The Architecture of PicCollage Server

http://godfat.org/slide/ 2013-01-08-PicCollage.pdf

Table of Contents

PicCollage, Software and Services

Table of Contents

- PicCollage, Software and Services
- Ruby Codes

Table of Contents

- PicCollage, Software and Services
- Ruby Codes
- Miscellaneous Extras

PicCollage, Software Services

Ruby Codes

- Gemfile
- Procfile
- Rakefile

- rainbows.rb
- concurrency.rb
- config.ru

Ruby Codes (Cont.)

- server.rb
- error rb
- eagerload.rbwaiter.rb

- newrelic.rb
- shell.rb

Ruby Codes (Cont.) (Cont.)

- image_util.rb
- run later.rb
- /search
- /create_and_share

Miscellaneous Extras

- crazy inspect
- who cares backtrace for 404
- assets pipeline and database connection
- fiber stack overflow
- fiber database connection pool
- thread-safe <u>kaminari patch</u>

PicCollage, Software Services

PicCollage

- Make collages of your photos
- Import photos from places
- Touch gestures to rotate, resize, delete
- Double-tap to edit, clip, etc
- Clip photos by outlining the area
- Lots of backgrounds and stickers

PicCollage, Software Services

• Ruby 1.8.7 -> 1.9.2 -> 1.9.3

- Ruby 1.8.7 -> 1.9.2 -> 1.9.3
- Rails 3.1.0.rc4 -> 3.1.0.rc5 -> 3.1.0 -> 3.1.3 -> 3.2.1 -> 3.2.2 -> 3.2.3 -> 3.2.5 -> 3.2.6 -> 3.2.7 -> 3.2.8 -> 3.2.9 -> 3.2.10

- Ruby 1.8.7 -> 1.9.2 -> 1.9.3
- Rails 3.1.0.rc4 -> 3.1.0.rc5 -> 3.1.0 -> 3.1.3 -> 3.2.1 -> 3.2.2 -> 3.2.3 -> 3.2.5 -> 3.2.6 -> 3.2.7 -> 3.2.8 -> 3.2.9 -> 3.2.10
- Sinatra -> Jellyfish

- Ruby 1.8.7 -> 1.9.2 -> 1.9.3
- Rails 3.1.0.rc4 -> 3.1.0.rc5 -> 3.1.0 -> 3.1.3 -> 3.2.1 -> 3.2.2 -> 3.2.3 -> 3.2.5 -> 3.2.6 -> 3.2.7 -> 3.2.8 -> 3.2.9 -> 3.2.10
- Sinatra -> Jellyfish
- Thin -> Zbatery

- Ruby 1.8.7 -> 1.9.2 -> 1.9.3
- Rails 3.1.0.rc4 -> 3.1.0.rc5 -> 3.1.0 -> 3.1.3 -> 3.2.1 -> 3.2.2 -> 3.2.3 -> 3.2.5 -> 3.2.6 -> 3.2.7 -> 3.2.8 -> 3.2.9 -> 3.2.10
- Sinatra -> Jellyfish
- Thin -> Zbatery
- Resque -> EM.next_tick or Thread.new -> Sidekiq (?)

PicCollage, Software Services

Github

- Github
- Heroku Bamboo -> Cedar

- Github
- Heroku Bamboo -> Cedar
- Heroku PostgreSQL 9.0 -> 9.2

- Github
- Heroku Bamboo -> Cedar
- Heroku PostgreSQL 9.0 -> 9.2
- MongoHQ -> MongoLab

- Github
- Heroku Bamboo -> Cedar
- Heroku PostgreSQL 9.0 -> 9.2
- MongoHQ -> MongoLab
- RedisToGo -> OpenRedis

- Github
- Heroku Bamboo -> Cedar
- Heroku PostgreSQL 9.0 -> 9.2
- MongoHQ -> MongoLab
- RedisToGo -> OpenRedis
- Memcache (?) -> Memcachier

- Github
- Heroku Bamboo -> Cedar
- Heroku PostgreSQL 9.0 -> 9.2
- MongoHQ -> MongoLab
- RedisToGo -> OpenRedis
- Memcache (?) -> Memcachier
- NewRelic

- Github
- Heroku Bamboo -> Cedar
- Heroku PostgreSQL 9.0 -> 9.2
- MongoHQ -> MongoLab
- RedisToGo -> OpenRedis
- Memcache (?) -> Memcachier
- NewRelic
- Exceptional

- Github
- Heroku Bamboo -> Cedar
- Heroku PostgreSQL 9.0 -> 9.2
- MongoHQ -> MongoLab
- RedisToGo -> OpenRedis
- Memcache (?) -> Memcachier
- NewRelic
- Exceptional
- SendGrid -> Mailgun

Ruby Codes

- Gemfile
- Procfile
- Rakefile

- rainbows.rb
- concurrency.rb
- config.ru

Gemfile

```
ruby '1.9.3'
source 'http://rubygems.org'
gem 'rails', '3.2.9'
gem 'jellyfish'
gem 'pg', :platform => :ruby
gem 'redis-objects'
gem 'dalli'
```

```
# service
gem 'newrelic rpm', :require => false
gem 'exceptional'
gem 'apns', :github => 'jpoz/APNS'
    # ^^^^ needs feedback patch
# server
platform : ruby do
  gem 'zbatery', :require => false
  gem 'em-http-request'
  gem 'yajl-ruby'
```

```
# upload
gem 'paperclip' # needs various patches
gem 'aws-sdk'

# plist
gem 'nokogiri'
```

```
# pagination
gem 'kaminari' # needs thread-safe patch
```

```
# javascript
gem 'jquery-rails'
```

make `rails server` work for zbatery
gem 'rack-handlers'

```
# respect rack logger
gem 'rack-rails-logger'
```

redirect from www to root domain
gem 'rack-rewrite'

```
# cache for API
gem 'rack-cache'
```

```
gem 'rest-core'
gem 'rest-more'

# bluebase
gem 'bb_more' , :path => 'bluebase'
gem 'bb_locales', :path => 'bluebase'
gem 'bb_auth' , :path => 'bluebase'
gem 'bb redis' , :path => 'bluebase'
```

api client

```
:assets do
group
      'sass-rails'
  gem
      'coffee-rails'
  gem
  gem 'uglifier'
end
group : cedar do
      '<u>ri</u>b'
  gem
      'bond'
  gem
      '<u>readline_buffer</u>',
  gem
       :platform => :ruby
end
```

Gemfile (Cont.) :test do

```
group
      'minitest'
  gem
      'rr'
  gem
       'webmock'
  gem
      'kramdown'
  gem
end
       :test do
group
       'capybara'
  gem
       'rspec'
  gem
       'rspec-rails'
  gem
end
```

Ruby Codes

- Gemfile
- Procfile
- Rakefile

- rainbows.rb
- concurrency.rb
- config.ru

Procfile

```
web: ruby -r bundler/setup -S zbatery
-c config/rainbows.rb
-p $PORT -E $RACK_ENV
```

Ruby Codes

- Gemfile
- Procfile
- Rakefile

- rainbows.rb
- concurrency.rb
- config.ru

Rakefile

```
task 'test:bluebase' => ['env:test'] do
    # remove bundler shit
    gemfile = ENV.delete('BUNDLE_GEMFILE')
    sh './bin/bluebase-test'
    ENV['BUNDLE_GEMFILE'] = gemfile
end
```

```
task :test => ['test:api']
```

```
task 'test:api' => ['env:test',
                     'environment'] do
  require './config/environments/test'
  # force rails loaded, otherwise it's
  # causing too much loading troubles,
  # by rails' autoloading mechanism
  Dir['./test/api/*.rb'].each do |test|
    require test
  end
 # ...
```

```
# this is only available in 1.9.3+
if Test::Unit.const_defined?(:RunCount)
   Test::Unit::RunCount.run_once{
        status = Test::Unit::Runner.new.run
        exit(status) if status != 0
    }
end
```

Ruby Codes

- Gemfile
- Procfile
- Rakefile

- rainbows.rb
- concurrency.rb
- config.ru

rainbows.rb

```
worker processes 2 # assume 2 CPU cores
preload app
                 true
l = ::Logger.new($stderr)
l.level = case ENV['RACK ENV']
          when 'production'
            ::Logger::WARN
          else
            ::Logger::DEBUG
          end
logger l
```

```
Rainbows! do
  use :EventMachine, :em client class =>
    lambda{
      require 'concurrency'
      client = Concurrency.
        eventmachine client
      l.info("Using #{client}")
      client
  worker connections 64
end
```

```
before_fork do |_, _|
  # don't hold connections on master
  unless defined?(Zbatery)
    l.info("Discon' PSQL and Redis...")
    AR::Base.connection.disconnect!
    Redis.current.quit
```

end end

```
after fork do |_, _|
          hold connections on workers
  #
  unless defined?(Zbatery)
    l.info("Conn' PSQL and Redis...")
    Redis.current.connect
    AR::Base.connection.
      establish connection
  end
```

```
EM.error handler do |e|
  puts "Err: #{e.inspect}" \
       " #{e.backtrace.inspect}"
  begin
    :: Exceptional:: Catcher.handle(e)
    :: NewRelic:: Agent.instance.
      error collector.notice error(e)
  rescue Exception => e
    puts "Err: Exceptional/NewRelic:" \
         " #{e.inspect}" \
           #{e.backtrace.inspect}"
```

Rainbows! EventMachine Thread Client

https://gist.github.com/4451322 https://github.com/godfat/rainbows/pull/2

```
class ::RainbowsEMThreadPoolClient <</pre>
        Rainbows::EventMachine::Client
  def app call input
    @deferred = true
    EM. defer do
      # Call the application here!
    end
  end
```

```
# Call the application here!
begin
  @env.merge!(RACK DEFAULTS)
  status, headers, body = APP.call(@env)
  @deferred = nil # EM.next tick?
  ev write response(status,
    headers, body, @hp.next?)
rescue Exception => e
  EM.instance variable get(
    :@error handler).call(e)
  handle error(e)
```

Ruby Codes

- Gemfile
- Procfile
- Rakefile

- rainbows.rb
- concurrency.rb
- config.ru

concurrency.rb

```
module Concurrency
  Model = Thread
  module function
  def eventmachine client
          Model == Thread
      RainbowsEMThreadPoolClient
    elsif Model == Fiber
      RainbowsEMFiberSpawnClient
    else
      Rainbows::EventMachine::Client
    end
```

concurrency.rb (Con

```
module Concurrency
  def wrap
    if Model == Thread
      EM. defer do
        # ...
      end
    elsif Model == Fiber
      Fiber.new{ yield }.resume
    else
     yield
    end
```

concurrency.rb (Con

```
def wrap
  begin
    yield
  rescue Exception => e
    EM.instance variable get(
      :@error handler).call(e)
  ensure
    AR::Base.clear active connections!
  end
```

Ruby Codes

- Gemfile
- Procfile
- Rakefile

- rainbows.rb
- concurrency.rb
- config.ru

config.ru

```
use Api::HostDispatcher, Api::Server
map '/api' do
  run Api::Server
end
map '/' do
  run PicCollage::Application
end
```

Ruby Codes (Cont.)

- server.rb
- error rb
- eagerload.rbwaiter.rb

- newrelic.rb
- shell.rb

server.rb

```
Api::Server = Rack::Builder.new do
    # Embrace Rack is the <u>Jellyfish</u> way
end
```

```
use Rack::CommonLogger, debug logger
use Rack::Deflater
use Rack::Static,
  :urls => ['/images'], :root =>
  "#{File.dirname( FILE )}/public"
use AR::ConnectionAdapters::
    ConnectionManagement
```

```
use Api::MiddleError
use Api::MiddleDevice
use Api::MiddleAuth
use Api::MiddleCache
use Rack::Cache, {} # [snipped]
run Api::ServerCore.new
```

```
class Api::ServerCore
  include Jellyfish
  handle exceptions false
  def controller; Controller; end
  class Controller < Api::Controller</pre>
    include Api::ServerMethods
    include Jellyfish::NewRelic
  end
end
```

```
class Api::ServerCore
  post '/collages/create_and_share' do
    render_collage create_collage{ |c|
        transfer "/collages/#{c.id}/share"
    }
  end
end
```

Ruby Codes (Cont.)

- server.rb
- error.rb
- eagerload.rbwaiter.rb

- newrelic.rb
- shell.rb

error.rb

```
class Api::MiddleError
  include Jellyfish
  handle exceptions true
  def controller; Api::Controller; end
  handle Api::Error do |error|
    render(error custom(error))
  end
end
```

error.rb (Cont.)

```
class Api::MiddleError
  handle Jellyfish::NotFound do
    render(error custom(
      Api::NotFound.new(
        "Path not found.")))
  end
  handle Exception do |error|
    render(error 500(error))
  end
```

Ruby Codes (Cont.)

- server.rb
- error rb
- eagerload.rbwaiter.rb

- newrelic.rb
- shell.rb

eagerload.rb

```
module Eagerload
  module function
  def eagerload
    if ENV['RAILS GROUPS'] == 'assets'
      $stderr.puts "Don't eagerload " \
         "for assets precompilation."
      return
    end
    # ...
```

eagerload.rb (Cont

```
def eagerload
    # ...
    require 'rest-more'
    RC.eagerload
    require 'aws-sdk'
    RC.eagerload(AWS::Core)
    RC.eagerload(AWS::S3)
```

eagerload.rb (Cont

```
def eagerload
# ...
%w[app/models/**/*
    app/controllers/**/*
    app/mailers/**/*].
    each{ |pattern|
        require_all(Dir[pattern], 2) }
```

eagerload.rb (Cont

```
def eagerload
    # ...
    require 'api/requirements'
    require_all(Dir['lib/**/*'], 1)
    $stderr.puts "Eagerly loaded."
    true
end
```

Ruby Codes (Cont.)

- server.rb
- error rb
- eagerload.rbwaiter.rb

- newrelic.rb
- shell.rb

newrelic.rb

```
ActiveSupport.on load(:active record) do
  require 'newrelic rpm'
  NewRelic::Agent.after fork(
    :force reconnect => true) if
    defined?(Unicorn)
```

```
require 'rest-core/engine/future/future'
RC::Future.module_eval do
  include NewRelic::Agent::MethodTracer
  alias_method :yield_original, :yield
  def yield
    method = env[RC::REQUEST_METHOD]
    path = env[RC::REQUEST_PATH]
    # ...
```

```
def yield
  metrics = ["External/#{path}/" \
     "#{self.class.name}/#{method}",
             "External/#{path}/all",
             "External/all",
             "External/allWeb"]
  self.class.
    trace execution scoped metrics do
      yield original
    end
```

Ruby Codes (Cont.)

- server.rb
- error rb
- eagerload.rbwaiter.rb

- newrelic.rb
- shell.rb

shell.rb

```
require 'waiter'
require 'concurrency'
module Shell
  module function
  # this is for cocaine
  def call command, env={}
    Shell.system(*command)
  end
  # e.g. Shell.system('identify', 'img')
  def system *args
    system multi(args).first
  end
```

```
# e.g. Shell.system multi(%w[identify
       image.png], %w[identify me.png])
def system multi *cmds
  if EM. reactor running?
    # ...
  else
    cmds.map{ | cmd |
      system blocking(cmd) }
  end
end
```

```
def system multi *cmds
  # . . .
  w = Waiter.new
  r = \{\}
  cmds.each.with index do | cmd, idx |
    system async(*cmd) do |output|
      r[idx] = output
      w.resume(r.sort.map(&:last)) if
        r.size == cmds.size
    end
  end
  w.yield
```

```
def system async *args
  tuple = system spawn(args)
  Thread.new do
    begin
      result = system wait(*tuple)
    rescue Exception => e
      EM.instance variable get(
        :@error handler).call(e)
    ensure
    # ...
ennnd
```

```
def system_async *args
    # ...
    if Concurrency::Model == Fiber
        EM.next_tick{ yield(result) }
        else
        yield(result)
        end
```

```
def system spawn args
  err = if %w[test production].
                include?(Rails.env)
          IO::NULL
        else
          Rails.logger.debug(
             "Shell: #{args.inspect}")
          $stderr
        end
```

```
def system_wait pid, rd, wr
  wr.close
  result = rd.read
  Process.waitpid(pid)
  result
end
```

Ruby Codes (Cont.)

- server.rb
- error rb
- eagerload.rbwaiter.rb

- newrelic.rb
- shell.rb

waiter.rb

```
require 'thread'
require 'fiber'
class Waiter
  def self.fiber around?
    RC::RootFiber != Fiber.current &&
      Thread.main == Thread.current
  end
end
```

waiter.rb (Cont.)

```
class Waiter
  def initialize
    if Waiter.fiber around?
      @fiber = Fiber.current
    else
      @condv = ConditionVariable.new
      @mutex = Mutex.new
    end
  end
end
```

waiter.rb (Cont.)

```
class Waiter
  def resume value
    if @fiber
      @fiber.resume(value)
    else
      @value = value
      @condv.broadcast
    end
  end
end
```

waiter.rb (Cont.)

```
class Waiter
  def yield
    if @fiber
      Fiber.yield
    else
      @mutex.synchronize{
        @condv.wait(@mutex) }
      @value
    end
  end
end
```

Ruby Codes (Cont.) (Cont.)

- image util.rb
- run later.rb
- /search
- /create_and_share

collage_web.rb

```
class CollageWeb < Collage</pre>
  extend ImageUtil
  def self.create from urls or url,
                         fallback urls=[]
    create(:image =>
      merge images (
        urls2images(urls or url,
                     fallback urls)))
  end
```

end

image_util.rb

require 'tempfile'

```
require 'shell'
module ImageUtil
  module function
  def urls2files urls or url,
                  fallback urls=[]
    urls2images(urls or url,
             fallback urls).map(&:file)
  end
end
```

```
module ImageUtil
  module function
  def urls2images urls or url,
                   fallback urls=[]
    urls = [urls or url].flatten
    w = Waiter.new
    results = {}
    # ...
    w.yield.sort.map(&:last).compact
  end
end
```

```
def urls2images ...
urls.zip([fallback urls].flatten).
  each.with index do | (url, fb), idx|
    ImageFile.from url(url, fb) do |img|
      results[idx] = img
      w.resume(results) if
        results.size == urls.size
    end
  end
w.yield.sort.map(&:last).compact
```

```
module ImageUtil — util.rb (Cont
  module function
  def merge images images
    #
  end
  ImageStruct =
    Struct.new(:width, :height, :file)
  class ImageFile < ImageStruct</pre>
    include ImageUtil
 #
```

```
module ImageUtil
  module_function
  def wget?
    @wget ||= !`which wget`.strip.empty?
  end
end
```

```
image util.rb (Cont
module ImageUtil
  module function
  def wget command path, url
    if wget?
      ['wget', '-0', path,
       '--no-check-certificate', url]
    else
      ['curl', '-o', path,
       '--insecure'
                               , url]
    end
  end
```

```
def identify command path
  ['identify', '-format', '%wx%h',
   path]
end
def identify result output
  if m = output.match(/(\d+)x(\d+)/)
    [m[1], m[2]].map(\&:to i)
  else
    nil
  end
```

Ruby Codes (Cont.) (Cont.)

- image_util.rb
- run later.rb
- /search
- /create_and_share

run later.rb

```
require 'concurrency'
module RunLater
  module function
  def with
    if EM. reactor running?
      EM.next tick{
        Concurrency.wrap{ yield } }
    else
     yield
    end
  end
```

Ruby Codes (Cont.) (Cont.)

- image_util.rb
- run later.rb
- /search
- /create_and_share

/search

```
get '/search' do
  params['q'].blank? &&
    missing arguments('q') # this raises
  images, total = if limit == 0
    [[], 0]
  else
    # . . .
  end
```

/search (Cont.)

```
params['engine']
case
when 'pinterest'; sch pinterest
    'pinterest api'; sch pinterest api
when
when
    'bing'; sch bing offset, limit
     'azure'; sch azure offset, limit
when
when 'google'; sch google offset, limit
when 'flickr'; sch flickr offset, limit
             ; sch google offset, limit
else
               # default google
end
```

/search (Cont.)

/search (Cont.)

```
def search from google offset, limit
  google = RC::GoogleImage.new(
    referer: request.url,
     userip: request.ip)
  search image and flatten(
    google, offset, limit) do [q, n, i]
      google.search image(
        q, rsz: n, start: i)
  end
```

Ruby Codes (Cont.) (Cont.)

- image_util.rb
- run later.rb
- /search
- /create_and_share

```
post '/collages' do
  render_collage create_collage
end
```

```
post '/collages/create_and_share' do
  render_collage create_collage{ |c|
    transfer "/collages/#{c.id}/share" }
end
```

```
def transfer path
  status, headers, body =
    jellyfish.call(
      env.merge('PATH_INFO' => path))
  self.status     status
  self.headers headers
  self.body     body
end
```

```
if image
    # ...
else
    missing_arguments('image')
end
```

```
c para = params.merge('plist' => plist)
c para.delete('image')
if rc tumblr &&
   for auths and clients.empty?
 # no user for tumblr
else # check user for any other cases
  c para['user'] =
    current user with check
end
```

```
c = Collage.create_with_params(c_para)
# make paperclip happy
image.extend(OriginalFilename)
image.original_filename =
  params['image'][:filename]
c.save_image(image, &upload_callback)
```

```
def save_image img, &block
  save_dimension img
  process_thumbnails img, &block
  self
end
```

```
def process thumbnails img, &blk
  if blk
    RunLater.with do
      process thumbnails inplace img
      if blk.arity == 1; blk.call(self)
                        ; blk.call; end
      else
    end
  else
    process thumbnails inplace img
  end
```

crazy inspect

- crazy inspect
- who cares backtrace for 404

- crazy inspect
- who cares backtrace for 404
- assets pipeline and database connection

- crazy inspect
- who cares backtrace for 404
- assets pipeline and database connection
- fiber stack overflow

- crazy inspect
- who cares backtrace for 404
- assets pipeline and database connection
- fiber stack overflow
- fiber database connection pool

- crazy inspect
- who cares backtrace for 404
- assets pipeline and database connection
- fiber stack overflow
- fiber database connection pool
- thread-safe <u>kaminari patch</u>

