Abstract—Coordination languages, as Reo, have emerged for the specification and implementation of interaction protocols among concurrent entities. In this paper, we propose a framework for generating partially-distributed, partially-centralized implementations of Reo connectors to improve 1) build-time compilation and 2) run-time throughput and parallelism. Our framework relies on the definition of a new formal product operator on constraint automata (Reo’s formal semantics), which enables the formally correct distribution of disjoint parts of a coordination scheme over different machines according to several possible motivations (e.g., performance, privacy, QoS constraints, resource availability, network topology). First, we describe the design and a proof-of-concept implementation of our framework. Then, in a case study, we show and explain how a generated connector implementation can be executed in the Cloud and supports Big Data coordination.

Keywords- Reo coordination language; distributed computation; Web services; Cloud; Big Data;

I. INTRODUCTION

A. Context

Coordination languages have emerged for the specification and implementation of interaction protocols among concurrent entities (components, services, threads, etc.). This class of languages includes Reo [1, 2], a graphical dataflow language for compositional construction of connectors: communication media through which entities can interact with each other. Figure 1 shows example connectors in their usual graphical syntax. Briefly, connectors consist of one or more channels, through which data items flow, and a number of (named) nodes, on which channel ends coincide. Through

![Figure 1: Example connectors: the Alternator family.](image)

connector composition (the act of gluing connectors together on their shared nodes), developers can construct arbitrarily complex connectors. While nodes have fixed dataflow behavior, Reo features an open-ended set of channels: developers can define their own channels with custom semantics.

Compared to general-purpose languages, Reo has several software engineering advantages as a domain-specific language for programming interaction protocols [3]. For instance, the use of Reo forces developers to separate their computation code from their protocol code instead of intermixing computations with protocols (as usually done when both are implemented in the same language). This separation facilitates verbatim reuse, independent modification, and compositional construction of protocol implementations (i.e., Reo connectors) in a straightforward way. Moreover, Reo has a formal foundation, which enables formal analysis of connectors (e.g., model checking [4]). This makes statically verifying that a given protocol does not deadlock, for instance, relatively easy. Such protocol analyses are much harder on lower-level general-purpose languages, especially when computations and protocols are intermixed.

B. Problem

To use connectors in real applications, one must derive implementations from their graphical specification, as pre-compiled executable code or using a run-time interpretation engine. Roughly two implementation approaches currently exist. In the distributed approach [5, 6, 7], one implements the behavior of each of the \( k \) constituents of a connector (its nodes and channels) and runs these \( k \) implementations concurrently as a distributed system; in the centralized approach [3, 8], one computes the behavior of a connector as a whole, implements this behavior, and runs this implementation sequentially as a centralized system. The distributed approach has the advantage of fast compilation at build-time and high parallelism at run-time. However, this comes at the cost of reduced throughput at run-time (because of communication necessary for executing distributed algorithms). In contrast, the centralized approach has the advantage of high throughput at run-time but at the cost of slow compilation and reduced parallelism. Moreover, the amount of code generated in the centralized approach may
be exponential in $k$, in which case the output is prohibitively big and the time to produce it prohibitively long. Proença et al. observe that a partially-distributed hybrid approach, where parts of a connector are compiled according to the centralized approach and deployed in a distributed fashion, is generally ideal [6, 7]: a hybrid approach strikes a middle ground between throughput and parallelism at run-time while achieving reasonably fast compilation at build-time.

In this paper, we address the problem that no systematic, formally justified way of automatically constructing connector implementations according to the hybrid approach exists.

C. Contribution

We use Reo’s notion of (a)synchronous regions and an extended recent result on connector composition to automatically construct formally sound hybrid connector implementations (using constraint automata as our semantic formalism [9]). We implement this construction on top of an existing Reo-to-Java centralized-code generator [3]. This enables a case study on distributed service orchestration by reusing an existing Reo-based orchestration framework [8]. This case study shows that the hybrid approach can improve the centralized approach not only in terms of run-time parallelism but also throughput.

We organize this paper as follows. In Sec. II, we give a concise overview of Reo. In Sec. III, we explain the formal theory behind our hybrid connector implementations and how to automatically generate them. In Sec. IV, we discuss the salient aspects of our implementation. In Sec. V, we exemplify our code generator with a distributed service orchestration scenario. In Sec. VI, we discuss related work, after which Sec. VII concludes this paper.

II. REO COORDINATION LANGUAGE

Reo is a language for compositional construction of concurrency protocols, manifested as connectors [1, 2]. Connectors consist of channels and nodes, organized in a graph-like structure. Every channel consists of two ends and a constraint that relates the timing and the contents of the data-flows at those ends. A channel end has one of two types: source ends accept data (i.e., a source end of a channel connects to that channel’s data source/producer), while sink ends dispense data (i.e., a sink end of a channel connects to that channel’s data sink/consumer). Reo makes no other assumptions about channels. This means, for instance, that Reo allows channels with two source ends. Table I shows the syntax and an informal description of common channels.

Entities communicating through a connector perform I/O operations—writes and takes—on its nodes. Reo features three types of nodes: source nodes on which only source ends coincide (see Figure 2a), sink nodes on which only sink ends coincide (see Figure 2b), and mixed nodes on which both types of channel end coincide (see Figure 2c). Informally, nodes behave as follows.

- A source node $n$ has replicator semantics. Once a communicating entity attempts to write a piece of data $d$ on $n$, this node first suspends that operation. Subsequently, $n$ notifies the channels whose source ends coincide on $n$ that it offers $d$. Once each of those channels has notified $n$ that it accepts $d$, $n$ resolves the write: the write operation terminates successfully and atomically, $n$ dispenses (a copy of) $d$ to each of its coincident source ends. Source nodes forbid takes.
- A sink node $n$ has merger semantics. Once a communicating entity attempts to take a piece of data from $n$, this node first suspends that operation. Subsequently, $n$ notifies the channels whose sink ends coincide on $n$ that it accepts a piece of data. Once at least one of these channels has notified $n$ that it offers a piece of data $d$, $n$ resolves the take: atomically, $n$ fetches $d$ from the appropriate channel end and dispenses it to the entity attempting to take. If multiple sink ends offer a data item, $n$ chooses one of them nondeterministically. Sink nodes forbid writes.
- A mixed node has pumping station semantics: the atomic execution of the replicator semantics and merger semantics discussed above. Mixed nodes forbid I/O.

Table I: Syntax and semantics of common channels.

<table>
<thead>
<tr>
<th>Name</th>
<th>Graphical syntax</th>
<th>Semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>sync</td>
<td></td>
<td>Atomically takes a data item $d$ from its source end and writes $d$ to its sink end.</td>
</tr>
<tr>
<td>lossysync-nd</td>
<td></td>
<td>Atomically takes a data item $d$ from its source end and nondeterministically either writes $d$ to its sink end or loses $d$.</td>
</tr>
<tr>
<td>syncdrain</td>
<td></td>
<td>Atomically takes data items from both its source ends and loses them.</td>
</tr>
<tr>
<td>fifo</td>
<td></td>
<td>Takes a data item $d$ from its source end, then writes $d$ to its sink end.</td>
</tr>
</tbody>
</table>

Figure 2: Node types.

Figure 3: Merger, Replicator, and Node for $n = 3$ and $m = 2$. 


So far, we explicitly distinguished between three language constructs in Reo: channels, nodes, and connectors. In the rest of this paper, for simplicity (but without loss of generality), we represent every channel \( c \) as its corresponding connector, which consists of two nodes connected by \( c \). To define the semantics of \( c \), it suffices to define the semantics of its corresponding connector. Furthermore, we assume that at most one source end and at most one sink end coincides on every node [9]. We model such binary nodes with ports. Ports occur either on the boundary between a connector and its environment or inside a connector. If a connector consists only of boundary ports, we call it a primitive; otherwise, we call it a composite. To simulate the semantics of nodes with more than two coincident channel ends, we use two ternary primitives: Merger and Replicator. These primitives can compose into a Node composite, which connects \( n \geq 0 \) input ports to \( m \geq 0 \) output ports and behaves as a node on which \( n \) source ends and \( m \) sink ends coincide (where \( n + m > 0 \)). Figure 3 shows an example.

Many formalisms exist for mathematically defining the semantics of connectors [10]. In this paper, we adopt the same formalism as the existing code generator that we use: constraint automata (CA) [9]. A constraint automaton consists of finite sets of states and transitions. States represent the internal configurations of a connector; transitions describe the atomic steps of the protocol specified by a connector. Formally, we represent a transition as a tuple of four elements: a source state, a synchronization constraint, a data constraint, and a target state. A synchronization constraint is a set that specifies on which ports a data item flows (i.e., which ports synchronize); a data constraint is a logical formula that specifies which particular data items flow on which of those ports.

Figure 4 shows example CA, where A and B refer to ports. Informally, the data constraint \( d(A) = d(B) \) means that the data item flowing on port A equals the data item flowing on port B; the data constraint \( \top \) means that it does not matter which particular data items flow; the data constraint \( x' = d(A) \) means that the value of \( x \), a private (to the connector) memory cell, after completing the transition equals the data flowing on A during the transition (i.e., \( x \) is read from in the end); finally, the data constraint \( d(B) = x \) means that the data flowing on B equals the value of private memory cell \( x \) before starting the transition (i.e., \( x \) is read from in the beginning).

Let \( \text{State} \), \( \text{Port} \), \( \text{Mem} \), and \( \text{DC} \), defined in [11, Appx. B], denote universes of states, ports, memory cells, and data constraints.

**Definition 1.** The universe of CA, denoted by \( \mathcal{CA} \), is the largest set of tuples \( (Q, P, M, \rightarrow, i) \) where:

- \( Q \subseteq \text{State} \) (states)
- \( P \subseteq \text{Port} \) (ports)
- \( M \subseteq \text{Mem} \) (memory cells)
- \( \rightarrow \subseteq Q \times \psi(P) \times \text{DC} \times Q \) (transitions)
- \( i \in Q \) (initial state)

If \( \alpha \) denotes a CA, let \( \text{State}(\alpha) \), \( \text{Port}(\alpha) \), \( \text{Mem}(\alpha) \), and \( \text{init}(\alpha) \), defined in [11, Appx. B], denote its states, ports, memory cells, and initial state. We adopt bisimilarity on \( \mathcal{CA} \) as behavioral equivalence [9]: if \( \alpha \) and \( \beta \) are bisimilar, denoted by \( \alpha \approx \beta \), \( \alpha \) can “simulate” every transition of \( \beta \) in every state and vice versa. Henceforth, we consider \( \mathcal{CA} \) up to bisimilarity.

Individual CA describe the behavior of individual connectors; the application of the existing product operator to such CA models connector composition [9]. The formal definition and a congruence lemma, proved in [9], follow below.

**Definition 2.** The product operator, denoted by \( \_ \boxtimes _\_ \), is the
operator on \( \mathbb{C} \alpha \times \mathbb{C} \alpha \) defined by the following equation:

\[
\alpha \boxtimes \beta = \left( \text{State}(\alpha) \times \text{State}(\beta), \text{Port}(\alpha) \cup \text{Port}(\beta), \text{Mem}(\alpha) \cup \text{Mem}(\beta), \rightarrow, (\text{init}(\alpha), \text{init}(\beta)) \right)
\]

where \( \rightarrow \) denotes the smallest relation induced by Rule (1), Rule (3), and Rule (4) in Figure 5.

**Lemma 1.** \( [\alpha \approx \beta \text{ and } \gamma \approx \delta] \) implies \( \alpha \boxtimes \gamma \approx \beta \boxtimes \delta \)

Henceforth, let \( \boxtimes \{\alpha_1, \ldots, \alpha_k\} \) denote \( \alpha_1 \boxtimes \cdots \boxtimes \alpha_k \). This is well-defined, because \( \boxtimes \) is associative and commutative.

### III. Design: Theoretical Justification

#### A. Hybrid Connector Implementations

We start with presenting our design of hybrid connector implementations; in Sec. III-B, we discuss the design of a tool that automatically generates corresponding code.

To first better explain the different implementation approaches for connectors with \( \mathbb{C} \alpha \) as formal semantics, let \( X = \{\alpha_1, \ldots, \alpha_k\} \) denote a set of “small” \( \mathbb{C} \alpha \), each of which models one of the \( k \) primitive constituents of the connector \( \text{Conn} \) to implement. We associate with \( X \) an interpretation, denoted by \( [X] \), which models the composition of the constituents in \( X \) (i.e., the full behavior of \( \text{Conn} \)):

\[
[X] = \boxtimes X
\]

Every implementation of \( \text{Conn} \)—be it distributed, centralized, or hybrid—must behave as \( [X] \).

With the distributed approach, one first writes code for every \( \alpha \in X \) and then deploys those \( k \) \( \mathbb{C} \alpha \)-implementations in \( k \) parallel processing units (e.g., processes, threads, actors). To ensure that those \( \mathbb{C} \alpha \)-implementations behave as \( [X] \), at run-time, they must communicate with each other to check which of their transitions can fire: generally, the enabledness of a transition of one \( \mathbb{C} \alpha \)-implementation depends on the enabledness of transitions of other \( \mathbb{C} \alpha \)-implementations—an example follows shortly. Essentially, with the distributed approach, the \( k \) \( \mathbb{C} \alpha \)-implementations compute the \( \boxtimes \) operators amongst them dynamically at run-time. With the centralized approach, in contrast, one first computes \( \boxtimes X \) (i.e., the full behavior of \( \text{Conn} \)), then writes code for the resulting “big” \( \mathbb{C} \alpha \), and finally deploys this single \( \mathbb{C} \alpha \)-implementation in a single processing unit. By its construction, this single \( \mathbb{C} \alpha \)-implementation behaves as \( [X] \). Finally, with the hybrid approach, one first constructs a partition \( \mathcal{A} = \{A_1, \ldots, A_\ell\} \) of \( X \), then computes \( \text{Conn} \) for every part \( A \in \mathcal{A} \), then writes code for the resulting “medium” \( \mathbb{C} \alpha \), and finally deploys those \( \ell \) \( \mathbb{C} \alpha \)-implementations in \( \ell \) concurrent processing units. We associate with \( \mathcal{A} \) an interpretation, denoted by \( [\mathcal{A}] \), by straightforwardly lifting the interpretation of sets of \( \mathbb{C} \alpha \):

\[
[\mathcal{A}] = \left\{ [A_1], \ldots, [A_\ell] \right\}
\]

One can easily show that the interpretation of \( X \) equals the interpretation of any of its partitions (i.e., \( [X] = [\mathcal{A}] \)). Essentially, with the hybrid approach, a code generator computes the \( \boxtimes \) operators embedded in the “inner” interpretations statically at build-time (as in the centralized approach), while the \( \ell \) \( \mathbb{C} \alpha \)-implementations compute the \( \boxtimes \) operators embedded in the “outer” interpretation dynamically at run-time (as in the distributed approach).

To carry out the second, third, and fourth steps of the hybrid approach we can use existing techniques from the distributed/centralized approach. The current challenge, thus, lies in the first step: finding a reasonable partition of \( X \) in a potentially huge search space.\(^2\) At one extreme, if we put every \( \alpha \in X \) in its own part (i.e., \( \mathcal{A} = \{\{\alpha_1\}, \ldots, \{\alpha_k\}\} \)), we get the distributed approach, whose communication overhead reduces throughput at run-time. At the other extreme, if we put every \( \alpha \in X \) in the same part (i.e., \( \mathcal{A} = \{\{\alpha_1, \ldots, \alpha_k\}\} \)), we get the centralized approach, which sequentializes all parallelism. The hybrid approach should avoid both these unwelcome phenomena. As a compromise, we therefore adopt the following guideline for constructing partitions:

Put those \( \mathbb{C} \alpha \) whose implementations would require “expensive” communication at run-time in the same part; separate those \( \mathbb{C} \alpha \) that will require only “cheap” or no communication in different parts.

To exemplify the meaning of “expensive” and “cheap,” let Alice, Bob, and Carol be three \( \mathbb{C} \alpha \)-implementations (see also [11]). In particular, let Alice, Bob, and Carol implement three Sync primitives in sequence between ports A, B, and C. Suppose that Alice’s input port A has a pending write operation coming from the environment and that she wants to fire her \( \{A, B\}\)-transition (see Figure 4a). Operationally, to fire this transition, Alice must atomically take the data item written on A from that port and write it to B (see Table I). However, to guarantee atomicity, she must first ascertain that Bob is in fact ready to take a data item from B. So, before Alice takes the data item from A, she first asks Bob if he is ready to take from B. But Bob cannot immediately answer that question: he, in turn, must first ask Carol if she is ready to take a data item from C. In this simple example, Carol can answer Bob’s question without further derivative communication by locally checking if D has a pending take operation. Generally, however, the chain of derivative communication can be much longer (ignoring, for the moment, the possibility of loops, which can be resolved).

This makes the initial communication between Alice and Bob potentially very expensive. In summary, if Bob initiates derivative communication to answer a question from Alice, we call communication between Alice and Bob “expensive”

\(^2\)Theoretically, the total number of partitions of a \( k \)-cardinality set equals the \( k \)-th Bell number, denoted by \( B_k \), which grows rapidly in \( k \). For instance, the number of possible partitions for Alternator in Figure 1a, which consists of only 6 constituents, is \( B_6 = 203 \).
(and “cheap” otherwise).

To identify which CA require only cheap communication, we extend the recent local product theory from CA without data constraints to unrestricted CA [13]. The idea is to introduce a new product operator on CA, called \( l \)-product, which models connector composition with cheap communication. This contrasts the existing product in Definition 2, which models standard connector composition with potentially expensive communication: \( l \)-product is generally not faithful to Reo’s semantics. The formal definition and a congruence lemma, proved in [11, Appx. B], follow below. Afterward, we address the issue of determining when substituting the existing product with \( l \)-product is sound (i.e., faithful to Reo’s semantics).

**Definition 3.** The \( l \)-product (where “l” stands for “local”), denoted by \( _l \otimes _l \), is the operator on \( CA \times CA \) defined by the following equation:

\[
\alpha \otimes \beta := \left( \text{State}(\alpha) \times \text{State}(\beta), \text{Port}(\alpha) \cup \text{Port}(\beta), \text{Mem}(\alpha) \cup \text{Mem}(\beta), \rightarrow, (\text{init}(\alpha), \text{init}(\beta)) \right)
\]

where \( \rightarrow \) denotes the smallest relation induced by Rule (2), Rule (3), and Rule (4) in Figure 5.

**Lemma 2.** \( [\alpha \approx \beta \text{ and } \gamma \approx \delta] \implies [\alpha \otimes \gamma \approx \beta \otimes \delta] \)

The difference with Definition 2 is that the transition relation is induced by Rule (2) in Figure 5 instead of Rule (1).

To construct a reasonable—according to our guideline—partition \( \mathcal{A} = \{A_1, \ldots, A_t\} \), we must ensure that only cheap communication occurs among (implementations of interpretations of) parts in \( \mathcal{A} \) at run-time. Using the \( l \)-product just introduced, we can formally state this property as follows:

\[
[X] = [A] = [[[B_1]], \ldots, \lbrack B_m \rbrack, \lbrack C_1 \rbrack, \ldots, \lbrack C_n \rbrack\\
= [B_1] \otimes \cdots \otimes [B_m] \otimes [C_1] \otimes \cdots \otimes [C_n]\\
\approx [B_1] \otimes \cdots \otimes [B_m] \otimes [C_1] \otimes \cdots \otimes [C_n]
\]

**B. Hybrid-Code Generator**

The new task of a hybrid-code generator, in addition to the tasks of a centralized-code generator [3], is dividing the CA of the \( k \) primitive constituents of the connector to implement over parts in a partition \( \mathcal{A} \) such that the two conditions in Lemma 3 hold; it can subsequently use existing techniques to generate code for every part in \( \mathcal{A} \). To carry out this new task, a hybrid-code generator can run the algorithm in Figure 6 with time complexity upper bound \( O(k^2) \). The following lemma, proved in [11, Appx. D], establishes its functional correctness.

**Lemma 4.** Partition(\( X \)) = \( (B \cup C) \)

1) \( B \cup C \) is a partition of \( X \)
2) \( \rightarrow [B] \) for all \( B \in B \)
3) \( C \not= C' \implies [C] \not\approx [C'] \) for all \( C, C' \in C \)

The reasonable partitions \( \mathcal{A} = B \cup C \) computed by the algorithm in Figure 6 correspond to the synchronous and

---

3The second line of the premise in Rule (1) states that, if two transitions agree on their shared ports, they can fire together. These transitions may involve any number of other, unshared ports, which causes expensive communication when computing \( \otimes \) at run-time (e.g., Alice asks Bob about their shared port, but Bob wants to involve a port not shared with Alice for which he asks Carol). In contrast, the second line of the premise in Rule (2) restricts the involvement of unshared ports: either one of the transitions fires only shared ports, or both transitions fire only unshared ports.

4Mathematically, one can formally compensate for non-associativity of \( \otimes \) along the same lines as [13, Sec. 5], manifested as proper locking schemes at run-time, as mentioned in Sec. IV-A.
the asynchronous regions of a connector [6, 7]. First, every part \( B \in B \) seems to represent an asynchronous region (e.g., a FIFO primitive): the fact that \( [[B]] \) has only singleton synchronization constraints (i.e., \( \{1\} [[B]] \) ) models that its ports cannot—neither intentionally nor coincidentally—synchronize at run-time. Dually, every part \( C \in C \) seems to represent a synchronous region (e.g., the sequence of three Sync primitives modeled by Alice, Bob, and Carol in a previous example): by the duality of (a)synchronous regions, a CA for a synchronous region has at least one transition with a nonsingleton synchronization constraint. When such a transition fires, synchronization between at least two ports takes place.

IV. IMPLEMENTATION: PRACTICAL REALIZATION
A. Hybrid Connectors

We extended the existing Reo-to-Java code generator [3]. This tool translates CA to implementations of Java’s Runnable interface. Every such a CA-implementation can run in its own Java thread. The control-flow inside the main \( \text{run}() \) method follows a conceptually simple (event-driven) state machine pattern—details appear elsewhere [3]. We focus on the process of firing a transition \( \frac{P, f}{q' \Leftarrow \bot} \). The following enumeration of steps summarizes this process.

1) Check synchronization constraint \( P \).
   a) Check the ports in \( P \) for pending \( t/o \). Let \( P' \subseteq P \) denote the set of ports without pending \( t/o \).
   b) Ask neighbors with ports in \( P' \) which data constraints must hold for them to be ready for \( t/o \) on those ports. Let \( F \) denote that set of data constraints.

2) Solve extended data constraint \( f \land \neg \land F \).

3) Commit and conclude.
   a) Distribute data items among pending take operations according to the solution found for \( f \land \neg \land F \).
   b) Mark pending \( t/o \) completed, and update state to \( q' \).
   c) Perform \( t/o \) on ports in \( P' \) according to the solution found for \( f \land \neg \land F \).
    d) Notify neighbors with ports in \( P' \) that they should fire their transitions involving those ports.
    e) Await the completion of those transition.

During steps (1) and (2), a CA-implementation can still abort the firing process, and it will do so if either the synchronization constraint does not hold—in which case \( F \) contains \( \bot \)—or the data constraint has no solution under the then-pending \( t/o \) operations. Once step (3) starts, however, the transition must run to completion. To ensure that the whole firing process runs atomically, that pending \( t/o \) operations do not timeout during the firing process, and that neighbors do not change state, a CA-implementation uses a two-phase locking scheme [14].

Ports are implemented as interfaces that provide access to concurrent data structures called synchronization points [3].

Essentially, synchronization points register pending write and take operations. There exist two kinds of port interfaces: input ports expose only \( \text{write}(\ldots) \) methods for performing write operations, while output ports expose only \( \text{take}(\ldots) \) methods for performing take operations. Application developers can use input and output ports for letting concurrent fragments of their computation code interact with each other via a CA-implementation (i.e., via a connector). Internal classes in the Reo run-time libraries, as well as generated CA-implementations, call also other methods on input and output ports (e.g., checking for pending \( t/o \) operations, communicating with neighbors via synchronization points).

The run-time libraries of the original implementation of the Reo-to-Java code generator contain only one shared-memory implementation of input and output ports. Figure 7a shows an infographic, where “LSP” stands for “local synchronization point.” This name reflects that the Java objects constituting such a synchronization point live on the same machine as the CA-implementations that access that synchronization point through input and output ports. Logically, a triple of an LSP, an input port, and an output port makes up a single port.

To enable developers to fully exploit the improved parallelism in the hybrid connector implementations generated by our tool extension (cf. the centralized approach), we implemented a new distributed-memory implementation of input and output ports. In particular, this allows developers to deploy connector implementations generated by our tool.
extension on different machines in a network. Figure 7b shows an infographic, where “RSP” stands for “remote synchronization point,” “-S” for “server,” “-IC” for “input client,” and “-OC” for “output client.” Deployment of a remote synchronization point starts with deploying an RSP-S on some machine in the network. Essentially, an RSP-S is a Web service, implemented using JAX-WS, whose operations provide its clients access to a “classical” LSP inside of it. Once an RSP-S has been deployed, one can construct input and output ports to access it, including input and output clients. During execution, those port interfaces use such clients for delegating, to the deployed RSP-S, those method calls that they cannot process locally (e.g., checking for pending I/O operations). The CA-implementations that call methods on port interfaces do not know whether those calls require network communication: whether synchronization points accessed through port interfaces run locally or remotely is completely transparent.

B. Hybrid Code Generator

Our extension to the Reo-to-Java code generator is— as the original—implemented in Java as an Eclipse plug-in. This plug-in depends on the Extensible Coordination Tools (ECT): a collection of Eclipse plug-ins that constitute an IDE for Reo. Our extension outputs \( \ell \) Java classes—one for every part in the reasonable partition computed using the algorithm in Figure 6—each of which implements the Runnable interface. Instances of those classes can be deployed on different machines and connected to each other via distributed-memory ports (including the required remote synchronization point servers and clients). For testing purposes, the code generator also generates a default main program which deploys an instance of each of the \( \ell \) classes and each of the required distributed-memory ports on the same machine.

For performance reasons, we simplified the implementation of the algorithm in Figure 6 and the actual code generation process (using ANTLR’s StringTemplate engine [15]) by exploiting the observation that the only primitive currently supported by the ECT that satisfies the first condition in Lemma 3 is FIFO. This, for instance, reduced checking the condition of the if-statement in Figure 6 from iterating over all transitions of a CA to checking that CA’s type.

V. CASE STUDY

In the following case study, we elaborate on the same example as given in [8], which implements a classical online-purchase scenario. This interaction involves four Web services (WS) named ClientBroker, StoreOffice, SalesOffice, and Bank. The ClientBroker service takes care of interfacing a client to the other services, which deal with: the information about the store (i.e., the StoreOffice service), the procedure to prepare the invoice (i.e., the SalesOffice service), and the effective payment management (i.e., the Bank service). See [8] for a more detailed description. Figure 7a shows a Reo Connector, named Orchestrator, for orchestrating the four WSs. A Reo expert designed this connector by hand. Alternatively, depending on the expertise available in an organization, one can specify the orchestration protocol as an automaton, as a BPEL program, or as a UML sequence/activity diagram, and use mechanical connector synthesis technology to obtain this (or a behaviorally equivalent) connector [16, 17].

In [8], Jongmans et al. generate a centralized implementation of Orchestrator and deploy that implementation on a single machine in a network. We improve on that by using our hybrid-code generator to obtain a hybrid implementation and deploy it across multiple machines.

First, our code generator establishes that Orchestrator consists of 43 channels and 42 nodes. It then matches each of those constituents with a CA describing that constituent’s behavior, which yields \( k = 85 \) small CA. For instance, Figure 9 shows seven CA describing the constituents in the largest colored region in Figure 8. Next, our code generator runs the algorithm in Figure 6 to obtain a reasonable partition. This partition consists of thirteen parts: seven singleton parts for asynchronous regions (i.e., FIFOs, which satisfy \( \rightarrow \)) and six parts for synchronous regions. For instance, the CA in Figure 9 form a complete part of the latter kind (none of those seven CA shares a port with any CA in any other part). The code generator then computes, for every part, the \( \boxplus \)-product of the CA in that part. This yields \( \ell = 13 \) medium CA. For instance, Figure 10 shows the \( \boxplus \)-product of the CA in Figure 9. Finally, the code generator compiles the thirteen medium CA to as many CA-implementations. For instance, at run-time, the generated implementation of the CA in Figure 10 has only one execution step (i.e., transition) that it repeats infinitely often: it atomically transports a piece of data from StoreOffice to SalesOffice and, simultaneously, transports a piece of data from the output port of one FIFO to the input port of another one. The latter ensures that
of input and output ports, which are shared with channels.

Next, we discuss the deployment of the generated code. Logically, we have five machines: called Red (for the actual client), Green (for ClientBroker), Blue (for StoreOffice), Cyan (for SalesOffice), and Magenta (for Bank). Although any distribution of the thirteen generated CA-implcations (and the ports between them, through which CA-implcations communicate with each other) over the five available machines would technically work, some of those distributions make more sense than others. Generally, the problem of (automatically) optimally distributing CA-implcations over machines is an interesting research challenge, which we regard as important future work (see also Sec. VII). For now, we adopted the following ad hoc approach: to minimize network traffic, we manually distributed CA-implcations in such a way that every piece of data goes over the network exactly once. In contrast, in the centralized implementation in [8], every piece of data goes over the network twice: first from the sending WS machine to the Orchestrator machine and then from the Orchestrator machine to the receiving WS machine. Thus, the hybrid approach can improve the centralized approach not only in terms of run-time parallelism but, as a consequence of less network traffic and no single-point contention, also in terms of throughput (especially with large data). Figure 8 shows the deployment of parts of the orchestration on machine Blue; see [11, Appx. A] for the full distribution.

For deploying ports, we first analyzed which ports are shared between CA-implcations running on the same machine. We deployed those ports as shared-memory ports (similar to [8]) and all other ports as distributed-memory ports. For every distributed-memory port p, we deployed its remote synchronization point server (RSP-S) on the same machine as the CA-implcation that uses p’s input interface. This means that pieces of data involved in write operations on p stay on that machine as long as no other CA-implcation (running on a different machine) wants to take that data from p’s RSP-S (via p’s output interface). Consequently, pieces of data travel over the network only if absolutely necessary.

To actually run this case study on five machines, we use Amazon Web Services (AWS), which is a collection of services that together make up a Cloud computing platform, offered over the Internet by Amazon.com. In particular, we take advantage of Elastic Compute Cloud (EC2): a WS that provides resizable compute capacity in the Cloud. It is designed to make Web-scale computing easier for developers, by allowing them to rent virtual machines on which to run their own applications. We rented five of those virtual machines and deployed the generated code for this case study as explained above. This shows that the framework underlying Reo can be integrated with Cloud technologies (i.e., “Reo goes to the Cloud”).

Actually, by using the Cloud in this way, we can position the example in [8] in the following new scenario. Suppose we run a (large) online business company that offers a purchase service through AWS: the motivation is that our company desires to scale its business in the Cloud, benefiting from a third-party infrastructure that can efficiently manage Big Data. For instance, combined with EC2, our company can use another Cloud service from the AWS platform called Simple Storage Service (S3): an online technology for managing large amounts of information at any time, capable of handling Big Data such as transaction logs of our clients’

---

http://aws.amazon.com
orders. Because the hybrid approach allows us to sensibly distribute CA-implementations over machines (minimizing network traffic, as explained above), it is—contrasting the centralized approach—suitable for Big Data coordination.

VI. Related Work

In this section, we list a selection of related work on Reo and distributed orchestration/workflow. For reasons of space, here, we merely describe related work without comparing it to our work; see [11] for additional comparative notes.

A. Related Work on Reo

Closest to ours is the work on splitting connectors into (a)synchronous regions for better performance. Proença developed the first implementation based on these ideas, demonstrated its merit through benchmarks, and invented a new automaton model to reason about split connectors [5, 6].

B. Related Work on Distributed Orchestration/Workflow

To put our previous case study in perspective, we conclude with related work on distributed orchestration/workflow.

In [18], Nanda et al. present a technique for partitioning a composite service written as a single BPEL [19] program into an equivalent set of decentralized processes. In their approach, Nanda et al. construct the dependency graph of a BPEL program with the aim of minimizing communication costs while maximizing throughput. In [20], Chafle et al. decentralize the orchestration of a FindRoute service by partitioning the BPEL code into four parts, which are executed by four distinct Java engines. Afterward, Chafle et al. compare the performance of the centralized and the decentralized implementation by using both service-time and message-size metrics. Chafle et al. also estimate the additional complexity in error recovery and fault handling in a decentralized orchestration. In [21], Mostarda et al. use a BPEL-based language for distributed orchestration in the context of pervasive computing. In their approach, Mostarda et al. automatically decompose a centralized workflow into implementations of finite state machines that, when synchronized using a consensus protocol, execute the original workflow. In [22], Fernández et al. introduce an execution model for distributed orchestration based on a metaphor from chemistry. In this approach, services communicate with each other through a distributed shared multiset ("chemical solution"), which contains both control-flow and data-flow information ("molecules"). A workflow engine then executes an orchestration protocol by applying formal rewrite rules ("chemical reactions"). Fernández et al. propose translating BPEL programs to chemical representations that the workflow engine they describe can process.

In [23], Barker et al. present an architecture (including a proxy API) for optimizing data-flow in workflow execution. In their architecture, control-flow messages are still sent to a centralized orchestration engine, while data-flow messages (i.e., the actual, potentially large, pieces of data) go directly from one service to another. The latter reduces network traffic, and it coincides with the heuristic we adopted for distributing the CA-implementations over machines in our case study. Similarly motivated as the work of Barker et al. (i.e., optimizing data-flow), in [24], Binder et al. propose a distributed orchestration methodology based on decomposing an orchestration protocol (or workflow) into a directed acyclic graph of service invocations, represented as triggers. Every trigger encodes the data dependencies of the invocation it represents (i.e., the parents and children of the corresponding node in the graph). Triggers act as proxies: they collect all input data before actually invoking the service, and they transmit all output data directly to the dependent triggers for subsequent service invocations.

In [25], Tretola and Zimeo present a technique for improving concurrency in workflow execution. Their technique works by concurrently executing otherwise sequential, data-dependent service invocations with dummy data until those services require the actual data. If service invocations have a data-independent initialization phase, for instance, the workflow engine of Tretola and Zimeo executes those initialization phases in parallel instead of in sequence. For workflows and services implemented on top of the framework of Tretola and Zimeo, this process happens automatically.

In [26], Pedraza and Estublier present FOCAS. This framework consumes as input an annotated APEL [27] specification of an orchestration scenario and produces as output a number of suborchestrations and a deployment plan (for distributing suborchestrations over machines).

Finally, in [28], Muth et al. use the formal semantics of state/activity charts to develop an algorithm for transforming centralized state/activity charts into equivalent partitioned ones, suitable for distributed execution by a workflow engine. Muth et al. subsequently refine their basic approach to also reduce communication overhead and exploit parallelism between parts in partitions.

VII. Conclusion

We presented a hybrid approach for the implementation of Reo connectors by partitioning them into several (a)synchronous regions at build-time. Every such region can be executed on a different machine at run-time. We use the term “hybrid” because our approach is neither purely centralized (regions run in parallel at run-time), nor purely distributed (the elements inside a region are compiled to a sequential program at build-time and require no distributed algorithms or communication at run-time). In this way, we have the benefits of both pure approaches: the high run-time

8Nowadays, these qualify as Big Data: “Wal-Mart handles more than one million customer transactions every hour, feeding databases estimated at more than 2.5 petabytes” (http://www.economist.com/node/15557443).

In [23], Barker et al. present an architecture (including a proxy API) for optimizing data-flow in workflow execution. In their architecture, control-flow messages are still sent to a centralized orchestration engine, while data-flow messages
throughput of a centralized scheme combined with the high parallelism and fast compilation of a distributed scheme.

In the future, we plan to design an algorithm to automatically find the best partitioning of CA-implementations according to user and system-defined constraints. These constraints may need to be either "crisply" or "softly" satisfied (in relation to their indispensability), and concern different criteria as hardware requirements of software, QoS/QoE/performance desiderata, and issues correlated with security (e.g., preventing attacks based on Business Process Discovery), privacy (of both data and workflow), and fault handling.

REFERENCES


