

Scalaris: Scalable Web Applications with a Transactional Key-Value Store

Nico Kruber Michael Berlin



Zuse Institut Berlin
Parallel and Distributed Systems

5th June @ Berlin Buzzwords 2012



The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreements n° 258862 (4CaaSt) and n° 257438 (Contrail).

Outline

Scalaris

Introduction

Scalability

Wiki on Scalaris

MediaWiki Introduction

Data Model

Optimisations

Strategies

Results

Scalaris - A Transactional Key-Value Store

Relational DB

- multiple relations/tables
- SQL
- Joins
- Transactions

most NoSQL stores

- key-value/document mapping
- scalable
- fault-tolerant (replication)
- eventual consistency

Scalaris

- key-value pairs
- scalable
- **transaction support**
- select, update/insert
- fault-tolerant (replication)
- **strong consistency**



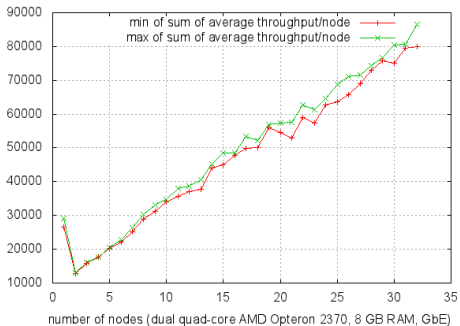
Scalaris Implementation

- Scalaris: ~23 000 lines of Erlang code
 - 2 200 for transactions
- APIs:
 - Java library (11 500 lines)
 - Python library (1 400 lines)
 - Ruby library (700 lines)
 - HTTP-JSON interface
- <http://scalis.googlecode.com>

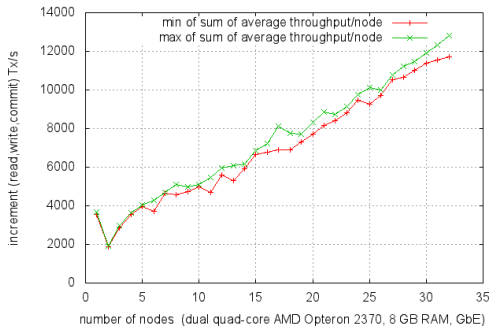


Scalaris - Scalability

Quorum-read benchmark on cumulus
(10 clients/node, 10000 quorum-reads/key)



Increment benchmark on cumulus
(10 clients/node, 1000 incs/key)





Wikitext A-Z



Page [Talk](#) [Read](#) [Change](#) [View history](#)

Berlin

You are not [logged in](#). Your [IP address](#) will be recorded in this page's [edit history](#).

```
'''Berlin''' is the [[Capital of
Germany|capital]] city of [[Germany]] and
is one of the 16 [[states of Germany]].
With a population of 3.45&nbsp;million
people, Berlin is Germany's largest city.
It is the second [[Largest cities of the
European Union by population within city
limits|most populous city proper]] and the
seventh [[Largest urban areas of the
European Union|most populous urban area]]
in the [[European Union]].
```

```
<!-- TEMPLATES -->
{{Navboxes
|list=
{{Berlin}}
{{Cities in Germany}}
}}
```

```
<!-- CATEGORIES -->
[[Category:Berlin| ]]
[[Category:States of Germany]]
[[Category:German state capitals]]
[[Category:Capitals in Europe]]
```

[Edit summary](#)

[Save page](#) [Show preview](#) [Show changes](#) [Cancel](#) [Editing help](#) (opens in new window)

[Main Page](#)
[New changes](#)
[Show any entry](#)
[Help](#)

[Toolbox](#)
[What links here](#)
[Special pages](#)





Wikitext A-Z



Page [Talk](#) [Read](#) [Change](#) [View history](#)

Berlin

page title including namespace

You are not logged in. Your IP address will be recorded in this page's [edit history](#).

```
'''Berlin''' is the [[Capital of
Germany|capital]] city of [[Germany]] and
is one of the 16 [[states of Germany]].
With a population of 3.45&nbsp;million
people, Berlin is Germany's largest city.
It is the second [[Largest cities of the
European Union by population within city
limits|most populous city proper]] and the
seventh [[Largest urban areas of the
European Union|most populous urban area]]
in the [[European Union]].
```

```
<!-- TEMPLATES -->
{{Navboxes
|list=
{{Berlin}}
{{Cities in Germany}}
}}
```

```
<!-- CATEGORIES -->
[[Category:Berlin| ]]
[[Category:States of Germany]]
[[Category:German state capitals]]
[[Category:Capitals in Europe]]
```

[Edit summary](#)

[Save page](#) [Show preview](#) [Show changes](#) [Cancel](#) [Editing help](#) (opens in new





Wikitext A-Z



Page [Talk](#) [Read](#) [Change](#) [View history](#)

Berlin

page title including namespace

You are not logged in. Your IP address will be recorded in this page's [edit history](#).

'''Berlin''' is the [\[\[Capital of Germany|capital\]\]](#) city of [\[\[Germany\]\]](#) and is one of the 16 [\[\[states of Germany\]\]](#). With a population of 3.45 million people, Berlin is Germany's largest city. It is the second [\[\[Largest cities of the European Union by population within city limits|most populous city proper\]\]](#) and the seventh [\[\[Largest urban areas of the European Union|most populous urban area\]\]](#) in the [\[\[European Union\]\]](#).

link to another page

```
<!-- TEMPLATES -->
{{Navboxes
|list=
{{Berlin}}
{{Cities in Germany}}
}}
```

```
<!-- CATEGORIES -->
[[Category:Berlin| ]]
[[Category:States of Germany]]
[[Category:German state capitals]]
[[Category:Capitals in Europe]]
```

[Edit summary](#)

[Save page](#) [Show preview](#) [Show changes](#) [Cancel](#) [Editing help](#) (opens in new window)





Wikitext A-Z



Page Talk Read Change View history Search

Berlin

page title including namespace

You are not logged in. Your IP address will be recorded in this page's [edit history](#).

'''Berlin''' is the [\[\[Capital of Germany|capital\]\]](#) city of [\[\[Germany\]\]](#) and is one of the 16 [\[\[states of Germany\]\]](#). With a population of 3.45 million people, Berlin is Germany's largest city. It is the second [\[\[Largest cities of the European Union by population within city limits|most populous city proper\]\]](#) and the seventh [\[\[Largest urban areas of the European Union|most populous urban area\]\]](#) in the [\[\[European Union\]\]](#).

link to another page

<!-- TEMPLATES -->

```
{ {Navboxes
|list=
{{Berlin}}
{{Cities in Germany}}
}}
```

template inclusion

<!-- CATEGORIES -->

```
[[Category:Berlin| ]]
[[Category:States of Germany]]
[[Category:German state capitals]]
[[Category:Capitals in Europe]]
```

Edit summary

Save page Show preview Show changes Cancel Editing help (opens in new





Wikitext A-Z



Page Talk Read Change View history Search

Berlin

page title including namespace

You are not logged in. Your IP address will be recorded in this page's [edit history](#).

'''Berlin''' is the [\[\[Capital of Germany|capital\]\]](#) city of [\[\[Germany\]\]](#) and is one of the 16 [\[\[states of Germany\]\]](#). With a population of 3.45 million people, Berlin is Germany's largest city. It is the second [\[\[Largest cities of the European Union by population within city limits|most populous city proper\]\]](#) and the seventh [\[\[Largest urban areas of the European Union|most populous urban area\]\]](#) in the [\[\[European Union\]\]](#).

link to another page

<!-- TEMPLATES -->

```

{{Navboxes
|list=
{{Berlin}}
{{Cities in Germany}}
}}
```

template inclusion

<!-- CATEGORIES -->

```

[[Category:Berlin| ]]
[[Category:States of Germany]]
[[Category:German state capitals]]
[[Category:Capitals in Europe]]
```

category the page belongs to

Edit summary

Save page Show preview Show changes Cancel Editing help (opens in new





Wiki Operations



Page [Talk](#) [Read](#) [Change](#) [View history](#)

Berlin

You are not [logged in](#). Your [IP address](#) will be recorded in this page's [edit history](#).

```
'''Berlin''' is the [[Capital of
Germany|capital]] city of [[Germany]] and
is one of the 16 [[states of Germany]].
With a population of 3.45&nbsp;million
people, Berlin is Germany's largest city.
It is the second [[Largest cities of the
European Union by population within city
limits|most populous city proper]] and the
seventh [[Largest urban areas of the
European Union|most populous urban area]]
in the [[European Union]].
```

```
<!-- TEMPLATES -->
{{Navboxes
|list=
{{Berlin}}
{{Cities in Germany}}
}}
```

```
<!-- CATEGORIES -->
[[Category:Berlin| ]]
[[Category:States of Germany]]
[[Category:German state capitals]]
[[Category:Capitals in Europe]]
```

[Edit summary](#)

[Editing help](#) (opens in new

[Main Page](#)
[New changes](#)
[Show any entry](#)
[Help](#)

[Toolbox](#)
[What links here](#)
[Special pages](#)





Wiki Operations

Page Talk **Read** Change View history Search

Berlin

You are not logged in. Your IP address will be recorded in this page's [edit history](#).

'Berlin' is the [\[\[Capital of Germany|capital\]\]](#) city of [\[\[Germany\]\]](#) and is one of the 16 [\[\[states of Germany\]\]](#). With a population of 3.45 million people, Berlin is Germany's largest city. It is the second [\[\[Largest cities of the European Union by population within city limits|most populous city proper\]\]](#) and the seventh [\[\[Largest urban areas of the European Union|most populous urban area\]\]](#) in the [\[\[European Union\]\]](#).

```
<!-- TEMPLATES -->
{{Navboxes
|list=
{{Berlin}}
{{Cities in Germany}}
}}
```

```
<!-- CATEGORIES -->
[[Category:Berlin| ]]
[[Category:States of Germany]]
[[Category:German state capitals]]
[[Category:Capitals in Europe]]
```

[Edit summary](#) [Save page](#) [Show preview](#) [Show changes](#) [Cancel](#) [Editing help \(opens in new window\)](#)

show page: read revisions



Wiki Operations

Page Talk Read Change View history Search

Berlin

You are not logged in. Your IP address will be recorded in this page's edit history.

'Berlin' is the [[Capital of Germany|capital]] city of [[Germany]] and is one of the 16 [[states of Germany]]. With a population of 3.45 million people, Berlin is Germany's largest city. It is the second [[Largest cities of the European Union by population within city limits|most populous city proper]] and the seventh [[Largest urban areas of the European Union|most populous urban area]] in the [[European Union]].

```
<!-- TEMPLATES -->
{{Navboxes
|list=
{{Berlin}}
{{Cities in Germany}}
}}
```

```
<!-- CATEGORIES -->
[[Category:Berlin| ]]
[[Category:States of Germany]]
[[Category:German state capitals]]
[[Category:Capitals in Europe]]
```

Edit summary


Save page Show preview Show changes Cancel Editing help (opens in new)

revision history overview

show page: read revisions



Wiki Operations



Scalaris

Main Page
New changes
Show any entry
Help

Toolbox
What links here
Special pages

Edit summary

Save page Show preview Show changes Cancel | [Editing help](#) (opens in new window)

Page Talk Read Change View history Search

Berlin

You are not logged in. Your IP address will be recorded in this page's edit history.

'Berlin' is the [[Capital of Germany|capital]] city of [[Germany]] and is one of the 16 [[states of Germany]]. With a population of 3.45 million people, Berlin is Germany's largest city. It is the second [[Largest cities of the European Union by population within city limits|most populous city proper]] and the seventh [[Largest urban areas of the European Union|most populous urban area]] in the [[European Union]].

```
<!-- TEMPLATES -->
{{Navboxes
|list=
{{Berlin}}
{{Cities in Germany}}
}}

<!-- CATEGORIES -->
[[Category:Berlin| ]]
[[Category:States of Germany]]
[[Category:German state capitals]]
[[Category:Capitals in Europe]]
```

search for articles

revision history overview

show page: read revisions



Wiki Operations

Page Talk Read Change View history Search

Scalaris

Berlin

You are not logged in. Your IP address will be recorded in this page's edit history.

'Berlin' is the [[Capital of Germany|capital]] city of [[Germany]] and is one of the 16 [[states of Germany]]. With a population of 3.45 million people, Berlin is Germany's largest city. It is the second [[Largest cities of the European Union by population within city limits|most populous city proper]] and the seventh [[Largest urban areas of the European Union|most populous urban area]] in the [[European Union]].

```
<!-- TEMPLATES -->
{{Navboxes
|list=
{{Berlin}}
{{Cities in Germany}}
}}
```

```
<!-- CATEGORIES -->
[[Category:Berlin| ]]
[[Category:States of Germany]]
[[Category:German state capitals]]
[[Category:Capitals in Europe]]
```

Edit summary

Save page Show preview Show changes Cancel | Editing help (opens in new)

Main Page
New changes
Show any entry
Help

Toolbox
What links here
Special pages

search for articles

revision history overview

show page: read revisions

redirect to random article





Wiki Operations

The screenshot shows a Wiki page for 'Berlin' with several UI elements highlighted by red circles and lines:

- Page** and **Talk** buttons are circled in red.
- Read** and **Change** buttons are circled in red.
- View history** and **Search** buttons are circled in red.
- Show any entry** and **What links here** buttons in the sidebar are circled in red.
- Red lines connect these buttons to their corresponding functions on the page:
 - 'Page' points to the article title 'Berlin'.
 - 'Talk' points to the 'Talk' link below the title.
 - 'Read' points to the main text of the article.
 - 'Change' points to the 'Edit' link below the title.
 - 'View history' points to the 'View history' link below the title.
 - 'Search' points to the search input field.
 - 'Show any entry' points to the 'Show any entry' button in the sidebar.
 - 'What links here' points to the 'What links here' button in the sidebar.

The main content of the page is as follows:

You are not logged in. Your IP address will be recorded in this page's edit history.

'Berlin' is the [[Capital of Germany|capital]] city of [[Germany]] and is one of the 16 [[states of Germany]]. With a population of 3.45 million people, Berlin is Germany's largest city. It is the second [[Largest cities of the European Union by population within city limits|most populous city proper]] and the seventh [[Largest urban areas of the European Union|most populous urban area]] in the [[European Union]].

```
<!-- TEMPLATES -->
{{Navboxes
|list=
{{Berlin}}
{{Cities in Germany}}
}}
```

```
<!-- CATEGORIES -->
[[Category:Berlin| ]]
[[Category:States of Germany]]
[[Category:German state capitals]]
[[Category:Capitals in Europe]]
```

At the bottom, there is an 'Edit summary' field and buttons for 'Save page', 'Show preview', 'Show changes', and 'Cancel | Editing help (opens in new window)'.

search for articles
 revision history overview
 show page: read revisions
 redirect to random article
 backlinks (“what links here”)



Wiki Operations

The screenshot shows a Wiki page for 'Berlin' with several red annotations:

- Page** and **Talk** buttons are circled in red.
- Read** and **Change** buttons are circled in red.
- View history** and **Search** buttons are circled in red.
- Show any entry** and **What links here** buttons in the sidebar are circled in red.
- The **Save page** button at the bottom is circled in red.

The main content area shows the text of the article and the wikitext source code below it:

```
<!-- TEMPLATES -->
{{Navboxes
|list=
{{Berlin}}
{{Cities in Germany}}
}}

<!-- CATEGORIES -->
[[Category:Berlin| ]]
[[Category:States of Germany]]
[[Category:German state capitals]]
[[Category:Capitals in Europe]]
```

search for articles

revision history overview

show page: read revisions


redirect to random article

backlinks (“what links here”)

save new revision



Editing Pages



[Main Page](#)
[New changes](#)
[Show any entry](#)
[Help](#)

Toolbox
[What links here](#)
[Special pages](#)

Page [Talk](#) [Read](#) [Change](#) [View history](#)

Berlin

You are not logged in. Your IP address will be recorded in this page's [edit history](#).

```

'''Berlin''' is the [[Capital of Germany|capital]] city of [[Germany]] and is one of the 16 [[states of Germany]]. With a population of 3.45&nbsp;million people, Berlin is Germany's largest city. It is the second [[Largest cities of the European Union by population within city limits|most populous city proper]] and the seventh [[Largest urban areas of the European Union|most populous urban area]] in the [[European Union]].

<!-- TEMPLATES -->
{{Navboxes
|list=
{{Berlin}}
{{Cities in Germany}}
}}

<!-- CATEGORIES -->
[[Category:Berlin| ]]
[[Category:States of Germany]]
[[Category:German state capitals]]
[[Category:Capitals in Europe]]
    
```

[Edit summary](#)

[Editing help \(opens in new window\)](#)

Do atomically:

1. store new revision
2. update page object
3. add new backlinks
4. remove old backlinks
5. add to new categories
6. remove from old categories
7. if new page:
 - a) add to page list
 - b) increase page counter





Editing Pages



Page Talk Read Change View history Search

Berlin

You are not logged in. Your IP address will be recorded in this page's [edit history](#).

'''Berlin''' is the [[Capital of Germany|capital]] city of [[Germany]] and is one of the 16 [[states of Germany]]. With a population of 3.45 million people, Berlin is Germany's largest city. It is the second [[Largest cities of the European Union by population within city limits|most populous city proper]] and the seventh [[Largest urban areas of the European Union|most populous urban area]] in the [[European Union]].

```
<!-- TEMPLATES -->
{{Navboxes
|list=
{{Berlin}}
{{Cities in Germany}}
}}
```

```
<!-- CATEGORIES -->
[[Category:Berlin| ]]
[[Category:States of Germany]]
[[Category:German state capitals]]
[[Category:Capitals in Europe]]
```

Edit summary

Do atomically:

1. store new revision
2. update page object
3. add new backlinks
4. remove old backlinks
5. add to new categories
6. remove from old categories
7. if new page:
 - a) add to page list
 - b) increase page counter





Editing Pages



Page [Talk](#) [Read](#) [Change](#) [View history](#)

Berlin

You are not logged in. Your IP address will be recorded in this page's [edit history](#).

```
'''Berlin''' is the [[Capital of
Germany|capital]] city of [[Germany]] and
is one of the 16 [[states of Germany]].
With a population of 3.45&nbsp;million
people, Berlin is Germany's largest city.
It is the second [[Largest cities of the
European Union by population within city
limits|most populous city proper]] and the
seventh [[Largest urban areas of the
European Union|most populous urban area]]
in the [[European Union]].
```

```
<!-- TEMPLATES -->
{{Navboxes
|list=
{{Berlin}}
{{Cities in Germany}}
}}
```

```
<!-- CATEGORIES -->
[[Category:Berlin| ]]
[[Category:States of Germany]]
[[Category:German state capitals]]
[[Category:Capitals in Europe]]
```

[Edit summary](#)

[Save page](#)

[Show preview](#)

[Show changes](#)

[Cancel](#) | [Editing help](#) (opens in new

Do atomically:

1. store new revision
2. update page object
3. add new backlinks
4. remove old backlinks
5. add to new categories
6. remove from old categories
7. if new page:
 - a) add to page list
 - b) increase page counter





Editing Pages



Page [Talk](#) [Read](#) [Change](#) [View history](#)

Berlin

You are not [logged in](#). Your [IP address](#) will be recorded in this page's [edit history](#).

```
'''Berlin''' is the [[Capital of
Germany|capital]] city of [[Germany]] and
is one of the 16 [[states of Germany]].
With a population of 3.45&nbsp;million
people, Berlin is Germany's largest city.
It is the second [[Largest cities of the
European Union by population within city
limits|most populous city proper]] and the
seventh [[Largest urban areas of the
European Union|most populous urban area]]
in the [[European Union]].
```

```
<!-- TEMPLATES -->
{{Navboxes
|list=
{{Berlin}}
{{Cities in Germany}}
}}
```

```
<!-- CATEGORIES -->
[[Category:Berlin| ]]
[[Category:States of Germany]]
[[Category:German state capitals]]
[[Category:Capitals in Europe]]
```

[Edit summary](#)

[Save page](#) [Show preview](#) [Show changes](#) [Cancel](#) | [Editing help](#) (opens in new

Do atomically:

1. store new revision
2. update page object
3. add new backlinks
4. remove old backlinks
5. add to new categories
6. remove from old categories
7. if new page:
 - a) add to page list
 - b) increase page counter





Editing Pages



Page [Talk](#) [Read](#) [Change](#) [View history](#)

Berlin

You are not [logged in](#). Your [IP address](#) will be recorded in this page's [edit history](#).

```
'''Berlin''' is the [[Capital of
Germany|capital]] city of [[Germany]] and
is one of the 16 [[states of Germany]].
With a population of 3.45&nbsp;million
people, Berlin is Germany's largest city.
It is the second [[Largest cities of the
European Union by population within city
limits|most populous city proper]] and the
seventh [[Largest urban areas of the
European Union|most populous urban area]]
in the [[European Union]].
```

```
<!-- TEMPLATES -->
{{Navboxes
|list=
{{Berlin}}
{{Cities in Germany}}
}}
```

```
<!-- CATEGORIES -->
[[Category:Berlin| ]]
[[Category:States of Germany]]
[[Category:German state capitals]]
[[Category:Capitals in Europe]]
```

[Edit summary](#)

[Save page](#) [Show preview](#) [Show changes](#) [Cancel](#) | [Editing help](#) (opens in new window)

Do atomically:

1. store new revision
2. update page object
3. add new backlinks
4. remove old backlinks
5. add to new categories
6. remove from old categories
7. if new page:
 - a) add to page list
 - b) increase page counter





Editing Pages



Page [Talk](#) [Read](#) [Change](#) [View history](#)

Berlin

You are not [logged in](#). Your [IP address](#) will be recorded in this page's [edit history](#).

```
'''Berlin''' is the [[Capital of
Germany|capital]] city of [[Germany]] and
is one of the 16 [[states of Germany]].
With a population of 3.45&nbsp;million
people, Berlin is Germany's largest city.
It is the second [[Largest cities of the
European Union by population within city
limits|most populous city proper]] and the
seventh [[Largest urban areas of the
European Union|most populous urban area]]
in the [[European Union]].
```

```
<!-- TEMPLATES -->
{{Navboxes
|list=
{{Berlin}}
{{Cities in Germany}}
}}
```

```
<!-- CATEGORIES -->
[[Category:Berlin| ]]
[[Category:States of Germany]]
[[Category:German state capitals]]
[[Category:Capitals in Europe]]
```

[Edit summary](#)

[Save page](#) [Show preview](#) [Show changes](#) [Cancel](#) [Editing help](#) (opens in new window)

Do atomically:

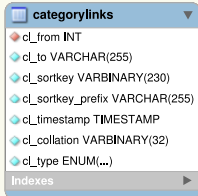
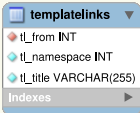
1. store new revision
2. update page object
3. add new backlinks
4. remove old backlinks
5. add to new categories
6. remove from old categories
7. if new page:
 - a) add to page list
 - b) increase page counter





MediaWiki Schema

Page text and associated information



Statistics and logging



Page containing the templatelink
tl_from - page_id

pl_from - page_id

Page defined as category member

rev_text_id
- old_id_id



adapted from https://commons.wikimedia.org/wiki/File:MediaWiki_database_schema_1-19_%28r102798%29.svg



Wiki on Scalaris - Data Layout

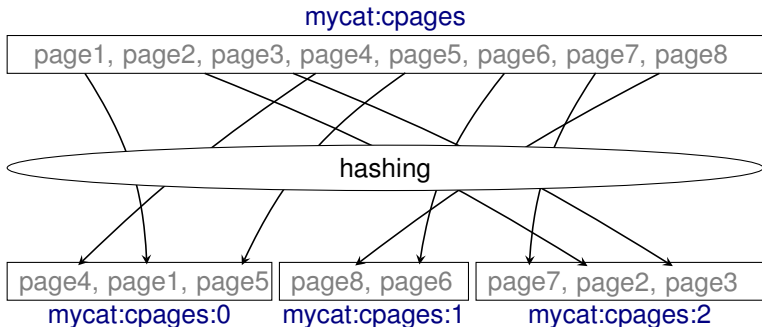
key	content
pages:<namespace>	list of pages in <namespace>
pages:<namespace>:count	number of pages in <namespace>
<title>:page	page information (JSON object)
<title>:rev:<id>	single page revision (JSON object)
<title>:revs	list of revision info for history
<title>:blpages	list of page titles linking to <title>
<cat_title>:cpages	list of page titles inside a category
<cat_title>:cpages:count	number of pages inside a category

Data Model - Issues?

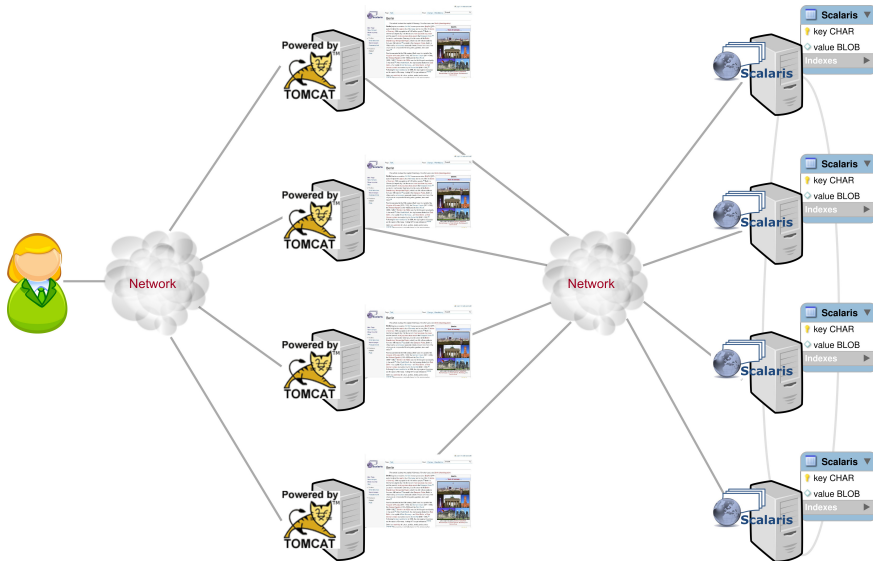
- non-scalable data model!
- Problem: page lists, counters

Data Model - Solution: Partitioning!

- a) spread randomly over partitions
 - suitable for e.g. counters
- b) spread list contents with hashing
 - allows efficient lookup and delete
 - suitable for e.g. lists:

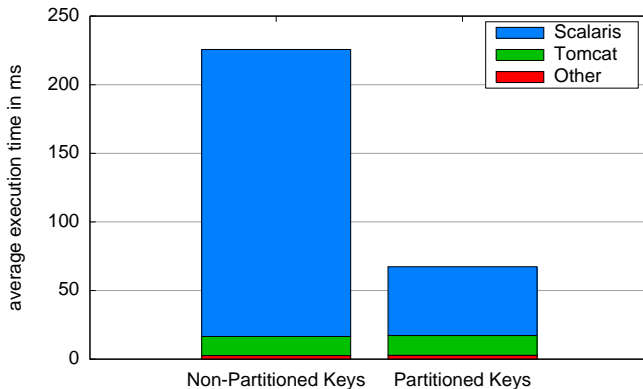


Wiki on Scalaris - Architecture



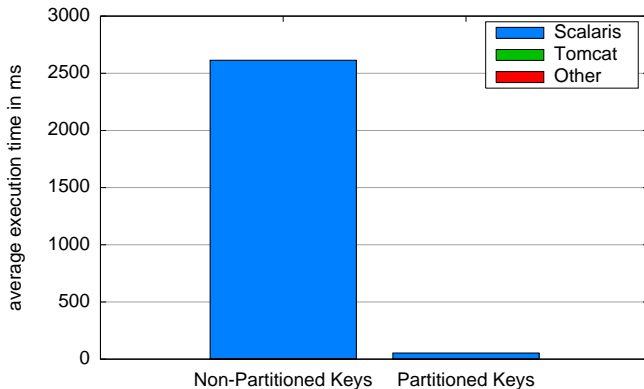
Wiki on Scalaris - Optimisation Results (I)

- effects on page edits:

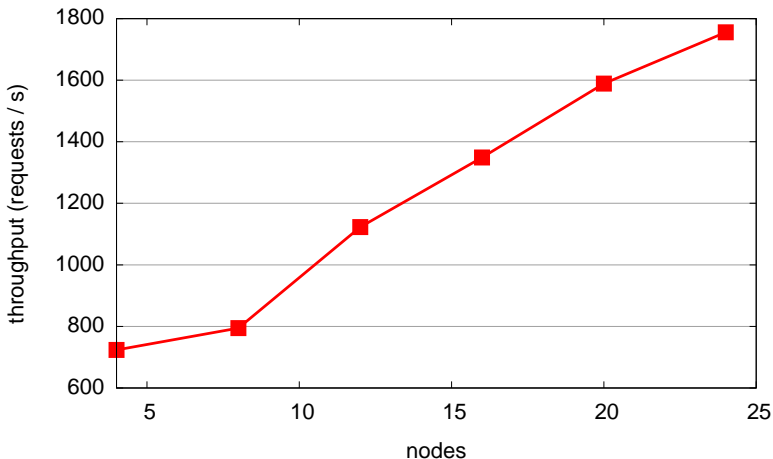


Wiki on Scalaris - Optimisation Results (II)

- effects on page creates:



Wiki on Scalaris - Scalability



Questions



<http://scalaris.googlecode.com>