Master’s Thesis

Original Idea

Thomas Fankhauser
Super Scale Systems
2012

SUPER SCALE SYSTEMS
HOW TO SCALE THE SOCIAL WEB

Media University
Stuttgart
Project Roadmap

Work Packages & Publications

WP1: Web Scaling Framework
- April 2013

WP2: Post-Processing
- November 2014

WP3: Thesis

April 2016

CONFERENCE PAPER
IEEE ICC14

JOURNAL PAPER
IEEE TSC

CONFERENCE PAPER

THESS
Work Package 1

Web Scaling Frameworks: IEEE ICC Conference Paper

Web Scaling Frameworks:
A novel class of frameworks for scalable web services in cloud environments

Thomas Fankhauser⇤, Student Member, IEEE, Qi Wang†, Member, IEEE, Ansgar Gerlicher⇤, Senior Member, IEEE, and Xinheng Wang†, Member, IEEE

Abstract—The social web and huge growth of mobile smart devices dramatically increases the performance requirements for web services. State-of-the-art Web Application Frameworks (WAFs) are often too expensive or unable to cope with the huge amounts of requests in aggregation. Handling the exponentially increasing global requests adds the requirements of being able to run multiple instances of an application for highly scalable web services. The major challenges that are introduced by this requirement are the management of the shared resources, the balancing of the requests among all instances and the decision when to open or terminate instances. These challenges are collectively referred to as horizontal scaling [13] [14] [16].

Our experiments have showed that WAFs have different strengths and weaknesses. A highly abstracted WAF like Ruby on Rails, for example, was slower than the very thin WAF node.js but more powerful regarding data management and interface rendering. If a web service needs to provide both a fast and slim JSON API and a full blown HTML website it is the best solution to combine both WAFs. As both the horizontal scaling and web service composition are very complex matters, it makes sense not to introduce them to WAFs but offload them to another layer - the Web Scaling Framework (WSF) proposed in this paper. Fig. 1 illustrates a WSF that incorporates multiple WAF applications.

I. INTRODUCTION

The continuous growth of smart mobile devices in combination with social web services increases the number of requests that need to be processed by modern web platforms in a timely fashion. Whereas cloud computing provides the ability to provision resources on-demand, the overhead of state-of-the-art Web Application Frameworks (WAFs) do not offer integrated scaling concepts to deal with automatic resource-provisioning and elastic scaling or ensure a guaranteed minimum response time.

They are rather designed to abstract common functionality needed for web application development including data management, interfacing, session-handling and response-generation. Today, users progressively access the social web from anywhere using their mobile smart devices, which leads to increased traffic. A single computing resource might not be able to satisfy such an amount of requests — only the conjunction of multiple computing resources, where each resource gets a small share of the total requests, allows to handle such huge amounts of requests in aggregation. Handling the exponentially increasing global requests adds the requirements of being able to run multiple instances of an application for highly scalable web services. The major challenges that are introduced by this requirement are the management of the shared resources, the balancing of the requests among all instances and the decision when to open or terminate instances. These challenges are collectively referred to as horizontal scaling [13] [14] [16].

Our experiments have showed that WAFs have different strengths and weaknesses. A highly abstracted WAF like Ruby on Rails, for example, was slower than the very thin WAF node.js but more powerful regarding data management and interface rendering. If a web service needs to provide both a fast and slim JSON API and a full blown HTML website it is the best solution to combine both WAFs. As both the horizontal scaling and web service composition are very complex matters, it makes sense not to introduce them to WAFs but offload them to another layer — the Web Scaling Framework (WSF) proposed in this paper. Fig. 1 illustrates a WSF that incorporates multiple WAF applications.

II. RELATED WORK

To comply to a proposed class of WSFs, a WSF should:

• separate the business logic in the web service from the hosting logic
• connect to and combine existing WAFs to a compound web service using standard HTTP requests
• introduce low overhead when added, whilst adding the incase ability to scale
• constantly adjust their infrastructures to fit the required performance at all times

Presented at IEEE Flagship Conference: IEEE International Conference on Communications 2014 in Sydney, Australia
Work Package 1

Web Scaling Frameworks: IEEE ICC Conference Paper

- $R_P$: Process Request
- $R_R$: Read Request
- $C_P$: Cache Push
- $R_S$: Subsystem Request

Diagram:
- LB (Load Balancer)
- S (Server)
- Q (Queue)
- C (Cache)
- W (Worker)
- A (App)
- DB (Database)
- PS (PubSub)

Connections:
- $R_P$, $R_R$, $C_P$, $R_S$
Work Package 1

Web Scaling Frameworks: IEEE ICC Conference Paper

Requests / Second (RPS)

Cache / Processing Ratio (CPR)

c = 320
u = 0
s = 30
d_a = 0.400
d_p = 0.012
d_s = 0.001
d_f = 0.050

BEP
CPR
S_{1,S}: RPS = 1441
S_{1,N}: RPS = 723

CPR = 0.7
d_n, min = 0.00
d_n, gain = 0.01
s = 50
d_q = 0.001
c = 6000

M\_N
M\_S
M\_Reg\_N
M\_Reg\_S

Maximum Request Flow per Second (RFPS\_max)

Total Machines (M)

\(d_{LB} = 0.0001\)
\(d_Q = 0.001\)
\(d_{PS} = 0.001\)
\(d_C = 0.001\)
\(d_S = 0.001\)
\(d_{w} = 0.5\)
\(d_{a} = 0.5\)
Project Roadmap

Work Packages & Publications

- **WP1**: Web Scaling Framework
  - April 2013
  - Literature Review
  - Report
  - Datasets
  - Charts
  - Design Documents
  - Program

- **WP2**: Post-Processing
  - November 2014
  - Design
  - Implementation
  - Literature Review
  - Report
  - Datasets
  - Charts
  - Design Documents
  - Program

- **WP3**: Thesis
  - April 2016
  - Design
  - Implementation
  - Literature Review
  - Report
  - Datasets
  - Charts
  - Design Documents
  - Program

The roadmap includes milestones such as:
- **WP1**
- **WP2**
- **WP3**
Web Package 1

Web Scaling Frameworks: IEEE TSC Journal Paper

Introducing Elastic Scalability to Web Services in the Cloud with Web Scaling Frameworks

Thomas Farhause, Student Member, IEEE, Qi Wang, Member, IEEE, Ansgar Gerlicher, Member, IEEE, Christina Grecos, Senior Member, IEEE, and XinHeng Wang, Member, IEEE

Abstract—The social web, web services have to accommodate a significant number of requests due to the high intensity of current applications. Applications have to be built in a scalable fashion so the number of machines can be adapted to highly dynamic traffic situations. In the deployment cycle, web service providers often need to focus initially on the business logic which prevents detailed scalability considerations. To overcome a critical state of customers is reached, providers need to be able to react immediately to web services to stay in business. State-of-the-Art Web Application Frameworks (WAFs) are based on the creation of application logic including data validation, view compositions and session handling. However, they often integrate low-level scalability features manually. As the creation of such scaling solutions is a very complex, we proposed in our recent work the concept of Web Scaling Frameworks (WSFs) in order to offload scaling to another layer of abstraction. WSFs are composed of traditional WAFs with multiple other components to provide scalability right from the launch of the deployment cycle. In this work, a detailed design for WSFs including necessary modules, interfaces and components is presented. A mathematical model used for performance rating is introduced. Application logic, data structures, data validation, view layer presentation and session handling are taken care of by the WAFs while the scaling logic is provided by the WSF. Fig. 1 illustrates the interplay between a WSF and multiple WAFs. To utilise an existing interface, the frameworks adapt to the WAFs while the scaling logic is provided by the WSF.

1 INTRODUCTION

The demand for modern web services increases due to the soaring social nature of the web and the upsurge of the total number of mobile devices. Web services have to deal with high levels of intensity in applications, which in turn introduces enormous amounts of requests. Static websites are replaced by dynamic and highly interactive applications. For instance, TV shows deploy apps that allow users to influence the course of the show, advertisements are not able to handle the increased load, applications have to be built in a scalable fashion so the number of machines can be adapted to highly dynamic traffic situations. Typically, web services providers need to react immediately to stay in business. Before this threshold is reached, providers need to focus on the business logic, which often prevents detailed scalability considerations.

State-of-the-Art Web Application Frameworks (WAFs) are designed to abstract common functionalities needed for the efficient implementation of web services. They focus on the creation of application logic, data structures, data validation, view layer presentation and session handling. They don’t offer integrated scaling concepts that handle the provisioning of resources, manage elastic caching or ensure guaranteed response time. Today, web service providers have to create custom-built systems that consider these scalability issues manually.

The creation of such scaling systems is a very complex task. Hence, we proposed the introduction of Web Scaling Frameworks (WSFs) in our recent work [2]. WSFs offload scaling to another layer of abstraction. They take over the responsibilities of scaling by embedding existing WAFs in a larger system. The application logic runs on the side of the WAFs while the scaling logic is provided by the WSFs. Fig. 1 illustrates the interplay between a WSF and multiple WAFs. To utilise an existing interface, the frameworks adapt to the WAFs while the scaling logic is provided by the WSF. The creation of such scaling systems is a very complex task. Hence, we proposed the introduction of Web Scaling Frameworks (WSFs) in our recent work [2]. WSFs offload scaling to another layer of abstraction. They take over the responsibilities of scaling by embedding existing WAFs in a larger system. The application logic runs on the side of the WAFs while the scaling logic is provided by the WSFs. Fig. 1 illustrates the interplay between a WSF and multiple WAFs. To utilise an existing interface, the frameworks adapt to the WAFs while the scaling logic is provided by the WSF.
Work Package 1

Web Scaling Frameworks: Pi-One Cluster
Work Package 1

Web Scaling Frameworks: IEEE TSC Journal Paper
Work Package 1

Web Scaling Frameworks: IEEE TSC Journal Paper
Work Package 1

Web Scaling Frameworks: IEEE TSC Journal Paper

[Graphs and diagrams illustrating performance comparison between hardware vs. software, Pi vs. server, and JavaScript vs. native C, with throughput vs. concurrency and evaluation metrics.]
Work Package 1

Web Scaling Frameworks: IEEE TSC Journal Paper

Measured
Optimal
Concurrency Range

RFPS of m=1..20
Normalized RFPS of m=1..20 / m

$f(x)$
$q_{f_{opt}, x}$
f_{opt, x}
$c_{low, x}$
$c_{opt, x}$
$c_{high, x}$
f_{range, x}
f_{mean, x}

$c_{x}$
Work Package 1

Web Scaling Frameworks: IEEE TSC Journal Paper

Normalized Residual Scatter and Mean Error $s=1..400kb$...
Work Package 1
Web Scaling Frameworks: IEEE TSC Journal Paper

Evaluation
Normal Version vs. Scaled Version
Web Scaling Framework

WP1

WP2
Post-Processing

WP3
Thesis

April 2013

November 2014

April 2016

Project Roadmap

Work Packages & Publications

CONFERENCE PAPER
IEEE ICC14

JOURNAL PAPER
IEEE TSC

CONFERENCE PAPER

THESIS

Literature Review

Design

Implementation

Literature Review

Report

Datasets

Charts

Design Documents

Program

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysation

Model Adaptation

Write Up

Data Collection

Analysa
Work Package 2

Reactive/Proactive Post-Processing: Conference Paper

Next focus:
Find data structures and algorithms that allow a fast, distributed and parallel processing of dependencies
Work Package 2

Reactive/Proactive Post-Processing: Conference Paper

Dependency Analysis
Data Structures
Declaration
Generation
Visualisation
Link Analysis
Work Package 2

Reactive/Proactive Post-Processing: Conference Paper

Pre/Post-Processing
Synchronous vs. Asynchronous
Optimisation
Parallelisation
Fragmentation
Eventual Cache
Work Package 3
Web Scaling Frameworks: Thesis

Thesis focus:
Create and evaluate a full-stack Web Scaling Framework with dependency declaration and optimised post-processing algorithms
Web Scaling Frameworks

PhD Transfer Event

Thomas Fankhauser
Qi Wang
Christos Grecos
Xinheng Wang
Ansgar Gerlicher