The Remote Monad Design Pattern

Andy Gill  Neil Sculthorpe  Justin Dawson  Aleksander Eskilson
Andrew Farmer  Mark Grebe  Jeffrey Rosenbluth
Ryan Scott  James Stanton
Information and Telecommunication Technology Center, University of Kansas, USA
first.last@ittc.ku.edu

Abstract
Remote Procedure Calls are expensive. This paper demonstrates how to reduce the cost of calling remote procedures from Haskell by using the remote monad design pattern, which amortizes the cost of remote calls. This gives the Haskell community access to remote capabilities that are not directly supported, at a surprisingly inexpensive cost.

We explore the remote monad design pattern through six models of remote execution patterns, using a simulated Internet of Things toaster as a running example. We consider the expressiveness and optimizations enabled by each remote execution model, and assess the feasibility of our approach. We then present a full-scale case study: a Haskell library that provides a Foreign Function Interface to the JavaScript HTML5 Canvas. Finally, we discuss existing instances of the remote monad design pattern found in Haskell libraries.

Categories and Subject Descriptors D.3.2 [Programming Languages]: Language Classifications—Applicative (functional) languages

Keywords Monads, Remote Procedure Call, FFI, Design Pattern.

1. Introduction
Remote Procedure Calls (RPCs) are expensive. This paper presents a way to make them considerably cheaper: the remote monad design pattern. Monads provide a general way of structuring composite computations. These monadic computations are first class: they can be passed to functions as arguments, returned from functions as results, and stored in data structures. But monadic computations are not typically executed remotely.

This paper investigates the ways that monadic computations can be serialized for the purposes of being sent to remote locations for external execution. The idea is that, rather than directly call a remote procedure, we instead give the remote procedure call a service-specific monadic type, and invoke the remote procedure call using a monadic "send" function.

Definition. A remote monad is a monad that has its evaluation function in a remote location, outside the local runtime system.

By factoring the RPC into sending invocation and service name, we can group together procedure calls, and amortize the cost of the remote call. To give an example, Blank Canvas, our library for remotely accessing the JavaScript HTML5 Canvas, has a send function, lineWidth and strokeStyle services, and our remote monad is called Canvas:

\[
\text{send} :: \text{Device} -> \text{Canvas a} \to \text{IO a}
\]
\[
\text{lineWidth} :: \text{Double} -> \text{Canvas ()}
\]
\[
\text{strokeStyle} :: \text{Text} -> \text{Canvas ()}
\]

If we wanted to change the (remote) line width, the lineWidth RPC can be invoked by combining send and lineWidth:

\[
\text{send device (lineWidth 10)}
\]

Likewise, if we wanted to change the (remote) stroke color, the strokeStyle RPC can be invoked by combining send and strokeStyle:

\[
\text{send device (strokeStyle "red")}
\]

The key idea of this paper is that remote monadic commands can be locally combined before sending them to a remote server. For example:

\[
\text{send device (lineWidth 10 >> strokeStyle "red")}
\]

The complication is that, in general, monadic commands can return a result, which may be used by subsequent commands. For example, if we add a monadic command that returns a Boolean,

\[
\text{isPointInPath :: (Double,Double) -> Canvas Bool}
\]

we could use the result as follows:

\[
\text{send device $ do}
\]
\[
\text{inside <- isPointInPath (0,0)}
\]
\[
\text{lineWidth (if inside then 10 else 2)}
\]
\[
\ldots
\]

The invocation of send can also return a value:

\[
\text{do res <- send device (isPointInPath (0,0))}
\]
\[
\ldots
\]

Thus, while the monadic commands inside send are executed in a remote location, the results of those executions need to be made available for use locally.
2. Haskell Calling the Real World

When Haskell users want an established graphics library, such as OpenGL, or a web-client library, such as cURL, they typically use a Haskell library that uses the Foreign Function Interface (FFI) capability to link Haskell code with the native capabilities of the desired external library. Rather than reimplement existing code, the rich FFI capability is used to tunnel to C, and onwards to established libraries, such as the C or C++ implementations of OpenGL and cURL. The FFI is well-supported, but there are three conceptual problems to be solved in crossing to native C libraries:

1. Control flow needs to transfer to the correct C function. Given the lowest level of the GHC runtime system is written in C, control-flow transfer is straightforward.
2. The data structures that are arguments and results of calls to C need to be marshalled into the correct format. For example, C strings are not the same as Haskell strings.
3. The abstractions of the target C library may not be idiomatic Haskell abstractions. For example, many C++ APIs assume OO-style class inheritance. Foreign abstractions can be simulated in Haskell, but this raises an obfuscation barrier.

Any time control flow leaves the Haskell eco-system, all three of these concerns need to be addressed. All three are supported by the Haskell FFI for C. There is a way of directly promoting a C function into Haskell-land; there is good support for marshalling Haskell data structures into C structures, as well as automatic memory-management support; and Haskell abstraction capabilities are used to build more Haskell-centric APIs on top of the FFI capability. However, what about functions that cannot be called directly, but must be invoked using a remote procedure call? All three identified FFI issues come into play:

1. The control flow needs to use an RPC mechanism to establish foreign control-flow transfer. This can add considerable costs. A typical mechanism would be to establish a TCP/IP connection with a published remote service.
2. The procedure arguments are sent to the remote location, typically over a TCP/IP channel, the remote command is executed, and the result is communicated back.
3. The remote nature of the call raises issues with presenting remoteness to the API user. What does this remote function look like to a Haskell programmer? Are the costs reflected in the API design?

In this paper we investigate a generalization of the remote procedure call, using monads as a descriptor of remote execution. Specifically, we make the following contributions:

- We document a DSL pattern for bundling RPCs: the remote monad design pattern. This pattern cuts the cost of RPCs in many cases by amortizing the remote aspect, and is a design point between classical monadic APIs, and using deeply embedded DSLs for offline remote execution.
- Toward understanding the remote monad, we give four complete models of remote execution (§5.1, §5.2, §5.3, §5.4) and sketch two useful variations of the strongest model (§5.5). We give a set of remote monad laws (§5.6), and observe that the pattern has been used in a weak form several times before (§5.7).
- We explore the design pattern in a real-world large-scale example, called Blank Canvas (§5.8). We quantify our experiences, and measure performance on some benchmarks, comparing native JavaScript and Blank Canvas, over two operating systems, and two web browsers (§5.9).

3. Modeling Remote Communication

For illustration purposes, we will remotely control an Internet of Things toaster, which includes a thermometer and a voice synthesis circuit. We can ask the toaster to say things, measure the temperature, or make toast for a given number of seconds. Using a monad called Remote, we have:

```haskell
say :: String -> Remote ()
temperature :: Remote Int
toast :: Int -> Remote ()
```

In the remote monad design pattern, we have a way of running our monad remotely, which we call send:

```haskell
send :: Device -> Remote a -> IO a
```

The key step in an efficient implementation of the remote monad is the choice of packaging of commands. Ideally, the entire monadic computation argument of send would be transmitted in a single packet; but in general this is not possible.

Towards understanding the choices, we are going to build several ways of interacting with our toaster:

- As a preamble, in §3.1 we build a basic asynchronous RPC, and in §3.2 we build a basic synchronous RPC.
- In §3.3 we build the simple version of a remote monad, called a weak remote monad, where a monadic API is utilized to conceptually bundle commands for remote execution.
- In §3.4 we build another remote monad, called a strong remote monad, where a monadic API is utilized to actually bundle commands for more efficient remote execution. This is our principal model of the key idea of this paper.
- In §3.5 we build a remote applicative functor, which exploits the restrictions of the applicative functor operations (relative to the more expressive monadic bind) to allow for better bundling.
- Finally, in §3.6 we use a deep-embedding technique to add support for remote binding of non-local values, again improving the bundling of remote primitives.

In each case, we implement the local behavior of the send function, as well as give a minimal implementation of remote behavior. In general, in the remote monad design pattern, the remote execution may be in any language. Here, for simplicity, we simulate remote execution from within Haskell, and use the built-in Show and Read classes to simulate serialization and deserialization. We highlight the dual execution nature of the remote monad by using distinct data structures locally and remotely. Using distinct data structures also allows us to sidestep the orthogonal challenge of GADT deserialization.

3.1 An Asynchronous Remote Command Call

This first model will be able to send single commands asynchronously to a remote device. This model is also known as the Command design pattern [19].

Definition. A remote command is a request to perform an action for remote effect, where there is no result value or temporal consequence.

We use a deep embedding of the commands that we want to send asynchronously. For this simple example, we only consider a single command:

```haskell
data Command = Say String
              deriving Show
```

We derive a Show instance so that we can serialize Command, in preparation for transmitting it to a remote interpreter.
We represent a remote device as a function from String to IO (), wrapped in a data structure called Device, which models the communication channel between local and remote execution.

\[
data \text{Device} = \text{Device} \{ \text{async} :: \text{String} \to \text{IO} () \}\]

We call the internal function async because it represents asynchronous communication. For asynchronous commands, there is no need for a back-channel on which to return values.

We can now define send, which serializes a Command and then transmits it to a remote device:

\[
\text{send} :: \text{Device} \to \text{Command} \to \text{IO} ()
\]

\[
\text{send} \ d \ m = \text{async} \ d \ (\text{show} \ m)
\]

Our Haskell simulation of the remote device requires a representation of commands on the remote device, an execution function for those commands, and a deserialization function that reads those commands. As a convenience, we use the same constructor name for the remote command data type, so that deserialization can be handled by a derived Read instance:

\[
data \text{RCommand} = \text{Say} \ \text{String} \\
\text{deriving Read}
\]

\[
\text{execRCommand} :: \text{RCommand} \to \text{IO} ()
\]

\[
\text{execRCommand} \ (\text{Say} \ \text{str}) = \text{putStrLn} \ (\text{"Remote: " ++ str})
\]

Our device, which simulates the remote interpreter handle, can then be defined as:

\[
\text{device} :: \text{Device}
\]

\[
\text{device} = \text{Device} \ (\text{execRCommand} \ . \ \text{read})
\]

This completes our model. Now we can test it by sending a Say command, which prints remotely:

\[
\text{GHCi > send device} \ (\text{Say} \ \text{"Do you want some toast?”})
\]

Remote: Do you want some toast?

In summary, this model is a direct implementation of asynchronous remote calls. Figure 1 shows the interactions in this example invocation using a sequence diagram. We take the liberty of expressing our construction of Command as a sequence process, and we give the serialized text sent to the remote device, using green for an asynchronous call.

### 3.2 A Synchronous Remote Call

In this subsection we build a synchronous remote call: a version of send that can receive a reply to a remotely transmitted procedure. That is, we will define a model of a remote procedure call.

**Definition.** A **remote procedure** is a request to perform an action for its remote effects, where there is a result value or temporal consequence.

When we execute a remote procedure, we either want to get a result back (e.g., the measured temperature), or know that a specific remote action has been completed (e.g., the toast is made). In this model, because we are interested in getting a reply, we represent remoteness using a function from String to IO String.

\[
data \text{Device} = \text{Device} \{ \text{sync} :: \text{String} \to \text{IO} \text{String} () \}\]

As with commands, we use a deep embedding of the procedures that we want to send. However, as we now expect to receive a result in reply, we use a GADT with a phantom type index denoting the expected result type:

\[
data \text{Procedure} :: \exists \ a \to \text{Procedure} \text{Int}
\]

\[
\text{Temperature} :: \text{Procedure} \text{Int}
\]

\[
\text{Toast} :: \text{Int} \to \text{Procedure} ()
\]

As with commands, we provide serialization using the Show class:

\[
\text{deriving instance Show (Procedure \ a)}
\]

For deserialization, we provide an auxiliary function that uses the phantom type index of Procedure to determine which Read instance should be used to parse the reply from the remote device:

\[
\text{readProcedureReply} :: \text{Procedure} \text{a} \to \text{String} \to \text{a}
\]

\[
\text{execRProcedureReply} \ (\text{Temperature} {}) \ i = \text{read} \ i
\]

\[
\text{execRProcedureReply} \ (\text{Toast} {}) \ i = \text{read} \ i
\]

The two read functions are each reading different types, as constrained by the specific Procedure constructor.

Now we can write our send command:

\[
\text{send} :: \text{Device} \to \text{Procedure} \text{a} \to \text{IO} \text{a}
\]

\[
\text{send} \ d \ m = \text{do}
\]

\[
\text{r} \leftarrow \text{sync} \ d \ (\text{show} \ m)
\]

\[
\text{return} \ (\text{readProcedureReply} \ m \ r)
\]

That is, send serializes the Procedure, sends it over a synchronous channel, and interprets the reply in terms of the type of the same Procedure. This time, send is polymorphic in the type parameter of Procedure and IO.

Our simulated remote Device is straightforward to construct. For conciseness we have merged serialization of the result value into the execution function, but these steps could be separated.

\[
data \text{RProcedure} = \text{Temperature} \\
| \text{Toast} \text{Int} \\
\text{deriving Read}
\]

\[
\text{execRProcedure} :: \text{RProcedure} \to \text{IO} \text{String}
\]

\[
\text{execRProcedure} \ (\text{Temperature} \ i) = \text{do}
\]

\[
\text{t} \leftarrow \text{randomRIO} \ (50, 100 :: \text{Int})
\]

\[
\text{return} \ (\text{show} \ t)
\]

\[
\text{execRProcedure} \ (\text{Toast} \ n) = \text{do}
\]

\[
\text{threadDelay} \ (1000 \cdot 1000 \cdot n)
\]

\[
\text{putStrLn} \ (\text{"Remote: Toasting..."})
\]

\[
\text{return} \ (\text{show} \ (\text{))}
\]

This completes our model. Now we can test it by sending a Temperature procedure, which returns the temperature locally:

\[
\text{GHCi > send device Temperature}
\]

We have taken a procedure, transmitted it to a remote interpreter, executed it remotely, and returned with the result — thus we have a basic model of a typical Remote Procedure Call. Figure 2 shows the sequence diagram for this interaction. We use red to highlight the synchronous communication call.
4. The Weak Remote Monad

Thus far, our arguments to send have not been instances of Monad. We will now address this. Here we join the asynchronous and synchronous models into a monad called Remote. In this section, we implement an initial version of the remote monad that does not amortize the cost of communication. We call this a weak remote monad.

Definition. A weak remote monad is a remote monad that sends each of its remote calls individually to a remote interpreter.

In this example, the Remote monad is implemented using the reader monad, where the environment is our Device, nested around the IO monad, using monad transformers [22, 27].

```haskell
newtype Remote a = Remote (ReaderT Device IO a)
deriving instance Monad Remote
```

This gives us access to the specific Device, which will allow us to send individual commands to the remote interpreter on the fly. We use deriving instance Monad Remote

Device now needs to support both sync and async. In a real implementation, it would be expected that there would be two entry points into the same communication channel, where the usage specifics depend on if we want to receive a reply.

```haskell
data Device = Device
  { sync :: String -> IO String,
    async :: String -> IO ()
  }
```

A send-style function, customized for our weak Remote monad, provides remote execution for Command. Observe that each command invokes the remote procedure call immediately.

```haskell
sendCommand :: Command -> Remote ()
sendCommand m = Remote $ do
  d <- ask
  liftIO (async d (show m))
  return ()
say :: String -> Remote ()
say txt = sendCommand (Say txt)
```

For procedures, we also provide a send-style function, again customized for our weak Remote monad. Again, each procedure invokes the remote procedure call immediately.

```haskell
sendProcedure :: Procedure a -> Remote a
sendProcedure m = Remote $ do
  d <- ask
  r <- liftIO (sync d (show m))
  return (readProcedureReply m r)
```

```haskell
temperature :: Remote Int
temperature = sendProcedure Temperature
```

```haskell
toast :: Int -> Remote ()
toast n = sendProcedure (Toast n)
```

Our main send function is now used to run the Remote monad, unboxing the reader function, and applying it to the Device:

```haskell
send :: Device -> Remote a -> IO a
send d (Remote m) = runReaderT m d
```

Finally, the virtual remote Device is a combination of the synchronous and asynchronous Devices:

```haskell
device :: Device
device = Device (execRProcedure . read) (execRCommand . read)
```

Now we can call send with a monadic argument, and chains of primitives, connected using the monad, will be executed. We have achieved our original goal: a (weak) remote monad where the primitive commands and procedures are executed in a remote location.

```
GHCi> t <- send device $ do
  say "Do you want some toast?"
  t <- temperature
  say (show t ++ "F")
  return t
Remote: Do you want some toast?
Remote: 56F
GHCi> when (t < 70) (send device (toast 120))
Remote: Toasting...
...sleeping for 120 seconds...
Remote: Done!
```

Figure 3 shows the sequence diagram for this example of the weak remote monad. For every primitive call invoked, there is a remote call to the device. Furthermore, the GHCi computation does not terminate until the toast is complete.
5. The Strong Remote Monad

We want to bundle monadic remote calls, and send them as packets of computations to be remotely executed. The strong remote monad does this. We have two classes of primitive remote calls: commands that do not require any specific result to be observed, and procedures that require a reply from the remote site.

- For commands, which are asynchronous and do not send a reply, we can queue up the command and commit to sending it later.
- For procedures, which are synchronous, we need to transmit them immediately. Thus we first send all outstanding queued commands, then send the procedure, and then await the procedure’s result.

At the end of executing a send, we flush the queue of any outstanding commands.

This is the key idea behind a strong remote monad: package the sequence of monadic actions into a list of commands, which are (in all but the final case) terminated by procedures. This design assumes that the only primitive remote calls in our remote monad are commands and procedures, and thus there is no way of locally pausing or stalling the pipeline of commands. The queuing of commands is simply a bundling strategy for commands and procedures that would be executed immediately after each other anyway.

**Definition.** A **strong remote monad** is a remote monad that bundles all of its remote calls into packets of commands, punctuated by procedures, for remote execution.

In the strong remote monad, as well as knowing what Device to talk to, we need to queue up the list of to-be-transmitted Commands. We use a reader monad for the Device (as with the weak remote monad), and a state monad for the command queue. Thus, our (strong) remote monad can be defined as:

```haskell
newtype Remote a = Remote (ReaderT Device (StateT [Command] IO) a)
```

We want to send a packet of Commands, terminated by a Procedure, and also to have a way of sending a packet of Commands without a terminating Procedure. For the latter case, we simply send [Command]. For the former case, we introduce a dedicated data type:

```haskell
data Packet a = Packet [Command] (Procedure a)
deriving Show
```

Now we can start providing monadic primitives. A utility function, `sendCommand`, appends Commands to our “to-be-sent” queue, and say supports the remote monad interface:

```haskell
sendCommand :: Command -> Remote ()
sendCommand cmd = Remote (modify (++ [cmd]))
say :: String -> Remote ()
say txt = sendCommand (Say txt)
```

Supporting procedures is more involved. We need to flush the queue of outstanding Commands, as well as to actually send a packet containing the Commands and procedure to the remote site:

```haskell
sendProcedure :: Procedure a -> Remote a
sendProcedure p = Remote $ do
d <- ask
cs <- get
r <- liftIO (sync d (show (Packet cs p)))
put []
return (readProcedureReply p r)
temperature :: Remote Int
temperature = sendProcedure Temperature
toast :: Int -> Remote ()
toast n = sendProcedure (Toast n)
```

Next, we provide our `send` function, which runs our inner monad, and flushes any remaining Commands to the remote interpreter:

```haskell
send :: Device -> Remote a -> IO a
send d (Remote m) = do
  (r,cs) <- runStateT (runReaderT m d) []
  when (not (null cs)) (async d (show cs))
  return r
```

Finally, we define our simulated remote device, now extended with an execution function for packets:

```haskell
device :: Device
device = Device (execRPacket . read) (mapM_ execRCommand . read)
data RPacket = Packet [RCommand] RProcedure
deriving Read
execRPacket :: RPacket -> IO String
execRPacket (Packet cs p) = do
  mapM_ execRCommand cs
  execProcedure p
```

Figure 4 shows the sequence diagram for the strong remote monad, on the same example as used for the weak remote monad. As can be seen, Packet combines Commands, punctuated by Procedures.
6. The Remote Applicative Functor

There is also a remote applicative functor.

**Definition.** A **remote applicative functor** is an applicative functor that has its evaluation function in a remote location, outside the local runtime system.

As with monads, there are two classes: the **weak** remote applicative functor, and the **strong** remote applicative functor. Without a bind operator, applicative functors [30] are fundamentally better suited to remoteness than monads are: subsequent applicative computations cannot depend on the results of prior computations, which in our context allows for bundling of procedures. In fact (using our terminology) a strong remote applicative functor is currently used by Facebook to bundle database requests [29].

Any weak remote monad is, by definition, (at least) a weak remote applicative functor, with primitive calls transmitting individually. More interesting is the **strong** remote applicative functor, which we consider here. Exploiting the independence of subsequent calls from the results of prior calls, we are going to bundle all the Commands and Procedures together in a single packet, which we represent by a list of Prims (primitive remote calls).

```
data Prim :: * = Command CommandToPrim Procedure ProcedureToPrim
  deriving instance Read Prim

newtype Remote a = Remote (WriterT (State String) Prim a)
```

This monad transformer is a combination of the writer monad, for accumulating queued primitive calls, and the state monad, for simultaneously parsing results. We will use lazy evaluation to "tie-the-knot" [2] between these two effects in `send`. We explicitly provide the instances for Applicative and Functor, but not Monad.

We handle **Commands** using the underlying writer monad:

```
sendCommand :: Command -> Remote ()
sendCommand cmd = Remote (tell [Command cmd])
```

Handling **Procedures** is a bit more involved. We use the remote monad to remember the `Procedure`, then read the result from the state, lazily pulling it off a list. Crucially, there are not any external effects inside our inner monad.

```
sendProcedure :: Show a => Procedure a -> Remote a
sendProcedure p = Remote $ do
  tell [Procedure p]
  r <- execRProcedure p
  put rs
  return (readProcedureReply p r)
```

We build `send` using recursive do notation [16] [17]. If our list of Prims contains only Commands, then we send it asynchronously to our remote device. Otherwise, we send it synchronously, and await the list of results from the remote device. We then recursively feed this result list back into the invocation of the applicative functor. Whether synchronous or asynchronous, we bundle the complete list of Prims as a single packet.

```
send :: Device -> Remote a -> IO a
send d (Remote m) = do
  let ((a,p),...) = runState $ runWriterT m r
  do async d (show ps) return []
    else do str <- sync d (show ps)
     return (read str)

return a
```

We are able to transmit Procedures as a bundle because we cannot examine the result on individual Procedures until we have the entire applicative functor computation result — a key property of applicative functors.

Our simulated remote device now needs to handle lists of intermingled commands and procedures:

```
data RPrim = Command RCommand
  | Procedure RProcedure
  deriving Read RPrim

execRPrims :: [RPrim] -> IO [String]
execRPrims [] = return []
execRPrims (Command c : ps) = do
  execRCommand c
  execRPrims ps
execRPrims (Procedure p : ps) = do
  r <- execRProcedure p
  rs <- execRPrims ps
  return (r:rs)
```

The device implementation uses `execRPrims` for both synchronous and asynchronous requests:

```
device :: Device
device = Device (liftM show . execRPrims . read)
          (void . execRPrims . read)
```

We can now transmit bundles containing both commands and procedures. For example, we can measure the temperature before and after toasting in one request:

```
GHCi> (t1,t2) <- send device $
  liftA2 ()
  (say "Good Morning" => temperature)
  (toast 120 => temperature)
Remote: Good Morning
Remote: Toasting...
...sleeping for 120 seconds...
Remote: Done!
GHCi> print (t1,t2)
(56,99)
```

Figure [5] shows the sequence diagram for this example. Notice how all three Procedures (and the Command) are sent to the toaster in a single packet.
7. Remote Binding

An alternative to combining procedures into one super-procedure with restrictions, would be to have the remote interpreter directly use the result of a procedure. We achieve this by introducing a remote binding mechanism, which locally acts like a remote procedure, but remotely acts like a command. To do so, we use some tricks for deeply embedded DSLs. To recap deep embeddings:

- We capture expressions by building data-structures that represent the operations inside the expression.
- We capture functions by extending our embedding with variables [12], and applying the function to a fresh variable.

As an example, towards building our remote monad model, consider appending the string "F" to a string:

\[
f :: \text{StringExpr} \to \text{StringExpr}
f t = t <> "F"
\]

We can capture \( f \) by using a small data structure for StringExpr that includes unique variables.

- type \( \text{Id} = \text{Int} \)
- data \( \text{StringExpr} = \text{Var Id} \mid \text{Lit String} \mid \text{Append StringExpr StringExpr} \)

With suitable IsString and Monoid instances for StringExpr, we can reify the function:

\[
\text{GHCi} \gg f \ (\text{Var 0}) = \text{Append} \ (\text{Var 0}) \ (\text{Lit } "F")
\]

Our remote monad will use the function-capture technique, by generating a new \( \text{Id} \) locally for each binding done remotely.

**Definition.** A remote binding is the combination of a request to perform an action for its remote effects, where there is an interesting result value, and the remote interpreter binding this result value to a name remotely. The remote action is called a remote bindee.

The result of a remote binding can be used immediately, without needing local interaction, because the result value resides remotely. Locally, we generate a new and unique \( \text{Id} \), allowing the bindee to be bundled with commands. This is the trick — remote binders are a form of remote procedure calls, but because they return handles with known identifiers, we can transmit them to the remote site as remote commands.

We want our Temperature procedure to now be a Bindee, so that we can directly use the result remotely. We constrain the type of an Bindee in the same way we constrained the return type of a procedure: by using a GADT.

- data Bindee :: * -> * where
  - Temperature :: Bindee StringExpr

We now define the Command data type such that Say takes a StringExpr argument, and with a Bind command that will remotely bind an \( \text{Id} \) to a Bindee:

- data Command where
  - Say :: StringExpr -> Command
  - Bind :: Id -> Bindee StringExpr -> Command

We have our utility functions say and temperature, and a function bindBindee that locally generates a name for the remotely bound value, before utilizing sendCommand.

- say :: StringExpr -> Remote ()
- say txt = sendCommand (Say txt)

For returning the results of remote expressions, we provide a Procedure that takes a StringExpr, remotely executes it, and returns the result:

- type Procedure :: * -> * where
  - Reply :: StringExpr -> Procedure String

This completes our set of transmittable primitives. The remaining definitions are almost identical to the strong remote monad model, except that the Remote monad now contains a numeric state that is used for fresh name generation, and the remote interpreter has an environment that can be updated. Using the above new machinery, and a suitable remote device, we can now run an example:

\[
\text{GHCi} \gg \text{bindBindee} \ :: \text{Bindee StringExpr} \to \text{Remote Id}
\]

\[
\text{bindBindee} \ e = \ do
  \ i \ <- \ \text{newId}
  \ \text{sendCommand} \ (\text{Bind} \ i \ e)
  \ \text{return} \ i
\]

\[
\text{temperature} :: \text{Remote StringExpr}
\]

\[
\text{temperature} = \ do
  \ i \ <- \ \text{bindBindee} \ \text{Temperature}
  \ \text{return} \ (\text{Var} \ i)
\]

For returning the results of remote expressions, we provide a Procedure that takes a StringExpr, remotely executes it, and returns the result:

- \[
\text{Reply} :: \text{StringExpr} \to \text{Procedure String}
\]

\[
\text{reply} :: \text{StringExpr} \to \text{Remote String}
\]

\[
\text{reply} \ e = \ \text{sendProcedure} \ (\text{Reply} \ e)
\]

This completes our set of transmittable primitives. The remaining definitions are almost identical to the strong remote monad model, except that the Remote monad now contains a numeric state that is used for fresh name generation, and the remote interpreter has an environment that can be updated. Using the above new machinery, and a suitable remote device, we can now run an example:

\[
\text{GHCi} \gg \text{send device} \ \& \ \text{do}
  \ t \ <- \ \text{temperature}
  \ \text{say} \ (\"\text{The temperature is } t \ \text{<> } "F\"")
  \ \text{reply} \ t
\]

\[
\text{Remote: The temperature is } 56\text{F} \ "56"
\]

Figure 6 shows the sequence diagram for this example. Observe that the local monadic bind has been transported to the device for remote binding.

Remote bindings provide an extension of traditional serialization. While procedures can return anything serializable, remote bindings can return handles to remote things that cannot be realistically transmitted back to Haskell. For example, some languages support objects that, when serialized, lose their identity. With a remote binding, we have a handle to the original object.

- newtype \( \text{Object} = \text{Object Int} \)

Haskell can pass around the abstract but serializable \( \text{Object} \), and the remote interpreter has an array of objects, indexed by an integer. However, a caveat with remotely bound objects is that, by creating a handle to a remote object, there is no (easy) way to know when you can garbage collect it.
8. The Remote Monad Laws

A remote monad is a monad that has its evaluation function in a remote location, outside the local runtime system. In the remote monad design pattern, this achieved by providing:

- a remote monad — a set of primitive commands and procedures that form a monad; and
- a \texttt{send} function that facilitates remote execution of the primitive commands and procedures.

Towards a more systematic understanding of the remote monad, we propose the \textit{remote monad laws}. Consider a \texttt{send} specialized to a specific destination or device \texttt{d}, with a remote monad \texttt{Remote}, and local monad \texttt{Local}. The \texttt{send} function is a natural transformation between \texttt{Remote} and \texttt{Local}:

\[
\text{send}_d :: \forall a . \text{Remote} a \rightarrow \text{Local} a
\]

As part of a design pattern, the \texttt{send} function itself says nothing about how the individual commands and procedures within the \texttt{Remote} monad are bundled into packets. Instead, as we have seen, \texttt{send} functions can implement different bundling algorithms.

We propose using the monad-transformer lift laws \cite{22,27}, also known as the monad homomorphism laws, as our remote monad laws, because we are lifting the \texttt{Remote} computation to the remote site, by \texttt{Local} effect.

\[
\text{send}_d (\text{return } a) = \text{return } a \quad (1)
\]

\[
\text{send}_d (m \gg\gg k) = \text{send}_d m \gg\gg (\text{send}_d . k) \quad (2)
\]

Assuming these laws, the monad laws, and the laws relating functions and applicative functors to monads, the following morphism laws can be derived:

\[
\text{send}_d (\text{pure } a) = \text{pure } a \quad (3)
\]

\[
\text{send}_d (m_1 \lll m_2) = \text{send}_d m_1 \lll \text{send}_d m_2 \quad (4)
\]

\[
\text{send}_d (\text{fmap } f m) = \text{fmap } f (\text{send}_d m) \quad (5)
\]

Laws \((1)\) and \((3)\) state that a \texttt{send} has no effect for pure computations. Laws \((2)\) and \((4)\) state that packets of remote commands and procedures preserve the ordering of their effects, and can be split and joined into different sized packets without side effects. Law \((5)\) is a reflection of the fact that \texttt{send} captures a natural transformation.

8.1 Infinite Remote Monads

From observation of the laws, it is straightforward to split any finite sequence of total \texttt{Remote} primitives into arbitrarily sized packets. It should be possible to lift the finiteness pre-condition as follows:

- In the weak remote monad, each primitive invokes an individual RPC. Therefore, it is completely reasonable to have an infinite stream of primitives in the \texttt{Remote} monad; in the same way that we can have infinite monadic computations in the \texttt{IO} monad. The \texttt{hArduino} library, discussed in \cite{11,13} is an example of this.
- In the strong remote monad, it would be possible to have an infinite stream of primitives in the \texttt{Remote} monad, \textit{provided the packets are themselves finite}. If necessary, this could be artificially manufactured by putting an upper bound on the number of consecutive commands before sending a packet.

We leave lifting the finiteness requirement of the remote applicative functor to future work, observing that an applicative functor will likely require \texttt{Alternative}, or something similar, to achieve an interesting infinite denotation.

8.2 Networking

Making networks robust is a hard problem. For example, UDP packets are not guaranteed to arrive, or arrive in order, so are unsuitable, without additional infrastructure, for use in a remote monad. The remote monad laws assume the network works. However, having all remote calls wrapped in a \texttt{send} allows for \texttt{send} to enforce a specific policy; for example, time-out after 3 seconds and raise an exception in the \texttt{Local} monad. Also, the \texttt{Remote} monad could have control structures for exception handling built in. The advantage is that by having a structured way of calling remote services, this gives hooks to include systematic recovery services as part of the remote service API.

8.3 Threading

We have assumed a single-threaded local user and a single-threaded remote service, and the remote monad laws reflect this assumption. We could lift the local single-threaded assumption trivially, by using a lock on invocations of \texttt{send}, keeping the remote interactions single threaded. Alternatively, we could allow the local actions to be threaded, and have the thread usage reflect into the remote site. The interleaving semantics will certainly depend on the specific remote monad, and, as when running any side-effecting command on any multi-threaded system, interference patterns need to be considered and accounted for.

9. Extended Example: Blank Canvas

In our toaster models, the remote interpreter was also written in Haskell. This will not be the case in general. Blank Canvas is our Haskell library that provides the complete HTML5 Canvas API, using a strong remote monad. In this section, we describe how we used the remote monad design pattern to build this full scale API that remotely calls JavaScript.

The HTML5 standard, implemented by all modern web browsers, supports a rich canvas element. This canvas element provides a low-level 2-dimensional interface to the in-browser widget. The API provides approximately 30 methods, most of which are void methods called for their effect, and almost 20 settable attributes. Table \ref{ExtendedExample-API} gives a complete list of the base API.

We name our remote monad \texttt{Canvas}, because it represents methods on the JavaScript \texttt{Canvas}. From examining Figure \ref{Canvas} all methods and all settable attributes can be transfigured into monadic Haskell functions in our remote monad. All methods and attributes are serializable values; there are no callbacks in this API.

Classifying the API as remote commands, bindings and procedures is all about deciding if a specific data structure is local or remote:

- \texttt{isPointInPath}, \texttt{toDataURL}, and \texttt{measureText}, are \texttt{procedures}, which return values to the Haskell runtime system, returning \texttt{Bool}, \texttt{Text} and \texttt{TextMetrics} respectively. (\texttt{TextMetrics} is simply a wrapper around a \texttt{Double}.)
- \texttt{createLinearGradient}, \texttt{createRadialGradient}, and \texttt{createPattern}, are \texttt{bindings}, because \texttt{CanvasGradient} and \texttt{CanvasPattern} are subtypes of an image class, which can only be used remotely.
- \texttt{ImageData} is a type that is local to Haskell. This is a design choice. \texttt{ImageData} is an array of RGBA byte values, intended for pixel-level manipulations of images before calling \texttt{putImageData} to render to a canvas. Instead of reflecting an entire deep embedding of arithmetic and array operations, we make \texttt{ImageData} a wrapper around a Haskell byte vector, to be constructed Haskell-side, and provided as a serializable argument to \texttt{putImageData}. Thus, \texttt{getImageData} is a \texttt{procedure}.
- Everything else, including setting attributes, are \texttt{commands}.
The taxonomy of remote primitives allowed the design choice to be an informed choice. We maximize the size of our packets going to the JavaScript interpreter, though we flush the commands at the end of every send. We also add two meta-primitives to our API: sync :: Canvas (), an empty procedure that ensures the pipeline is flushed; and async :: Canvas (), an empty command that asynchronously flushes the pending commands.

We used a deep embedding of the Canvas monad in Blank Canvas. This is more of a historical accident: at the time we wrote Blank Canvas we thought the strong remote monad required such an embedding. However, as we can see from our models, it is possible to build a remote monad with a shallow embedding. The send function interprets the embedding of the Canvas GADT.

To give an example, consider drawing a line using Blank Canvas. All the required primitives are commands. So send bundles them up as a single packet, and transmits the following JavaScript (without the whitespace) to the browser, for evaluation:

```haskell
send context $ do
    rect(50,50,100,100)
    strokeStyle "red"
    stroke()
    moveTo(50,50)
```

The function $. kc . reply, from the package kansas-comet [24], sends a JavaScript value back to the Haskell program. The first argument is the transaction id, and the second argument is the returned (JavaScript) value. We then embed JavaScript’s isPointInPath inside a Blank Canvas supporting function.

```javascript
function IsPointInPath(x, y) {
    return function (u, c) {
        $.kc.reply(u, c.isPointInPath(x, y));
    }
}
```

For bindings, we allocate a global variable, and remember the unique number inside the proxy object.

```javascript
send context $ do
    try{
        var v_42 = ![BINDEE] ...
        )catch(e){
            alert('...');
        }
    }
```

In summary, the JavaScript generated is a direct transliteration of the function calls made in Haskell, using the remote monad as the guiding design pattern.

## 10. Blank Canvas in Practice

Out of the box, Blank Canvas is pac-man complete — it is a platform for simple graphics, classic video games, and building more powerful abstractions that use graphics. The key question is how well does Blank Canvas perform in practice. In this section, we give both empirical and anecdotal evidence.

### 10.1 Empirical Evidence

An important question is the cost of using the remote monad design pattern, albeit over a machine-local network. At first glance the cost of using the remote monad to implement Blank Canvas seem prohibitive. Blank Canvas transliterates each packet to a Text string, and then sends the packet over a network, where it is parsed by the JavaScript virtual machine. Normally, in a JavaScript application, the JavaScript program is parsed and compiled once, and the inner loops are run many times. But a loop in Blank Canvas will repeatedly transmit the contents of the loop body, requiring re-parsing and re-compilation each time. Even with the use of the remote monad design pattern, we expect to take a significant performance hit, above and beyond the cost of networking.
In order to quantify our ideas, we have measured the performance of Blank Canvas on several benchmark programs and compared them to native JavaScript. We have two classes of benchmarks: command benchmarks, which simply render to the canvas; and procedure benchmarks, where the inner loop of the benchmark invokes some form of query that requires a round-trip from server, to client, and back to the server. The JavaScript versions of the benchmarks were written in idiomatic JavaScript. The Blank Canvas tests were run using criterion [35], and the JavaScript tests used the criterion mechanism to estimate confidence.

We ran our benchmarks on both Chrome and Firefox, and on both Linux and OSX. The relative performance of our command benchmarks varied widely, depending on browser and benchmark, but on average, the cost of using Haskell and the Blank Canvas API was between a factor of 2 and 10 relative to native JavaScript. This is surprising and encouraging! However, the performance cost of our procedure benchmarks was significantly higher, up to a factor of 600 relative to native JavaScript. We expect that adding a strong remote applicative functor capability would significantly reduce the cost of the procedure benchmarks, because it would also allow Blank Canvas to bundle procedures. We plan to perform a more systematic study of the costs of the remote monad, and the benefits of the remote applicative functor, in the future.

10.2 Anecdotal Evidence
Blank Canvas has now been used by the students in four separate instances of our functional programming class. Students find it easy to understand, given the analog between the IO monad and the remote Canvas monad, with students often choosing to use Blank Canvas for their end-of-semester project. To give two examples, one end-of-semester project was Omar Bari and Dain Vermaak’s Isometric Tile Game, that can be rotated in 3D in real-time; another project was Blankeroids, a playable asteroids clone, written by Mark Grebe, on top of Yampa [9] and yampa-canvas [39]. Both are shown here with the students’ permission.

In summary, Blank Canvas, using the remote monad, is a viable way for students to draw pictures and write games in Haskell. Rendering graphics uses many more commands than procedures, so it (retrospectively) turns out to be an ideal candidate for the remote monad. Given that we have access to whole new capabilities, we consider the overheads of using the remote monad reasonable.

11. Related Work
Once the remote monad design pattern is understood, many instances of its use, or of related patterns, can be observed in the wild. In this section we discuss some of the existing instances, and present the type of the send analog, with the natural transformation highlighted in red. This list is not intended to be comprehensive, but rather to give a flavor of existing uses, or close uses, of the remote monad and related ideas.

11.1 Terminal Interaction
- The ncurses [33] package provides two levels of weak remote monads, in order to provide an API for moving the cursor around a terminal, and other terminal-specific services. The outer send has the type
  \[
  \text{runCurses :: Curses a \rightarrow IO a} 
  \]
  and the inner send has the type
  \[
  \text{updateWindow :: Window \rightarrow Update a \rightarrow Curses a} 
  \]
  This package compiles commands into ANSI-style command sequences to be execute directly on the terminal.

11.2 Database Access
- In the Haxl DSL, Marlow [29] uses the properties of applicative functors to issue database queries in parallel. The send function has the type:
  \[
  \text{runHaxl :: Env u \rightarrow GenHaxl u a \rightarrow IO a} 
  \]
  Haxl is a DSL with a weak remote monad, with a strong remote applicative functor.
- The mongoDB [23] is an API into the popular MongoDB NoSQL database that uses a weak remote monad. The send function has the type:
  \[
  \text{access :: MonadIO m \Rightarrow Pipe} \\
  \text{\rightarrow AccessMode} \\
  \text{\rightarrow Database} \\
  \text{\rightarrow Action m a \rightarrow m a} 
  \]

11.3 Browser / Server
- The rather ingenious Haste.App library [10,11] is an instance of the remote monad design pattern. The Haste.App uses two Haskell compilers, a Haskell to JavaScript compiler for the local monad, and GHC for the remote monad evaluator, with the glue code between the two generated artifacts being automatically generated. The send function has the type:
  \[
  \text{onServer :: Binary a \rightarrow Remote \((Server a) \rightarrow Client a} 
  \]
  The Remote is a wrapper to help stage the difference between the client and server compilations. The Server monad is the remote monad. The remoteness here is running on the server; the main program runs on the client browser.
- Sunroof is a compiler for a monadic JavaScript DSL [4,29], developed by the University of Kansas. The compiler, which is considerably more involved than Blank Canvas, could be used stand-alone, as a part of a strong remote monad. There are three separate send commands:
  \[
  \text{asyncJS :: SunroofEngine \rightarrow JS t a \rightarrow IO a} \\
  \text{syncJS :: SunroofEngine \rightarrow JS t a \rightarrow IO (ResultOf a) } \\
  \text{rsyncJS :: SunroofEngine \rightarrow JS t a \rightarrow IO a} 
  \]
  Using what we have learned from studying the remote monad design pattern, these three send functions could be combined into a single send.

11.4 Embedded Systems
- The Levent Erkő’s hårduino package [14] uses the weak remote monad design pattern. The send-command withArduino is a one-shot operation, and does not have the ability to return values. Instead, the monad represents the whole computation to be executed.
  \[
  \text{withArduino :: Bool \rightarrow FilePath \rightarrow Arduino () \rightarrow IO ()} 
  \]
• Ben Gamari’s bus-pirate package [18] allows access to the I²C or SPI embedded device protocols via a serial port, using monadic primitives. The `send` command implements the remote monad directly.

    runBusPirate :: FilePath -> BusPirateM a -> IO (Either String a)

• The University of Kansas used an (as yet unpublished) strong remote applicative functor in the design of an α-bridge, an interface between Haskell and an FPGA, built on top of the open Wishbone [35] bus protocol.

    send :: Board -> BusCmd a -> IO (Maybe a)

The applicative functor structure allows multiple reads and writes to be combined into a bus transaction, and executed on an FPGA.

11.5 GUIs

• Heinrich Apfelmus’ threepenny-gui [1] uses a user-interface element monad UI as a weak remote monad.

    runUI :: Window -> UI a -> IO a

11.6 GPGPUs

• The popular accelerate package [7] uses a remote array design pattern to program a GPGPU. The `send` function takes an Acc, and returns a pure result.

    runIn :: Arrays a => Context -> Acc a -> a

This version of `send` does not use a local monad, but it can be considered to use the Identity monad. The `send` function can be purely functional specifically because the computations that are performed remotely are purely functional. Like other uses of the remote design pattern, Acc encodes the computation to be performed, in this case on the GPGPU. Acc is a DSL which provides array-based operators; such as zipWith and map. There is no applicative apply, or monadic bind, for Acc.

11.7 SMT solvers

• Levent Erk¨ok’s sbv package [15] provides access to SMT solvers, and uses a weak remote monad as its lowest level API.

    runSymbolic' :: SBVRunMode -> Symbolic a -> IO (a, Result)

11.8 Games

• Douglas Burke’s mcpi package [5] uses a weak remote monad to allow Haskell users to interact with a Minecraft server.

    runMCPI :: MCPI a -> IO a

11.9 Other Related Works

This work is the marriage of Domain Specific Languages and Remote Procedure Calls. Both have a long and rich history.

Sun Microsystems implemented the first widely-distributed RPC implementation [31,32]. The ideas were based on Birrell and Nelson’s seminal work in this area [3]. All modern operating systems include RPC capabilities, building on these initial implementations. There is now an array of high-level libraries and frameworks for almost any modern language, offering RPC services such as D-Bus [28], which operates within a single machine, and Java’s Jini, now called Apache River (http://river.apache.org/). The Haskell cloud-computing effort [13] includes support for RPCs, and can be considered such a middleware solution. There are also many low-level protocols for executing RPCs, such as JSON-RPC (http://www.jsonrpc.org/).

Domain Specific Languages are Haskell’s forte. There are shallower embedded DSLs [25,26], with direct execution, and deeply embedded DSLs [12], with a generated structure representing the computation and a paired evaluator. There are also tagless-final DSLs [6], where there is no early commitment to a specific embedding, but instead class overloading in used to specify the API. Several Haskell DSLs, including Feldspar [37] and Obsidian [8], reify the structure of monadic code [40,41] to generate external imperative code. Other DSLs, including CoPilot [38] and Ivory [24], use a shallow embedding of statements to capture the monadic structure. All of these DSLs could leverage the remote monad design pattern.

12. Conclusion and Future Work

This work was inspired by the question of what would an online evaluator of a deeply embedded DSL outside the Haskell heap look like. However, as we have seen, other DSL technologies can be used to generate the remote packets; the remote monad ideas are orthogonal to specific DSL implementation technologies. We thought we were writing a paper about deeply embedded DSLs and interesting GADT encodings. Instead, we discovered a design pattern and a small language for RPC APIs.

We have built a number of remote monads and remote applicative functions. Aside from the examples already documented above, we have also reimplemented the Minecraft API found in mcpi, but with a strong remote monad, and a general JSON-RPC framework in Haskell. In particular, the JSON-RPC protocol supports multiple batched calls, as well as individual calls. Currently, the user needs to choose between the monadic and applicative functor API. We will use this prototype as a test-bench to explore the combination of simultaneously being a strong remote monad and a strong remote applicative functor. The recent push to add applicative functor do-notation as a GHC-extension is something that we can take advantage of here.

Our simulated remote interpreters used regular data structures, rather than sharing the GADT used by `send`. This was for clarity, but also to avoid the complications of reifying GADTs. This is not a fundamental limitation: we have also implemented all our models using GADTs for the remote procedures, using a wrapper GADT called `Transport` to enable deserialization.

    data Transport (m :: * -> *) = forall a . (Show a) => Transport (m a)

By providing a `Read` instance for `Transport`, we can hide the phantom type index from the `Read` instance, while ensuring that we have a `Show` instance for the phantom type. We expect that any full-scale implementation would use similar techniques when the remote interpreter uses Haskell.

A remote monad is a sweet spot between an RPC and a deeply embedded DSL, offering the possibility of rich FFIs for little effort and runtime cost. The remote monad design pattern allows a progression from weak, to strong, to deep embedding, giving a gentler pathway to implement deeply embedded DSLs. We have used the remote monad design pattern many times, and hope others will find the pattern as useful as we have.

Acknowledgments

We would like to thank Heinrich Apfelmus, Conal Elliott, and Levent Erk¨ok for their useful discussions and insights, and the anonymous reviewers for their useful and constructive comments. This material is based upon work supported by the National Science Foundation under Grant No. 1117569 and Grant No. 1350901. Aleksander Eskilson, Ryan Scott and James Stanton were supported by the NSF REU initiative. The icons in the sequence diagrams were created by Tomoyuki Miyano, and made available online, at http://iconarchive.com/.
References