

DEVELOPMENT OF A PARAMETRIC-DECOMPOSITION  
METHODOLOGY FOR SOLVING QUEUEING NETWORKS WITH  
SIMULTANEOUS RESOURCE POSSESSION UNDER CAPACITY  
RESTRICTIONS

By

RONNY PACHECO

Licenciature in Industrial Engineering  
University of Costa Rica  
San José, Costa Rica  
2000

Master of Science in Industrial Engineering  
University of Costa Rica  
San José, Costa Rica  
2012

Submitted to the Faculty of the  
Graduate College of  
Oklahoma State University  
in partial fulfillment of  
the requirements for  
the Degree of  
DOCTOR OF PHILOSOPHY  
December, 2020

COPYRIGHT ©

By

RONNY PACHECO

December, 2020

DEVELOPMENT OF A PARAMETRIC-DECOMPOSITION  
METHODOLOGY FOR SOLVING QUEUEING NETWORKS WITH  
SIMULTANEOUS RESOURCE POSSESSION UNDER CAPACITY  
RESTRICTIONS

Dissertation Approved:

Dr. Manjunath Kamath

---

Dissertation Advisor

Dr. Tieming Liu

---

Dr. Fazard Yousefian

---

Dr. James Pappas

---

*Dedicated to Hernán, Rafaela, Kathy and Amanda.*

---

The dedication reflects the views of the author and are not endorsed by committee members or Oklahoma State University.

## ACKNOWLEDGMENTS

Pursuing a PhD is an intellectual and personal challenge. In this journey, I was lucky enough to find people who made all the effort worth it, even more than the title itself. And maybe, this acknowledgement is not long enough to include all the people who deserve to be thanked; I beg your pardon.

First, I would like to thank Dr. Kamath, my tutor, for believing in me. His patience, guidance and support were invaluable for me, and his knowledge and expertise have been inspiring. This dissertation would not have been possible without his direction; I owe him a lot. Thank you as well to my committee members for their support and advice. Thank you Dr. Yousefian, Dr. Liu and Dr. Pappas.

I would like to thank Dr. DeYong for always being there. Her support and advice has made all this experience special. Next, I would like to thank Dr. Collins who always has a friendly and supportive comment to inspire me. Thank you to Dr. Baski and Dr. Buchanan for allowing me to sneak into their optimization classes and learn a lot from them and for the nice talks over lunch. Thank you Dr. Heragu for all your support and advice. I owe a special thank you to Laura Brown and Regina Henry who were always there to help. Without their help and smiles this process would have been really different.

---

Acknowledgments reflects the views of the author and are not endorsed by committee members or Oklahoma State University.

I would like to recognize some of my friends: thank you Babak, Utara, Mostafa, Akash, Kushal and Harshal for sharing your time and advice with me. Thank you Hernán, Miguel, Adriana, Juan, Shahrouz, April and Oscar, as well. You made my life much more easier, colorful, and joyful.

Thank you Gabby, Ryan, Dr. Green, and Deborah. I can't thank you enough for all you did for me and my family. Thank you to Sofi and Rheiny, to the Hispanic community in Stillwater and the Costa Rican community in Tulsa. Thank you Tim and Amy. Thank you all for making my family and me part of yours. We feel at home around you.

I need to thank the University of Costa Rica for all its support in this process.

Finally, I give a very special thank you to my wife; I am thankful for her love, comfort and endless support. I would like to thank my parents and my brother and his family for all their support.

---

Acknowledgments reflects the views of the author and are not endorsed by committee members or Oklahoma State University.

Name: RONNY PACHECO

Date of Degree: December, 2020

Title of Study: DEVELOPMENT OF A PARAMETRIC-DECOMPOSITION METHODOLOGY FOR SOLVING NETWORKS OF QUEUES WITH SIMULTANEOUS RESOURCE POSSESSION UNDER CAPACITY RESTRICTIONS

Major Field: INDUSTRIAL ENGINEERING AND MANAGEMENT

Motivated by applications in health care systems, this study focused on queueing network models with instances of simultaneous resource possession under capacity restrictions in different stages. Customers during service at certain nodes in the first-level or primary system may simultaneously need service from a second-level server. In this simultaneous resource possession situation, the time the customer spends waiting for the second-level server will impact the overall service time at first-level server and consequently the performance of the entire system. Closed queueing networks and fork/join approximations for one and two-stage systems were used to capture the capacity restriction effects in different sections of the primary system. An iterative algorithm was developed to incorporate the effect of the second-level server on the performance of the entire system. This study used a modular approach rooted in the parametric-decomposition method and two-moment approximations. Different systems showing the proposed building blocks and their combinations to solve more complex systems illustrated the application of the proposed modeling approach.

To evaluate the performance prediction accuracy of the solution algorithms, the analytical results were compared to simulation estimates for several configurations with single and multi-server nodes, a wide range of service time variability, and different levels of capacity constraints. The analytical results tracked the simulation estimates well with errors less than 10% in more than 80% of the configurations simulated. Higher errors in the 15% to 20% range were observed for system with low capacity limits, high service time variability (SCV=2), and high demand ( $\rho=0.75$ ) for the second-level server.

## TABLE OF CONTENTS

Chapter	Page
<b>I INTRODUCTION</b>	<b>1</b>
1.1 Motivation . . . . .	3
1.2 The Solution Approach . . . . .	8
1.3 Outline of the Dissertation . . . . .	10
<b>II BACKGROUND AND LITERATURE REVIEW</b>	<b>12</b>
2.1 The Machine Interference and the Extended Machine Interference Problems .	13
2.2 The Simultaneous Resource Possession (SRP) Problem . . . . .	14
2.3 Other Approaches for the SRP Problem . . . . .	15
2.4 Semi-Open Queueing Networks and Open Queueing Networks with Population Constraints . . . . .	16
2.5 Literature Gap . . . . .	18
<b>III STATEMENT OF RESEARCH</b>	<b>20</b>
3.1 Problem Description . . . . .	20
3.2 Technical Challenges . . . . .	22
3.3 Objectives and Tasks . . . . .	23
<b>IV THE SRP BUILDING BLOCK</b>	<b>27</b>
4.1 The SRP Solution Approach . . . . .	28
4.2 Notation and System Parameters . . . . .	30



Chapter	Page
4.3 Mean and SCV of Equivalent Service Time for a Server at a First-Level Node Affected by SRP . . . . .	32
4.3.1 Expectation of the Modified Service Time at the Node Influenced by the SRP system . . . . .	33
4.3.2 SCV of the Modified Service Time at the Node Influenced by the SRP	33
4.4 Calculating Node-Level Performance Measures . . . . .	34
4.5 Corrections for the Average Waiting Time in Closed Networks . . . . .	38
4.6 Variance of Waiting Time in Queue . . . . .	40
4.7 General Algorithm for Solving the SRP System . . . . .	44
4.8 SRP Building Block with Multiple Sources . . . . .	47
<b>V CLOSED QUEUEING NETWORK (CQN) WITH MULTI-SERVER NODES</b>	<b>50</b>
5.1 Multi-server Extension . . . . .	51
5.1.1 The Original PDMCQN-2 Algorithm . . . . .	51
5.1.2 Modifications to the PDMCQN-2 Algorithm to Include Multi-Server Nodes . . . . .	53
5.2 General CQN Results . . . . .	54
5.2.1 Error Calculation . . . . .	56
5.2.2 Determination of Simulation Warm Up Period and Run Length . . .	58
5.2.3 Numerical Results . . . . .	58
<b>VI DOUBLE AND SINGLE FORK/JOIN SYSTEMS WITH MULTI-SERVER NODES</b>	<b>63</b>
6.1 The Fork/Join System Solution . . . . .	66
6.1.1 The Fork/Join System Solution with One Synchronization Station . .	75
6.2 Fork-Join Numerical Experiments . . . . .	77

Chapter	Page
6.2.1 Double Fork-Join Results . . . . .	81
6.2.2 Single Fork-Join Results . . . . .	86
<b>VII LINKING TWO STAGES WITH INDIVIDUAL CAPACITY RESTRICTIONS</b>	<b>91</b>
7.1 The Solution Approach . . . . .	92
7.2 Numerical Experiments for the Two-Stage System . . . . .	95
7.2.1 Numerical results for the Two-Stage System . . . . .	99
<b>VIII SOLVING CLOSED QUEUEING NETWORKS WITH SRP</b>	<b>106</b>
8.1 Closed Queueing Network With Simultaneous Resource Possession . . . . .	106
8.2 The Solution Approach for CQN with SRP . . . . .	108
8.3 CQN with SRP Numerical Experiment . . . . .	111
8.3.1 Experimental Set Up . . . . .	111
8.4 Numerical Results . . . . .	113
<b>IX SINGLE-STAGE FORK/JOIN SYSTEM WITH SRP</b>	<b>127</b>
9.1 Algorithm for Solving the Single-Stage Capacity Restricted System with SRP	128
9.2 Numerical Experiments for Single-Stage Fork/Join System with SRP . . . . .	131
9.2.1 Parameter selection . . . . .	131
<b>X TWO CAPACITY-RESTRICTED STAGES WITH A SHARED SRP SYSTEM</b>	<b>148</b>
10.1 The Solution Approach . . . . .	149
10.2 Experimental Set Up . . . . .	152
10.2.1 Experimental Parameters . . . . .	152
10.2.2 Experiment Results . . . . .	156

Chapter	Page
<b>XI CONCLUSIONS AND FUTURE WORK</b>	<b>180</b>
11.1 Accuracy of Performance Prediction . . . . .	185
11.2 Contributions . . . . .	186
11.3 Future Work . . . . .	187
<b>A Verification of Fork/Join Algorithm Implementation</b>	<b>198</b>
<b>B Additional Results for the Two Capacity-Restricted Stages with a Shared     SRP System</b>	<b>201</b>

## LIST OF TABLES

Table	Page
4.1 Second Level System Notation . . . . .	31
4.2 Rules for multi-server waiting time correction ( $\theta$ ) . . . . .	40
5.1 CQN Configurations . . . . .	56
5.2 CQN Throughput Results . . . . .	59
5.3 CQN Utilization Results . . . . .	60
5.4 CQN Average Number in Node Results . . . . .	61
6.1 Additional notation needed for the fork/join system . . . . .	65
6.2 Double FJ Experimental Configurations . . . . .	78
6.3 Arrival Parameters for the Double FJ Experiments . . . . .	79
6.4 Percentages of CQN Throughput Used to Set the DFJ Arrival Rates . . . . .	80
6.5 Double Fork-Join Throughput Results . . . . .	82
6.6 Double Fork-Join Utilization Results . . . . .	83
6.7 Double Fork-Join Average Number in Node Results . . . . .	84
6.8 Double FJ Mean Number in Queue for Fork-Join Nodes . . . . .	85
6.9 Single Fork-Join Throughput Results . . . . .	87
6.10 Single Fork-Join Utilization Results . . . . .	88
6.11 Single Fork-Join Average Number in Node Results . . . . .	89
6.12 Single FJ Number in Queue for Fork-Join Node . . . . .	90
7.1 Arrival Process Parameters for the Two-stage Configurations . . . . .	96

Table	Page
7.2 Two-stage Experiments: Node Parameters for the First Stage . . . . .	97
7.3 Two-Stage Experiments: Node Parameters for the Second Stage . . . . .	98
7.4 Throughput Results for the Two-Stage System . . . . .	100
7.5 Utilization Results for Stage 1 . . . . .	101
7.6 Utilization Results for Stage 2 . . . . .	102
7.7 Average Number in Node Results for Stage 1 . . . . .	103
7.8 Average Number in Node Results for Stage 2 . . . . .	104
7.9 Average Number of Customers in Queue at the Fork-Join Nodes . . . . .	105
8.1 Additional Notation for the CQN-SRP Model . . . . .	108
8.2 CQN with SRP Parameters for the First Level Nodes . . . . .	112
8.3 CQN with SRP Parameters for the SRP Node . . . . .	113
8.4 Throughput Results for $p = 25\%$ . . . . .	114
8.5 Throughput Results for $p = 50\%$ . . . . .	115
8.6 Throughput Results for $p = 75\%$ . . . . .	116
8.7 First-Level Node Utilization Results for $p = 25\%$ . . . . .	117
8.8 First-Level Node Utilization Results for $p = 50\%$ . . . . .	118
8.9 First-Level Node Utilization Results for $p = 75\%$ . . . . .	119
8.10 First Level Average Number in Node Results $p = 0.25$ . . . . .	120
8.11 First Level Average Number in Node Results $p = 0.5$ . . . . .	121
8.12 First Level Average Number in Node Results $p = 0.75$ . . . . .	122
8.13 Utilization and average number of customers in node performance metrics for the SRP node $p = 0.25$ . . . . .	123
8.14 Utilization and average number of customers in node performance metrics for the SRP node $p = 0.5$ . . . . .	124

Table	Page
8.15 Utilization and average number of customers in node performance metrics for the SRP node $p = 0.75$ . . . . .	125
9.1 External Arrival Process Parameters . . . . .	132
9.2 Throughput Results for $p=0.25$ . . . . .	134
9.3 Throughput Results for $p=0.5$ . . . . .	135
9.4 Throughput Results for $p=0.75$ . . . . .	136
9.5 First Level Node Utilization Results for $p=0.25$ . . . . .	137
9.6 First Level Node Utilization Results for $p=0.50$ . . . . .	138
9.7 First Level Node Utilization Results for $p=0.75$ . . . . .	139
9.8 Average Number in Node Results for $p = 0.25$ . . . . .	140
9.9 Average Number in Node Results for $p = 0.50$ . . . . .	141
9.10 Average Number in Node Results for $p = 0.75$ . . . . .	142
9.11 SRP Node and Synchronization Node Results for $p=0.25$ . . . . .	143
9.12 SRP Node and Synchronization Node Results for $p=0.50$ . . . . .	144
9.13 SRP Node and Synchronization Node Results for $p=0.75$ . . . . .	145
10.1 Arrival Rates . . . . .	153
10.2 Experimental Parameters for the First Stage . . . . .	154
10.3 Experimental Parameters for the Second Stage . . . . .	155
10.4 Probability of First-Level Nodes Using the SRP Subsystem . . . . .	156
10.5 Throughput Results for First Set of Configurations . . . . .	158
10.6 Throughput Results for Second Set of Configurations . . . . .	159
10.7 Throughput Results for the Third Set of Configurations . . . . .	160
10.8 Utilization Results for Stage 1, First Set of Configurations . . . . .	161
10.9 Utilization Results for Stage 1, Second Set of Configurations . . . . .	162

Table	Page
10.10 Utilization Results for Stage 1, Third Set of Configurations . . . . .	163
10.11 Average Number in Node Results for Stage 1, Set of Configurations 1 . . . . .	164
10.12 Average Number in Node Results for Stage 1, Set of Configurations 2 . . . . .	165
10.13 Average Number in Node Results for Stage 1, Set of Configurations 3 . . . . .	166
10.14 Utilization Results for Stage 2, First Set of Configurations . . . . .	167
10.15 Utilization Results for Stage 2, Second Set of Configurations . . . . .	168
10.16 Utilization Results for Stage 2, Third Set of Configurations . . . . .	169
10.17 Average Number in Node Results for Stage 2, Set of Configurations 1 . . . . .	170
10.18 Average Number in Node Results for Stage 2, Set of Configurations 2 . . . . .	171
10.19 Average Number in Node Results for Stage 2, Set of Configurations 3 . . . . .	172
10.20 Average Number in Node Results for Fork/Join nodes, Set of Configurations 1	173
10.21 Average Number in Node Results for Fork/Join nodes, Set of Configurations 2	174
10.22 Average Number in Node Results for Fork/Join nodes, Set of Configurations 3	175
10.23 Results for the SRP node, Set of Configurations 1 . . . . .	176
10.24 Results for the SRP node, Set of Configurations 2 . . . . .	177
10.25 Results for the SRP node, Set of Configurations 3 . . . . .	178
A.1 Double Fork/Join Algorithm Verification for $\lambda_D$ Results . . . . .	199
A.2 Double Fork/Join Algorithm Verification for $L_3$ Results . . . . .	200
B.1 Throughput Results for First Set of Configurations . . . . .	201
B.2 Throughput Results for Second Set of Configurations . . . . .	202
B.3 Throughput Results for the Third Set of Configurations . . . . .	203
B.4 Utilization Results for Stage 1, First Set of Configurations . . . . .	204
B.5 Utilization Results for Stage 1, Second Set of Configurations . . . . .	205
B.6 Utilization Results for Stage 1, Third Set of Configurations . . . . .	206

Table	Page
B.7 Average Number in Node Results for Stage 1, Set of Configurations 1 . . . . .	207
B.8 Average Number in Node Results for Stage 1, Set of Configurations 2 . . . . .	208
B.9 Two Stages FJ With Shared SRP Node. Average Number in Node Results for Stage 1, Set of Configurations 3 . . . . .	209
B.10 Utilization Results for Stage 2, First Set of Configurations . . . . .	210
B.11 Utilization Results for Stage 2, Second Set of Configurations . . . . .	211
B.12 Utilization Results for Stage 2, Third Set of Configurations . . . . .	212
B.13 Average Number in Node Results for Stage 2, Set of Configurations 1 . . . . .	213
B.14 Average Number in Node Results for Stage 2, Set of Configurations 2 . . . . .	214
B.15 Average Number in Node Results for Stage 2, Set of Configurations 3 . . . . .	215
B.16 Average Number in Node Results for Fork/Join nodes, Set of Configurations 1	216
B.17 Average Number in Node Results for Fork/Join nodes, Set of Configurations 2	217
B.18 Average Number in Node Results for Fork/Join nodes, Set of Configurations 3	218
B.19 Results for the SRP node, Set of Configurations 1 . . . . .	219
B.20 Results for the SRP node, Set of Configurations 2 . . . . .	220
B.21 Results for the SRP node, Set of Configurations 3 . . . . .	221



## LIST OF FIGURES

Figure	Page
3.1 Example System Configuration 1 . . . . .	21
3.2 Example System Configuration 2 . . . . .	22
4.1 A First Level Network Connected to a SRP Subsystem . . . . .	27
4.2 Approximation used in the node affected by the SRP system . . . . .	29
4.3 Representation of the second-level system . . . . .	31
4.4 SRP System Solution Approach . . . . .	45
4.5 Multi-Input SRP . . . . .	48
5.1 Network Structure Used For Testing Purposes . . . . .	55
6.1 Double Fork/Join Example . . . . .	63
6.2 Fork/Join Notation . . . . .	65
7.1 Two-stage Fork/Join Example . . . . .	92
7.2 System's Stability Check . . . . .	93
8.1 CQN With SRP Example . . . . .	107
9.1 Single-Stage Fork/Join System With Simultaneous Resource Possession . . .	128
10.1 Two Capacity-Restricted Stages with a Shared SRP System . . . . .	148
10.2 Two-Stage with SRP System Solution Approach . . . . .	150
11.1 Some Possible Problem Extension . . . . .	188

## CHAPTER I

### INTRODUCTION

It is common in service and manufacturing environments to find people, parts, or requests waiting in queues. The most basic schema contains a queue of customers waiting in front of a server. People all over the world experience these types of systems daily, for instance, when utilizing public or private transportation services, visiting an ATM, or having a medical appointment.

Let us imagine a scenario where a patient is received in an emergency room (ER). The patient is first assigned to a bed, then the nurse and the ER physician provide medical care to the patient. So far, the patient has had to wait for the bed, the nurse and the physician to become available, given they could have been busy with other patients at the moment the new patient arrived. First, the nurse and the physician work to diagnose the patient and stabilize him/her. Note that the patient is holding, at least, three resources already (bed, nurse, and ER physician). Now, the team could find a more complex situation, and ask for an expert opinion. In this case, they all (patient, nurse, and ER physician) may have to wait for the specialist's answer or intervention. All the waiting times discussed so far can be seen as time the patient is waiting in queues in order to receive different services.

Depending on the patient's condition, the nurse and the ER physician may have to wait along with the patient for the specialist, and this waiting prevents them from attending to other patients. If this happens, the patient will be holding simultaneously at least four

resources: bed, nurse, ER physician, and specialist. This case is different from other networks of queues because once the customer receives service from a server, he/she does not release it and move to the next server, but rather moves with the original server to the next server, waits in queue, and receives the second-level service, everything without releasing the resources acquired first. This prolongs the first-level resource's service time by an amount of time that cannot be known in advance.

In this example, there are at least two levels of service. The ER physician and the nurse are the servers at the first level and the specialist at the second level. This document addresses systems in which more than one server may be required simultaneously in order to provide the service. The first level server will be called non-overlapped server, and the second level server will be called overlapped server, using names suggested by Jacobson and Lazowska [26]. Another way to address this system is first layer server and second and higher layer servers, following the layered queueing networks' nomenclature [19]. In this document, such nomenclature will be used interchangeably.

Many articles have been written that address problems related to requirements of resources in different layers. In these systems, the customers acquire one non-overlapped resource and while holding it, look for one or more overlapped ones. Some examples are machine interference problem (MIP), machine repairman problem (MRP), extended machine interference problem (EMIP), simultaneous resource possession (SRP), and layered queueing networks (LQN).

This study focused on cases of overlapped requirements of resources, that have two fundamental challenges: first, customers need so called passive resources in order to get accepted into the different system sections. There is no service time associated with the

passive resource. Examples of passive resources are beds and rooms in hospitals and pallets in production systems. A customer needs one to get entry into a system or sub-system and holds it as long as the customer is in the sub-system, and simultaneously receives service from active resources during this time. A customer may need to wait for a passive resource because of their limited number. The second challenge occurs once the customer is in the system and holding passive resources. This customer could find situations that require more than one active resource simultaneously to receive a specific service, which implies a SRP condition. An active resource is a resource that has a positive service time.

The variations of this problem can be diverse. Therefore, the goal of this study was to identify and explain how to solve a group of the problem's configurations that can be utilized as building blocks to represent a broader range of possible real-life problems.

## **1.1 Motivation**

The systems studied in this document are likely to be found in many different real-life environments, for instance, hospitals, factories, warehouses, and computer systems. These systems are characterized by two main features. First, there exists a resource that is passive and limits the physical capacity of the system; this could be beds or rooms in a hospital, kanban cards on the shop floor, or storage locations in a warehouse. These resources are typically limited and difficult to increase, because of cost or space constraints. Consequently, their proper utilization becomes a priority for the system operation. Secondly, instances of simultaneous resource possession with the overlapping of resources in different levels, makes the analysis of the overall system challenging. These overlapped resources can be physicians waiting for equipment or other physicians in the hospital; machines waiting for tools, personnel or material handling equipment in the factory; and material handling equipment waiting for

personnel in the warehouse.

In a capacity restricted environment, the system's throughput, customers' waiting times, and resources' utilization are difficult metrics to obtain given the presence of high variability in such systems. The time the resource limiting the capacity remains busy is dependent on the other resources' service times, and the number of customers in each station. Consequently, the customers' flow into the system, which is a combination of the external arrivals and the availability of the resource limiting the capacity, becomes a challenging process to model. On the other hand, the presence of SRP conditions involves multiple layers of servers, and servers waiting in lines belonging to other servers. The superposition of both systems, makes this problem particularly challenging. This combination is at the heart of the real-life problems discussed in this section.

All the complications related with the analysis of these systems could make them appear as rare exceptions. However, they are relatively common. Both characteristics explained previously are quite frequent. For instance, hospital beds, hospital rooms, containers for shipping, and parking or storage spaces are typically limited. On the other hand, as resources become specialized, they become expensive and scarce and have to be shared across multiple sub-systems giving rise to complex server-to-server interactions.

In spite of all these different applications, this document was highly motivated by health care applications and the societal benefits produced by improving health care services in many communities.

### **Relevance for the Health Care Environment**

In health care environments, resources are scarce, costly, and difficult to substitute,

mostly because of their highly specialized functionality. This combination generates resources with busy schedules and high demand, impacting negatively on other resources and patients, that usually have to wait for them to be available, in order to continue their activities.

There exists a trade-off between return on investment and customer satisfaction that makes these types of problems interesting to study and valuable for society. It is easy to understand that managers and investors are interested in achieving high utilization of costly resources. This way, they assure a better return on investments. However, this policy could have some drawbacks such as extending patient's length of stay (LoS) or delaying diagnostics that could have a negative impact on the patient's health or recovery time. These effects are a consequence of the increased waiting times related to highly utilized resources. Patient's health and life style, health care provider's image, and society's productivity could be negatively impacted under these circumstances.

The pressure for an appropriate use of the resources combined with the previously explained problem's sophistication, makes the decision making process in the hospital environment, a difficult one. In the following paragraphs, some examples of these situations are discussed, in order to illustrate how frequently these problems can be found in the hospital environment.

Emergency rooms (ER) are known to be busy places. The emergency room's capacity is usually evaluated in terms of "number of beds" or "treatment rooms". Hence, the number of beds limits the number of patients that can be admitted, some examples can be seen in [38], [24], [14], [59], [1]. Then, the bed or room becomes the capacity-restricting passive resource, necessary to be obtained by the patient in order to be admitted to the ER.

Once the patient is admitted to the ER, she is likely to require service from different

resources, for instance: physicians, nurses, life support equipment, technicians, or diagnosis equipment (ultrasound, X-rays). All of the previously mentioned resources have two characteristics in common. First, all of them are shared by different patients in the department. Secondly, the tasks to be done, or part of them, frequently require more than one resource to be present simultaneously. These conditions make the waiting of one resource for another resource, unavoidable.

Similar conditions can be found in surgical theaters, where patients are waiting for operation rooms to become available. Once the patient is assigned a room, her pathway moves through the surgical theater and to the recovery room or intensive care unit (ICU) and requires the synchronization of a series of resources in different steps of the process. Resources that are often shared by different operation rooms and patients, such as anesthesiologists, nurses, surgeons, in-patient physicians, technicians, stretcher-bearers, cleaning teams, or ICU teams. Some examples are contained in [4] and [37], [9], [60].

In a scenario where urgent surgeries would be needed, all the benefits provided by scheduling tools in order to foresee the resources' availability are lost, and the scarce resources that are shared by the patients become critical. First, the patient has to wait for a surgery room to be vacant; then, according to the other resources' availability, wait until the room is clean, and the surgery team is complete. Hence, from all resources needed for the surgery, some will become available earlier than others, consequently, the first resource to arrive will have to wait until the team is complete to perform the surgery. Here, the operation room is the passive resource and later the other resources such as anesthesiologist, surgeons, nurses, technicians and equipment, become the active resources that the patient holds simultaneously.

We can recall the ER patient example presented in the introduction of this document and

realize that many situations other than those represented in that example could occur. The patient could require other resources such as technicians, lab exams, and high tech equipment. The ER physician could leave the room and see other patients before the specialist arrives. The nurse could have left the room as well, and served other patients. The nurse could have arrived first and waited for the ER physician to arrive, presenting another SRP problem. There is a vast number of scenarios involving the multiple resources needed to bring the service to a patient, all of this embedded in an environment where the capacity at each section of the pathway is a real constraint. All these considerations, make evident the necessity for having a framework, flexible enough, that allows us to model different possible scenarios in the health care environment.

### **Relevance in Other Fields**

The original motivation for studying services with different layers was found in the manufacturing environment, specifically in MRP and MIP. However, the EMIP and SRP development has been highly influenced by the computer science community. For instance, computer science problems were the motivation for approaching SRP from the very first attempts to solve it. Some examples can be seen in [46] and [26]. Nowadays, computer science leads the research in this topic under the name of layered queueing networks [19], [21]. Similar research has been done to study the performance of systems such as web applications [47], data centers [2], and cloud systems [3]. From these examples, it is evident that this topic is relevant in studying the performance of other complex systems.

In computer science, it is common to find scenarios where a customer's request has to wait for a memory partition, and once it is granted, it has to queue for software and hardware resources that are required simultaneously. In this way, this becomes an example of a behavior similar to the hospital case.



Despite the enormous interest in these topics from the computer science field, some applications to other knowledge areas can be found. Authors such as Basnet and Kamath [6], De Almeida and Keller [13], and Suri and Desiraju [49] utilized SRP to model the blocking effect produced by the transportation equipment or material handling equipment in a manufacturing environment. Kamath and Sanders [28] utilized the SRP to model operator/machine interference in asynchronous automatic assembly systems.

Under a kanban production system or any other production environment with controlled work in process (WIP), for instance CONWIP production systems, the process load is controlled by the mechanism in place to control the WIP. Consequently, this becomes a capacity restricted process, meaning the production amount allowed in the production system simultaneously has a limit. Additionally, different SRP situations can be present downstream. In this case, the WIP control mechanism that allows jobs to enter the system, could be modeled as a passive resource. Then the operators/machine relations present when the operators load/unload machines, or release a stocked part in the production process are examples of SRP. Similar situation occurs when the repairmen support machines; or specific molds and tools are required as a complement for machining operations. All these are examples where multiple resources are required simultaneously to perform the required service.

## **1.2 The Solution Approach**

In the context of analytical approaches, the solution for this problem is based on approximations. Approximations are required because the structure of the layered problem does not allow an exact solution for the general case.

The earlier approximations looked for an equivalent flow model and solved the problem for passive resources in the first level [46]. Later, decomposition approximations have been successfully utilized in solving the simultaneous resource problem, for instance in [26] and [15]. Similarly, Dorsman, Boxma and Vlasiou [16] proposed an approach based on server vacations for studying EMIP in the case of two layers of servers. This method follows a procedure similar to the decomposition approximations in terms of segmenting the problem into complementary subsystems. Later, the subsystems are solved recurrently based on the results obtained in the complementary subnetworks and continued until convergence.

In computer science applications there is significant work based on Mean Value Analysis (MVA) tools, for instance in [40], [19]. These applications have been exclusively designed to model computer science scenarios and consequently, they are highly specialized for these environments and conditions.

Another approach for passive resources modeling is the semi-open queueing networks (SOQN) or open queueing networks with capacity restrictions (OQN-CR). These developments are characterized by the capacity restriction effect on the production system that is modeled by passive resources. The solutions to the SOQN have been dominated by MVA like solutions and the matrix-geometric method (MGM) as explained in [27] and [41].

The method developed to approach this problem is somewhat different from the methods explained so far. The health care services are specially susceptible to the variability effects. Instead of unsatisfied customers as a result of production lines delays, customers in the tails of service time distribution in health care services could face irreversible health consequences. Therefore, a solution approach that incorporates the variability as an inherent part of the method could bring additional benefits to the analysis, in comparison with other methods

based on mean values or exponential distributions.

The two-moment framework proposed by Whitt in [54] and improved in [55] is based on queueing approximations involving the mean and squared coefficient of variation (SCV) as parameters. It is utilized to represent systems with non-Markovian arrivals or service times. An approach based on the two-moment framework, known as the parametric-decomposition method, was used by Kamath and Sanders [28] to model MIP and by Krishnamurthy and Suri [32] to model kanban production systems. The Krishnamurthy and Suri [32] fork/join approach to model passive resources (kanban), and other tools from the parametric-decomposition method formed the foundational building blocks on which we developed the solution approaches for the situations addressed by this study.

### **1.3 Outline of the Dissertation**

In the next chapter, a summary of the relevant literature is presented. In Chapter 3, the research statement is introduced, including the problem description, the main technical challenges, and the dissertation's objectives.

In chapters 4, 5, and 6 the building blocks on which the different solutions are constructed are presented. These three chapters present the development, solution, and testing/evaluation of the building blocks to model SRP and capacity restriction. Chapter 4 deals with the basic SRP building block, where a second-level resource receives service requests from multiple servers at a single node or multiple nodes in the first-level network.

Chapter 5 presents a standard closed queueing network (CQN) model with single and multi-server nodes representing situations where most of the passive resources are always

occupied or in use. Chapter 6 presents the fork/join (FJ) approach, and expands it to be used with multi server nodes. The synchronization nodes at the beginning and end of each stage allow us to pair the arriving customers with the capacity restricted resources in order to grant the customers access to each stage when capacity is available. We also show how the fork-join structure can be modified to model a single-stage capacity-restricted network model of a service system with a customer arrival process to a synchronization node and no synchronization node at the end. The fork/join structure allows us to model a capacity-restricted situation where the passive resources are not always occupied. In Chapter 7, the fork/join structure was extended to model a two-stage system. A synchronization station connects two capacity-restricted networks to model customer movement from one capacity-restricted subsystem to another.

In chapter 8, 9, and 10 these building blocks are combined to solve different configurations involving capacity-restrictions and SRP. In chapters 8 and 9 the CQN and FJ systems are combined with the SRP subsystem, respectively. Chapter 10 models a two-stage system, where nodes in each stage share a single SRP node. Finally, the conclusions and future work ideas are presented in Chapter 11.

## CHAPTER II

### BACKGROUND AND LITERATURE REVIEW

Many complex systems can be represented as queueing networks. The flow through these networks faces a series of delays caused by different mechanisms embedded in them. Messages transmitted in telecommunication networks, passengers moving through airports, part flow on the shopfloor, and patients receiving care in hospitals are some examples. For many years, queueing theory has been used to represent and evaluate the performance of these systems. One of the earliest and probably the most famous works in queueing networks was published by Jackson [25], where he established the basic cornerstones for steady-state analysis of networks of queues.

In spite of this early start and the prolific amount of developments in the field, many problems still remain difficult to solve, or not solvable under Jackson's product-form solution. A class of problems that have remained difficult to solve includes cases where overlapped possession of resources is present, such as SRP, MIP, and EMIP.

The literature related to this class of problems is vast and complex. In the next few sections, a summary of the literature review conducted is presented.

## 2.1 The Machine Interference and the Extended Machine Interference Problems

Originally called the machine-repairman problem or machine-operator interference problem, the machine interference problem is a well-studied problem, as can be seen in [48] and [23]. These two surveys, each one with hundreds of papers cited, explain the nature of this problem and its variants. In the MIP, the main concern is the interference generated in the overlapped resource, when different non-overlapped resources arrive for service, for instance [28] and [51]. Consequently, over the years, researchers have been focused on analyzing the performance metrics for the overlapped resources or resources in the second level.

The MIP evolved to include the performance measurements for the non-overlapped servers. Based on this new feature, the problem was called EMIP. Sauer, first in [46] and later in [45], proposed one of the first approaches to find the performance measurement for the EMIP. Sauer relied on Norton's Theorem applied to queueing networks. The work by Chandy, Hezog and Woo [10] set the basis for Sauer's decomposition/approximation approach.

In the MIP case, the only queue studied is the second level queue, and consequently, traditional queueing solutions can be found in the literature, including product-form solutions. On the other hand, when the queues in both levels of the service are studied, the problem becomes much more challenging to analyze. In [23], many solutions from the traditional queueing literature can be found for the MIP problem, for instance, M/M/1, M/M/r, M/G/r or G/G/r approaches. In contrast, in EMIP the main methods for solution have been more complex approximations, similar to the solutions mentioned in previous paragraphs from Sauer's publications [46], [45].

Kamath and Sanders [28] evaluated the EMIP in a network of queues for non-exponential service times in the non-overlapped and overlapped servers. They also offered performance measurements for the case of non-exponential inter-arrival times of the non-overlapped servers. In order to achieve their goal, these authors developed an algorithm based on the parametric-decomposition method proposed by Whitt [54].

Dorsman, Boxma and Vlasiou [16] analyzed the correlation effect produced in the non-overlapped servers' queues, by sharing the overlapped server queue in the EMIP. These authors explore the effect in two cases where two independent machines, working in parallel or tandem, share one repairman. Dorsman et al [16] allowed the non-overlapped machines to have different uptime and repair-time distributions. These authors followed an approach that connects the model for a single server queue with the EMIP, by introducing the effect of successive vacation periods to model the visits to the overlapped resource. This approach proved to be successful in obtaining the complete marginal queue length distribution for the queues in the non-overlapped resources.

The SRP and the EMIP have been utilized as synonyms and in many cases as exchangeable concepts [46], [17]. In spite of this, the next section has been separated from the current one, and it focuses on problems the authors have named SRP.

## **2.2 The Simultaneous Resource Possession (SRP) Problem**

Sauer's decomposition/approximation approach discussed in the EMIP section solves a network of queues with multiple classes of customers. At some point in the network route, customers require one resource, from a set of homogeneous passive resources, in order to have access to an overlapped resource, with the customer holding both resources simultaneously [45]. The method relies on the solution of two sub-networks that exchange information based

on an “equivalent flow computation.”

Jacobson and Lazowska [26] extended Sauer’s work by allowing the non-overlapped servers to be non-homogeneous, and they proposed the “Method of Surrogates.” This method splits the problem into two networks and iterates between them until a convergence solution is found. This approach can be applied to passive and active non-overlapped resources. This inclusion of active resources makes an important difference in the applicability of the method.

The network decomposition approach that characterizes the research explained so far, was utilized by De Souza e Silva and Muntz [15] to solve networks of queues with SRP. This approach incorporates customers from different network chains. These customers are allowed to hold non-overlapped resources that belong to different sets of passive resources and could hold more than one passive resource simultaneously. Later, De Almeida and Keller [13] expanded this research including a numerical solution for which the fictitious server included in this approach could have non-exponential service times.

### **2.3 Other Approaches for the SRP Problem**

Other approaches can be found in the literature for solving SRP. The main two alternative approaches found in the literature to the decomposition methods are the Mean Value Analysis (MVA) and Petri nets/Markov chain approach. The Petri nets/Markov chain approach could be used to derive exact solutions usually under exponential assumptions. However, the number of states to be analyzed in order to solve a non-trivial network of queues with SRP makes these methods intractable [20], [34], [36], [18].

MVA is the main approach utilized for solving layered queueing networks (LQN). LQN is



a powerful tool for evaluating the systems performance in computer science and telecommunications. The main characteristic of the systems solved by the LQN models is that the service requires at least two nested resources. The technology advances and trends in decentralized services have impacted the complexity of these systems, which have become more complex with a deep series of layers containing many client/server relationships.

In order to solve LQN models, the researchers have developed solvers such as the method of layers (MOL) [40] and simultaneous rendezvous networks (SRVN) [57] [58]. In order to solve these complex models, LQN relies on the integration of the previous tools with others, such as simulation tools. Consequently, this methodology has become a hybrid tool set for solving deep instances of SRP. For a more general explanation about the tools utilized in solving LQN, see [20] and [19]. The MVA was utilized, by Suri and Desiraju in [49] to model a network of queues with SRP and state-dependent routing.

The paper by Kaufman and Wong [30] explains the case of a capacity restricted network of queues and some of the servers in the network chain receiving additional external arrivals. The authors explained how their analysis does not pursue the approximation for all mean performance values for the SRP problem, but it could be a starting point for research related to this configuration.

## **2.4 Semi-Open Queueing Networks and Open Queueing Networks with Population Constraints**

Open queueing networks with population constraints (OQN-PC), open queueing networks with restricted capacity (OQN-RC), and semi-open queueing networks (SOQN) are some of the names that are given to networks of queues that have restrictions on the populations

allowed to be in the system at one time. This capacity restriction on the system can be modeled as a requirement for a passive resource. The customer needs to seize this passive resource in order to enter the system, and hold it until she leaves the system. At the end, the passive resource returns to the pool of free passive resources and from there to the first station completing a closed loop.

Dillard [12] and Ballard and Dillard [7] proposed methodologies utilizing aggregation techniques and Marie's [35] method to solve different variations of this problem. Later, Buithenhek et al. [8] utilized MVA methods to approach the same problem. Jia and Heragu [27] named this problem SOQN and solved it using the matrix-geometric method (MGM). In Roy [?] a recent survey in SOQN can be found.

Jia and Heragu [27] made a key assumption regarding the capacity constraint, which could limit the applicability of their approach. They assume that "the number of pallets,  $N$ , is sufficient large so that the external job queue will never explode." [27, p. 393]. They also use the MGM method which "requires the service time and inter-arrival time distribution to be phase-type, which may not hold in practice" [41, p. 1748].

In the framework of this study, the approximations developed by Krishnamurthy, Suri and Vernon [33] and Krishnamurthy and Suri [32], and later extended by Ramakrishnan and Krishnamurthy [39] played a key role. In these studies, the authors utilized the fork/join technique based on the parametric-decomposition method to model the capacity restriction in semi-open queueing networks. In [33], the authors developed two-moment approximations for the fork/join synchronization station. In [32], the authors model a series of production stations where kanban cards control the number of parts in the system. In [39], the authors extend the approximation to fork/join nodes with more than two input queues.

## 2.5 Literature Gap

In recent years, the MVA approach has become the basis for the analysis of layered networks, mainly focusing on computer science applications. In spite of these applications' high success, their analysis is based on the knowledge and representation of only mean values, and many fields of study have different necessities, mainly related to how the model represents and explains the variability present in system. Health care is one of these fields where the variability plays an important role. The service time variability could dramatically affect the experience of patients whose health conditions could be particularly sensitive to extreme sojourn time values. Consequently, how the model represents the service and inter-arrival time random variables becomes an important part of the analysis.

The parametric decomposition method introduced by Whitt [54] is well-suited for directly including the component's variability in the performance analysis of the system. However, there is not enough research available that allows us to model complex systems that include, simultaneously, capacity restrictions that control the movement between segments in the customer pathway and nodes that also require service from a second-level server giving rise to SRP instances.

The fork/join approach developed by Krishnamurthy and Suri [32] to model kanban systems could be adapted to develop a building block to model the capacity restriction in service systems. However, the only method found that utilized the two-moment approach to solve SRP like problems was the EMIP solution proposed by Kamath and Sanders [28] in the context of automatic assembly systems. Their approach could serve as a starting point for developing a building block to model SRP instances.

In summary, there is need to extend the parametric decomposition methodology to solve queueing network models with SRP instances in the presence of capacity restrictions.

## CHAPTER III

### STATEMENT OF RESEARCH

#### 3.1 Problem Description

As discussed in the previous chapters, this research study was motivated by challenges faced while analyzing queueing network models of hospital systems that typically include instances of SRP and capacity restrictions in many parts. The models' main characteristics are the following:

1. The patient pathway is formed by one or more sub-networks. Each sub-network could have its own capacity restriction. These capacity restrictions could be equal to or different from the capacity restrictions in other parts of the system.
2. Inside of these sub-networks, there could be SRP situations. That is, once the patient gets her access to the sub-network, one or more servers require at least one additional resource to provide the service. Consequently, the patient holds more than one active resource, simultaneously. This situation generates a contention effect affecting the overall service time of the non-overlapped resource. The resource first held by the customer or resource in the first level is called the non-overlapped server.
3. The solution proposed should be able to include the service and arrival variability as a structural part of the model, so that the performance results can easily include the effect of these variabilities on system performance.

In order to gain a better understanding of these systems and the problem under consider-

ation, some example system configurations are presented next.

In Figure 3.1, the customer needs the passive resource “A” in order to enter the system. Once in the system, the customer advances to join a queue and waits for service from one of the servers numbered 1. Once she reaches this server, and uses it for a given amount of time, the customer may wait for the shared support sub-system to receive additional service before finishing the service in server 1. Then the customer is able to advance to the next station. There exists a second network of queues that has a regulated capacity because of resource “B”. After visiting this second network, the customer leaves the system.

It is common in health care for patients to have different pathways through the hospital and to move outside the pathway at any point of the process. Consequently, a patient’s exit was added after the first network in order to highlight this feature. A probabilistic routing to the exit could be modeled at the end of the first sub-network.

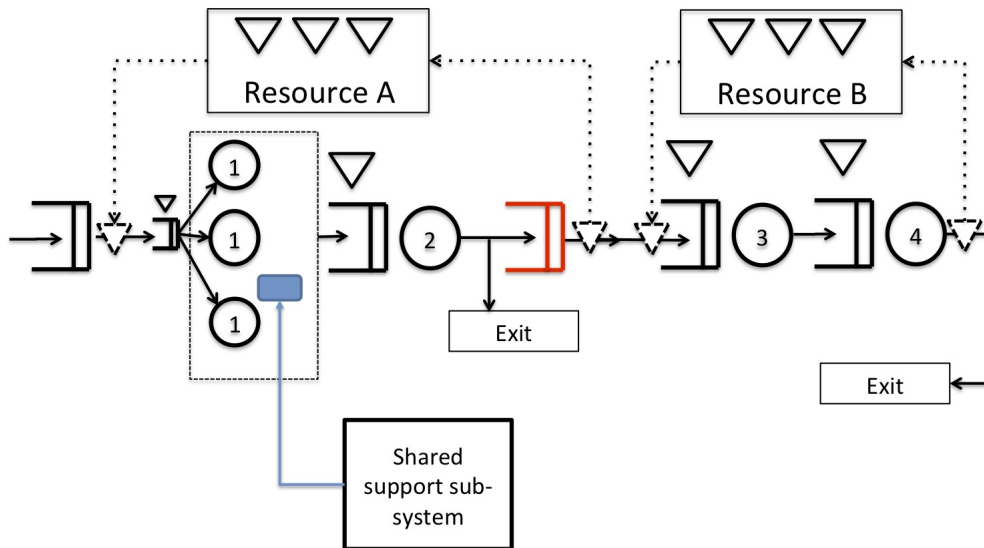


Figure 3.1: Example System Configuration 1

A possible variation of the same problem is a system where the shared support subsystem receives customers holding first level servers belonging to different networks of queues. Although both non-overlapped resources (1 and 3) belong to different subnetworks, both of them are in the same customer pathway as presented in Figure 3.2.

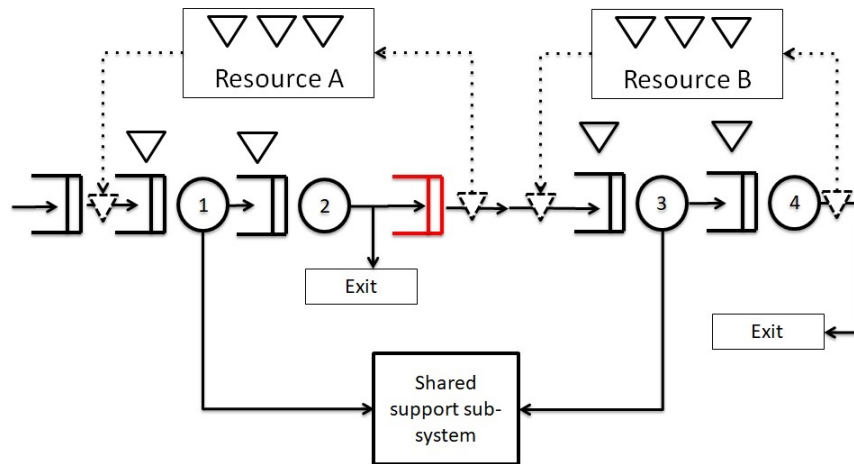


Figure 3.2: Example System Configuration 2

Based on the above discussion, the problem could be stated as:

“How to extend parametric-decomposition techniques to model queueing networks with simultaneous resource possession under capacity restriction?”

### 3.2 Technical Challenges

Despite the fact that part of our solution methodology utilize the previous work of Krishnamurthy and Suri [32], Kamath and Sanders [28] and Satyam, Krishnamurthy and Kamath [42], significant technical difficulties had to be addressed in order to achieve the goals of this study.

1. To be consistent with the underlying principles of the parametric-decomposition methodology, the SRP extension had to be developed in a modular or building block manner

with input parameters clearly identified and the solution method self-contained.

2. The SRP model and its solution methodology should be independent of the type of the first-level network (closed or semi-open) and the number of stages in the first-level network that generate requests for the overlapped server.
3. Proper algorithmic structure had to be devised while embedding the SRP model into the fork-join solution proposed by Krishnamurthy and Suri [32] or other first level solutions, depending on the system characteristics, in order to model capacity restricted processes. Appropriate tolerance levels had to be decided as the fork-join approximations involve open-ended iterations.
4. Limited research has been done to develop correction factors to obtain accurate approximations for the waiting time in queue for multi-server nodes in CQN systems. The only correction identified was based on the MVA approach and did not necessarily meet our research requirements. This imposed an additional technical challenge given this metric is not only necessary by itself, but also essential for mutiserver-nodes that are part of the first and second level networks.

### **3.3 Objectives and Tasks**

Two main objectives were completed in this dissertation, and they are as follows:

1. Developed a methodology to model SRP situations in a capacity constrained network of queues using a building-block approach, so that the models can be easily adapted to various real-life scenarios.
2. Validated the models and the methodology by modeling different system configurations and evaluating the performance prediction accuracy by comparing the analytical results with simulation estimates.



In achieving these objectives, the following tasks were completed:

1. Tasks related to Objective 1.

1.1. Identified the various building blocks that would be needed to model the typical system configurations considered in this study. The building blocks identified are as follows.

1.1.1. A second-level resource that receives requests for service from multiple servers at a single node or stage in the first-level network.

1.1.2. A second-level resource that receives requests for service from multiple servers at multiple nodes or stages in the first-level network.

1.1.3. A closed queueing network model for a capacity-restricted situation where most of the passive resources are almost always occupied or in use.

1.1.4. A capacity-restricted network model for a situation where the passive resources are not always occupied.

1.1.5. Linkages (fork-join stations) to connect two capacity restricted networks to model situations where a customer moves from one to the other.

1.2. Identified and developed proper corrections to improve model accuracy where necessary. An example is corrections for waiting time at multi-server nodes in a capacity-restricted network.

1.3. Developed new or modified existing solution algorithms to solve models of the building blocks identified in Task 1.1. All of the algorithms followed the parametric decomposition methodology and used two-moment approximations.

1.4. Combined the building blocks to form representative system configurations and developed complete solution algorithms for each to illustrate how the building blocks could be connected.

- 1.4.1. A closed queueing network model at the first level in which a server at a multi-server node may require the services of a second-level server with a specified probability.
- 1.4.2. A single capacity-restricted network with a patient arrival process to a synchronization node to model a situation where the passive resources are not always occupied. A server at a multi-server node in this network may require the services of a second-level server with a specified probability.
- 1.4.3. Two capacity restricted networks in tandem with a fork-join station linking the two to model the capacity restriction at each of the individual networks. Servers at a multi-server node in each of the two networks may request service from the same shared second-level resource.

## 2. Tasks related to Objective 2.

- 2.1. Developed Arena (Simulation software by Rockwell Automation, Inc. Version 16.00.00) simulation models to simulate each of the building blocks and the representative system configurations and conducted warm-up analysis using Welch's [52, 53] method to determine the simulation parameters such as the warm-up period, run length, and the number of replications to obtain statistically accurate estimates.
- 2.2. Designed and executed numerical experiments to evaluate the performance prediction accuracy of the building block models as well as the models of the representative configurations. The experiments tested the accuracy of the analytical models at different utilization levels, capacity restrictions, and different levels service time variability.

These objectives and tasks were implemented using a two-phase strategy. In the first phase, the solution algorithms for the building blocks were selected and adjusted to fit the

anticipated needs for solving hospital system models. These building blocks included the SRP system and its interactions with the first-level network, CQN with multi-server nodes, and the fork/join solution. The fork/join building block had three variants. The first is the single-stage system with two fork/joins, which was extended to include multi-server nodes. Then, the single-stage system with a single fork/join station was solved. This modification was required to capture the “push” nature of the service system in health care. The original approach was designed for “pull” production systems controlled by Kanban cards. Finally, the two-stage fork/join system to model two connected capacity-restricted stages.

In the second phase, the building blocks were combined to connect the SRP subsystem with a first-level network. The focus initially was on single-stage systems at the first level. For these combinations, the SRP subsystem solution was embedded in the solution of the first level network - a CQN, a double fork/join system, or a single fork/join system. Finally, the case where a SRP subsystem is shared by different stages of a multi-stage network in the first level is considered. In this final case, there is an interaction not only between the two stages but also with the SRP system, where the throughput, and consequently performance metrics of each stage, are intimately related with the other stages performance metrics. The knowledge gained in solving the building blocks played a key role in devising an appropriate approach to find a solution to the final, more complex system.

## CHAPTER IV

### THE SRP BUILDING BLOCK

The SRP building block is a key component in the systems studied in this document. For instance, in chapters VIII, IX, and X the SRP system is included as a second-level system for different capacity-restricted systems. Hence, this chapter is devoted to the solution of the SRP subsystem and its connection with the first-level system.

In Figure 4.1 the SRP subsystem and its connection with the first-level system is presented.

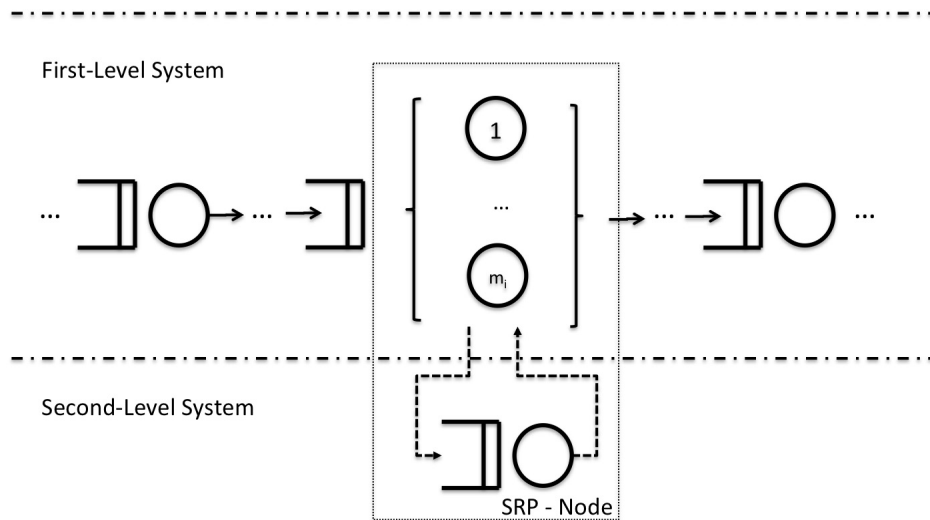


Figure 4.1: A First Level Network Connected to a SRP Subsystem

The SRP solution approach for the single and multi server SRP nodes is presented in the next sections. This solution is based on recognizing the capacity-restricted conditions in the SRP subsystem and computing this node's performance metrics accordingly. The relation

between the first and second level systems is fundamental in the overall system solution. Consequently, the method to connect both systems is a critical part of the overall system solution, and it is discussed in next sections as well, including cases where servers at multiple nodes in the first-level system request service from the SRP node.

#### 4.1 The SRP Solution Approach

The demand for additional service from a second-level server is best modeled in a probabilistic manner. Only a proportion of the customers in the first-level network require service from the second level sub-system. For instance, resources such as specialist or specialized equipment in hospital environments are not required for all the patients in the system. Actually, for some of these cases only a small percentage of customers require the additional server that generates the SRP condition. Therefore, the problem solution needs to account for variable proportions of the customers requiring the SRP server. As discussed previously, the first level system for the capacity-restricted versions of the problem can be solved using a CQN or a fork/join structure depending on the system characteristics.

The overall service time for the server where the SRP interference occurs needs to be computed before solving the first-level system, and it depends directly on the SRP interference time. The interference time represents the first obstacle for finding the system's performance metrics. Therefore, the main idea behind the proposed solution is to compute the SRP node's response time (waiting time in queue plus service time at the SRP node) and combine it to the service time in the first-level node. Figure 4.2 shows a graphical representation of the idea.

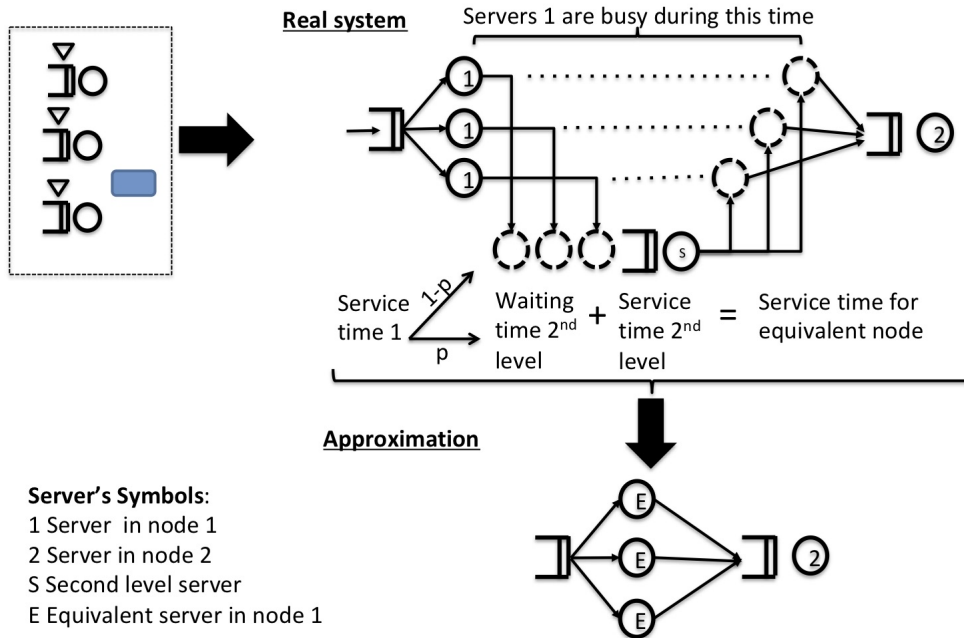


Figure 4.2: Approximation used in the node affected by the SRP system

The algorithm aims to solve the first-level system once the effect of the SRP server has been incorporated into the first level server's service time that is affected by it. In order to obtain this, the sub-system composed of the SRP server in the second level and the server affected by it in the first level, will be substituted by an "Equivalent Node" in the first level.

The equivalent node's service time parameters should include the mean and SCV of the original service time in the first level and probabilistically, the mean and SCV of the response time from the SRP server in the second level. The SRP node's mean and SCV of response time are estimated in the second level system. For computing these, the arrival rate and SCV of inter-arrival times of requests from the first-level network to the SRP system are needed. The proposed algorithm iterates between these two systems until the SCV of the arrival process to the SRP node ( $c_{a,SRP}^2$ ) converges as determined by a specified tolerance level.

## 4.2 Notation and System Parameters

Let the random variable (rv)  $S_r$  representing the original service time of the first-level server affected by the SRP instance.  $E[S_r] = \tau_r$  and  $Var[S_r] = \sigma_r^2$  are known service time parameters.

Let  $S_r^I$  be the rv represent the modified service time of server “ $r$ ” in the first level, where the SRP effect is included.  $Z$  is an indicator rv, with  $Z = 1$  if a customer needs service from the second level and  $Z = 0$  otherwise.

$$S_r^I = S_r + ZR \quad (4.1)$$

where  $R$  is the rv representing the SRP response time.  $E[Z]=p$  and  $p$  represents the probability that a customer requires service from the SRP system. In other words, a customer after receiving regular service at node  $r$  in the first-level will need service from the second-level server with probability  $p$ .

The second-level system or SRP system’s main objective is to estimate the mean and variance of the SRP node’s response time. This system has some specific characteristics that are explained below.

- The maximum number of customers is limited and equal to the number of servers at node  $r$  in the first level.
- The throughput in the second-level system must be equal to the throughput in the first-level system, times the probability  $p$ .

Given these characteristics, the second-level system emulates a CQN where the number of customers is equal to the number of servers in the first-level node  $r$ , namely,  $m_r$ . A customer in the second-level system represents a customer in the first-level system in service at a server

at node  $r$ . There can only be a maximum  $m_r$  of such customers. The idea is to compute the waiting time in queue in the SRP node by taking advantage of the fact that the arrival parameters can be estimated from the first level. The SRP node belongs to a CQN with  $m_r$  customers and whose throughput is known. (See Figure 4.3)

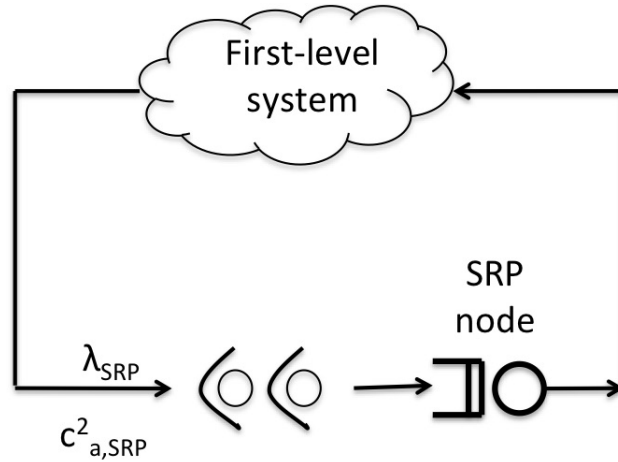


Figure 4.3: Representation of the second-level system

The parameters specifically used for the SRP node computations are presented in Table 4.1

Table 4.1: Second Level System Notation

$\lambda_{SRP}$	Arrival rate to the SRP node.
$c^2_{a,SRP}$	SCV of the inter-arrival time at the SRP node.
$\tau_{SRP}$	Service time of the SRP server.
$c^2_{s,SRP}$	SCV of the service time at the SRP server.
$m_{s,SRP}$	Number of servers in the SRP node.
$n_{2L}$	Number of customers in the second-level system or SRP System.**

\*\* If there is only a single node in the first-level that is affected by the SRP node, then  $n_{2L} = m_r$ , where  $m_r$  is the number of servers in node  $r$  (Node in the first level affected by



the SRP node).

The main result expected from this system is  $W_{q,SRP}$  and  $\sigma_{q,SRP}^2$ , the mean waiting time in queue and variance of the waiting time in queue, respectively for the SRP node.

The idea for the arrival parameters computation comes from the splitting of flows approach introduced by Whitt [54]. Only a proportion  $p$  of customers flowing through node  $r$  in the first-level require service from the SRP node. As an approximation, we treat the arrival process to the SRP node as a splitting of the departure process from node  $r$ . As explained previously, if  $p$  is the expected proportion of customers going into the SRP system, the arrival rate to the second level,  $\lambda_{SRP}$  is equal to

$$\lambda_{SRP} = \lambda_r * p \tag{4.2}$$

where  $\lambda_r$  is the departure rate of node  $r$  in the first-level.

Similarly, following the splitting approach, the SCV for the arrival process to the SRP system is obtained as follows

$$c_{a,SRP}^2 = p * c_{d,r}^2 + 1 - p \tag{4.3}$$

Where  $c_{d,r}^2$  is the inter-departure time SCV at node  $r$ .

### **4.3 Mean and SCV of Equivalent Service Time for a Server at a First-Level Node Affected by SRP**

As described in Section 4.1, the overall solution idea is to compute a response time for the SRP node and incorporate it in the service time of the node in the first-level system that is

directly influenced by the SRP system.

#### 4.3.1 Expectation of the Modified Service Time at the Node Influenced by the SRP system

Following the idea presented in Equation 4.1, the modified service time at node  $r$  or node influenced by the SRP system, is  $S_r^I = S_r + ZR$ , where  $Z$  is an indicator rv or a Bernoulli with probability  $p$  and  $R$  is the rv representing the SRP node response time.

It is important to remember that the time a customer spends in the SRP node is composed of the time the customer spends waiting at the SRP node to be served plus the time in service at the SRP node. Hence, we have

$$E[R] = W_{q,SRP} + \tau_{SRP} \quad (4.4)$$

And the expected modified service time for the server influenced by the SRP in the first level will be,

$$E[S_r^I] = \tau_r^I = E[S_r] + E[Z] * E[R]$$

then,

$$E[S_r^I] = \tau_r + p * [W_{q,SRP} + \tau_{SRP}] \quad (4.5)$$

#### 4.3.2 SCV of the Modified Service Time at the Node Influenced by the SRP

Following the same logic, the variability estimation is based on Equation 4.1 and the fact that,

$$Var[R] = \sigma_{q,SRP}^2 + \sigma_{SRP}^2 \quad (4.6)$$

where  $\sigma_{q,SRP}^2$  is the variance of the waiting time at the SRP node. Then,

$$Var[S_r^I] = Var[S_r] + Var[ZR]$$

As  $Z$  and  $R$  are independent random variables, we have

$$Var[S_r^I] = \sigma_r^2 + \sigma_Z^2 * \sigma_R^2 + \sigma_Z^2 * [E[R]]^2 + [E[Z]]^2 * \sigma_R^2$$

$$\begin{aligned} Var[S_r^I] = \sigma_r^2 + p(1-p) * [\sigma_{q,SRP}^2 + \sigma_{SRP}^2] + p(1-p) * [W_{q,SRP} + \tau_{SRP}]^2 \\ + p^2 * [\sigma_{q,SRP}^2 + \sigma_{SRP}^2] \end{aligned}$$

and finally,

$$Var[S_r^I] = \sigma_r^2 + p * [\sigma_{q,SRP}^2 + \sigma_{SRP}^2] + p(1-p) * [W_{q,SRP} + \tau_{SRP}]^2 \quad (4.7)$$

#### 4.4 Calculating Node-Level Performance Measures

In this section, we present several node-level calculations for both single-server and multi-server nodes. The general approach is to first calculate the performance measures for a node in an open network context and then apply correction factors to account for capacity restrictions. We present the corrections in subsequent sections. These calculations are used for nodes in both the first-level and second-level systems.

The utilization for any node is give by Equation 9 in [54], and that adapted to our notation is

$$\rho = \frac{\lambda * \tau}{m} \quad (4.8)$$

To compute the inter-departure time SCV from a node, the following equations taken from Whitt [54] were used. Equations 38 and 39 are for single and multi-server nodes, respectively.

$$c_d^2 = \rho^2 * c_s^2 + (1 - \rho^2)c_a^2 \quad (4.9)$$

$$c_d^2 = 1 + (1 - \rho^2)(c_a^2 - 1) + \frac{\rho^2}{\sqrt{m}}(c_s^2 - 1) \quad (4.10)$$

Mean waiting time in a queue is a key performance measure. For single-server nodes, the approach suggested by Kramer and Lagenbach-Belz [31] was implemented, as shown in Equation 4.11.

$$\bar{W}q^{GI/G/1} \approx \left( \frac{\lambda_a \tau^2}{1 - \lambda_a \tau^2} \right) \left( \frac{c_s^2 + c_a^2}{2} \right) * G \quad (4.11)$$

if  $c_a^2 < 1$ , then  $G$  is

$$\exp\left(\frac{-2(1-\rho)(1-c_a^2)^2}{3\rho(c_s^2+c_a^2)}\right)$$

On the other hand, if  $c_a^2 \geq 1$ ,  $G$  is

$$\exp\left(\frac{-(1-\rho)(c_a^2-1)}{(c_s^2+4*c_a^2)}\right)$$

In order to estimate the waiting time in queue for multi-server nodes, Equation (2.24) in [56] was originally utilized. This equation adjusted to the notation in this document is as follows,

$$Wq(\rho, c_a^2, c_s^2, m) = \phi(\rho, c_a^2, c_s^2, m) \left( \frac{c_a^2 + c_s^2}{2} \right) Wq(M/M/m) \quad (4.12)$$

where  $Wq_S(M/M/m)$  refers to the waiting time for the case of Poisson arrivals and exponential service times. Equation 4.13 was used to obtain this value,

$$Wq = \frac{1}{\lambda} \left( \frac{(\lambda\tau)^m \lambda\mu}{(m-1)!(m\mu - \lambda)^2} \right) P_0 \quad (4.13)$$

Where  $\mu$  and  $\lambda$  are the service and arrival rates, and  $P_0$  is the probability that there are no customers at the  $M/M/m$  node, given by formula 4.14

$$P_0 = \left( \left( \sum_{n=0}^{m-1} \frac{1}{n!} (\lambda\tau)^n \right) + \frac{1}{m!} (\lambda\tau)^m \left( \frac{m\mu}{m\mu - \lambda} \right) \right)^{-1} \quad (4.14)$$

and  $\phi(\rho, c_a^2, c_s^2, m)$  comes from Equation (2.25) in [56]. There are two cases, the first one when  $c_a^2 \geq c_s^2$ ,

$$\phi(\rho, c_a^2, c_s^2, m) = \left( \frac{4(c_a^2 - c_s^2)}{4c_a^2 - 3c_s^2} \right) \phi_1(m\rho) + \left( \frac{c_s^2}{4c_a^2 - 3c_s^2} \right) \Psi(((c_a^2 - c_s^2)/2), m, \rho) \quad (4.15)$$

When  $c_a^2 \leq c_s^2$  Equation 4.16 is recommended,

$$\phi(\rho, c_a^2, c_s^2, m) = \left( \frac{c_s^2 - c_a^2}{2c_a^2 + 2c_s^2} \right) \phi_3(m, \rho) + \left( \frac{c_s^2 + 3c_a^2}{2c_a^2 + 2c_s^2} \right) \Psi(((c_a^2 - c_s^2)/2), m, \rho) \quad (4.16)$$

It is important to note that author in [56] recommend the use of  $\Psi((c^2, m, \rho) = 1$  when  $c^2 \geq 1$ . On the other hand, when  $0 \leq c^2 \leq 1$  the next expression is suggested

$$\Psi((c^2, m, \rho) = \phi_4(m, \rho)^{2(1-c^2)} \quad (4.17)$$

where

$$\phi_4(m, \rho) = \min\{1, ((\phi_1(m, \rho) + \phi_3(m, \rho))/2)\} \quad (4.18)$$

$$\phi_1(m, \rho) = 1 + \gamma(m, \rho) \quad (4.19)$$

$$\gamma(m, \rho) = \min \left[ 0.24, (1 - \rho)(m - 1) \left( \frac{(4 + 5m)^{1/2} - 2}{16m\rho} \right) \right] \quad (4.20)$$

$$\phi_3(m, \rho) = \phi_2(m, \rho)e^{-2(1-\rho)/3\rho} \quad (4.21)$$

and finally,

$$\phi_2(m, \rho) = 1 - 4\gamma(m, \rho) \quad (4.22)$$

It is important to note the average waiting time equations presented in this section apply to non-capacity-restricted environments. The corrections required for the capacity-restricted conditions are based on these equations and will be presented later in this chapter.

In order to have a more complete set of performances measures for the systems under study, the average number of customers in queue was computed by applying Little's Law.

$$Lq = Wq\lambda \quad (4.23)$$

and average number of customers in the node

$$L = \lambda(Wq + \tau) \quad (4.24)$$

## 4.5 Corrections for the Average Waiting Time in Closed Networks

In this section, the average waiting time correction for closed networks is presented. The approaches presented in this section will be used in order to compute the average waiting time at the SRP node, and to compute the average waiting times for single and multi-server nodes in the capacity restricted first level networks. Two approaches are presented, for single and multi-servers nodes respectively.

### Single-server Case

In the parametric decomposition literature, the general approach to computing the mean waiting time in a single-server queue is to compute the waiting time in queue for a GI/G/1 queue and then adjust this for the presence of the server in a closed queueing network. The algorithms in [42] use the correction factor developed by Kamath, Suri and Sanders [29].  $K$  is the population of the network in which this node is present

$$\beta = \left( \frac{K-1}{K} \right) \left( \frac{\tau * K}{\tau * K + \bar{W}_q^{GI/G/1}} \right) \quad (4.25)$$

The average waiting time in node is computed as shown in equation 4.26

$$W_q = \bar{W}_q^{GI/G/1} * \beta \quad (4.26)$$

### Multi-server Case

Equation 4.12 is meant for the  $W_q$  computation of a multi-server node in a open queueing environment. The only correction factor for a closed system that was found in the literature was the one developed by Suri et al. [50] for the mean-value analysis (MVA) approach.

We decided to experiment with a factor that extends the single-server correction factor, heuristically, for the multi-server case and the correction is as follows,

$$\theta^A = \left(\frac{K-1}{K}\right) \left(\frac{1}{1+Wq/(\tau * K)}\right) \left(\frac{K-m}{K}\right) \quad (4.27)$$

The first two terms of the correction factor came from the correction factor developed by Kamath et al. [29]. The first term could be seen as addressing the population constraint  $K$  and is one that is inspired by the arrival theorem [5] which says that an arriving customer at a node sees a network with one less customer in equilibrium. The second term together with first makes the mean waiting time exact for a balanced CQN with exponential service times. The third term is motivated by the fact that the mean waiting time at a multi-server node is zero if the network population is less than or equal to the number of servers in the node. It should be noted that all these terms tend to 1 as  $K$  tends to  $\infty$ , making the correction factor tend to 1.

The second approach considered was the correction proposed by Suri et al. in [50]. This empirical correction was developed for the multi-server node case while solving the CQN with the MVA approach,

$$\theta^B = \frac{1}{m} \rho^{4.464(m^{0.676}-1)} \quad (4.28)$$

The last approach was to combine the above two factors with equal weightage. This was based on the observation that in some cases, when one factor caused an over estimation, the other caused the approach to underestimate the mean waiting time.

The different waiting time factors for multi-server nodes in a CQN were tested for a series of closed networks configurations in order to have a better understanding of how each



performed. Based on our numerical experiments the following scheme for the average waiting time correction was derived. We will call the correction obtained by using the following rules as  $\theta$

Table 4.2: Rules for multi-server waiting time correction ( $\theta$ )

Utilization level	Utilization value	$c_s^2$ value	Correction suggested
Low and Medium	$\rho < 90\%$	$c_s^2 < 2$	$\theta^A$
		$c_s^2 \geq 2$	$0.5*\theta^A + 0.5*\theta^B$
High	$90\% \leq \rho < 97\%$	$c_s^2 \leq 0.5$	$\theta^A$
		$c_s^2 > 0.5$	$\theta^B$
Very High	$\rho \geq 97\%$	All	$\theta^B$

Consequently, the mean waiting time at multi-server nodes in the system is computed by multiplying the value of  $Wq$  by  $\theta$ , as shown in Equation 4.29.

$$Wq = Wq(\rho, c_a^2, c_s^2, m) * \theta \quad (4.29)$$

#### 4.6 Variance of Waiting Time in Queue

The waiting time's variability in the SRP node is required for computing the first level system parameters as explained in Section 4.3.

The mean and variance of the waiting time in the SRP node are the main results required from the second-level system used to study the SRP effect. The mean waiting time will be computed using the CQN approach presented in previous sections. Next, the steps to compute the variance of the waiting time are presented.

Whitt presents the approaches for estimating the variance of the waiting time in queue for the single-server case in [54] and the multi-server case in [56]. We first use his approaches for estimating the SCV,  $c_w^2$  of waiting time in queue and then use our corrected average waiting times to calculate the variance of the waiting times in both the single and multi-server cases. Our assumption here is that the correction for the capacity restriction will result in appropriate correction for the variance of waiting time in queue.

### Single-server Case

The variance of the waiting time computed using Equation (54) in [54]. This equation adjusted to the notation in this document is

$$\sigma_q^2 = (Wq)^2 c_w^2 \quad (4.30)$$

where  $c_w^2$  is the SCV at the waiting time and is computed using

$$c_w^2 = \frac{c_D^2 + 1 - \hat{\sigma}}{\hat{\sigma}} \quad (4.31)$$

We use our own corrected  $Wq$  for the mean waiting time in Equation 4.31.

The value of  $\hat{\sigma}$  is computed as [54]

$$\hat{\sigma} = \rho + (c_a^2 - 1)\rho(1 - \rho)h(\rho, c_a^2, c_s^2) \quad (4.32)$$

According to Whitt [54],  $h(\rho, c_a^2, c_s^2)$  is computed differently depending on the value  $c_s^2$ . If  $c_s^2 \leq 1$  the the equation is as follows,

$$h(\rho, c_a^2, c_s^2) = \frac{1 + c_a^2 + \rho * c_s^2}{1 + \rho(c_s^2 - 1) + \rho^2(4c_a^2 + c_s^2)} \quad (4.33)$$

On the other hand, if  $c_s^2 \geq 1$  the equation to be used is

$$h(\rho_s, c_a^2, c_s^2) = \frac{4\rho}{c_a^2 + \rho^2(4c_a^2 + c_s^2)} \quad (4.34)$$

The value  $c_D^2$  is computed based on the equation (50) in [54] as

$$c_D^2 = 2\rho - 1 + \frac{4(1 - \rho)d_s^3}{3(c_s^2 - 1)^2} \quad (4.35)$$

The expression  $d_s^3$  is computed differently based on the  $c_s^2$  value. If  $c_s^2 \geq 1$  the equation (51) from [54] is used, otherwise, equation (52).

when  $c_s^2 \geq 1$ ,

$$d_s^3 = 3c_s^2(1 + c_s^2) \quad (4.36)$$

if  $c_s^2 < 1$ ,

$$d_s^3 = (2c_s^2 + 1)(c_s^2 + 1) \quad (4.37)$$

In summary, we used the procedure outlined in [56] to compute the SCV of the waiting time in queue,  $c_w^2$ , and then used our own corrected  $Wq$  to estimate the variance of the waiting time.

## Multi-Server Case

To obtain the SCV of the waiting time at nodes with multiple servers, the approach proposed by Whitt [56] has been followed. Whitt's approach is presented next using our notation

$$\sigma_q^2 = (Wq)^2 c_w^2 \quad (4.38)$$

and

$$c_w^2 = \frac{c_D^2 + 1 - P(w_q > 0)}{P(w_q > 0)} \quad (4.39)$$

In order to compute  $P(w_q > 0)$ , equations 3.9 to 3.11 in Whitt [56] were used

$$P(w_q(\rho, c_a^2, c_s^2, m) > 0) \approx \min\{1, \pi\} \quad (4.40)$$

And  $\pi$  corresponds to:

- $\pi_1$ , if  $m \leq 6$  or  $\gamma \leq 0.5$  or  $c_a^2 \geq 1$
- $\pi_2$ , if  $m \geq 7$  and  $\gamma \geq 1$  and  $c_a^2 < 1$
- $\pi_3$ , if  $m \geq 7$  and  $c_a^2 < 1$  and  $0.5 < \gamma < 1$

and

$$\pi_1 = \rho^2 \pi_4 + (1 - \rho^2) \pi_5 \quad (4.41)$$

$$\pi_2 = c_a^2 \pi_1 + (1 - c_a^2) \pi_6 \quad (4.42)$$

$$\pi_3 = 2(1 - c_a^2)(\gamma - 0.5) \pi_2 + (1 - [2(1 - c_a^2)(\gamma - 0.5)]) \pi_1 \quad (4.43)$$

$$\pi_4 = \min\left\{1, \frac{1 - \phi((1 + c_a^2)(1 - \rho)m^{1/2}/(c_a^2 + c_s^2))}{1 - \phi((1 - \rho)m^{1/2})} P(w_q(M/M/m) > 0)\right\} \quad (4.44)$$

$$\pi_5 = \min\left\{1, \frac{1 - \phi(2(1 - \rho)m^{1/2}/(1 + c_a^2))}{1 - \phi((1 - \rho)m^{1/2})} P(w_q(M/M/m) > 0)\right\} \quad (4.45)$$

$$\pi_6 = 1 - \phi((m - m\rho - 0.5)/\sqrt{m\rho z}) \quad (4.46)$$

$$\gamma \equiv \gamma(m, \rho, z) = (m - m\rho - 0.5)/(m\rho z)^{1/2} \quad (4.47)$$

$$z = (c_a^2 + c_s^2)/(1 + c_s^2) \quad (4.48)$$

and  $\phi$  is the normal continuous density function (CDF) as explained in Whitt [56], page 136.

#### 4.7 General Algorithm for Solving the SRP System

Following the general approach outlined in section 4.1, the algorithm will iterate between the solutions of the two levels until convergence is reached. This iterative process is shown in figure 4.4.

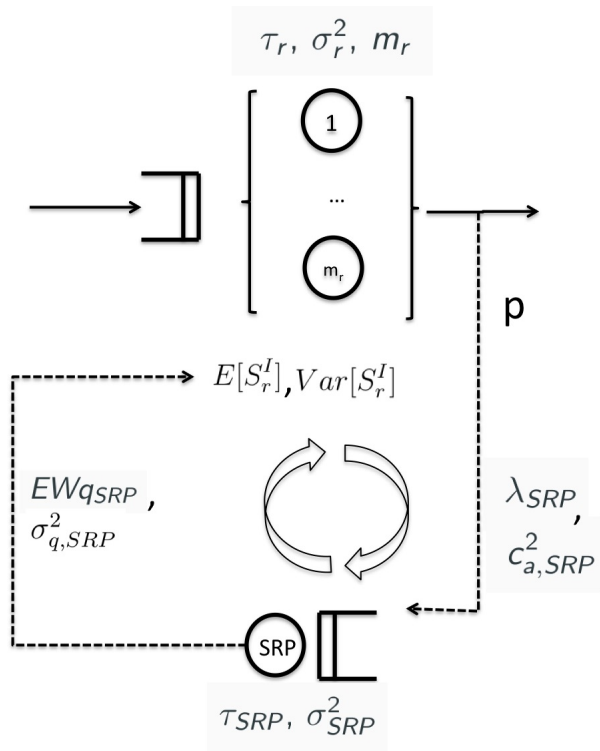


Figure 4.4: SRP System Solution Approach

Where  $S_r^I$  is the equivalent service time. The General Algorithm is presented next.

---

**Algorithm 1** SRP Solution Algorithm

---

**Require:**  $\lambda_{1L}, c_a^2, \tau, m, c_s^2, \tau_{SRP}, m_{SRP}, c_{s,SRP}^2$

**Ensure:**  $W_{q,SRP} = 0, \sigma_{q,SRP}^2 = 0, c_{a,SRP}^{2(0)} = 1000, \epsilon = 0.001, j = 1.$

**Step 1**

Using equations 4.5 and 4.7 compute the equivalent node mean service time and SCV.

**Step 2**

Solve the first level system including the equivalent using the appropriate approach.

**Step 3**

**while**  $\delta = |c_{a,SRP}^{2(j)} - c_{a,SRP}^{2(j-1)}| \geq \epsilon$  **do**

**Step 3.1**

Compute the population constraint,  $n_{2L}$  and SRP arrival parameters  $\lambda_{SRP}$  and  $c_{a,SRP}^2$  using equations 4.49, 4.50 and 4.51 respectively

**if**  $m_{SRP} > 1$  **then**

    Compute  $W_{q,SRP}$  and  $\sigma_{q,SRP}^2$  using equation 4.29 and 4.38 respectively

**else**

    Compute  $W_{q,SRP}$  and  $\sigma_{q,SRP}^2$  using equations 4.26 and 4.30 respectively

**end if**

**Step 3.2**

Set  $c_{a,SRP}^{2(j)} = c_{a,SRP}^2$

**Step 3.3**

Using equations 4.5 and 4.7 compute the equivalent node service time and SCV with the new  $W_{q,SRP}$  and  $\sigma_{q,SRP}^2$ .

**Step 3.4**

Solve the first level system with the new equivalent node parameter with the appropriate approach.

**Step 3.5**

Compute  $\delta = |c_{a,SRP}^{2(j)} - c_{a,SRP}^{2(j-1)}|$

**Step 3.6**

Set  $c_{a,SRP}^{2(j-1)} = c_{a,SRP}^{2(j)}$

**Step 3.7**

Set  $j = j + 1$

**end while**

---

In Step 4.1 of the algorithm,  $n_{2L}$  refers to the population constraint in the SRP node solution approached based on a closed queueing network.

#### 4.8 SRP Building Block with Multiple Sources

In this section, we show the extension of the SRP building block to include multiple sources. This configuration captures situations where requests arrive to the SRP node from multiple nodes in the first-level network and expands the range of real time applications that could be captured with this building block.

Each node that requires service from the SRP node will become a source. The structure and solution of the SRP system would remain the same, so will modifications needed to the service time of each of the affected servers in the first level. The basic idea is to merge the arrivals from multiple sources to derive the rate and SCV of the combined arrival process to the SRP system. In Figure 4.5 a representation of this situation can be found.



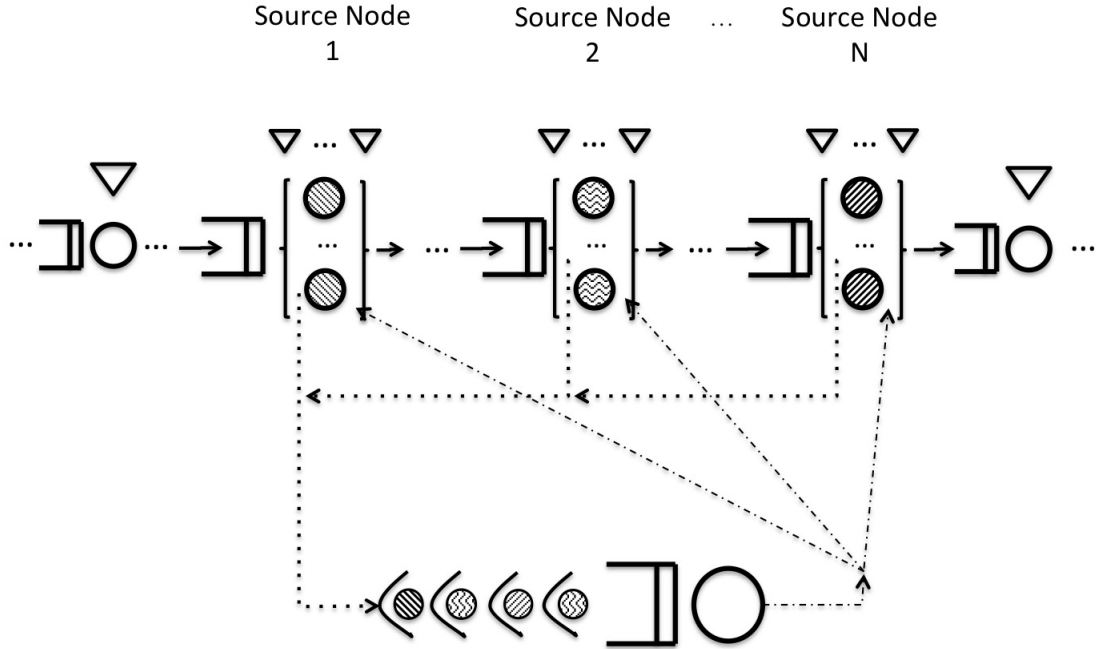


Figure 4.5: Multi-Input SRP

Following the solution approach for the single input SRP building block, the solution will depend on the estimation of the second-level expected waiting time and variance of the waiting time.

As previously explained, the second level or SRP node will be considered as part of a CQN. The population  $n_{2L}$  of this CQN can be obtained as follows

$$n_{2L} = \sum_{i=1}^N m_i \quad (4.49)$$

where  $m_i$  is the number of servers at source node  $i$  and  $N$  is the number of source nodes that use the SRP system.

Following the merging of flow results in Whitt [54], the arrival rate and SCV of combined

arrival process can be obtained using the following equations

$$\lambda_{SRP} = \sum_{i=1}^N \lambda_i * p_i \quad (4.50)$$

$$c_{a,SRP}^2 = \sum_{i=1}^N \frac{\lambda_i * p_i}{\lambda_{SRP}} \{p_i * c_{d,eq,i}^2 + 1 - p_i\} \quad (4.51)$$

Every source node in the first level will be replaced by an equivalent node, and its expected service time and SCV will be obtained as explained in section 4.3. The solution algorithm is the same as in section 4.7.

## CHAPTER V

### CLOSED QUEUEING NETWORK (CQN) WITH MULTI-SERVER NODES

CQN is a fundamental building block for solving capacity-restricted systems operating near or at full capacity. While solving CQN models has been an active area of research for several decades, a parametric decomposition based approach was published only a few years ago [42]. We extended the CQN solution approach presented in [42] to include multi-server nodes, a variation that was not included in the original paper. Modeling multi-server nodes is important in many hospital systems, where at many stages there could be multiple care providers available for patients.

The CQN as a tool is particularly useful when the passive resources controlling the system's capacity are working near saturation. When the passive resources along with the customers finish their path through the system, they are released. The passive resources immediately return to the beginning of the system where customers are already waiting for them. Under this condition, the passive resource is in use nearly all the time and rarely idle. The number of passive resources becomes the population constraint in the CQN model.

Another important CQN role for the systems under study is to identify the maximum throughput a capacity restricted system can deliver. This knowledge will allow the researcher or practitioner to anticipate if systems that are capacity restricted, but not at or near full capacity all the time, will be able to handle the external demand without experiencing stability issues.

## 5.1 Multi-server Extension

The original PDMCQN-2 algorithm was introduced in [42] as a parametric decomposition based approach to evaluate the performance of general multi-class, closed queueing networks. In the original work, the authors presented a general version for multi-class applications and a version for single class networks. However, the networks only included single-server nodes. In this section, the modifications developed for the algorithm to work with multi-servers nodes in the single class version are presented.

### 5.1.1 The Original PDMCQN-2 Algorithm

The original single-class version of the PDMCQN-2 algorithm is presented in this section. This is the version of the algorithm presented in the original publication [42] as the single-class closed queueing network (SCQN) algorithm.

This algorithm requires as inputs the number of nodes in the network  $N$ , number customers in the system  $K$ , the service time's mean and SCV at each node,  $\tau_S$  and  $c_S^2$ , respectively and the routing probabilities  $q_{v,s} = (s, v = 1, 2, \dots, N)$ . In the original algorithm, the main performance measure of interest evaluated is the mean waiting time at a node,  $\bar{W}q_S$ .

---

**Algorithm 2** PDMCQN-2 Single class version [42]

---

**Require:**  $N, K, \tau_S, c_{aS}^2, S = 1, \dots, N$  and  $q_{v,s}, v = 1, \dots, N$  and  $S = 1, \dots, N$

**Step 1**

Find relative utilization  $y_S$  of node  $S$ , relative to node 1 by solving traffic equations

$$\lambda_v = \sum_{S=1}^N \lambda_S * q_{vS} \text{ for } v=1, \dots, N \text{ with } \lambda_1 = 1/\tau_1, y_S = \lambda_S \tau_S, S = 2, \dots, N.$$

**Step 2**

Initialize: Low=0 and High=1

**Step 3**

Find the index of the node with the highest utilization,  $n = \arg \max y_S$

Begin loop

For  $iter=1$  to 19

**Step 4.1**

Calculate utilization,  $\rho_S$ , and throughput,  $\lambda_S$  at each node using following equations

$$\rho_n = (low + high)/2$$

$$\rho_S = \frac{y_S}{y_n} * \rho_n, S = 1, \dots, N$$

$$\lambda_S = \rho_S / \tau_S, S = 1, \dots, N$$

**Step 4.2**

Solve the traffic variability equations to calculate  $c_{aS}^2$

$$c_{aS}^2 = a_j + \sum_{S=1}^N c_{aS}^2 d_{vS} *$$

**Step 4.3**

Calculate the mean waiting time at node  $\bar{W}q_S$  using the equation 4.26

**Step 4.4**

Calculate  $\delta = \sum_{S=1}^N \lambda_S (\bar{W}q_S + \tau_S) - K$

if  $\delta < 0$ , then set  $Low = \rho_n$  else set  $High = \rho_n$

End Loop

---

\*  $a_S = 1 + \eta_S [-1 + \sum_{S=1}^N p_{vS} (1 - q_{vS} + \gamma_v q_{vS} \rho_v^2 x_v)]$ ,  $d_{vS} = \eta_S p_{vS} q_{vS} \gamma_v (1 - \rho_v^2)$ ,  $x_v = 1 + [\max(c_{aS}^2, 0) - 1]$ ,  $\eta_S = [1 + 4(1 - \rho_S^2)(\gamma_S - 1)]^{-1}$ ,  $\gamma_S = [\sum_{v=1}^N p_{vS}^2]^{-1}$  and  $p_{vS} = \lambda_{a_{vS}} / \lambda_{a_S}$ .

### 5.1.2 Modifications to the PDMCQN-2 Algorithm to Include Multi-Server Nodes

Each modification will be explained following the steps in the algorithm presented in Section 5.1.1.

In finding the relative utilization of the nodes, the traffic equations remain the same.

$$\lambda_v = \sum_{S=1}^N \lambda_v q_{vS}, \quad v = 1, 2, \dots, N \quad (5.1)$$

The initialization is  $\lambda_1 = m_1/\tau_1$ , and

$$y_S = \frac{\lambda_S \tau_S}{m_S}, \quad S = 1, 2, \dots, N \quad (5.2)$$

The steps two and three of the algorithm remain the same, initializing the levels  $Low = 0$  and  $High = 1$ , and finding the node of highest utilization,  $n = \text{argmax } y_S$ .

In Step 4.1, the utilization and throughput are computed. Since the utilization computations are relative in this algorithm, their computation remains unaltered. However, the throughput computation is modified to include the multi-server case.

$$\rho_n = (low + high)/2 \quad (5.3)$$

$$\rho_n = \frac{y_S}{y_n} \rho_n, \quad S = 1, 2, \dots, N \quad (5.4)$$

$$\lambda_S = \frac{\rho_S m_S}{\tau_S}, \quad S = 1, 2, \dots, N \quad (5.5)$$

In Step 4.2,  $c_{a,S}^2$  is computed. For the cases where it is required, the departure variability

for a multi-server node case is computed with Equation 4.10. The traffic variability equations will have to be modified accordingly.

In Step 4.3 of the PDMCQN-2 algorithm, the waiting time in queue at a node is computed. For the single server nodes, the main waiting time in queue is computed following the original algorithm. For multi-server nodes, we need an expression for the mean waiting time in queue that takes into account the closed nature of the system. The multi-server average waiting time is computed using the Equation 4.29 and the approach explained in Chapter IV.

Following the PDMCQN-2 algorithm, in Step 4.4 the value  $\delta$  is computed as follows,

$$\delta = \sum_{S=1}^N \lambda_S (Wq_S + \tau_S) - K \quad (5.6)$$

Based on the  $\delta$  value, the algorithm modifies the values of low =  $\rho_n$  if  $\delta < 0$  or high =  $\rho_n$  if  $\delta > 0$ .

The PDMCQN-2 algorithm's main metric of interest is the mean waiting time in queue. However, for comparison and completeness purposes the average number of customers per node is also computed in this case. For computing the average number of customers in queue per node,  $Lq_s$  and the average number of customers in the node  $L_S$ , equations 4.23 and 4.24 are used.

## 5.2 General CQN Results

In order to evaluate the performance of the CQN algorithm for networks with multi-server nodes, some test configurations were utilized. Our main focus here was to test how well the new corrections for the waiting time in a multi-server node performed. The example

configurations were designed to have a sequential flow to keep the routing simple and also represent the general flow of patients through the hospital system. The basic model consisted of five nodes, each emulating a key step in health system process.

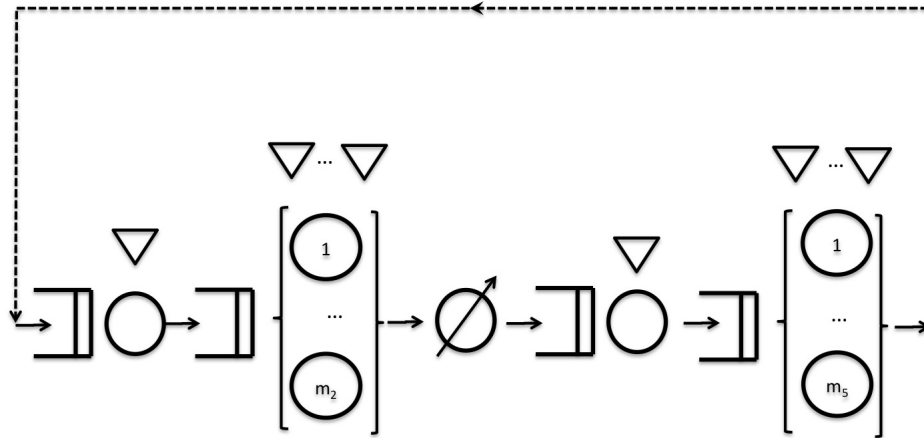


Figure 5.1: Network Structure Used For Testing Purposes

The configurations proposed for testing follow a tandem network structure that starts with a single-server node followed by a multi-server node emulating the patient's initial evaluation and the physician's service. The third node is a delay representing recovery in the medical process, where there is no queue for the server, but just a delay representing the desired recovery time. This effect was modeled by making the number of servers equal to the number of passive resources. The last two nodes represent the final medical evaluation and the support personnel escorting the patient out of the system, represented by a single and multi-server nodes, respectively. The parameters for the different configurations studied are shown in Table 5.1.



Table 5.1: CQN Configurations

Configuration	K	$\tau_1$	$c_{s,1}^2$	$m_1$	$\tau_2$	$c_{s,2}^2$	$m_2$	$\tau_3$	$c_{s,3}^2$	$m_3$	$\tau_4$	$c_{s,4}^2$	$m_4$	$\tau_5$	$c_{s,5}^2$	$m_5$
1	20	10.0	0.5	1	45.0	0.5	6.0	30.0	0.5	20	10.0	0.5	1.0	25.0	1.0	3
2	20	10.0	1.0	1	45.0	0.5	6.0	30.0	1.0	20	10.0	0.5	1.0	25.0	1.0	3
3	20	10.0	2.0	1	45.0	0.5	6.0	30.0	2.0	20	10.0	0.5	1.0	25.0	1.0	3
4	10	18.0	0.5	1	75.0	1.0	6.0	30.0	0.5	10	10.0	0.5	1.0	25.0	1.0	3
5	10	18.0	1.0	1	75.0	1.0	6.0	30.0	1.0	10	10.0	0.5	1.0	25.0	1.0	3
6	10	18.0	2.0	1	75.0	1.0	6.0	30.0	2.0	10	10.0	0.5	1.0	25.0	1.0	3
7	20	13.0	0.5	1	35.0	2.0	3.0	30.0	0.5	20	10.0	0.5	1.0	25.0	1.0	3
8	20	13.0	1.0	1	35.0	2.0	3.0	30.0	1.0	20	10.0	0.5	1.0	25.0	1.0	3
9	20	13.0	2.0	1	35.0	2.0	3.0	30.0	2.0	20	10.0	0.5	1.0	25.0	1.0	3
10	10	15.0	0.5	1	60.0	0.5	6.0	30.0	0.5	10	10.0	0.5	1.0	25.0	1.0	3
11	10	15.0	1.0	1	60.0	0.5	6.0	30.0	1.0	10	10.0	0.5	1.0	25.0	1.0	3
12	10	15.0	2.0	1	60.0	0.5	6.0	30.0	2.0	10	10.0	0.5	1.0	25.0	1.0	3
13	10	15.0	0.5	1	40.0	2.0	3.0	30.0	0.5	10	10.0	0.5	1.0	25.0	1.0	3
14	10	15.0	1.0	1	40.0	2.0	3.0	30.0	1.0	10	10.0	0.5	1.0	25.0	1.0	3
15	10	15.0	2.0	1	40.0	2.0	3.0	30.0	2.0	10	10.0	0.5	1.0	25.0	1.0	3

The configurations have been designed to test the accuracy of the approximations under different population and variability conditions. As mentioned previously, it is important to note that node 3 is a delay node, and nodes 2 and 5 are multi-server nodes. The variability conditions change as well in the system, combining nodes with SCVs that vary from 0.5 to 2 in the same configuration.

### 5.2.1 Error Calculation

The analytical results were compared with estimates obtained via simulation. The simulation results were obtained using a model created in Arena Software (Copyright ©2019 Rockwell Automation Technologies, Inc. Version 16.00.00000), using gamma and exponential distributions for representing the service times. The main performance values reported were node

utilization, throughput, and average number of customers in a node. The accuracy of the analytical results was evaluated using the relative percentage error for utilization, defined by

$$Error = \frac{analytical - simulation}{simulation} * 100 \quad (5.7)$$

For cases where the denominator is too small, the difference is used to show the result's accuracy, instead of the relative percentage error. This convention is assumed in order to avoid large error percentages due to division by very small numbers.

Errors in the average number of customers are presented as a proportion of the number of passive resources in the network, in order to normalize the error, following the method used by Krishnamurthy and Suri [32]

$$Error(\bar{L}_S) = \frac{\bar{L}_S^{Analytical} - \bar{L}_S^{Simulation}}{K} * 100 \quad (5.8)$$

It is important to note the errors calculated are not absolute values, rather, each error computation shows the sign of the error in addition to its magnitude. Reporting the errors in this fashion, allows the researchers and readers to find patterns in the errors that facilitate the refinement of the approximation, for instance over and under estimations. Consequently, a negative error value implies that the analytical model is under estimating, and positive error values mean the analytical model is over estimating with respect to the simulation results.

Finally, throughput values which could be much smaller than 1 because of the base time unit, the error was defined as just the difference between analytical and simulation values following the suggestion by Whitt [54]

### 5.2.2 Determination of Simulation Warm Up Period and Run Length

The simulation experiments were set for 10 independent replications, each one with a run time of 400,000 hours plus a warm up period of 150,000 hours (24,000,000 and 9,000,000 minutes, respectively. Minute is the base time period). The warm-up period was established using the Welch's procedure [52, 53]. The run length was established to guarantee tight 95% confidence intervals.

### 5.2.3 Numerical Results

Following the error metrics described in Section 5.2.1 the main results can be seen in tables 5.2, 5.3, and 5.4.

As can be seen in Table 5.2, despite the widely varying CQN configurations, the multi-server algorithm accurately predicts the system throughput. All variations registered between the analytical and simulation results fall below 0.002 showing a very good prediction performance.

Table 5.2: CQN Throughput Results

Configuration	Simulation (S)	Analytical (A)	Error
1	0.092	0.092	-0.001
2	0.091	0.090	-0.001
3	0.088	0.087	0.000
4	0.050	0.048	-0.001
5	0.048	0.046	-0.002
6	0.046	0.044	-0.002
7	0.073	0.074	0.001
8	0.073	0.073	0.001
9	0.070	0.072	0.001
10	0.057	0.056	-0.001
11	0.056	0.054	-0.002
12	0.053	0.051	-0.002
13	0.056	0.057	0.001
14	0.055	0.056	0.001
15	0.053	0.054	0.001

\* S= Simulation, A= Analytical

The utilization measure is also very well predicted by the method, as can be seen in Table 5.3, the widest gap found between the simulation and analytical results is about 4%. This should not be surprising as utilization is directly related to the throughput. Slightly larger errors were observed for smaller network population (K) values.

Table 5.3: CQN Utilization Results

Conf.	Node 1			Node 2			Node 3			Node 4			Node 5		
	S*	A*	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
1	92.5	91.72	-1.0%	69.4	68.79	-1.0%	13.9	13.76	-1.0%	92.5	91.72	-1.0%	76.6	76.44	0.0%
2	90.3	90.07	0.0%	68.1	67.55	-1.0%	13.5	13.51	0.0%	90.9	90.07	-1.0%	75.3	75.06	0.0%
3	87.5	87.36	0.0%	65.8	65.52	0.0%	13.1	13.1	0.0%	87.8	87.36	0.0%	72.7	72.8	0.0%
4	89.1	86.53	-3.0%	61.5	60.09	-2.0%	14.8	14.42	-3.0%	49.5	48.07	-3.0%	41.0	40.06	-2.0%
5	86.3	83.44	-3.0%	59.9	57.94	-3.0%	14.4	13.91	-3.0%	48.2	46.35	-4.0%	39.9	38.63	-3.0%
6	82.9	79.44	-4.0%	57.4	55.17	-4.0%	13.8	13.24	-4.0%	46.2	44.13	-4.0%	38.3	36.78	-4.0%
7	95.4	96.46	1.0%	85.5	86.57	1.0%	11.0	11.13	1.0%	73.4	74.20	1.0%	60.8	61.84	2.0%
8	93.8	95.37	2.0%	84.4	85.59	1.0%	10.8	11.00	2.0%	72.6	73.36	1.0%	60.1	61.13	2.0%
9	91.4	93.44	2.0%	82.0	83.86	2.0%	10.5	10.78	3.0%	70.5	71.88	2.0%	58.4	59.90	3.0%
10	85.8	84.27	-2.0%	57.2	56.18	-2.0%	17.2	16.85	-2.0%	57.2	56.18	-2.0%	47.4	46.82	-1.0%
11	83.0	81.13	-2.0%	55.7	54.09	-3.0%	16.6	16.23	-2.0%	55.7	54.09	-3.0%	46.1	45.07	-2.0%
12	79.7	77.10	-3.0%	53.2	51.40	-3.0%	15.9	15.42	-3.0%	53.3	51.4	-4.0%	44.1	42.84	-3.0%
13	84.6	85.96	2.0%	75.0	76.41	2.0%	16.9	17.19	2.0%	56.4	57.31	2.0%	46.7	47.76	2.0%
14	82.1	83.75	2.0%	73.3	74.45	2.0%	16.4	16.75	2.0%	55.1	55.83	1.0%	45.6	46.53	2.0%
15	79.0	80.47	2.0%	70.2	71.53	2.0%	15.8	16.09	2.0%	52.8	53.65	2.0%	43.7	44.71	2.0%

\* S= Simulation, A= Analytical

Another performance measure presented is the average number of customers in a node. This measure also indicates the prediction accuracy for the mean waiting time at node because of the relationship through Little's Law. In Table 5.4, the results for the different nodes in the various test configurations can be seen. In general, the prediction accuracy is very good (errors less than 10%) in 12 out of 15 configurations.

For configurations 7, 8, and 9 the prediction error in the first node is around 13% and around 11% for Node 2. The next highest error is around 7% can be found in configurations 13, 14, and 15. All these six configuration have one thing in common,  $c_{s,2}^2 = 2$ , a very high

variability of service time at Node 2. In a CQN an error in the average number for one node will affect accuracy for other nodes in the system, because the number of customers in the network is constant. In the experiments, the multi-server Node 2 with high variability is introducing an error that is mainly manifested in the number in node for that node and for Node 1, the node in the network with highest utilization. Node 1 by itself could also have high prediction errors given it is the one with highest utilization in the system.

Table 5.4: CQN Average Number in Node Results

Conf.	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
1	4.74	5.54	3.97%	4.42	4.13	-1.47%	2.78	2.75	-0.14%	4.66	5.17	2.55%	3.41	2.42	-4.94%
2	5.01	5.88	4.38%	4.46	4.05	-2.03%	2.70	2.70	0.01%	4.52	5.0	2.42%	3.30	2.36	-4.72%
3	5.51	6.36	4.23%	4.49	3.93	-2.76%	2.62	2.62	0.00%	4.25	4.82	2.83%	3.12	2.27	-4.28%
4	2.69	3.05	3.58%	3.77	3.61	-1.61%	1.48	1.44	-0.38%	0.78	0.7	-0.75%	1.28	1.2	-0.80%
5	2.86	3.26	3.95%	3.69	3.48	-2.10%	1.44	1.39	-0.49%	0.76	0.71	-0.44%	1.25	1.16	-0.88%
6	3.14	3.55	4.15%	3.56	3.31	-2.49%	1.38	1.32	-0.56%	0.73	0.71	-0.16%	1.20	1.10	-0.94%
7	7.36	9.97	13.06%	5.85	3.6	-11.26%	2.2	2.23	0.13%	2.32	2.33	0.05%	2.27	1.87	-1.97%
8	7.30	10.03	13.65%	5.99	3.6	-11.95%	2.16	2.20	0.20%	2.30	2.31	0.07%	2.24	1.85	-1.94%
9	7.46	10.16	13.52%	6.05	3.56	-12.44%	2.10	2.16	0.28%	2.22	2.30	0.41%	2.16	1.81	-1.75%
10	2.33	2.65	3.18%	3.46	3.37	-0.93%	1.72	1.69	-0.35%	0.98	0.89	-0.91%	1.51	1.41	-1.04%
11	2.53	2.87	3.37%	3.39	3.25	-1.44%	1.66	1.62	-0.37%	0.95	0.91	-0.42%	1.47	1.35	-1.15%
12	2.83	3.17	3.42%	3.26	3.08	-1.76%	1.59	1.54	-0.48%	0.91	0.92	0.05%	1.40	1.29	-1.18%
13	2.58	3.25	6.66%	3.19	2.51	-6.75%	1.69	1.72	0.29%	1.02	1.08	0.57%	1.51	1.43	-0.72%
14	2.73	3.44	7.08%	3.16	2.43	-7.25%	1.64	1.68	0.35%	1.00	1.06	0.57%	1.47	1.40	-0.72%
15	2.97	3.70	7.33%	3.08	2.31	-7.67%	1.58	1.61	0.29%	0.96	1.03	0.71%	1.41	1.34	-0.64%

\* S= Simulation, A= Analytical

In summary, a CQN is an important building block for modeling capacity restrictions. The parametric decomposition approach available in the literature did not include multi-server

nodes. In extending the published approach [42], the main contribution was the development of a correction factor for mean waiting time at a multi-server node. Numerical experiments indicated that the multi-server approximation performs well in a wide variety of system configurations. This extension makes the parametric decomposition approach applicable to a much broader class of CQN configurations.

## CHAPTER VI

### DOUBLE AND SINGLE FORK/JOIN SYSTEMS WITH MULTI-SERVER NODES

The fork/join system studied in this chapter is characterized by synchronization stations that pair demand arriving from different sources. In our system, there are two sources. The first arrival comes from passive resources such as beds and rooms released by patients and become available. The second source is patients that arrive and request services. None of these arrivals can go through the system without the other, and consequently, we require the synchronization stations. In Figure 6.1 a double fork/join example is presented.

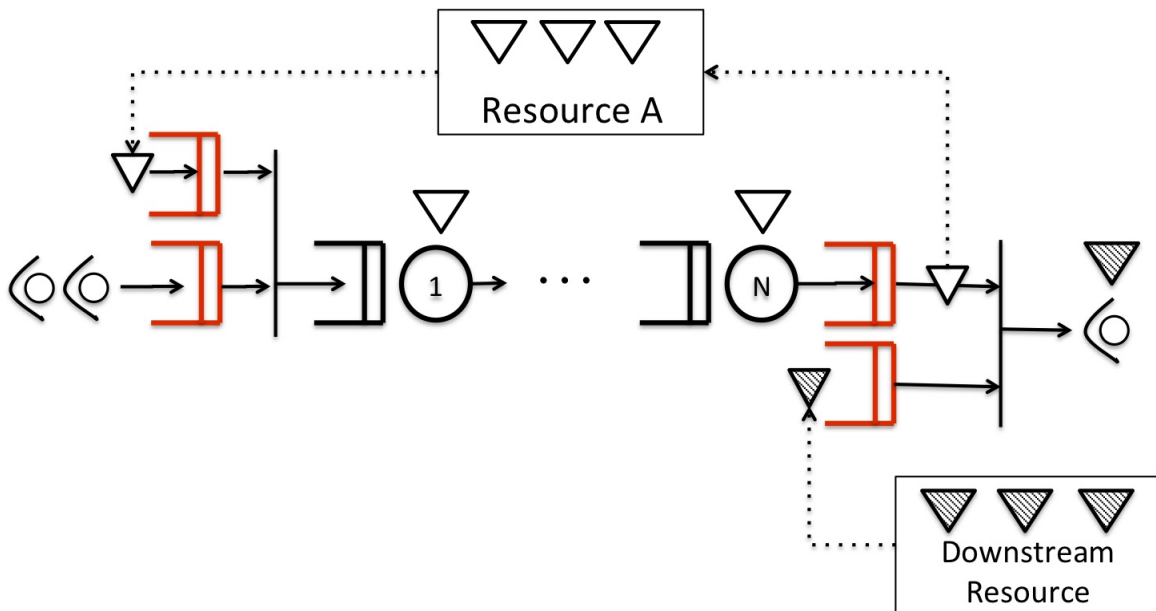


Figure 6.1: Double Fork/Join Example



The main difference between this approach and the CQN model presented in the previous chapter, is that in a CQN it was assumed the passive resources, in our case the beds or rooms, are under high demand, and patients are always waiting for them. Then the bed/room with a patient becomes a customer who is perpetually circulating in the system. On the other hand, in the fork/join system, the passive resources are expected not be under high demand and consequently are more likely to be sometimes idle. Under situations where it is possible to find passive resources idle, the synchronizing stations are vital to better model the complete system behavior.

The double fork/join (FJ) system contains these synchronization nodes at the beginning and at the end of the system. Customers at the end need to wait for resources or demand from another system downstream in order to continue their pathway. For the single FJ system, customers do not need to wait for anything else at the end and just leave the system. In this system, there is no need for a fork/join node at the end of the system and the passive resources become free and wait at the beginning for new customers to arrive to the system. While the double fork/join system is designed to model a pull-based production system, the single fork/join version is well suited for service systems, where the customers can leave the system after service completion.

The notation previously introduced needs to be expanded with some additional values for the fork/join case. The notation used in this document is based on the notation suggested by Krishnamurthy and Suri [32] and is shown in Table 6.1. Figure 6.2 helps in explaining the notation for the fork/join system

Table 6.1: Additional notation needed for the fork/join system

---

$J_i$	Fork/Join station at the end of stage $i$ and the beginning of stage $i + 1$ . ( $J_0$ is the fork/join station at the beginning of Stage 1)
$\lambda_{P,i}$	Customer arrival rate to the fork/join node $J_i$ .
$c_{P,i}^2$	Customer inter-arrival time's SCV at the fork/join node $J_i$ .
$\lambda_{F,i}$	Free capacity-restricting resource arrival rate to the fork/join node $J_{i-1}$
$c_{F,i}^2$	Free capacity-restricting resource inter-arrival time's SCV to the fork/join node $J_{i-1}$ .
$\lambda_{D,i}$	Departure rate from the fork/join node $J_i$ .
$c_{D,i}^2$	Inter-departure time's SCV at the fork/join node $J_i$ .
$\lambda_{d,i}$	Departure rate from the network in Stage $i$ .
$c_{d,i}^2$	Inter-departure time's SCV for the network in Stage $i$ .

---

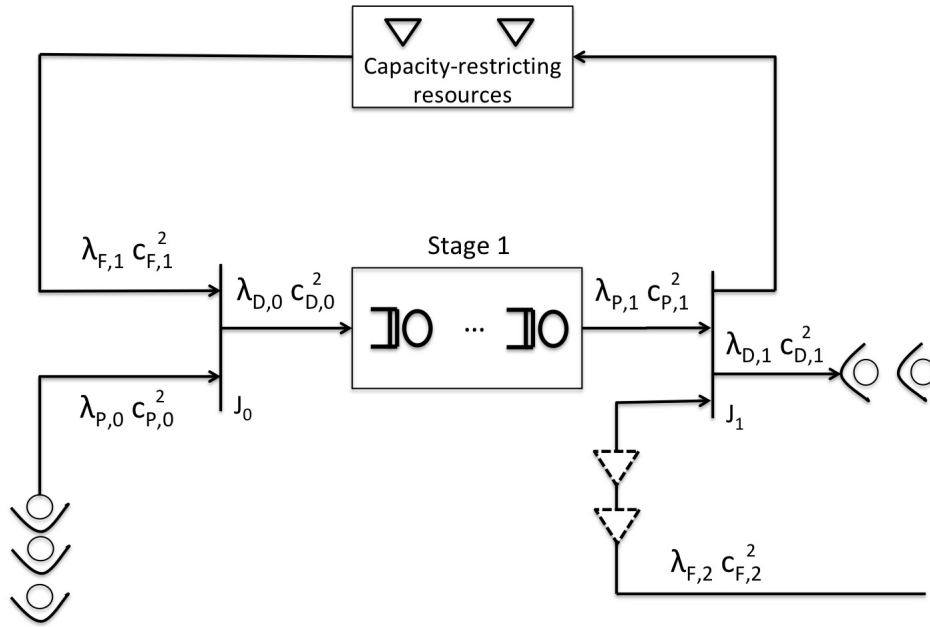


Figure 6.2: Fork/Join Notation

## 6.1 The Fork/Join System Solution

In order to solve the fork/join system, the solution approach proposed by Krishnamurthy and Suri [32] was utilized. As previously explained, the fork/join system is utilized in cases where the first level system's passive resources are not utilized always or very heavily. In this case, synchronization stations are required to harmonize the customer arrivals with the passive resources once they are available.

The solution proposed by Krishnamurthy and Suri [32] follows the parametric decomposition method. According to the authors, the approach has four steps: decomposition, characterization, linkage and solution. Perhaps the more complex step in this case is the linkage. This linkage, step takes care of possible arrival shut downs generated by the probability that all the passive resources find themselves at the synchronization station, and consequently no more arrivals become possible. This situation produces a condition not compatible with the common arrival process studied in the queueing literature. The authors approach this challenge by introducing stochastic transformations that deal with the variation in the arrival process and emulate the process required.

In this section, the original algorithm with two synchronization stations, one at the beginning and another at the end, is presented. In Section 6.1.1, the adjustment needed for the single stage version, where only one synchronization station is required, will be explained. In [32], Krishnamurthy and Suri proposed three linkage procedures that will be explained in the following paragraphs.

The first linking procedure connects the departure from the nodes in the system to the synchronization station at the end of the system. Consequently, this process relates  $(\lambda_{P,i}^{-1}, c_{P,i}^2)$

to  $(\lambda_{d,i}^{-1}, c_{d,i}^2)$ , the main equations used are

$$\lambda_{P,i} = \frac{\lambda_{d,i}}{1 - \pi_{P,i}} \quad (6.1)$$

$$c_{P,i}^2 = \frac{c_{d,i}^2}{(1 - \pi_{P,i})^2} - \left( \frac{\pi_{P,i}}{(1 - \pi_{P,i})^2} \right) \left( \frac{\lambda_{d,i}}{\lambda_{F,i+1}} \right) \left( \frac{2c_{F,i+1}^2}{1 + c_{F,i+1}^2} \right) \quad (6.2)$$

and

$$\lambda_{D,i} = \lambda_{d,i} \quad (6.3)$$

Where  $\pi_{P,i}$  is the long run proportion of time the customer arrivals to fork/join node  $J_i$  is shut down. According to Krishnamurthy and Suri [32], this shut down effect is introduced by the possibility that, at some time, all the passive resources in stage  $i$  queue in the fork/join  $J_i$ . Consequently, at these periods there is no customer departure from the fork/join station  $J_{i-1}$ , generating the shut down referenced here.

From the characterization step,  $\lambda_{D,i}$  can be computed as follows,

If  $\omega_i \neq 1$ , ( $\omega_i = \lambda_{P,i}/\lambda_{F,i+1}$ )

$$\lambda_{D,i} = \lambda_{P,i} \left[ \frac{1 - \omega_i^{K_i + K_{i+1}}}{1 - \omega_i^{K_i + K_{i+1} + 1}} \right] \left[ 1 - 0.5(c_i^2 - 1) \left( \frac{(1 - \omega_i)\omega_i^{K_i + K_{i+1}}}{1 - \omega_i^{2(K_i + K_{i+1}) + 1}} \right) \right] \quad (6.4)$$

where  $c_i^2 = 0.5(c_{P,i}^2 + c_{F,i+1}^2)$ .

If  $\omega_i = 1$ ,

$$\lambda_{D,i} = \lambda_{P,i} \left( \frac{K_i + K_{i+1}}{K_i + K_{i+1} + 1} \right) \left( 1 - \frac{0.5(c_i^2 - 1)}{2(K_i + K_{i+1}) + 1} \right) \quad (6.5)$$

In order to solve this first linkage, the following algorithm was proposed by Krishnamurthy and Suri [32],

---

**Algorithm 3** Krishnamurthy and Suri's Algorithm 1 [32]

---

**Require:**  $\lambda_{d,i}$ ,  $c_{d,i}^2$ ,  $K_i$ ,  $\lambda_{F,i+1}$ ,  $c_{F,i+1}^2$ , and  $K_{i+1}$ .

**Ensure:**  $Low = 0$ ,  $High = 1$ ,  $g = 1$  and  $\epsilon$ .

**Begin loop**

**while**  $|\delta| > \epsilon$  **do**

**Step 1**

Set  $\pi_{P,i}^{(g)} = (Low + High)/2$

**Step 2**

Compute  $\lambda_{P,i}$  by setting  $\pi_{P,i} = \pi_{P,i}^{(g)}$  in equation 6.1

**Step 3**

Compute  $c_{P,i}^2$  using Equation 6.2

**Step 4**

Compute  $\lambda_{D,i}$  using Equations 6.4 and 6.5 and set  $\lambda_{D,i}^{(g)} = \lambda_{D,i}$

**Step 5**

Compute  $\delta = \lambda_{D,i}^{(g)} - \lambda_{d,i}$

**Step 6**

If  $\delta < -\epsilon$  then set  $Low = \pi_{P,i}^{(g)}$

If  $\delta > \epsilon$  then set  $High = \pi_{P,i}^{(g)}$

If  $|\delta| \leq \epsilon$  then STOP.

Set  $g = g + 1$

**end while**

---

In Algorithm 3,  $\pi_{P,i}^{(g)}$  represents the  $\pi_{P,i}$  estimation in the  $g^{th}$  iteration. Another stochastic transformation is required in order to link the departure process from fork/join node  $J_i$  at

the end of the stage  $i$ , to the arrival process at fork/join node  $J_{i-1}$  at the beginning of the stage  $i$ . Following the same logic, Krishnamurthy and Suri proposed the following equations to compute the values required.

$$\lambda_{F,i} = \frac{\lambda_{D,i}}{1 - \pi_{F,i}} \quad (6.6)$$

$$c_{F,i}^2 = \frac{c_{D,i}^2}{(1 - \pi_{F,i})^2} - \left( \frac{c_{P,i-1}^2}{1 + c_{P,i-1}^2} \right) \left( \frac{\lambda_{D,i}}{\lambda_{P,i-1}} \right) \left( \frac{2\pi_{F,i}}{(1 - \pi_{F,i})^2} \right) \quad (6.7)$$

and

$$\lambda_{D,i-1} = \lambda_{D,i} \quad (6.8)$$

Similarly, they used an iterative algorithm to compute the values required in this new transformation. The algorithm is as follows.

---

**Algorithm 4** Krishnamurthy and Suri's Algorithm 2 [32]

---

**Require:**  $\lambda_{D,i}$ ,  $c_{D,i}^2$ ,  $K_i$ ,  $\lambda_{P,i-1}$ ,  $c_{P,i-1}^2$ , and  $K_{i-1}$ .

**Ensure:**  $Low = 0$ ,  $High = 1$ ,  $g = 1$  and  $\epsilon$ .

**Begin loop**

**while**  $|\delta| > \epsilon$  **do**

**Step 1**

Set  $\pi_{F,i}^{(g)} = 0.5(Low + High)$

**Step 2**

Compute  $\lambda_{F,i}$  by setting  $\pi_{F,i} = \pi_{F,i}^{(g)}$  in Equation 6.6

**Step 3**

Compute  $c_{F,i}^2$  using Equation 6.7

**Step 4**

Compute  $\lambda_{D,i-1}$  using equations 6.4 and 6.5 and set  $\lambda_{D,i-1}^{(g)} = \lambda_{D,i-1}$

**Step 5**

Compute  $\delta = \lambda_{D,i-1}^{(g)} - \lambda_{D,i}$

**Step 6**

If  $\delta < -\epsilon$  then set  $Low = \pi_{F,i}^{(g)}$

If  $\delta > \epsilon$  then set  $High = \pi_{F,i}^{(g)}$

If  $|\delta| \leq \epsilon$  then STOP.

Set  $g = g + 1$

**end while**

---

The final linking process is characterized by the equations that synchronize the departure from  $J_{i-1}$  to the arrivals to the network in Stage  $i$ . For this specific linking, the authors rely in the flow conservation property, and use the following equations:

$$\lambda_{a,i} = \lambda_{D,i-1} \quad (6.9)$$

and

$$c_{a,i}^2 = c_{D,i-1}^2 \quad (6.10)$$

Before solving the fork/join system, we need discuss the solution of the network in Stage  $i$ . Krishnamurthy and Suri [32] used a tandem network structure. The network structure can be analyzed using the standard parametric-decomposition method. The solution of the traffic rate equations is simplified for the tandem structure with the departure rate of a node equal to its arriving rate because of the flow conservation, which finally results in the following.

$$\lambda_{d,i} = \lambda_{a,i} \quad (6.11)$$

The interdeparture time SCV for each of the nodes can be computed using equation 4.9.

In Krishnamurthy and Suri's [32] original paper, equations for computing utilization, mean waiting time and the mean number of customers at node, were presented only for single-server nodes. Also, the mean waiting time in a GI/G/1 queue was computed using the simple approximation that does not include the factor proposed by Kramer and Lagenbach-Belz [31] and presented in equation 4.11 in Section 4.4. We included this approach in our implementation as it improves the overall accuracy of performance prediction, as shown by the numerical results included in Appendix A.

The importance of including multi-server nodes in the first-level network models was explained in Chapter V. Our FJ implementation includes multi-server nodes and we used the correction factor introduced in Table 4.2 for the average waiting time computation in



capacity-restricted systems.

Following Krishnamurthy and Suri's fork/join approach, there are additional equations characterizing the synchronizing nodes that need to be computed in order to obtain the desired metrics. Already, the characterizing equations 6.4 and 6.5 were introduced. The other characterizing equations are also classified based on the relation  $\omega$  as explained earlier:

If  $\omega \neq 1$ ,

$$\bar{L}_{P,i} = \left[ \left( \frac{\omega_i^{K_{i+1}}}{1 - \omega_i} \right) \left( \frac{1 - \omega_i^{K_i}}{1 - \omega_i^{K_i + K_{i+1} + 1}} \right) - \left( \frac{K_i \omega_i^{K_i + K_{i+1} + 1}}{1 - \omega_i^{K_i + K_{i+1} + 1}} \right) \right] \\ * \left[ 1 + \left( \frac{1 - \omega_i}{1 + \omega_i} \right) \left( \frac{\omega_i^4}{1 + \omega_i^8} \right) (c_i^2 - 1) \right] \quad (6.12)$$

$$\bar{L}_{F,i+1} = \left[ \left( \frac{K_{i+1}}{1 - \omega_i^{K_i + K_{i+1} + 1}} \right) - \left( \frac{\omega_i}{1 - \omega_i} \right) \left( \frac{1 - \omega_i^{K_{i+1}}}{1 - \omega_i^{K_i + K_{i+1} + 1}} \right) \right] \\ * \left[ 1 + \left( \frac{1 - \omega_i}{1 + \omega_i} \right) \left( \frac{\omega_i^4}{1 + \omega_i^8} \right) (c_i^2 - 1) \right] \quad (6.13)$$

On the other hand, if  $\omega = 1$ ,

$$\bar{L}_{P,i} = \left( \frac{K_i}{2} \right) \left( \frac{K_i + 1}{K_i + K_{i+1} + 1} \right) \quad (6.14)$$

$$\bar{L}_{F,i+1} = \left( \frac{K_{i+1}}{2} \right) \left( \frac{K_{i+1} + 1}{K_i + K_{i+1} + 1} \right) \quad (6.15)$$

Independent of the  $\omega$  ratio, the SCV is computed with the following equation.

$$\begin{aligned}
c_{D,i}^2 = & \left[ \left( \frac{\lambda_{P,i}^5}{\lambda_{P,i}^5 + \lambda_{F,i+1}^5} \right) c_{F,i+1}^2 + \left( \frac{\lambda_{F,i+1}^5}{\lambda_{P,i}^5 + \lambda_{F,i+1}^5} \right) c_{P,i}^2 \right] \\
& * \left[ 1 - \frac{1}{K_i + K_{i+1} + 1} - \frac{1}{(K_i + K_{i+1} + 1)^2} \right] \\
& * \left[ \frac{\sqrt{\lambda_{P,i}^4 + \lambda_{F,i+1}^4}}{\lambda_{P,i}^2 + \lambda_{F,i+1}^2} \right]
\end{aligned} \tag{6.16}$$

These equations and the the former linking algorithms are integrated in a final algorithm by Krishnamurthy and Suri [32]. The algorithm is as follows,

---

**Algorithm 5** Krishnamurthy and Suri's Algorithm 3 [32]

---

**Require:**  $\lambda_{P,i-1}$ ,  $c_{P,i-1}^2$ ,  $K_{i-1}$ ,  $K_i$ ,  $\lambda_{F,i+1}$ ,  $c_{F,i+1}^2$ ,  $K_{i+1}$ , and  $\mu_{S,i}$ , which is the maximum node service rate in the network.

**Ensure:**  $Low = 0$ ,  $High = \min(\lambda_{P,i-1}, \mu_{S,i}, \lambda_{F,i+1})$ ,  $g = 1$ .

**Begin loop 1**

**while**  $|\delta_1| > \epsilon$  **do**

**Step 1**

$\lambda_{d,i}^{(g)} = (Low + High)/2$  and  $c_{d,i}^{2(g)} = 1$  (say)

**Begin loop 2**

**Step 2**

**while**  $|\delta_2| > \epsilon$  **do**

**Step 2.1**

Solve for  $\pi_{P,i}$ ,  $\lambda_{P,i}$ , and  $c_{P,i}^2$  using Krishnamurthy and Suri's [32] algorithm 1 setting  $\lambda_{d,i} = \lambda_{d,i}^{(g)}$  and  $c_{d,i}^2 = c_{d,i}^{2(g)}$  in linking Eqs. 6.1, 6.2 and 6.3

**Step 2.2**

Compute  $\lambda_{D,i}$ ,  $c_{D,i}^2$ ,  $\bar{L}_{P,i}$ , and  $\bar{L}_{F,i+1}$  using characterization Eqs. 6.4, 6.5, 6.12, 6.13, 6.14, 6.15, and 6.16.

**Step 2.3**

Solve for  $\pi_{F,i}$ ,  $\lambda_{F,i}$ , and  $c_{F,i}^2$  using algorithm 2 and linking Eqs. 6.6, 6.7 and 6.8

**Step 2.4**

Compute  $\lambda_{D,i-1}$ ,  $c_{D,i-1}^2$ ,  $\bar{L}_{P,i-1}$ , and  $\bar{L}_{F,i}$  using characterization Eqs. 6.4, 6.5, 6.12, 6.13, 6.14, 6.15, and 6.16.

**Step 2.5**

Compute  $\lambda_{a,i}$ ,  $c_{a,i}^2$  using linking Eqs. 6.9 and 6.10.

**Step 2.6**

Solve the network and compute  $\lambda_{d,i}$ ,  $c_{d,i}^2$ , and  $\bar{L}_{S,i}$ , the average number of customers in the network.

**Step 2.7**

Compute  $\delta_2 = |c_{d,i}^2 - c_{d,i}^{2(g)}|$ . Set  $c_{d,i}^{2(g)} = c_{d,i}^2$

**end while**

**Step 3**

Compute  $\delta_1 = \bar{L}_{F,i} + \bar{L}_{S,i} + \bar{L}_{P,i} - K_i$ .

**Step 4**

If  $\delta < -\epsilon$  then set  $Low = \lambda_{d,i}^{(g)}$

If  $\delta > \epsilon$  then set  $High = \lambda_{d,i}^{(g)}$

Set  $g = g + 1$

**end while**

---

### 6.1.1 The Fork/Join System Solution with One Synchronization Station

In the previous section, the overall Krishnamurthy and Suri [32] algorithm for solving the fork/join system was explained. However, the variation presented in this section has a significant conceptual difference. The kanban system that motivated the work of Krishnamurthy and Suri [32] corresponds to a pull system and the production units completed wait for demand to arrive and consume them. The motivation for our study is health care services. In health care services, as in all other services, there is no concept equivalent to a finished product. Hence, the inherent system in services is a push system. The service starts when the customer arrives and when the customer pathway ends, the customer simply leaves. For modeling purposes, there is no synchronization node at the end. The synchronization station in the beginning is needed to match a free capacity-restricting resource to a new patient.

In hospitals, for instance, once patients complete their service in the last node, they can exit the system immediately. Consequently, the last synchronization station in the Krishnamurthy and Suri [32] algorithm would not be required. In these cases, some adjustment to the algorithm discussed in the previous sections is required.

The system under study in this section is close to what in literature is called Open Queuing Networks with Restricted Capacity (OQN-RC) or Semi-Open Queueing Networks. However, they are not exactly the same. The main difference is that in this section's case, customers arrive to the system controlled by a capacity restriction belonging to the previous process section, while in the OQN-RC they arrive freely. In other words, customers in the case of this section, cannot accumulate infinitely in the synchronization station at the entry to the system. They can accumulate up to a specific limit.

Basically, the four steps in the original Krishnamurthy and Suri [32] algorithm remain the same. However, there are adjustments required in some of them. The decomposition and characterization steps remain unaltered. The linking step is where modifications are required. Later, the solution step will change based on the adjustments made in the linking part.

The linking step includes three stochastic transformations, in two of the three linking steps, this transformation includes the use of an specific algorithm for this purpose. According to Krishnamurthy and Suri [32], the stochastic transformations are required in order to properly model the departure times from the synchronization stations, including the arrival shut downs motivated by the probability that all the passive resources could be in the synchronization station in some specific moment, and no more arrivals are possible in these periods.

In the original Krishnamurthy and Suri solution, Algorithm 1 takes the departure rate from the service nodes and the external demand, and transforms them into the departing rate from the second synchronization station. This step is not required in the variation under study, given there does not exist any external demand, which obviates the need for a second synchronization station. Consequently, in the correction for this case, the values are simply  $\lambda_{D,i} = \lambda_{d,i}$  and  $c_{D,i}^2 = c_{d,i}^2$ . Since the second synchronization station is not required anymore, there will not be any patients waiting at this station and consequently the value  $L_{P,i} = 0$ . Once the previous values are defined, the Krishnamurthy and Suri's algorithm 1 is not required anymore. The other aspects of the solution remain the same.

### **Other Potential Approaches**

As previously explained, the single fork/join station system considered in this section is a system similar to the SOQN. Jia and Heragu [27] pointed out that the number of passive resources in the system must be large enough to avoid the external queue from exploding.

However, no method was provided in order to compute a threshold for the number of passive resources. On the other hand, Roy [41] pointed the arrival rate should not be higher than the maximum throughput the system is capable of delivering, which can be computed with a CQN version of the problem.

Additionally, the main techniques used to solve the SOQN are the matrix-geometric method, and aggregation method that is based on the Mean Value Analysis (MVA). None of these techniques is compatible with the parametric decomposition method that is the basis for solving the other building-blocks in this document.

## 6.2 Fork-Join Numerical Experiments

The experiments for the fork/join systems follow the structure used for the CQN experiments, where five nodes constitute the network under study, emulating health care system as explained in Chapter V.

The FJ system analyzed is basically the same system studied in the CQN chapter and can be seen in Table 6.2. However, there are two customer arrival processes, one at the end and one at the beginning. As explained earlier in this chapter, these two arrival processes and the nodes are required for synchronizing the different arrivals.

In order to test the FJ system, arrivals rates were set to 75% to 85% of the throughput found for the CQN system. The arrival rates are shown in Table 6.3.

Table 6.2: Double FJ Experimental Configurations

Conf.	$K_1$	$\tau_1$	$c_{s,1}^2$	$m_1$	$\tau_2$	$c_{s,2}^2$	$m_2$	$\tau_3$	$c_{s,3}^2$	$m_3$	$\tau_4$	$c_{s,4}^2$	$m_4$	$\tau_5$	$c_{s,5}^2$	$m_5$
1	20	10.0	0.5	1	45.0	0.5	6	30.0	0.5	20	10.0	0.5	1.0	25.0	1.0	3
2	20	10.0	1.0	1	45.0	0.5	6	30.0	1.0	20	10.0	0.5	1.0	25.0	1.0	3
3	20	10.0	2.0	1	45.0	0.5	6	30.0	2.0	20	10.0	0.5	1.0	25.0	1.0	3
4	10	18.0	0.5	1	75.0	1.0	6	30.0	0.5	10	10.0	0.5	1.0	25.0	1.0	3
5	10	18.0	1.0	1	75.0	1.0	6	30.0	1.0	10	10.0	0.5	1.0	25.0	1.0	3
6	10	18.0	2.0	1	75.0	1.0	6	30.0	2.0	10	10.0	0.5	1.0	25.0	1.0	3
7	20	13.0	0.5	1	35.0	2.0	3	30.0	0.5	20	10.0	0.5	1.0	25.0	1.0	3
8	20	13.0	1.0	1	35.0	2.0	3	30.0	1.0	20	10.0	0.5	1.0	25.0	1.0	3
9	20	13.0	2.0	1	35.0	2.0	3	30.0	2.0	20	10.0	0.5	1.0	25.0	1.0	3
10	10	15.0	0.5	1	60.0	0.5	6	30.0	0.5	10	10.0	0.5	1.0	25.0	1.0	3
11	10	15.0	1.0	1	60.0	0.5	6	30.0	1.0	10	10.0	0.5	1.0	25.0	1.0	3
12	10	15.0	2.0	1	60.0	0.5	6	30.0	2.0	10	10.0	0.5	1.0	25.0	1.0	3
13	10	15.0	0.5	1	40.0	2.0	3	30.0	0.5	10	10.0	0.5	1.0	25.0	1.0	3
14	10	15.0	1.0	1	40.0	2.0	3	30.0	1.0	10	10.0	0.5	1.0	25.0	1.0	3
15	10	15.0	2.0	1	40.0	2.0	3	30.0	2.0	10	10.0	0.5	1.0	25.0	1.0	3

Table 6.3: Arrival Parameters for the Double FJ Experiments

Configurations	$\lambda_{P,0}$	$c_{P,0}^2$	$K_0$	$\lambda_{F,2}$	$c_{F,2}^2$	$K_2$
1	0.079	1.0	20	0.079	1.0	20
2	0.077	1.0	20	0.077	1.0	20
3	0.075	1.0	20	0.075	1.0	20
4	0.042	1.0	10	0.042	1.0	10
5	0.041	1.0	10	0.041	1.0	10
6	0.039	1.0	10	0.035	1.0	10
7	0.062	1.0	20	0.055	1.0	20
8	0.062	1.0	20	0.054	1.0	20
9	0.060	1.0	20	0.053	1.0	20
10	0.049	1.0	10	0.043	1.0	10
11	0.042	1.0	10	0.047	1.0	10
12	0.040	1.0	10	0.045	1.0	10
13	0.042	1.0	10	0.048	1.0	10
14	0.041	1.0	10	0.047	1.0	10
15	0.040	1.0	10	0.045	1.0	10



Table 6.4: Percentages of CQN Throughput Used to Set the DFJ Arrival Rates

Configuration	$\lambda_{P,0}$	$\lambda_{F,2}$
1	85.0%	85.0%
2	85.0%	85.0%
3	85.0%	85.0%
4	85.0%	85.0%
5	85.0%	85.0%
6	85.0%	75.0%
7	85.0%	75.0%
8	85.0%	75.0%
9	85.0%	75.0%
10	85.0%	75.0%
11	75.0%	85.0%
12	75.0%	85.0%
13	75.0%	85.0%
14	75.0%	85.0%
15	75.0%	85.0%

As can be seen in Table 6.4, the algorithm is tested using different arrival configurations, from balanced arrivals to bigger arrivals to the first synchronization node or to the second synchronization node.

As mentioned earlier, in our implementation we used the Kramer and Lagenbach-Belz [31] Equation 4.11 for the GI/G/1 waiting time approximation. To show the improvement in the prediction accuracy and validate our implementation of the Double FJ algorithm, results for the configurations studied by Krishnamurthy and Suri are presented in Appendix A.

### 6.2.1 Double Fork-Join Results

The system throughput, node utilizations, and average number in node values were utilized to study the accuracy of the analytical results, for both single and double FJ systems. The number in node in the synchronization stations is shown as a new result. The results were compared with simulation results and error estimations were computed using the methods presented in Chapter V. An important aspect to keep in mind is that the error in the average number of customers per node has been normalized with the number of passive resources in the stage, for all nodes including the SRP. This approach was used in Krishnamurthy and Suri [32].

In Table 6.5 the throughput values obtained can be found.

Table 6.5: Double Fork-Join Throughput Results

Configuration	Simulation (S)	Analytical (A)	Error
1	0.077	0.076	-0.001
2	0.076	0.075	-0.001
3	0.073	0.072	-0.001
4	0.040	0.038	-0.002
5	0.039	0.037	-0.002
6	0.034	0.033	-0.001
7	0.055	0.055	0.000
8	0.055	0.054	-0.001
9	0.053	0.053	-0.001
10	0.043	0.041	-0.002
11	0.042	0.040	-0.001
12	0.040	0.038	-0.001
13	0.042	0.041	-0.001
14	0.041	0.040	-0.001
15	0.039	0.038	-0.001

The throughput values for this system must be very close to the external arrival rate of customers. However, Krhishnamurthy and Suri [32] approach requires the specification of a  $K$  value for external customer queue as they assume that these arrivals also happen from another capacity-restricted system. Setting  $K$  to be large is approximately equivalent to not having a limit on the external customer queue size at the first synchronization node. In our experiments we set this value to be equal to the first level population constraint following examples in [32]. As a result, there could be a small fraction of external arrivals that may be lost if this queue becomes full.

As can be seen in Table 6.5, the throughput results are very close to the throughput values obtained from simulation, with the biggest difference being 0.002.

In Table 6.6 results for the utilization metrics for all the five nodes are shown. It can be seen the algorithm's ability to predict the node utilizations is good, even with the modifications to include multi-servers nodes.

Table 6.6: Double Fork-Join Utilization Results

Conf.	Node 1			Node 2			Node 3			Node 4			Node 5		
	S*	A*	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
1	77.4	76.03	-1.77%	58.1	57.02	-1.85%	11.6	11.4	-1.68%	77.4	76.03	-1.77%	64.1	63.36	-1.16%
2	75.6	74.56	-1.37%	57.0	55.92	-1.89%	11.3	11.18	-1.02%	76.0	74.56	-1.89%	63.0	62.13	-1.37%
3	73.1	71.88	-1.67%	55.0	53.91	-1.98%	11.0	10.78	-1.98%	73.3	71.88	-1.94%	60.8	59.90	-1.48%
4	72.6	69.28	-4.57%	50.1	48.11	-3.96%	12.1	11.55	-4.57%	40.3	38.49	-4.49%	33.4	32.08	-3.96%
5	70.2	67.25	-4.20%	48.8	46.7	-4.30%	11.7	11.21	-4.2%	39.3	37.36	-4.93%	32.5	31.14	-4.20%
6	61.8	59.69	-3.41%	42.8	41.45	-3.15%	10.3	9.95	-3.41%	34.5	33.16	-3.88%	28.5	27.64	-3.03%
7	72.0	71.37	-0.87%	64.4	64.05	-0.54%	8.3	8.24	-0.78%	55.4	54.9	-0.9%	45.9	45.75	-0.32%
8	70.8	70.49	-0.44%	63.8	63.26	-0.85%	8.2	8.13	-0.81%	54.8	54.22	-1.05%	45.4	45.18	-0.47%
9	68.9	68.36	-0.78%	61.8	61.35	-0.72%	8.0	7.89	-1.4%	53.1	52.59	-0.96%	44.0	43.82	-0.40%
10	64.2	61.94	-3.52%	42.8	41.29	-3.52%	12.8	12.39	-3.22%	42.8	41.29	-3.52%	35.5	34.41	-3.07%
11	62.1	60.21	-3.04%	41.6	40.14	-3.50%	12.4	12.04	-2.88%	41.6	40.14	-3.5%	34.5	33.45	-3.04%
12	59.4	57.47	-3.25%	39.7	38.31	-3.50%	11.9	11.49	-3.41%	39.7	38.31	-3.5%	32.9	31.93	-2.96%
13	63.3	61.51	-2.82%	56.1	54.68	-2.53%	12.7	12.3	-3.13%	42.2	41.01	-2.82%	34.9	34.17	-2.08%
14	61.4	59.97	-2.33%	54.8	53.31	-2.72%	12.3	11.99	-2.49%	41.2	39.98	-2.96%	34.1	33.32	-2.30%
15	58.9	57.41	-2.54%	52.4	51.03	-2.62%	11.8	11.48	-2.70%	39.4	38.27	-2.87%	32.6	31.89	-2.17%

\* S= Simulation, A= Analytical

Similarly, the results for the average number of customers per node are accurate, and the analytical model results track the simulation results, for both, single and multi-server nodes.

Table 6.7: Double Fork-Join Average Number in Node Results

Conf.	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
1	2.45	2.04	-2.0%	3.59	3.5	0.0%	2.32	2.28	0.0%	2.41	1.96	-2.0%	2.43	2.32	-1.0%
2	2.73	2.38	-2.0%	3.57	3.46	-1.0%	2.26	2.24	0.0%	2.36	2.04	-2.0%	2.37	2.29	0.0%
3	3.22	2.87	-2.0%	3.53	3.36	-1.0%	2.20	2.16	0.0%	2.27	2.08	-1.0%	2.27	2.24	0.0%
4	1.67	1.42	-2.0%	3.04	2.9	-1.0%	1.21	1.15	-1.0%	0.58	0.50	-1.0%	1.03	0.99	0.0%
5	1.81	1.66	-2.0%	2.97	2.83	-1.0%	1.17	1.12	0.0%	0.57	0.51	-1.0%	1.00	0.97	0.0%
6	1.74	1.65	-1.0%	2.6	2.51	-1.0%	1.03	0.99	0.0%	0.48	0.46	0.0%	0.87	0.86	0.0%
7	2.04	1.81	-1.0%	2.84	2.71	-1.0%	1.66	1.65	0.0%	1.11	1.02	0.0%	1.52	1.53	0.0%
8	2.33	2.14	-1.0%	2.89	2.75	-1.0%	1.64	1.63	0.0%	1.10	1.03	0.0%	1.50	1.52	0.0%
9	2.82	2.60	-1.0%	2.9	2.78	-1.0%	1.60	1.58	0.0%	1.08	1.03	0.0%	1.45	1.49	0.0%
10	1.29	1.08	-2.0%	2.58	2.48	-1.0%	1.28	1.24	0.0%	0.64	0.54	-1.0%	1.10	1.06	0.0%
11	1.40	1.26	-1.0%	2.51	2.42	-1.0%	1.24	1.20	0.0%	0.62	0.56	-1.0%	1.07	1.04	0.0%
12	1.59	1.49	-1.0%	2.4	2.31	-1.0%	1.19	1.15	0.0%	0.59	0.56	0.0%	1.02	1.01	0.0%
13	1.30	1.18	-1.0%	2.04	1.95	-1.0%	1.27	1.23	0.0%	0.65	0.59	-1.0%	1.09	1.07	0.0%
14	1.41	1.32	-1.0%	2.01	1.92	-1.0%	1.23	1.20	0.0%	0.63	0.59	0.0%	1.06	1.05	0.0%
15	1.60	1.52	-1.0%	1.95	1.87	-1.0%	1.18	1.15	0.0%	0.60	0.57	0.0%	1.01	1.01	0.0%

\* S= Simulation, A= Analytical

Finally, the mean number of customers waiting in the queue for the synchronization stations or fork-join stations was computed. The results show a good agreement between the simulation and the analytical models.

Table 6.8: Double FJ Mean Number in Queue for Fork-Join Nodes

Configuration	Initial Fork-Join Node			Final Fork-Join Node		
	S	A	Error	S	A	Error
11	4.75	6.45	-8.46%	3.41	3.96	-2.74%
2	4.84	6.62	-8.91%	3.38	3.83	-2.23%
3	5.12	6.81	-8.43%	3.35	3.69	-1.72%
4	2.94	3.95	-10.08%	1.25	1.52	-2.76%
5	3.02	4.05	-10.31%	1.27	1.47	-1.99%
6	5.20	5.66	-4.56%	2.77	2.94	-1.68%
7	13.25	13.36	-0.56%	10.45	10.87	-2.13%
8	13.21	13.38	-0.84%	10.17	10.51	-1.72%
9	13.29	13.41	-0.57%	9.80	10.10	-1.52%
10	5.16	5.59	-4.34%	2.64	2.97	-3.38%
11	1.00	2.35	-13.45%	0.52	0.61	-0.93%
12	1.14	2.38	-12.43%	0.56	0.61	-0.44%
13	0.94	2.00	-10.59%	0.61	0.68	-0.7%
14	0.98	2.04	-10.6%	0.63	0.67	-0.42%
15	1.10	2.07	-9.7%	0.66	0.66	-0.03%

\* S= Simulation, A= Analytical

Despite the complexity of the test configurations where all five nodes represent a different challenge and the arrival rates to the synchronization nodes varied to include both homogeneous and heterogeneous cases, the algorithm, including the adaptations and modifications, performed in a satisfactory manner.

### 6.2.2 Single Fork-Join Results

For testing the single fork-join modification to the algorithm, the same configurations shown in Tables 6.2 for the node parameters and 6.3 for the arrival parameters, were used. The only exceptions are the values for the arrivals to the second fork-join node, that are not needed, since there is no second fork/join in this version of the problem.

The analytical model captures the throughput values in a very good way. In Table 6.9 the comparison between simulation and analytical results is presented. The difference between these two values is less than 0.002 in all the cases.

Table 6.9: Singe Fork-Join Throughput Results

Configuration	Simulation (S)	Analytical (A)	Error
1	0.0791	0.0777	-0.0014
2	0.0777	0.0762	-0.0014
3	0.0749	0.0735	-0.0015
4	0.0420	0.0401	-0.0019
5	0.0408	0.0389	-0.0019
6	0.0389	0.0371	-0.0018
7	0.0627	0.0619	-0.0008
8	0.0620	0.0611	-0.0009
9	0.0601	0.0592	-0.0009
10	0.0486	0.0463	-0.0023
11	0.0419	0.0405	-0.0014
12	0.0400	0.0386	-0.0014
13	0.0425	0.0414	-0.0011
14	0.0415	0.0403	-0.0011
15	0.0397	0.0386	-0.0011

\* S= Simulation, A= Analytical

The utilization metric is well predicted by the analytical model as well, as shown in Table 6.10. The error percentage is small in most of the cases.



Table 6.10: Single Fork-Join Utilization Results

Conf.	Node 1			Node 2			Node 3			Node 4			Node 5		
	S*	A*	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
1	79.10	77.7	-1.77%	59.3	58.28	-1.73%	11.9	11.66	-2.06%	79.10	77.70	-1.77%	65.5	64.75	-1.14%
2	77.20	76.21	-1.29%	58.2	57.15	-1.8%	11.6	11.43	-1.46%	77.6	76.21	-1.80%	64.30	63.5	-1.24%
3	74.80	73.47	-1.78%	56.2	55.1	-1.95%	11.2	11.02	-1.60%	74.90	73.47	-1.91%	62.10	61.23	-1.41%
4	75.60	72.22	-4.48%	52.2	50.15	-3.93%	12.6	12.04	-4.48%	42.00	40.12	-4.48%	34.80	33.43	-3.93%
5	73.10	70.04	-4.19%	50.8	48.64	-4.26%	12.20	11.67	-4.32%	40.90	38.91	-4.87%	33.80	32.42	-4.07%
6	69.80	66.72	-4.42%	48.3	46.33	-4.08%	11.60	11.12	-4.14%	38.90	37.06	-4.72%	32.20	30.89	-4.08%
7	81.60	80.51	-1.34%	73.0	72.25	-1.03%	9.40	9.29	-1.18%	62.70	61.93	-1.23%	52.00	51.61	-0.75%
8	80.10	79.42	-0.85%	72.1	71.27	-1.14%	9.20	9.16	-0.39%	62.00	61.09	-1.46%	51.40	50.91	-0.95%
9	77.90	76.92	-1.25%	70.0	69.03	-1.38%	9.00	8.88	-1.38%	60.01	59.17	-1.54%	49.80	49.31	-0.98%
10	72.90	69.48	-4.69%	48.6	46.32	-4.69%	14.60	13.9	-4.82%	48.60	46.32	-4.69%	40.30	38.6	-4.21%
11	62.50	60.78	-2.75%	41.9	40.52	-3.29%	12.50	12.16	-2.75%	41.90	40.52	-3.29%	34.7	33.77	-2.69%
12	59.80	57.97	-3.07%	40.0	38.64	-3.39%	12.00	11.59	-3.39%	40.00	38.64	-3.39%	33.10	32.20	-2.71%
13	63.70	62.03	-2.63%	56.5	55.13	-2.42%	12.80	12.41	-3.08%	42.50	41.35	-2.70%	35.2	34.46	-2.11%
14	61.80	60.47	-2.15%	55.2	53.75	-2.62%	12.40	12.09	-2.47%	41.50	40.31	-2.86%	34.40	33.60	-2.34%
15	59.40	57.85	-2.61%	52.8	51.42	-2.61%	11.90	11.57	-2.77%	39.70	38.57	-2.86%	32.90	32.14	-2.31%

\* S= Simulation, A= Analytical

The mean number of customers per node is an important result for this algorithm. The results for the mean number of customers per node confirm the algorithm’s performance prediction accuracy as can be seen in Table 6.11

Table 6.11: Single Fork-Join Average Number in Node Results

Conf.	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
1	2.67	2.21	-2.31%	3.67	3.58	-0.46%	2.38	2.33	-0.24%	2.62	2.12	-2.5%	2.53	2.41	-0.61%
2	2.96	2.57	-1.94%	3.66	3.54	-0.59%	2.32	2.29	-0.17%	2.56	2.20	-1.83%	2.47	2.37	-0.48%
3	3.47	3.08	-1.95%	3.63	3.44	-0.92%	2.24	2.20	-0.18%	2.47	2.24	-1.15%	2.36	2.31	-0.23%
4	1.85	1.59	-2.58%	3.17	3.03	-1.39%	1.26	1.20	-0.56%	0.62	0.53	-0.85%	1.07	1.03	-0.43%
5	1.99	1.85	-1.46%	3.1	2.95	-1.47%	1.22	1.17	-0.53%	0.60	0.54	-0.55%	1.04	1.01	-0.36%
6	2.22	2.15	-0.65%	2.96	2.82	-1.39%	1.16	1.11	-0.48%	0.57	0.54	-0.29%	0.99	0.97	-0.22%
7	3.23	2.76	-2.38%	3.82	3.57	-1.25%	1.88	1.86	-0.11%	1.5	1.34	-0.81%	1.80	1.79	-0.03%
8	3.61	3.23	-1.90%	3.88	3.66	-1.13%	1.84	1.83	-0.04%	1.49	1.35	-0.69%	1.77	1.78	0.05%
9	4.18	3.81	-1.82%	3.92	3.71	-1.06%	1.80	1.78	-0.12%	1.45	1.34	-0.53%	1.72	1.74	0.15%
10	1.65	1.41	-2.45%	2.93	2.79	-1.41%	1.46	1.39	-0.7%	0.77	0.64	-1.25%	1.26	1.20	-0.6%
11	1.43	1.29	-1.42%	2.53	2.44	-0.88%	1.25	1.22	-0.34%	0.63	0.57	-0.58%	1.08	1.05	-0.26%
12	1.62	1.52	-1.01%	2.42	2.33	-0.93%	1.20	1.16	-0.41%	0.6	0.57	-0.27%	1.03	1.02	-0.11%
13	1.34	1.20	-1.32%	2.07	1.97	-0.96%	1.28	1.24	-0.39%	0.65	0.60	-0.53%	1.10	1.08	-0.18%
14	1.44	1.35	-0.98%	2.04	1.95	-0.94%	1.24	1.21	-0.31%	0.64	0.59	-0.43%	1.07	1.06	-0.13%
15	1.64	1.54	-0.96%	1.98	1.89	-0.83%	1.19	1.16	-0.33%	0.61	0.58	-0.3%	1.02	1.02	-0.06%

\* S= Simulation, A= Analytical

The analytical prediction related with the average number in node for the single fork/join system is accurate. Table 6.11 shows error values near 3% in the worst cases.

The final performance metric used to demonstrate the algorithm's accuracy is the mean number of customers in the fork/join node. In order to evaluate it in perspective, this prediction error was computed as an proportion of the total number of customers allowed to be in the system following the method used in the literature [32]. In Table 6.12 the comparison between the analytical and simulation results is presented.

Table 6.12: Single FJ Number in Queue for Fork-Join Node

Configuration	Initial Fork-Join Node		
	S	A	Error
1	0.70	3.21	-12.56%
2	0.82	3.45	-13.15%
3	1.18	3.69	-12.55%
4	1.24	2.63	-13.85%
5	1.33	2.76	-14.22%
6	1.50	2.84	-13.37%
7	0.68	2.32	-8.23%
8	0.85	2.66	-9.02%
9	1.24	3.02	-8.89%
10	1.27	2.67	-14.02%
11	0.49	1.89	-13.92%
12	0.61	1.91	-12.98%
13	0.45	1.54	-10.85%
14	0.48	1.58	-10.90%
15	0.60	1.60	-10.05%

\* S= Simulation, A= Analytical

In summary, this chapter presented an implementation of Krishnamurthy and Suri's [32] fork/join system with modifications that included multi-server nodes and a single fork/join node to model situations commonly found in service systems.

## CHAPTER VII

### LINKING TWO STAGES WITH INDIVIDUAL CAPACITY RESTRICTIONS

The previous chapter focused on the fork/join approach proposed by Krishnamurthy and Suri [32] in order to implement synchronization nodes to handle systems with multiple arrival sources that need to be paired. In the previous chapter, some modifications were introduced to the original algorithm in order to include multi-server nodes. A modification was also suggested that shows the removal of a synchronization node at the end of the system to model how hospital systems and other service systems operate. In this chapter the fork/join approaches' capability to work with multiple stages, each with its own capacity-restricting passive resources will be developed and tested.

Krishnamurthy and Suri [32] studied only a single-stage, double fork/join system, which they correctly portrayed as a key building block for multi-stage kanban systems. However, they presented only a brief conceptual discussion on how the single-stage models could be linked together to model larger multistage systems. There has been only one study that has explored a solution approach in the context of a multi-product, multi-stage pull system [43, 44]. The iterative algorithm with forward and backward passes was specifically developed for a pull-based manufacturing system and works only with a demand process pulling items from a finished goods store. As explained in the introductory chapters of this document, a patient's pathway could span multiple, capacity-restricted stages such as the Emergency Department, ICU, Surgery and Recovery. When we consider multiple, capacity-restricted stages in a hospital,, the system functions more like a push system, but still needs the FJ

station between two capacity-restricted stages to ensure that a passive resource is available in a downstream stage before a patient can be transferred from the immediate upstream stage. No FJ station is needed at the end of the last stage.

A system with two stages is studied in this chapter. The first stage has two synchronization stations, one at the beginning and another at the end. The second synchronization node in the first stage is shared by the second stage, where this stage starts. Figure 7.1 shows an example of the system under study. The second FJ node ensures that a Resource B is available before a customer is allowed into the second stage. After service in stage 2, a customer can leave the system and release Resource B.

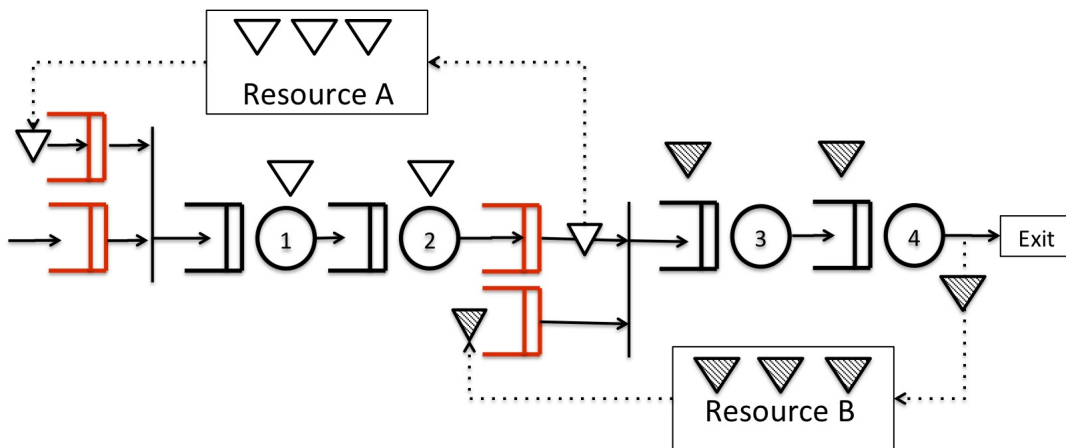


Figure 7.1: Two-stage Fork/Join Example

## 7.1 The Solution Approach

The solution to this problem is an extension of the approach presented in Chapter VI. Given there is an external arrival to this system, it is important to verify the system stability. An external arrival rate that surpasses the maximum throughput the system is capable of will lead to an unstable system. Consequently, cases where the arrival rate is bigger than the

system's maximum throughput rate are not in the scope of this work. Figure 7.2 helps in understanding this condition.

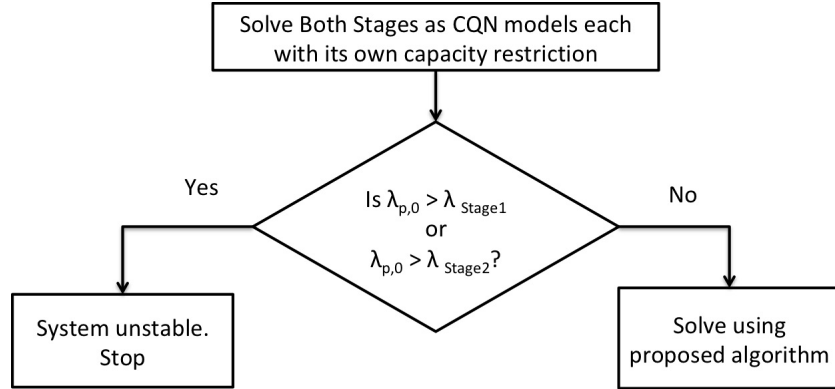


Figure 7.2: System's Stability Check

Once the stability condition has been checked, the algorithm 6 is used to solve the two-stage system. The basic idea to solve the second stage as CQN to get an estimate of the arrival of free resources to the second FJ node. This would allow us to solve the first stage as a double FJ. Finally, we solve the second stage as a single FJ system.

The initial attempts to model the two-stage system followed the logic seen so far for the approximation methodologies, where an iterative approach is followed until convergence is obtained in some carefully selected metric based on some threshold value. However, in solving the two-stage system, it was seen that pursuing the open-ended iteration between the two stages was a difficult process.

The natural information shared between the two stages are the arrival rates (customers or free kanbans) and these values were used as the basis for iterating. However, in this process, they became an upper limit for the other system throughput computation. After

a few attempts, the results revealed that this iteration process made the computation of the systems throughput decrease monotonically. Additionally, it was observed that in the first few iterations the throughput decreased in big steps and went below the real system throughput by a big margin, making the approach to systematically underestimate the system performance measures.

Another observation was that after computing the CQN throughput value in the second stage, and using it as the kanban arrival rate to the second FJ node in the first stage, the first stage throughput computation was very accurate (in the first iteration), and from there, the computation of the second stage performance measures were as well. This approach was used in this section for computing the two-stage system performance measures with the algorithm presented next.

---

**Algorithm 6** Solution for the Two-Stage Fork/join System

---

**Require:**  $\lambda_{P,0}$ ,  $c_{a,P,0}^2$ ,  $K_0$ ,  $K_1$ ,  $K_2$ , and the parameter tuple  $\{\tau, c_s^2, m\}$  for all the nodes in both stages.

**Ensure:**  $\epsilon = 0.01$ ,  $\delta = 100$ .

**Step 1**

Solve stage 2 as CQN with approximation in Chapter V

Set  $\lambda_{F,2}$ ,  $c_{a,F,2}^2$  as  $\lambda_{CQN}$  and  $c_{CQN}^2$

**Step 2**

Solve the first stage using the Double fork/join algorithm in Chapter VI.

**Step 3**

Use  $\lambda_{P,1}$ ,  $c_{a,P,1}^2$  and  $K_1$  as previous stage information to solve the Second-Stage System.

**Step 4**

Solve the second stage using the Single fork/join algorithm in Chapter VI.

---

## 7.2 Numerical Experiments for the Two-Stage System

Based on the experiments used in previous chapters, the two-stage configurations were determined in order to test the solution approach. The configurations in Chapter V were used as stages in this chapter. The number of passive resources for each stage and the order in which the Chapter V's configuration were selected as stages were changed to yield different configurations.

The arrival rate parameters were selected based on 85% of the slower CQN throughput obtained using the approach in Chapter V for the two stages in each configuration. This arrival process parameters are the following.



Table 7.1: Arrival Process Parameters for the Two-stage Configurations

Configuration	$\lambda_{P,0}$	$c_{P,0}^2$
1	0.079	1.0
2	0.077	1.0
3	0.075	1.0
4	0.042	1.0
5	0.041	1.0
6	0.039	1.0
7	0.062	1.0
8	0.062	1.0
9	0.060	1.0
10	0.048	1.0
11	0.047	1.0
12	0.045	1.0
13	0.048	1.0
14	0.047	1.0
15	0.045	1.0

Similarly, the node parameters chosen for the first and second stages can be seen in tables 7.2 and 7.3. As explained earlier, each stage contains single and multi server nodes, and one delay node, thus, testing a wide variety of possible node characteristics in order to test the approach thoroughly with each configuration.

Table 7.2: Two-stage Experiments: Node Parameters for the First Stage

Conf.	$K_1$	$\tau_{1,1}$	$c_{s,1,1}^2$	$m_{1,1}$	$\tau_{1,2}$	$c_{s,1,2}^2$	$m_{1,2}$	$\tau_{1,3}$	$c_{s,1,3}^2$	$m_{1,3}$	$\tau_{1,4}$	$c_{s,1,4}^2$	$m_{1,4}$	$\tau_{1,5}$	$c_{s,1,5}^2$	$m_{1,5}$
1	20.0	10.0	0.5	1.0	45.0	0.5	6.0	30.0	0.5	20.0	10.0	0.5	1.0	25.0	1.0	3.0
2	20.0	10.0	1.0	1.0	45.0	0.5	6.0	30.0	1.0	20.0	10.0	0.5	1.0	25.0	1.0	3.0
3	15.0	10.0	2.0	1.0	45.0	0.5	6.0	30.0	2.0	15.0	10.0	0.5	1.0	25.0	1.0	3.0
4	10.0	18.0	0.5	1.0	75.0	1.0	6.0	30.0	0.5	10.0	10.0	0.5	1.0	25.0	1.0	3.0
5	10.0	18.0	1.0	1.0	75.0	1.0	6.0	30.0	1.0	10.0	10.0	0.5	1.0	25.0	1.0	3.0
6	12.0	18.0	2.0	1.0	75.0	1.0	6.0	30.0	2.0	12.0	10.0	0.5	1.0	25.0	1.0	3.0
7	20.0	13.0	0.5	1.0	35.0	2.0	3.0	30.0	0.5	20.0	10.0	0.5	1.0	25.0	1.0	3.0
8	20.0	13.0	1.0	1.0	35.0	2.0	3.0	30.0	1.0	20.0	10.0	0.5	1.0	25.0	1.0	3.0
9	15.0	13.0	2.0	1.0	35.0	2.0	3.0	30.0	2.0	15.0	10.0	0.5	1.0	25.0	1.0	3.0
10	10.0	15.0	0.5	1.0	60.0	0.5	6.0	30.0	0.5	10.0	10.0	0.5	1.0	25.0	1.0	3.0
11	10.0	15.0	1.0	1.0	60.0	0.5	6.0	30.0	1.0	10.0	10.0	0.5	1.0	25.0	1.0	3.0
12	12.0	15.0	2.0	1.0	60.0	0.5	6.0	30.0	2.0	12.0	10.0	0.5	1.0	25.0	1.0	3.0
13	10.0	15.0	0.5	1.0	40.0	2.0	3.0	30.0	0.5	10.0	10.0	0.5	1.0	25.0	1.0	3.0
14	10.0	15.0	1.0	1.0	40.0	2.0	3.0	30.0	1.0	10.0	10.0	0.5	1.0	25.0	1.0	3.0
15	8.0	15.0	2.0	1.0	40.0	2.0	3.0	30.0	2.0	8.0	10.0	0.5	1.0	25.0	1.0	3.0

Table 7.3: Two-Stage Experiments: Node Parameters for the Second Stage

Conf.	$K_2$	$\tau_{2,1}$	$c_{s,2,1}^2$	$m_{2,1}$	$\tau_{2,2}$	$c_{s,2,2}^2$	$m_{2,2}$	$\tau_{2,3}$	$c_{s,2,3}^2$	$m_{2,3}$	$\tau_{2,4}$	$c_{s,2,4}^2$	$m_{2,4}$	$\tau_{2,5}$	$c_{s,2,5}^2$	$m_{2,5}$
1	20.0	10.0	0.5	1.0	45.0	0.5	6.0	30.0	0.5	20.0	10.0	0.5	1.0	25.0	1.0	3.0
2	15.0	10.0	1.0	1.0	45.0	0.5	6.0	30.0	1.0	15.0	10.0	0.5	1.0	25.0	1.0	3.0
3	20.0	10.0	2.0	1.0	45.0	0.5	6.0	30.0	2.0	20.0	10.0	0.5	1.0	25.0	1.0	3.0
4	10.0	18.0	0.5	1.0	75.0	1.0	6.0	30.0	0.5	10.0	10.0	0.5	1.0	25.0	1.0	3.0
5	12.0	18.0	1.0	1.0	75.0	1.0	6.0	30.0	1.0	12.0	10.0	0.5	1.0	25.0	1.0	3.0
6	10.0	18.0	2.0	1.0	75.0	1.0	6.0	30.0	2.0	10.0	10.0	0.5	1.0	25.0	1.0	3.0
7	20.0	13.0	0.5	1.0	35.0	2.0	3.0	30.0	0.5	20.0	10.0	0.5	1.0	25.0	1.0	3.0
8	15.0	13.0	1.0	1.0	35.0	2.0	3.0	30.0	1.0	15.0	10.0	0.5	1.0	25.0	1.0	3.0
9	20.0	13.0	2.0	1.0	35.0	2.0	3.0	30.0	2.0	20.0	10.0	0.5	1.0	25.0	1.0	3.0
10	10.0	15.0	0.5	1.0	40.0	2.0	3.0	30.0	0.5	10.0	10.0	0.5	1.0	25.0	1.0	3.0
11	12.0	15.0	1.0	1.0	40.0	2.0	3.0	30.0	1.0	12.0	10.0	0.5	1.0	25.0	1.0	3.0
12	10.0	15.0	2.0	1.0	40.0	2.0	3.0	30.0	2.0	10.0	10.0	0.5	1.0	25.0	1.0	3.0
13	10.0	15.0	0.5	1.0	60.0	0.5	6.0	30.0	0.5	10.0	10.0	0.5	1.0	25.0	1.0	3.0
14	8.0	15.0	1.0	1.0	60.0	0.5	6.0	30.0	1.0	8.0	10.0	0.5	1.0	25.0	1.0	3.0
15	10.0	15.0	2.0	1.0	60.0	0.5	6.0	30.0	2.0	10.0	10.0	0.5	1.0	25.0	1.0	3.0

As can be seen from tables 7.2 and 7.3 there are homogeneous configurations such as configuration 1 where both stages are identical. There are configurations where each configuration has the same values in the node parameters but different number of capacity restricting resources, such as configurations 2 and 3. Moreover, there are some configurations, which more heterogeneous, where not only the number of passive resources change, but some of the node parameters as well. For instance, configurations 12 and 13.

In some configurations, the stage constraining the overall throughput is the first, and in some cases it is the second stage. This set of configurations provided a comprehensive set of numerical test cases.

### 7.2.1 Numerical results for the Two-Stage System

The system throughput, node utilization, and average number in node values for both stages are presented to study the accuracy of the analytical results. The number in node in the synchronization stations is presented also. The results were compared with simulation results and error estimations were computed following the methods presented in Chapter V. It is important to keep in mind that the error in the average number of customers per node has been normalized with the number of passive resources in the stage, for all nodes including the SRP. This approach was used in Krishnamurthy and Suri [32].

Table 7.4 presents the analytical and simulation results for the throughput. In general, the first stage throughput estimates are better than the second stage ones. However, the difference between the simulation and the analytical values is less than 0.005. The throughput values for both stages should be the same. Small differences are observed in the analytical estimates.

Table 7.4: Throughput Results for the Two-Stage System

Configuration	Stage 1			Stage 2		
	Simulation (S)	Analytical (A)	Error	Simulation (S)	Analytical (A)	Error
1	0.079	0.078	-0.001	0.079	0.077	-0.002
2	0.077	0.076	-0.002	0.077	0.073	-0.004
3	0.074	0.071	-0.003	0.074	0.07	-0.004
4	0.042	0.040	-0.002	0.042	0.038	-0.003
5	0.041	0.039	-0.002	0.041	0.038	-0.003
6	0.039	0.038	-0.001	0.039	0.036	-0.003
7	0.063	0.062	-0.001	0.063	0.061	-0.001
8	0.062	0.061	-0.001	0.062	0.06	-0.002
9	0.059	0.058	-0.001	0.059	0.057	-0.002
10	0.048	0.045	-0.002	0.048	0.044	-0.004
11	0.047	0.044	-0.002	0.047	0.044	-0.003
12	0.045	0.043	-0.001	0.045	0.042	-0.003
13	0.048	0.046	-0.002	0.048	0.044	-0.003
14	0.045	0.043	-0.002	0.045	0.04	-0.005
15	0.042	0.040	-0.002	0.042	0.039	-0.004

The throughput values for this system must be very close to the external arrival rate of customers. However, as explained in Chapter VI, Krhishnamurthy and Suri [32] approach requires the specification of a  $K$  value for external customer queue as they assume that these arrivals also happen from another capacity-restricted system. Setting  $K$  to be large is approximately equivalent to not having a limit on the external customer queue size at the first synchronization node. In our experiments we set this value to be equal to the first level population constraint following examples in [32]. As a result, there could be a small fraction of external arrivals that may be lost if this queue becomes full.

An interesting aspect to note is that the test configurations presented in Chapter VI

basically use exponential arrivals, In the two-stage cases, since the arrivals to stage 1 are departures from stage 2, they have different variability characteristics.

In tables 7.5 and 7.6 results for the utilization metric are shown. It can be seen the algorithm's ability to predict utilization is good, despite the modifications to include multi-servers nodes and extension to multiple stages.

Table 7.5: Utilization Results for Stage 1

Conf.	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
2	76.9	75.54	-1.77%	58.0	56.66	-2.32%	11.5	11.33	-1.47%	77.4	75.54	-2.40%	64.1	62.95	-1.79%
3	73.5	71.04	-3.34%	55.3	53.28	-3.65%	14.7	14.21	-3.34%	73.7	71.04	-3.60%	61.0	59.2	-2.95%
4	75.2	71.52	-4.89%	51.9	49.67	-4.3%	12.5	11.92	-4.64%	41.8	39.73	-4.94%	34.6	33.11	-4.30%
5	72.8	69.88	-4.01%	50.6	48.53	-4.09%	12.1	11.65	-3.74%	40.7	38.82	-4.61%	33.7	32.35	-4.00%
6	69.8	67.93	-2.68%	48.3	47.17	-2.34%	9.7	9.43	-2.74%	38.9	37.74	-2.99%	32.2	31.45	-2.34%
7	81.5	80.51	-1.21%	73.0	72.26	-1.02%	9.4	9.29	-1.17%	62.7	61.93	-1.22%	51.9	51.61	-0.56%
8	79.9	79.2	-0.88%	71.9	71.07	-1.15%	9.2	9.14	-0.67%	61.8	60.92	-1.42%	51.2	50.77	-0.85%
9	76.7	75.16	-2.01%	68.9	67.45	-2.11%	11.8	11.56	-2.01%	59.1	57.81	-2.18%	49.0	48.18	-1.68%
10	71.3	68.2	-4.34%	47.6	45.47	-4.48%	14.3	13.64	-4.61%	47.6	45.47	-4.48%	39.4	37.89	-3.83%
11	69.3	66.7	-3.76%	46.5	44.47	-4.38%	13.9	13.34	-4.03%	46.5	44.47	-4.38%	38.5	37.05	-3.76%
12	66.7	65.14	-2.34%	44.5	43.43	-2.41%	11.1	10.86	-2.19%	44.5	43.43	-2.41%	36.9	36.19	-1.92%
13	71.2	68.45	-3.86%	63.2	60.84	-3.73%	14.2	13.69	-3.59%	47.5	45.63	-3.93%	39.4	38.03	-3.48%
14	67.1	64.23	-4.28%	59.9	57.09	-4.69%	13.4	12.85	-4.14%	45.0	42.82	-4.85%	37.3	35.68	-4.34%
15	63.2	60.49	-4.29%	56.2	53.77	-4.33%	15.8	15.12	-4.29%	42.3	40.33	-4.67%	35.0	33.61	-3.98%

\* S= Simulation, A= Analytical

Table 7.6: Utilization Results for Stage 2

Conf.	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
1	79.1	76.87	-2.81%	59.3	57.66	-2.77%	11.9	11.53	-3.1%	79.1	76.87	-2.81%	65.5	64.06	-2.2%
2	76.9	73.39	-4.57%	58.0	55.04	-5.1%	15.4	14.68	-4.69%	77.4	73.39	-5.19%	64.1	61.15	-4.59%
3	73.5	69.98	-4.79%	55.3	52.49	-5.09%	11.0	10.5	-4.57%	73.7	69.98	-5.05%	61.0	58.32	-4.40%
4	75.1	69.07	-8.03%	51.9	47.96	-7.58%	12.5	11.51	-7.91%	41.7	38.37	-7.98%	34.6	31.98	-7.58%
5	72.8	68.27	-6.22%	50.6	47.41	-6.31%	10.1	9.48	-6.12%	40.7	37.93	-6.81%	33.7	31.61	-6.21%
6	69.8	65.16	-6.65%	48.3	45.25	-6.32%	11.6	10.86	-6.38%	38.9	36.2	-6.94%	32.2	30.17	-6.32%
7	81.5	79.93	-1.92%	73.0	71.74	-1.73%	9.4	9.22	-1.88%	62.7	61.49	-1.93%	51.9	51.24	-1.27%
8	79.8	77.6	-2.76%	71.9	69.64	-3.14%	12.3	11.94	-2.94%	61.8	59.69	-3.41%	51.2	49.74	-2.85%
9	76.7	74.17	-3.3%	68.8	66.56	-3.25%	8.9	8.56	-3.84%	59.2	57.06	-3.62%	49.0	47.55	-2.97%
10	71.3	66.09	-7.31%	63.3	58.74	-7.2%	14.3	13.22	-7.57%	47.5	44.06	-7.25%	39.4	36.71	-6.82%
11	69.3	65.32	-5.74%	61.9	58.06	-6.2%	11.6	10.89	-6.15%	46.5	43.55	-6.35%	38.5	36.29	-5.74%
12	66.6	62.87	-5.6%	59.2	55.88	-5.6%	13.3	12.57	-5.46%	44.5	41.91	-5.81%	36.9	34.93	-5.35%
13	71.2	66.03	-7.26%	47.5	44.02	-7.33%	14.3	13.21	-7.65%	47.5	44.02	-7.33%	39.3	36.68	-6.66%
14	67.1	60.05	-10.5%	45.0	40.04	-11.03%	16.8	15.01	-10.63%	45.0	40.04	-11.03%	37.3	33.36	-10.55%
15	63.3	57.84	-8.62%	42.3	38.56	-8.84%	12.7	11.57	-8.91%	42.3	38.56	-8.84%	35.0	32.13	-8.19%

\* S= Simulation, A= Analytical

As can be expected, configurations where stages have more heterogeneity, have a tendency to show higher error levels as in configurations 14 and 15. On the other hand, configurations such as 1 and 2, with more homogeneous parameters tend to produce better results.

The results for the average number of customers at the nodes are shown in tables 7.7 and 7.8.

Table 7.7: Average Number in Node Results for Stage 1

Conf.	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
1	2.64	2.21	-2.16%	3.67	3.58	-0.46%	2.38	2.33	-0.25%	2.61	2.12	-2.45%	2.53	2.41	-0.6%
2	2.79	2.5	-1.47%	3.63	3.51	-0.63%	2.30	2.27	-0.17%	2.45	2.13	-1.57%	2.44	2.34	-0.49%
3	2.82	2.63	-1.30%	3.48	3.29	-1.27%	2.21	2.13	-0.49%	2.05	1.95	-0.69%	2.22	2.16	-0.39%
4	1.76	1.55	-2.11%	3.15	3.00	-1.48%	1.25	1.19	-0.58%	0.61	0.52	-0.86%	1.07	1.02	-0.46%
5	1.96	1.84	-1.19%	3.08	2.94	-1.39%	1.21	1.16	-0.45%	0.59	0.54	-0.52%	1.04	1.01	-0.35%
6	2.43	2.36	-0.65%	2.99	2.88	-0.87%	1.16	1.13	-0.27%	0.58	0.56	-0.16%	1.00	1.00	-0.02%
7	3.19	2.76	-2.15%	3.81	3.57	-1.21%	1.88	1.86	-0.11%	1.5	1.34	-0.78%	1.79	1.79	0.0%
8	3.45	3.2	-1.27%	3.82	3.63	-0.96%	1.84	1.83	-0.06%	1.46	1.34	-0.61%	1.76	1.78	0.08%
9	3.46	3.30	-1.04%	3.46	3.35	-0.72%	1.77	1.73	-0.24%	1.32	1.24	-0.52%	1.66	1.67	0.07%
10	1.56	1.34	-2.14%	2.87	2.74	-1.31%	1.43	1.36	-0.66%	0.74	0.62	-1.17%	1.23	1.18	-0.54%
11	1.73	1.60	-1.36%	2.81	2.68	-1.28%	1.39	1.33	-0.56%	0.73	0.65	-0.72%	1.2	1.16	-0.41%
12	2.17	2.07	-0.85%	2.72	2.63	-0.76%	1.33	1.3	-0.24%	0.71	0.69	-0.2%	1.16	1.16	0.02%
13	1.61	1.52	-0.93%	2.43	2.31	-1.2%	1.42	1.37	-0.51%	0.77	0.70	-0.68%	1.24	1.21	-0.36%
14	1.57	1.55	-0.27%	2.25	2.14	-1.18%	1.34	1.28	-0.55%	0.71	0.65	-0.55%	1.17	1.13	-0.35%
15	1.64	1.62	-0.24%	2.03	1.97	-0.74%	1.26	1.21	-0.68%	0.64	0.61	-0.36%	1.08	1.06	-0.29%

\* S= Simulation, A= Analytical



Table 7.8: Average Number in Node Results for Stage 2

Conf.	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
1	2.54	2.13	-2.05%	3.67	3.54	-0.64%	2.38	2.31	-0.37%	2.56	2.04	-2.62%	2.52	2.37	-0.78%
2	2.47	2.18	-1.93%	3.6	3.38	-1.47%	2.31	2.20	-0.72%	2.21	1.88	-2.17%	2.37	2.2	-1.13%
3	3.21	2.64	-2.85%	3.54	3.25	-1.44%	2.20	2.10	-0.50%	2.28	1.92	-1.80%	2.28	2.15	-0.65%
4	1.75	1.41	-3.34%	3.15	2.90	-2.54%	1.25	1.15	-0.99%	0.61	0.50	-1.09%	1.07	0.98	-0.83%
5	2.13	1.78	-2.93%	3.1	2.88	-1.89%	1.21	1.14	-0.62%	0.60	0.53	-0.64%	1.04	0.98	-0.50%
6	2.23	2.03	-1.99%	2.96	2.75	-2.10%	1.16	1.09	-0.74%	0.57	0.52	-0.48%	0.99	0.95	-0.47%
7	3.25	2.68	-2.87%	3.83	3.5	-1.66%	1.88	1.84	-0.18%	1.51	1.32	-0.96%	1.79	1.77	-0.10%
8	3.17	2.81	-2.37%	3.56	3.32	-1.60%	1.84	1.79	-0.36%	1.41	1.25	-1.06%	1.75	1.71	-0.25%
9	4.02	3.36	-3.32%	3.78	3.36	-2.08%	1.78	1.71	-0.34%	1.40	1.23	-0.82%	1.68	1.66	-0.08%
10	1.59	1.39	-2.00%	2.43	2.18	-2.50%	1.43	1.32	-1.08%	0.77	0.66	-1.06%	1.24	1.16	-0.82%
11	1.88	1.65	-1.88%	2.5	2.24	-2.19%	1.39	1.31	-0.71%	0.77	0.68	-0.80%	1.22	1.16	-0.50%
12	2.07	1.87	-2.04%	2.35	2.16	-1.89%	1.33	1.26	-0.73%	0.73	0.66	-0.68%	1.16	1.12	-0.46%
13	1.56	1.24	-3.16%	2.86	2.65	-2.14%	1.43	1.32	-1.09%	0.74	0.59	-1.47%	1.23	1.14	-0.90%
14	1.45	1.22	-2.85%	2.71	2.41	-3.73%	1.34	1.20	-1.79%	0.66	0.55	-1.40%	1.15	1.03	-1.49%
15	1.78	1.51	-2.70%	2.56	2.32	-2.39%	1.27	1.16	-1.13%	0.64	0.57	-0.76%	1.09	1.01	-0.75%

\* S= Simulation, A= Analytical

The results show acceptable error levels. Most of the error values are well below 10% with a few exceptions for multi-server nodes in highly heterogeneous configurations.

Table 7.9: Average Number of Customers in Queue at the Fork-Join Nodes

Conf.	Initial Fork-Join Node			Final Fork-Join Node		
	S	A	Error	S	A	Error
1	0.89	3.37	-12.37%	0.36	3.02	-13.27%
2	2.46	4.97	-12.52%	2.19	6.11	-19.63%
3	3.55	5.13	-10.53%	0.37	2.13	-11.75%
4	1.78	3.02	-12.41%	0.51	2.21	-16.93%
5	1.59	2.90	-13.06%	0.26	1.52	-12.64%
6	1.99	3.06	-8.94%	1.29	3.05	-14.61%
7	0.94	2.43	-7.45%	0.57	2.20	-8.15%
8	2.04	3.34	-6.47%	2.07	4.73	-13.32%
9	3.07	4.12	-7.02%	0.62	1.78	-7.73%
10	1.74	2.88	-11.39%	0.51	2.01	-15.01%
11	1.53	2.80	-12.73%	0.28	1.37	-10.95%
12	1.79	2.73	-7.85%	1.21	2.74	-12.75%
13	1.70	2.73	-10.31%	0.60	2.21	-16.13%
14	3.35	3.95	-6.00%	1.94	3.34	-14.03%
15	2.65	3.10	-5.58%	0.47	1.53	-13.30%

\* S= Simulation, A= Analytical

The results for the number of customers waiting for a passive resource to become available is presented in Table 7.9 for the available two fork/join nodes. Because of the complex nature of approximations needed for the synchronization nodes, the prediction accuracy is not expected to be as good as that for the other metrics.

## CHAPTER VIII

### SOLVING CLOSED QUEUEING NETWORKS WITH SRP

The previous chapters explained the different approaches for solving each of the building blocks models. This chapter goes one step forward in developing a solution approach for systems that combine two of the building block. We consider a CQN system that contains a SRP subsystem, and the overall system performance needs to be evaluated.

#### 8.1 Closed Queueing Network With Simultaneous Resource Possession

The CQN with SRP is one of the capacity-restricted problems studied in this research. For simplicity we consider a CQN system that includes one node affected by an SRP subsystem. This mean, that the servers in the first-level node affected by the SRP need service from another server in a second level to complete the service to the original customer. Figure 8.1 presents an example of such a system.

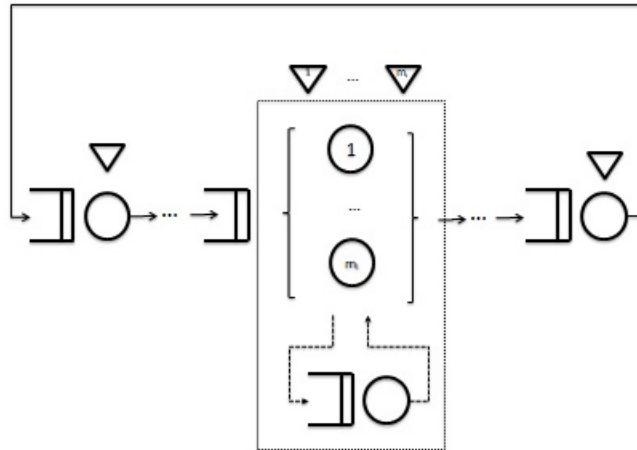


Figure 8.1: CQN With SRP Example

As mentioned previously, the CQN approach can be really helpful for solving systems with external arrivals of customers, where the capacity restricting resources work under high utilization conditions. Additional notation needed for the CQN-SRP model is shown in Table 8.1.

Table 8.1: Additional Notation for the CQN-SRP Model

---

$K$	Number of passive resources.
$N_{1L}$	Number of nodes in first-level.
$p$	Probability that the First-level node require the SRP Node service.
$\tau_S$	Mean service time in node $S$ .
$\mu_S$	Service rate in node $S$ . ( $\mu_S = \tau_S^{-1}$ ).
$c_{s,S}^2$	SCV of service time in node $S$ .
$m_S$	Number of servers in node $S$ .
$\tau_{SRP}$	Mean service time of the SRP server.
$\mu_{SRP}$	Service rate of the SRP server ( $\mu_{SRP} = \tau_{SRP}^{-1}$ ).
$c_{s,SRP}^2$	service time SCV in the SRP server.
$\sigma_S^2$	service time variance in node $S$ .
$\sigma_{SRP}^2$	service time variance in the SRP server.
$c_{a,S}^2$	SCV of inter-arrival time at node $S$ .
$c_{d,S}^2$	SCV of inter-departure time at node $S$ .

---

## 8.2 The Solution Approach for CQN with SRP

The general idea for solving this system was developed in Chapter IV. Basically, with the first-level throughput estimation, the SRP subsystem response time is computed and used to calculate the mean and SCV of an equivalent service time in the first-level node affected by the SRP sub system. Once the equivalent time mean and SCV for the first-level node are computed, the cycle starts again with solving the CQN.

The solution approach presented in this chapter takes advantage of the structure in the PDMCQN-2 Algorithm presented in Chapter V. This algorithm allows us to embed the SRP response time estimation in the heart of the CQN solution and carry on with the CQN

system solution.

The algorithm developed for solving the CQN-SRP problem is presented next. With in each iteration of the CQN solution, we have another iterative loop to compute the SRP response time parameter needed to compute the equivalent service time parameters for the first-level node affected. This inner loop continues till the arrival process SCV to the SRP system converges based on some specified tolerance level.

---

**Algorithm 7** CQN-SRP Case

---

**Require:**  $K$ ,  $\tau_{SRP}$ ,  $c_{s,SRP}^2$ ,  $m_{SRP}$ ,  $p$ , and the parameter's tuple  $\{\tau, c_s^2, m\}$  for all nodes in the system.

**Ensure:**  $W_{qSRP} = 0$ ,  $\sigma_{q,SRP}^2 = 0$ ,  $\epsilon=0.001$ ,  $j=1$ .

Compute  $\tau_r^I$  and  $c_{s,r}^{2I}$  based on Equations 4.5 and 4.7

Solve the first-level system (CQN) including the equivalent node and obtain  $\rho_1$ ,  $c_{a,r}^2$ , and  $\lambda^{(j)}$

Set  $High = \rho_1$ ,  $Low = 0$ ,

**while**  $j < 19$  **do**

$\rho_1^{(j)} = (High + Low)/2$

$\lambda^{(j)} = \rho_1^{(j)} * m_1/\tau_1$

Validate SRP utilization  $\rho_{SRP} = \lambda^{(j)} * p * \tau_{SRP} / m_{SRP}$ . If  $\rho_{SRP} > 0.95$ , adjust  $\lambda^{(j)} = 0.95 * m_{SRP}/(\tau_{SRP} * p)$

$control = 0$

**while**  $control = 0$  **do**

Set  $\delta_1=100$  and  $c_{d,r}^{2,(0)} = 0$

**while**  $\delta_1 > \epsilon$  **do**

$\rho_r^I = \lambda^{(j)} * \tau_r^I / m_r$

Compute  $c_{d,r}^2$ , using equation 4.10.

Set  $c_{d,r}^{2,(j)} = c_{d,r}^2$

Compute  $W_{qSRP}$  and  $\sigma_{q,SRP}^2$ , using the SRP algorithm in Section 4.7

Based on the new  $W_{qSRP}$  and  $\sigma_{q,SRP}^2$  values, compute  $\tau_r^I$  and  $c_{s,r}^{2I}$  using equations 4.5 and 4.7

Compute  $\delta_1 = c_{d,r}^{2,(j)} - c_{d,r}^{2,(j-1)}$ , and set  $c_{d,r}^{2,(j-1)} = c_{d,r}^{2,(j)}$

**end while**

Compute  $\rho = \lambda^{(j)} * \tau / m$  for all nodes in the system, including the new equivalent node. If  $\rho > 0.95$  for any node  $S$ , set  $\lambda^{(j)} = 0.95 * m_S/\tau_S$ , otherwise, set  $control = 1$

**end while**

Compute  $c_{a,S}^2$  by solving traffic variability equations. See algorithm in Chapter V

**for** all nodes  $S$  **do**

Compute  $W_{qS}$  using equations in Section 4.5.

Compute  $L_{qS}$  and  $L_S$  using equations 4.23 and 4.24, respectively.

**end for**

Compute  $\delta_0 = (\sum_{S=1}^N L_S) - K$ , if  $\delta_0 < 0$  state  $low = \rho_1$  otherwise  $high = \rho_1$

Set  $j = j + 1$

**end while**

---

The subscript  $r$  stands for the node in the first-level influenced by the SRP system. The superscript  $I$  refers to the equivalent values,  $(j)$  refers to the iteration number and  $S$  a node in the first-level network.

In order to assess the CQN-SRP algorithm's performance prediction accuracy, a series of numerical experiments were performed as explained in the next section.

### **8.3 CQN with SRP Numerical Experiment**

In this section, a series of configurations are defined to represent conditions representative of real-life situations. These configurations are designed to challenge the algorithm in order to test the robustness of its predictions accuracy. Analytical results are compared with simulation estimates. The parameters for the first level CQN-SRP system are presented next.

#### **8.3.1 Experimental Set Up**

As explained previously, the parameters for the configurations presented in this and other sections of this document, have been selected based on numerical examples from published health care articles for instance [11], [22], and are an attempt to represent possible real health care scenarios.

A total of forty five configurations have been grouped to three sets of fifteen. Each set of fifteen configurations has the same first and second level node parameters' values. Only the requirement for the SRP system is different for each set, 25%, 50% and 75%, respectively.



Table 8.2: CQN with SRP Parameters for the First Level Nodes

Conf	K	$\tau_1$	$c_{s,1}^2$	$m_1$	$\tau_2$	$c_{s,2}^2$	$m_2$	$\tau_3$	$c_{s,3}^2$	$m_3$	$\tau_4$	$c_{s,4}^2$	$m_4$	$\tau_5$	$c_{s,5}^2$	$m_5$
1	20	10.0	0.5	1	45.0	0.5	6	30	0.5	20	10.0	0.5	1	25.0	1.0	3
2	20	10.0	1.0	1	45.0	0.5	6	30	1.0	20	10.0	0.5	1	25.0	1.0	3
3	20	10.0	2.0	1	45.0	0.5	6	30	2.0	20	10.0	0.5	1	25.0	1.0	3
4	10	18.0	0.5	1	75.0	1.0	6	30	0.5	10	10.0	0.5	1	25.0	1.0	3
5	10	18.0	1.0	1	75.0	1.0	6	30	1.0	10	10.0	0.5	1	25.0	1.0	3
6	10	18.0	2.0	1	75.0	1.0	6	30	2.0	10	10.0	0.5	1	25.0	1.0	3
7	20	13.0	0.5	1	35.0	2.0	3	30	0.5	20	10.0	0.5	1	25.0	1.0	3
8	20	13.0	1.0	1	35.0	2.0	3	30	1.0	20	10.0	0.5	1	25.0	1.0	3
9	20	13.0	2.0	1	35.0	2.0	3	30	2.0	20	10.0	0.5	1	25.0	1.0	3
10	10	15.0	0.5	1	60.0	0.5	6	30	0.5	10	10.0	0.5	1	25.0	1.0	3
11	10	15.0	1.0	1	60.0	0.5	6	30	1.0	10	10.0	0.5	1	25.0	1.0	3
12	10	15.0	2.0	1	60.0	0.5	6	30	2.0	10	10.0	0.5	1	25.0	1.0	3
13	10	15.0	0.5	1	40.0	2.0	3	30	0.5	10	10.0	0.5	1	25.0	1.0	3
14	10	15.0	1.0	1	40.0	2.0	3	30	1.0	10	10.0	0.5	1	25.0	1.0	3
15	10	15.0	2.0	1	40.0	2.0	3	30	2.0	10	10.0	0.5	1	25.0	1.0	3

For each set of fifteen configurations, the SRP node parameter values are the following.

Table 8.3: CQN with SRP Parameters for the SRP Node

Configuration	$\tau_{SRP}$	$c_{s,SRP}^2$	$m_{SRP}$
1	15.0	1.0	1
2	15.0	1.0	1
3	15.0	1.0	1
4	15.0	1.0	1
5	15.0	1.0	1
6	15.0	0.5	1
7	15.0	0.5	1
8	15.0	0.5	1
9	15.0	0.5	1
10	15.0	0.5	1
11	15.0	2.0	1
12	15.0	2.0	1
13	15.0	2.0	1
14	15.0	2.0	1
15	15.0	2.0	1

As can be seen, the variability increased for the last few configurations. The solution approach is tested in a variability range that goes for  $c_{s,SRP}^2 = 0.5$  to  $c_{s,SRP}^2 = 2$ .

## 8.4 Numerical Results

For the experimental configurations, the throughput results for these sets of fifteen configurations are presented in tables 8.4, 8.5 and 8.6. The three sets correspond to three different probabilities - 0.25, 0.5, 0.75 - of requiring additional service from the SRP system.

Table 8.4: Throughput Results for  $p = 25\%$

Configuration	First Level System			SRP System		
	Simulation (S)	Analytical (A)	Error	Simulation (S)	Analytical (A)	Error
1	0.092	0.090	-0.002	0.023	0.022	0.000
2	0.090	0.088	-0.002	0.022	0.022	0.000
3	0.087	0.085	-0.002	0.022	0.021	0.000
4	0.049	0.047	-0.002	0.012	0.012	0.000
5	0.048	0.045	-0.002	0.012	0.011	-0.001
6	0.046	0.043	-0.002	0.011	0.011	-0.001
7	0.071	0.070	-0.001	0.018	0.018	0.000
8	0.070	0.067	-0.002	0.017	0.017	-0.001
9	0.068	0.066	-0.002	0.017	0.016	0.000
10	0.056	0.055	-0.001	0.014	0.014	0.000
11	0.055	0.053	-0.002	0.014	0.013	0.000
12	0.052	0.050	-0.002	0.013	0.013	-0.001
13	0.054	0.052	-0.002	0.014	0.013	-0.001
14	0.053	0.051	-0.002	0.013	0.013	-0.001
15	0.051	0.049	-0.002	0.013	0.012	0.000

Table 8.5: Throughput Results for  $p = 50\%$

Configuration	First Level System			SRP System		
	S	A	Error	S	A	Error
16	0.088	0.087	-0.001	0.044	0.043	-0.001
17	0.086	0.085	-0.001	0.043	0.042	-0.001
18	0.083	0.082	-0.001	0.042	0.041	-0.001
19	0.048	0.046	-0.001	0.024	0.023	-0.001
20	0.047	0.045	-0.002	0.023	0.022	-0.001
21	0.045	0.043	-0.002	0.022	0.021	-0.001
22	0.066	0.062	-0.004	0.033	0.031	-0.002
23	0.065	0.061	-0.004	0.032	0.031	-0.002
24	0.063	0.061	-0.003	0.032	0.030	-0.001
25	0.055	0.054	-0.001	0.027	0.027	0.000
26	0.053	0.052	-0.001	0.026	0.026	-0.001
27	0.050	0.049	-0.001	0.025	0.024	-0.001
28	0.051	0.048	-0.003	0.026	0.024	-0.002
29	0.050	0.047	-0.003	0.025	0.024	-0.002
30	0.048	0.046	-0.002	0.024	0.023	-0.001

\* S= Simulation, A= Analytical

Table 8.6: Throughput Results for  $p = 75\%$

Configuration	First Level System			SRP System		
	S	A	Error	S	A	Error
31	0.077	0.066	-0.011	0.058	0.050	-0.008
32	0.077	0.067	-0.009	0.057	0.051	-0.007
33	0.075	0.070	-0.005	0.056	0.053	-0.004
34	0.046	0.045	-0.001	0.035	0.034	-0.001
35	0.045	0.043	-0.002	0.034	0.033	-0.001
36	0.044	0.042	-0.002	0.033	0.031	-0.001
37	0.058	0.052	-0.006	0.043	0.039	-0.004
38	0.058	0.052	-0.006	0.043	0.039	-0.004
39	0.057	0.052	-0.005	0.043	0.039	-0.004
40	0.053	0.052	-0.001	0.040	0.039	0.000
41	0.050	0.049	-0.001	0.037	0.037	0.000
42	0.048	0.047	-0.001	0.036	0.035	-0.001
43	0.047	0.042	-0.005	0.036	0.032	-0.004
44	0.047	0.042	-0.005	0.035	0.032	-0.003
45	0.045	0.041	-0.004	0.034	0.031	-0.003

\* S= Simulation, A= Analytical

As can be expected, the proportion of customers that require service from the SRP node affects the overall system performance. First, the throughput diminishes from one configuration set to the next, given the only parameter changing is the probability of the SRP node service requirement. As  $p$  increases there is a greater influence of the SRP system on the first-level and this also challenges the performance prediction capability of the algorithm. Overall the algorithm predicts the throughput accurately in more than 90% of the configurations.

The throughput estimation is the first step in obtaining other performance metrics. Tables 8.7, 8.8 and 8.9 contains the first level node utilization results.

Table 8.7: First-Level Node Utilization Results for  $p = 25\%$

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
1	91.6	89.88	-2.0%	76.7	75.13	-2.0%	13.7	13.48	-2.0%	91.6	89.88	-2.0%	75.8	74.9	-1.0%
2	89.4	88.08	-1.0%	75.2	73.61	-2.0%	13.4	13.21	-1.0%	89.9	88.08	-2.0%	74.5	73.4	-1.0%
3	86.5	85.13	-2.0%	72.6	71.13	-2.0%	13.0	12.77	-2.0%	86.7	85.13	-2.0%	71.8	70.94	-1.0%
4	87.9	85.03	-3.0%	64.2	62.49	-3.0%	14.6	14.17	-3.0%	48.8	47.24	-3.0%	40.4	39.36	-3.0%
5	85.1	81.88	-4.0%	62.6	60.17	-4.0%	14.2	13.65	-4.0%	47.5	45.49	-4.0%	39.4	37.91	-4.0%
6	81.8	77.84	-5.0%	59.9	57.09	-5.0%	13.6	12.97	-5.0%	45.6	43.25	-5.0%	37.7	36.04	-4.0%
7	92.0	91.19	-1.0%	92.6	92.23	0.0%	10.6	10.52	-1.0%	70.8	70.15	-1.0%	58.7	58.46	0.0%
8	90.3	87.62	-3.0%	91.4	88.54	-3.0%	10.4	10.11	-3.0%	69.9	67.40	-4.0%	57.9	56.17	-3.0%
9	88.1	85.64	-3.0%	88.8	86.51	-3.0%	10.2	9.88	-3.0%	67.8	65.88	-3.0%	56.2	54.9	-2.0%
10	84.4	82.6	-2.0%	60.4	59.01	-2.0%	16.9	16.52	-2.0%	56.3	55.07	-2.0%	46.6	45.89	-2.0%
11	81.4	79.33	-3.0%	59.1	57.15	-3.0%	16.3	15.87	-3.0%	54.6	52.89	-3.0%	45.2	44.07	-2.0%
12	78.2	75.29	-4.0%	56.6	54.21	-4.0%	15.6	15.06	-3.0%	52.2	50.19	-4.0%	43.3	41.83	-3.0%
13	81.4	78.11	-4.0%	80.5	77.43	-4.0%	16.3	15.62	-4.0%	54.3	52.08	-4.0%	45.0	43.4	-4.0%
14	79.2	76.43	-3.0%	78.7	75.73	-4.0%	15.8	15.29	-3.0%	53.0	50.96	-4.0%	43.9	42.46	-3.0%
15	76.2	73.76	-3.0%	75.5	73.05	-3.0%	15.3	14.75	-4.0%	51.0	49.17	-4.0%	42.2	40.98	-3.0%

\* S= Simulation, A= Analytical

Table 8.8: First-Level Node Utilization Results for  $p = 50\%$

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
16	87.7	86.66	-1.0%	87.8	88.04	0.0%	13.2	13.0	-2.0%	87.7	86.66	-1.0%	72.7	72.22	-1.0%
17	85.7	84.86	-1.0%	86.2	85.89	0.0%	12.9	12.73	-1.0%	86.2	84.86	-2.0%	71.4	70.72	-1.0%
18	83.1	82.09	-1.0%	83.2	82.77	-1.0%	12.5	12.31	-1.0%	83.4	82.09	-2.0%	69.1	68.40	-1.0%
19	86.0	83.50	-3.0%	67.8	66.01	-3.0%	14.3	13.92	-3.0%	47.8	46.39	-3.0%	39.6	38.66	-2.0%
20	83.3	80.46	-3.0%	66.1	63.63	-4.0%	13.9	13.41	-4.0%	46.5	44.70	-4.0%	38.6	37.25	-3.0%
21	80.3	76.67	-5.0%	63.0	60.17	-4.0%	13.4	12.78	-5.0%	44.7	42.59	-5.0%	37.0	35.50	-4.0%
22	85.2	80.19	-6.0%	98.0	94.40	-4.0%	9.8	9.25	-6.0%	65.5	61.68	-6.0%	54.3	51.40	-5.0%
23	83.9	79.81	-5.0%	97.1	93.93	-3.0%	9.7	9.21	-5.0%	64.9	61.4	-5.0%	53.8	51.16	-5.0%
24	82.3	79.09	-4.0%	94.9	93.03	-2.0%	9.5	9.13	-4.0%	63.5	60.84	-4.0%	52.6	50.70	-4.0%
25	82.4	80.98	-2.0%	64.5	63.01	-2.0%	16.5	16.20	-2.0%	54.9	53.98	-2.0%	45.5	44.99	-1.0%
26	78.4	77.30	-1.0%	64.0	62.28	-3.0%	15.7	15.46	-2.0%	52.6	51.54	-2.0%	43.6	42.95	-2.0%
27	75.5	73.49	-3.0%	61.3	59.13	-4.0%	15.1	14.70	-3.0%	50.4	48.99	-3.0%	41.8	40.83	-2.0%
28	76.8	71.62	-7.0%	86.2	81.43	-6.0%	15.4	14.32	-7.0%	51.2	47.75	-7.0%	42.4	39.79	-6.0%
29	74.8	70.57	-6.0%	84.4	80.16	-5.0%	15.0	14.11	-6.0%	50.2	47.05	-6.0%	41.6	39.21	-6.0%
30	72.3	68.75	-5.0%	81.2	77.99	-4.0%	14.5	13.75	-5.0%	48.3	45.83	-5.0%	40.0	38.20	-5.0%

\* S= Simulation, A= Analytical

Table 8.9: First-Level Node Utilization Results for  $p = 25\%$

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
31	77.2	66.5	-14.0%	97.3	82.1	-16.0%	11.6	9.97	-14.0%	77.2	66.5	-14.0%	64.0	55.41	-13.0%
32	76.2	67.47	-11.0%	96.2	84.64	-12.0%	11.4	10.12	-11.0%	76.6	67.47	-12.0%	63.5	56.22	-11.0%
33	74.9	70.15	-6.0%	93.8	91.37	-3.0%	11.2	10.52	-6.0%	75.0	70.15	-6.0%	62.2	58.46	-6.0%
34	83.2	81.11	-3.0%	72.3	70.56	-2.0%	13.9	13.52	-3.0%	46.3	45.06	-3.0%	38.3	37.55	-2.0%
35	80.6	78.29	-3.0%	70.4	68.14	-3.0%	13.4	13.05	-3.0%	45.1	43.5	-4.0%	37.3	36.25	-3.0%
36	78.2	74.93	-4.0%	66.9	64.2	-4.0%	13.0	12.49	-4.0%	43.6	41.63	-5.0%	36.1	34.69	-4.0%
37	75.3	67.95	-10.0%	99.8	95.0	-5.0%	8.7	7.84	-10.0%	57.9	52.27	-10.0%	47.9	43.55	-9.0%
38	74.7	67.78	-9.0%	99.6	94.74	-5.0%	8.6	7.82	-9.0%	57.8	52.14	-10.0%	47.9	43.45	-9.0%
39	74.1	67.46	-9.0%	98.5	94.25	-4.0%	8.6	7.78	-9.0%	57.2	51.89	-9.0%	47.4	43.24	-9.0%
40	79.4	78.49	-1.0%	69.9	68.3	-2.0%	15.9	15.7	-1.0%	53.0	52.33	-1.0%	43.9	43.61	-1.0%
41	74.0	73.89	0.0%	69.7	69.22	-1.0%	14.8	14.78	0.0%	49.7	49.26	-1.0%	41.1	41.05	0.0%
42	71.5	70.51	-1.0%	67.0	65.85	-2.0%	14.3	14.10	-1.0%	47.8	47.01	-2.0%	39.6	39.17	-1.0%
43	71.2	63.56	-11.0%	91.3	84.17	-8.0%	14.2	12.71	-10.0%	47.4	42.37	-11.0%	39.3	35.31	-10.0%
44	69.6	63.00	-9.0%	89.7	83.35	-7.0%	13.9	12.60	-9.0%	46.6	42.00	-10.0%	38.6	35.00	-9.0%
45	67.5	61.96	-8.0%	86.7	81.84	-6.0%	13.5	12.39	-8.0%	45.2	41.31	-9.0%	37.4	34.42	-8.0%

\* S= Simulation, A= Analytical

The prediction accuracy obtained in the throughput computation is translated to the utilization estimation, where the higher demand for the SRP node makes it a bit more challenging to estimate the utilization accurately.

It is important to consider the wide range of utilization values obtained in these configurations, from nodes with really high utilization, around 95% to nodes with low utilization, about 10% and all this range of values were estimated very well in most of the cases with the proposed approach.

As the demand for SRP system increases the overall service time at node 2, the node



affected by the SRP system, increases and node 2 becomes the bottleneck node with very high utilizations.

Another important aspect to be notice is the extreme conditions of variability coupled with high utilizations would make some configurations challenging for any solution approach. Despite of all these factors, the results are very good overall.

In tables 8.10, 8.11 and 8.12 the results for the average number of customers in node for first level are presented.

Table 8.10: First Level Average Number in Node Results  $p = 0.25$

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
1	4.41	4.71	1.5%	5.18	4.89	-1.4%	2.74	2.7	-0.2%	4.35	4.47	0.6%	3.31	3.24	-0.4%
2	4.71	5.06	1.8%	5.23	4.92	-1.6%	2.68	2.64	-0.2%	4.18	4.26	0.4%	3.21	3.12	-0.4%
3	5.23	5.54	1.50%	5.26	4.90	-1.8%	2.60	2.55	-0.2%	3.89	4.04	0.8%	3.02	2.97	-0.3%
4	2.56	2.82	2.6%	3.95	3.84	-1.1%	1.46	1.42	-0.4%	0.76	0.69	-0.7%	1.26	1.23	-0.3%
5	2.74	3.03	2.9%	3.87	3.72	-1.5%	1.42	1.36	-0.6%	0.74	0.69	-0.5%	1.23	1.2	-0.3%
6	3.02	3.32	3.0%	3.73	3.55	-1.9%	1.36	1.30	-0.6%	0.71	0.69	-0.2%	1.18	1.15	-0.2%
7	5.9	5.71	-1.0%	7.78	7.99	1.1%	2.12	2.10	-0.1%	2.05	2.01	-0.2%	2.14	2.19	0.2%
8	6.03	5.25	-3.9%	7.79	8.86	5.4%	2.08	2.02	-0.3%	1.99	1.80	-1.0%	2.10	2.07	-0.2%
9	6.39	5.99	-2.0%	7.66	8.25	3.0%	2.04	1.98	-0.3%	1.90	1.76	-0.7%	2.02	2.02	0.0%
10	2.22	2.44	2.3%	3.67	3.58	-0.9%	1.69	1.65	-0.4%	0.95	0.86	-0.9%	1.48	1.46	-0.2%
11	2.4	2.64	2.4%	3.62	3.49	-1.3%	1.63	1.59	-0.4%	0.92	0.87	-0.5%	1.43	1.41	-0.2%
12	2.69	2.94	2.4%	3.49	3.32	-1.7%	1.56	1.51	-0.5%	0.88	0.87	-0.1%	1.37	1.36	-0.1%
13	2.33	2.28	-0.5%	3.63	3.83	2.0%	1.63	1.56	-0.7%	0.96	0.91	-0.5%	1.44	1.42	-0.3%
14	2.49	2.51	0.2%	3.59	3.69	1.0%	1.58	1.53	-0.5%	0.94	0.89	-0.4%	1.41	1.39	-0.2%
15	2.73	2.82	0.9%	3.5	3.48	-0.1%	1.53	1.48	-0.5%	0.9	0.87	-0.3%	1.35	1.34	-0.1%

\* S= Simulation, A= Analytical

Table 8.11: First Level Average Number in Node Results  $p = 0.5$

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
16	3.6	3.71	0.5%	7.13	7.07	-0.3%	2.64	2.6	-0.2%	3.61	3.63	0.1%	3.03	2.99	-0.2%
17	3.96	4.09	0.6%	7.1	7.11	0.0%	2.58	2.55	-0.2%	3.44	3.39	-0.2%	2.93	2.87	-0.3%
18	4.53	4.64	0.5%	7.01	6.99	-0.1%	2.5	2.46	-0.2%	3.18	3.18	0.0%	2.78	2.74	-0.2%
19	2.4	2.63	2.3%	4.2	4.09	-1.1%	1.43	1.39	-0.4%	0.73	0.67	-0.6%	1.23	1.21	-0.2%
20	2.58	2.84	2.6%	4.12	3.97	-1.5%	1.39	1.34	-0.5%	0.71	0.67	-0.4%	1.2	1.17	-0.3%
21	2.88	3.15	2.7%	3.95	3.77	-1.8%	1.34	1.28	-0.6%	0.68	0.67	-0.2%	1.15	1.13	-0.2%
22	3.94	2.89	-5.3%	10.61	11.99	6.9%	1.96	1.85	-0.5%	1.59	1.44	-0.7%	1.9	1.83	-0.3%
23	4.30	3.39	-4.5%	10.34	11.51	5.8%	1.94	1.84	-0.5%	1.55	1.43	-0.6%	1.87	1.82	-0.2%
24	4.95	4.22	-3.6%	9.86	10.72	4.3%	1.9	1.82	-0.4%	1.48	1.43	-0.3%	1.81	1.81	0.0%
25	2.07	2.27	2.0%	3.94	3.85	-0.9%	1.65	1.62	-0.3%	0.9	0.83	-0.7%	1.44	1.43	-0.1%
26	2.20	2.41	2.1%	3.98	3.84	-1.4%	1.57	1.55	-0.2%	0.87	0.83	-0.4%	1.38	1.37	-0.1%
27	2.50	2.72	2.2%	3.84	3.66	-1.8%	1.51	1.47	-0.4%	0.83	0.83	-0.1%	1.32	1.32	0.0%
28	2.02	1.79	-2.3%	4.21	4.70	4.9%	1.54	1.43	-1.1%	0.88	0.79	-0.9%	1.35	1.28	-0.6%
29	2.18	2.00	-1.8%	4.15	4.54	3.9%	1.50	1.41	-0.9%	0.85	0.78	-0.7%	1.32	1.27	-0.5%
30	2.44	2.33	-1.1%	4.03	4.29	2.6%	1.45	1.38	-0.7%	0.82	0.77	-0.5%	1.27	1.24	-0.3%

\* S= Simulation, A= Analytical

Table 8.12: First Level Average Number in Node Results  $p = 0.75$

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
31	2.32	1.39	-4.6%	10.63	6.07	-22.8%	2.32	2.0	-1.6%	2.33	1.39	-4.7%	2.41	1.91	-2.5%
32	2.67	1.75	-4.6%	10.39	6.97	-17.1%	2.28	2.02	-1.3%	2.28	1.49	-3.9%	2.38	1.96	-2.1%
33	3.3	2.58	-3.6%	10.0	11.52	7.6%	2.24	2.1	-0.7%	2.16	1.71	-2.2%	2.3	2.09	-1.1%
34	2.19	2.37	1.8%	4.54	4.45	-0.9%	1.39	1.35	-0.4%	0.7	0.65	-0.5%	1.19	1.18	-0.1%
35	2.37	2.58	2.1%	4.45	4.33	-1.2%	1.34	1.3	-0.4%	0.68	0.65	-0.3%	1.16	1.14	-0.2%
36	2.7	2.92	2.2%	4.24	4.09	-1.5%	1.3	1.25	-0.5%	0.65	0.64	-0.1%	1.12	1.1	-0.2%
37	2.25	1.72	-2.6%	13.33	14.23	4.5%	1.74	1.57	-0.9%	1.10	1.01	-0.5%	1.58	1.48	-0.5%
38	2.66	2.00	-3.3%	12.94	13.96	5.1%	1.72	1.56	-0.8%	1.10	1.01	-0.5%	1.58	1.48	-0.5%
39	3.40	2.50	-4.5%	12.25	13.47	6.1%	1.72	1.56	-0.8%	1.08	1.00	-0.4%	1.56	1.47	-0.4%
40	1.88	2.04	1.6%	4.31	4.21	-1.0%	1.59	1.57	-0.2%	0.84	0.80	-0.5%	1.38	1.38	0.0%
41	1.96	2.09	1.3%	4.46	4.37	-0.9%	1.48	1.48	0.0%	0.81	0.77	-0.4%	1.29	1.3	0.1%
42	2.24	2.4	1.5%	4.32	4.18	-1.4%	1.43	1.41	-0.2%	0.77	0.76	-0.1%	1.24	1.26	0.1%
43	1.70	1.36	-3.4%	4.86	5.59	7.3%	1.42	1.27	-1.5%	0.78	0.66	-1.2%	1.24	1.12	-1.1%
44	1.85	1.52	-3.3%	4.78	5.45	6.7%	1.39	1.26	-1.3%	0.76	0.65	-1.1%	1.21	1.11	-1.0%
45	2.11	1.80	-3.1%	4.64	5.21	5.8%	1.35	1.24	-1.1%	0.73	0.65	-0.8%	1.17	1.10	-0.7%

\* S= Simulation, A= Analytical

The average number of customers per node estimation shows that the analytical model tracks the simulation results very well, and the difference in the estimation in proportion with the total number of customers in the system is low. Only 2 instances out of the 225 instances exceed the 10% error.

In tables 8.13, 8.14 and 8.15 the utilization and average number of customers per node performance metrics for the SRP node are presented.

Table 8.13: Utilization and average number of customers in node performance metrics for the SRP node  $p = 0.25$

Configuration	Utilization			Avg Num in Node		
	S	A	Error	S	A	Error
1	34.1	33.71	-1.2%	0.48	0.46	-0.1%
2	33.5	33.03	-1.4%	0.47	0.45	-0.1%
3	32.3	31.92	-1.2%	0.45	0.44	-0.1%
4	18.2	17.71	-2.7%	0.22	0.21	-0.1%
5	17.7	17.06	-3.6%	0.21	0.2	-0.1%
6	17.1	16.22	-5.2%	0.2	0.18	-0.1%
7	26.5	26.3	-0.7%	0.31	0.31	0.0%
8	26.2	25.27	-3.5%	0.3	0.3	0.0%
9	25.4	24.7	-2.7%	0.29	0.29	0.0%
10	21.1	20.65	-2.1%	0.25	0.24	-0.1%
11	20.4	19.83	-2.8%	0.27	0.26	-0.2%
12	19.5	18.82	-3.5%	0.26	0.24	-0.2%
13	20.3	19.53	-3.8%	0.25	0.24	-0.1%
14	19.8	19.11	-3.5%	0.24	0.23	-0.1%
15	19.1	18.44	-3.5%	0.23	0.22	-0.1%

\* S= Simulation, A= Analytical

Table 8.14: Utilization and average number of customers in node performance metrics for the SRP node  $p = 0.5$

Configuration	Utilization			Avg Num in Node		
	S	A	Error	S	A	Error
16	65.4	64.99	-0.6%	1.32	1.38	0.3%
17	64.3	63.65	-1.0%	1.29	1.33	0.2%
18	62.2	61.56	-1.0%	1.24	1.27	0.2%
19	35.6	34.79	-2.3%	0.51	0.48	-0.3%
20	34.7	33.53	-3.4%	0.49	0.47	-0.3%
21	33.5	31.95	-4.6%	0.45	0.42	-0.3%
22	49.1	46.26	-5.8%	0.65	0.67	0.1%
23	48.7	46.05	-5.4%	0.65	0.67	0.1%
24	47.6	45.63	-4.1%	0.63	0.66	0.2%
25	41.2	40.49	-1.7%	0.58	0.54	-0.3%
26	39.4	38.65	-1.9%	0.68	0.64	-0.4%
27	37.7	36.74	-2.5%	0.65	0.61	-0.4%
28	38.3	35.81	-6.5%	0.54	0.53	-0.1%
29	37.5	35.28	-5.9%	0.53	0.52	-0.1%
30	36.1	34.38	-4.8%	0.51	0.51	0.0%

\* S= Simulation, A= Analytical

Table 8.15: Utilization and average number of customers in node performance metrics for the SRP node  $p = 0.75$

Configuration	Utilization			Avg Num in Node		
	S	A	Error	S	A	Error
31	86.4	74.81	-13.4%	2.36	1.93	-2.1%
32	85.7	75.90	-11.4%	2.33	2.04	-1.4%
33	83.9	78.92	-5.9%	2.25	2.33	0.4%
34	51.7	50.70	-1.9%	0.89	0.85	-0.4%
35	50.4	48.93	-2.9%	0.86	0.83	-0.4%
36	49.0	46.83	-4.4%	0.77	0.73	-0.4%
37	65.1	58.80	-9.7%	0.97	1.02	0.2%
38	65.0	58.66	-9.8%	0.97	1.02	0.2%
39	64.3	58.38	-9.2%	0.96	1.01	0.3%
40	59.6	58.87	-1.2%	1.01	0.96	-0.6%
41	55.7	55.42	-0.5%	1.2	1.2	-0.1%
42	53.6	52.89	-1.3%	1.15	1.13	-0.2%
43	53.2	47.67	-10.4%	0.85	0.83	-0.2%
44	52.4	47.25	-9.8%	0.83	0.82	-0.1%
45	50.7	46.47	-8.3%	0.8	0.8	0.0%

\* S= Simulation, A= Analytical

As expected, the SRP node utilization and average number in node increases as the demand for this service increases. The SRP node utilization results exhibit a wide range of values that validates the approach’s ability to estimate performance accurately over a wide range of parameter values.

It is important to notice that in general, the estimation error increases when the demand for the SRP node increases. This can be seen with the change in the error intervals between tables 8.4 and 8.6. Additionally, there are some characteristics that need special attention

about the error behavior. In general, the first two configurations from each set have higher errors than other configurations in the system. After them, configurations 7 and 8, and 13 and 14 are the ones with higher errors.

If we compare configuration 1, 2 and 3, with configurations 10, 11 and 12 they are similar<sup>1</sup>. The main differences are the number of passive resources in the system and the service time in node 2. Both differences result in customers arriving to the second level node more often, and consequently, demanding more from the SRP system. This finding is consistent with the observation that the error increases when the demand for the second level system increases, as discussed previously.

On the other hand, configurations 7, 8, 9, 13, 14 and 15 have some similarities. For instance, the second node in the first level for these configurations, has high service time variability and smaller number of servers in comparison with the other configurations. This is of special interest since the node 2 is the node influenced by the SRP system and these two parameters are used directly to estimate the parameters for the second level system solution approach. This results in high variability in the arrival process to the SRP node and the small number of servers in node 2, which are essentially the customers in the SRP subsystem, together have a negative impact on the prediction accuracy.

In summary, we demonstrated the viability of the building block approach by embedding the SRP system calculations in the CQN solution algorithm. Extensive numerical experimentation confirmed the performance prediction accuracy of the combined CQN-SRP algorithm.

---

<sup>1</sup>Note that configurations 1,2 and 3 in the first set of configurations are the same as 16, 17 and 18; or 31, 32 and 33 in the other sets of configurations. Consequently, the comments presented here also extend to those configurations.

## CHAPTER IX

### SINGLE-STAGE FORK/JOIN SYSTEM WITH SRP

The fork/join systems are differentiated from the CQN, as discussed previously, in their capacity restricting resource utilization. In contrast to the CQN case, where it is assumed a capacity restricting resource is occupied immediately, almost every single time it is released at the end of the path, in the fork/join case it is likely that the capacity restricting resource will sometimes find no customer waiting for it. This means that the number of customers using the active resources in the system varies within a maximum limit set by the number of passive resources.

The possibility for idle time increases the probability that all the capacity restricting resources become idle at the same time, and consequently, these resources' arrival process to the synchronization station could shut down. This possible arrivals shut down effect is one of the fork/join study's principal concerns.

In Figure 9.1 a fork/join system example is presented, with a synchronization node at the beginning of the system and not at the end, given that the customers can depart from the system as soon as they end the service at the last node in the system.



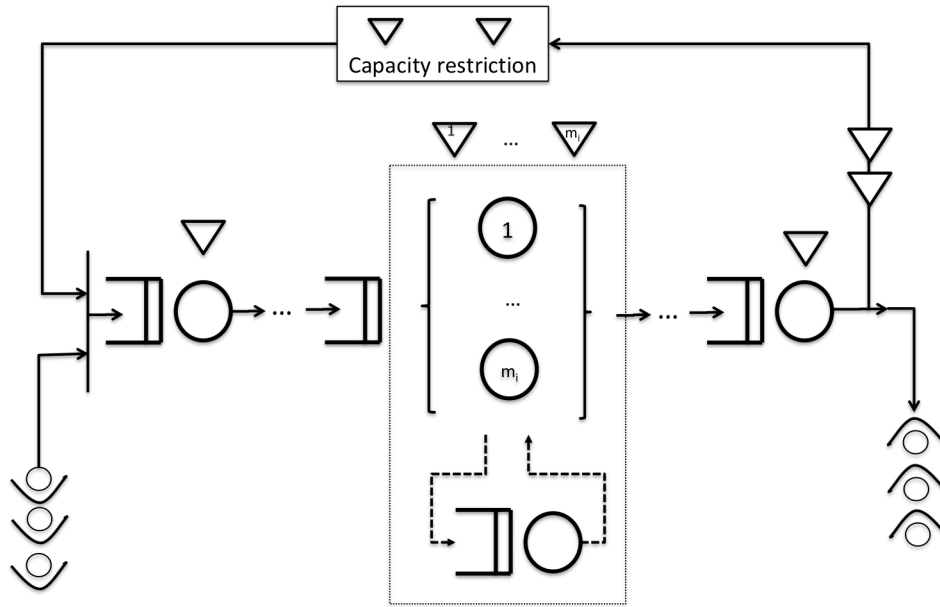


Figure 9.1: Single-Stage Fork/Join System With Simultaneous Resource Possession

The fork/join system solution approach follows the considerations discussed in Chapter VI. However, it is important to keep in mind that this system's maximal throughput cannot be exceeded by the external arrival process, otherwise stability problems could arise. This topic was discussed in previous chapters and depicted in Figure 7.2 for the two-stage case.

### 9.1 Algorithm for Solving the Single-Stage Capacity Restricted System with SRP

In previous sections, algorithms for solving the first and second level systems were presented. Similarly, the general approach for computing the parameters require for each subsystem were introduced and explained. Based on these explanations, an iterative solution is presented in order to solve a single-stage fork/join system with SRP.

The algorithm computes the SRP response time for each iteration of the first level solution. At each iteration of the solution, results from the first level system feed the second level, and vice versa to obtain results from complementary systems.

There are two iterative processes, the second-level system computation is an iterative process that is executed within each iteration in the first-level system. In the second-level system or SRP system, the iterative process is executed until the  $c_{a,SRP}^2$  value converges based on some specified threshold error. The first-level iterative process will carry-on following the original criteria specified by Krishnamurthy and Suri [32] for the fork/join approach.

The overall algorithm for solving a single-stage fork/join system with SRP is presented next. This algorithm integrates the single-stage system with one synchronization station algorithm presented in Chapter VI with the SRP calculations presented in Chapter IV.

---

**Algorithm 8** Single-Stage Fork/Join With SRP Algorithm

---

**Require:**  $\lambda_{P,i-1}$ ,  $c_{P,i-1}^2$ ,  $K_{i-1}$ ,  $\mu_{S,i}$ ,  $c_{S,i}^2$ ,  $K_i$ ,  $\lambda_{F,i+1}$ ,  $c_{F,i+1}^2$ ,  $K_{i+1}$ .

**Ensure:**  $Low = 0$ ,  $High = \min(\lambda_{P,i-1}, \mu_{S,i})$ ,  $\lambda_{F,i+1}$ ,  $j=1, \delta_1 = 100, \epsilon_1 = 0.001$ ,  $\epsilon_2 = 0.001$ ,  $\epsilon_3 = 0.001$ .

**Step 1**

Compute  $\tau_r^I$  and  $c_{s,r}^{2I}$  based on results in equations 4.5 and 4.7

**Step 2**

Solve the first-level FJ system with corresponding algorithm in chapter VI. Obtain  $\lambda_a$  and  $c_{a,s}^2$

Set  $High = \lambda_a$

**Begin Loop 1. Step 2.1**

**while**  $|\delta_1| \geq \epsilon_1$  **do**

$\lambda_{d,i}^{(j)} = (Low + High)/2$  and  $c_{d,i}^{2(j)} = 1$  (say)

**Begin Loop 2. Step 2.2**

Set  $\delta_2 = 100$

**while**  $|\delta_2| > \epsilon_2$  **do**

**Step 2.2.1**

Solve steps **Step 2.1** to **Step 2.5** from the main algorithm in Chapter VI

**Step 2.2.2****Begin Loop 3**

Set  $\delta_3 = 100$

**while**  $\delta_3 > \epsilon_3$  **do**

Set  $\rho_r^I = \lambda^{(j)} * \tau_r^I / m_r$

Validate  $\rho_r^I$ , if  $\rho_r^I > 0.9995$  then adjust  $\lambda_{d,i}^{(j)} = 0.9995 * m_r / (\tau_r^I)$

Compute  $c_{d,r}^2$ , using Equation 4.10.

Set  $c_{d,r}^{2,(j)} = c_{d,r}^2$

Compute  $W_{qSRP}$  and  $\sigma_{q,SRP}^2$ , using the SRP algorithm in Section 4.7

Based on the new  $W_{qSRP}$  and  $\sigma_{q,SRP}^2$  results, compute  $\tau_r^I$  and  $c_{s,r}^{2I}$  using equations 4.5 and 4.7.

Compute  $\delta_3 = c_{d,r}^{2,(j)} - c_{d,r}^{2,(j-1)}$ , and then set  $c_{d,r}^{2,(j-1)} = c_{d,r}^{2,(j)}$

**end while**

**Step Step 2.2.3**

Compute  $\lambda_{a,i}$ ,  $c_{a,i}^2$  using linking equations 6.9 and 6.10.

**Step 2.2.4**

Compute  $\lambda_{d,i}$ ,  $c_{d,i}^2$ , and  $\bar{L}_{S,i}$  using characterization Eqs. 6.11, 4.9, and 4.24.

**Step 2.2.5**

Compute  $\delta_2 = |c_{d,i}^2 - c_{d,i}^{2,(j)}|$ . Set  $c_{d,i}^{2,(j)} = c_{d,i}^2$

**end while**

**Step 2.3**

Complete steps 3 and 4 from the original algorithm.

Set  $j = j + 1$

**end while**

---

The main fork/join algorithm iterates on the throughput value. Given the SRP response time calculations were embedded in the Fork/Join algorithm, the throughput value in an iteration and the SRP response time added to the first-level node service time, could cause the utilization value for the equivalent first-level node to exceed 100%. We have to include logic to detect such situations in an iterative approach to avoid invalid computations. The algorithm includes an adjustment for the lambda value to levels that do not cause utilization to exceed 99% in order to keep the computations valid, and not affecting the bisection algorithm that governs the main fork/join algorithm.

## **9.2 Numerical Experiments for Single-Stage Fork/Join System with SRP**

For testing this single-stage fork/join system with SRP we used a system with a single synchronization stage that joined the external customer arrival process with the capacity restricting resources once they are released at the end of the system.

The internal network structure in the fork/join system is similar to the one utilized in the CQN example, with five nodes and the same node parameters. The main variation is in the external arrivals that happen in the fork/join system and the choice of the arrival process. In the next section, the parameter selection will be explained.

### **9.2.1 Parameter selection**

The first and second level node parameters were chosen to be identical to the parameters used in the CQN with SRP Chapter VIII, and can be seen in tables 8.2 and 8.3. This parameter selection was retained based on two aspects. First, they have yielded challenging configurations to test the algorithms. Second, by keeping the same configurations allows direct

comparison and could be useful in gaining additional insights into the problem. As before, a total of forty five configurations were used and divided in sets of fifteen. Configurations in each set had the same probability of requiring service from the SRP system.

The only difference here is the external customer arrival process to the synchronization station. As discussed previously, this arrival rate must be less than the throughput of the CQN with SRP version of the problem, in order to ensure stability. The external arrival rates were chosen to be 85% of the throughput estimated for the corresponding configuration in Chapter VIII and a SCV of 1. Table 9.1 shows the parameters used for the external arrival process.

Table 9.1: External Arrival Process Parameters

Configuration	Arrival Rate	Configuration	Arrival Rate	Configuration	Arrival Rate
1	0.07640	16	0.07366	31	0.05652
2	0.07486	17	0.07213	32	0.05735
3	0.07236	18	0.06977	33	0.05963
4	0.04015	19	0.03943	34	0.03830
5	0.03866	20	0.03800	35	0.03697
6	0.03676	21	0.03621	36	0.03539
7	0.05962	22	0.05243	37	0.04443
8	0.05729	23	0.05219	38	0.04432
9	0.05599	24	0.05171	39	0.04411
10	0.04681	25	0.04589	40	0.04448
11	0.04496	26	0.04380	41	0.04187
12	0.04267	27	0.04164	42	0.03996
13	0.04426	28	0.04058	43	0.03602
14	0.04331	29	0.03999	44	0.03570
15	0.04180	30	0.03896	45	0.03511

## Experimental Results

The system throughput, node utilization, and average number in node values were computed to study analytical results. The number in node in the synchronization station is presented also. The results were compared with simulation results and error estimations were computed following the method presented in Chapter V. It is important to keep in mind that the average number of customers in node has been normalized with the number of passive resources for all nodes, including the SRP node. This approach was used in Krishnamurthy and Suri [32].

In tables 9.2, 9.3 and 9.4 the throughput results for the first and second level systems are presented. These three tables correspond to the SRP subsystem usage probability of 25%, 50% and 75% respectively.

Table 9.2: Throughput Results for  $p=0.25$

Configuration	First Level System			SRP System		
	S	A	Error	S	A	Error
1	0.0768	0.0755	-0.0013	0.0192	0.0189	-0.0003
2	0.0753	0.0739	-0.0013	0.0188	0.0185	-0.0003
3	0.0727	0.0713	-0.0014	0.0182	0.0178	-0.0003
4	0.0402	0.0385	-0.0017	0.0101	0.0096	-0.0004
5	0.0387	0.0370	-0.0017	0.0097	0.0093	-0.0004
6	0.0367	0.0351	-0.0016	0.0092	0.0088	-0.0004
7	0.0599	0.0594	-0.0006	0.015	0.0148	-0.0001
8	0.0576	0.0570	-0.0006	0.0144	0.0143	-0.0001
9	0.0563	0.0556	-0.0007	0.0141	0.0139	-0.0002
10	0.0469	0.0447	-0.0022	0.0117	0.0112	-0.0005
11	0.045	0.0429	-0.0021	0.0113	0.0107	-0.0005
12	0.0425	0.0406	-0.0019	0.0106	0.0102	-0.0005
13	0.0442	0.0431	-0.0011	0.0111	0.0108	-0.0003
14	0.0432	0.0421	-0.0012	0.0108	0.0105	-0.0003
15	0.0416	0.0404	-0.0012	0.0104	0.0101	-0.0003

\* S= Simulation, A= Analytical

Table 9.3: Throughput Results for  $p=0.5$

Configuration	First Level System			SRP System		
	S	A	Error	S	A	Error
16	0.0741	0.0727	-0.0014	0.0370	0.0363	-0.0007
17	0.0725	0.0711	-0.0014	0.0363	0.0356	-0.0007
18	0.0701	0.0687	-0.0014	0.0351	0.0343	-0.0007
19	0.0395	0.0377	-0.0018	0.0197	0.0188	-0.0009
20	0.038	0.0363	-0.0018	0.019	0.0181	-0.0009
21	0.0361	0.0345	-0.0016	0.0181	0.0172	-0.0008
22	0.0527	0.0523	-0.0004	0.0264	0.0262	-0.0002
23	0.0525	0.0519	-0.0006	0.0262	0.0260	-0.0003
24	0.052	0.0513	-0.0007	0.026	0.0256	-0.0004
25	0.0459	0.0436	-0.0024	0.023	0.0218	-0.0012
26	0.0438	0.0416	-0.0022	0.0219	0.0208	-0.0011
27	0.0415	0.0395	-0.0020	0.0207	0.0197	-0.0010
28	0.0406	0.0396	-0.0010	0.0203	0.0198	-0.0005
29	0.0400	0.0389	-0.0011	0.0200	0.0194	-0.0006
30	0.0389	0.0377	-0.0012	0.0194	0.0188	-0.0006

\* S= Simulation, A= Analytical



Table 9.4: Throughput Results for  $p=0.75$

Configuration	First Level System			SRP System		
	S	A	Error	S	A	Error
31	0.0568	0.0563	-0.0006	0.0426	0.0422	-0.0004
32	0.0577	0.0570	-0.0007	0.0433	0.0428	-0.0005
33	0.0600	0.0590	-0.0009	0.045	0.0443	-0.0007
34	0.0383	0.0365	-0.0018	0.0287	0.0274	-0.0014
35	0.0370	0.0352	-0.0018	0.0277	0.0264	-0.0013
36	0.0353	0.0336	-0.0017	0.0265	0.0252	-0.0013
37	0.0447	0.0443	-0.0004	0.0335	0.0332	-0.0003
38	0.0446	0.0442	-0.0004	0.0334	0.0331	-0.0003
39	0.0444	0.0439	-0.0005	0.0333	0.0329	-0.0004
40	0.0445	0.0421	-0.0025	0.0334	0.0315	-0.0018
41	0.0418	0.0396	-0.0021	0.0313	0.0297	-0.0016
42	0.0398	0.0378	-0.0020	0.0298	0.0284	-0.0015
43	0.0361	0.0352	-0.0009	0.0271	0.0264	-0.0007
44	0.0358	0.0348	-0.0010	0.0268	0.0261	-0.0007
45	0.0352	0.0341	-0.0011	0.0264	0.0256	-0.0008

\* S= Simulation, A= Analytical

The throughput values for the first-level system must be very close to the external arrival rate of customers. However, as explained in Chapter VI, the Krishnamurthy and Suri [32] approach requires the specification of a  $K$  value for external customer queue as they assume that these arrivals also happen from another capacity-restricted system. Setting  $K$  to be large is approximately equivalent to not having a limit on the external customer queue size at the first synchronization node. In our experiments we set this value to be equal to the first level population constraint following examples in [32]. As a result, there could be a small

fraction of external arrivals that may be lost if this queue becomes full.

In tables 9.5, 9.6 and 9.7 the utilization performance metric for nodes in the first level system are shown for the SRP system usage of 25%, 50% and 75% respectively.

Table 9.5: First Level Node Utilization Results for  $p=0.25$

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S*	A*	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
1	76.8	75.51	-1.68%	64.0	67.45	5.39%	11.5	11.33	-1.51%	76.8	75.51	-1.68%	63.7	62.93	-1.22%
2	74.8	73.92	-1.17%	62.7	66.03	5.31%	11.2	11.09	-1.0%	75.2	73.92	-1.7%	62.4	61.6	-1.28%
3	72.5	71.34	-1.6%	60.6	63.71	5.14%	10.9	10.7	-1.82%	72.7	71.34	-1.87%	60.2	59.45	-1.24%
4	72.4	69.32	-4.25%	52.9	53.26	0.68%	12.1	11.55	-4.51%	40.2	38.51	-4.2%	33.3	32.09	-3.62%
5	69.3	66.64	-3.85%	50.9	51.2	0.58%	11.5	11.11	-3.43%	38.7	37.02	-4.34%	32.1	30.85	-3.9%
6	65.8	63.16	-4.01%	48.1	48.46	0.75%	11.0	10.53	-4.3%	36.7	35.09	-4.38%	30.4	29.24	-3.81%
7	77.9	77.18	-0.92%	78.3	85.0	8.55%	9.0	8.91	-1.05%	59.9	59.37	-0.88%	49.6	49.48	-0.25%
8	74.4	74.15	-0.34%	75.3	81.64	8.41%	8.6	8.56	-0.52%	57.6	57.04	-0.98%	47.7	47.53	-0.36%
9	73.0	72.31	-0.94%	73.5	79.63	8.34%	8.4	8.34	-0.67%	56.3	55.63	-1.2%	46.6	46.35	-0.53%
10	70.3	66.98	-4.72%	50.2	50.54	0.69%	14.1	13.4	-4.99%	46.9	44.65	-4.79%	38.8	37.21	-4.1%
11	67.1	64.28	-4.21%	48.6	48.81	0.43%	13.4	12.86	-4.07%	45.0	42.85	-4.78%	37.3	35.71	-4.27%
12	63.7	60.9	-4.39%	45.9	46.23	0.72%	12.7	12.18	-4.09%	42.6	40.6	-4.69%	35.2	33.83	-3.88%
13	66.4	64.69	-2.57%	65.4	69.19	5.79%	13.3	12.94	-2.72%	44.2	43.13	-2.42%	36.7	35.94	-2.07%
14	64.5	63.1	-2.16%	64.0	67.48	5.44%	12.9	12.62	-2.16%	43.3	42.07	-2.84%	35.8	35.06	-2.07%
15	62.2	60.58	-2.6%	61.5	64.77	5.32%	12.5	12.12	-3.07%	41.6	40.39	-2.91%	34.5	33.66	-2.44%

\* S= Simulation, A= Analytical

Table 9.6: First Level Node Utilization Results for p=0.50

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S*	A*	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
16	74.1	72.66	-1.95%	72.6	79.58	9.62%	11.1	10.9	-1.82%	74.1	72.66	-1.95%	61.3	60.55	-1.23%
17	72.1	71.11	-1.38%	71.0	77.81	9.59%	10.8	10.67	-1.24%	72.6	71.11	-2.06%	60.1	59.25	-1.41%
18	69.9	68.69	-1.73%	68.5	75.12	9.67%	10.5	10.3	-1.87%	70.1	68.69	-2.01%	58.1	57.24	-1.47%
19	71.1	67.81	-4.62%	55.7	57.77	3.71%	11.8	11.3	-4.22%	39.5	37.67	-4.62%	32.7	31.4	-3.99%
20	68.0	65.27	-4.02%	53.6	55.64	3.8%	11.3	10.88	-3.74%	38.0	36.26	-4.58%	31.5	30.22	-4.08%
21	64.8	62.03	-4.27%	50.6	52.63	4.02%	10.8	10.34	-4.27%	36.1	34.46	-4.53%	29.9	28.72	-3.95%
22	68.5	68.03	-0.69%	78.2	90.17	15.3%	7.9	7.85	-0.64%	52.8	52.33	-0.89%	43.7	43.61	-0.21%
23	67.8	67.49	-0.46%	77.9	89.49	14.88%	7.8	7.79	-0.17%	52.5	51.91	-1.12%	43.5	43.26	-0.55%
24	67.4	66.65	-1.12%	77.2	88.49	14.62%	7.8	7.69	-1.41%	52.0	51.27	-1.41%	43.1	42.72	-0.87%
25	68.9	65.34	-5.17%	53.6	55.74	4.0%	13.8	13.07	-5.3%	45.9	43.56	-5.1%	38.0	36.3	-4.47%
26	65.3	62.36	-4.5%	52.6	54.52	3.65%	13.0	12.47	-4.06%	43.8	41.57	-5.08%	36.3	34.64	-4.56%
27	62.1	59.23	-4.62%	49.8	51.76	3.94%	12.4	11.85	-4.46%	41.5	39.49	-4.85%	34.4	32.91	-4.34%
28	60.9	59.35	-2.54%	67.8	75.69	11.64%	12.2	11.87	-2.7%	40.6	39.57	-2.54%	33.6	32.97	-1.87%
29	59.6	58.31	-2.17%	66.8	74.37	11.33%	11.9	11.66	-2.01%	40.0	38.87	-2.82%	33.1	32.39	-2.14%
30	58.2	56.51	-2.9%	64.9	72.1	11.09%	11.6	11.3	-2.57%	38.9	37.67	-3.15%	32.2	31.39	-2.5%

\* S= Simulation, A= Analytical

Table 9.7: First Level Node Utilization Results for  $p=0.75$

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S*	A*	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
31	56.9	56.29	-1.08%	65.6	73.81	12.51%	8.5	8.44	-0.67%	56.8	56.29	-0.91%	47.1	46.9	-0.42%
32	57.4	57.04	-0.63%	66.9	75.49	12.84%	8.6	8.56	-0.52%	57.7	57.04	-1.15%	47.8	47.53	-0.56%
33	59.8	59.05	-1.26%	70.6	79.74	12.94%	9.0	8.86	-1.59%	60.0	59.05	-1.59%	49.7	49.21	-1.0%
34	69.0	65.68	-4.82%	59.0	62.19	5.41%	11.5	10.95	-4.82%	38.3	36.49	-4.73%	31.7	30.41	-4.08%
35	66.2	63.34	-4.32%	56.8	60.13	5.86%	11.0	10.56	-4.03%	37.0	35.19	-4.89%	30.6	29.32	-4.17%
36	63.4	60.5	-4.58%	53.4	56.94	6.64%	10.6	10.08	-4.88%	35.3	33.61	-4.79%	29.2	28.01	-4.09%
37	58.1	57.63	-0.81%	75.5	89.8	18.94%	6.7	6.65	-0.75%	44.7	44.33	-0.83%	37.0	36.94	-0.16%
38	57.6	57.45	-0.27%	75.3	89.64	19.05%	6.6	6.63	0.43%	44.6	44.19	-0.92%	36.9	36.82	-0.21%
39	57.5	57.06	-0.77%	75.0	89.29	19.05%	6.6	6.58	-0.25%	44.4	43.89	-1.15%	36.8	36.58	-0.61%
40	66.8	63.09	-5.56%	57.6	60.89	5.72%	13.4	12.62	-5.84%	44.5	42.06	-5.49%	36.9	35.05	-5.02%
41	62.3	59.44	-4.59%	57.2	60.39	5.58%	12.4	11.89	-4.13%	41.8	39.63	-5.2%	34.6	33.02	-4.56%
42	59.5	56.71	-4.69%	54.3	57.62	6.12%	11.9	11.34	-4.69%	39.8	37.8	-5.01%	32.9	31.5	-4.24%
43	54.2	52.84	-2.52%	68.3	79.18	15.93%	10.8	10.57	-2.15%	36.1	35.22	-2.43%	29.9	29.35	-1.83%
44	53.4	52.23	-2.18%	67.7	78.33	15.7%	10.7	10.45	-2.37%	35.8	34.82	-2.73%	29.6	29.02	-1.96%
45	52.6	51.13	-2.8%	66.5	76.77	15.44%	10.5	10.23	-2.62%	35.2	34.08	-3.17%	29.1	28.4	-2.39%

\* S= Simulation, A= Analytical

The results in the first level node utilizations show that the analytical results track the simulation estimates quite well. Most errors reported are below 10%. Additionally, increases in the utilization for the first level node influenced by the SRP subsystem, namely, Node 2 can be seen with an increase in  $p$  as more and more customers spend additional time at this node.

Another important aspect to notice is the node 2 utilization computations accuracy. Given that node 2 is the node under the SRP sub-system's influence, it is important in our analysis. These results validate the approach used for SRP node calculations and the subsequent modifications to the node 2 service time, considering the wide range of utilization results the Node 2 exhibits in the experiments conducted.

Tables 9.8, 9.9 and 9.10 present the average number in node performance metric for the 3 sets of fifteen configurations with  $p=0.25$ ,  $0.50$  and  $0.75$ .

Table 9.8: Average Number in Node Results for  $p = 0.25$

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
1	2.42	1.98	-2.18%	4.07	4.2	0.64%	2.3	2.27	-0.17%	2.37	1.86	-2.57%	2.41	2.29	-0.61%
2	2.66	2.29	-1.84%	4.05	4.15	0.51%	2.24	2.22	-0.11%	2.28	1.9	-1.93%	2.34	2.24	-0.49%
3	3.15	2.77	-1.9%	4.02	4.07	0.22%	2.18	2.14	-0.2%	2.19	1.94	-1.27%	2.24	2.18	-0.28%
4	1.67	1.36	-3.1%	3.22	3.21	-0.04%	1.21	1.16	-0.55%	0.58	0.48	-1.0%	1.03	0.98	-0.41%
5	1.77	1.57	-1.99%	3.1	3.09	-0.09%	1.15	1.11	-0.39%	0.55	0.49	-0.64%	0.99	0.95	-0.37%
6	1.96	1.84	-1.2%	2.94	2.94	-0.08%	1.1	1.05	-0.47%	0.52	0.49	-0.33%	0.94	0.91	-0.23%
7	2.67	2.01	-3.34%	4.55	3.59	-4.8%	1.8	1.78	-0.09%	1.33	0.94	-1.93%	1.68	1.61	-0.34%
8	2.74	2.22	-2.62%	4.24	3.59	-3.23%	1.72	1.71	-0.04%	1.22	0.92	-1.51%	1.6	1.55	-0.27%
9	3.36	2.82	-2.67%	4.31	3.95	-1.81%	1.68	1.67	-0.06%	1.2	0.97	-1.16%	1.56	1.52	-0.2%
10	1.52	1.27	-2.6%	3.03	3.05	0.14%	1.41	1.34	-0.7%	0.72	0.59	-1.3%	1.21	1.15	-0.58%
11	1.62	1.44	-1.81%	2.95	2.95	-0.02%	1.34	1.29	-0.54%	0.69	0.61	-0.86%	1.16	1.11	-0.49%
12	1.81	1.68	-1.32%	2.8	2.8	-0.01%	1.27	1.22	-0.52%	0.65	0.61	-0.45%	1.09	1.07	-0.26%
13	1.44	1.14	-2.95%	2.59	2.32	-2.66%	1.33	1.29	-0.36%	0.7	0.55	-1.45%	1.15	1.11	-0.41%
14	1.56	1.33	-2.31%	2.55	2.33	-2.22%	1.29	1.26	-0.28%	0.68	0.56	-1.16%	1.12	1.08	-0.34%
15	1.79	1.6	-1.87%	2.49	2.34	-1.51%	1.25	1.21	-0.38%	0.65	0.58	-0.75%	1.08	1.05	-0.28%

\* S= Simulation, A= Analytical

Table 9.9: Average Number in Node Results for  $p = 0.50$

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
16	2.12	1.75	-1.84%	5.08	5.31	1.11%	2.22	2.18	-0.2%	2.12	1.68	-2.19%	2.26	2.16	-0.53%
17	2.35	2.02	-1.67%	5.01	5.28	1.31%	2.16	2.13	-0.13%	2.03	1.66	-1.81%	2.2	2.1	-0.46%
18	2.81	2.44	-1.85%	4.94	5.21	1.32%	2.1	2.06	-0.2%	1.91	1.67	-1.21%	2.1	2.04	-0.31%
19	1.59	1.29	-2.99%	3.4	3.5	0.91%	1.18	1.13	-0.5%	0.56	0.47	-0.97%	1.01	0.96	-0.44%
20	1.69	1.49	-2.03%	3.28	3.37	0.91%	1.13	1.09	-0.42%	0.54	0.47	-0.65%	0.97	0.93	-0.38%
21	1.89	1.75	-1.4%	3.11	3.2	0.93%	1.08	1.03	-0.46%	0.51	0.47	-0.34%	0.92	0.89	-0.25%
22	1.76	1.36	-2.0%	4.48	3.21	-6.37%	1.58	1.57	-0.05%	0.97	0.75	-1.12%	1.42	1.39	-0.17%
23	2.04	1.64	-1.98%	4.56	5.1	2.69%	1.56	1.56	-0.01%	0.97	0.76	-1.06%	1.41	1.38	-0.18%
24	2.62	2.15	-2.38%	4.76	5.88	5.57%	1.56	1.54	-0.11%	0.97	0.78	-0.95%	1.4	1.37	-0.18%
25	1.45	1.2	-2.55%	3.25	3.37	1.2%	1.38	1.31	-0.73%	0.7	0.57	-1.23%	1.18	1.12	-0.6%
26	1.52	1.34	-1.86%	3.23	3.31	0.8%	1.3	1.25	-0.53%	0.67	0.58	-0.9%	1.13	1.08	-0.52%
27	1.71	1.56	-1.48%	3.07	3.15	0.81%	1.24	1.18	-0.55%	0.63	0.58	-0.52%	1.07	1.03	-0.34%
28	1.2	0.96	-2.38%	2.79	2.72	-0.69%	1.22	1.19	-0.33%	0.61	0.49	-1.22%	1.04	1.01	-0.32%
29	1.33	1.11	-2.12%	2.77	2.76	-0.08%	1.19	1.17	-0.24%	0.6	0.5	-1.04%	1.03	1.0	-0.32%
30	1.55	1.35	-2.01%	2.73	2.8	0.7%	1.16	1.13	-0.3%	0.59	0.51	-0.78%	1.0	0.97	-0.28%

\* S= Simulation, A= Analytical

Table 9.10: Average Number in Node Results for  $p = 0.75$

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
30	1.55	1.35	-2.01%	2.73	2.8	0.7%	1.16	1.13	-0.3%	0.59	0.51	-0.78%	1.0	0.97	-0.28%
31	1.11	0.95	-0.84%	4.58	4.77	0.95%	1.7	1.69	-0.06%	1.1	0.93	-0.86%	1.55	1.53	-0.12%
32	1.32	1.15	-0.82%	4.77	5.02	1.27%	1.72	1.71	-0.04%	1.14	0.99	-0.76%	1.58	1.56	-0.1%
33	1.86	1.62	-1.21%	5.39	5.83	2.18%	1.8	1.77	-0.14%	1.26	1.13	-0.67%	1.67	1.65	-0.11%
34	1.48	1.2	-2.79%	3.63	3.78	1.46%	1.15	1.09	-0.55%	0.54	0.45	-0.91%	0.97	0.93	-0.42%
35	1.58	1.38	-2.06%	3.51	3.66	1.57%	1.1	1.06	-0.44%	0.52	0.45	-0.64%	0.94	0.9	-0.37%
36	1.8	1.64	-1.58%	3.3	3.48	1.85%	1.06	1.01	-0.52%	0.49	0.45	-0.32%	0.89	0.87	-0.26%
37	1.17	0.93	-1.22%	3.92	4.53	3.04%	1.34	1.33	-0.05%	0.69	0.58	-0.59%	1.16	1.15	-0.05%
38	1.34	1.1	-1.17%	3.98	4.98	4.98%	1.32	1.33	0.03%	0.7	0.58	-0.57%	1.16	1.15	-0.05%
39	1.72	1.43	-1.45%	4.18	5.77	7.96%	1.32	1.32	-0.02%	0.7	0.6	-0.51%	1.16	1.14	-0.06%
40	1.35	1.11	-2.4%	3.51	3.7	1.84%	1.34	1.26	-0.78%	0.66	0.55	-1.12%	1.14	1.08	-0.61%
41	1.39	1.2	-1.85%	3.58	3.69	1.13%	1.24	1.19	-0.51%	0.63	0.54	-0.92%	1.07	1.02	-0.5%
42	1.57	1.41	-1.63%	3.41	3.53	1.25%	1.19	1.13	-0.56%	0.59	0.53	-0.59%	1.02	0.98	-0.34%
43	0.96	0.78	-1.8%	2.86	3.01	1.48%	1.08	1.06	-0.23%	0.52	0.43	-0.93%	0.92	0.9	-0.24%
44	1.07	0.9	-1.7%	2.86	3.07	2.1%	1.07	1.04	-0.25%	0.51	0.43	-0.83%	0.91	0.89	-0.23%
45	1.27	1.09	-1.85%	2.86	3.16	2.97%	1.05	1.02	-0.27%	0.51	0.44	-0.7%	0.9	0.87	-0.25%

\* S= Simulation, A= Analytical

Calculating the average number at Node 2 involves multiple layers of approximations, first, the variance of waiting time at the SRP node needs to be computed. This is not a typical calculation and involves several approximations [54].

This variance is used in computing the variance of the equivalent service at Node2. Finally, as explained in Chapter IV, additional correction factor were developed for waiting time at a multi-server node. Despite the many different approximations, it should be noted that the analytical model is able to accurately capture the average number at Node 2, except for very few cases where the errors are greater than 15%.

Other results of interest are the SRP node performance metrics and the synchronization node results. In tables 9.11, 9.12, and 9.13 the SRP node utilization and average number in queue are presented. Similarly the number of customers in queue for the synchronization node.

Table 9.11: SRP Node and Synchronization Node Results for  $p=0.25$

Conf	SRP Node Utilization			SRP Avg. Number in Node			Avg. External Queue Length		
	S	A	Error	S	A	Error	S	A	Error
1	28.7	28.32	-1.34%	0.39	0.37	-0.1%	0.55	3.17	13.09%
2	28.1	27.72	-1.35%	0.38	0.36	-0.09%	0.61	3.32	13.54%
3	27.1	26.75	-1.28%	0.36	0.34	-0.09%	0.91	3.54	13.13%
4	15.0	14.44	-3.72%	0.17	0.16	-0.11%	0.97	2.44	14.64%
5	14.4	13.88	-3.59%	0.17	0.16	-0.1%	0.95	2.46	15.06%
6	13.8	13.16	-4.64%	0.15	0.14	-0.1%	1.06	2.47	14.15%
7	22.5	22.26	-1.05%	0.26	0.25	-0.03%	0.68	1.58	4.53%
8	21.6	21.39	-0.98%	0.25	0.24	-0.04%	0.51	1.61	5.48%
9	21.1	20.86	-1.14%	0.24	0.23	-0.03%	0.79	2.1	6.56%
10	17.6	16.74	-4.86%	0.2	0.19	-0.15%	1.07	2.64	15.69%
11	16.9	16.07	-4.92%	0.22	0.2	-0.19%	1.05	2.63	15.78%
12	15.9	15.23	-4.24%	0.2	0.19	-0.18%	1.12	2.61	14.86%
13	16.5	16.17	-1.98%	0.2	0.19	-0.09%	0.91	1.78	8.72%
14	16.2	15.78	-2.62%	0.19	0.18	-0.1%	0.95	1.9	9.45%
15	15.6	15.15	-2.91%	0.19	0.18	-0.1%	1.11	2.06	9.52%

\* S= Simulation, A= Analytical



Table 9.12: SRP Node and Synchronization Node Results for  $p=0.50$

Conf	SRP Node Utilization			SRP Avg. Number in Node			Avg. External Queue Length		
	S	A	Error	S	A	Error	S	A	Error
16	55.3	54.49	-1.46%	1.03	0.96	-0.33%	0.68	3.53	14.26%
17	54.1	53.33	-1.42%	0.99	0.94	-0.3%	0.69	3.63	14.68%
18	52.3	51.52	-1.49%	0.96	0.9	-0.28%	0.91	3.8	14.45%
19	29.4	28.26	-3.89%	0.4	0.36	-0.41%	1.01	2.59	15.79%
20	28.3	27.19	-3.91%	0.38	0.35	-0.35%	0.98	2.59	16.14%
21	27.1	25.85	-4.62%	0.34	0.31	-0.27%	1.07	2.59	15.19%
22	39.5	39.25	-0.64%	0.51	0.48	-0.12%	0.28	0.89	3.04%
23	39.4	38.94	-1.18%	0.5	0.48	-0.13%	0.33	1.83	7.52%
24	39.0	38.45	-1.41%	0.5	0.48	-0.13%	0.53	2.57	10.21%
25	34.5	32.67	-5.3%	0.46	0.4	-0.55%	1.12	2.81	16.97%
26	32.7	31.18	-4.65%	0.53	0.46	-0.63%	1.14	2.79	16.48%
27	31.0	29.62	-4.46%	0.5	0.44	-0.57%	1.19	2.75	15.58%
28	30.4	29.68	-2.38%	0.42	0.39	-0.25%	0.75	1.74	9.96%
29	29.9	29.15	-2.5%	0.41	0.38	-0.23%	0.79	1.87	10.72%
30	29.1	28.26	-2.9%	0.4	0.37	-0.24%	0.96	2.05	10.9%

\* S= Simulation, A= Analytical

Table 9.13: SRP Node and Synchronization Node Results for  $p=0.75$

Conf	SRP Node Utilization			SRP Avg. Number in Node			Avg. External Queue Length		
	S	A	Error	S	A	Error	S	A	Error
31	63.6	63.32	-0.44%	1.38	1.26	-0.6%	0.08	1.54	7.34%
32	64.5	64.17	-0.52%	1.42	1.32	-0.49%	0.12	1.84	8.6%
33	67.1	66.43	-1.0%	1.54	1.46	-0.37%	0.36	2.76	11.97%
34	42.9	41.05	-4.32%	0.68	0.58	-0.98%	1.03	2.69	16.66%
35	41.4	39.59	-4.38%	0.65	0.57	-0.8%	1.0	2.69	16.86%
36	39.7	37.81	-4.76%	0.57	0.52	-0.56%	1.07	2.69	16.24%
37	50.3	49.87	-0.85%	0.7	0.64	-0.3%	0.07	0.98	4.53%
38	50.2	49.71	-0.97%	0.7	0.65	-0.28%	0.08	1.23	5.71%
39	49.9	49.38	-1.05%	0.7	0.65	-0.26%	0.13	1.74	8.03%
40	50.1	47.31	-5.56%	0.79	0.66	-1.31%	1.16	2.95	17.9%
41	46.9	44.58	-4.95%	0.93	0.8	-1.28%	1.27	2.89	16.14%
42	44.7	42.53	-4.85%	0.88	0.76	-1.11%	1.29	2.84	15.52%
43	40.5	39.63	-2.15%	0.61	0.57	-0.38%	0.53	1.59	10.67%
44	40.2	39.18	-2.55%	0.6	0.57	-0.39%	0.56	1.71	11.47%
45	39.5	38.34	-2.93%	0.59	0.56	-0.37%	0.7	1.9	11.96%

\* S= Simulation, A= Analytical

The SRP node utilization shows good agreement between the simulation and analytical results for a wide variety of utilization levels. Similarly, both analytical and simulation results match well for the average number of customers in the SRP node. On the other hand, the number of customers waiting in the external queue at the synchronization is not as accurate. Reasons for larger error at this node have been explained in earlier chapters.

Despite the overall good accuracy in performance prediction, configurations 4, 5, 6, 10, 11, and 12 for the first set of configurations, and their counterpart in others sets; show higher errors. These six configurations have two things in common. First, the number of passive

resources in the system is low and the number of servers in the second node, the node affected by the SRP system, is high.

First of all, it is known that in capacity-restricted systems, small numbers of passive resources could affect the prediction accuracy in average waiting time and average number in node. This is caused by the fact that these values are derived from computations based on non-restricted systems ([29, 28]) and consequently the higher the number of the customers in the system, closer is the system to an open system or a system with infinite capacity. Now, in this case, the first level system solution proposed by Krishnamurthy and Suri [32] iterates based on the average number in node, as the criterion to stop the iteration method. Hence, the number of passive resources can be expected an impact on the overall performance prediction.

Second, as the SRP demand increases, the second node, the node affected by the SRP system, becomes more preponderant in the overall system behavior. Mainly, in the six configurations under discussion, where node two has six servers and the system has 10 passive resources. This means, as the second node saturation increases, motivated by a higher demand of the SRP node, more customers are going to get occupy this node and the effects of small errors in the computations for this node could result in magnified errors in performance predictions for the overall system.

The number in node for the second node, that is one of the metrics affected by the overall number of passive resources discussed earlier, and crucial in Suri and Krishnamurthy's algorithm [32], will be affected by diminished number of passive resources in the system, given it is a multi-server node. The effect observed in the six configuration under discussion is because in these configurations node 2 is such an important piece for the systems (node

two could contain up to 6 of 10 passive resources being served simultaneously, plus some more in queue all present in node 2 simultaneously) that small errors in the approximations related with this node, result in big impacts on the overall system performance prediction.

It is interesting to note that these findings are not completely aligned with the findings for the CNQ with SRP system. However, this is not unusual at all, since both systems have different first-level solution approaches and consequently, the overall system accuracy depends on depending on the specific building block's characteristics and capabilities.

In this chapter we showed how the SRP building block can be integrated with a first-level network that is capacity-restricted, but may not be at full capacity all the time. Both the SRP node and the single-stage fork/join system involve multiple approximations. The ability of the overall model to estimate system performance reasonably accurately is a testament of the robustness of the two-moment approximations and the versatility of the parametric-decomposition method.

## CHAPTER X

### TWO CAPACITY-RESTRICTED STAGES WITH A SHARED SRP SYSTEM

The problem studied in this chapter is a two-stage capacity restricted system at the first level with a shared SRP system at the second level. The shared SRP system was introduced in Chapter IV and the approach to derive the required parameters was discussed. On the other hand, the approach to solve the two-stage system was developed and tested with numerical examples in Chapter VII. In this chapter, the steps to combine both building blocks will be detailed and tested using extensive numerical experimentation. In Figure 10.1 an example of this system can be seen.

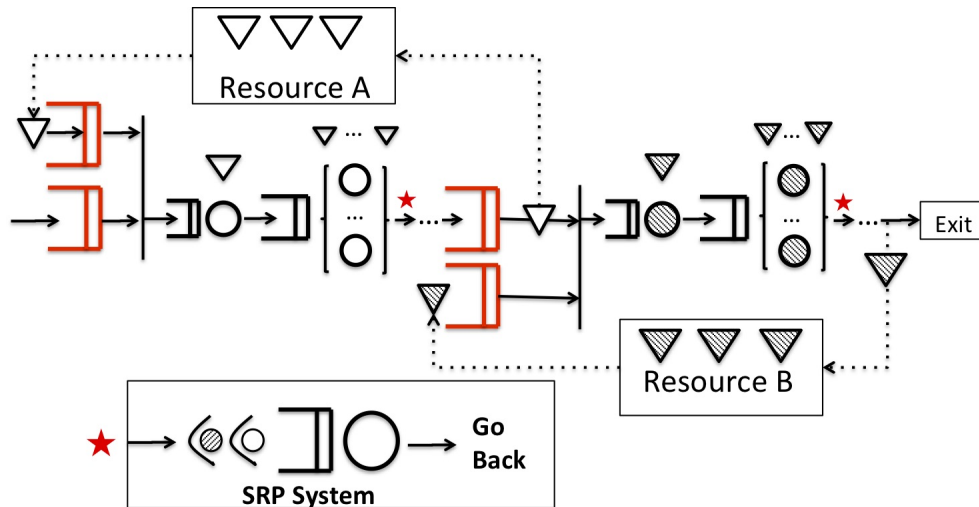


Figure 10.1: Two Capacity-Restricted Stages with a Shared SRP System

As can be seen, the SRP node is shared by customers from both stages at the first level. Consequently, the arrival process to the SRP node must be characterized by the merging

of arrival processes from the two stages, as shown in equations 4.50 and 4.51. In the next section, the proposed solution approach is discussed.

## 10.1 The Solution Approach

The solution to this problem involves linking the solutions to the building blocks. The basis for the fork/join system solution is the approach proposed by Krishnamurthy and Suri [32]. Then in Chapter VI, the inclusion of multi-server nodes was discussed. In Chapter VII, we presented the extension of the Krishnamurthy and Suri's to solve a two-stage system. In Chapter IV the SRP subsystem solution was discussed first for single source SRP system, and later extended for multiple sources. All of these solution approaches will be combined in this chapter to solve the problem under discussion.

Given there is an external arrival to this system, it is important to verify the system stability. This topic was formerly discussed, and the graphical representation of the procedure can be seen in Figure 7.2. Once the stability condition has been checked, the following algorithm can be used to solve the two-stage fork/join problem with SRP. Krishnamurthy and Suri [32] approach does require the specification of a limit for the size of the queue of external arrivals at the synchronization node. Even with this safety feature, we still believe it is prudent to check if the stability condition holds.

The overall solution approach is shown in figure 10.2. The solution algorithm is presented at aggregate level in algorithm 9. Many of steps involve solving building blocks presented in previous chapters.

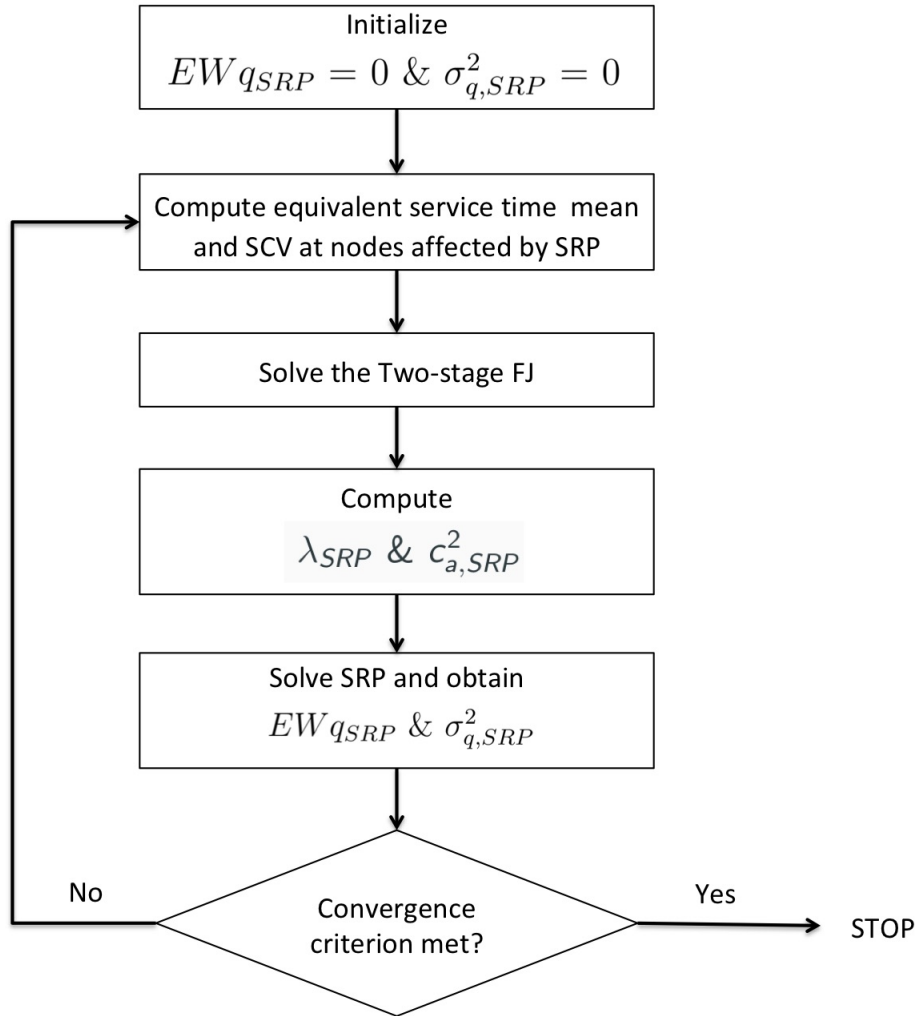


Figure 10.2: Two-Stage with SRP System Solution Approach

---

**Algorithm 9** Two Capacity-Restricted Stages With a Shared SRP System

---

**Require:**  $\lambda_{P,0}$ ,  $c_{a,P,0}^2$ ,  $K_0$ ,  $K_1$ ,  $K_2$ ,  $\tau_{SRP}$ ,  $c_{s,SRP}^2$ ,  $m_{SRP}$ , the parameter's tuple  $\tau_{s,S,i}$ ,  $c_{s,S,i}^2$ ,  $m_{S,i}$  for all nodes  $S$  in the stages  $i$ .

**Ensure:**  $\epsilon = 0.01$ ,  $\delta = 100$ ,  $W_{qSRP} = 0$   $\sigma_{qSRP}^2 = 0$ ,  $j=1$

**Step 1**

**while**  $|\delta| > \epsilon$  **do**

**Step 1.1**

Based on equations 4.5 and 4.7, compute the equivalent  $\tau_{r,1}^I$ ,  $\tau_{r,2}^I$ , and  $c_{s,r,1}^{2I}$  and  $c_{s,r,2}^{2I}$  for the nodes affected by the SRP in each stage.

**Step 1.2**

Solve Stage 1 and Stage 2 with the equivalent times, using CQN with approximation in Chapter V

**if**  $\lambda_{P,0} > \lambda_{CQN,STAGE1}$  **or**  $\lambda_{P,0} > \lambda_{CQN,STAGE2}$  **then**

    End the algorithm

    Report "System is not stable"

**end if**

**Step 1.3**

Solve the two stages system using approach in Chapter VII.

Set  $\lambda_{d,1}^{(j)} = \lambda_{d,1}$

**Step 1.4**

Based on the throughput obtained in the two-stage system solution, compute the  $\lambda_{SRP}$ ,  $c_{a,SRP}^2$  and  $n_{SRP}$  using equations 4.50, 4.51 and 4.49.

**Step 1.5**

Using the parameters computed in step 1.4, compute  $W_{qSRP}$  and  $\sigma_{W_{qSRP}}^2$  using equation 4.26 and 4.30 if the SRP node is a single server node or 4.29 and 4.38 if the SRP is a multi-server node.

Compute  $\delta = \lambda_{d,1}^{(j)} - \lambda_{d,1}$

Set  $\lambda_{d,1}^{(j-1)} = \lambda_{d,1}^{(j)}$

Set  $j = j + 1$

**end while**

---



## 10.2 Experimental Set Up

The experimental set up used here closely follows that used in Chapter VII. However, a few configurations were modified in order to obtain the throughput necessary to keep the system stable once the SRP node is introduced. The experimental configurations and results are presented and discussed in next sections.

### 10.2.1 Experimental Parameters

Forty five experiments were performed to test the approach in Algorithm 9, organized in three sets of fifteen each. The parameters for the basic set of fifteen are the same configurations presented in Chapter VII, however the last two configurations of each set had to be adjusted, as was previously explained.

The external arrival rates were computed as the 75% of the throughput for the configuration with the smallest throughput between the two in the stages for each configuration. The arrival rates are presented in Table 10.1, all of them with a SCV of one.

Table 10.1: Arrival Rates

Conf	$\lambda_{P,0}$	Conf	$\lambda_{P,0}$	Conf	$\lambda_{P,0}$
1	0.0687	16	0.0658	31	0.0579
2	0.0674	17	0.0647	32	0.0575
3	0.0650	18	0.0626	33	0.0563
4	0.0366	19	0.0358	34	0.0347
5	0.0357	20	0.0349	35	0.0338
6	0.0342	21	0.0335	36	0.0327
7	0.0531	22	0.0491	37	0.0434
8	0.0524	23	0.0487	38	0.0433
9	0.0509	24	0.0476	39	0.0429
10	0.0407	25	0.0384	40	0.0356
11	0.0398	26	0.0376	41	0.035
12	0.0382	27	0.0362	42	0.0339
13	0.0407	28	0.0384	43	0.0356
14	0.0398	29	0.0376	44	0.0350
15	0.0382	30	0.0362	45	0.0339

The node parameters are presented in Tables 10.2 and 10.3 for the first and second stages, respectively. Note the three sets of experiments share the same first and second-level node parameters, and only the demand probability for the SRP node varies from one set to the other.

Table 10.2: Experimental Parameters for the First Stage

Conf	$K_1$	$\tau_{1,1}$	$c_{s,1,1}^2$	$m_{1,1}$	$\tau_{1,2}$	$c_{s,1,2}^2$	$m_{1,2}$	$\tau_{1,3}$	$c_{s,1,3}^2$	$m_{1,3}$	$\tau_{1,4}$	$c_{s,1,4}^2$	$m_{1,4}$	$\tau_{1,5}$	$c_{s,1,5}^2$	$m_{1,5}$
1	20	10.0	0.5	1	45.0	0.5	6	30.0	0.5	20	10.0	0.5	1	25.0	1.0	3
2	20	10.0	1.0	1	45.0	0.5	6	30.0	1.0	20	10.0	0.5	1	25.0	1.0	3
3	15	10.0	2.0	1	45.0	0.5	6	30.0	2.0	15	10.0	0.5	1	25.0	1.0	3
4	10	18.0	0.5	1	75.0	1.0	6	30.0	0.5	10	10.0	0.5	1	25.0	1.0	3
5	10	18.0	1.0	1	75.0	1.0	6	30.0	1.0	10	10.0	0.5	1	25.0	1.0	3
6	12	18.0	2.0	1	75.0	1.0	6	30.0	2.0	12	10.0	0.5	1	25.0	1.0	3
7	20	13.0	0.5	1	35.0	2.0	3	30.0	0.5	20	10.0	0.5	1	25.0	1.0	3
8	20	13.0	1.0	1	35.0	2.0	3	30.0	1.0	20	10.0	0.5	1	25.0	1.0	3
9	15	13.0	2.0	1	35.0	2.0	3	30.0	2.0	15	10.0	0.5	1	25.0	1.0	3
10	10	15.0	0.5	1	60.0	0.5	6	30.0	0.5	10	10.0	0.5	1	25.0	1.0	3
11	10	15.0	1.0	1	60.0	0.5	6	30.0	1.0	10	10.0	0.5	1	25.0	1.0	3
12	12	15.0	2.0	1	60.0	0.5	6	30.0	2.0	12	10.0	0.5	1	25.0	1.0	3
13	10	15.0	0.5	1	40.0	2.0	3	30.0	0.5	10	10.0	0.5	1	25.0	1.0	3
14	12	15.0	1.0	1	40.0	2.0	3	30.0	1.0	12	10.0	0.5	1	25.0	1.0	3
15	10	15.0	2.0	1	40.0	2.0	3	30.0	2.0	10	10.0	0.5	1	25.0	1.0	3

Table 10.3: Experimental Parameters for the Second Stage

Conf	$K_2$	$\tau_{2,1}$	$c_{s,2,1}^2$	$m_{2,1}$	$\tau_{2,2}$	$c_{s,2,2}^2$	$m_{2,2}$	$\tau_{2,3}$	$c_{s,2,3}^2$	$m_{2,3}$	$\tau_{2,4}$	$c_{s,2,4}^2$	$m_{2,4}$	$\tau_{2,5}$	$c_{s,2,5}^2$	$m_{2,5}$
1	20	10.0	0.5	1	45.0	0.5	6	30.0	0.5	20	10.0	0.5	1	25.0	1.0	3
2	15	10.0	1.0	1	45.0	0.5	6	30.0	1.0	15	10.0	0.5	1	25.0	1.0	3
3	20	10.0	2.0	1	45.0	0.5	6	30.0	2.0	20	10.0	0.5	1	25.0	1.0	3
4	10	18.0	0.5	1	75.0	1.0	6	30.0	0.5	10	10.0	0.5	1	25.0	1.0	3
5	12	18.0	1.0	1	75.0	1.0	6	30.0	1.0	12	10.0	0.5	1	25.0	1.0	3
6	10	18.0	2.0	1	75.0	1.0	6	30.0	2.0	10	10.0	0.5	1	25.0	1.0	3
7	20	13.0	0.5	1	35.0	2.0	3	30.0	0.5	20	10.0	0.5	1	25.0	1.0	3
8	15	13.0	1.0	1	35.0	2.0	3	30.0	1.0	15	10.0	0.5	1	25.0	1.0	3
9	20	13.0	2.0	1	35.0	2.0	3	30.0	2.0	20	10.0	0.5	1	25.0	1.0	3
10	10	15.0	0.5	1	40.0	2.0	3	30.0	0.5	10	10.0	0.5	1	25.0	1.0	3
11	12	15.0	1.0	1	40.0	2.0	3	30.0	1.0	12	10.0	0.5	1	25.0	1.0	3
12	10	15.0	2.0	1	40.0	2.0	3	30.0	2.0	10	10.0	0.5	1	25.0	1.0	3
13	10	15.0	0.5	1	60.0	0.5	6	30.0	0.5	10	10.0	0.5	1	25.0	1.0	3
14	10	15.0	1.0	1	60.0	0.5	6	30.0	1.0	10	10.0	0.5	1	25.0	1.0	3
15	12	15.0	2.0	1	60.0	0.5	6	30.0	2.0	12	10.0	0.5	1	25.0	1.0	3

The difference between each of the sets is the probability of use of the SRP system. The different probabilities are as explained in Table 10.4.

Table 10.4: Probability of First-Level Nodes Using the SRP Subsystem

First Subset			Second Subset			Third Subset		
Conf	P Stage 1*	P Stage 2*	Conf	P Stage 1	P Stage 2	Conf	P Stage 1	P Stage 2
1	0.25	0.25	16	0.5	0.25	31	0.75	0.25
2	0.25	0.5	17	0.5	0.5	32	0.75	0.5
3	0.25	0.75	18	0.5	0.75	33	0.75	0.75
4	0.25	0.25	19	0.5	0.25	34	0.75	0.25
5	0.25	0.5	20	0.5	0.5	35	0.75	0.5
6	0.25	0.75	21	0.5	0.75	36	0.75	0.75
7	0.25	0.5	22	0.5	0.5	37	0.75	0.5
8	0.25	0.75	23	0.5	0.75	38	0.75	0.75
9	0.25	0.25	24	0.5	0.25	39	0.75	0.25
10	0.25	0.5	25	0.5	0.5	40	0.75	0.5
11	0.25	0.75	26	0.5	0.75	41	0.75	0.75
12	0.25	0.25	27	0.5	0.25	42	0.75	0.25
13	0.25	0.5	28	0.5	0.5	43	0.75	0.5
14	0.25	0.75	29	0.5	0.75	44	0.75	0.75
15	0.25	0.25	30	0.5	0.25	45	0.75	0.25

\* Note. “P Stage 1” and “P Stage 2” stand for “probability for the use of the SRP node in Stage 1” and “probability for the use of the SRP node in Stage 2”.

In appendix, an extra set of 45 experiments are presented. These experiments have the same parameters introduced in this section but slightly higher arrival rates, specifically the same as presented in Table 7.1 for each set of fifteen configurations.

### 10.2.2 Experiment Results

The two-stages with shared SRP problem has several nodes and many performance metrics involved. In order to help the reader follow the results, the experimental results will be

organized as follows.

- Throughput results
- First-stage results
- Second-stage results
- Fork/join results
- SRP results.

The analytical results were compared with simulation results. The error estimations were computed following the methods presented in Chapter V. It is important to note that the average number of customers at a node has been normalized with the number of passive resources in a stage for all nodes in that stage. In this case, since the SRP node is shared by two stages, these results were normalized by the number of servers at nodes affected by SRP, which happens to be the number of customers in the second level system,  $N_{2L}$ . The use of the stage's number of passive resources to normalize the error was used in Krishnamurthy and Suri [32].

## **Throughput Results**

In tables 10.5, 10.6, and 10.7, the throughput results are presented for the two capacity-restricted stages and the SRP node.

Table 10.5: Throughput Results for First Set of Configurations

Conf	Stage 1			Stage 2			SRP Subsystem		
	Simulation	Analytical	Difference	Simulation	Analytical	Difference	Simulation	Analytical	Difference
1	0.0691	0.0683	-0.0007	0.0691	0.068	-0.001	0.0345	0.0342	0.0004
2	0.0678	0.0670	-0.0008	0.0678	0.0661	-0.0017	0.0508	0.0503	0.0006
3	0.0653	0.0636	-0.0016	0.0653	0.0631	-0.0022	0.0653	0.0636	0.0016
4	0.0368	0.0356	-0.0011	0.0368	0.0349	-0.0019	0.0184	0.0178	0.0006
5	0.0358	0.0347	-0.0011	0.0358	0.0342	-0.0016	0.0268	0.026	0.0008
6	0.0343	0.0336	-0.0006	0.0343	0.0328	-0.0015	0.0343	0.0336	0.0006
7	0.0534	0.0530	-0.0004	0.0534	0.0529	-0.0005	0.0401	0.0397	0.0003
8	0.0527	0.0523	-0.0005	0.0527	0.0518	-0.0009	0.0527	0.0523	0.0005
9	0.0511	0.0503	-0.0008	0.0511	0.0500	-0.0011	0.0256	0.0251	0.0004
10	0.0409	0.0397	-0.0011	0.0409	0.039	-0.0019	0.0307	0.0298	0.0009
11	0.0400	0.0388	-0.0011	0.0400	0.0383	-0.0016	0.0399	0.0388	0.0011
12	0.0384	0.0377	-0.0006	0.0384	0.0371	-0.0013	0.0192	0.0189	0.0003
13	0.0409	0.0399	-0.0010	0.0409	0.039	-0.0018	0.0307	0.0299	0.0008
14	0.0399	0.0394	-0.0006	0.0399	0.0385	-0.0014	0.0399	0.0394	0.0006
15	0.0383	0.0374	-0.0009	0.0383	0.0368	-0.0015	0.0191	0.0187	0.0005

Table 10.6: Throughput Results for Second Set of Configurations

Conf	Stage 1			Stage 2			SRP Subsystem		
	Simulation	Analytical	Difference	Simulation	Analytical	Difference	Simulation	Analytical	Difference
16	0.0662	0.0655	-0.0006	0.0662	0.0653	-0.0009	0.0496	0.0491	0.0005
17	0.0650	0.0644	-0.0007	0.065	0.0637	-0.0014	0.0651	0.0644	0.0007
18	0.0629	0.0614	-0.0014	0.0629	0.061	-0.0019	0.0786	0.0768	0.0018
19	0.0360	0.0349	-0.0010	0.0360	0.0342	-0.0017	0.0270	0.0262	0.0008
20	0.0351	0.0340	-0.0011	0.0351	0.0336	-0.0015	0.0351	0.034	0.0011
21	0.0337	0.0330	-0.0006	0.0337	0.0323	-0.0014	0.0421	0.0413	0.0008
22	0.0494	0.0491	-0.0004	0.0494	0.049	-0.0004	0.0494	0.0491	0.0004
23	0.0490	0.0486	-0.0003	0.049	0.0484	-0.0006	0.0612	0.0608	0.0004
24	0.0478	0.0472	-0.0006	0.0478	0.0471	-0.0008	0.0359	0.0354	0.0005
25	0.0386	0.0377	-0.0009	0.0386	0.0372	-0.0014	0.0386	0.0377	0.0009
26	0.0378	0.0369	-0.0009	0.0378	0.0365	-0.0013	0.0472	0.0461	0.0011
27	0.0364	0.0359	-0.0006	0.0364	0.0354	-0.001	0.0273	0.0269	0.0004
28	0.0386	0.0378	-0.0008	0.0386	0.0371	-0.0014	0.0386	0.0378	0.0008
29	0.0377	0.0373	-0.0004	0.0377	0.0367	-0.0010	0.0471	0.0466	0.0005
30	0.0363	0.0355	-0.0008	0.0363	0.0351	-0.0012	0.0272	0.0267	0.0006



Table 10.7: Throughput Results for the Third Set of Configurations

Conf	Stage 1			Stage 2			SRP Subsystem		
	Simulation	Analytical	Difference	Simulation	Analytical	Difference	Simulation	Analytical	Difference
31	0.0583	0.0578	-0.0004	0.0583	0.0577	-0.0005	0.0583	0.0578	0.0004
32	0.0578	0.0573	-0.0004	0.0578	0.057	-0.0007	0.0722	0.0717	0.0005
33	0.0566	0.0557	-0.0009	0.0566	0.0554	-0.0012	0.0849	0.0835	0.0014
34	0.0349	0.0339	-0.0009	0.0349	0.0333	-0.0015	0.0349	0.0339	0.0009
35	0.0340	0.033	-0.0009	0.034	0.0326	-0.0013	0.0425	0.0413	0.0012
36	0.0328	0.0323	-0.0006	0.0328	0.0316	-0.0012	0.0492	0.0484	0.0008
37	0.0437	0.0434	-0.0003	0.0437	0.0434	-0.0003	0.0546	0.0542	0.0003
38	0.0436	0.0433	-0.0003	0.0436	0.0432	-0.0004	0.0654	0.0649	0.0005
39	0.0431	0.0427	-0.0004	0.0431	0.0426	-0.0005	0.0431	0.0427	0.0004
40	0.0358	0.0351	-0.0007	0.0358	0.0347	-0.001	0.0447	0.0439	0.0009
41	0.0352	0.0345	-0.0007	0.0352	0.0342	-0.001	0.0528	0.0517	0.0011
42	0.034	0.0336	-0.0004	0.034	0.0333	-0.0007	0.034	0.0336	0.0004
43	0.0357	0.0352	-0.0006	0.0357	0.0347	-0.001	0.0447	0.044	0.0007
44	0.035	0.0348	-0.0002	0.035	0.0343	-0.0007	0.0525	0.0522	0.0004
45	0.0339	0.0333	-0.0006	0.0339	0.033	-0.0009	0.0339	0.0333	0.0006

The results show that the analytical results track the simulations estimates very closely. Most of the differences are around the 0.001. Another interesting observation is that throughput values for the second stage seem to be a bit less accurate in general, than throughput values for the first stage.

### First Stage Results

The first stage of the two-stage first-level system is characterized by having two synchronization stations, one at the beginning and one at the end. The utilization results for the first-level nodes in this first stage and the average number of customers in a node are presented in this section. It is important to notice the demand for the SRP node increases from 25 % to 50 % to 75 % respectively, for the configuration sets.

In tables 10.8,10.9, and 10.10 utilization results for the first-level nodes in the first stage are presented.

Table 10.8: Utilization Results for Stage 1, First Set of Configurations

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
1	69.1	68.33	-1.11%	54.9	53.53	-2.5%	10.4	10.25	-1.44%	69.1	68.33	-1.11%	57.2	56.95	-0.45%
2	67.4	67.03	-0.55%	54.5	52.51	-3.66%	10.1	10.05	-0.45%	67.8	67.03	-1.13%	56.2	55.86	-0.61%
3	65.1	63.64	-2.24%	53.3	49.85	-6.47%	13.0	12.73	-2.09%	65.3	63.64	-2.54%	54.1	53.04	-1.97%
4	66.2	64.13	-3.12%	47.1	45.72	-2.92%	11.0	10.69	-2.83%	36.8	35.63	-3.18%	30.5	29.69	-2.65%
5	64.1	62.44	-2.59%	46.0	44.52	-3.23%	10.7	10.41	-2.75%	35.8	34.69	-3.11%	29.7	28.91	-2.67%
6	61.6	60.55	-1.7%	44.0	43.17	-1.88%	8.5	8.41	-1.06%	34.3	33.64	-1.92%	28.4	28.03	-1.29%
7	69.4	68.89	-0.74%	66.8	65.36	-2.16%	8.0	7.95	-0.64%	53.4	52.99	-0.77%	44.2	44.16	-0.09%
8	68.2	67.94	-0.38%	66.4	64.46	-2.93%	7.9	7.84	-0.77%	52.7	52.26	-0.83%	43.7	43.55	-0.34%
9	66.3	65.38	-1.38%	63.5	62.03	-2.31%	10.2	10.06	-1.38%	51.1	50.29	-1.58%	42.3	41.91	-0.92%
10	61.3	59.62	-2.75%	42.6	41.07	-3.59%	12.3	11.92	-3.06%	40.9	39.74	-2.83%	33.9	33.12	-2.3%
11	59.6	58.22	-2.31%	42.1	40.11	-4.73%	11.9	11.64	-2.15%	39.9	38.81	-2.72%	33.1	32.34	-2.28%
12	57.4	56.61	-1.38%	40.0	39.0	-2.51%	9.6	9.43	-1.72%	38.4	37.74	-1.72%	31.8	31.45	-1.11%
13	61.3	59.86	-2.36%	58.2	55.86	-4.01%	12.3	11.97	-2.67%	40.9	39.9	-2.44%	33.9	33.25	-1.91%
14	59.6	59.07	-0.88%	62.9	56.78	-9.74%	9.9	9.85	-0.55%	39.9	39.38	-1.3%	33.1	32.82	-0.85%
15	57.3	56.03	-2.22%	56.8	53.85	-5.2%	11.4	11.21	-1.71%	38.3	37.35	-2.48%	31.7	31.13	-1.81%

\* S= Simulation, A= Analytical

Table 10.9: Utilization Results for Stage 1, Second Set of Configurations

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
16	66.2	65.52	-1.02%	56.6	53.51	-5.46%	9.9	9.83	-0.72%	66.2	65.52	-1.02%	54.8	54.6	-0.36%
17	64.7	64.39	-0.48%	57.2	52.58	-8.07%	9.7	9.66	-0.43%	65.0	64.39	-0.94%	53.9	53.66	-0.45%
18	62.7	61.44	-2.01%	57.3	50.17	-12.43%	12.5	12.29	-1.7%	62.9	61.44	-2.32%	52.1	51.2	-1.73%
19	64.7	62.89	-2.8%	47.7	46.0	-3.56%	10.8	10.48	-2.95%	36.0	34.94	-2.95%	29.8	29.12	-2.3%
20	62.7	61.21	-2.37%	46.7	44.78	-4.12%	10.5	10.2	-2.84%	35.1	34.01	-3.11%	29.0	28.34	-2.28%
21	60.4	59.48	-1.52%	44.9	43.51	-3.1%	8.4	8.26	-1.65%	33.7	33.04	-1.94%	27.9	27.54	-1.3%
22	64.3	63.79	-0.79%	66.6	63.79	-4.22%	7.4	7.36	-0.53%	49.4	49.07	-0.67%	40.9	40.89	-0.02%
23	63.3	63.2	-0.16%	67.1	63.2	-5.81%	7.3	7.29	-0.1%	49.0	48.62	-0.78%	40.6	40.51	-0.21%
24	62.0	61.37	-1.02%	63.6	61.37	-3.51%	9.5	9.44	-0.62%	47.8	47.21	-1.24%	39.6	39.34	-0.66%
25	57.9	56.5	-2.41%	42.0	40.18	-4.33%	11.6	11.3	-2.58%	38.6	37.67	-2.41%	32.0	31.39	-1.91%
26	56.4	55.31	-1.93%	42.4	39.33	-7.23%	11.3	11.06	-2.1%	37.8	36.88	-2.45%	31.3	30.73	-1.82%
27	54.5	53.8	-1.28%	39.8	38.26	-3.87%	9.1	8.97	-1.46%	36.4	35.87	-1.46%	30.2	29.89	-1.02%
28	57.9	56.69	-2.1%	59.1	55.43	-6.22%	11.6	11.34	-2.27%	38.6	37.79	-2.1%	32.0	31.49	-1.59%
29	56.2	55.97	-0.41%	70.4	57.83	-17.85%	9.4	9.33	-0.77%	37.7	37.31	-1.03%	31.2	31.09	-0.34%
30	54.3	53.29	-1.85%	60.6	55.07	-9.13%	10.9	10.66	-2.21%	36.3	35.53	-2.12%	30.1	29.61	-1.64%

\* S= Simulation, A= Analytical

Table 10.10: Utilization Results for Stage 1, Third Set of Configurations

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
31	58.2	57.81	-0.67%	53.8	49.14	-8.66%	8.7	8.67	-0.33%	58.3	57.81	-0.84%	48.2	48.17	-0.05%
32	57.4	57.35	-0.09%	55.8	48.74	-12.64%	8.6	8.6	0.02%	57.8	57.35	-0.78%	47.9	47.79	-0.23%
33	56.5	55.67	-1.46%	57.8	47.32	-18.13%	11.3	11.13	-1.46%	56.6	55.67	-1.64%	46.9	46.39	-1.08%
34	62.8	61.06	-2.77%	48.0	45.8	-4.59%	10.5	10.18	-3.07%	34.9	33.92	-2.8%	28.9	28.27	-2.18%
35	60.8	59.45	-2.22%	47.2	44.59	-5.53%	10.1	9.91	-1.89%	34.0	33.03	-2.86%	28.1	27.52	-2.05%
36	58.9	58.06	-1.43%	45.6	43.54	-4.51%	8.2	8.06	-1.67%	32.8	32.25	-1.67%	27.2	26.88	-1.18%
37	56.8	56.41	-0.69%	63.3	59.3	-6.32%	6.5	6.51	0.13%	43.6	43.39	-0.48%	36.2	36.16	-0.12%
38	56.3	56.29	-0.03%	64.7	59.17	-8.54%	6.5	6.49	-0.09%	43.6	43.3	-0.7%	36.1	36.08	-0.05%
39	55.9	55.49	-0.73%	61.4	58.34	-4.99%	8.6	8.54	-0.73%	43.1	42.69	-0.96%	35.7	35.57	-0.36%
40	53.7	52.63	-2.0%	40.8	38.59	-5.41%	10.7	10.53	-1.63%	35.8	35.09	-2.0%	29.6	29.24	-1.22%
41	52.4	51.69	-1.36%	42.1	37.9	-9.97%	10.5	10.34	-1.55%	35.2	34.46	-2.11%	29.2	28.72	-1.66%
42	50.9	50.43	-0.92%	39.3	36.98	-5.9%	8.5	8.41	-1.12%	34.0	33.62	-1.12%	28.2	28.02	-0.65%
43	53.6	52.76	-1.57%	59.0	53.93	-8.59%	10.7	10.55	-1.39%	35.7	35.17	-1.48%	29.6	29.31	-0.98%
44	52.2	52.17	-0.07%	76.2	57.67	-24.32%	8.7	8.69	-0.07%	35.0	34.78	-0.64%	29.0	28.98	-0.07%
45	50.8	50.02	-1.53%	63.9	55.3	-13.46%	10.2	10.0	-1.92%	33.9	33.35	-1.63%	28.1	27.79	-1.1%

\* S= Simulation, A= Analytical

For most of the configurations the error percentages show a very good prediction capability from the analytical approach. In node 2 utilizations, the increase in the SRP node demand can be seen as the node utilization increases from one configuration set to the next. Despite the overall good results, some of the configurations show bigger error values. This is more pronounced when the demand for the SRP node is at the highest level

Tables 10.11, 10.12 and 10.13 show the average number of customers in a node results for the first stage, first-level nodes.

Table 10.11: Average Number in Node Results for Stage 1, Set of Configurations 1.

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
1	1.79	1.47	-1.57%	3.35	3.27	-0.4%	2.08	2.05	-0.15%	1.74	1.43	-1.53%	2.03	1.97	-0.3%
2	1.97	1.73	-1.21%	3.95	3.22	-3.65%	2.02	2.01	-0.05%	1.69	1.49	-1.0%	1.98	1.95	-0.17%
3	2.14	1.95	-1.28%	3.36	3.05	-2.07%	1.95	1.91	-0.27%	1.52	1.43	-0.6%	1.87	1.85	-0.14%
4	1.38	1.19	-1.98%	3.06	2.76	-3.02%	1.1	1.07	-0.31%	0.51	0.45	-0.62%	0.94	0.91	-0.25%
5	1.52	1.39	-1.25%	2.87	2.69	-1.8%	1.07	1.04	-0.29%	0.5	0.46	-0.36%	0.91	0.89	-0.19%
6	1.86	1.76	-0.8%	3.32	2.62	-5.81%	1.02	1.01	-0.09%	0.48	0.47	-0.09%	0.87	0.88	0.04%
7	1.84	1.63	-1.04%	2.14	2.78	3.18%	1.6	1.59	-0.05%	1.02	0.94	-0.39%	1.45	1.46	0.07%
8	2.07	1.92	-0.77%	2.85	2.81	-0.16%	1.58	1.57	-0.06%	1.01	0.95	-0.29%	1.43	1.45	0.12%
9	2.36	2.18	-1.19%	2.06	2.71	4.33%	1.53	1.51	-0.14%	0.97	0.92	-0.3%	1.38	1.4	0.14%
10	1.18	1.0	-1.81%	2.81	2.47	-3.41%	1.23	1.19	-0.38%	0.6	0.51	-0.84%	1.05	1.02	-0.27%
11	1.29	1.17	-1.21%	2.71	2.41	-2.9%	1.19	1.16	-0.26%	0.58	0.53	-0.49%	1.02	1.0	-0.19%
12	1.59	1.48	-0.87%	2.9	2.35	-4.6%	1.15	1.13	-0.17%	0.57	0.55	-0.14%	0.99	0.99	0.06%
13	1.21	1.11	-0.95%	2.01	2.0	-0.12%	1.23	1.2	-0.33%	0.61	0.57	-0.47%	1.05	1.04	-0.16%
14	1.36	1.3	-0.48%	2.67	2.13	-4.45%	1.19	1.18	-0.05%	0.61	0.58	-0.29%	1.03	1.03	0.03%
15	1.51	1.43	-0.81%	1.88	2.01	1.28%	1.14	1.12	-0.19%	0.57	0.55	-0.26%	0.98	0.98	-0.02%

\* S= Simulation, A= Analytical

Table 10.12: Average Number in Node Results for Stage 1, Set of Configurations 2.

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
16	1.59	1.32	-1.36%	3.42	3.26	-0.79%	1.98	1.97	-0.07%	1.55	1.29	-1.31%	1.91	1.86	-0.23%
17	1.76	1.55	-1.05%	3.88	3.22	-3.3%	1.94	1.93	-0.04%	1.51	1.34	-0.83%	1.86	1.84	-0.12%
18	1.95	1.78	-1.13%	3.56	3.07	-3.24%	1.88	1.84	-0.21%	1.38	1.31	-0.46%	1.77	1.76	-0.07%
19	1.32	1.14	-1.87%	3.06	2.77	-2.85%	1.08	1.05	-0.32%	0.5	0.44	-0.59%	0.91	0.89	-0.2%
20	1.45	1.33	-1.19%	2.89	2.7	-1.87%	1.05	1.02	-0.3%	0.48	0.45	-0.34%	0.89	0.87	-0.15%
21	1.78	1.69	-0.77%	3.3	2.64	-5.48%	1.01	0.99	-0.14%	0.47	0.46	-0.08%	0.86	0.86	0.04%
22	1.49	1.34	-0.77%	2.04	2.61	2.82%	1.48	1.47	-0.04%	0.87	0.81	-0.26%	1.31	1.33	0.07%
23	1.69	1.58	-0.55%	2.42	2.66	1.21%	1.46	1.46	-0.01%	0.86	0.82	-0.17%	1.3	1.33	0.12%
24	2.0	1.85	-1.02%	1.98	2.61	4.19%	1.43	1.42	-0.06%	0.84	0.81	-0.18%	1.27	1.29	0.15%
25	1.07	0.9	-1.64%	2.69	2.42	-2.75%	1.16	1.13	-0.3%	0.55	0.48	-0.73%	0.98	0.96	-0.21%
26	1.17	1.06	-1.09%	2.67	2.37	-2.99%	1.13	1.11	-0.24%	0.54	0.5	-0.45%	0.96	0.95	-0.14%
27	1.43	1.33	-0.82%	2.74	2.31	-3.58%	1.09	1.08	-0.13%	0.53	0.51	-0.13%	0.93	0.94	0.06%
28	1.08	1.0	-0.88%	1.95	1.96	0.19%	1.16	1.13	-0.26%	0.56	0.52	-0.42%	0.99	0.98	-0.11%
29	1.2	1.16	-0.37%	2.65	2.17	-4.03%	1.13	1.12	-0.07%	0.56	0.53	-0.3%	0.97	0.97	0.05%
30	1.36	1.28	-0.75%	1.94	2.06	1.21%	1.09	1.07	-0.24%	0.53	0.51	-0.23%	0.93	0.93	-0.01%

\* S= Simulation, A= Analytical

Table 10.13: Average Number in Node Results for Stage 1, Set of Configurations 3

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
31	1.18	1.0	-0.91%	3.23	2.98	-1.24%	1.74	1.73	-0.03%	1.16	0.98	-0.87%	1.6	1.58	-0.1%
32	1.33	1.18	-0.73%	3.47	2.97	-2.52%	1.72	1.72	0.0%	1.15	1.04	-0.55%	1.59	1.58	-0.02%
33	1.55	1.42	-0.89%	3.5	2.88	-4.15%	1.7	1.67	-0.17%	1.1	1.06	-0.29%	1.55	1.55	0.05%
34	1.25	1.07	-1.8%	3.03	2.76	-2.74%	1.05	1.02	-0.32%	0.48	0.42	-0.55%	0.88	0.86	-0.18%
35	1.36	1.25	-1.14%	2.9	2.69	-2.06%	1.01	0.99	-0.19%	0.46	0.43	-0.31%	0.86	0.85	-0.13%
36	1.68	1.59	-0.75%	3.24	2.64	-5.01%	0.98	0.97	-0.14%	0.45	0.44	-0.06%	0.83	0.84	0.04%
37	1.12	1.02	-0.51%	1.91	2.26	1.76%	1.3	1.3	0.01%	0.68	0.65	-0.15%	1.14	1.15	0.07%
38	1.28	1.2	-0.36%	2.05	2.32	1.35%	1.3	1.3	-0.01%	0.69	0.67	-0.07%	1.13	1.16	0.11%
39	1.57	1.46	-0.76%	1.86	2.33	3.16%	1.29	1.28	-0.06%	0.69	0.68	-0.07%	1.12	1.15	0.16%
40	0.94	0.79	-1.44%	2.55	2.32	-2.27%	1.07	1.05	-0.17%	0.49	0.43	-0.62%	0.91	0.89	-0.13%
41	1.02	0.93	-0.92%	2.6	2.28	-3.17%	1.05	1.03	-0.16%	0.49	0.45	-0.41%	0.9	0.88	-0.12%
42	1.24	1.16	-0.66%	2.57	2.23	-2.87%	1.02	1.01	-0.09%	0.48	0.47	-0.12%	0.87	0.87	0.07%
43	0.95	0.87	-0.77%	1.87	1.88	0.1%	1.07	1.06	-0.15%	0.51	0.47	-0.34%	0.91	0.9	-0.05%
44	1.04	1.0	-0.29%	2.6	2.14	-3.82%	1.04	1.04	-0.01%	0.51	0.48	-0.25%	0.9	0.9	0.05%
45	1.19	1.12	-0.67%	1.98	2.05	0.63%	1.02	1.0	-0.2%	0.48	0.46	-0.2%	0.86	0.87	0.02%

\* S= Simulation, A= Analytical

The average number of customers per node shows, in general, a good prediction from the analytical results compared with simulation results. Most of the errors show small percentages, as can be seen in tables 10.11, 10.12 and 10.13.

## Second Stage Results

The results for the first-level nodes in the second stage are presented in this section. Tables 10.14, 10.15 and 10.16 show the node utilization results.

Table 10.14: Utilization Results for Stage 2, First Set of Configurations

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
1	69.1	68.03	-1.55%	54.9	53.29	-2.93%	10.4	10.2	-1.88%	69.1	68.03	-1.55%	57.2	56.69	-0.89%
2	67.4	66.1	-1.93%	58.1	53.98	-7.09%	13.5	13.22	-2.08%	67.8	66.1	-2.51%	56.2	55.08	-1.99%
3	65.0	63.08	-2.95%	61.4	53.62	-12.67%	9.8	9.46	-3.45%	65.3	63.08	-3.4%	54.1	52.57	-2.83%
4	66.1	62.74	-5.08%	47.1	44.73	-5.03%	11.0	10.46	-4.94%	36.8	34.85	-5.28%	30.5	29.05	-4.77%
5	64.0	61.52	-3.88%	47.5	45.0	-5.26%	8.9	8.54	-4.0%	35.8	34.18	-4.53%	29.6	28.48	-3.78%
6	61.5	58.97	-4.12%	47.0	44.23	-5.91%	10.3	9.83	-4.59%	34.3	32.76	-4.49%	28.4	27.3	-3.88%
7	69.5	68.72	-1.12%	71.2	68.72	-3.48%	8.0	7.93	-0.88%	53.4	52.86	-1.01%	44.2	44.05	-0.34%
8	68.1	67.37	-1.07%	75.8	70.83	-6.56%	10.5	10.37	-1.29%	52.7	51.83	-1.66%	43.7	43.19	-1.17%
9	66.3	65.05	-1.89%	63.5	61.71	-2.82%	7.7	7.5	-2.53%	51.1	50.04	-2.08%	42.4	41.7	-1.66%
10	61.3	58.54	-4.51%	60.9	57.24	-6.02%	12.3	11.71	-4.82%	40.9	39.02	-4.58%	33.9	32.52	-4.07%
11	59.6	57.49	-3.53%	64.8	58.77	-9.3%	9.9	9.58	-3.21%	40.0	38.33	-4.18%	33.1	31.94	-3.5%
12	57.4	55.61	-3.11%	54.2	51.91	-4.23%	11.5	11.12	-3.28%	38.4	37.08	-3.45%	31.8	30.9	-2.84%
13	61.3	58.56	-4.47%	44.8	41.64	-7.05%	12.3	11.71	-4.78%	40.9	39.04	-4.55%	33.9	32.53	-4.03%
14	59.6	57.75	-3.11%	54.6	44.75	-18.03%	11.9	11.55	-2.95%	39.9	38.5	-3.52%	33.1	32.08	-3.08%
15	57.3	55.15	-3.76%	41.3	38.76	-6.16%	9.5	9.19	-3.25%	38.3	36.76	-4.01%	31.7	30.64	-3.36%

\* S= Simulation, A= Analytical



Table 10.15: Utilization Results for Stage 2, Second Set of Configurations

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
16	66.2	65.3	-1.36%	53.2	51.15	-3.85%	9.9	9.79	-1.06%	66.2	65.3	-1.36%	54.8	54.42	-0.7%
17	64.7	63.7	-1.55%	57.2	52.02	-9.06%	12.9	12.74	-1.25%	65.1	63.7	-2.16%	53.9	53.08	-1.52%
18	62.7	60.99	-2.72%	62.0	51.84	-16.38%	9.4	9.15	-2.67%	62.9	60.99	-3.03%	52.1	50.83	-2.44%
19	64.8	61.64	-4.88%	46.2	43.94	-4.88%	10.8	10.27	-4.88%	36.0	34.24	-4.88%	29.8	28.54	-4.24%
20	62.8	60.39	-3.84%	46.8	44.17	-5.61%	8.7	8.39	-3.59%	35.1	33.55	-4.41%	29.0	27.96	-3.59%
21	60.4	58.05	-3.89%	46.5	43.54	-6.37%	10.1	9.68	-4.21%	33.7	32.25	-4.3%	27.9	26.88	-3.68%
22	64.3	63.71	-0.92%	66.7	63.71	-4.48%	7.4	7.35	-0.66%	49.4	49.01	-0.79%	41.0	40.84	-0.39%
23	63.3	62.91	-0.62%	71.7	66.13	-7.77%	9.7	9.68	-0.23%	49.0	48.39	-1.25%	40.6	40.33	-0.68%
24	62.0	61.17	-1.34%	59.7	58.03	-2.79%	7.2	7.06	-1.97%	47.8	47.05	-1.56%	39.6	39.21	-0.98%
25	57.9	55.72	-3.77%	57.9	54.48	-5.91%	11.6	11.14	-3.93%	38.6	37.15	-3.77%	32.0	30.95	-3.27%
26	56.4	54.78	-2.86%	62.7	56.0	-10.68%	9.4	9.13	-2.86%	37.8	36.52	-3.38%	31.3	30.44	-2.76%
27	54.5	53.09	-2.58%	51.8	49.55	-4.34%	10.9	10.62	-2.58%	36.4	35.39	-2.76%	30.2	29.5	-2.33%
28	57.9	55.71	-3.78%	42.7	39.62	-7.22%	11.6	11.14	-3.95%	38.6	37.14	-3.78%	32.0	30.95	-3.28%
29	56.3	54.99	-2.33%	54.2	42.62	-21.37%	11.2	11.0	-1.8%	37.7	36.66	-2.76%	31.2	30.55	-2.08%
30	54.3	52.63	-3.08%	39.7	36.99	-6.83%	9.0	8.77	-2.54%	36.3	35.09	-3.34%	30.1	29.24	-2.86%

\* S= Simulation, A= Analytical

Table 10.16: Utilization Results for Stage 2, Third Set of Configurations

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
30	54.3	52.63	-3.08%	39.7	36.99	-6.83%	9.0	8.77	-2.54%	36.3	35.09	-3.34%	30.1	29.24	-2.86%
31	58.2	57.72	-0.83%	47.2	45.21	-4.21%	8.7	8.66	-0.49%	58.3	57.72	-1.0%	48.3	48.1	-0.42%
32	57.5	57.04	-0.8%	51.8	46.58	-10.07%	11.5	11.41	-0.8%	57.8	57.04	-1.31%	47.9	47.53	-0.76%
33	56.4	55.44	-1.71%	57.9	47.12	-18.62%	8.5	8.31	-2.17%	56.6	55.44	-2.06%	46.9	46.2	-1.5%
34	62.8	59.99	-4.47%	44.9	42.77	-4.74%	10.5	10.0	-4.77%	34.9	33.33	-4.5%	28.9	27.77	-3.89%
35	60.8	58.75	-3.37%	45.6	42.98	-5.75%	8.4	8.16	-2.86%	34.0	32.64	-4.0%	28.2	27.2	-3.55%
36	58.9	56.81	-3.54%	45.6	42.61	-6.56%	9.8	9.47	-3.38%	32.8	31.56	-3.77%	27.2	26.3	-3.3%
37	56.8	56.38	-0.74%	59.3	56.38	-4.93%	6.5	6.5	0.08%	43.6	43.37	-0.53%	36.2	36.14	-0.17%
38	56.3	56.18	-0.22%	64.7	59.06	-8.72%	8.7	8.64	-0.66%	43.6	43.21	-0.89%	36.1	36.01	-0.25%
39	55.9	55.41	-0.88%	54.0	52.57	-2.66%	6.5	6.39	-1.64%	43.1	42.62	-1.11%	35.7	35.52	-0.51%
40	53.7	52.11	-2.96%	54.1	50.95	-5.82%	10.7	10.42	-2.6%	35.8	34.74	-2.96%	29.7	28.95	-2.52%
41	52.5	51.34	-2.21%	59.4	52.48	-11.65%	8.7	8.56	-1.65%	35.2	34.23	-2.76%	29.2	28.52	-2.32%
42	50.9	49.97	-1.83%	48.7	46.64	-4.24%	10.2	9.99	-2.02%	34.0	33.31	-2.02%	28.2	27.76	-1.56%
43	53.6	52.09	-2.82%	39.9	37.04	-7.17%	10.7	10.42	-2.64%	35.8	34.73	-3.0%	29.6	28.94	-2.24%
44	52.2	51.51	-1.33%	52.3	39.92	-23.67%	10.4	10.3	-0.95%	35.0	34.34	-1.89%	29.0	28.61	-1.33%
45	50.8	49.56	-2.44%	37.6	34.83	-7.37%	8.5	8.26	-2.83%	33.9	33.04	-2.54%	28.1	27.53	-2.02%

\* S= Simulation, A= Analytical

A review of the utilization results for stage 2 shows the overall prediction accuracy over the three sets and all the nodes is reasonably good with higher errors observed for the second node which is affected by the SRP system. These errors reflect the estimation errors in the throughput of both stages and in the mean equivalent service time at node 2.

Tables 10.17, 10.18 and 10.19 present the average number of customers per node in the stage 2 of the first-level, for the three configuration sets.

Table 10.17: Average Number in Node Results for Stage 2, Set of Configurations 1

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
1	1.71	1.45	-1.27%	3.37	3.25	-0.61%	2.08	2.04	-0.2%	1.71	1.42	-1.48%	2.02	1.96	-0.33%
2	1.79	1.62	-1.13%	3.63	3.3	-2.17%	2.02	1.98	-0.28%	1.59	1.4	-1.23%	1.95	1.89	-0.43%
3	2.3	1.97	-1.67%	4.12	3.32	-3.97%	1.96	1.89	-0.34%	1.57	1.42	-0.76%	1.89	1.84	-0.22%
4	1.36	1.13	-2.33%	2.85	2.7	-1.55%	1.1	1.05	-0.54%	0.51	0.44	-0.75%	0.94	0.89	-0.45%
5	1.59	1.37	-1.81%	2.9	2.72	-1.49%	1.07	1.03	-0.36%	0.5	0.46	-0.39%	0.91	0.88	-0.24%
6	1.73	1.6	-1.35%	2.88	2.68	-1.95%	1.03	0.98	-0.47%	0.47	0.45	-0.25%	0.87	0.85	-0.2%
7	1.85	1.62	-1.17%	3.5	3.04	-2.27%	1.6	1.59	-0.07%	1.02	0.94	-0.39%	1.45	1.45	0.04%
8	1.98	1.8	-1.19%	3.89	3.26	-4.17%	1.57	1.55	-0.13%	0.96	0.91	-0.36%	1.42	1.42	0.0%
9	2.61	2.25	-1.83%	3.05	2.76	-1.48%	1.54	1.5	-0.19%	0.99	0.93	-0.34%	1.39	1.4	0.06%
10	1.18	1.06	-1.23%	2.27	2.06	-2.16%	1.23	1.17	-0.59%	0.61	0.55	-0.61%	1.05	1.01	-0.39%
11	1.35	1.22	-1.04%	2.69	2.23	-3.9%	1.19	1.15	-0.32%	0.6	0.55	-0.45%	1.03	1.0	-0.22%
12	1.54	1.41	-1.32%	2.05	1.9	-1.5%	1.15	1.11	-0.38%	0.58	0.54	-0.36%	0.99	0.97	-0.13%
13	1.18	0.97	-2.14%	2.7	2.51	-1.99%	1.23	1.17	-0.59%	0.6	0.5	-0.98%	1.05	1.0	-0.46%
14	1.29	1.15	-1.36%	3.39	2.7	-6.92%	1.19	1.15	-0.35%	0.59	0.52	-0.64%	1.02	0.99	-0.3%
15	1.61	1.4	-1.73%	2.52	2.34	-1.53%	1.14	1.1	-0.31%	0.57	0.53	-0.33%	0.98	0.96	-0.15%

\* S= Simulation, A= Analytical

Table 10.18: Average Number in Node Results for Stage 2, Set of Configurations 2

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
16	1.53	1.31	-1.1%	3.26	3.11	-0.74%	1.98	1.96	-0.11%	1.53	1.28	-1.27%	1.9	1.85	-0.25%
17	1.62	1.47	-0.98%	3.58	3.17	-2.67%	1.94	1.91	-0.16%	1.45	1.29	-1.07%	1.84	1.8	-0.31%
18	2.1	1.81	-1.45%	4.2	3.2	-5.03%	1.88	1.83	-0.25%	1.42	1.31	-0.56%	1.79	1.76	-0.13%
19	1.31	1.09	-2.23%	2.79	2.65	-1.47%	1.08	1.03	-0.53%	0.5	0.43	-0.7%	0.91	0.87	-0.39%
20	1.52	1.32	-1.71%	2.86	2.67	-1.55%	1.04	1.01	-0.31%	0.49	0.44	-0.37%	0.89	0.86	-0.22%
21	1.68	1.54	-1.33%	2.84	2.64	-2.07%	1.01	0.97	-0.43%	0.46	0.44	-0.24%	0.85	0.84	-0.19%
22	1.49	1.33	-0.78%	2.98	2.6	-1.9%	1.48	1.47	-0.05%	0.87	0.81	-0.29%	1.32	1.32	0.04%
23	1.64	1.51	-0.86%	3.42	2.8	-4.15%	1.46	1.45	-0.02%	0.84	0.8	-0.25%	1.3	1.31	0.06%
24	2.16	1.91	-1.23%	2.66	2.43	-1.13%	1.44	1.41	-0.14%	0.87	0.82	-0.21%	1.28	1.3	0.11%
25	1.07	0.96	-1.04%	2.11	1.91	-1.97%	1.16	1.11	-0.46%	0.56	0.51	-0.51%	0.99	0.96	-0.29%
26	1.21	1.11	-0.82%	2.54	2.06	-4.02%	1.13	1.1	-0.27%	0.56	0.51	-0.36%	0.97	0.95	-0.14%
27	1.4	1.28	-1.15%	1.92	1.78	-1.41%	1.09	1.06	-0.28%	0.54	0.51	-0.28%	0.93	0.93	-0.08%
28	1.07	0.88	-1.88%	2.58	2.38	-1.94%	1.16	1.11	-0.46%	0.55	0.47	-0.85%	0.98	0.95	-0.36%
29	1.17	1.04	-1.22%	3.37	2.57	-8.03%	1.12	1.1	-0.2%	0.54	0.49	-0.53%	0.96	0.94	-0.18%
30	1.45	1.27	-1.46%	2.42	2.23	-1.58%	1.08	1.05	-0.23%	0.53	0.5	-0.25%	0.93	0.92	-0.1%

\* S= Simulation, A= Analytical

Table 10.19: Average Number in Node Results for Stage 2, Set of Configurations 3

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
31	1.14	0.99	-0.74%	2.87	2.73	-0.66%	1.74	1.73	-0.04%	1.15	0.98	-0.85%	1.6	1.58	-0.12%
32	1.26	1.15	-0.76%	3.2	2.82	-2.52%	1.73	1.71	-0.09%	1.13	1.01	-0.76%	1.58	1.56	-0.13%
33	1.62	1.45	-0.86%	3.83	2.88	-4.75%	1.7	1.66	-0.18%	1.12	1.06	-0.28%	1.55	1.56	0.03%
34	1.24	1.03	-2.06%	2.71	2.58	-1.38%	1.05	1.0	-0.5%	0.48	0.41	-0.65%	0.88	0.85	-0.35%
35	1.42	1.24	-1.51%	2.78	2.6	-1.51%	1.01	0.98	-0.24%	0.47	0.43	-0.33%	0.86	0.84	-0.21%
36	1.59	1.47	-1.23%	2.79	2.58	-2.07%	0.98	0.95	-0.33%	0.45	0.43	-0.2%	0.83	0.82	-0.16%
37	1.11	1.02	-0.43%	2.34	2.1	-1.19%	1.3	1.3	0.01%	0.69	0.66	-0.18%	1.14	1.15	0.05%
38	1.25	1.18	-0.5%	2.75	2.27	-3.22%	1.3	1.3	-0.06%	0.69	0.66	-0.16%	1.13	1.15	0.1%
39	1.65	1.51	-0.66%	2.17	2.04	-0.69%	1.3	1.28	-0.11%	0.71	0.69	-0.11%	1.13	1.16	0.13%
40	0.93	0.85	-0.82%	1.92	1.74	-1.75%	1.07	1.04	-0.28%	0.5	0.46	-0.39%	0.91	0.89	-0.19%
41	1.05	0.98	-0.65%	2.32	1.87	-3.8%	1.04	1.03	-0.14%	0.5	0.47	-0.28%	0.9	0.89	-0.1%
42	1.23	1.14	-0.92%	1.75	1.63	-1.25%	1.02	1.0	-0.21%	0.48	0.46	-0.21%	0.87	0.87	-0.01%
43	0.94	0.78	-1.57%	2.4	2.23	-1.79%	1.07	1.04	-0.28%	0.5	0.43	-0.7%	0.91	0.88	-0.22%
44	1.03	0.92	-1.03%	3.24	2.4	-8.4%	1.04	1.03	-0.1%	0.49	0.45	-0.41%	0.89	0.88	-0.09%
45	1.26	1.13	-1.13%	2.28	2.1	-1.56%	1.02	0.99	-0.24%	0.48	0.45	-0.18%	0.86	0.86	-0.04%

\* S= Simulation, A= Analytical

The overall results for the average number at a node are very good with few values surpassing the 10% of error. Again the estimation errors were expressed as a percentage of the network population following the approach recommended by Krishnamurthy and Suri [32].

The results presented so far confirm that our approach, which combines the solution approaches for the building blocks, performs well overall in predicting the first-level performance.

## Results for the Synchronization Nodes

The fork/join approach allow us to estimate synchronization nodes performance as well. The significance of these results for health case systems can be explained as follows. The average

number in the customer queue at the node before stage 1 represents the patients waiting to enter the hospital system. More importantly, the average number at the entry to stage 2, represents patients on hold between the two stages.

Tables 10.20, 10.21 and 10.22 show the average number of customers waiting at the synchronization stations. The metric  $Lq_{P,0}$  represents the average number of customers at the first synchronization station, that is customers waiting for entering the system. On the other hand, the metric  $Lq_{P,1}$  represents customers between stages, waiting for free passive resources at stage 2 in order to move on.

Table 10.20: Average Number in Node Results for Fork/Join nodes, Set of Configurations 1

Conf	$Lq_{P,0}$			$Lq_{P,1}$		
	S	A	Error	S	A	Error
1	0.09	1.71	-8.11%	0.05	1.68	-8.13%
2	0.24	1.95	-8.57%	0.68	3.9	-16.12%
3	1.02	3.02	-13.35%	0.17	1.42	-8.36%
4	0.54	1.85	-13.03%	0.23	1.61	-13.83%
5	0.54	1.84	-13.07%	0.11	1.08	-9.75%
6	0.6	1.59	-8.24%	0.68	2.21	-12.75%
7	0.1	0.94	-4.21%	0.14	1.02	-4.39%
8	0.26	1.14	-4.4%	0.85	2.63	-8.9%
9	0.64	1.94	-8.64%	0.15	0.85	-4.63%
10	0.48	1.72	-12.34%	0.26	1.37	-11.19%
11	0.49	1.72	-12.28%	0.18	0.94	-7.62%
12	0.4	1.4	-8.35%	0.51	1.66	-9.62%
13	0.51	1.48	-9.76%	0.27	1.57	-13.08%
14	0.63	1.24	-5.05%	0.78	2.18	-11.7%
15	0.68	1.55	-8.74%	0.18	1.01	-8.38%

\* S= Simulation, A= Analytical

Table 10.21: Average Number in Node Results for Fork/Join nodes, Set of Configurations 2

Conf	$Lq_{P,0}$			$Lq_{P,1}$		
	S	A	Error	S	A	Error
16	0.06	1.48	-7.14%	0.03	1.39	-6.82%
17	0.14	1.62	-7.42%	0.45	3.3	-14.27%
18	0.86	2.68	-12.12%	0.12	1.22	-7.31%
19	0.48	1.75	-12.7%	0.2	1.52	-13.24%
20	0.48	1.77	-12.89%	0.09	1.01	-9.23%
21	0.52	1.49	-8.16%	0.6	2.08	-12.26%
22	0.04	0.66	-3.1%	0.05	0.65	-3.01%
23	0.1	0.77	-3.36%	0.4	1.78	-6.88%
24	0.37	1.5	-7.53%	0.07	0.59	-3.46%
25	0.3	1.45	-11.46%	0.17	1.13	-9.62%
26	0.34	1.47	-11.25%	0.12	0.77	-6.44%
27	0.25	1.16	-7.57%	0.35	1.35	-8.37%
28	0.37	1.24	-8.7%	0.17	1.34	-11.65%
29	0.76	1.04	-2.36%	0.54	1.84	-10.8%
30	0.62	1.37	-7.47%	0.12	0.85	-7.32%

\* S= Simulation, A= Analytical

Table 10.22: Average Number in Node Results for Fork/Join nodes, Set of Configurations 3

Conf	$Lq_{P,0}$			$Lq_{P,1}$		
	S	A	Error	S	A	Error
31	0.01	0.9	-4.46%	0.0	0.81	-4.04%
32	0.03	0.98	-4.75%	0.12	2.08	-9.77%
33	0.4	1.85	-9.66%	0.04	0.8	-5.06%
34	0.4	1.62	-12.22%	0.15	1.39	-12.34%
35	0.4	1.64	-12.39%	0.07	0.91	-8.48%
36	0.42	1.37	-7.92%	0.51	1.91	-11.68%
37	0.01	0.36	-1.77%	0.01	0.33	-1.59%
38	0.02	0.42	-2.02%	0.11	0.98	-4.36%
39	0.14	0.97	-5.51%	0.02	0.33	-2.06%
40	0.17	1.17	-10.04%	0.1	0.88	-7.8%
41	0.22	1.21	-9.91%	0.07	0.58	-5.12%
42	0.13	0.92	-6.51%	0.21	1.04	-6.88%
43	0.24	0.99	-7.45%	0.1	1.08	-9.8%
44	0.88	0.83	0.39%	0.31	1.47	-9.66%
45	0.55	1.15	-6.02%	0.07	0.68	-6.08%

\* S= Simulation, A= Analytical

From previous chapters, we know that the average number in node results for the synchronization stations are not as accurate as the regular node estimations. Additionally, there are configurations where specific conditions make the configurations more challenging, for instance, configuration 44 where both stages use the SRP node 75% of the time.

### SRP Results

The SRP node results are of special interest since the special condition of receiving arrivals for 2 different sources has not been seen so far in other examples.

Tables 10.23, 10.24 and 10.25 show the utilization and average number in node results for



the SRP node.

Table 10.23: Results for the SRP node, Set of Configurations 1

Configuration	SRP Node Utilization			Average Num in Node		
	S	A	Error	S	A	Error
1	27.5	27.33	0.6%	0.37	0.36	0.11%
2	40.4	40.22	0.45%	0.65	0.63	0.2%
3	51.9	50.91	1.9%	1.01	0.97	0.28%
4	14.6	14.25	2.38%	0.17	0.16	0.06%
5	21.4	20.81	2.75%	0.27	0.26	0.11%
6	27.4	26.91	1.78%	0.35	0.34	0.07%
7	32.0	31.79	0.64%	0.41	0.41	0.04%
8	42.2	41.81	0.93%	0.58	0.6	-0.19%
9	20.4	20.12	1.38%	0.24	0.23	0.08%
10	24.5	23.85	2.67%	0.29	0.29	0.11%
11	31.9	31.05	2.66%	0.48	0.49	-0.06%
12	15.3	15.1	1.34%	0.19	0.19	0.06%
13	24.5	23.94	2.28%	0.35	0.33	0.27%
14	51.8	51.2	1.17%	1.17	1.11	0.71%
15	24.8	24.28	2.11%	0.36	0.34	0.17%

\* S= Simulation, A= Analytical

Table 10.24: Results for the SRP node, Set of Configurations 2

Configuration	SRP Node Utilization			Average Num in Node		
	S	A	Error	S	A	Error
16	39.5	39.31	0.47%	0.63	0.6	0.27%
17	51.7	51.51	0.37%	1.01	0.96	0.44%
18	62.5	61.44	1.7%	1.5	1.44	0.52%
19	21.5	20.96	2.5%	0.27	0.26	0.12%
20	27.9	27.21	2.49%	0.38	0.36	0.18%
21	33.7	33.04	1.94%	0.46	0.45	0.11%
22	39.5	39.26	0.62%	0.55	0.54	0.12%
23	49.0	48.62	0.78%	0.75	0.75	-0.1%
24	28.7	28.32	1.31%	0.36	0.35	0.07%
25	30.9	30.13	2.48%	0.4	0.38	0.25%
26	37.7	36.88	2.19%	0.65	0.63	0.15%
27	21.8	21.52	1.28%	0.31	0.29	0.15%
28	30.8	30.23	1.84%	0.48	0.46	0.28%
29	61.1	60.63	0.77%	1.59	1.58	0.09%
30	35.3	34.64	1.87%	0.58	0.58	-0.04%

\* S= Simulation, A= Analytical

Table 10.25: Results for the SRP node, Set of Configurations 3

Configuration	SRP Node Utilization			Average Num in Node		
	S	A	Error	S	A	Error
31	46.3	46.25	0.11%	0.82	0.76	0.45%
32	57.4	57.35	0.09%	1.25	1.18	0.62%
33	67.6	66.81	1.17%	1.85	1.77	0.63%
34	27.7	27.14	2.02%	0.38	0.35	0.21%
35	33.8	33.03	2.28%	0.5	0.47	0.25%
36	39.4	38.7	1.77%	0.58	0.57	0.12%
37	43.7	43.39	0.71%	0.63	0.62	0.21%
38	52.3	51.96	0.66%	0.84	0.84	0.02%
39	34.5	34.15	1.02%	0.45	0.46	-0.03%
40	35.8	35.09	2.0%	0.49	0.45	0.43%
41	42.1	41.35	1.78%	0.79	0.76	0.38%
42	27.2	26.9	1.12%	0.42	0.4	0.22%
43	35.7	35.17	1.48%	0.59	0.58	0.18%
44	68.1	67.81	0.42%	1.93	2.08	-1.66%
45	44.0	43.35	1.47%	0.78	0.85	-0.73%

\* S= Simulation, A= Analytical

The results are in general very good, the SRP node utilization and average number in node performance metrics are predicted with very good accuracy for all 45 configurations.

In general, the ideas discussed at the end of chapter IX about the factors influencing the approach's accuracy, apply here as well, mainly because Suri and Krishnamurthy's [32] approach is a key building block for both solutions. However, in this case, there are two systems demanding service from the SRP system and their combined effect needs to be considered. For instance, it could be seen that, in general, configuration 44 corresponds to the highest error among all the configurations, and it is consistent with results discussed previously, since both systems demand service from the SRP system with a probability of 75%, implying a big contention effect for the SRP system that has an impact on the overall

system. Additionally, both systems has relatively small number of passive resources (12 and 10) in this configuration. Both conditions, as discussed in previous chapters, are factors that could have and adverse effect on the performance prediction accuracy.

In this chapter, we considered the most complex system in this study with a two capacity-restricted stages in the first level and each sharing the same SRP system at the second level. The successful performance prediction for this system has validated our building block approach.

## CHAPTER XI

### CONCLUSIONS AND FUTURE WORK

The problems studied in this dissertation were motivated by the health care environment. In hospital systems, patients typically need to acquire a passive resource such as a bed or a room before receiving service and hold it as long as they are in the system. Passive resources can impose capacity restrictions on systems by limiting the number of customers that can be simultaneously present in the system. Sometimes customers need to visit only one stage (section controlled by only one set of capacity restricting set of resources), and at other times customers need to go through two or more stages. Physicians, nurses, technicians, and specialists are active resources that provide service or care. Scarce and expensive active resources are typically shared by multiple operations in the hospital. Sometimes service operations require these scarce and costly resources in an concurrent manner. The challenge then becomes combining all these process characteristics, or different combinations of them, in a single analytical model and figuring out how to solve it.

The overall approach we adopted was to solve one problem at a time, and doing so in a manner which will allow them to be used as building blocks to solve more complex problems and a wider variety of problems in a modular fashion. Since it is expected that the problems' solution approaches to be compatible and exchangeable, the proposed solutions must be generated under the same general approach and the parametric decomposition methodology proposed by Whitt [54] was selected for this purpose.

In health care problems, the process time variability is a key determinant in the success of care or service. Consequently, the parametric decomposition methodology's ability to include variability as a parameter for solving the problems using two-moment approximations is a key advantage.

The first building block we considered was the simultaneous resource possession problem that was introduced in Chapter IV. In order to calculate the SRP node response time, a sub-system was established that emulates the SRP node conditions, which includes determining the arrival process characteristics.

Based on SRP subsystem study, there are two characteristics in the approach that can be considered key. First, there is only a limited number of customers that can be in the SRP system at any single time. This characteristic makes the SRP subsystem to work like a CQN where the number of customers is defined by the number of servers in the first level nodes affected by the SRP node. Secondly, the SRP arrival process characteristics depend on the departure rate and variability of the departure process of the node(s) in the first-level affected by the SRP system and the probability of a customer at this node requiring additional service from the SRP sub-system.

The influence of the SRP system on the first-level nodes affected is captured by modifying the service times at these nodes. The mean and variance of the SRP response time is needed in order to obtain the parameters for a equivalent node in the first level system that emulates the combined first and second level nodes' behavior. With this information, a new node's service time and SCV are included in the first level system that is solved to obtain the overall system performance metrics. We developed an iterative algorithm to capture the SRP system influence on the first-level nodes affected.

The CQN system was the second building block studied. The PDMCQN-2 algorithm presented by Satyam et al.[42] was utilized for solving the CQN version of the problem, and it was adapted in two ways. First, the approach for multi-servers nodes using the parametric decomposition method proposed by Whitt [56] was used in order to compute the mean and variance of the waiting time corresponding to the multi-servers nodes. Secondly, a correction for the waiting time in queue for CQN nodes with multi-servers was introduced. The correction is a combination of a correction introduced by Suri et al. [50] for multi-server nodes' waiting time in a CQN and a modification of the single server node correction in CQN environments presented by Kamath et al. [29] for its use in multi-server nodes. The development of this correction was a key step in adequately modeling the waiting time in queue for multi-server systems in capacity restricted environments.

Despite the fact the system under study has an external arrival process, the CQN system solution is a key building block for two reasons. Firstly, the CQN system allow us to identify the maximum throughput the capacity restricted systems are capable of producing. Identifying the maximum throughput that can be handled is vital in determining if the system can handle the arrival rate specified or not. Secondly, the CQN mode can be directly used when the resources limiting the capacity are heavily utilized. Freed passive resources will almost every single time find customers waiting for them when they finish their path with a customer. As a result, they will work in a fashion similar to customers in a CQN arrangement.

The fork/join system was the third building block. The approach proposed by Krishnamurthy and Suri [32] to solve a capacity restricted system with two synchronization stations was the starting point. One at the beginning that allows the external arrival of customers to wait for passive resources in order to gain entry into the system. The second synchroniza-

tion station allows the customers wait for downstream resources to gain the access to the downstream stages. This building block is useful for connecting different stages and studying systems where the capacity restricting resources are idle with more regularity.

The second synchronization station is not needed for the stage where the customers finish their service in a service system. In Chapter VI with some simple modifications, we showed how to deal with a single synchronization station system. Additional modification to the fork/join algorithm included better approximations for single-server nodes and the approximations for multi-server nodes with the new correction factor.

The last building block presented in this document corresponds to the two-stage fork/join system. Two stages with individual capacity restrictions are connected by a synchronization node. The solution approach is based on Krishnamurthy and Suri [32] work on the single-stage system. The only implementation for a multi-stage system was in the context of a multi-product, kanban-based pull system [44, 43]. The method relies heavily on the demand process that pulls items from the finished goods store. Consequently, the algorithm for solving the two-stage system in a service system context is an additional contribution of this study.

The general idea, as explained previously, was to solve the building blocks in a modular fashion, and to combine them to generate solutions to a wider variety of problems. In Chapter VIII, the CQN and SRP solutions building blocks were combined. The proposed solution approach took advantage of the CQN algorithm structure to embed the SRP solution in the key iterative step in the CQN solution.

In Chapter IX we showed how the SRP solution can be integrated in the solution of single and double fork/join systems. The linkage is achieved without altering the basic structure of



the fork/join algorithm.

Extensive numerical experimentation was performed to test the building blocks as well as their combinations. The configurations included a mix of single-server, multi-server, and delay nodes, different levels of variability - from low ( $SCV=0.25$ ) to high ( $SCV=2$ ), and different capacity limits. This was necessary to stress test the solution approaches.

The last problem solved included the two-stage problem with a SRP system that is shared by two first level nodes, one in the first stage and one in the second stage. Additional steps were needed to join arrivals from the two sources to the SRP system. The solution approach for this problem changed from the approach used in the CQN-SRP and FJ-SRP systems solution approaches, given the SRP solution now depends on two different system conditions (stage 1 and stage 2), the option of embedding the SRP solution became very difficult. Consequently, the solution approach iterates between solving the first level system, collecting the required results, solving the second level system, and passing the required solutions back to the first level. In an iterative fashion, this information exchange follows until the difference in the variability in the arrival to the SRP system between iterations goes below some threshold value.

Another important remark is that the solution approach used for the two stages with a SRP system that includes iterating between the first and second level solutions instead of solving the SRP system embedded in the first level solution, can be used in the CQN-SRP and FJ-SRP solutions in most of the cases. It was observed that when the embedded approach was used to solve the SRP system, there was more stability in the arrival rate used to solve the SRP system and this brought stability to the algorithm.

## 11.1 Accuracy of Performance Prediction

The main approach developed in this study is concerned with the use of building blocks as a modular way to consider a wider variety of problems. As a result, the prediction accuracy for each system solution will be highly related with the building blocks used to solve each specific problem. Nevertheless, all the systems used are capacity restricted systems, and accordingly, they share some characteristics that need to be taken into consideration as well. Two key building blocks were used to solve the first-level system, and their characteristics and interactions with the second-level system determine the performance prediction accuracy.

In the CQN with SRP system, the variability in the service time of the node influenced by the SRP system and the number of servers in this node, have a significant effect on the solution accuracy. Specifically, the higher the service time variability and smaller the number of servers, greater is the impact on the solution accuracy.

For the single stage FJ with SRP system, the number of passive resources is a key factor to take into account when analyzing the system. When the number of passive resources decreases, the solution accuracy decreases. This effect is more pronounced when the node influenced by the SRP system has many servers and these servers in their numbers represent a significant proportion of the overall number of passive resources. This effect become more important as the second node utilization increased influenced by the increased SRP participation in the service process.

In the case of the double stage FJ with SRP system, the conditions discussed in the previous paragraph remain valid, since Suri and Krishnamurthy's [32] approach is a key part of both solutions. However, in the two-stage case, these factors must be considered more

cautiously, since there are two stages, and both interacting with the SRP node. In spite of the robustness of the approach, the high demand for the SRP node and a small number of passive resources in configuration 44, for instance, made it very challenging among all the configurations tested.

In summary, the overall solution algorithm involves several layers of approximations starting with the solution of the SRP system, effect of the SRP system on the first-level node's equivalent service time and the solution of the first-level network. In spite of this stacking of approximations, the performance prediction in the worst cases can still be considered as being in the acceptable range for planning purposes.

## 11.2 Contributions

The main contribution of this dissertation effort is a building block approach using the parametric decomposition methodology for solving capacity restricted problems that includes simultaneous resource possession instances. However, in the journey to achieve this dissertation's goals, a series of small contributions were made, as shown below.

- A methodology for solving general SRP systems, including arrivals from single and multiple sources.
- Extension of the PDMCQN-2 algorithm presented by Satyam et al [42] for solving CQNs and FJ algorithm presented by Krishnamurthy and Suri [32] to include multi-server nodes.
- Modification of FJ algorithm presented by Krishnamurthy and Suri [32] for systems with only one synchronization station at the beginning that allows it to work under push system conditions.

- Extension of FJ algorithm presented by Krishnamurthy and Suri [32] to a two-stage system in a service environment.
- A new hybrid correction factor that improves the waiting time approximations for multi-server nodes in CQN systems.
- Two different approaches for combining the capacity restricting system solutions with the SRP system solution.

### 11.3 Future Work

There are some extensions to the building blocks presented in this document that can be explored in the future. For instance, different routes, more complex than the tandem or sequential used in this study, including routing inside a stage or different routes between stages. Another natural extension is the inclusion of different customer types, each with its own SRP requirement and service time characteristics. These two extensions would enrich the practical applications of the models presented.

The waiting time computation under CQN conditions is still a topic that can be improved for multi-server nodes. Despite the improvement realized in this work, there is still scope for additional work.

There are some additional building blocks that could be developed in order to expand the overall modeling capabilities. Some ideas are listed below:

1. Consider variations to the fork/join representation in order to model patient/customer exit points in different sections of the process. How this modification influences the flow in the capacity restricted network and in the SRP system needs to be addressed.

2. Similarly, external customers arrivals at different entry points, for instance, at the synchronization stations between stages can be considered.
3. Consider other variations in the fork/join-SRP model. For instance, consider a third resource that is shared by two different stages.

Some of these variations can be seen in Figure 11.1. The customer, before accessing the system, requires two passive resources, one from the first sub-system and the second one from the second sub-system. In figure 11.1 there is an external arrival of customers to the second sub-system, that could make the solution more challenging.

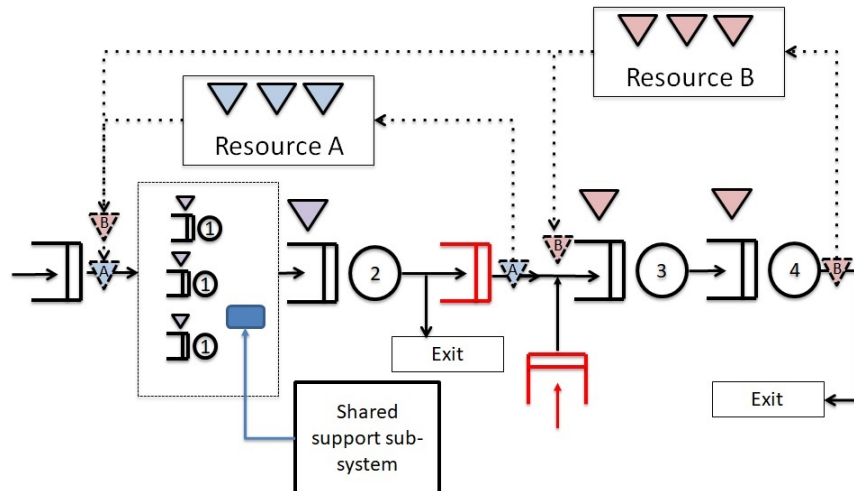


Figure 11.1: Some Possible Problem Extension

Another possible research venue is related with the effect of the SRP subsystem in open or non-restricted environments. Simulation results for open network configurations have revealed an additional correlation effect introduced by the presence of the SRP node, which could severely impact the first-level node's waiting time. However, the correlation effect's magnitude and behavior are not well known.

Some research in the planning and management of health care system under different circumstances can be done based on the models developed and solved in this study. For instance, the ER departments and ICU departments in hospitals under the COVID-19 pandemic, are examples of capacity restricted stages used by patients affected by the novel corona virus. Hospital/department preparedness to face different pandemic conditions, the congestion effects produced by shared resources or other process conditions could be of interest to plan the resources' capacity and assignment.

Even under regular circumstances, proper capacity management in health care organizations is crucial. Patients with life threatening conditions can find their pathway through the hospital blocked because some resource is saturated and cannot serve the patient at the required time. The effect of different resources over the patient experience and the response time must be evaluated and managed. For instance, how the variability in the second level resources' service time affects the overall time the patient is in the system. What are the effects of different resources such as specialist, technicians, equipment, or even bed carriers on the time-to-doctor, cycle time, or other metrics?

If the health care facilities are under extreme circumstances, such as the pandemic conditions that were present around the world in 2020 caused by the COVID-19, natural disasters or war conditions, proper capacity planning in hospitals may have a significant impact on the well-being of the general population. The outcome from the event (war, pandemic, or natural disaster) could be drastically different if the right decisions are made about capacity of passive resources, specialists, equipment, etc.

## BIBLIOGRAPHY

- [1] E. Akcali, M.J. Côté, and C. Lin. A network flow approach to optimizing hospital bed capacity decisions. *Health Care Management Science*, 9(4):391–404, 2006.
- [2] D. Ansaloni, L.Y. Chen, E. Smirni, and W. Binder. Model-driven consolidation of java workloads on multicores. In *Dependable Systems and Networks (DSN), 2012 42nd Annual IEEE/IFIP International Conference on*, pages 1–12. IEEE, 2012.
- [3] D.A. Bacigalupo, J. Van Hemert, X. Chen, A. Usmani, A.P. Chester, L. He, D.N. Dillenberger, G.B. Wills, L. Gilbert, and S.A. Jarvis. Managing dynamic enterprise and urgent workloads on clouds using layered queuing and historical performance models. *Simulation Modelling Practice and Theory*, 19(6):1479–1495, 2011.
- [4] S. Barbagallo, L. Corradi, J.V. de Goyet, M. Iannucci, I. Porro, N. Rosso, E. Tanfani, and A. Testi. Optimization and planning of operating theatre activities: an original definition of pathways and process modeling. *BMC medical informatics and decision making*, 15(1):38, 2015.
- [5] Forest Baskett, K Mani Chandy, Richard R Muntz, and Fernando G Palacios. Open, closed, and mixed networks of queues with different classes of customers. *Journal of the ACM (JACM)*, 22(2):248–260, 1975.
- [6] C. Basnet and M. Kamath. A two-moment queueing network model for flexible manufacturing systems with transport vehicles. In *Proc. of the First Joint Conf. of the OR Society of New Zealand and the New Zealand Production & Inventory Control Soc*, pages 7–12, 1991.

- [7] B. Baynat and Y. Dallery. Approximate techniques for general closed queueing networks with subnetworks having population constraints. *European Journal of Operational Research*, 69(2):250–264, 1993.
- [8] R. Buitenhek, G.J. van Houtum, and H. Zijm. Amva-based solution procedures for open queueing networks with population constraints. *Annals of Operations Research*, 93(1-4):15–40, 2000.
- [9] B. Cardoen, E. Demeulemeester, and J. Beliën. Operating room planning and scheduling: A literature review. *European journal of operational research*, 201(3):921–932, 2010.
- [10] K.M. Chandy, U. Herzog, and L. Woo. Parametric analysis of queueing networks. *IBM Journal of Research and Development*, 19(1):36–42, 1975.
- [11] J.K. Cochran and K.T. Roche. A multi-class queueing network analysis methodology for improving hospital emergency department performance. *Computers & Operations Research*, 36(5):1497–1512, 2009.
- [12] Y. Dallery. Approximate analysis of general open queueing networks with restricted capacity. *Performance Evaluation*, 11(3):209–222, 1990.
- [13] D. De Almeida and P. Kellert. An analytical queueing network model for flexible manufacturing systems with a discrete handling device and transfer blockings. *International Journal of Flexible Manufacturing Systems*, 12(1):25–57, 2000.
- [14] A.M. De Bruin, A.C. Van Rossum, M.C. Visser, and G.M. Koole. Modeling the emergency cardiac in-patient flow: an application of queueing theory. *Health Care Management Science*, 10(2):125–137, 2007.
- [15] E. de Souza e Silva and R.R. Muntz. Approximate solutions for a class of non-product form queueing network models. *Performance Evaluation*, 7(3):221–242, 1987.



- [16] J.L. Dorsman, O.J. Boxma, and M. Vlasiou. Marginal queue length approximations for a two-layered network with correlated queues. *Queueing Systems*, 75(1):29–63, 2013.
- [17] J.L. Dorsman, R.D. van der Mei, and M. Vlasiou. Analysis of a two-layered network by means of the power-series algorithm. *Performance Evaluation*, 70(12):1072–1089, 2013.
- [18] J.P.L. Dorsman, S. Bhulai, and M. Vlasiou. Dynamic server assignment in an extended machine-repair model. *IIE Transactions*, 47(4):392–413, 2015.
- [19] G. Franks, T. Al-Omari, M. Woodside, O. Das, and S. Derisavi. Enhanced modeling and solution of layered queueing networks. *IEEE Transactions on Software Engineering*, 35(2):148–161, 2009.
- [20] G. Franks, A. Hubbard, S. Majumdar, J. Neilson, D. Petriu, J. Rolia, and M. Woodside. A toolset for performance engineering and software design of client-server systems. *Performance evaluation*, 24(1-2):117–136, 1995.
- [21] G. Franks and M. Woodside. Multiclass multiservers with deferred operations in layered queueing networks, with software system applications. In *The IEEE Computer Society’s 12th Annual International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunications Systems, 2004.(MASCOTS 2004). Proceedings.*, pages 239–248. IEEE, 2004.
- [22] Matthew Gombolay, Toni Golen, Neel Shah, and Julie Shah. Queueing theoretic analysis of labor and delivery. *Health care management science*, 22(1):16–33, 2019.
- [23] L. Haque and M.J. Armstrong. A survey of the machine interference problem. *European Journal of Operational Research*, 179(2):469–482, 2007.
- [24] P.J.H. Hulshof, N. Kortbeek, R.J. Boucherie, E.W. Hans, and P.J.M. Bakker. Taxonomic

- classification of planning decisions in health care: a structured review of the state of the art in or/ms. *Health systems*, 1(2):129–175, 2012.
- [25] J.R. Jackson. Networks of waiting lines. *Operations research*, 5(4):518–521, 1957.
- [26] P.A. Jacobson and E.D. Lazowska. Analyzing queueing networks with simultaneous resource possession. *Communications of the ACM*, 25(2):142–151, 1982.
- [27] J. Jia and S.S. Heragu. Solving semi-open queueing networks. *Operations research*, 57(2):391–401, 2009.
- [28] M. Kamath and J.L. Sanders. Modeling operator/workstation interference in asynchronous automatic assembly systems. *Discrete Event Dynamic Systems*, 1(1):93–124, 1991.
- [29] Manjunath Kamath, Rajan Suri, and Jerry L Sanders. Analytical performance models for closed-loop flexible assembly systems. *International Journal of Flexible Manufacturing Systems*, 1(1):51–84, 1988.
- [30] J.S. Kaufman and W.S. Wong. Approximate analysis of a cyclic queueing network with applications to a simultaneous resource possession problem. *Performance Evaluation*, 11(3):187–200, 1990.
- [31] Wolfgang Krämer and M Langenbach-Belz. Approximate formulae for the delay in the queueing system  $gi/g/l$ . *Congressbook, 8th ITC, Melbourne*, pages 235–1, 1976.
- [32] A. Krishnamurthy and R. Suri. Performance analysis of single stage kanban controlled production systems using parametric decomposition. *Queueing Systems*, 54(2):141–162, 2006.

- [33] A. Krishnamurthy, R. Suri, and M. Vernon. Analysis of a fork/join synchronization station with inputs from coxian servers in a closed queuing network. *Annals of Operations Research*, 125(1-4):69–94, 2004.
- [34] T. Kurasugi and I. Kino. Approximation methods for two-layer queueing models. *Performance Evaluation*, 36:55–70, 1999.
- [35] R.A. Marie. An approximate analytical method for general queueing networks. *IEEE Transactions on Software Engineering*, (5):530–538, 1979.
- [36] T. Omari, S. Derisavi, and G. Franks. Deriving distribution of thread service time in layered queueing networks. In *Proceedings of the 6th International Workshop on Software and Performance*, pages 66–77. ACM, 2007.
- [37] Y.A. Ozcan, E. Tànfani, and A. Testi. Improving the performance of surgery-based clinical pathways: a simulation-optimization approach. *Health care management science*, 20(1):1–15, 2017.
- [38] F.M. Pajouh and M. Kamath. Applications of queueing models in hospitals. *Midwest Association for Information (MWAIS) 2010*, 2010.
- [39] Ram Ramakrishnan and Ananth Krishnamurthy. Performance evaluation of a synchronization station with multiple inputs and population constraints. *Computers & operations research*, 39(3):560–570, 2012.
- [40] J.A. Rolia and K.C. Sevcik. The method of layers. *IEEE transactions on software engineering*, 21(8):689–700, 1995.
- [41] D. Roy. Semi-open queuing networks: a review of stochastic models, solution methods and new research areas. *International Journal of Production Research*, 54(6):1735–1752, 2016.

- [42] K. Satyam, A. Krishnamurthy, and M. Kamath. Solving general multi-class closed queuing networks using parametric decomposition. *Computers & Operations Research*, 40(7):1777–1789, 2013.
- [43] Kumar Satyam. *Analytical models for multi-product systems with pull type material control strategies*. ProQuest, 2007.
- [44] Kumar Satyam and Ananth Krishnamurthy. Effect of product variety and loop length in multi-stage pull systems. In *IIE Annual Conference. Proceedings*, page 1284. Institute of Industrial and Systems Engineers (IISE), 2007.
- [45] C.H. Sauer. Approximate solution of queueing networks with simultaneous resource possession. *IBM Journal of Research and Development*, 25(6):894–903, 1981.
- [46] C.H. Sauer and E.A. MacNair. Simultaneous resource possession in queueing models of computers. *ACM SIGMETRICS Performance Evaluation Review*, 7(1-2):41–52, 1978.
- [47] Y. Shoaib and O. Das. Web application performance modeling using layered queueing networks. *Electronic notes in theoretical computer science*, 275:123–142, 2011.
- [48] K.E. Stecke and J.E. Aronson. Review of operator&sol; machine interference models. *International Journal of Production Research*, 23(1):129–151, 1985.
- [49] R. Suri and R. Desiraju. Performance analysis of flexible manufacturing systems with a single discrete material-handling device. *International Journal of Flexible Manufacturing Systems*, 9(3):223–249, 1997.
- [50] Rajan Suri, Sushanta Sahu, and Mary Vernon. Approximate mean value analysis for closed queuing networks with multiple-server stations. In *Proceedings of the 2007 Industrial Engineering Research Conference*, pages 1–6, 2007.

- [51] Kuo-Hsiung Wang, Cheng-Dar Liou, and Yu-Hsueh Lin. Comparative analysis of the machine repair problem with imperfect coverage and service pressure condition. *Applied Mathematical Modelling*, 37(5):2870–2880, 2013.
- [52] Peter D Welch. On the problem of the initial transient in steady-state simulation. *IBM Watson Research Center*, 1981.
- [53] Peter D Welch. The statistical analysis of simulation results. *The computer performance modeling handbook*, 22:268–328, 1983.
- [54] W. Whitt. The queueing network analyzer. *Bell System Technical Journal*, 62(9):2779–2815, 1983.
- [55] W. Whitt. Towards better multi-class parametric-decomposition approximations for open queueing networks. *Annals of Operations Research*, 48(3):221–248, 1994.
- [56] Ward Whitt. Approximations for the gi/g/m queue. *Production and Operations Management*, 2(2):114–161, 1993.
- [57] C.M. Woodside. Throughput calculation for basic stochastic rendezvous networks. *Performance Evaluation*, 9(2):143–160, 1989.
- [58] M. Woodside, J.E. Neilson, D.C. Petriu, and S. Majumdar. The stochastic rendezvous network model for performance of synchronous client-server-like distributed software. *IEEE Transactions on Computers*, 44(1):20–34, 1995.
- [59] H. Xie, T. Chausalet, and M. Rees. A semi-open queueing network approach to the analysis of patient flow in healthcare systems. In *Twentieth IEEE International Symposium on Computer-Based Medical Systems (CBMS'07)*, pages 719–724. IEEE, 2007.

- [60] M.E. Zonderland, R.J. Boucherie, N. Litvak, and C.L. Vleggeert-Lankamp. Planning and scheduling of semi-urgent surgeries. *Health Care Management Science*, 13(3):256–267, 2010.

## APPENDIX A

### Verification of Fork/Join Algorithm Implementation

The implementation of Krishnamurthy and Suri's [32] fork/join algorithms in this study included some modifications and enhancements listed below.

- Correction factor Equation 4.11 for mean waiting time in a GI/G/1 queue
- Extension to include multi-server nodes
- Modifications for a single FJ to model service systems.

In order to verify the implementation of the original algorithm, the experiments conducted by Krishnamurthy and Suri [32] and shown in their Table 6 were replicated. In these experiments the authors tested their algorithm's performance in predicting throughput and mean queue length measurements. They used system configurations where the node service time and SCV varies from one node to the other.

The experiments presented in Table 6 in Krishnamurthy and Suri [32] were replicated and the results are shown in Tables A.1 and A.2

Table A.1: Double Fork/Join Algorithm Verification for  $\lambda_D$  Results

Configuration	K&S Results	Simulation	Error	New Results	Error
1	0.504	0.490	2.85%	0.488	-0.4%
2	0.648	0.629	3.02%	0.629	0.0%
3	0.713	0.695	2.59%	0.695	0.04%
4	0.749	0.734	2.04%	0.733	-0.12%
5	0.785	0.776	1.16%	0.774	-0.3%
6	0.497	0.489	1.64%	0.488	-0.2%
7	0.639	0.628	1.75%	0.629	0.16%
8	0.704	0.695	1.29%	0.695	0.04%
9	0.741	0.733	1.09%	0.733	0.01%
10	0.779	0.777	0.26%	0.774	-0.43%
11	0.491	0.489	0.41%	0.488	-0.2%
12	0.629	0.628	0.16%	0.629	0.16%
13	0.695	0.695	0.0%	0.695	0.04%
14	0.733	0.735	-0.27%	0.733	-0.26%
15	0.773	0.777	-0.51%	0.774	-0.43%

The throughput computations show significant improvement in the results with the modifications in our implementation. Actually, in fifteen experiments, our implementation produced better results for twelve cases compared to the original algorithm.

Table A.2 presents the results for the average number in node measure. It is important to remember that the error percentage was computed in proportion to the number of passive resources in the system.



Table A.2: Double Fork/Join Algorithm Verification for  $L_3$  Results

Configurations	K&S Results	Simulation	Error	New Results	Error
1	1.005	0.989	0.320%	1.008	0.373%
2	2.047	1.947	0.667%	1.976	0.190%
3	3.182	2.947	1.567%	2.959	0.078%
4	4.415	3.972	2.215%	3.992	0.100%
5	7.083	6.232	2.837%	6.222	-0.034%
6	1.029	1.013	0.320%	1.032	0.379%
7	2.116	2.025	0.607%	2.104	0.528%
8	3.252	3.094	1.053%	3.203	0.729%
9	4.434	4.199	1.175%	4.327	0.642%
10	6.922	6.618	1.013%	6.672	0.181%
11	1.074	1.096	-0.44%	1.066	-0.597%
12	2.368	2.354	0.093%	2.343	-0.076%
13	3.842	3.720	0.813%	3.808	0.586%
14	5.445	5.149	1.480%	5.411	1.311%
15	8.940	8.294	2.153%	8.899	2.017%

In predicting the average number at a node, our implementation produced better results for twelve out of the fifteen cases.

The numerical results confirm the correct implementation of the fork/join algorithm and the improvement obtained as result of the Kramer and Lagenbach-Belz [31] correction factor for the waiting time in queue.

## APPENDIX B

### Additional Results for the Two Capacity-Restricted Stages with a Shared SRP System

Table B.1: Throughput Results for First Set of Configurations

Conf	Stage 1			Stage 2			SRP Subsystem		
	Simulation	Analytical	Difference	Simulation	Analytical	Difference	Simulation	Analytical	Difference
1	0.079	0.078	-0.001	0.079	0.077	-0.002	0.040	0.039	0.001
2	0.077	0.075	-0.002	0.077	0.073	-0.004	0.058	0.057	0.001
3	0.073	0.071	-0.002	0.073	0.070	-0.004	0.073	0.071	0.002
4	0.042	0.040	-0.002	0.042	0.038	-0.003	0.021	0.020	0.001
5	0.041	0.039	-0.002	0.041	0.038	-0.003	0.030	0.029	0.001
6	0.039	0.038	-0.001	0.039	0.036	-0.003	0.039	0.038	0.001
7	0.063	0.062	-0.001	0.063	0.061	-0.001	0.047	0.046	0.001
8	0.061	0.060	-0.001	0.061	0.058	-0.003	0.061	0.060	0.001
9	0.059	0.058	-0.001	0.059	0.057	-0.002	0.029	0.029	0.001
10	0.047	0.045	-0.002	0.047	0.043	-0.004	0.035	0.034	0.002
11	0.046	0.044	-0.002	0.046	0.043	-0.003	0.046	0.044	0.002
12	0.044	0.043	-0.001	0.044	0.042	-0.003	0.022	0.022	0.001
13	0.047	0.045	-0.002	0.047	0.044	-0.004	0.035	0.034	0.001
14	0.046	0.045	-0.001	0.046	0.043	-0.002	0.046	0.045	0.001
15	0.044	0.043	-0.001	0.044	0.041	-0.003	0.022	0.021	0.001

Table B.2: Throughput Results for Second Set of Configurations

Conf	Stage 1			Stage 2			SRP Subsystem		
	Simulation	Analytical	Difference	Simulation	Analytical	Difference	Simulation	Analytical	Difference
16	0.079	0.078	-0.002	0.079	0.077	-0.002	0.059	0.058	0.001
17	0.077	0.075	-0.002	0.077	0.073	-0.004	0.077	0.075	0.002
18	0.072	0.071	-0.002	0.072	0.07	-0.003	0.091	0.088	0.002
19	0.042	0.04	-0.002	0.042	0.038	-0.004	0.031	0.030	0.002
20	0.041	0.039	-0.002	0.041	0.038	-0.003	0.041	0.039	0.002
21	0.039	0.038	-0.001	0.039	0.036	-0.003	0.048	0.047	0.001
22	0.062	0.062	-0.001	0.062	0.061	-0.001	0.062	0.062	0.001
23	0.06	0.06	-0.000	0.06	0.058	-0.002	0.075	0.075	0.000
24	0.059	0.057	-0.001	0.059	0.057	-0.002	0.044	0.043	0.001
25	0.047	0.045	-0.002	0.047	0.043	-0.004	0.047	0.045	0.002
26	0.046	0.044	-0.002	0.046	0.043	-0.003	0.057	0.055	0.002
27	0.044	0.043	-0.001	0.044	0.042	-0.003	0.033	0.032	0.001
28	0.047	0.045	-0.002	0.047	0.044	-0.003	0.047	0.045	0.002
29	0.044	0.045	0.001	0.044	0.043	-0.001	0.055	0.056	-0.001
30	0.043	0.042	-0.001	0.043	0.041	-0.002	0.032	0.032	0.001

Table B.3: Throughput Results for the Third Set of Configurations

Conf	Stage 1			Stage 2			SRP Subsystem		
	Simulation	Analytical	Difference	Simulation	Analytical	Difference	Simulation	Analytical	Difference
31	0.079	0.077	-0.002	0.079	0.077	-0.002	0.079	0.077	0.002
32	0.076	0.075	-0.001	0.076	0.073	-0.004	0.096	0.094	0.002
33	0.070	0.071	0.001	0.07	0.07	-0.001	0.106	0.106	-0.000
34	0.042	0.04	-0.002	0.042	0.038	-0.003	0.042	0.040	0.002
35	0.041	0.039	-0.002	0.041	0.038	-0.003	0.051	0.048	0.002
36	0.039	0.037	-0.001	0.039	0.036	-0.003	0.058	0.056	0.002
37	0.062	0.061	-0.000	0.062	0.061	-0.001	0.077	0.077	0.000
38	0.059	0.060	0.001	0.059	0.058	-0.001	0.088	0.089	-0.001
39	0.058	0.057	-0.001	0.058	0.056	-0.002	0.058	0.057	0.001
40	0.047	0.045	-0.002	0.047	0.043	-0.004	0.059	0.056	0.003
41	0.045	0.044	-0.002	0.045	0.043	-0.003	0.068	0.066	0.002
42	0.044	0.043	-0.001	0.044	0.042	-0.003	0.044	0.043	0.001
43	0.046	0.045	-0.001	0.046	0.043	-0.003	0.058	0.056	0.002
44	0.041	0.045	0.004	0.041	0.043	0.002	0.061	0.067	-0.005
45	0.042	0.042	-0.000	0.042	0.041	-0.001	0.042	0.042	0.000

Table B.4: Utilization Results for Stage 1, First Set of Configurations

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
1	79.1	77.62	-1.9%	63.0	60.8	-3.5%	11.9	11.64	-2.2%	79.1	77.62	-1.9%	65.5	64.69	-1.2%
2	76.7	75.33	-1.8%	62.4	59.01	-5.4%	11.5	11.3	-1.8%	77.2	75.33	-2.4%	63.9	62.77	-1.8%
3	73.0	70.88	-2.9%	60.3	55.52	-7.9%	14.6	14.18	-2.9%	73.3	70.88	-3.3%	60.7	59.07	-2.7%
4	75.0	71.36	-4.9%	53.5	50.87	-4.9%	12.5	11.89	-4.9%	41.7	39.64	-4.9%	34.5	33.03	-4.3%
5	72.8	69.74	-4.2%	52.3	49.72	-4.9%	12.1	11.62	-3.9%	40.6	38.74	-4.6%	33.7	32.29	-4.2%
6	69.6	67.7	-2.7%	49.9	48.27	-3.3%	9.7	9.4	-3.1%	38.8	37.61	-3.1%	32.1	31.34	-2.4%
7	81.3	80.39	-1.1%	78.4	76.26	-2.7%	9.4	9.28	-1.3%	62.6	61.83	-1.2%	51.8	51.53	-0.5%
8	78.6	78.0	-0.8%	77.0	74.0	-3.9%	9.1	9.0	-1.1%	60.9	60.0	-1.5%	50.4	50.0	-0.8%
9	76.3	74.78	-2.0%	73.2	70.94	-3.1%	11.7	11.5	-1.7%	58.9	57.52	-2.3%	48.8	47.93	-1.8%
10	71.0	67.61	-4.8%	49.3	46.57	-5.5%	14.2	13.52	-4.8%	47.3	45.07	-4.7%	39.2	37.56	-4.2%
11	68.8	66.29	-3.7%	48.8	45.66	-6.4%	13.8	13.26	-3.9%	46.2	44.19	-4.4%	38.2	36.83	-3.6%
12	66.5	64.76	-2.6%	46.4	44.62	-3.9%	11.1	10.79	-2.8%	44.4	43.18	-2.8%	36.8	35.98	-2.2%
13	70.8	68.13	-3.8%	67.6	63.59	-5.9%	14.2	13.63	-4.0%	47.2	45.42	-3.8%	39.1	37.85	-3.2%
14	68.0	67.71	-0.4%	72.9	65.08	-10.7%	11.3	11.29	-0.1%	45.6	45.14	-1.0%	37.8	37.62	-0.5%
15	65.7	63.71	-3.0%	65.3	61.23	-6.2%	13.1	12.74	-2.7%	43.9	42.47	-3.3%	36.3	35.4	-2.5%

Table B.5: Utilization Results for Stage 1, Second Set of Configurations

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
16	79.1	77.57	-1.9%	68.5	63.35	-7.5%	11.9	11.63	-2.2%	79.1	77.57	-1.9%	65.5	64.64	-1.3%
17	76.5	75.26	-1.6%	69.3	61.46	-11.3%	11.5	11.29	-1.8%	77.0	75.26	-2.3%	63.8	62.72	-1.7%
18	72.3	70.73	-2.2%	68.2	57.77	-15.3%	14.5	14.15	-2.4%	72.5	70.73	-2.4%	60.0	58.95	-1.8%
19	74.9	71.21	-4.9%	55.3	52.09	-5.8%	12.5	11.87	-5.1%	41.6	39.56	-4.9%	34.5	32.97	-4.5%
20	72.7	69.61	-4.3%	54.3	50.92	-6.2%	12.1	11.6	-4.1%	40.6	38.67	-4.7%	33.7	32.23	-4.4%
21	69.4	67.63	-2.5%	51.8	49.47	-4.5%	9.7	9.39	-3.2%	38.7	37.57	-2.9%	32.1	31.31	-2.5%
22	81.0	80.23	-1.0%	85.0	80.23	-5.6%	9.3	9.26	-0.5%	62.3	61.71	-0.9%	51.6	51.43	-0.3%
23	77.8	77.81	0.0%	83.9	77.81	-7.3%	9.0	8.98	-0.3%	60.2	59.85	-0.6%	49.9	49.88	-0.1%
24	75.9	74.4	-2.0%	78.3	74.4	-5.0%	11.7	11.45	-2.2%	58.6	57.23	-2.3%	48.5	47.69	-1.7%
25	70.8	67.47	-4.7%	51.6	47.98	-7.0%	14.2	13.49	-5.0%	47.2	44.98	-4.7%	39.1	37.48	-4.1%
26	68.4	66.15	-3.3%	52.1	47.04	-9.7%	13.7	13.23	-3.4%	45.9	44.1	-3.9%	38.0	36.75	-3.3%
27	66.4	64.7	-2.6%	48.9	46.01	-5.9%	11.1	10.78	-2.9%	44.4	43.13	-2.9%	36.8	35.94	-2.3%
28	70.4	67.87	-3.6%	72.7	66.36	-8.7%	14.1	13.57	-3.7%	46.9	45.25	-3.5%	38.9	37.71	-3.1%
29	65.5	67.38	2.9%	85.1	69.63	-18.2%	10.9	11.23	3.0%	43.9	44.92	2.3%	36.4	37.44	2.8%
30	64.7	63.28	-2.2%	73.1	65.39	-10.5%	12.9	12.66	-1.9%	43.2	42.19	-2.3%	35.8	35.16	-1.8%

Table B.6: Utilization Results for Stage 1, Third Set of Configurations

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
31	79.0	77.51	-1.9%	76.4	65.88	-13.8%	11.8	11.63	-1.5%	79.0	77.51	-1.9%	65.4	64.59	-1.2%
32	75.9	75.19	-0.9%	79.5	63.91	-19.6%	11.4	11.28	-1.1%	76.4	75.19	-1.6%	63.3	62.66	-1.0%
33	70.2	70.59	0.6%	79.0	60.0	-24.1%	14.0	14.12	0.8%	70.5	70.59	0.1%	58.3	58.82	0.9%
34	74.8	71.08	-5.0%	57.5	53.31	-7.3%	12.5	11.85	-5.2%	41.5	39.49	-4.9%	34.4	32.91	-4.3%
35	72.5	69.47	-4.2%	56.8	52.1	-8.3%	12.1	11.58	-4.3%	40.5	38.59	-4.7%	33.6	32.16	-4.3%
36	69.5	67.56	-2.8%	54.3	50.67	-6.7%	9.6	9.38	-2.3%	38.7	37.54	-3.0%	32.0	31.28	-2.3%
37	80.1	79.98	-0.2%	92.4	84.08	-9.0%	9.2	9.23	0.3%	61.6	61.52	-0.1%	51.0	51.27	0.5%
38	76.0	77.51	2.0%	91.1	81.49	-10.6%	8.8	8.94	1.6%	58.8	59.62	1.4%	48.7	49.69	2.0%
39	75.2	73.92	-1.7%	83.9	77.71	-7.4%	11.6	11.37	-2.0%	58.0	56.86	-2.0%	48.1	47.38	-1.5%
40	70.6	67.31	-4.7%	54.4	49.36	-9.3%	14.1	13.46	-4.5%	47.0	44.87	-4.5%	39.0	37.39	-4.1%
41	67.9	66.01	-2.8%	56.2	48.41	-13.9%	13.6	13.2	-2.9%	45.5	44.01	-3.3%	37.7	36.67	-2.7%
42	66.3	64.61	-2.5%	52.0	47.38	-8.9%	11.0	10.77	-2.1%	44.3	43.08	-2.8%	36.7	35.9	-2.2%
43	69.7	67.57	-3.1%	78.2	69.07	-11.7%	13.9	13.51	-2.8%	46.4	45.05	-2.9%	38.5	37.54	-2.5%
44	61.1	66.97	9.6%	93.9	74.04	-21.2%	10.2	11.16	9.4%	40.9	44.65	9.2%	33.9	37.21	9.8%
45	62.9	62.79	-0.2%	81.2	69.42	-14.5%	12.6	12.56	-0.3%	42.1	41.86	-0.6%	34.8	34.88	0.2%

Table B.7: Average Number in Node Results for Stage 1, Set of Configurations 1.

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
1	2.62	2.21	-2.1%	4.17	3.76	-2.1%	3.57	2.33	-6.2%	2.58	2.11	-2.3%	2.52	2.4	-0.6%
2	2.71	2.47	-1.2%	6.54	3.67	-14.3%	3.45	2.26	-6.0%	2.37	2.1	-1.3%	2.41	2.32	-0.4%
3	2.75	2.61	-0.9%	4.13	3.45	-4.5%	4.38	2.13	-15.0%	1.98	1.92	-0.4%	2.19	2.15	-0.3%
4	1.74	1.54	-2.0%	3.74	3.08	-6.6%	3.75	1.19	-25.6%	0.61	0.52	-0.8%	1.06	1.02	-0.5%
5	1.94	1.82	-1.2%	3.42	3.02	-4.1%	3.63	1.16	-24.7%	0.59	0.54	-0.5%	1.04	1.0	-0.4%
6	2.4	2.33	-0.6%	4.47	2.95	-12.7%	2.91	1.13	-14.8%	0.57	0.56	-0.1%	0.99	0.99	0.0%
7	3.1	2.73	-1.8%	3.59	4.07	2.4%	2.82	1.85	-4.8%	1.47	1.33	-0.7%	1.78	1.78	0.0%
8	3.16	3.0	-0.8%	6.71	3.97	-13.7%	2.73	1.8	-4.6%	1.38	1.28	-0.5%	1.72	1.74	0.1%
9	3.38	3.23	-1.0%	2.88	3.74	5.7%	3.51	1.73	-11.9%	1.29	1.22	-0.5%	1.65	1.66	0.1%
10	1.53	1.32	-2.1%	3.62	2.81	-8.1%	4.26	1.35	-29.1%	0.73	0.61	-1.2%	1.22	1.17	-0.6%
11	1.69	1.57	-1.2%	3.46	2.76	-7.0%	4.14	1.33	-28.1%	0.72	0.65	-0.7%	1.19	1.15	-0.4%
12	2.15	2.04	-0.9%	4.11	2.71	-11.7%	3.33	1.29	-17.0%	0.71	0.68	-0.2%	1.16	1.16	0.0%
13	1.57	1.5	-0.7%	2.72	2.47	-2.5%	4.26	1.36	-29.0%	0.76	0.7	-0.6%	1.23	1.2	-0.3%
14	1.76	1.81	0.4%	4.21	2.72	-12.4%	3.39	1.35	-17.0%	0.75	0.72	-0.3%	1.19	1.21	0.1%
15	1.97	1.91	-0.6%	2.39	2.5	1.2%	3.93	1.27	-26.6%	0.7	0.67	-0.3%	1.14	1.13	-0.1%



Table B.8: Average Number in Node Results for Stage 1, Set of Configurations 2.

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
16	2.59	2.2	-1.9%	4.48	3.94	-2.7%	3.57	2.33	-6.2%	2.56	2.11	-2.3%	2.52	2.4	-0.6%
17	2.64	2.46	-0.9%	7.06	3.85	-16.0%	3.45	2.26	-6.0%	2.31	2.09	-1.1%	2.4	2.32	-0.4%
18	2.65	2.59	-0.4%	4.58	3.61	-6.5%	4.35	2.12	-14.9%	1.87	1.9	0.2%	2.15	2.14	0.0%
19	1.73	1.53	-2.0%	3.83	3.15	-6.8%	3.75	1.19	-25.6%	0.6	0.52	-0.8%	1.06	1.02	-0.5%
20	1.92	1.81	-1.1%	3.53	3.09	-4.4%	3.63	1.16	-24.7%	0.59	0.54	-0.5%	1.04	1.0	-0.4%
21	2.38	2.32	-0.5%	4.56	3.03	-12.8%	2.91	1.13	-14.9%	0.57	0.55	-0.1%	0.99	0.99	0.0%
22	3.0	2.71	-1.5%	3.87	4.7	4.2%	2.79	1.85	-4.7%	1.43	1.32	-0.5%	1.77	1.78	0.0%
23	3.0	2.95	-0.2%	7.03	4.54	-12.4%	2.7	1.8	-4.5%	1.31	1.26	-0.2%	1.69	1.72	0.2%
24	3.31	3.15	-1.0%	2.97	4.19	8.1%	3.51	1.72	-12.0%	1.24	1.19	-0.4%	1.63	1.64	0.1%
25	1.51	1.31	-2.0%	3.75	2.89	-8.5%	4.26	1.35	-29.1%	0.73	0.61	-1.1%	1.22	1.17	-0.5%
26	1.66	1.56	-0.9%	3.68	2.84	-8.4%	4.11	1.32	-27.9%	0.71	0.64	-0.7%	1.19	1.15	-0.3%
27	2.13	2.03	-0.8%	4.25	2.8	-12.1%	3.33	1.29	-17.0%	0.71	0.68	-0.2%	1.16	1.15	0.0%
28	1.53	1.48	-0.6%	2.83	2.64	-1.9%	4.23	1.36	-28.7%	0.76	0.69	-0.6%	1.22	1.19	-0.3%
29	1.61	1.77	1.4%	4.21	3.08	-9.4%	3.27	1.35	-16.0%	0.71	0.71	-0.1%	1.15	1.2	0.4%
30	1.9	1.87	-0.3%	2.56	2.8	2.4%	3.87	1.27	-26.0%	0.69	0.66	-0.3%	1.12	1.12	0.0%

Table B.9: Two Stages FJ With Shared SRP Node. Average Number in Node Results for Stage 1, Set of Configurations 3

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
31	2.52	2.19	-1.6%	4.92	4.13	-3.9%	3.54	2.33	-6.1%	2.5	2.1	-2.0%	2.5	2.4	-0.5%
32	2.52	2.45	-0.3%	7.55	4.04	-17.5%	3.42	2.26	-5.8%	2.21	2.07	-0.7%	2.35	2.31	-0.2%
33	2.44	2.57	0.9%	5.13	3.77	-9.0%	4.2	2.12	-13.9%	1.7	1.87	1.2%	2.04	2.13	0.6%
34	1.7	1.52	-1.8%	3.93	3.23	-7.1%	3.75	1.19	-25.7%	0.6	0.52	-0.8%	1.06	1.01	-0.5%
35	1.9	1.8	-1.0%	3.67	3.17	-5.0%	3.63	1.16	-24.7%	0.59	0.54	-0.5%	1.04	1.0	-0.4%
36	2.36	2.31	-0.5%	4.7	3.11	-13.3%	2.88	1.13	-14.6%	0.56	0.55	-0.1%	0.99	0.99	0.0%
37	2.8	2.66	-0.7%	4.0	5.54	7.7%	2.76	1.85	-4.6%	1.33	1.31	-0.1%	1.73	1.77	0.2%
38	2.71	2.9	0.9%	6.86	5.25	-8.0%	2.64	1.79	-4.3%	1.19	1.24	0.3%	1.62	1.71	0.4%
39	3.18	3.07	-0.8%	3.05	4.71	11.1%	3.48	1.71	-11.8%	1.17	1.15	-0.1%	1.6	1.62	0.1%
40	1.49	1.3	-1.8%	3.9	2.98	-9.2%	4.23	1.35	-28.8%	0.72	0.61	-1.1%	1.22	1.16	-0.5%
41	1.62	1.55	-0.7%	3.94	2.93	-10.1%	4.08	1.32	-27.6%	0.7	0.64	-0.6%	1.18	1.15	-0.3%
42	2.11	2.02	-0.8%	4.41	2.88	-12.7%	3.3	1.29	-16.7%	0.71	0.68	-0.2%	1.15	1.15	0.0%
43	1.49	1.46	-0.3%	2.92	2.82	-1.0%	4.17	1.35	-28.2%	0.74	0.69	-0.5%	1.21	1.19	-0.2%
44	1.39	1.74	2.9%	3.85	3.5	-2.8%	3.06	1.34	-14.3%	0.64	0.7	0.4%	1.06	1.19	1.1%
45	1.79	1.82	0.3%	2.72	3.12	4.0%	3.78	1.26	-25.2%	0.65	0.64	-0.1%	1.08	1.11	0.2%

Table B.10: Utilization Results for Stage 2, First Set of Configurations

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
1	79.1	76.78	-2.9%	63.0	60.14	-4.5%	11.9	11.52	-3.2%	79.1	76.78	-2.9%	65.5	63.98	-2.3%
2	76.7	72.99	-4.8%	66.7	59.61	-10.6%	15.3	14.6	-4.6%	77.2	72.99	-5.5%	63.9	60.82	-4.8%
3	73.1	69.71	-4.6%	70.2	59.25	-15.6%	11.0	10.46	-4.9%	73.3	69.71	-4.9%	60.7	58.09	-4.3%
4	75.0	68.84	-8.2%	53.4	49.08	-8.1%	12.5	11.47	-8.2%	41.7	38.24	-8.3%	34.5	31.87	-7.6%
5	72.8	68.03	-6.5%	54.0	49.77	-7.8%	10.1	9.45	-6.4%	40.6	37.8	-6.9%	33.7	31.5	-6.5%
6	69.6	64.74	-7.0%	53.3	48.55	-8.9%	11.6	10.79	-7.0%	38.8	35.97	-7.3%	32.1	29.97	-6.6%
7	81.3	79.58	-2.1%	83.7	79.58	-4.9%	9.4	9.18	-2.3%	62.6	61.21	-2.2%	51.9	51.01	-1.7%
8	78.7	75.81	-3.7%	88.2	79.7	-9.6%	12.1	11.66	-3.6%	60.9	58.32	-4.2%	50.4	48.6	-3.6%
9	76.3	73.73	-3.4%	73.2	69.95	-4.4%	8.8	8.51	-3.3%	58.9	56.72	-3.7%	48.8	47.27	-3.1%
10	71.0	65.28	-8.1%	70.6	63.83	-9.6%	14.2	13.06	-8.1%	47.3	43.52	-8.0%	39.2	36.27	-7.5%
11	68.8	64.63	-6.1%	75.4	66.06	-12.4%	11.5	10.77	-6.3%	46.1	43.09	-6.5%	38.3	35.9	-6.3%
12	66.5	62.43	-6.1%	62.8	58.27	-7.2%	13.3	12.49	-6.1%	44.4	41.62	-6.3%	36.8	34.68	-5.7%
13	70.9	65.56	-7.5%	52.0	46.62	-10.4%	14.2	13.11	-7.7%	47.2	43.7	-7.4%	39.1	36.42	-6.9%
14	68.0	64.85	-4.6%	63.9	50.26	-21.3%	13.6	12.97	-4.6%	45.6	43.24	-5.2%	37.8	36.03	-4.7%
15	65.6	61.95	-5.6%	47.6	43.54	-8.5%	10.9	10.32	-5.3%	43.9	41.3	-5.9%	36.4	34.42	-5.5%

Table B.11: Utilization Results for Stage 2, Second Set of Configurations

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
16	79.1	76.72	-3.0%	64.0	60.1	-6.1%	11.9	11.51	-3.3%	79.1	76.72	-3.0%	65.5	63.94	-2.4%
17	76.5	72.94	-4.7%	69.2	59.57	-13.9%	15.3	14.59	-4.7%	77.0	72.94	-5.3%	63.8	60.78	-4.7%
18	72.3	69.58	-3.8%	74.4	59.14	-20.5%	10.8	10.44	-3.4%	72.5	69.58	-4.0%	60.1	57.98	-3.5%
19	74.9	68.72	-8.3%	53.6	48.99	-8.6%	12.5	11.45	-8.4%	41.6	38.18	-8.2%	34.5	31.81	-7.8%
20	72.7	67.93	-6.6%	54.4	49.69	-8.7%	10.1	9.43	-6.6%	40.6	37.74	-7.1%	33.6	31.45	-6.4%
21	69.5	64.69	-6.9%	53.7	48.52	-9.6%	11.6	10.78	-7.1%	38.7	35.94	-7.1%	32.1	29.95	-6.7%
22	81.0	79.44	-1.9%	85.1	79.44	-6.7%	9.4	9.17	-2.5%	62.3	61.1	-1.9%	51.6	50.92	-1.3%
23	77.8	75.68	-2.7%	90.0	79.56	-11.6%	12.0	11.64	-3.0%	60.2	58.21	-3.3%	49.9	48.51	-2.8%
24	75.9	73.4	-3.3%	73.4	69.64	-5.1%	8.8	8.47	-3.8%	58.6	56.46	-3.6%	48.5	47.05	-3.0%
25	70.8	65.17	-8.0%	71.3	63.72	-10.6%	14.2	13.03	-8.2%	47.2	43.45	-8.0%	39.1	36.2	-7.4%
26	68.4	64.51	-5.7%	77.2	65.95	-14.6%	11.4	10.75	-5.7%	45.9	43.01	-6.3%	38.0	35.84	-5.7%
27	66.4	62.38	-6.0%	63.4	58.23	-8.2%	13.3	12.48	-6.2%	44.4	41.59	-6.3%	36.8	34.66	-5.8%
28	70.4	65.35	-7.2%	52.5	46.47	-11.5%	14.1	13.07	-7.3%	46.9	43.56	-7.1%	38.9	36.3	-6.7%
29	65.5	64.61	-1.4%	66.1	50.07	-24.2%	13.1	12.92	-1.4%	43.9	43.07	-1.9%	36.4	35.9	-1.4%
30	64.7	61.59	-4.8%	47.7	43.29	-9.3%	10.8	10.27	-4.9%	43.2	41.06	-4.9%	35.8	34.22	-4.4%

Table B.12: Utilization Results for Stage 2, Third Set of Configurations

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
31	79.0	76.67	-2.9%	65.4	60.06	-8.2%	11.9	11.5	-3.4%	79.0	76.67	-2.9%	65.5	63.89	-2.5%
32	75.9	72.89	-4.0%	72.8	59.53	-18.2%	15.2	14.58	-4.1%	76.4	72.89	-4.6%	63.3	60.74	-4.0%
33	70.2	69.45	-1.1%	79.1	59.03	-25.4%	10.5	10.42	-0.8%	70.4	69.45	-1.3%	58.4	57.88	-0.9%
34	74.8	68.61	-8.3%	53.6	48.92	-8.7%	12.5	11.44	-8.5%	41.5	38.12	-8.1%	34.4	31.77	-7.7%
35	72.5	67.8	-6.5%	54.7	49.6	-9.3%	10.1	9.42	-6.8%	40.5	37.67	-7.0%	33.6	31.39	-6.6%
36	69.4	64.64	-6.9%	54.2	48.48	-10.5%	11.6	10.77	-7.1%	38.7	35.91	-7.2%	32.0	29.93	-6.5%
37	80.1	79.22	-1.1%	86.1	79.22	-8.0%	9.2	9.14	-0.6%	61.6	60.94	-1.1%	51.0	50.78	-0.4%
38	76.0	75.47	-0.7%	91.1	79.34	-12.9%	11.7	11.61	-0.8%	58.8	58.05	-1.3%	48.7	48.38	-0.7%
39	75.2	72.99	-2.9%	73.3	69.24	-5.5%	8.7	8.42	-3.2%	58.0	56.14	-3.2%	48.1	46.79	-2.7%
40	70.6	65.05	-7.9%	72.1	63.6	-11.8%	14.1	13.01	-7.7%	47.0	43.36	-7.7%	39.0	36.14	-7.3%
41	67.9	64.4	-5.2%	79.4	65.83	-17.1%	11.3	10.73	-5.0%	45.5	42.93	-5.6%	37.7	35.78	-5.1%
42	66.3	62.32	-6.0%	64.0	58.17	-9.1%	13.2	12.46	-5.6%	44.3	41.55	-6.2%	36.7	34.62	-5.7%
43	69.6	65.11	-6.5%	52.7	46.3	-12.1%	13.9	13.02	-6.3%	46.4	43.41	-6.5%	38.5	36.17	-6.0%
44	61.1	64.3	5.2%	65.1	49.84	-23.4%	12.2	12.86	5.4%	41.0	42.87	4.6%	33.9	35.72	5.4%
45	62.9	61.18	-2.7%	47.3	43.0	-9.1%	10.5	10.2	-2.9%	42.1	40.79	-3.1%	34.9	33.99	-2.6%

Table B.13: Average Number in Node Results for Stage 2, Set of Configurations 1

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
1	2.52	2.12	-2.0%	3.94	3.71	-1.17%	2.38	2.3	-0.38%	2.54	2.03	-2.56%	2.52	2.36	-0.79%
2	2.41	2.14	-1.79%	4.27	3.69	-3.85%	2.29	2.19	-0.7%	2.14	1.84	-1.97%	2.34	2.18	-1.1%
3	3.08	2.59	-2.44%	5.01	3.75	-6.3%	2.2	2.09	-0.54%	2.1	1.86	-1.21%	2.24	2.12	-0.57%
4	1.73	1.4	-3.3%	3.24	2.96	-2.8%	1.25	1.15	-1.03%	0.61	0.5	-1.11%	1.06	0.98	-0.83%
5	2.1	1.76	-2.86%	3.33	3.03	-2.55%	1.21	1.13	-0.65%	0.6	0.52	-0.63%	1.04	0.98	-0.53%
6	2.19	1.99	-2.0%	3.29	2.96	-3.29%	1.16	1.08	-0.81%	0.56	0.51	-0.47%	0.99	0.94	-0.5%
7	3.14	2.62	-2.63%	5.57	4.57	-4.99%	1.88	1.84	-0.22%	1.45	1.29	-0.8%	1.79	1.76	-0.14%
8	2.92	2.56	-2.45%	5.65	4.59	-7.09%	1.81	1.75	-0.44%	1.27	1.16	-0.77%	1.69	1.64	-0.32%
9	3.91	3.27	-3.19%	4.32	3.74	-2.85%	1.76	1.7	-0.29%	1.35	1.2	-0.75%	1.66	1.64	-0.11%
10	1.55	1.34	-2.11%	2.85	2.45	-3.91%	1.42	1.31	-1.14%	0.76	0.65	-1.07%	1.23	1.14	-0.9%
11	1.81	1.59	-1.77%	3.55	2.75	-6.66%	1.38	1.29	-0.73%	0.75	0.66	-0.79%	1.21	1.14	-0.56%
12	2.05	1.83	-2.22%	2.57	2.3	-2.67%	1.33	1.25	-0.81%	0.72	0.65	-0.72%	1.16	1.11	-0.51%
13	1.52	1.23	-3.0%	3.15	2.81	-3.4%	1.42	1.31	-1.09%	0.73	0.59	-1.46%	1.22	1.13	-0.92%
14	1.64	1.49	-1.51%	4.04	3.04	-9.92%	1.36	1.3	-0.63%	0.71	0.62	-0.87%	1.18	1.13	-0.53%
15	2.11	1.82	-2.39%	2.94	2.64	-2.47%	1.31	1.24	-0.58%	0.7	0.63	-0.54%	1.14	1.1	-0.38%

Table B.14: Average Number in Node Results for Stage 2, Set of Configurations 2

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
16	2.5	2.11	-1.97%	4.03	3.71	-1.59%	2.38	2.3	-0.39%	2.54	2.03	-2.54%	2.52	2.36	-0.79%
17	2.38	2.14	-1.65%	4.5	3.69	-5.4%	2.29	2.19	-0.71%	2.11	1.84	-1.82%	2.33	2.18	-1.04%
18	2.96	2.58	-1.88%	5.56	3.74	-9.12%	2.16	2.09	-0.36%	2.0	1.85	-0.75%	2.19	2.12	-0.37%
19	1.72	1.4	-3.26%	3.26	2.96	-2.98%	1.25	1.15	-1.05%	0.6	0.5	-1.1%	1.06	0.98	-0.85%
20	2.09	1.75	-2.81%	3.36	3.02	-2.81%	1.21	1.13	-0.67%	0.6	0.52	-0.63%	1.04	0.98	-0.51%
21	2.17	1.98	-1.91%	3.31	2.96	-3.58%	1.16	1.08	-0.82%	0.56	0.51	-0.46%	0.99	0.94	-0.51%
22	3.07	2.6	-2.33%	5.88	4.54	-6.71%	1.88	1.83	-0.23%	1.43	1.29	-0.72%	1.77	1.75	-0.08%
23	2.8	2.54	-1.72%	6.01	4.56	-9.63%	1.8	1.75	-0.36%	1.23	1.16	-0.52%	1.67	1.64	-0.18%
24	3.81	3.22	-2.95%	4.32	3.7	-3.12%	1.76	1.69	-0.33%	1.33	1.19	-0.69%	1.65	1.63	-0.07%
25	1.54	1.34	-2.0%	2.88	2.45	-4.38%	1.42	1.3	-1.17%	0.75	0.65	-1.06%	1.23	1.14	-0.89%
26	1.78	1.59	-1.57%	3.73	2.74	-8.26%	1.37	1.29	-0.65%	0.75	0.66	-0.76%	1.2	1.14	-0.49%
27	2.04	1.82	-2.16%	2.61	2.3	-3.08%	1.33	1.25	-0.82%	0.72	0.65	-0.71%	1.16	1.11	-0.52%
28	1.5	1.22	-2.88%	3.18	2.8	-3.82%	1.41	1.31	-1.03%	0.73	0.59	-1.42%	1.21	1.13	-0.89%
29	1.52	1.47	-0.52%	4.21	3.03	-11.76%	1.31	1.29	-0.18%	0.67	0.62	-0.53%	1.13	1.12	-0.09%
30	2.04	1.8	-2.03%	2.94	2.62	-2.66%	1.3	1.23	-0.53%	0.68	0.63	-0.43%	1.12	1.09	-0.26%

Table B.15: Average Number in Node Results for Stage 2, Set of Configurations 3

Conf	Node 1			Node 2			Node 3			Node 4			Node 5		
	S	A	Error	S	A	Error	S	A	Error	S	A	Error	S	A	Error
31	2.48	2.11	-1.86%	4.14	3.71	-2.19%	2.38	2.3	-0.4%	2.5	2.02	-2.41%	2.51	2.36	-0.76%
32	2.32	2.13	-1.22%	4.85	3.69	-7.72%	2.28	2.19	-0.62%	2.05	1.83	-1.44%	2.3	2.17	-0.84%
33	2.69	2.57	-0.65%	6.21	3.73	-12.39%	2.1	2.08	-0.08%	1.79	1.84	0.21%	2.08	2.11	0.14%
34	1.71	1.39	-3.18%	3.26	2.95	-3.03%	1.25	1.14	-1.06%	0.6	0.49	-1.09%	1.06	0.98	-0.83%
35	2.06	1.74	-2.69%	3.38	3.02	-3.03%	1.21	1.13	-0.68%	0.6	0.52	-0.62%	1.04	0.98	-0.53%
36	2.17	1.98	-1.87%	3.35	2.95	-3.95%	1.16	1.08	-0.83%	0.56	0.51	-0.44%	0.99	0.94	-0.47%
37	2.85	2.57	-1.39%	6.12	4.5	-8.12%	1.84	1.83	-0.06%	1.38	1.28	-0.51%	1.74	1.75	0.04%
38	2.57	2.52	-0.37%	6.27	4.52	-11.71%	1.75	1.74	-0.09%	1.16	1.15	-0.09%	1.61	1.64	0.14%
39	3.66	3.16	-2.48%	4.28	3.64	-3.18%	1.74	1.68	-0.28%	1.29	1.17	-0.57%	1.63	1.62	-0.03%
40	1.52	1.33	-1.85%	2.93	2.44	-4.95%	1.41	1.3	-1.09%	0.75	0.65	-1.02%	1.22	1.14	-0.88%
41	1.73	1.58	-1.28%	3.95	2.73	-10.14%	1.36	1.29	-0.57%	0.74	0.66	-0.69%	1.19	1.14	-0.42%
42	2.03	1.82	-2.11%	2.64	2.29	-3.48%	1.32	1.25	-0.74%	0.72	0.65	-0.69%	1.15	1.11	-0.49%
43	1.47	1.21	-2.61%	3.2	2.79	-4.07%	1.39	1.3	-0.88%	0.71	0.58	-1.32%	1.2	1.12	-0.79%
44	1.34	1.46	1.12%	4.14	3.02	-11.21%	1.22	1.29	0.66%	0.61	0.61	0.07%	1.05	1.12	0.68%
45	1.92	1.77	-1.27%	2.91	2.6	-2.57%	1.26	1.22	-0.3%	0.65	0.62	-0.23%	1.09	1.08	-0.06%



Table B.16: Average Number in Node Results for Fork/Join nodes, Set of Configurations 1

Conf	$Lq_{P,0}$			$Lq_{P,1}$		
	S	A	Error	S	A	Error
1	0.99	3.5	-12.6%	0.39	3.12	-13.7%
2	3.38	5.37	-9.9%	2.8	6.43	-18.2%
3	4.21	5.28	-7.2%	0.51	2.31	-12.0%
4	1.89	3.1	-12.1%	0.53	2.25	-17.2%
5	1.72	2.97	-12.5%	0.28	1.59	-13.1%
6	2.27	3.24	-8.1%	1.48	3.22	-14.5%
7	1.79	2.76	-4.8%	1.24	2.81	-7.8%
8	5.2	5.67	-2.3%	4.4	5.89	-7.4%
9	3.54	4.51	-6.5%	0.69	1.89	-8.0%
10	2.08	3.17	-11.0%	0.66	2.16	-15.0%
11	2.04	3.02	-9.8%	0.53	1.6	-10.7%
12	2.0	3.05	-8.8%	1.33	2.81	-12.4%
13	2.08	2.9	-8.2%	0.69	2.32	-16.2%
14	3.45	2.98	3.9%	2.03	3.35	-11.0%
15	2.23	2.87	-6.4%	0.43	1.59	-11.6%

Table B.17: Average Number in Node Results for Fork/Join nodes, Set of Configurations 2

Conf	$Lq_{P,0}$			$Lq_{P,1}$		
	S	A	Error	S	A	Error
16	1.19	3.63	-12.2%	0.37	3.11	-13.7%
17	4.22	5.49	-6.3%	2.9	6.41	-17.5%
18	5.22	5.41	-1.3%	0.49	2.29	-12.0%
19	2.0	3.16	-11.6%	0.51	2.23	-17.2%
20	1.84	3.03	-11.9%	0.27	1.58	-13.1%
21	2.34	3.29	-7.9%	1.46	3.21	-14.6%
22	2.82	3.15	-1.7%	1.32	2.77	-7.3%
23	6.97	5.99	4.9%	4.51	5.8	-6.5%
24	4.12	4.87	-5.0%	0.62	1.83	-8.1%
25	2.22	3.24	-10.2%	0.65	2.15	-15.0%
26	2.33	3.09	-7.5%	0.56	1.59	-10.3%
27	2.13	3.11	-8.2%	1.31	2.8	-12.4%
28	2.45	3.03	-5.8%	0.64	2.29	-16.5%
29	5.2	3.23	16.3%	1.66	3.3	-13.7%
30	2.88	3.1	-2.2%	0.37	1.56	-11.9%

Table B.18: Average Number in Node Results for Fork/Join nodes, Set of Configurations 3

Conf	$Lq_{P,0}$			$Lq_{P,1}$		
	S	A	Error	S	A	Error
31	1.7	3.76	-10.3%	0.33	3.1	-13.8%
32	5.96	5.6	1.8%	2.78	6.38	-18.0%
33	7.27	5.54	11.5%	0.39	2.27	-12.5%
34	2.12	3.22	-11.0%	0.48	2.22	-17.4%
35	1.99	3.1	-11.1%	0.26	1.57	-13.1%
36	2.5	3.35	-7.0%	1.45	3.2	-14.6%
37	5.1	3.71	6.9%	1.23	2.71	-7.4%
38	9.78	6.46	16.6%	4.12	5.67	-7.7%
39	4.96	5.3	-2.2%	0.53	1.76	-8.2%
40	2.4	3.31	-9.1%	0.63	2.13	-15.0%
41	2.76	3.15	-4.0%	0.57	1.57	-10.1%
42	2.31	3.18	-7.2%	1.29	2.79	-12.5%
43	2.94	3.17	-2.4%	0.57	2.26	-16.9%
44	7.16	3.54	30.1%	1.03	3.24	-18.4%
45	3.83	3.36	4.7%	0.29	1.52	-12.3%

Table B.19: Results for the SRP node, Set of Configurations 1

Configuration	SRP Node Utilization			Average Num in Node		
	S	A	Error	S	A	Error
1	31.4	31.05	1.1%	0.45	0.43	0.2%
2	46.0	45.2	1.7%	0.8	0.76	0.3%
3	58.3	56.7	2.7%	1.23	1.21	0.2%
4	16.6	15.86	4.5%	0.2	0.18	0.1%
5	24.2	23.25	3.9%	0.32	0.29	0.2%
6	31.0	30.09	2.9%	0.41	0.4	0.1%
7	37.5	37.1	1.1%	0.5	0.5	0.0%
8	48.7	48.0	1.4%	0.71	0.74	-0.5%
9	23.5	23.01	2.1%	0.28	0.27	0.1%
10	28.4	27.04	4.8%	0.35	0.33	0.2%
11	36.8	35.35	3.9%	0.58	0.59	-0.1%
12	17.7	17.27	2.4%	0.23	0.22	0.1%
13	28.3	27.25	3.7%	0.43	0.39	0.4%
14	59.2	58.68	0.9%	1.47	1.46	0.1%
15	28.4	27.61	2.8%	0.43	0.41	0.2%

Table B.20: Results for the SRP node, Set of Configurations 2

Configuration	SRP Node Utilization			Average Num in Node		
	S	A	Error	S	A	Error
16	47.2	46.54	1.4%	0.83	0.79	0.4%
17	61.2	60.21	1.6%	1.38	1.31	0.6%
18	72.1	70.73	1.9%	2.03	2.07	-0.3%
19	24.8	23.74	4.3%	0.32	0.3	0.2%
20	32.3	30.94	4.2%	0.47	0.43	0.3%
21	38.7	37.57	2.9%	0.56	0.54	0.2%
22	49.9	49.37	1.1%	0.75	0.77	-0.2%
23	60.2	59.85	0.6%	1.01	1.09	-1.3%
24	35.1	34.34	2.2%	0.46	0.46	0.0%
25	37.7	35.98	4.6%	0.52	0.48	0.4%
26	45.8	44.1	3.7%	0.86	0.85	0.1%
27	26.6	25.88	2.7%	0.4	0.38	0.3%
28	37.5	36.2	3.5%	0.64	0.6	0.4%
29	71.2	73.0	-2.5%	2.13	2.52	-4.4%
30	42.0	41.13	2.1%	0.74	0.77	-0.3%

Table B.21: Results for the SRP node, Set of Configurations 3

Configuration	SRP Node Utilization			Average Num in Node		
	S	A	Error	S	A	Error
31	62.8	62.01	1.3%	1.4	1.35	0.3%
32	75.9	75.19	0.9%	2.27	2.35	-0.7%
33	84.0	84.71	-0.8%	3.15	3.92	-6.4%
34	33.0	31.59	4.3%	0.47	0.43	0.4%
35	40.3	38.59	4.2%	0.65	0.59	0.5%
36	46.4	45.04	2.9%	0.75	0.72	0.2%
37	61.6	61.52	0.1%	1.05	1.14	-1.5%
38	70.5	71.55	-1.5%	1.36	1.61	-4.2%
39	46.4	45.49	2.0%	0.67	0.69	-0.4%
40	47.0	44.87	4.5%	0.72	0.65	0.9%
41	54.5	52.81	3.1%	1.21	1.19	0.2%
42	35.3	34.46	2.4%	0.61	0.58	0.4%
43	46.3	45.05	2.7%	0.87	0.87	0.0%
44	79.7	87.06	-9.2%	2.63	4.54	-21.2%
45	54.6	54.42	0.3%	1.07	1.3	-2.6%

## VITA

Ronny Pacheco Segura

Candidate for the Degree of

Doctor of Philosophy

Dissertation: DEVELOPMENT OF A PARAMETRIC-DECOMPOSITION METHODOLOGY FOR SOLVING NETWORKS OF QUEUES WITH SIMULTANEOUS RESOURCE POSSESSION UNDER CAPACITY RESTRICTIONS

Major Field: Industrial Engineering and Management

Biographical:

### Education:

Completed the requirements for the Ph.D. degree in Industrial Engineering and Management at Oklahoma State University in December, 2020

Received the M.S. degree in Industrial Engineering from University of Costa Rica in June, 2012

Received the Licenciature degree in Industrial Engineering from Universtiy of Costa Rica, Costa Rica, 2000

### Experience:

Strategy and ABC Costing consultant. Van der Leer S.A. Costa Rica (March 2000 - September 2002).

Project Engineer, Process Engineer, Quality Assurance Department Manager. Bticino S.A. Costa Rica (September 2002 - March 2008).

Freelance Consultant. Costa Rica (March 2008 - August 2015)

Lecturer. University of Costa Rica. (March 2001 - August 2015)

Graduate Research/Teaching Assistant - School of Industrial Engineering and Management, Oklahoma State University (March 2015 - May 2020).

Lecturer. University of Costa Rica. (June 2020 - Today)