Comparing Agent-based Control Architectures For Next Generation Telecommunication Network Infrastructures

Marco Pérez Hernández, Duncan McFarlane,
Manuel Herrera, Amit Kumar Jain, Ajith Kumar Parlikad

Institute for Manufacturing, Department of Engineering
University of Cambridge, United Kingdom

[mep53, dm114, amh226, akj29, aknp2] @cam.ac.uk

Abstract: Multi-agent systems have been an effective choice for designing control systems that are flexible and agile. However, few attention has been given to the evaluation of the architectures of such systems. This becomes critical with the emerging requirements in complex domains such as digital network infrastructures. In this paper, we propose an approach for the evaluation of agent-based control architectures and introduce three multi-agent based architectures for the supervisory control of network service operations of the next generation of digital infrastructures. With the proposed approach, we evaluated the architectures and the implemented control systems prototypes under a realistic network infrastructure environment. Our approach has been effective to evaluate the candidate architectures. The results of communication overhead and reaction time, have shown that agent-based hierarchical and heterarchical-ring architectures have outperformed the heterarchical-complete network architecture.

Keywords: Optimization and control of large-scale network systems; Multiagent systems; Decentralized and distributed control

1. INTRODUCTION

The digital telecommunication network infrastructures are fundamental for the realisation of the developments in fields such as the internet of things, autonomous cars, virtual and augmented reality, among others. These infrastructures should support different control domains and deal with vast amounts of operational data. In addition, they should tolerate various demand profiles and resource constraints while delivering high performance services. Given these characteristics, a distributed approach is sensible for the supervisory control of these infrastructures.

Agent-based control systems offer flexibility and inherent distribution of control. Although these systems have been proposed in different domains, it is difficult to evaluate how the architectures of these systems are suitable for the described infrastructures. As control architectures are tailored to the domain and problem at hand, a criteria for comparing such architectures and an approach for their evaluation are still open challenges. The aim of this paper is to assess the suitability of different agent-based supervisory control architectures for the service operations in the context of the next generation telecommunications infrastructures. We focus on: Q1: Which of these architectures perform better in this context? and Q2: What are the key aspects to consider in the assessment these architectures?

We address these questions by identifying a set of key control requirements for service operations of digital infrastructures and by proposing the criteria for evaluation of architectures from the design and operational perspectives. These are our contributions that are presented in the remaining of paper as follows. First, in section 2, we review the relevant reference control architectures and key previous efforts for evaluation of agent-based control architectures. Then, in section 3, we summarise the identified control system requirements. Next, in section 4, we introduce the three agent-based control candidate architectures for evaluation. The design perspective of Q2 is tackled in section 5 and the operational perspective in section 6. The results and discussion addressing Q1 are in section 7. Finally, concluding remarks are in section 8.

2. DISTRIBUTED CONTROL SYSTEMS

We review the reference control architectures and key evaluation approaches for agent-based control architectures.

2.1 Control System Architectures

The reference architectures present the essence of functionalities, components, their relations and structure of the control systems. In addition to centralised architectures, there are three main distributed control architectural approaches: hierarchical, heterarchical and holonic (Bongaerts et al., 2000). The hierarchical and heterarchical
are the two ends of the control spectrum whereas holonic control is an intermediate solution. In this study we focus on the hierarchical and heterarchical approaches.

The hierarchical architectures are organised vertically across various levels. The supervision function is clear with low-level supervisors commanded by high-level ones (Cai and Wonham, 2016). Dilts et al. (1991) identified two forms of this architecture: the proper hierarchical, with decision-making at higher levels and a command/response communication down to lower levels; and the modified hierarchical, with coordination among subordinates that might interact with others at the same level, without constant instructions from higher levels (Bongaerts et al., 2000). These architectures usually have limited flexibility to adapt to changes and disturbances (Dilts et al., 1991).

The heterarchical architectures do not have a direct controlling component. The supervision is cooperatively carried out and spread across the system. This brings the benefits of reduced complexity and higher flexibility (Duffie and Piper, 1987; Cai and Wonham, 2016). However, this horizontal distribution implies a lack of global view and brings unpredictability to the system, which is why it is still controversial whether this is a supervisory control.

2.2 Evaluation of Agent-based Control Architectures

Few works have addressed the problem of evaluating or selecting among a set of control architectures. Liang et al. (2008) focus on transaction-based communication cost and the conflict-resolution lead time to briefly analysed performance of the architectures for Virtual Manufacturing Enterprises. Brintrup et al. (2011) proposed the metrics of response time, stability, scalability, and overall optimality to evaluate agent platforms that enable the vision of intelligent self-service assets. Kruger and Basson (2013) compared the implementation time of multi-agent systems and the IEC 61499 function blocks control approach. Drakaki and Tzionas (2016) used Petri net models to evaluate performance of hierarchical, heterarchical and hybrid systems for warehouse resource allocation. Kruger and Basson (2019) evaluated control systems for a manufacturing cell with attention to the implementation, particularly, development time, code complexity and code reuse, among other metrics. Palau et al. (2019) introduced a cost-component methodology to assess optimality of four agent-based architectures for collaborative prognostics. They compared the total cost of each architecture and processing cost in a simulated environment.

3. SERVICE OPERATIONS FOR DIGITAL INFRASTRUCTURES

The next generation digital infrastructures (NGDIs) are multi-layer networks that enable the provision of customised and heterogeneous digital services (see Fig 1). Beyond the convergent mobile-fixed access networks (Leitão et al., 2016), the physical and the software-defined networks (SDN) (İçer et al., 2013), other networked systems such as the cyber-physical ones (Serpanos, 2018) are also embedded in the NGDIs making them complex to manage.

![Fig. 1. Next Generation Digital Infrastructure](http://www.ng-cdi.org/)

3.1 Service Operations

Service operations refers to the processes required to ensure that both the service experience and the service outcomes, such as “products”, benefits or emotions, meet customer and business expectations (Johnston et al., 2012). In telecommunications, services such as telephony or on-demand video streaming, are a type of user services (Stavdas, 2010). These are supported by other network services, that refer to data transfer modes (Stavdas, 2010), e.g. a highly secured point-to-point circuit at 100MB/second. Likewise, these network services are supported by IT services enabling the management of resources, events and other related operational activities.

Network service operations include the reactive management of user requests, incidents and root-cause problem investigation as well as the proactive data analytics to trigger the adaptation of operations e.g. asset maintenance re-planning. Existing IT Service Management frameworks such as the IT Infrastructure Library (Orr et al., 2011), provide a thorough description of processes, roles and activities involved in the IT service operation. However, different levels of control and automation are required to timely execute expected activities, considering the complexity of the NGDIs, the service quality requirements, the necessary volumes of assets and data and particularly, the demanded fast response times.

3.2 Control Requirements for Service Operations Control

In service operations, the supervisory control system manages the collective strategy, of the telecommunication assets, to deliver services as requested by the user. Together with academic and industrial partners of the NG-CDI1 programme, we identified the requirements for a supervisory control system in the context of service operations. Below, we describe the requirements we focus on, which are aligned to those identified in the literature (Wang et al., 2017; Qian et al., 2018; Saadon et al., 2019).

Req1 Multi-service flexibility - The control system should enable configuration of multiple types of services with different properties, workflows, resources, assets, performance indicators and target levels.

Req2 Service visibility and aggregation - The control system should enable continuous evaluation of service performance against agreed levels. As service delivery brings multiple resources, assets, processes and roles, creating and maintaining an end-to-end service view implies aggregation of multiple data sources.

Req3 Online anomaly management - The control system should detect anomalies during operation and manage them effectively. An effective anomaly management should lead to better and timely fault detection.

Req4 Multi-demand profile tolerance - The control system should be aware of variations in the service activity and therefore in the operational measurements due to natural demand profiles – e.g. seasonal fluctuations.

In addition to these requirements, next generation digital infrastructures bring particular challenges, that given the complexity of current architectures and demands of new services, are still open (Qian et al., 2018; Steiner et al., 2016; Saadon et al., 2019). These challenges lead to other desired features for the control system such as support to multiple timescales and control levels, independence of the network under control and a bounded resource usage.

4. AGENT-BASED CONTROL CANDIDATES

The inherent distribution of operational data, resources and decision points, make this service operations control, suitable for an agent-based approach. The agents provide flexibility to the control architecture while enable integration of different data sources and control domains. We designed the following candidate architectures for evaluation.

Hierarchical (HR): As shown in Fig. 2, it has two main agents: Service Manager (SM) that maps one-to-one with services and Digital Asset (DA) that maps one-to-one with network elements. SM performs service aggregation, service view generation and supervision control decision-making. Inbound and outbound interfaces to low-level network controllers enable real-time data collection and actuation, respectively. A Function Provisioner (FP) agent works as a resource mediator between DAs and SMs.

Heterarchical-Ring (HT-R): This architecture has a ring topology as shown in Fig. 3. There are only DA agents also mapped one-to-one to the network elements. They perform data collection, time-based aggregation and keep partial view of the services. Each DA is connected directly to other two neighbour agents with whom DAs exchange oper-
d) Interaction model: The hierarchical architecture enables a simple master-slave communication and coordination model (Bongaerts et al., 2000). However, this is more complex in the heterarchical ones and opens up various design choices (Bedrouni et al., 2009), e.g. based on available connectivity.

5.2 Design Criteria & Candidate Evaluation

The evaluation of candidates from a design perspective is summarised in Table 1 and discussed below.

Table 1. Design Evaluation

<table>
<thead>
<tr>
<th>Criteria / Candidate Architecture</th>
<th>HR</th>
<th>HT-R</th>
<th>HT-CN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Req1 Multi-service flexibility</td>
<td>★★</td>
<td>★★</td>
<td>★★</td>
</tr>
<tr>
<td>Req2 Service visibility &amp; aggregation</td>
<td>★★★</td>
<td>★</td>
<td>★★★</td>
</tr>
<tr>
<td>Req3 Online anomaly management</td>
<td>★★★</td>
<td>★</td>
<td>★★★</td>
</tr>
<tr>
<td>Req4 Multi-demand tolerance</td>
<td>★★★</td>
<td>★</td>
<td>★★★</td>
</tr>
<tr>
<td>Clear traceability with controlled digital infrastructure</td>
<td>★★★</td>
<td>★★★</td>
<td>★★★</td>
</tr>
<tr>
<td>Distribution of operational responsibilities</td>
<td>★</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Distribution of control decision making</td>
<td>★★★</td>
<td>★</td>
<td>★★★</td>
</tr>
<tr>
<td>Communication/coordination overhead*</td>
<td>★</td>
<td>★</td>
<td>★★★</td>
</tr>
<tr>
<td>Runtime flexibility</td>
<td>★★★</td>
<td>★★★</td>
<td>★★★</td>
</tr>
</tbody>
</table>

**Requirements Coverage** The main criteria are based on how well the design addresses the requirements identified. The multi service flexibility is better addressed by the HR candidate with one SM per service. It provides proper organisation, separation of concerns and is fully aligned with the service view. As heterarchical architectures do not have a master, each DA agent combines monitoring of individual asset and supervision of one or several services simultaneously. In the HR candidate, a new service can be added without overloading the DA agents. The opposite occurs in heterarchical candidates imposing a constraint in the number of services. The HR and HT-CN candidates bring best visibility of services enabling aggregation of data from each asset. Thanks to this visibility, the anomaly management is simple in the HR candidate and does not require further coordination. However, the HT-CN requires all the agents connected and dealing with a high number of messages between them which leads to scalability constraints. The multi demand tolerance is harder to achieve in HT-R candidate, as additional monitoring data comes only from neighbours, then the observed behaviour might not reflect overall service conditions.

**Asset Traceability** The mapping between agents and the assets varies from one-to-one to one-to-many or many-to-many. The granularity of this relationship depends on how independent the asset is in its behaviour from others and its influence over the overall infrastructure. To favour the clarity of traceability, for the candidates proposed, we selected the one-to-one mapping.

**Distribution of Responsibilities** Distribution favours flexibility as changes in one function not necessarily affect others. For example, using multiple anomaly detection algorithms not necessary requires to change the way that assets are monitored. The HR candidate provides higher distribution of responsibilities with more agents involved.

![Fig. 5. Emulated Network Infrastructure](image)

**Distribution of Control** Multiple agents with different control decision-making capabilities provide better separation for the control levels that the infrastructure requires. So, these control decisions can be taken at different time scales and contexts with no interference to each other.

**Communication Overhead** The master/slave structure of the HR candidate is enough to ensure coordination. In the heterarchical candidates, the initial decision is the communication topology and the model to maintain the links with each other agent. Furthermore, the coordination approach enables agents to reach consensus on actions to take, given the condition of the system. The latter approaches are more complex from a design perspective and then require the design of protocols and mechanisms to enable the communication and coordination, which brings additional workload to DA agents.

**Runtime flexibility** The agent-based control approach intrinsically provides runtime flexibility, however this is increased when agents do not depend fully on each other. In the case of the HR candidate, the SMs and DAs depend on the FP agent in order to locate each other. In the heterarchical candidates agents are more independent and therefore do not rely on a particular agent that enables communication and coordination among them.

6. CONTROL OPERATION EVALUATION

This evaluation is based on the control system prototypes we built with each architecture. The evaluation scenario and the experimental design are presented below.

6.1 Problem Description

We aim to measure the performance of different agent-based supervisory control systems while controlling a digital network infrastructure in operation. Suppose that a temporal network is provisioned for the video streaming of a sports event. It requires a high throughput data transmission between the content servers and clients. Such network is illustrated in Fig. 5, forming a ring topology where switches under supervision are rounded rectangles and hosts are circles. Servers h1 and h4 serve contents to clients h9 and h7, respectively. We assume end-users use h4 and h7 as cache servers for video streaming. The control
system monitors switches $s1$ to $s10$ to avoid any reduction of the throughput and trigger the corrective action timely. In this case this action is to serve contents from another available server. The low-level network control is enough to guarantee service provision but the supervisory control ensures that target throughput levels are met. This problem gives room to demonstrate operation considering the requirements identified in section 3.2.

6.2 Experimental Design

We simulated traffic from content providers to clients and evaluate the performance of the control systems built from the candidate architectures. The experimental set up is summarised in Table 2. Each system went through periods of standard operation and simulated failures. The digital network topology and the traffic were kept stable.

We focused evaluation on three characteristics: the resource usage, measured in MB of RAM used by the agents, the communication overhead, given by the bytes transmitted within the system, and the reaction time, measured between point $t_0$, when the traffic is sensed, and the point $t_1$, when an anomaly is detected. For each characteristic, we calculated a performance indicator relative to the best observation as follows. First, we scaled the observations of each characteristic using the absolute maximum ($\text{MaxAbs}$) method (Bonaccorso, 2017). Then, we calculated the vector $D$ with the differences of each observation $X_i$ from the best observation, in this case, $X_{\text{min}}$. Finally, as the lowest difference indicates better performance and we wanted a value in the $[0 - 100]$ interval, we applied the equation 1.

$$P(D) = (1 - D_i) \cdot 100$$

6.3 Prototypes and Emulation Environment

For each architecture described in section 4 an agent-based control system prototype has been developed in Scala. We used the actor model abstractions as the fundamental building blocks. Actors have been widely used in engineering of distributed system and have shown benefits for handling concurrency and managing programs running in different network nodes (De Koster et al., 2016). We used a toolset² that included akka actors for building containerised agents, with mininet providing a realistic emulated network and Ryu as the low-level SDN controller.

Fig. 6 shows the performance of the prototypes of each architecture. Overall, the $HR$ architecture outperforms the other candidates, especially in the reaction time. The $HR$ candidate is slightly better in RAM usage and shows a high performance in communication overhead. The second best is the $HT-R$ architecture, with performance above 97% in RAM usage and nearly 10% more in communication overhead than the $HR$ option. We now discuss the results obtained for each evaluated characteristic.

Table 3. Reaction Time (Seconds)

<table>
<thead>
<tr>
<th>Candidate</th>
<th>Mean</th>
<th>Std</th>
<th>Min</th>
<th>Max</th>
<th>% Perf.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hierarchical (HR)</td>
<td>6.00</td>
<td>0.15</td>
<td>5.88</td>
<td>6.25</td>
<td>99.22</td>
</tr>
<tr>
<td>Heterarchical-Ring (HT-R)</td>
<td>11.88</td>
<td>0.52</td>
<td>11.25</td>
<td>12.38</td>
<td>62.5</td>
</tr>
<tr>
<td>Heterarchical-Complete Network (HT-CN)</td>
<td>14.57</td>
<td>1.36</td>
<td>12.33</td>
<td>16.00</td>
<td>45.68</td>
</tr>
</tbody>
</table>

Reaction Time In Fig. 6 a higher performance means a lower reaction time. The Table 3 shows that best mean reaction time along the runs is 6 seconds, for the prototype with the $HR$ architecture. This is explained, as DAs have predefined SMs they need to report their readings to and also the SMs are specialised in deciding and triggering the actions for the anomalies. This is contrary to the heterarchical architectures where DAs have several responsibilities, including network elements monitoring, data aggregation, decision-making and triggering of actions.

Table 4. RAM Usage (MB)

<table>
<thead>
<tr>
<th>Candidate</th>
<th>Mean</th>
<th>Std</th>
<th>Min</th>
<th>Max</th>
<th>% Perf.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hierarchical (HR)</td>
<td>287.83</td>
<td>0.68</td>
<td>286.89</td>
<td>288.51</td>
<td>99.68</td>
</tr>
<tr>
<td>Heterarchical-Ring (HT-R)</td>
<td>295.56</td>
<td>1.20</td>
<td>293.65</td>
<td>296.64</td>
<td>97.08</td>
</tr>
<tr>
<td>Heterarchical-Complete Network (HT-CN)</td>
<td>291.89</td>
<td>2.69</td>
<td>289.37</td>
<td>295.91</td>
<td>98.31</td>
</tr>
</tbody>
</table>

RAM Usage Table 4 summarises RAM usage. The three candidates perform similarly, with the $HR$ candidate using slightly less RAM than the others. Considering the available RAM (350MB), the differences among prototypes are low as shown by the performance index with values above 98%. There is no substantial difference in RAM usage among the architectures evaluated.

Table 5. Communication Overhead (MB)

<table>
<thead>
<tr>
<th>Candidate</th>
<th>Mean</th>
<th>Std</th>
<th>Min</th>
<th>Max</th>
<th>% Perf.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hierarchical (HR)</td>
<td>1.24</td>
<td>0.02</td>
<td>1.20</td>
<td>1.27</td>
<td>88.62</td>
</tr>
<tr>
<td>Heterarchical-Ring (HT-R)</td>
<td>1.04</td>
<td>0.02</td>
<td>1.01</td>
<td>1.07</td>
<td>98.50</td>
</tr>
<tr>
<td>Heterarchical-Complete Network (HT-CN)</td>
<td>1.91</td>
<td>0.11</td>
<td>1.75</td>
<td>2.04</td>
<td>55.69</td>
</tr>
</tbody>
</table>

Fig. 6. Performance Comparison of Architectures

<table>
<thead>
<tr>
<th>Candidate</th>
<th>Mean</th>
<th>Std</th>
<th>Min</th>
<th>Max</th>
<th>% Perf.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hierarchical (HR)</td>
<td>1.24</td>
<td>0.02</td>
<td>1.20</td>
<td>1.27</td>
<td>88.62</td>
</tr>
<tr>
<td>Heterarchical-Ring (HT-R)</td>
<td>1.04</td>
<td>0.02</td>
<td>1.01</td>
<td>1.07</td>
<td>98.50</td>
</tr>
<tr>
<td>Heterarchical-Complete Network (HT-CN)</td>
<td>1.91</td>
<td>0.11</td>
<td>1.75</td>
<td>2.04</td>
<td>55.69</td>
</tr>
</tbody>
</table>
Communication Overhead  Table 5 summarises the mean of bytes transferred per agents in each architecture. The mean is particularly high for the $HT-CN$ architecture, which almost double the others. This is explained as DAs in $HT-CN$ manage connections with each other therefore requiring extensive exchange of messages. The other candidates show similar figures under 1.25 MB transmitted. Note the standard deviation is also low among the runs per architecture. The $HT-CN$ architecture performs around a 42% worse than the $HT-R$.

The results of the prototype evaluation are aligned with our initial design evaluation. The dispersion across the runs is low, which give us confidence especially in the communication overhead and reaction time metrics. Ordered from best to worst performance the architectures are hiearchical, heterarchical-ring and heterarchical-complete network. The hierarchical candidate easily provides a global service view that is more complex and expensive to achieve in the heterarchical-complete network candidate.

8. CONCLUSIONS

We evaluated three agent-based control candidate architectures for the supervision of digital infrastructures. In addition, we proposed an approach for evaluating these type of architectures from a design and operational perspective. In the design perspective, we considered: requirements-coverage, asset traceability, distribution of responsibilities and control. For the operational evaluation, we built three prototypes, one per each candidate architecture and set a realistic environment based on an emulated network. We measured resource-usage, with focus in RAM, communication overhead and the reaction time. We built a performance index based on the difference of observations of each evaluated characteristic to the best value obtained. Our results show that the hierarchical (HR) followed by the heterarchical-ring (HT-R) candidate architectures perform better in the anomaly detection scenario, over a digital infrastructure, in comparison to the heterarchical-complete network (HT-CN) architecture.

REFERENCES


