Visualizing Interfirm Collaborations in the Microservices Ecosystem

Rahul C. Basole
Georgia Institute of Technology
basole@gatech.edu

Abstract

The shift from monolithic software solutions to a microservices architecture is fundamentally changing the way software is developed, deployed, and managed. In this paper, we aim to uncover the collaborative fabric of the microservices ecosystem using a data-driven visualization approach of 2,608 software firms. Our visual analysis reveals a core-periphery structure with several subcommunities, suggesting both complementary and competing arrangements between software vendors. Theoretically, our paper contributes to our understanding of interfirm relationships in a software context. Managerially, our results show that there are wide range of partnership strategies that shape the microservices ecosystem. Methodologically, we demonstrate how a data-driven ecosystem visualization approach can help decision makers augment their sensemaking capability of emerging software ecosystems. The paper concludes with opportunities for future research.

1. Introduction

The shift from monolithic software solutions to a microservices architecture is fundamentally changing the way software is developed, deployed, and managed [1]. The fundamental idea behind microservices is that software applications are broken down into smaller, modular, and loosely connected components that are easier to understand, test, and maintain [2]. This novel approach enables organizations to achieve greater software engineering agility, accelerate the time to production, and facilitate better coordination of geographically dispersed developers [3].

The idea of decomposing applications into smaller parts is not new; a service-oriented architecture, for instance, followed a similar programming paradigm [1]. Yet, technological advances, such as containerization, application programming interfaces (APIs), and load balancing, along with the growing expectation of seamless digital experiences, have led to many novel types of software development tools and techniques needed for the changing business environment [4, 5]. Given the many benefits of microservices, it is not surprising that many organizations are embracing this new model1.

While microservices are technical artifacts that can be (almost freely) coupled, the utility and extent of these linkages can be influenced and enhanced by the organizational arrangements between the providing vendors. Ultimately, no software vendor is an island. In the past, vendors frequently provided complete solutions themselves, resulting in closed software environments. For a vendor to create sustained value today, it must be able to integrate, consume, and share services with other vendors. Building, joining, and cultivating a broad partner ecosystem can lead to new value opportunities and is thus an important prerequisite for success.

To understand these new organizational arrangements and economic activities, scholars are increasingly using an ecosystem lens [6, 7, 8]. This lens is particularly applicable to microservices, which can be characterized as a complex, evolving socio-technical system of highly interdependent actors who co-create value [9]. As in many other technology domains (e.g., [10, 11]), interfirm collaboration is also of essence in microservices. Collaboration allows organizations to offer new services, scale operations through new channels, enter global markets, provide training and expertise, and accelerate innovation [12].

Despite the widely acknowledged benefits of collaboration in a software context, our understanding of the nature of partnerships in microservices is lagging. In this paper, we aim to uncover the collaborative fabric of the microservices ecosystem using a data-driven approach. Specifically, we focus our visual analysis on interfirm relationships formed by microservices vendors. We first identify and curate a dataset of 2,608 companies (vendors and their partners) from multiple unstructured data sources. Given the emerging nature of

the microservices ecosystem, our dataset contains both publicly-listed and privately-held firms. We use a visual analysis approach as it has been shown to be particularly valuable for discovery, exploration, and sensemaking for relatively unknown contexts and complex datasets [11, 10, 13, 14].

Theoretically, our study contributes to our general understanding of interfirm relationships in a software context revealing the existence of a core-periphery structure. Managerially, our results show that there are wide range of partnership strategies employed by vendors – including differences in partner network size, partner network composition, and emphasis on core and peripheral vendors. Methodologically, we demonstrate how a data-driven ecosystem visualization approach can help decision makers augment their sensemaking capability of emerging software ecosystems. As microservices continue to evolve, such a capability will allow decision makers to prepare, anticipate, or mimic successful partnership strategies.

The remainder of the paper is organized as follows. Section 2 reviews related work. Section 3 presents our methodology, including our data curation process, metrics, and visualizations. Section 4 presents the results. We discuss our findings from the vendor, partner, and developer perspectives in Section 5. We conclude with implications and future research opportunities in Section 6.

2. Related Work

For our study of the microservices ecosystem, we draw on three traditionally separate research streams, namely interfirm relationships, software ecosystems, and microservices.

2.1. Interfirm Relationships

Firms face a difficult challenge to effectively compete on their own in an increasingly dynamic, turbulent, and hypercompetitive business environments [15]. Resources, knowledge, and capabilities necessary to innovate and succeed are often distributed among other firms [6, 8]. It is this reality that requires firms, irrespective of size, financial resources, and industry, to engage in interorganizational relationships with a heterogeneous set of partners to gain access to novel and complementary capabilities, knowledge and resources [16]. Broadly considered, interfirm networks can be considered complex webs of ties interconnecting firms within and across industries [17]. Each firm in the network is endowed with a diverse set of capabilities and resources. It has been argued that networks act as both channels and prisms through which these resources flow and knowledge is shared [18, 19]. Relationships are formed for exploratory or exploitative reasons. Exploratory relationships are formed to create new while exploitative relationships are formed to execute existing knowledge, functions, and activities [20].

2.2. Software Ecosystems

Interfirm relationships play a particularly critical role in the software industry [21, 22, 23, 24]. In the early days of software engineering, the development of a software product was commonly the result of a single independent software vendor creating a “monolithic product” [25]. Modern cloud-based software products on the other hand rely on many different components and infrastructures provided by a range of third party vendors and open source suppliers. It is this change in business models that has transformed the software development landscape into a highly interconnected ecosystem characterized by a myriad of complex relationships between software vendors [12, 21].

Similar to its biological metaphor, value in a software ecosystem is co-created by a range of actors. The success of a software vendor is consequently co-dependent on the relationships it maintains with other players in the ecosystem. Since firms can operate in one or more layers of the software stack, it is not unusual for two software vendors to collaborate on one activity level and compete in another.

Research on software ecosystems has been gaining prominence since the seminal work by Messerschmitt and Szyperski [26]. Prior studies have focused on a variety of topics and themes, including the architecture of the underlying software, the management and governance of software ecosystems, and the structure and dynamics of interfirm relationships in software ecosystems. Research on software architecture is concerned with the technical design and configuration of the interfaces and components. Studies include an exploration of the requirements of software as well as the importance of modularity and flexibility. Research on the management and governance of software ecosystems focuses on how software ecosystems are organized, monitored, and orchestrated, including an exploration of the rules and mechanisms that shape healthy software ecosystems. The third focus area involves an examination of roles actors assume in software ecosystems and relationships between them. Prior work has identified many different roles, including orchestrators or keystones, niche players, and external actors. It is the interaction between these actors that define the trajectory of software ecosystems. Different types of ecosystems use different interaction models.
2.3. Microservices

The scholarly investigation of microservices has primarily been a technical endeavor limited to the software engineering, service engineering, and cloud computing disciplines [4]. Prominent topics have included the design, testing, configuration and management, migration, identification, and deployment of microservices [27, 28, 1, 29].

To the best of our knowledge, studies on the broader organizational strategies and implications associated explicitly with microservices have yet to emerge. However, IS scholars are increasingly recognizing the importance of services, service-oriented architectures, and service management for digital transformation [30, 31, 32, 33, 34]. Our study aims to fill this gap by examining supply-side partnering strategies by major vendors.

3. Methodology

Following [35], we use a data-driven approach to identify and curate a dataset of interfirm collaboration in the microservices ecosystem (see Figure 1). To the best of our knowledge, there are no comprehensive data sources that capture microservices vendors. Our identification of core vendors and their partners (Step 1) thus started with extracting firms from two prominent microservice vendor market maps provided by Sequoia Capital (2016)3 and Nanoscale (2017)4 and a list of vendors participating in the Cloud Native Computing Foundation landscape (2017)5. Together, these maps/lists provide an excellent starting point as they organize vendors into broad categories2 and provide a triangulated perspective of relevant focal actors in the microservices ecosystem. Rather than using all vendors, we used firms that co-appeared in the three lists for cross-validation purposes.

After removing duplicates and consolidating names by the ultimate parent, our core dataset contained 131 unique vendors. For each of these vendors, if available, we identified the partner page on their corporate website and extracted all of the partner names using a semi-automated approach. If only logos were provided, and the company name was not directly discernible, we used a reverse image search using the Google Cloud Vision API7 to determine the partner name. If available, we also noted the type of partnership level between the two companies (e.g., technology, strategic, channel/reseller, system integrator, etc.). Once we had curated a list of vendor-partner pairs, we used a fuzzy match algorithm to match and resolve any ambiguities in partner names. This resulted in a total of 2,583 unique partner names.

Next, we modeled the microservices ecosystem as an undirected network, with nodes representing firms (vendors and their partners) and edges representing partnerships between these firms (Step 2). We scaled edges proportional to the number of relationship types (each type receiving the same weight) between two firms. We then computed various network metrics to gain insight into the topological structure of the ecosystem (Step 3). Specifically, following [10, 21, 36] we measured density (the proportion of ties that are realized in the network relative to the hypothetical maximum possible), average degree (the average number of partnerships each firm has), average weighted degree (average of sum of weights of edges connected to nodes in the network), modularity (number of subcommunities in the network defined as firms that are more densely connected together than to the rest of the network)6, network diameter (shortest distance between the two most distant firms in the network), and average path length (the average number of steps along the shortest paths for all possible pairs of firms).

We used Gephi 0.92, an open-source graph visualization software, to visualize the network [38] (Step 4). We used OpenORD, a cluster emphasizing layout algorithm, to depict the overall ecosystem structure [39] and applied the NoOverlap heuristic to minimize node occlusion, thereby improving the overall readability and aesthetics of the graph. In all of our visualization, we sized nodes proportionally to their betweenness centrality (the positional prominence of a firm in the network)40.

Lastly, to make sense of the resulting visualizations and to understand implications for practice, we conducted semi-structured interviews with software developers, vendors, and consultants (Step 5). Our five respondents provided qualitative feedback on the visualizations associated with their understanding of the industry.

2A keyword search of “microservices” in studies published in the IS Senior Scholar’s Basket of Journals revealed no results.

3https://www.sequoiacap.com/article/build-us-microservices/

4https://www.nanoscale.io/ecosystem/

5https://github.com/cncf/landscape

6The maps/lists organize vendors by many different categories, including Orchestration, API Management, Platform Management, Load Balancing, Container Registry, Monitoring, and Log Analysis. However, there is no consensus between these categories and many category labeling differences exist. Moreover, vendors can appear in one or more categories. We did not consider these categories in our study and further analysis on these categorical assignments are needed, but beyond the scope of this paper.

7https://cloud.google.com/vision/

8We applied Louvain’s community detection algorithm to identify subcommunities in the ecosystem [37].
4. Results

Table 1 shows summary statistics of key structural metrics for the entire microservices partner ecosystem (2,608 firms) and for core vendors only (131 firms). In total, there are 3,510 unique partnerships, with 239 among vendors only. The average number of partners (average degree) each vendor has is 6.128. The average path length, which indicates the average number of hops each firm has to take from another is 2.546. The density, which is the ratio of existing partnerships over all possible partnerships, is 8% for the core vendor network. Together, these metrics suggest that vendors form a highly interconnected core and operate in a small world-like context.

<table>
<thead>
<tr>
<th>Ecosystem Metric</th>
<th>Overall</th>
<th>Vendors Only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nodes</td>
<td>2,608</td>
<td>131</td>
</tr>
<tr>
<td>Edges</td>
<td>3,510</td>
<td>239</td>
</tr>
<tr>
<td>Density</td>
<td>0.001</td>
<td>0.08</td>
</tr>
<tr>
<td>Average Degree</td>
<td>2.612</td>
<td>6.128</td>
</tr>
<tr>
<td>Average Weighted Degree</td>
<td>2.873</td>
<td>6.59</td>
</tr>
<tr>
<td>Modularity</td>
<td>0.259</td>
<td>0.293</td>
</tr>
<tr>
<td>Network Diameter</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>Average Path Length</td>
<td>3.807</td>
<td>2.546</td>
</tr>
</tbody>
</table>

Figure 2 shows the corresponding visualization of the microservices ecosystem. For readability purposes, we only depict the main component (largest-connected subgraph) and focus the labels on the core vendors. This eliminates vendors that have no partners or are not connected to the core activities in the ecosystem. We differentiate core vendors and partners using two-hue color-encoding. Microservices vendors are depicted in red; partners are depicted in gray. Nodes (and node labels) are sized proportionally to their level of betweenness centrality, a graph theoretic measure that captures the relative importance of an entity in the broader ecosystem. Several interesting observations can be made from this graphical depiction.

First, there are several centrally prominent vendors in the microservices ecosystem. These include Atlassian, OpenShift, Kubernetes, and Docker. Each of them have a rich partner network that resembles a subecosystem. More peripheral players, with significant number of partners, include Mulesoft, Nagios, New Relic, and Palo Alto. Most vendors operate relatively small to moderate partner network sizes. However, some players have significant partner collaborations, including Splunk, Atlassian, Nagios, and Mulesoft. Some vendors are only weakly connected to the main component, including Tibco and Pubnub.

Another important observation that can be made from the visualization is the large number of vendors that occupy a relatively central position but are not necessarily prominent as measured by their partner network. We have labeled these as the ecosystem core. These vendors play an important coordinating or enabling role with other microservices vendors and come from key segments like the public cloud, infrastructure, and monitoring. Examples include AWS, Microsoft Azure, and Google Cloud.

To more deeply understand the collaborative structure among vendors only, Figure 3 filters out partner nodes. We then rerun the force-directed layout algorithm on this reduced dataset to identify the structure of core and peripheral players and size nodes again by their structural prominence. We clearly notice the relative size of four vendor actors: Kubernetes, OpenShift, Docker, and Microsoft Azure. We also observe some deeper relationships (indicated by line thickness) between a few players, in particular Splunk and its partners.

Furthermore, we use the Louvain community detection algorithm to distinguish subgroups and color-encode them according to subcommunity membership. Our visualization reveals the presence of six subcommunities that are more closely interconnected. The central (light green) core consists of actors such as Google Cloud, Docker, Microsoft Azure, Oracle, IBM, Mesosphere, and Dreamfactory (indicated by A). Kubernetes and its partners form a large second subecosystem (indicated by B). The largest subcommunity is depicted in purple, which includes OpenShift, jfrog, Smologic, Hashicorp, and Atlassian, among many others.
5. Discussion

Our study contributes to the growing yet nascent literature adopting a visual analytic view of emerging software ecosystems. The visualizations provide important insights from three complementary perspectives.

5.1. Vendor Perspective

Software vendors have the choice and opportunity to architect high-performing partner ecosystems by ensuring that relationships with software partners are chosen and managed “wisely”, that partners are selected appropriately such that it provides the opportunity for greater collaboration, and that the “right” relationships are built, maintained, and cultivated. We postulate that it may be more judicious to have a few good relationships, i.e., establish relationships with the right partners, who provide complementary or value-added capabilities and networks, rather than simply create a lot of partnerships that do not provide significant value. Moreover, vendors can collaborate with both partners and other major vendors to create subecosystems that can create compelling value propositions to companies and developers. In doing so, they can create important
fluid lock-in effects. The creation of such ecosystems is important particularly as new entrants emerge and existing power structure is shifted.

5.2. Partner Perspective

An equally important dual perspective includes the strategies that partners use to position themselves in the microservices ecosystem. Our visual analysis reveals that many small partner firms tend to align themselves almost exclusively to a single vendor. On other other hand, we also identify a few partner firms that position themselves more centrally between vendors, either hedging their risks or acting as important bridges and gatekeepers. This result underlines the importance and relevance of niche players advocated in prior ecosystem work (e.g. [6]).

5.3. Developer Perspective

While our study takes a vendor-centric view of interfir collaboration, it needs to be acknowledged that developers play a crucial role in the overall evolution of the microservices ecosystem. To understand our interfir collaboration maps from a developers’ perspective, we thus discussed our visualizations with a group of active microservices developers. One of the key insights that developers commented on...
is the surprising distribution of vendors across the core-periphery structure of the ecosystem. In particular, it was commented that, while disconnected from an interfirm collaboration perspective, developers tend to use and integrate microservices from both core and peripheral vendors. In general, it was argued that developers do not think much about which vendors collaborate, but rather whether the underlying services/technologies actually work together seamlessly. As such the real importance to developers is technology integration. Yet, developers were surprised to observe that some of the vendors were not collaborating closer given the natural connection/integration between the value offerings they provide. The findings of the core vendor collaboration however aligned well with their perspectives of the ecosystem. In general, developers suggested that vendors should be more cognizant of developers’ needs and form interfirm strategies accordingly.

6. Concluding Remarks

The shift from monolithic software solutions to a microservices architecture is fundamentally changing the way software is developed, deployed, and managed. In this paper, we uncovered the collaborative fabric of the microservices ecosystem using a data-driven visualization approach of nearly 3,000 software firms. Our analysis revealed a clear core-periphery structure, with significant collaborations between core vendors. We also discovered that some players occupy a much more prominent position and sub-clusters exist. Theoretically, our paper contributes to our understanding of interfirm relationships in a software context. Managerially, our results show that there are wide range of partnership strategies that shape the microservices ecosystem. Methodologically, we demonstrate how a data-driven ecosystem visualization approach can help decision makers augment their sensemaking capability of emerging software ecosystems.

Our paper provides a foundation for several future research avenues. First, as with all ecosystems, the microservices ecosystem is highly dynamic. New entrants emerge and new relationships are formed constantly. An important future research direction thus would include a temporal, evolutionary perspective on the ecosystem. Second, interfirm partnerships are only one form of value creation that actors in the microservices ecosystem can pursue. As more startups emerge, we will see increased venture funding and acquisition activities. An examination of these alternative relationships would be fruitful. Third, a natural extension of our work would be an examination of the association of the collaborative structure and position a vendor occupies with respect to firm performance. Do the positions that firms occupy in the ecosystem actually lead to differential performance (e.g., market share, revenue, traffic, etc.)? Fourth, another interesting area of research would be to examine the structural footprint of different segments in the microservices ecosystem. Do vendors concentrate or distribute their efforts across different value segments? Fifth, microservices are geography-agnostic, yet most vendors may be located in North America. Yet, given the global nature of software development and cross-border data policies, it would be interesting to explore what geographic differences may exist. Lastly, from the conversation with developers it appears that while interfirm collaboration provide novel insights, an equally important visualization would focus on microservices integrations.

References


