Parallel all the way

Clojure Parallelism and Beyond

LDN Functionals - Renzo Borgatti @reborg

About @reborg

- ~ 5 years of Clojure full time
- Organiser of the Papers We Love Meetup
- Clojure Pills on YouTube
- Currently at Droit Fintech

Agenda

1. Rationale
2. A Task Oriented Approach
3. Examples and Challenges
4. Conclusions

Context

- Circa 2007-2008
- CPU top speed achieved
- Cores to increase further
- Push to leverage multiple cores
- Functional Programming return

Clojure Parallelism

- Basic set of parallel APIs:
- `pmap` `pvalues` `pcalls` (lazy, sequential, chunked)
- `reducers/fold` (work-stealing, fork-join)
- Custom with `future`, `agent`, etc.
- `core.async` pipelines (external lib)

Problems

- Powerful, but somewhat low level
- Not necessarily easy to use
- Even more to use correctly
- Inconsistencies with stateful xforms
- Chunk size dependency (`pmap`)

Can we approach it differently?

- Task oriented API
- Predictable semantic
- Easy to use
- Documented

The Parallel Library

- Experimenting ideas in a library
- https://github.com/reborg/parallel
- Documented, tested and benchmarked
- No dependencies (other than Clojure itself)

At a glance

- Modeled on existing functions from the stdlib
- Drop-in replacement (when possible)
- A few brand new functions
- Some specific transducers support

Current Line-up 1/4

Name	Description
`p/let`	Parallel `let` binding.
`p/slurp`	Parallel slurping files.
`p/count`	Transducer-aware parallel `core/count`.
`p/frequencies`	Parallel `core/frequencies`
`p/group-by`	Parallel `core/group-by`

Current Line-up 2/4

Name	Description
`p/update-vals`	Updates values in a map in parallel.
`p/external-sort`	Memory efficient, file-based, parallel merge-sort.
`p/sort`	Parallel `core/sort`.
`p/fold`	Transducer-aware `r/fold`.

Current Line-up 3/4

Name	Description
`p/min and p/max`	Parallel `core/min` and `core/max` functions.
`p/distinct`	Parallel version of `core/distinct`
`p/amap`	Parallel array transformation.
`p/armap`	Parallel array reversal with transformation.

Current Line-up 4/4

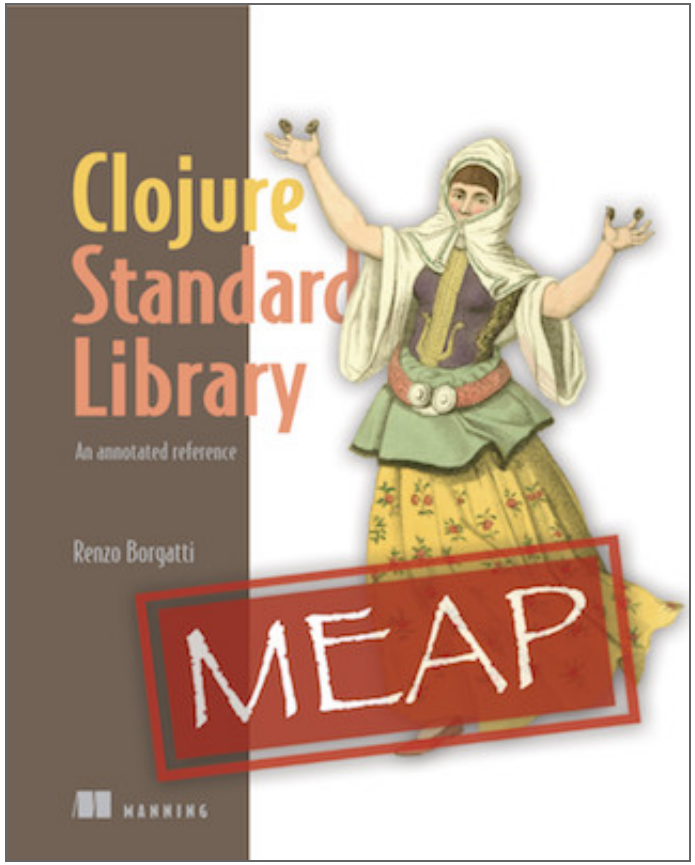
Name	Description
`xf/interleave`	Like `core/interleave`, transducer version.
`xf/pmap`	Like `core/pmap`, transducer version.
`xf/identity`	Alternative identity transducer to `core/identity`

DEMO TIME!

Gotchas

- There is hope!
- Not suitable for trivial computation
- Same for small collections
- Go mutable as an option
- Always use a profiler!

Quick Plug


Clojure
Standard
Library
An annotated reference
Renzo Borgatti
MEAP
MANNING

Resources

- The **parallel** library enables consistent stateful xforms in parallel.
- **A Java fork-join framework** paper by Doug Lea
- **Clojure Applied** book contains chapters dedicated to Transducers with core.async pipelines examples.
- **Standard Library book**, Chapter 7 Reducers and Transducers