HOrc
An Implementation of Orc in Haskell

Marco Devesas Campos
Nordic DataGrid Facility, Denmark

Luís Soares Barbosa
Universidade do Minho, Portugal
"[...] if you introduce recursion, you lose some people; if you introduce pointers, you lose some people; if you introduce event-driven models, you lose some people; **when you introduce concurrency, you lose virtually everybody.**"
Russ Daniels, CTO and VP of Cloud Computing HP.

"ATI believes that to write efficient parallel code requires a level of skill that's an order of magnitude higher than what your typical programmer can do. [..] If you're a brilliant programmer, you know exactly what you should view as your next frontier"

"Massively multithreaded programming isn't just a "hard problem" — rather, **it's a generation's worth of Ph.D. dissertations** that have yet to be written."
Jon Stokes @ http://arstechnica.com/news.ars/post/20081207-analysis-more-than-16-cores-may-well-be-pointless.html

The State of the Union
4 Cores + Hyperthreading = 8 Simultaneous Threads

In our CPU’s/Internetz
I wanna watch a film!

Cloud computing
An Old Problem

Timeline

1962 - Petri Nets
1965 - Semaphores
1973 - Actor Model
1974 - Monitors
1978 - CSP
1980 - CCS
CoordInspect
To Play is To Know
Today’s menu

- HOrc Intro
- Philosophy
- Components
- Examples
- Some benchmarks
HOrc’s Features

- Compositionality
  - Sites are the computational element
  - 3 operators that combine the results of their sub-orchestrations
**HOrc as a DSL**

- Every HOrc program is a valid Haskell Program
- Values represented by Haskell Values
- Variables Represented by Haskell Variables -- *Important (see paper)*
- Orchestrations are typed
- Different syntax from Orc
HOrc as DSL

• Every HOrc program represents a value of type:

\[
\text{Action } a = \text{Channel } (\text{Maybe } a) \to \text{IO } ()
\]

• That the Haskell runtime runs when you call “run” on it.
Sites

IO Functions
- Pure Functions
- User Interaction
- Database Queries
- Remote Procedure Calls
Creating Sites

- **liftIO**: `IO a -> Action a`
  - Always publishes

- **liftMaybe**: `IO (Maybe a)`
  - May fail (return `Nothing`) and not publish
Two Simple Examples

- ozero never publishes
  ozero :: Action a
  ozero = liftMaybe (return Nothing)

- oprint displays its argument on screen
  oprint :: (Show a) => a -> Action ()
  oprint x = liftIO (print x)
Parallel Composition

\[ \text{f \ `mplus` g \:: Action a} \]
Example: Garbled Output
Sequential Composition

\[
f \circ \ \lambda x \rightarrow g \ x
\]

\[
g \ x_1 \circ \ldots \circ g \ x_n
\]
Example: Garbled Output(2)
Pruned Seq. Composition
Hadoop, a Free Software Program, Finds Uses Beyond Search

Christophe Bisciglia, Amr Awadallah, Jeff Hammerbacher and Mike Olson started their company Cloudera, around Hadoop.

By ASHLEE VANCE
Published: March 16, 2009

BURLINGAME, Calif. — In the span of just a couple of years, Hadoop, a free software program named after a toy elephant, has taken over some of the world's biggest Web sites. It controls
Google Trends

Searches

Websites

cloud computing

Search Volume index

News reference volume

Monday, 6 July 2009