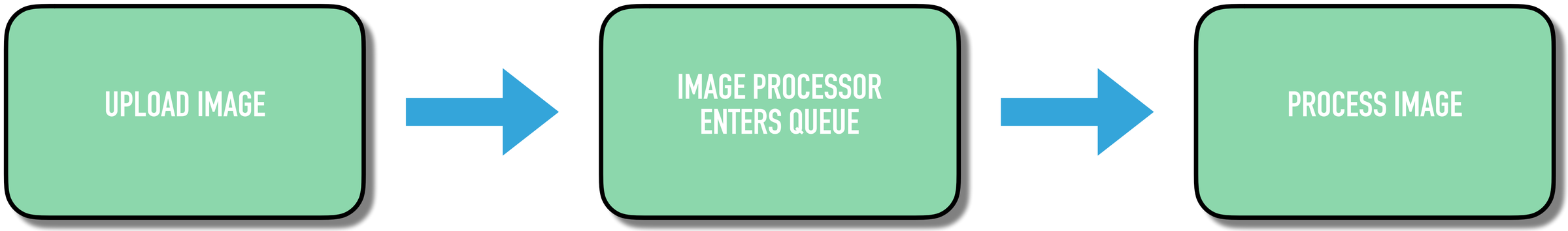
CHUCK REEVES @MANCHUCK

BACKGROUND JOB PROCESSING

DO'S AND DON'TS

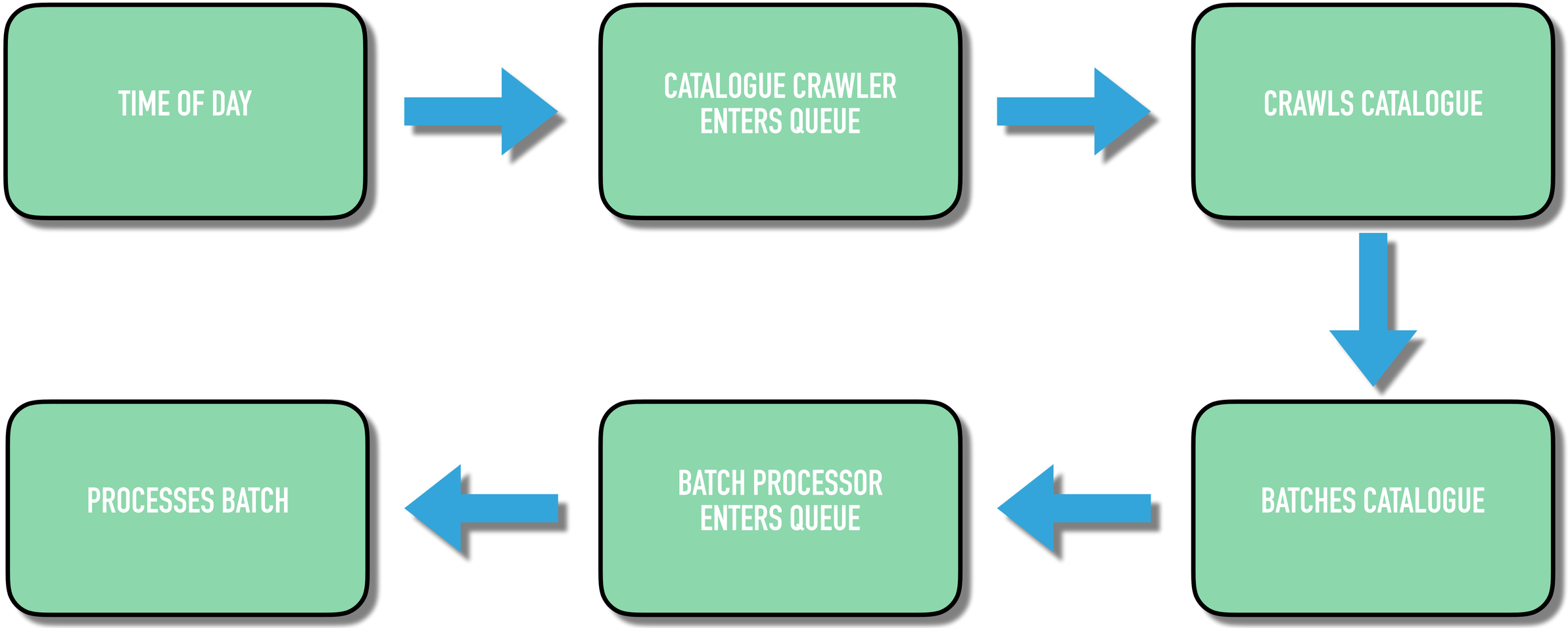
BACKGROUND JOB PROCESSING - DO'S AND DON'TS



THE ~~PROBLEM~~ CHALLENGE

- Crawl Google Merchant Center
- Crawl Google Ad Words
- Crawl Google Analytics
- Match Data
- Analyze Data
- Make Decisions
- POST/PUT Google Ad Words
- Within 4 hours

BACKGROUND JOB PROCESSING - DO'S AND DON'TS




Time is an illusion.
Lunchtime doubly so.
Douglas Adams
quotefancy

REAL TIME IS BULLSHIT

REALTIME (ACCORDING TO POSIX)

Ability of a system to provide a service within a **bounded** response time

~~PROBLEMS~~ CHALLENGES YOU WILL FACE

- Memory
- IO/Socket Blocks
- Network stack
- Timing of sequential jobs
- Queue

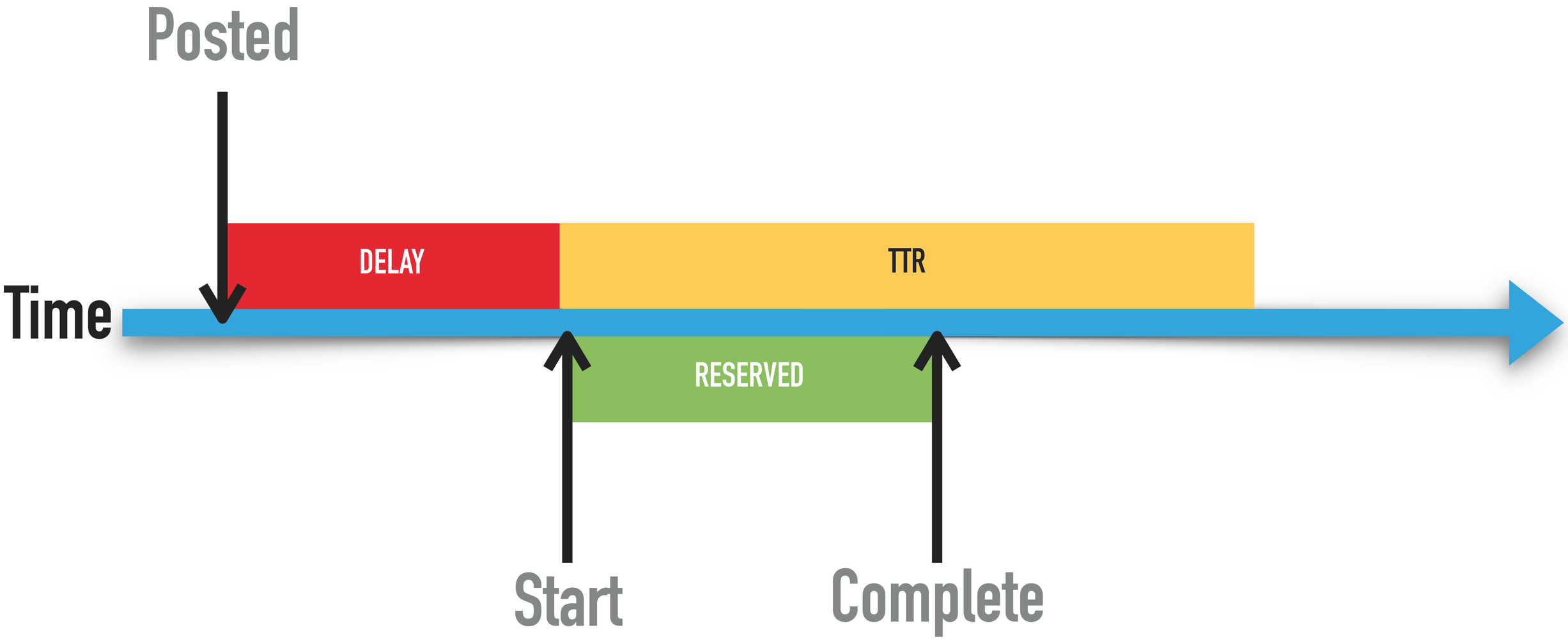
QUEUES

- Push/Pull (poll)
- Priority
- Time To Live
- Delay
- Tubes

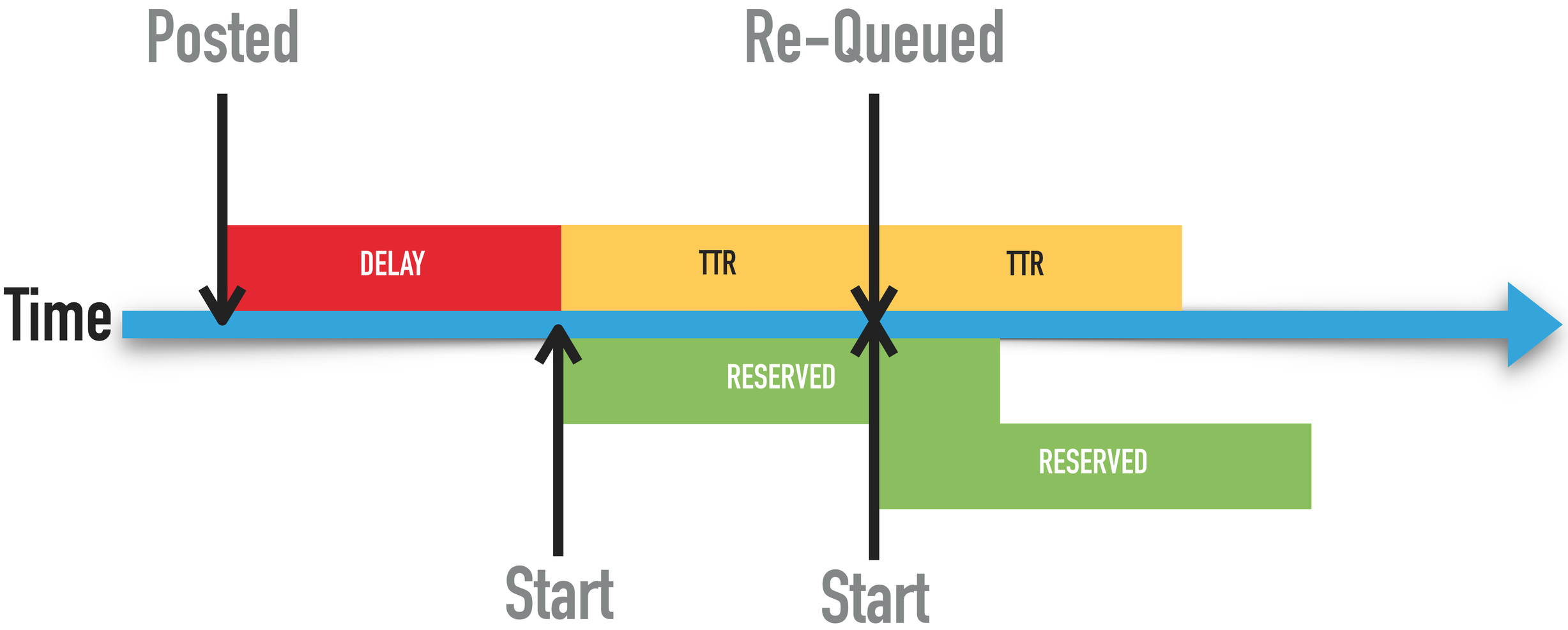
TERMS

- Ready
- Reserved
- Delayed
- Buried
- Touch

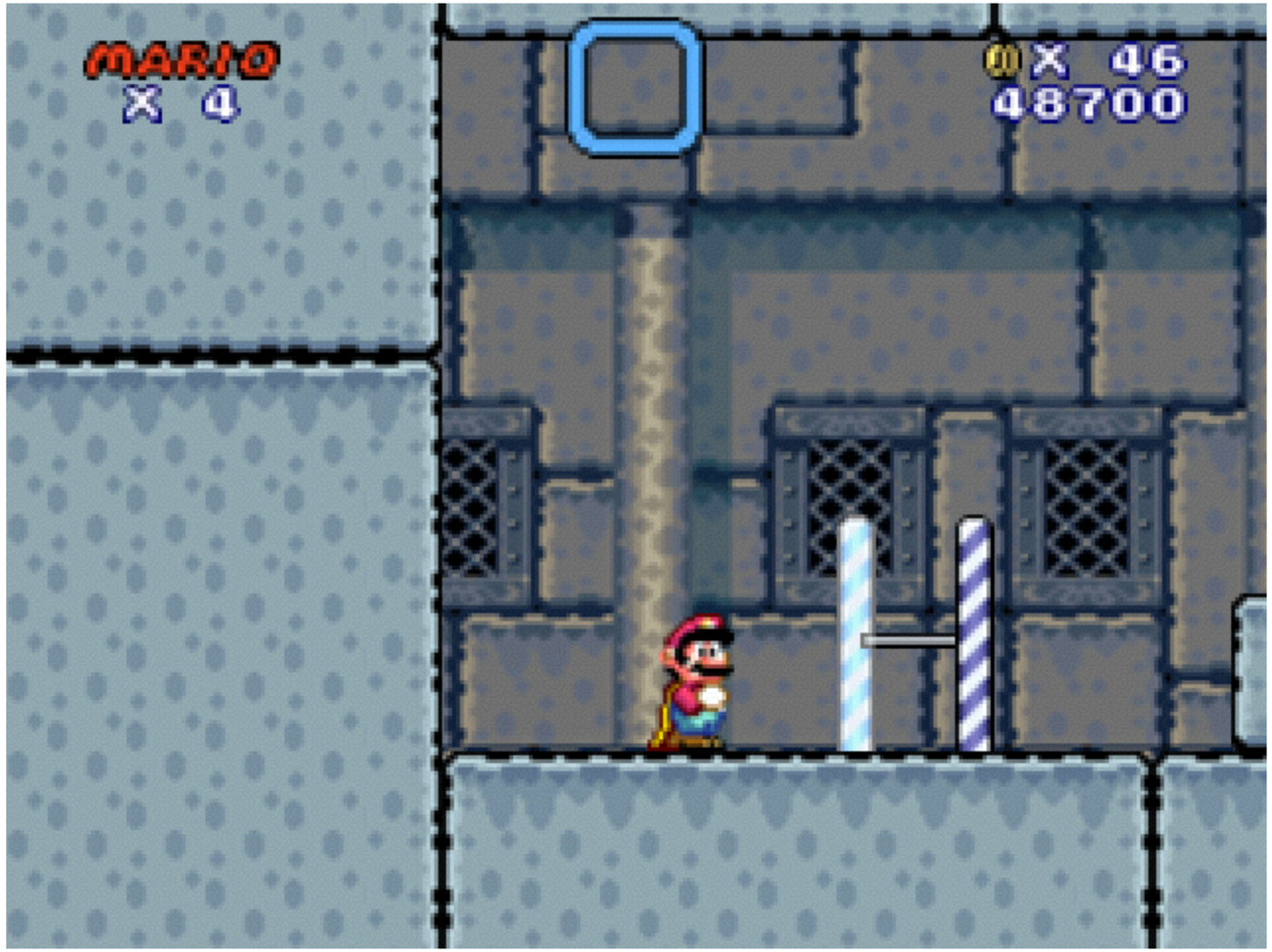
JOB LIFE CYCLE



JOB LIFE CYCLE GOTCHA



CHECKPOINTS



PUSH OR PULL?

PUSH OR PULL?

PUSH

- Hosted
- Queue Starts worker
- Scales with size of queue
  - Limits number of workers
- Limited control over worker environment

PULL

- Self-Hosted
- Process needed to check for new queue
  - Resources required for polling
- Additional delay
- Control of the Environment

BEANSTALKD – PULL

PROS

- In Memory
- Lightweight text protocol
- Useful reporting

CONS

- Cannot scale

GEARMAN - PULL

PROS

- In Memory
- Distributed
- Workers can specify their capabilities
- Persistent with MySQL Plugin
  - But Don't use it

CONS

- PHP Lib has Limited reporting
- Workers connecting to multiple job servers can cause delays in jobs
- Job types can be confusing
- Simple Priorities

RESQUE - PULL

PROS

- Redis
- Easy Protocol
- Management GUI

CONS

- Not really a Job Queue
- Limited reporting
  - Roll your own if you know Redis

RABBITMQ – PULL

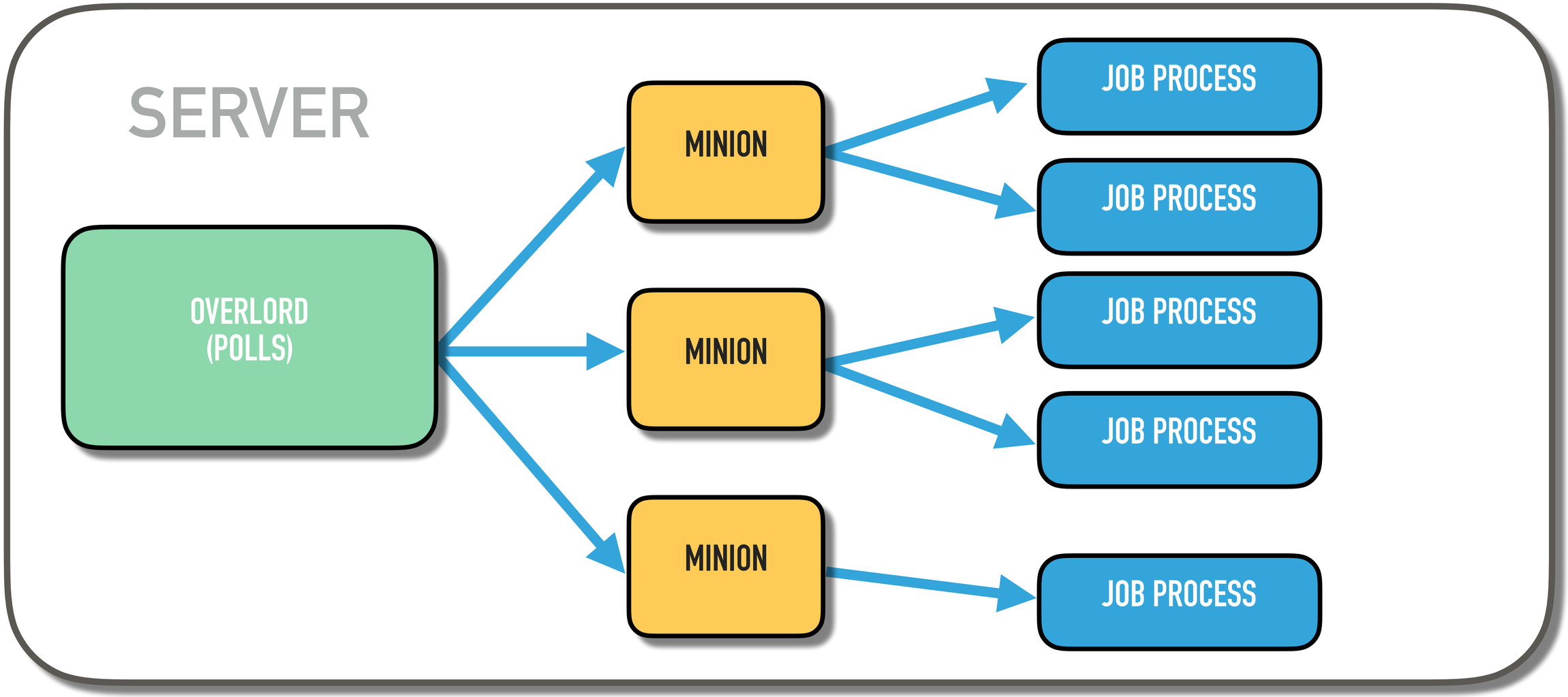
PROS

- Easy Protocol
- Management GUI
- Persistent

CONS

- Not really a Job Queue
- Limited reporting
  - Roll your own if you know Redis
- Memory error

WHAT YOU SHOULD DO – POLLING QUEUES



WHAT YOU SHOULD NOT DO – POLLING QUEUES

- Process every job with one process
- Have worker processes talk to the queue
- Use PCNTL (well...)
  - PHP 7 Could be better
    - But not on Windows

JOB

```php
interface JobInterface
{
    public function perform();

    public function getArrayCopy();

    public function exchangeArray(array $data);
}
```

JOB

```php
public function perform()
{
    if (!$this->getUser() instanceof UserInterface) {
        throw new \RuntimeException();
    }
    /**@var SuggestionContainer $masterContainer*/
    $masterContainer = new SuggestionContainer();

    $this->getFilterSuggestions($masterContainer);
    $this->applyRules($masterContainer);
    if (count($masterContainer) > SuggestionEngine::MAX_CAPACITY) {
        $masterContainer->exchangeArray(
            array_rand($masterContainer->getArrayCopy(), SuggestionEngine::MAX_CAPACITY)
        );
    }
    $this->attachSuggestions($masterContainer);
}
```

JOB AGGREGATE

```php
interface JobAggregateInterface extends JobInterface
{
    public function createJob(JobService $jobService);
}
```

JOB SERVICE

```php
interface JobService
{
    public function __construct(QueueAdapterInterface $adapter);

    public function reserveJob(JobInterface $job);

    public function touchJob(JobInterface $job);

    public function buryJob(JobInterface $job);
}
```

JOB HANDLER

```php
interface JobHandlerInterface
{
    public function __construct(QueueAdapterInterface $adapter);

    public function work($tube = 'default');
}
```

JOB HANDLER

```php
public function workAction()
{
    $request = $this->getRequest();
    if (!$request instanceof ConsoleRequest) {
        throw new \RuntimeException('Invalid Request');
    }

    $queue    = [$request->getParam('queue', 'default')];
    $interval = $request->getParam('interval', 5);
    $adapter  = new ResqueAdapter($queue, $this->services);
    $adapter->setLogger($this->getLogger());
    $adapter->setInterval($interval);

    $handler = new JobHandler($adapter)
    $this->getLogger()->notice('Starting Worker');
    $handler->work();
}
```

JOB HANDLER

```php
public function work($tube = 'default')
{
    while($this->adapter->work()) {
        $job = $this->adapter->getJob();
        $fullCommand = $this->phpPath . ' '
            . APPLICATION_PATH . '/public/index.php '
            . $job->getId();
        $this->getLogger()->notice('Executing: ' . $fullCommand);
        system($fullCommand, $exitCode);

        if ($exitCode != 0) {
            $this->getLogger()->notice(sprintf('Job %s failed to execute');
        }

        $this->jobService->buryJob($job)
    }
}
```

AMAZON SQS/SNS/LAMBDA

- Scales Easily
- Easy Protocol (API)
- Management GUI
- Workers must be pre-defined
- Jobs can take between 10 seconds to 15 min to pop out
- Python, Java or NodeJS

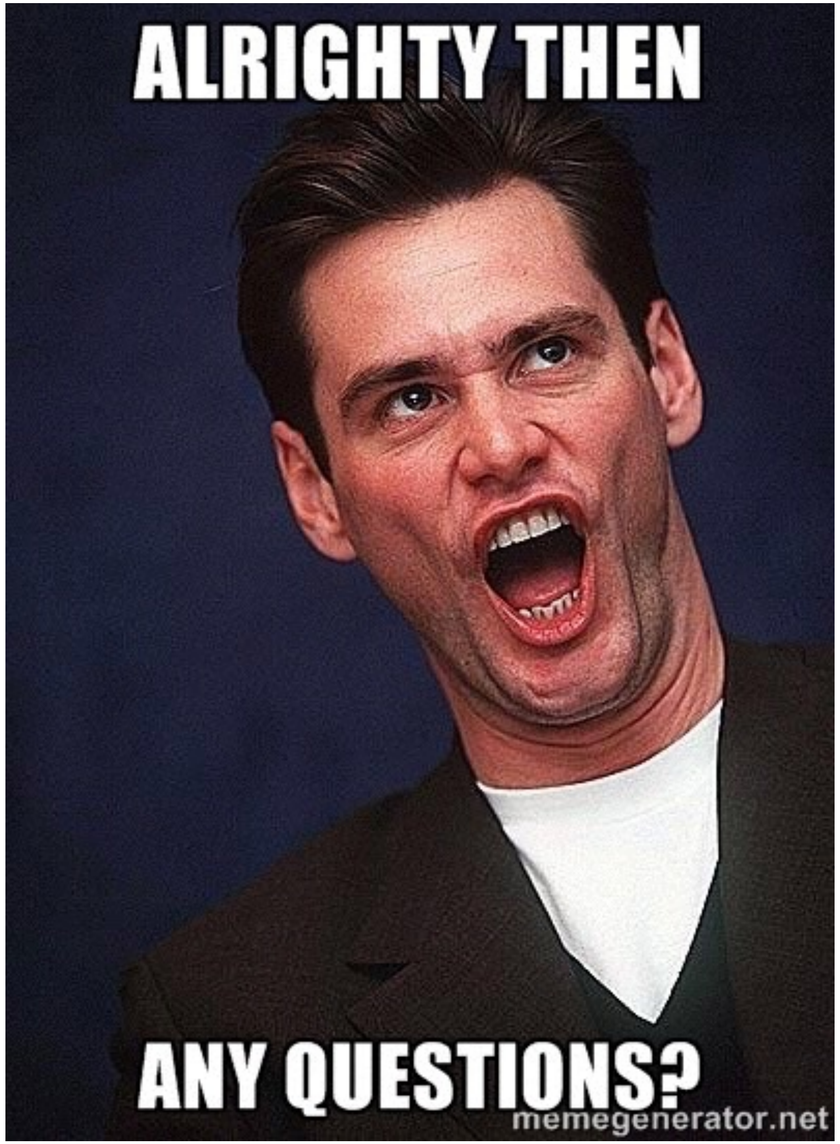
GOOGLE TASK QUEUE

- Scales Easily
- Queues can be programmatically created or through API
- Supports push and pull queues
- Schedule Tasks
- On Demand TTR

FINIAL THOUGHTS

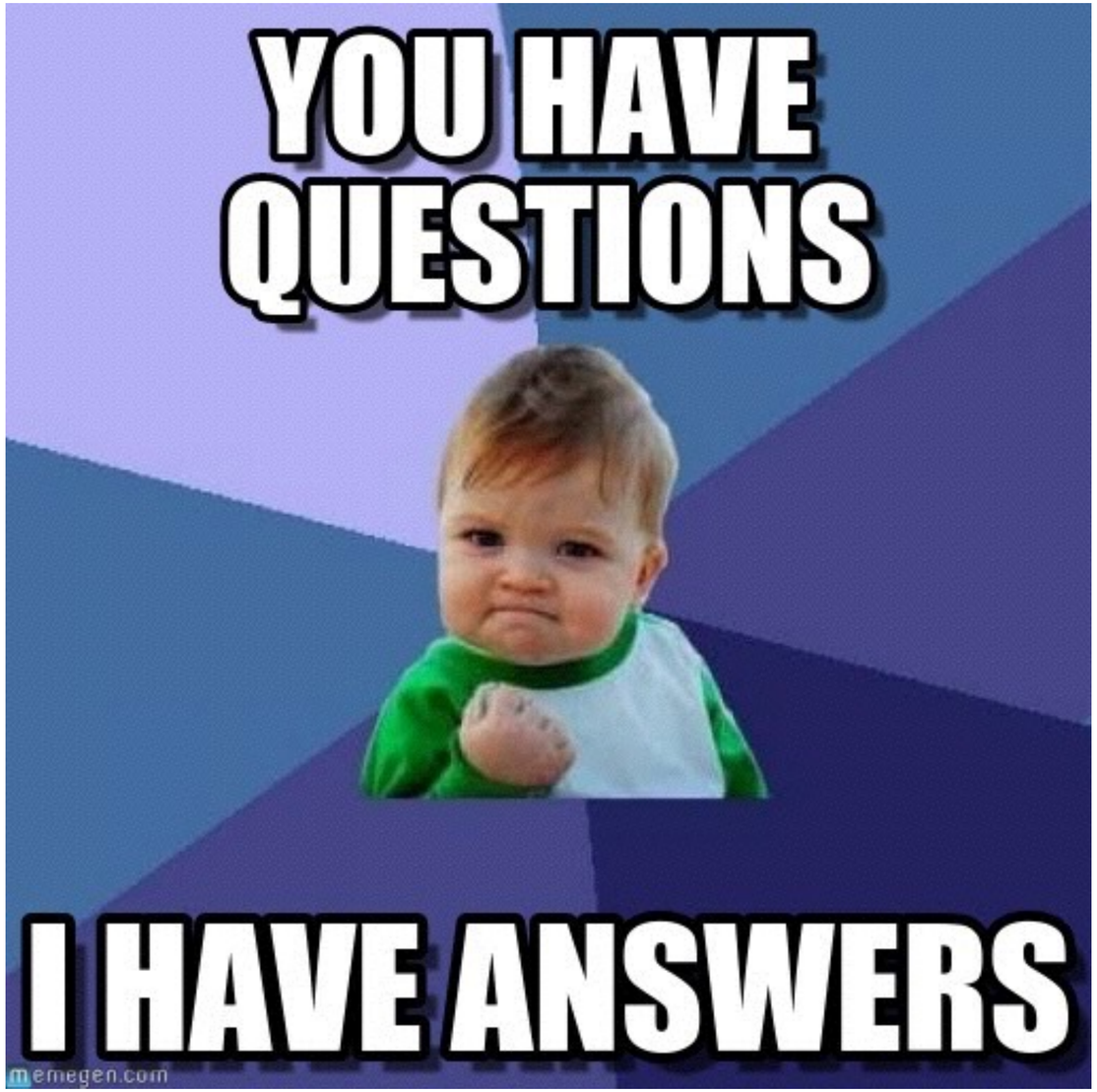
- Limit Connections to the Database
- Validate data going into a job
- Handel Errors
  - Remove job on error
- Central Logging

BACKGROUND JOB PROCESSING - DO'S AND DON'TS



BACKGROUND JOB PROCESSING - DO'S AND DON'TS

NEW
MORE ESSENTIAL OILS
AVALON ORGANICS
SHAMPOO
EXTRA MOISTURIZING
FRAGRANCE FREE
OLIVE & GRAPE SEED
Organic Olive & Grape Seed Essential Oils, Borage, Organic Aloe and Vitamins deeply hydrate, balance and restore dry, brittle, damaged hair for strong, healthy-looking shine.
MADE WITH 70% ORGANIC CONTENT
CONSCIOUSNESS IN COSMETICS®
SAFE, NATURAL, EFFECTIVE.
PRO-ORGANIC BODY CARE WITHOUT SYNTHETIC FRAGRANCES, PHTHALATES, PARABENS OR SULFATES.
325 mL e 11 fl oz


YOU HAVE QUESTIONS
I HAVE ANSWERS
memegen.com

CHUCK REEVES @MANCHUCK

THANKS