globality. We propose a set of properties a consistent globality metric should have for the types of OLTP workload Amadeus support. Then we provide a measurement algorithm which satisfies many of these properties. In short, we provide a *consistent* method for measuring the globality of OLTP workloads on EDCs.

Globality vs Centrality: We exploit the spectral properties of trees [8]. EDC OLTP globality measurement provides a fresh application domain for many of these techniques. Centrality identifies the most important vertices in a graph. Popular approaches include: Degree Centrality [9] (the number of links incident on a node); Closeness centrality (the more central a node is, the lower its total distance to all other nodes); Freedman’s Betweenness Centrality [10] (quantifies the number of times a node acts as a bridge along the shortest path between two nodes); Eigenvector Centrality [11] (exploits the idea that connections to high-scoring nodes contribute more to the score of the node than connections to low-scoring

Dr de Fréin is affiliated with TSSG, Waterford Institute of Technology, Ireland. This work was supported by 1) an ELEVATE Irish Research Council International Career Development Fellowship, EOLAS, co-funded by Marie Cure Actions award: ELEVATEPD/2014/62 and 2) the European Commission via the FP7 IAPP project SOLAS (grant no. 612480).

nodes); and so on. In general, a centrality measure that is optimal for one application is sub-optimal for another, which gives rise to a myriad of definitions [9]. **To measure globality of EDC OLTP workloads we need a centrality-like measure that allows us to specify that the use of certain vertices and edges in a graph is more costly than other vertices and edges.** Specifically, (1) we need to be able to define where the core (and other layers) of the EDC lie(s) in the graph, and (2) that the usage of core (and other selected) links is undesirable. To this end, we extend the notion of centrality, and call this new measure *globality*. Globality is measured using a template and a measurand matrix, which capture the cost of using different network resources. The centrality measures listed above do not provide this flexibility. Globality measurement gives a good indicator, from the *operational analytics* perspective [12], of the extent of a work-flow’s coverage of the EDC. For *hypothesis testing* and *failure scenario simulation* [13], it provides a performance metric for the resilience of the EDC. The ability to measure the globality of OLTP workload provides a crucial step towards predicting failure scenarios using Machine Learning [14], [15].

Contributions: We introduce a 1) matrix-structure for encapsulating OLTP/EDC statistics which are suitable for globality measurement; 2) framework of properties for comparing globality measures. These properties are of independent interest because as the needs of EDCs change, so will globality measures. These properties may form the basis for globality measurement of future applications. 3) We introduce an algorithm for computing relative globality measures. This paper does not report on an evaluation of the new globality measure using real traces from Amadeus’ EDC because: 1) this EDC generates 10^5 tps and thus the results from the 100 significant OLTP traces, in terms of impact on the globality of the EDC’s OLTP, occurring in a 5 minute sampling period would be lost in the noise of presenting the results for all of the transactions occurring in a 5 minute interval; 2) the aim of this paper is to introduce a new globality measure and to demonstrate that other measures are in-appropriate and when this is the case. We present a number of these cases here, in a simplified form. In Section II we describe previous work on OLTP-EDC cross-layer modeling. In Section III we describe the test topology used for our model. In Section IV we review an aggregate globality measurement framework. We define global and local workloads and give examples of OLTP workload types. Section V describes the matrices used for higher-order globality measurement and Section VI evaluates them.

II. RELATED WORK

Globality measurement has been examined before in [16] although the authors of [17] note that the lack of publicly available traces has hampered progress for EDC OLTP workloads. EDC OLTP workloads of Amadeus’ form are unique and are not publically available due to the correlation between high globality and commercially sensitive OLTP delay. Traces arising from Map-Reduce-like [18] workloads are not suitable substitutes for performance evaluation here as EDC OLTP workloads generally consist of many small tree-like data transfers as opposed to a few large data transfers. Traditional methods measure globality in an aggregated way –we show that important correlation behaviour may be lost and also that existing methods may not be sensitive enough with respect to

the OLTP and EDC being measured. In other words, if two workloads are to be compared, does one aspect of one workload *unfairly* impact the globality measure over measurement of the second workload?

This work is underpinned by a Cross-layer Performance Analysis (CPA) system, [16], that fuses and analyzes trace data collected from multiple levels of the software stack in an operational EDC. Data was collected from Amadeus’ production EDC in [16]. CPA was used to assess how a topology-aware load balancing procedure, coupled with carefully assigned VLANs [19], could eliminate most of the unnecessary OLTP network traffic in the EDC. We go significantly beyond this as we address the question of how to measure the globality of OLTP traffic, given the CPA system in [16] and an exemplar EDC architecture (which is based on an architecture which is prevalent in most of today’s EDCs [20]). EDC network, Enterprise Service Bus (ESB) and application state discovery, e.g. connectivity information, node configuration such as IP and MAC addresses and VLAN membership, is not described here. We refer the reader to [16] for information about the time taken to construct OLTPs. In summary, the CPA system consists of an EDC topology model which is enriched with information about the application and ESB placement. A hierarchical structure consisting of transaction trees, transactions, messages and host-to-host hops is used to capture the OLTP workload. This data-set is queried using a MapReduce framework. We assume that the network paths between two hosts has been inferred using [16], along with other topology and configuration queries and we focus on measurement.

OLTP globality measurement is not directly applicable to Cloud Data Centers (CDC). Software Oriented Architecture (SOA) EDCs carry three types of traffic: OLTP, which is typically the main source of traffic; Batch, which arises from applications such as MapReduce; and File Distribution, which arises due to running updates for all machines of a certain type. CDCs are typically designed for batch-like workloads [6]. Typical OLTP workloads stress the EDC with many small transactions as opposed to the fewer larger tasks found on CDC. This claim is supported by examining some Google traces [21]. Recent work on optimizing Virtual Machine (VM) fairness by modeling data-center bandwidth allocation as a cooperative game does not explicitly consider the problem of reducing the number of packets that traverse the core [22].

What causes global OLTPs? In many cases global transaction trees can be attributed to the multi-layer network architecture of traditional EDCs [6]. This conclusion is supported in [23]. Many previous studies of data center traffic have focused on CDCs [24], [25], [26], [27], [28], batch-like MapReduce workloads, and not OLTP workloads. Globality of OLTP workloads is important as many companies maintain OLTP oriented EDCs and are unwilling to move their workload to the cloud due to cost, privacy and security reasons. The lack of publically available EDC OLTP traces is another factor that contributes to the under-representation of OLTP workloads in these studies [17]. The results presented in this paper were generated by constructing OLTP workloads using TPC-C’s OLTP-like workload generating models [29]. The TPC-C benchmark is a popular yardstick for comparing OLTP performance on various hardware and software configurations [30], [31]. We generalize the TPC-C benchmark and Stochastic

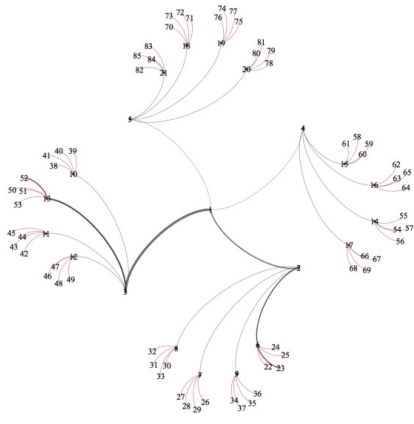


Fig. 1. A simple symmetric EDC topology is illustrated. Numbers indicate graph vertices, which represent switches, web-servers, application-servers in the EDC. This figure also illustrates that the aggregate usage of the EDC is sub-optimal. Dark colors and thick links illustrate that the traffic is heavy on certain links ($\{1, 3\}, \{1, 2\}, \{3, 13\}$ etc.), and that this traffic causes OLTP workload to traverse the core of the EDC. [Color figure is available.]

Block-model in [32] and introduce a *Simple Stochastic Layered Block Model for Data Center topology and OLTP workload Creation* (SSLBMDC) which allows us to generate ensembles of workloads that traverse the core with a tunable probability, and thus we can thoroughly evaluate our globality measure. Globality measurement is important: HDFS was designed with rack-awareness in mind [33] to reduce traffic on higher level switches; topology aware load balancing procedures have been designed for peer-to-peer networks [34]. Machine placement has been identified [6] as a crucial factor in ensuring that the sum of traffic does not saturate any links.

III. SYSTEM MODEL

Our EDC topology data-set is represented as an unoriented graph $G = (V, E)$ where V is the set of vertices in the graph, for example a switch, (web-) server, application server, Enterprise Service Bus (ESB), middle-ware machine, mainframe machine and data-base machine. There are N vertices (machines/servers) in the graph (EDC). Communication between this set of vertices is captured by the set of edges, E , for example, the path between two hosts. The bandwidth usage of each edge is typically determined by the set of ESBs, which is the primary source of network traffic. In Fig. 1 $N = 85$, and the set of nodes $\{1, 2, 3, 4, 5\}$ corresponds to the Core, DMZ, MainFrame, WAN and Production switches. The remaining nodes are applications, access switches etc.

Problem: Given the plurality of machines involved in a single root request, it is likely that some of the machines required to process the transaction are located in different layer-2 network domains. This is further exacerbated by the fact that most current EDCs employ VLANs [19] to fragment physical layer 2 domains into several mutually isolated layer-2 domains. Further fragmentation is achieved by Private VLANs [35]. This scenario is undesirable. EDC managers would prefer that workloads are isolated within separate partitions of the graph G . The goal of this paper is to measure the globality of OLTP workload; the extent to which workload spans different partitions. We develop the ability to quantify the dependencies between OLTP workload and the physical EDC it is running

on. From the perspective of workload isolation, the underlying problem is the hierarchical nature of the EDC.

Architecture: The first layer of the hierarchy of our exemplar EDC consists of a central core switch and four distribution switches, production (PRD), DMZ, Mainframe and WAN. They form a layer-3 network and are connected by a set of edges. The next layer of the hierarchy consists of layer-2 links. For example, the PRD switch is connected with access switches, each with some number of ports. Finally, the layer-2 network is sub-divided into a two-levels. On the first level are the devices directly connected to the access switches, i.e. blade switches, VMWare virtual switches and some machines [36], [37]. On the second level are the machines that are connected through the blade or virtual switches to the access layer.

EDC Tree: A Tree T in the EDC is an undirected graph in which any two vertices are connected by exactly one simple path. Any connected graph without simple cycles is a tree. Many distribution switches have the ability to compute trees to aid routing etc. The EDC can therefore be modelled as a family of overlaid trees $T_k = (V_k, E_k)$ where the index k denotes the unique identifier of the tree. This family is denoted

$$\mathcal{T} = \{T_k | V_k \subseteq V \text{ and } E_k \subseteq E, \quad k = 1, 2, \dots, K\} \quad (1)$$

Fig. 1 has one unique tree that traverses all nodes $K = 1$.

OLTP Overlay: A transaction tree W_l , indexed by l , represents an ordered set of messages which are spawned by a client request. Similar to the EDC, it consists of a set of machines $M \subseteq V$ which are traversed by the transaction tree, and a set of edges $P \subseteq E$ which are traversed, $W_l = (M_l, P_l)$. The cardinality of the set of machines traversed is denoted L . The number of messages in a tree is T_l . From a measurement perspective the EDC is effectively a set of trees \mathcal{T} ; and OLTP traffic can be viewed as a tree which is overlaid onto the EDC underlay. This dependence is made explicit: each transaction tree is supported by a member of the set \mathcal{T} .

$$\exists T_k \text{ such that } W_l \subseteq T_k \implies M \subseteq V_k, P \subseteq E_k \quad (2)$$

This is sensible as W_l would not have existed had the underlying substrate T_k not existed to support this workload. This implies that given some globality function $g(\cdot)$ which takes as its input an OLTP workload and EDC we can compute the set of *conditional* globality scores $g(W_l|T_k)$. In other words, the globality of the workload tree W_l given T_k where T_k has the property that $W_l \subseteq T_k$. Moreover, we are interested in the worst-case globality score, and thus, we compute

$$g(W_l|G) = \max_k g(W_l|T_k). \quad (3)$$

IV. GLOBALITY MEASUREMENT

We discuss aggregated globality measurement and introduce a higher dimensional measure. We define global and local workloads using examples.

A. Aggregate Measurement Framework

The EDC in Fig. 1 consists of access-to-access (aa), distribution-to-access (da) and distribution-to-core (dc) links. Each transaction tree W_l may be characterized by the contribution its messages make to these three types of links. The

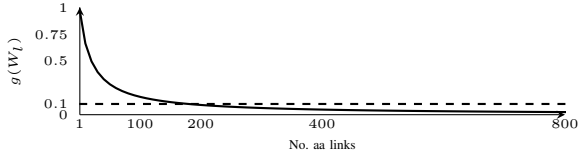


Fig. 2. If one dc or da is used in a transaction tree, how many aa do we need to achieve a low globality, $g(W_l) = 0.1$? For a globality measure $g(W_l) = 0.1$ the number of aa links required in a transaction tree is ≈ 200 . For a globality measure $g(W_l) = 0.2$ the number of aa required is 100. Although the globality measure takes values in the range $0 < g(W_l) < 1$, da and dc links *unfairly* dominate the measure.

Weighted Good-put contributed by the transaction tree W_l is computed as follows for each type of link $\lambda = \{aa, da, dc\}$.

$$w_\lambda = \sum_t s_t \hat{\alpha}_\lambda \delta(t, \lambda) \quad (4)$$

The parameter $\hat{\alpha}_\lambda$ is the Over Subscription Ratio (OSR) [38] for link type $\lambda \in \{aa, da, dc\}$, s_t is the message size in bytes, and $\delta(t, \lambda)$ is an indicator function which equals 1 if the t -th message was sent over link type λ . We compute this sum over the ordered messages t in the transaction tree W_l for all λ . The role of $\hat{\alpha}_\lambda$ is to build the notion that different layers have different costs associated with them into the globality measure, for example, OSR. Cisco recommend in [38] that da links are typically oversubscribed by 20:1 and that the dc links are oversubscribed by 4:1. In [16] relative (to the access layer) OSRs are computed: the da layers are typically oversubscribed by 20:1 and the dc layers are oversubscribed by 80:1. An aggregate globality measure, $g: \mathbb{R}^3 \mapsto \mathbb{R}^1$, of OLTP W_l is

$$g(W_l) = \frac{w_{da} + w_{dc}}{w_{aa} + w_{da} + w_{dc}}. \quad (5)$$

Remark: This measure is reasonable because a transaction tree that predominantly traverses dc and da links has a globality $g(W_l) \approx 1$. We make the following observations: 1) The measure is EDC independent. The aggregates of three terms $\{w_{aa}, w_{da}, w_{dc}\}$ are used to evaluate globality. This triple may assume many values and give the same globality score. 2) The sensitivity of the score is heavily reliant on the dominant term in the triple. If OSRs are used $\hat{\alpha}_{dc} \gg \hat{\alpha}_{da} \gg \hat{\alpha}_{aa}$ and thus in $g(W_l)$ the relation $w_{aa} \ll w_{da}$ and $w_{aa} \ll w_{dc}$ holds. If transaction tree W_l has one dc or da link, most transaction trees evaluated using this measure will give a high globality score $g(W_l) \approx 1$. This is because the numerator is normalized by the sum $w_{aa} + w_{da} + w_{dc}$ to ensure that the measure is not adversely affected by comparing transaction trees of different lengths. Aggregate globality scores are likely to be un-informative.

B. Properties of Global OLTP

Sensitivity: Fig. 2 demonstrates that for any practical EDC, any transaction tree with a da or dc link will be global according to (Eqn. 5). This is because hundreds of aa links are required to balance-out the measure to get $g(W_l) \approx .1$. It is unreasonable to assume that transactions trees will have enough aa links in order counteract the effects of the da and dc links. Indeed, many existing EDCs will have a number of da links. We formalize the description of these problems by introducing a set of properties. They include correlation and balance, behavior which is lost by the triplet used above. What

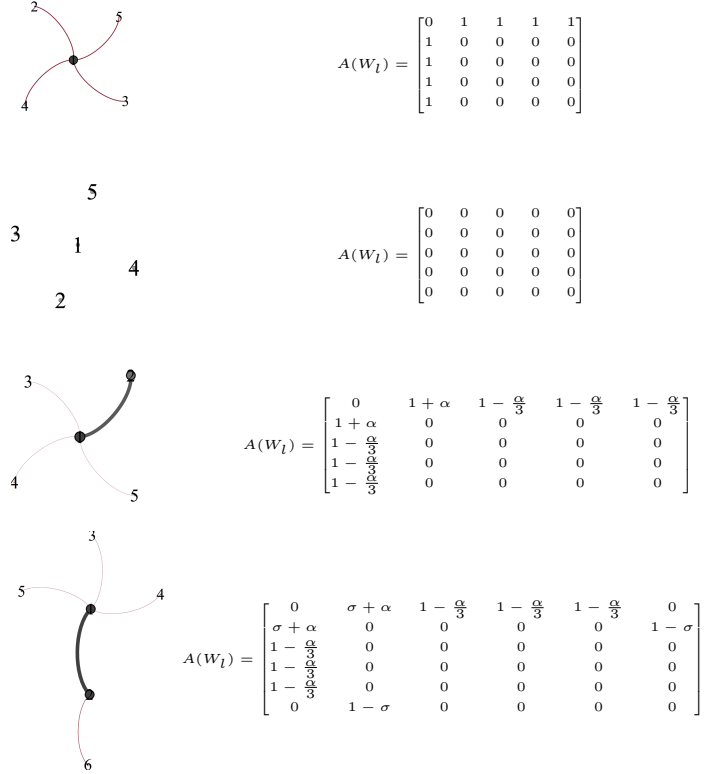


Fig. 3. Workload-type illustrations: The presence of traffic is illustrated by an edge on the graphs, and a positive entry in the adjacency matrices on the RHS. Thick dark (grey-black) links indicate that the load is heavy. Large vertex sizes indicate that the weighted degree (by the number of incoming messages) of the vertex is large.

is required is a structure that accounts for the higher order dependencies: correlation between the usage of links (used or not used); relative usage of links (to the worst OLTP workload scenario); locality of EDC usage; and finally, cognizance of the symmetries in EDC usage (correctly determines if two scenarios are symmetric or mirrored versions of each other).

Structure: Many of the symmetry and correlation issues are addressed by considering the adjacency matrices of the EDC tree set \mathcal{T} . We denote the adjacency matrix of the EDC's k -th tree, $A(T_k)$. Pairs of the vertices of T_k are assigned a weight $a_{i,j}$ with the properties $a_{i,j} = a_{j,i}$; $a_{i,j} \neq 0$, if and only if i, j are adjacent in T_k ; $a_{i,j} \geq 0$ if $i, j \in V_k$ the vertex set of V . The degree of a vertex in the EDC's k -th tree is $d(v_i) = \sum_j a_{i,j}$. A diagonal matrix composed by placing these degrees along the diagonal is denoted $D(T_k)$. We introduce the properties, **p1-5**, a globality measure $g(\mathcal{W}|\mathcal{T})$ should have, where \mathcal{W} is the family of transaction trees in the workload under test and \mathcal{T} is the EDC. The Laplacian of the k -th tree is defined as

$$L_k = L(T_k) = D(T_k) - A(T_k). \quad (6)$$

p1 Bounded: $g(\mathcal{W}|\mathcal{T})$ should lie in a bounded interval

$$0 \leq g(\mathcal{W}|\mathcal{T}) \leq 1. \quad (7)$$

It is important that a globality measure is just one number as opposed to a list of performance indicators because in large-scale deployments this measure may well be communicated between different parts of an organization. A globality

$g(\mathcal{W}|\mathcal{T}) = 1$ implies that the traffic is global. A measure $g(\mathcal{W}|\mathcal{T}) = 0$ implies that the workload is completely local.

p2 programmable: $g(\mathcal{W}|\mathcal{T})$ should be amenable to different parametrization. An example of a parametrization is the encoding of Cisco’s OSR into the measure, or the variation of packet sizes, s_t , into the measure.

p3 0-multiplicity: globality measurement should be able to indicate that there are up to N connected but isolated OLTP work-flows in the EDC topology.

p4 relative: the EDC, G must be fixed if two different measurements are to be compared. This claim is justified as it is unlikely that two large organizations will have exactly the same infrastructure. What is important is how globality compares with previous measures on their EDC.

p5 spatial correlation: The globality measure must take into account the spatial correlation of link usage in the EDC. In a 2-link EDC, if one transaction tree uses one link 99% of the time, and the other link 1% of the time, the globality measure should be able to distinguish between this workload and when the usage is evenly balanced. Similarly, if there is a discrepancy between message sizes, the measure should account for this.

C. Exemplar Workloads

Property 1 motivates the question, what is a completely global transaction tree on an EDC? We define the properties of a global OLTP workload. We then consider a set of workloads where the number of messages (times message size s_t) sent is the same for all, but the usage of links is non-uniform, where uniformity is parametrized by $\sigma, \alpha \geq 0$. In each example the number of messages is scaled to equal T . The adjacency matrix of the l -th workload, $A(W_l)$, is constructed by placing a count of the usage of each link in the corresponding entries of $A(W_l)$, e.g. $a_{ij} = a_{ji} = 4$ if link ij is traversed 4 times by a W_l , and $a_{ij} = a_{ji} = 0$ if ij is not traversed.

Global OLTP Workload: A global workload \mathcal{W} on the EDC corresponds to an ensemble of transaction trees which are 1) order-1 simple trees (one row or column of its adjacency matrix has $L - 1$ non-zero entries), including the core switch; and 2) where all links are equally weighted –they have the same weight (traffic) on each link from the core to the children nodes, where the weight can be the portion of traffic $T/(L - 1)$ on a link, the OSR, the bandwidth usage on a link, or the multiplicity of messages on a link. 3) When $A(W_l) \in \{0, 1\}$, W_l is an order-1 simple tree with $L - 1$ links. The corresponding eigenvalues of its Laplacian are $\lambda_1 = L, \lambda_2 = \dots = \lambda_{L-1} = 1$ and $\lambda_N = 0$. 4) Kirchhoff’s matrix tree theorem gives the number of trees in this workload as: $K = \frac{1}{L} \lambda_1 \lambda_2 \dots \lambda_{L-1} = L/L = 1$. 5) The fact that $\lambda_{L-1} = 1$ implies that the workload is connected. Fig. 3, row 1, illustrates the link usage (LHS) and workload adjacency matrix (RHS). When the traffic is fixed for all order-1 W_l , but the number of vertices used by the workload (L) changes, the maximum globality is when $L = 2$. Globality decreases as L increases.

Local OLTP Workload: A local L -vertex workload \mathcal{W} has no entries in the adjacency matrix between the core and the distribution switches. The number of trees in this graph is $K = L$ (cf. Fig. 3, row 2.). The number of messages in this case is 0 and not T (for illustration purposes).

TABLE I. ALGORITHM FOR COMPUTING WORKLOAD GLOBALITY

| | |
|----|--|
| 1. | Normalize workload: $C = 2(c-1) \frac{C}{\sum_{ij} C_{ij}}$ |
| 2. | Add workload to template: $\bar{C} = T(T_k) + C$ |
| 3. | Normalize to sum-to-one: $\bar{C} = \frac{\bar{C}}{4(c-1)}$ |
| 4. | Compute largest eigenvalue of the Laplacian of \bar{C} : $g = \max_k g(W_l T_k)$ |

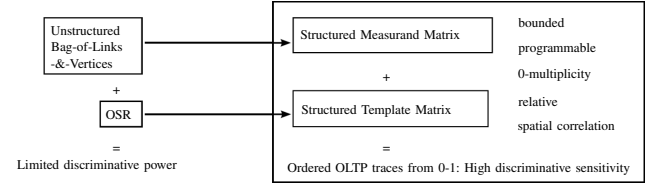


Fig. 4. The (LHS) aggregate globality approach is improved by considering a structure Measurand and Template matrix (RHS box). The idea of an OSR is generalized and encoded in a structured matrix form.

Nonuniform Workload: If the portion of messages on all links from the distribution switches ($\{2, 3, 4, 5\}$) to the core ($\{1\}$) is not uniformly equal, $G(\mathcal{W}|\mathcal{T})$ should be less than one. In Fig. 3, row 3, a portion $\frac{\alpha}{3}$ of the workload is subtracted from links $\{1, 3\}, \{1, 4\}, \{1, 5\}$ and added to link $\{1, 2\}$ causing imbalance in the workload when $\alpha > 0$.

Workload Appendages: If the distribution switches ($\{2, 3, 4, 5\}$) receive messages from access switches, the globality of the workload \mathcal{W} should be less than one. In Fig. 3, row 4, $(1 - \sigma)$ workload is added to link $\{2, 6\}$ by subtracting some portion of workload from $\{1, 2\}$. The globality should be less than zero for any value of $\alpha \geq 0$.

Centrality Measures do not Suffice: Consider the EDCs in rows 3 and 4 of Fig. 3. The parameters α, σ are chosen such that links $\{1, 2\}$ and $\{1, 3\}$ are heavily loaded compared to the rest. Both graphs are trees; any node could be the root, the *center/core* of the EDC. We know that the roles of vertices $1, 2, \dots, 5$ are to be the core and distribution switches, and the usage of these links is costly. If node 6 was to be considered the core switch by an (inappropriate) centrality measure – the link $\{6, 2\}$ is lightly loaded and less-costly to use than $\{2, 1\}$ – this would give a low globality score. In fact the score should be high as the traffic from distribution switch, 2, to the true core switch, 1, is costly. We desire a *globality* score for the workload, which incorporates knowledge of the roles of vertices $1, 2, \dots, 5$. Traditional centrality scores do not allow us to assign varying importance to different nodes in the EDC. In the next section we describe a *globality* score that allows us to encode the locations of costly links. This measure is cognizant of the fact that node 6 is not the core switch.

V. HIGHER ORDER GLOBALITY MEASUREMENT

We introduce our measurement algorithm and demonstrate that it meets each of the properties above (p1-p5). To this end we consider the properties of the weighted Laplacian of a tree. We introduce two special matrix measurement structures, an EDC template matrix and a workload measurand matrix, and an associated theorem. The theorem leverages the properties of the weighted Laplacian to bound the globality measure. To give this measure context Fig. 4 illustrates how our work extends previous measures. Instead of using the aggregate measure (Eqn. 5), the indiscriminate *bag-of-links-and-vertices* approach, we consider a structured measurand matrix which

gives us the five properties listed in Fig. 4. Fig. 4 also illustrates that instead of considering the OSR on each link in isolation (in a 1-dimensional way), we encode a similar concept in the form of a structured (2-dimensional) Template matrix. This leads to improved discriminate sensitivity –our new measure can detect fine-scale differences. Our new measure is a *higher order globality measure* as the structured template and measurand matrices allow us to consider higher-dimensional relationships, e.g. correlations, between vertices and edges of the workload.

Template: The template matrix is a special type of adjacency matrix. Its role is to act like an *irritation* function: a large entry in this matrix is used to encode the fact that the use of the corresponding link in the EDC to support OLTP workload has a high globality cost; a small entry encodes the fact that the use of this link has a small cost. The template is *programmable*, in the sense that we can pick-and-choose which links to penalize the workload for using. It is more flexible than an OSR as it allows us to encode spatial correlation behaviour (the OSR does not). We use the topology in Fig. 1 as the EDC reference, T_k , in a running example which demonstrates the construction of the template. There is 1 core and 4 distribution vertices, $c = 1 + 4 = 5$ and a total of $N = 85$ vertices in the topology.

Definition 1: The core and distribution sub-graph form a star tree $S_c = (V_c, E_c)$, a c -vertex sub-tree of the topology, e.g. $V_5 = \{1, 2, 3, 4, 5\}$ and $E_5 = \{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{1, 5\}\}$. In this sub-tree the $c - 1$ distribution vertices have 1 link. The parameter c is crucial as it plays a role in normalizing the workload; it depends on the selected template, in this case S_c .

Definition 2: The most global workload on this EDC is the star tree S_2 , e.g. V_2 includes node 1 and one other element of V_5 , and the associated link $E_2 = \{1, j\}$ where $j \in V_5$. The entire transaction occurs on this link. Each of the $c - 1$ distribution vertices roots an access network sub-tree consisting of $\gamma = (N - c)/(c - 1)$ vertices. These sub-trees are denoted $D_d = (V_d \subset V_k, E_d \subset E_k \text{ and } E_d \not\subseteq E_k)$, for $d = \{2, 3, 4, 5\}$.

Definition 3: The template matrix for the k -th ($K = 1$, the EDC topology here is a simple tree) EDC tree is constructed by weighting the sub-tree links of the distribution vertices

$$\bar{T}(T_k) = A(T_k) - A(S_c) \quad (8)$$

by β , where $\beta > 0$ but small, e.g. $\bar{T}(T_k)\beta$. We call this process *connecting the template*. Connecting the template is important as we want the first $L - 1$ eigenvalues of the template plus measurand to be greater than zero. The core to distribution links compensate for these small weights by subtracting the sum of the weights in each sub-tree from the associated distribution to core link

$$T(T_k) = \bar{T}(T_k)\beta - \gamma\beta A(S_c) + A(S_c) \quad (9)$$

We have formed a template adjacency matrix by using a global sub-tree $A(S_c)$ (cf. Fig. 3, row 1) and adding weighted (by $\beta \approx 0$) appendages for each of the sub-trees rooted by the distribution vertices (similar to Fig. 3, row 4). The expression for the template is easily generalized for unbalanced topologies (where γ_d is the size of the sub-tree of each distribution vertex). The choice of template is arbitrary, and therefore, allows the user to program (cf. **p2**) the metric by selecting an appropriate template/*irritation* function. We recommend that for T_k , β is approximately machine precision.

TABLE II. NUMBER OF DIFFERENT DC LINKS SHARING THE OLTP WORKLOAD ($\beta = 10^{-6}$)

| dc links used | {1, 2} | {1, 2}, {1, 3} | {1, 2}, {1, 3}, {1, 4} | {1, 2}, {1, 3}, {1, 4}, {1, 5} |
|---------------|--------|----------------|------------------------|--------------------------------|
| $g(W_l T_1)$ | 1.00 | 0.89 | 0.86 | 0.84 |
| $g(W_l)$ | 1 | 1 | 1 | 1 |

Connected: The template $T(T_k)$ is a connected EDC. It is the **most global workload (for a given $\beta > 0$) with a nonzero weight on each core-distribution link**. All eigenvalues of (Eqn. 9) are greater than zero save $\lambda_N = 0$. The purpose of adding β weights is to ensure that $N - 1$ eigenvalues are nonnegative so that when workloads are considered, lack of connectivity does not adversely affect results, because the first $N - 1$ eigenvalues are always positive when $\beta > 0$. The template has large weights on all dc links and smaller weights on all other links, which is analogous to the OSR weights.

Normalized: The sum of the template adjacency matrix is $\mathbf{1}^T T(T_k) \mathbf{1} = 2(c - 1)$, where $\mathbf{1}$ is a vector of ones. If we scale the measurand weights so that the sum of its adjacency matrix is $2(c - 1)$ we compare normalized workloads and avoid unfairly penalizing workloads due to the number of messages.

Measurand: Tab. I lists how to compute the workload globality measure. The measurand is constructed from the adjacency matrix of the OLTP workload W_l , and is denoted $C = A(W_l)$. The adjacency matrix records the counts of the numbers of messages (they could be weighted by the message size s_l and relative OSR $\hat{\alpha}_\lambda$) that traversed link $\{ij\}$ in the $C_{ij} = C_{ji}$ entries, e.g. C_{ij} equals the number messages on link $\{ij\} \in W_l$. This adjacency matrix is scaled to sum to $2(c - 1)$. The crucial step is that we add the measurand to the template, normalize the sum, and then compute its largest eigenvalue $\lambda_1(\cdot)$. This construction has a number of appealing properties.

Normalized and Positive semi-definite: The linear combination $T(T_k) + C$ is normalized: $\mathbf{1}^T (\bar{C}) \mathbf{1} = 4(c - 1)$. Both $L(T(T_k))$ and $L(C)$ are positive semi-definite (psd), their sum is psd, which implies that $\lambda_n \geq 0, \forall n$. This is proved by appealing to the positive semi-definiteness of Gram matrices.

Theorem 1. Globality Measurement of the OLTP workload W_l , is performed by computing

$$g = \max_k g(W_l|T_k) = \max_k \frac{\lambda_1(L(\bar{C})) - \lambda_1(L(\bar{C}_{min}))}{\lambda_1(L(\bar{C}_{max})) - \lambda_1(L(\bar{C}_{min}))} \quad (10)$$

This measure satisfies **p1-5**. \bar{C}_{max} and \bar{C}_{min} are derived by setting C to equal the most global and most local workload, e.g. a workload with all the traffic on one dc link, $C = A(S_2)$, and all of the traffic equally shared between all links other than the distribution-core links, $C = \bar{T}(T_k)$, respectively.

VI. NUMERICAL EVALUATION

We compute the globality of a series of simple, parametrized synthetic workloads, in (1a) and (1b), using the examples in the previous sections. These first experiments demonstrate that the measure works as expected on easily verifiable small-scale traces. (1a) investigates how the balance of traffic in the core affects the globality measure. (1b) investigates how distributing some of the dc traffic onto a da link reduces the globality score of the workload. Our higher-order measure outperforms (Eqn. 5). In (2a) we explore the dynamic

TABLE III. ORDERING INDEX, GLOBALITY OF A SUBSET OF THE OLTP WORKLOAD PATHS (GIVEN $A(S_5)$), CORRESPONDING PATH AND THE GLOBALITY USING THE TEMPLATE B .

| Sorted Transitions 1 – 8 | | | Sorted Transitions 9 – 16 | | | | |
|--------------------------|--------------------|------------|---------------------------|------|--------------------|------------------|---------------|
| Ind. | $\lambda_1 A(S_5)$ | Path | $\lambda_1 B$ | Ind. | $\lambda_1 A(S_5)$ | Path | $\lambda_1 B$ |
| 1 | 1.00 | 1,5 | 0.825 | 884 | 0.41 | 32,8,2,1,3 | 0.356 |
| 8 | 1.00 | 3,1 | 0.825 | 1012 | 0.35 | 23,6,2,1 | 0.487 |
| 20 | 0.81 | 5,1,3 | 0.444 | 2548 | 0.33 | 36,9,2,1,3,13 | 0.335 |
| 52 | 0.58 | 15,4,1 | 0.444 | 2596 | 0.29 | 19,5,18 | 0.444 |
| 148 | 0.56 | 10,3,12 | 0.487 | 5668 | 0.27 | 63,16,4,1,2,6,23 | 0.279 |
| 180 | 0.53 | 19,5 | 0.394 | 5796 | 0.19 | 66,17,4 | 0.444 |
| 308 | 0.43 | 35,9 | 0.305 | 5988 | 0.14 | 28,7,27 | 0.444 |
| 500 | 0.42 | 13,3,1,2,6 | 0.356 | 6372 | 0.10 | 53,13,3,12 | 0.243 |

range of the globality measure. Our higher-order measure has sufficient discriminative sensitivity to be able to sort all paths between any two nodes in terms of their globality score (from 0 to 1). Some of the paths generated traverse the core of the network; other paths are completely isolated in the access part of the EDC; and finally, some paths are combinations of many aa, da and dc messages. Can our globality score distinguish between them? This test is important as the aggregate measure (Eqn. 5) does not have sufficient discriminative power to do this. Up to this point, we have used one template matrix, our proxy for Cisco’s OSR, in our experiments. In (2b) we test how changing the cost of using a link, its OSR, by considering a new template matrix, changes the globality score of the same traffic traces as in (2a). In a real deployment changing the OSR in this fashion could correspond to the network manager upgrading the value of certain links, in order to reduce the vulnerability of the EDC to failure, for example. (3) We move onto carefully generated large scale traces that reflect the usage of Amadeus’ EDC and perform a statistical analysis of the measure.

(1a) Core Star Cardinality (p1,p5): We measure the globality of a set of parametrized *global OLTP* workloads, where the parameters control the portion of the workload messages that traverse a link. We encode OSRs, in these experiments, in the template matrix. In the first case, these OLTP workloads only traverse the dc links of the topology in Fig. 1. This class of OLTPs is unfeasible in practice but useful for evaluation purposes. The number of different *dc* links traversed is changed in this experiment from $1 < c - 1 < 4$. For each c , the traffic is uniformly divided across the available links and the adjacency matrix is appropriately scaled. Tab. II illustrates that as the number of different dc links used to support the work load increases the globality score $g(W_i|T_1)$ is decreased. The measure in Eqn. 5 does not differentiate between the number of different links used. The higher-order matrix structure of $g(W_i|T_1)$ allows the globality measure to distinguish between the correlation and relative usage of the dc links (cf. Table II). The measure $g(W_i|T_k)$ determines that the usage of 1 dc link to support all of the workload, e.g. $C = A(S_2)$ is more global than sharing the workload over all 4 dc links, $C = A(S_5)$. Tab. II supports the claim made by **p1**, $g(W_i|T_k)$ is upper bounded by one. In addition **p5** is supported by the fact that $g(W_i|T_k)$ distinguishes between different spatial correlations of the workload in Tab. II. **In conclusion, Tab. II demonstrates that the higher-order structure of our globality measure gives the measure a higher discriminative sensitivity than (Eqn. 5).**

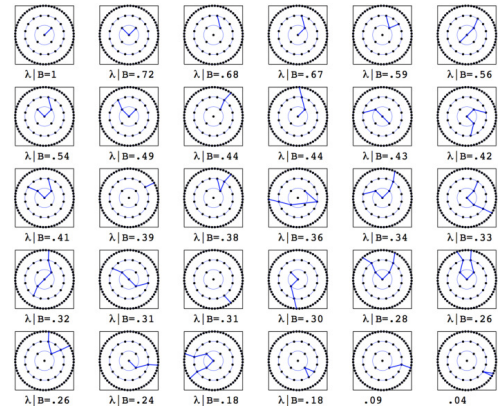


Fig. 5. OLTPs (blue lines), access, distribution and core network regions (illustrated with circles of radius 1,2,3 etc), machines/servers (dots). The globality measure ($\lambda|B$) for a subset of the OLTP workload paths using the template B is indicated for each path.

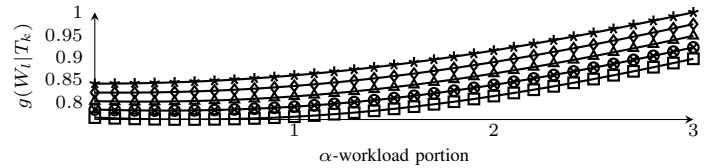


Fig. 6. $\sigma = \{.2, .4, .6, .8, 1\}$ = square, circle, triangle, diamond, star.

(1b) Appendages (p1,p5): In a second parametrization (α, σ) of the workload, we examine the effects of allocating different portions of workload on the four dc links for $0 \leq \alpha \leq 3$ and one da link for $0 \leq \sigma \leq 1$ using the parametrized workload model in Fig. 3 row 4. We plot the globality scores for these parametrized workloads in Fig. 6. The workload quantity is fixed for all parametrizations, and thus the only change is the spatial distribution of workload (which is controlled by the pair α, σ). As σ increases from .2 to .8 the globality of the workload generally increases as α increases. This means that the globality score gives a higher score, for a given value of α , when there is less traffic on the da link and more traffic on the dc links; this is exactly the required behaviour (cf. **p5**). When the traffic is predominantly on the dc links, the globality score should be high. The measure $g(W_i|T_k)$ respects the symmetries in the workloads because different values of α can yield a higher globality for two different σ parametrizations. For completeness we also illustrate the score yielded by the measure (Eqn. 5) in Fig. 7. Here, when $\sigma = 1$ (Eqn. 5) is insensitive to changes in the distribution of the workload. On examination of the case $\sigma = .2$, (Eqn. 5) is over sensitive to changes in α which make (Eqn. 5) an unattractive optimization objective. (Eqn. 5) achieves a minimum value of .75 in these experiments. There is just one appendage in this OLTP and yet the the globality score has fallen by 25%.

(2a) Path Evaluation (p1,p2,p5): We generate workloads using TPC-C’s OLTP-like workload models [29]. We evaluate the globality score by computing the globality of the paths between every pair of nodes in Fig. 1 in order to evaluate the dynamic range of $g(W_i|T_k)$ (cf. **p1**). There are ≈ 7000 feasible transaction-tree workloads in this data-set. We also evaluated the globality on a set of random OLTPs which are composed

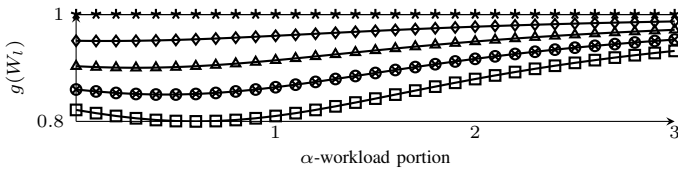


Fig. 7. $\sigma = \{.2, .4, .6, .8, 1\}$ = square, diamond, triangle, diamond, star.

of an arbitrary subset of the paths described above.

(2b) Programmability: We evaluate the programmability of the $g(W_i|T_k)$ by considering an arbitrary template matrix using the following irritation function instead of $A(S_5)$ above. The symmetric adjacency matrix B has ones on the links $\{2, 6\}$, $\{2, 7\}$, $\{2, 8\}$, $\{1, 2\}$, $\{1, 3\}$, $\{1, 4\}$ and $\{1, 5\}$. This irritation function is interesting because the cost of traversing a subset of the distribution to access links $\{2, 6\}$, $\{2, 7\}$, $\{2, 8\}$ is high. The irritation function penalizes links *asymmetrically*. It penalizes workload on a subset of the distribution to access links more than others, and also the workload on the access to access links which are descendants of $\{6, 7, 8\}$. Tab. III illustrates the globality measure of a subset of the OLTP workload paths, the path index and the corresponding path using the template matrix based on $A(S_5)$. In the rightmost column we plot the globality score using B as the template to illustrate the effect of using a different template function. The symmetry of the topology in Fig. 1 means that many paths have the same globality score. Tab. III tabulates the scores and paths associated with transitions in the sorted globality scores for all paths. The crucial result is that transactions that traverse dc links exclusively, have a larger globality measure than transaction trees that traverse the dc and da links. Moreover transaction trees that traverse the dc, da and aa links have a lower globality measure than transactions that traverse da and aa links. **In short Tab. III gives evidence that $g(W_i|T_k)$ provides a consistent ordering of the EDC and OLTP statistics to the real number line interval $(0, 1)$. Comparing the rightmost column of Tab. III, $\lambda_1|B$ with $\lambda_1|A(S_5)$ demonstrates the role the template has in the ordering of the workloads. Some notable re-orderings due to the change in the template are the decrease in the globality score of the path $1 \mapsto 5$ when B is used. This re-ordering has occurred because the template matrix was programmed to penalize the usage of the EDC differently.** Fig. 5 illustrates the paths associated with transitions in the sorted globality scores (overlayed onto the topology in Fig. 1) for all paths using the template B . Firstly, there are more transitions due to the *asymmetry* of B . Secondly, path $1 \mapsto 2$ has the highest globality due to the higher cost associated with links from 2 to $\{6, 7, 8\}$. The *center of mass* of this measure is different due to the *asymmetry* of B . Finally, node 2 is more prevalent in Tab. III compared to Fig. 5 when the globality scores are high. The measure $g(W_i|T_k)$ may be programmed and the resulting measure reflects the level of irritation specified by the encoder. In future work we will consider the application of the globality measure to other EDC architectures, as many EDCs are deploying more Ethernet based VPNs (MPLS or BGP based), which offer a much more flexible network environment.

(3) Simple Stochastic Layered Block Model for Data Center topology and OLTP workload Creation: We have

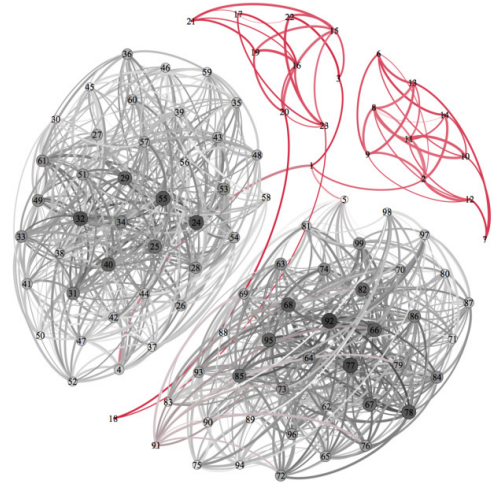


Fig. 8. A randomly generated EDC is created using the simple stochastic layered block model. For illustration purposes $N = 99$ and $K = 4$. The assignment of vertices to each layer-2 network follows the portions $p_k \in \{.1, .1, .4, .4\}$ which results in two layer-2 networks which are significantly larger. The probability of a link between two vertices is $a_k = .3, \forall k$.

verified that the measure works when it should on small examples. We now demonstrate that for large scale traces (up to 100 messages in a transaction tree) that reflect the scenario on Amadeus' EDC, the measure behaves well. To generate a sufficient number of EDC topologies and workloads we introduce a simple stochastic layered block model for data center topology and OLTP workload creation.

Simple Stochastic Layered Block Model for EDC topologies: We fix the number of vertices in the EDC to be N . We define the number of distribution switches to be K ($N = 1000$ and $K = 4$). The distribution and core vertices have edges with probability 1. We generate the layer-2 network by assigning each of the remaining $N - K - 1$ vertices at random to one of the distribution switches $k = 2, 3, 4, 5$ with the proportion p_k , such that $\sum_k p_k = 1$. We connect any two edges within the k th layer-2 network according to the assignment probability a_k . An example EDC is illustrated in Fig. 8 with $N = 99$ and $K = 4$. This model produces K layer-2 networks with different counts of vertices. Each layer-2 network can have different probabilities of any two nodes being connected. This model gives us the freedom to generate a large family of EDC topologies at random. Care must be taken to choose $\{a_k, p_k\} \forall k$ so that each vertex is connected.

Generating random OLTP Workloads: Random OLTP traffic is required to test the globality score. For each EDC topology above, we generate an ensemble of OLTP workloads by: (1) generating Minimum Spanning Trees, which are rooted at different vertices (in the generated EDC). We then prune these trees so that they have different numbers of edges in them; (2) we generate OLTP workloads which are the shortest path between every pair of nodes in the network. Finally we combine all of these OLTP workloads into two sets. In the first set the OLTP workloads have the characteristic that they traverse the core of the EDC. In the second set, the OLTP workloads do not traverse the core of the EDC. Fig. 8 illustrates how a sample ensemble of OLTP workload traces use the EDC resources. It is this usage that we want to quantify.

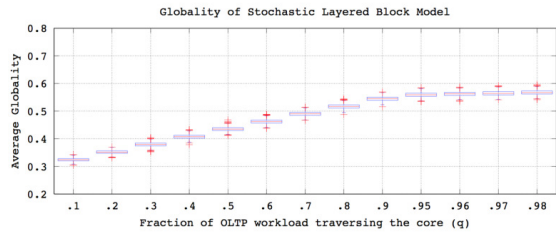


Fig. 9. Aggregate Globality Measurement over a range of EDC topologies and types of OLTP workloads (for a given q).

Dark thick links indicate large numbers of messages. Dark large vertices indicate that the vertex has a large number of incoming/outgoing messages. Thin red links indicate that the link delivers few messages. Red small vertices indicate that the vertex receives/sends few messages. Note that vertices 1, 2, 3, 4, 5 are small and red and that the links between them are thin and red. This indicates that the load on the links between them is small. The two larger clusters of nodes have thicker darker links and vertices. The OLTP workload is well isolated within these layers of the EDC.

To generate a *test* OLTP workload, we draw with replacement 100 OLTP workloads such that proportion q of the OLTP workloads is drawn from the set that does not traverse the core, and the remaining $1 - q$ OLTP workloads do traverse the core. Each test OLTP workload is parametrized by q . We generate a test OLTP world for each instance of a range of the parametrizations of the EDC above $\{p_k, a_k\} \forall k$ and we compute the average globality score of each of these workloads given its associated EDC. Fig. 9 illustrates box-plots of these globality scores. These box-plots demonstrate that we cannot just use q the percentage of OLTP workloads that traverse the core as a globality measure, as the corresponding globality scores have a wide range, e.g. up-to 10% of the median globality score. This 10% may be the difference between detecting whether or not the EDC is in a vulnerable state or not. It may also be occlude whether or not the current OLTP workload has step-changed into a new performance regime, e.g. a globality score of $\approx .5 \pm 0.05$ (which corresponds to $q = .7$) could be confused with a globality score of $\approx .55 \pm 0.055$ (which corresponds to $q = .8$). As we average our globality measurements over many EDC topologies, it is reasonable to assume that the more extreme globality scores are averaged out, and that—especially for smaller values of q —the range of globality scores for a particular EDC may be more extreme. What is clear is that the range of the average globality scores is reduced to $0.3 \leq g(\cdot) \leq .6$. We posit that this is an artifact of averaging the globality scores over different EDC parametrizations. In future work we will give a more detailed treatment of the range of potential globality scores achievable for a given EDC (as opposed to considering the average globality scores over a range of EDCs) by considering the effects of $\{p_k, a_k\}$.

VII. CONCLUSION

The starting point for designing an EDC OLTP workload optimization procedure is to have the ability to measure how sub-optimal the current (or any other candidate) EDC OLTP workload is in way which is informative. In this paper

we argued that the notion of workload isolation, locality or globality has not been rigorously defined in the literature. Questions such as, what is the worst (or best) case OLTP workload in terms of its isolation, locality or globality, have not been answered. We presented a measurement algorithm that maps a high dimensional set of EDC and OLTP statistics to the real number line, in way that allows the user to encode prior knowledge of what a bad workload is, into the measurement function. We demonstrated that this mapping is consistent. It respects the set of properties globality measures should have.

REFERENCES

- [1] N. Bitar, S. Gringeri, and T. Xia, “Technologies and protocols for data center and cloud networking,” *Comms Mag., IEEE*, vol. 51, no. 9, pp. 24–31, 2013.
- [2] K. Li, H. Zheng, and J. Wu, “Migration-based virtual machine placement in cloud systems,” in *IEEE 2nd CloudNet*, 2013, pp. 83–90.
- [3] R. Jain and S. Paul, “Network virtualization and software defined networking for cloud computing: a survey,” *Comms Mag., IEEE*, vol. 51, no. 11, pp. 24–31, 2013.
- [4] Ruairí de Fréin, C. Olariu, Y. Song, R. Brennan, P. McDonagh, A. Hava, C. Thorpe, J. Murphy, L. Murphy, and P. French, “Integration of QoS Metrics, Rules and Semantic Uplift for Advanced IPTV Monitoring,” *J. Net. and Sys. Man.*, pp. 1–36, 2014.
- [5] Ruairí de Fréin, “Effect of system load on video service metrics,” in *IEEE Irish Sig. Sys. Conf. (ISSC)*, 2015, pp. 1–6.
- [6] A. Greenberg, J. Hamilton, D. Maltz, and P. Patel, “The cost of a cloud: Research problems in data center networks,” *SIGCOMM Comput. Commun. Rev.*, vol. 39, no. 1, pp. 68–73, 2008.
- [7] M. Banikazemi, D. Olshefski, A. Shaikh, J. Tracey, and G. Wang, “Meridian: an SDN platform for cloud network services,” *Comms Mag., IEEE*, vol. 51, no. 2, pp. 120–7, 2013.
- [8] I. Gutman, “The Star is the Tree with Greatest Laplacian Eigenvalue,” *Kragujevac J. Math.*, vol. 24, pp. 61–5, 2002.
- [9] M. Benzi and C. Klymko, “On the limiting behavior of parameter-dependent network centrality measures,” *SIAM Journal on Matrix Analysis and Applications*, vol. 36, no. 2, pp. 686–706, 2015.
- [10] S. Borgatti and M. Everett, “A graph-theoretic perspective on centrality,” *Social Networks*, vol. 28, no. 4, pp. 466–84, 2006.
- [11] P. Bonacich, “Power and Centrality: A Family of Measures,” *Amer. J. Sociology*, vol. 92, no. 5, pp. 1170–82, 1987.
- [12] R. Birke, L. Chen, and E. Smirni, “Data centers in the cloud: A large scale performance study,” in *IEEE 5th Cloud Comp.*, 2012, pp. 336–43.
- [13] Y. Watanabe, H. Otsuka, M. Sonoda, S. Kikuchi, and Y. Matsumoto, “Online failure prediction in cloud datacenters by real-time message pattern learning,” in *IEEE CloudCom*, 2012, pp. 504–11.
- [14] Ruairí de Fréin, “Formal concept analysis via atomic priming,” in *Formal Concept Analysis*, ser. LNCS, P. Cellier, F. Distel, and B. Ganter, Eds. Springer Berlin Heidelberg, 2013, vol. 7880, pp. 92–108.
- [15] Ruairí de Fréin, “Ghostbusters: A parts-based NMF algorithm,” in *IET Irish Sig. Sys. Conf. (ISSC)*, 2013, pp. 1–8.
- [16] C. Niemz, “Network Modeling of Enterprise Datacenters,” Master’s thesis, Systems Group, Dep. Comp. Sc., ETH Zürich, 2013.
- [17] M. Al-Fares, S. Radhakrishnan, B. Raghavan, N. Huang, and A. Vahdat, “Hedera: Dynamic Flow Scheduling for Data Center Networks,” in *7th USENIX Conf. Net. Sys. Design & Implem.*, 2010, pp. 1–15.
- [18] B. Xu, Ruairí de Fréin, E. Robson, and M. Ó Foghlí, “Distributed formal concept analysis algorithms based on an iterative mapreduce framework,” *Formal Concept Analysis*, vol. 7278, pp. 292–308, 2012.
- [19] *IEEE 802.1Q, IEEE Standard for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks*, IEEE, 2012.
- [20] *Cisco data center infrastructure 2.5 design guide*, Cisco, 2005.
- [21] C. Reiss, A. Tumanov, G. Ganger, R. Katz, and M. Kozuch, “Towards understanding heterogeneous clouds at scale: Google trace analysis,” Intel Sc. & Tech. Center for Cloud Computing, Tech. Rep., 2012.

- [22] J. Guo, F. Liu, J. Lui, and H. Jin, "Fair Network Bandwidth Allocation in IaaS Datacenters via a Cooperative Game Approach," *Networking, IEEE/ACM Trans.*, vol. PP, no. 99, pp. 1–14, 2015.
- [23] T. Benson, A. Akella, and D. Maltz, "Network traffic characteristics of data centers in the wild," in *10th ACM SIGCOMM IMC*, 2010, pp. 267–280.
- [24] S. Kandula, S. Sengupta, A. Greenberg, P. Patel, and R. Chaiken, "The nature of data center traffic: Measurements & analysis," in *Proc. 9th ACM SIGCOMM IMC*, 2009, pp. 202–8.
- [25] S. Kandula, J. Padhye, and P. Bahl, "Flyways to de-congest data center networks." in *HotNets*. ACM SIGCOMM, 2009.
- [26] A. Greenberg, J. Hamilton, N. Jain, S. Kandula, C. Kim, P. Lahiri, D. Maltz, P. Patel, and S. Sengupta, "VI2: A scalable and flexible data center network," *ACM SIGCOMM Conf. Data Comm.*, pp. 51–62, 2009.
- [27] T. Benson, A. Anand, A. Akella, and M. Zhang, "Understanding data center traffic characteristics," *SIGCOMM Comput. Commun. Rev.*, vol. 40, no. 1, pp. 92–9, 2010.
- [28] T. Benson, A. Anand, A. Akella, and M. Zhang, "The case for fine-grained traffic engineering in data centers," in *Int. Net. Man. Conf. on R. Ent. Net.* USENIX, 2010, pp. 1–6.
- [29] S. Leutenegger and D. Dias, "A Modeling Study of the TPC-C Benchmark," in *Proc. 1993 ACM SIGMOD*, 1993, pp. 22–31.
- [30] B. Mozafari, C. Curino, A. Jindal, and S. Madden, "Performance and Resource Modeling in Highly-concurrent OLTP Workloads," in *Proc. ACM SIGMOD*, 2013, pp. 301–12.
- [31] Y. Chen, F. Raab, and R. Katz, "From TPC-C to Big Data Benchmarks: A Functional Workload Model," EECS Dep., University of California, Berkeley, Tech. Rep. UCB/EECS-2012-174, 2012.
- [32] K. Rohe, S. Chatterjee, and B. Yu, "Spectral clustering and the high-dimensional stochastic blockmodel," *Ann. Stat.*, vol. 39, no. 4, pp. 1878–915, 2011.
- [33] *Hadoop Rack Awareness*, Apache SW Found. [Online]. Available: http://hadoop.apache.org/docs/r1.2.1/cluster_setup.html#Hadoop+Rack+Awareness
- [34] Z. Xu, C. Tang, and Z. Zhang, "Building topology-aware overlays using global soft-state," in *23rd IEEE Int. Conf. Dist. Comp. Sys.*, 2003.
- [35] S. HomChaudhuri and M. Foschiano, *Private VLANs: Addressing VLAN scalability and security issues in a multi-client environment*, 2008. [Online]. Available: <http://tools.ieft.org/html/draft-sanjib-private-vlan-10>
- [36] *VMWare virtual networking concepts*, VMWare, 2007. [Online]. Available: http://www.vmware.com/files/pdf/virtual_networking_concepts.pdf
- [37] B. Pfaff, J. Pettit, T. Koponen, K. Amidon, M. Casado, and S. Shenker, "Extending networking into the virtualization layer," *HotNets-VIII*, 2009.
- [38] *Campus Network for High Availability Design Guide*, Cisco, 2008.