

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions

Spilgames Storage Platform

Enrique Paz

Senior Backend Developer

21/03/2013

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions

Introduction

About Me

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions



- Passionate Erlang developer
- Testing enthusiast
- Love beautiful code!

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions

- Gaming Platform
- Serving data to 190+ countries world-wide
- 200+ million unique users per month
- Multiple Platforms: Desktop, Mobile Native & Web
- 300+ employees
- Offices in The Netherlands & China
- Revenue: Advertising & EUM

Gamesgames.com



Play Free Online Games on GamesGames.com

JOIN FOR FREE or

Sign In



Search for Games

Search for...

HOME

ACTION

ADVENTURE

BOARD & CARD

GIRLS

MULTIPLAYER

PUZZLE

RACING

SKILL

MORE

GAMES

Home

New Games

Popular Games

My Favorites

Social Games

Award games

Creation games

Highscore games

Game Albums

Avatar Games

Mobile games

Videos

Cash Games (King)

Action



Slotomania

Roll up and reel in the jackpot in slot-machine paradise!

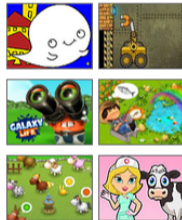
1 2 3

PLAY NOW

POPULAR

1. Slotomania
2. Family Barn
3. Butterfly Kyodal
4. Elements Makeover: F...
5. Galaxy Life
6. Dream Pet Link
7. My "Dear" Boss: Physi...
8. Chicken Soup: Sara's ...
9. Ancient Jewels
10. Thrill Rush 2

RECENTLY PLAYED



ADVERTISEMENT



NEW GAMES



Snoring 2: Winter...



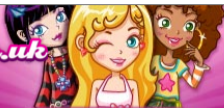
Celebrity Spa



Accurate Slapshot

GirlsGoGames.co.uk

Free online Games for girls on GGG.co.uk



GGG.co.uk

SIGN UP FOR FREE! OR SIGN IN

Search Games

SEARCH

Home

Games for girls

Popular

Game Albums

Social Games

Award Games

Creation Games

High-Score Games

Mobile games

Sara's Cooking Class

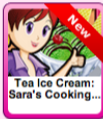
Adventure Games

Animal Games

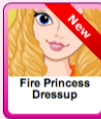
Cooking Games

Games for girls

Like 8.6k



Tea Ice Cream:
Sara's Cooking...



Fire Princess
Dressup



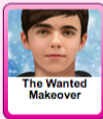
Selena Gomez
Fan Room



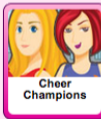
Angel Cat
Dressing Up



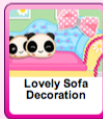
Rainbow
Unicorn#



The Wanted
Makeover



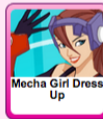
Cheer
Champions



Lovely Sofa
Decoration



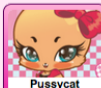
Chicken Soup:
Sara's Cooking...



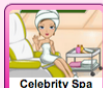
Mecha Girl
Dress Up



Sailor Senshi



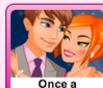
Pussycat



Celebrity Spa



Tiny Dragon



Once a

Rainbow Unicorn#



Harness all the colours of the rainbow to make this unicorn glow!

1 2 3

PLAY NOW!

Popular

1 Celebrity Spa

2 Chicken Soup: Sara's Cooking...

Games

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions

GirlsGoGames.com

SARA'S COOKING CLASS

Pancakes

Play

Recipe

Instructions

More Games

English (US)

Add To Your Site

Galaxy Life

Gratis Vrienden Uitnodigen Berichten Logboek Klassement

Mooi zo! Met deze Starinators winnen we vast!
Zoek Firebit in het planetenstelsel.

STELSEL MALTMEAT (779,635)

OPHOUDEN OPHOUDEN OPHOUDEN OPHOUDEN OPHOUDEN TYLAFSTIER ELDERDY

PLANEEET (779,635)

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions

Another Storage

Another Storage

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions



LAMP Stack & MySql

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

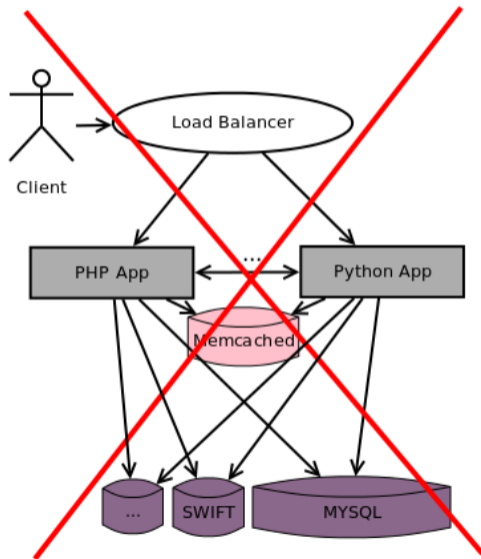
Scenarios

Lessons Learned

Currents Status

Contributions

- Not all developers are DB experts
- Difficult to shard the databases
- Storage model all over the place
- Security
- Performance
- Caching
- ...



Our Ambition

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datcenters

Disaster

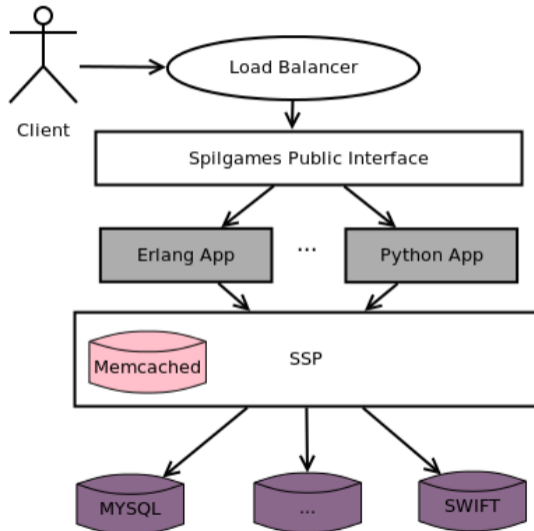
Scenarios

Lessons Learned

Currents Status

Contributions

- Transparent sharding layer
- Sharding on data ownership
- High availability
- Centralized caching layer
- Storage engine agnostic
- One strict data model
- Transparent storage changes
- Scaling geographically



Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions

System Properties

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions

- Be always available
- Avoid global locks
- Accept change as the only constant
- Embrace inconsistencies
 - ▶ Hardware breaks down (power failures)
 - ▶ Version mismatches (upgrading system not atomic)
 - ▶ State mismatches (adding new machines)

A Key-Value Store With Schema

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datcenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions

- **Buckets**
 - ▶ Largely generated OTP applications
 - ▶ Offer a CRUD-like interface (with filters)
- **GIDs** (64 bits) identify the data owner
 - ▶ user
 - ▶ game
- Buckets can use different **storage engines**
 - ▶ Several MySQL tables in different databases
 - ▶ Just a binary storage (SWIFT)
 - ▶ ...
- Data for a bucket/GID is cached
- Requests can be atomic per bucket/GID

Optimistic Operations

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions

- Speed > Consistency
- Losing some updates in case of crash is affordable
- Act first on cache and then on disk
- No warranties of eventual consistency upon crashes
 - ▶ i.e. Activity feeds
 - ▶ i.e. Popular games list

Pessimistic Operations

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions

- Consistency is key and confirmation is required
- Dealing with critical data
- Persist data and, upon success, update cache
 - ▶ i.e. Payments
 - ▶ i.e. Personal information

How does it work?

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions



System Components

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

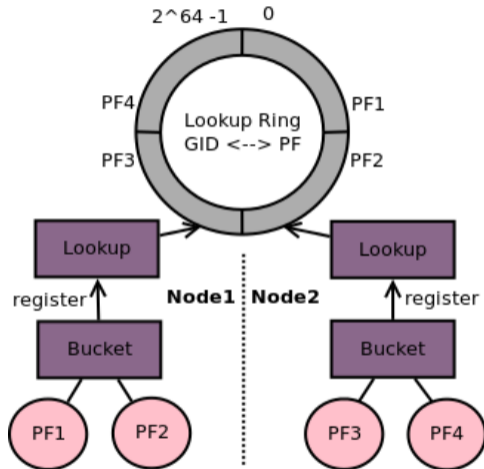
Scaling
Datacenters

Disaster
Scenarios

Lessons Learned

Currents Status
Contributions

- **lookup** application in all nodes. Uses a **hashing ring** (mnesia):
 - ▶ replicated in all nodes
 - ▶ ram_copies
 - ▶ dirty reads
 - ▶ transactional writes
- **Buckets** have Pipeline Factories
- Buckets register **PFs** in lookup



Tracing Pessimistic Operations

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

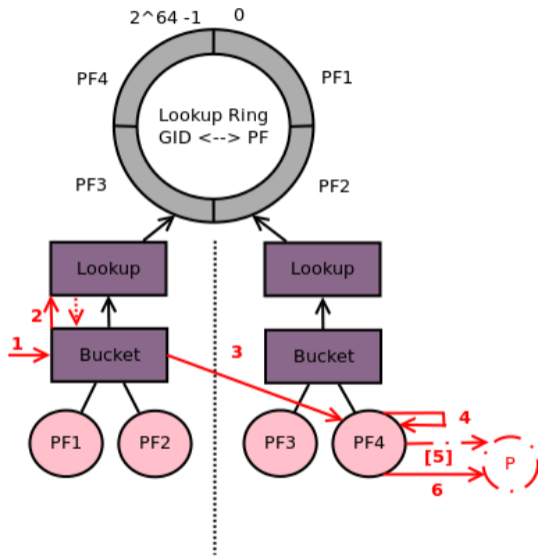
Scaling
Datacenters

Disaster
Scenarios

Lessons Learned

Currents Status
Contributions

1. Bucket/GID request in a node
2. Local lookup to find a PF
3. PF receives request
4. PF builds job
5. PF ensures Pipeline for GID
6. PF queues operation in Pipeline



Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datcenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions

- “Why do we need pipelines?”
- “Sequential == Bottleneck !!!”
- “Don’t you guys know Erlang is about **parallelizing** work?”



Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions

- **CONS**

- ▶ Sequential (read) access for hotspots
i.e. Popular games
- ▶ Optimization: read from SSP cache in PF

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions

- **CONS**

- ▶ Sequential (read) access for hotspots
i.e. Popular games
- ▶ Optimization: read from SSP cache in PF

- **PROS**

- ▶ No need for storage engines to support global locks
- ▶ A bucket can combine several engines

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons

Learned

Currents Status

Contributions

- **CONS**
 - ▶ Sequential (read) access for hotspots
i.e. Popular games
 - ▶ Optimization: read from SSP cache in PF
- **PROS**
 - ▶ No need for storage engines to support global locks
 - ▶ A bucket can combine several engines
- Requests to most GIDs (users) are evenly distributed

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions

- Schema Versions determine allowed operations and storage(s)
- Client is not aware of them
- Max 2 schema versions of a bucket at the time
- Schema version can be changed at bucket/GID level

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions

- Useful for partitioning big blocks of data
- Shard points to the physical location of the data
- Sharding rules are bucket specific. Default is $GID \% Shards$
- bucket/GID combinations can be migrated between shards

Working With Versions & Shards

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datcenters

Disaster

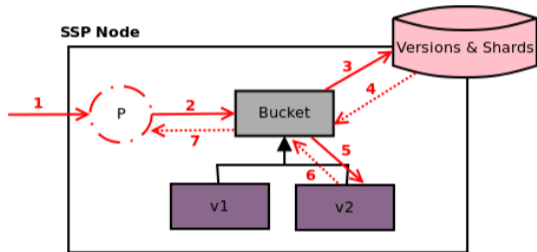
Scenarios

Lessons Learned

Currents Status

Contributions

1. *insert(Bucket, Gid, Value)*
2. *insert(Gid, Value)*
3. *get_vs(Bucket, Gid)*
4. *{v2, Shard1}*
5. *build_job(insert, Gid, Shard1)*
6. *{ok, InsertJob}*
7. *{ok, InsertJob}*



One API To Rule Them All

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions

- “Don’t care where it is, just want my data!!!”

One API To Rule Them All

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling
Datacenters

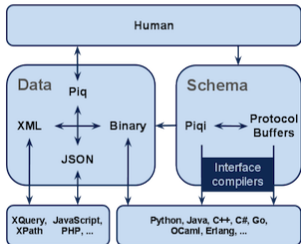
Disaster
Scenarios

Lessons Learned

Currents Status

Contributions

- “Don’t care where it is, just want my data!!!”
- **PIQI** helps with the API



- ▶ Erlang client + Protocol Buffers
- ▶ HTTP + JSON
- ▶ HTTP + Protocol Buffers

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions

Worldwide

Worldwide

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions



Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling
Datacenters

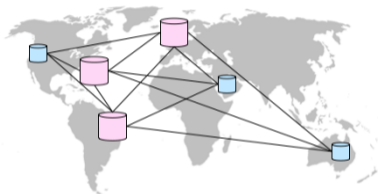
Disaster
Scenarios

Lessons Learned

Currents Status

Contributions

- **Master Datacenters**
 - ▶ Have persistent storage
 - ▶ Can own GIDs
 - ▶ GIDs can be migrated between Master DCs
- **Satellite Datacenters**
 - ▶ Don't have persistent storage
 - ▶ Easy to setup and decommission
 - ▶ Virtual/Cloud-based



Start With One DC

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

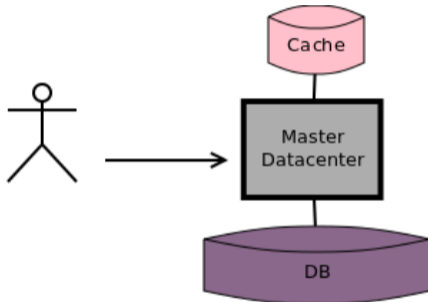
Disaster

Scenarios

Lessons Learned

Currents Status

Contributions



Scale Up A Satellite Where Needed

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

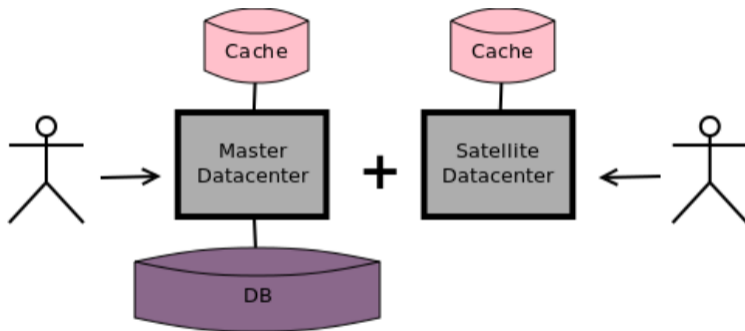
Scaling
Datacenters

Disaster
Scenarios

Lessons Learned

Currents Status

Contributions



Turn Satellites Into Masters When Ready

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

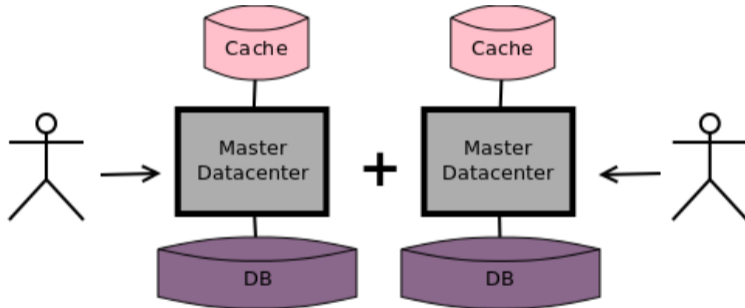
Scaling
Datacenters

Disaster
Scenarios

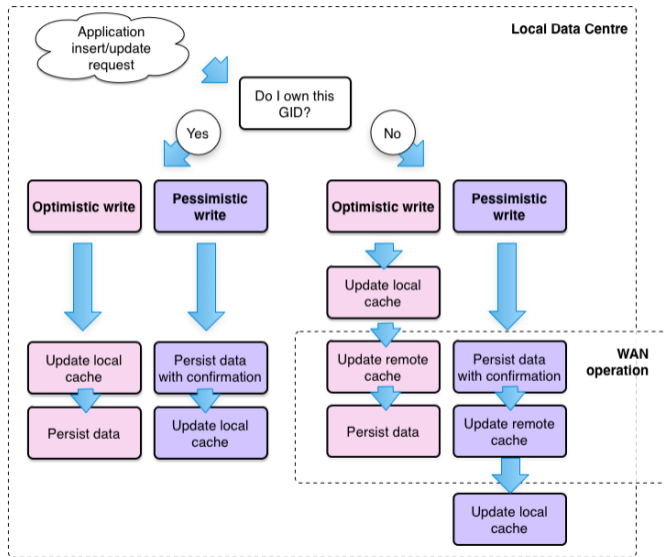
Lessons Learned

Currents Status

Contributions



Working With Multiple Datacenters



Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling
Datacenters

Disaster
Scenarios

Lessons Learned

Currents Status

Contributions

Disaster Scenarios

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions



Losing A Satellite Datacenter

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

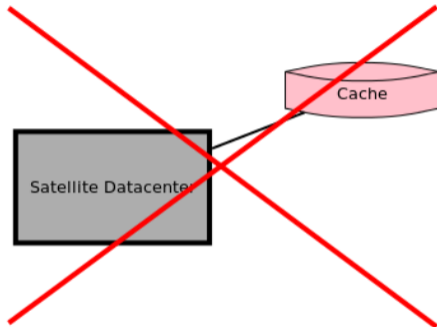
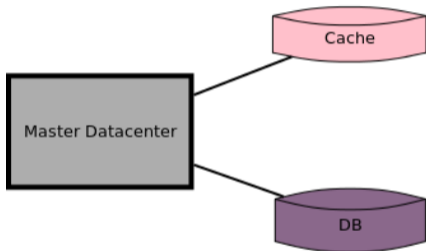
Disaster

Scenarios

Lessons Learned

Currents Status

Contributions



Losing A Master Datacenter

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

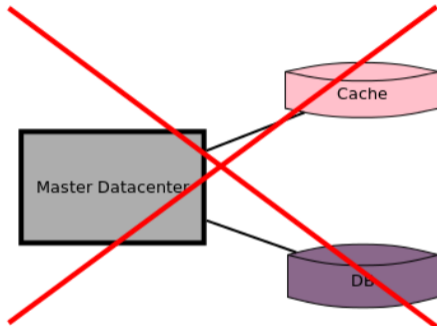
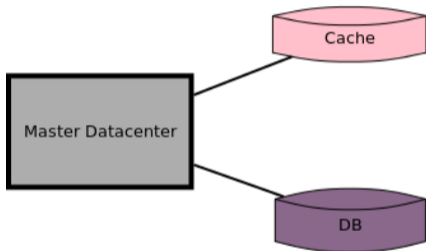
Disaster

Scenarios

Lessons Learned

Currents Status

Contributions



Losing A Master Datacenter

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datcenters

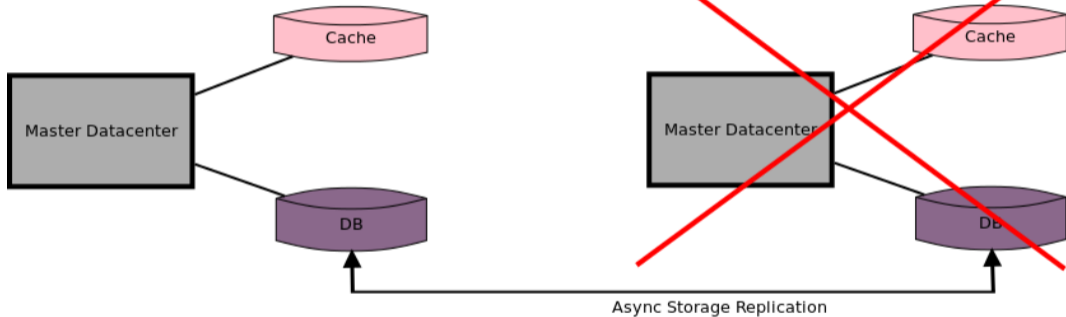
Disaster

Scenarios

Lessons Learned

Currents Status

Contributions



Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions

Lessons Learned

Where We Are

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datcenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions

- Using simple buckets on LIVE in one DC
- Added relup support for bucket only updates
- Hammering SSP using property based testing
- Integrating restricted search capabilities
- Testing the WAN protocol for inter DC communication
- More buckets to go live in H1 2013
- Satellite DCs coming on H2 2013

What We've Used

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions

- Emysql
 - ▶ (+) multi-database transaction support
 - ▶ (*) multi-timezone support
- Eep0018/Jiffy
- Estatsd
- PropEr
- Poolboy
- Lager
- Rebar
 - ▶ (*) semantic versioning, i.e. [">=1.3.1", "<2.0.0"]
 - ▶ (*) shared dependencies
 - ▶ (*) xref fixes
- Piqi
- BashoBench
 - ▶ (+) Several tests on the same plot support

Questions?

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions



Thanks!

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions

Introduction

Spilgames & Me

Another Storage

Motivation

System Properties

Overview

Internals

Versions & Shards

Worldwide

Scaling

Datacenters

Disaster

Scenarios

Lessons Learned

Currents Status

Contributions

**WE'RE
HIRING!**

<http://www.spilgames.com/careers/job-openings/>