Fiber-local storage

Fiber-local storage (fls) is a mechanism for associating a fiber with some implicit, local state. There are three core operations. The first creates an empty local storage; the second sets the fls for the host vproc; the third gets the fls on the host vproc.

type fls
val new : unit -> fls
val set : fls -> unit
val get : unit -> fls

To make fls useful, we provide some basic dictionary operations. Keys into the dictionary are represented by the \texttt{tag} type. The argument to the tag type is a phantom type, \textit{i.e.},

type 'a tag = int
val key1 : key1 ty tag = 1
val key2 : key2 ty tag = 2
...

We have operations for adding and finding elements.
val add : (fls * ('a tag * 'a)) -> fls
val find : (fls * 'a tag) -> 'a option

Fiber-group storage

In certain cases, we wish to associate some implicit, shared state with a "group" of fibers. For read-only access, our fls mechanism is already sufficient. But for read-write access, we need synchronized memory.

Example: scheduler initialization for futures

Our architecture allows each thread to use multiple schedulers. To avoid unnecessary overhead, we delay initialization until the last possible moment, when the first spawn operation occurs.

Unfortunately, because of our architecture, we need to synchronize the process of initializing schedulers. Suppose that we have two versions of futures, one subject to a breadth-first scheduling policy and the other to a depth first policy. It is easy to imagine a program that uses both schedulers, such as the one below. Unfortunately, we cannot know which call to \texttt{futureDFS} happens first, so we need some synchronized memory.

\[ \text{futureBFS}(\text{fn}() \Rightarrow \text{futureDFS}(\text{fn}() \Rightarrow \text{e1})) + \text{futureDFS}(\text{fn}() \Rightarrow \text{e2}) \]

The function below handles the initialization of our DFS scheduler. As a convenience, it returns the ready queue of the scheduler, hence the name. This function starts by finding the shared scheduler state in the fls. It then uses \texttt{Set-once memory} to start the scheduler and get the handle on the ready queue.

\[
\begin{align*}
\text{fun getDFSReadyQ () = let} \\
\quad \text{val fls = FLS.get()} \\
\quad \text{val SOME schedState = FLS.find(fls, tag(futureDFS))} \\
\quad \text{in} \\
\quad \text{SetOnceMem.set(schedState, fn () \Rightarrow let} \\
\quad \quad \text{val readyQ = LockedQueue.new()} \\
\quad \quad \text{in} \\
\quad \quad \text{startDFSSched readyQ; readyQ} \\
\quad \text{end)} \\
\text{end)} \\
\text{end)
\end{align*}
\]

Now it is easy to spawn a future on the DFS scheduler. We allocate the future, put it on the ready queue, and return the handle.

\[
\begin{align*}
\text{fun futureDFS thunk = let} \\
\quad \text{val fut = alloc(EMPTY_FUTURE, thunk)} \\
\quad \text{in} \\
\quad \text{LockedQueue.enqueue(getDFSReadyQ(), fut); fut} \\
\text{end)
\end{align*}
\]

Thread capabilities

Because threads consist of one or more fibers, they necessarily have read-only access to fls. Thus, any local storage entry must be added during the \texttt{spawn} operation. To address this issue, we use the notion of "thread capabilities". For example, a given thread could have a capability for depth-first or breadth-first futures.

\[
\begin{align*}
\text{spawn(f1, FutureDFS.capability);} \\
\text{spawn(f2, FutureBFS.capability)}
\end{align*}
\]

Capabilities are just entries in the fiber-local storage dictionary. Before spawning a thread, we add all the capabilities to its local storage.
type 'a capability = 'a tag * 'a
fun spawn (thunk, capabilities) = let
  val fls = List.foldl FLS.add (FLS.new()) capabilities
  in
  VProcQueue.enqueue(flsl, Control.fiber thunk);
  fls
end

Our implementation defines a set of default capabilities for threads. See <code>src/lib/basis/runtime/utils/default-thread-capabilities.pml</code> for the complete list.