

**Internet Briefing. Real-Time Web. Technische
Standards und Herausforderungen. Namics.**

Jürg Stuker. CEO. Partner.

7. April 2010

**“The sum of the expertise
of the people in the
audience is greater than
the sum of expertise of
the person on stage.”**

Quelle: <http://scripting.wordpress.com/2006/03/05/what-is-an-unconference/>

Nutrition Facts

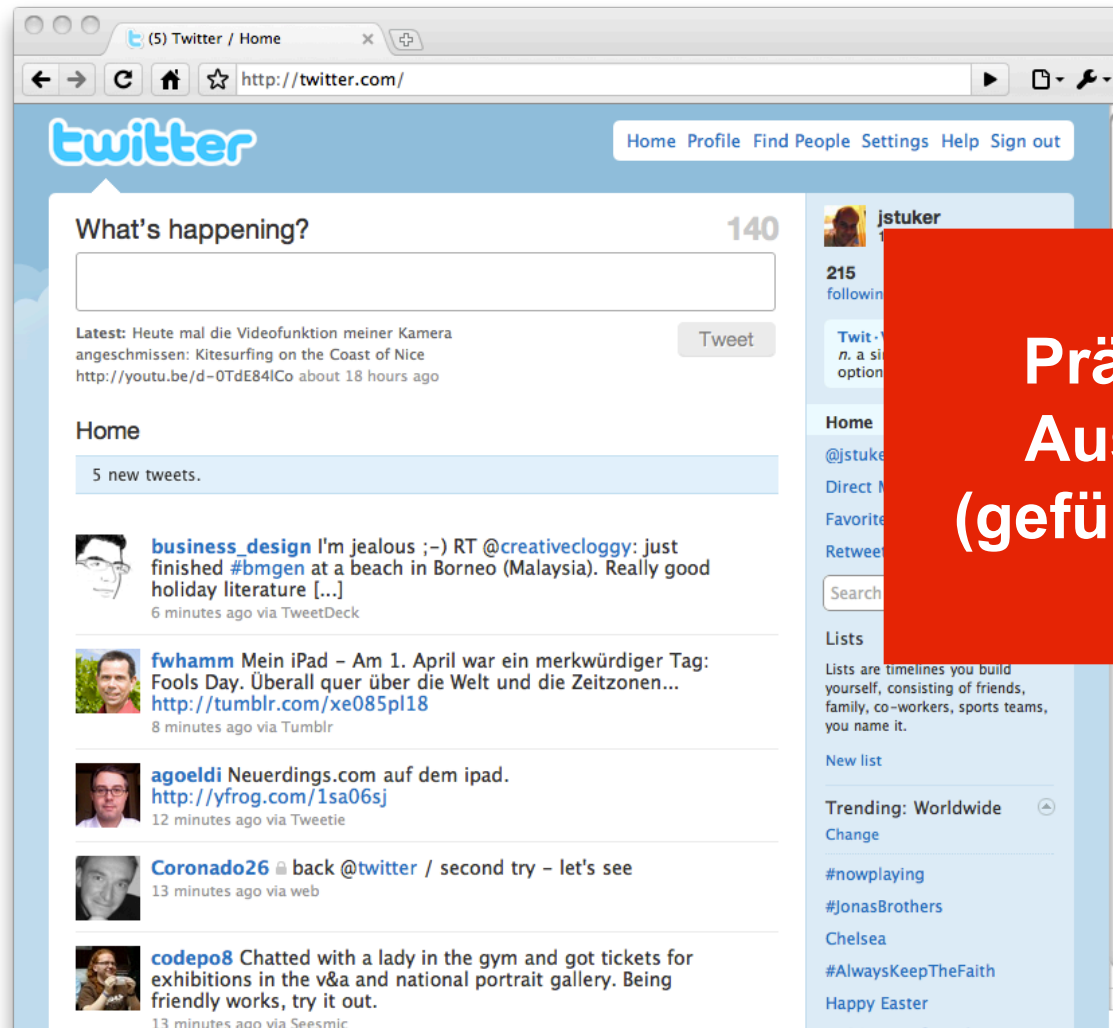
Serving Size less than 1 Hour (1 RU = Reto Unit)

	% of Hourly Value
Answers	0%
Effects	10%
Principles and Tipps	40%
Your Experience	50%

No significant source of Backend.

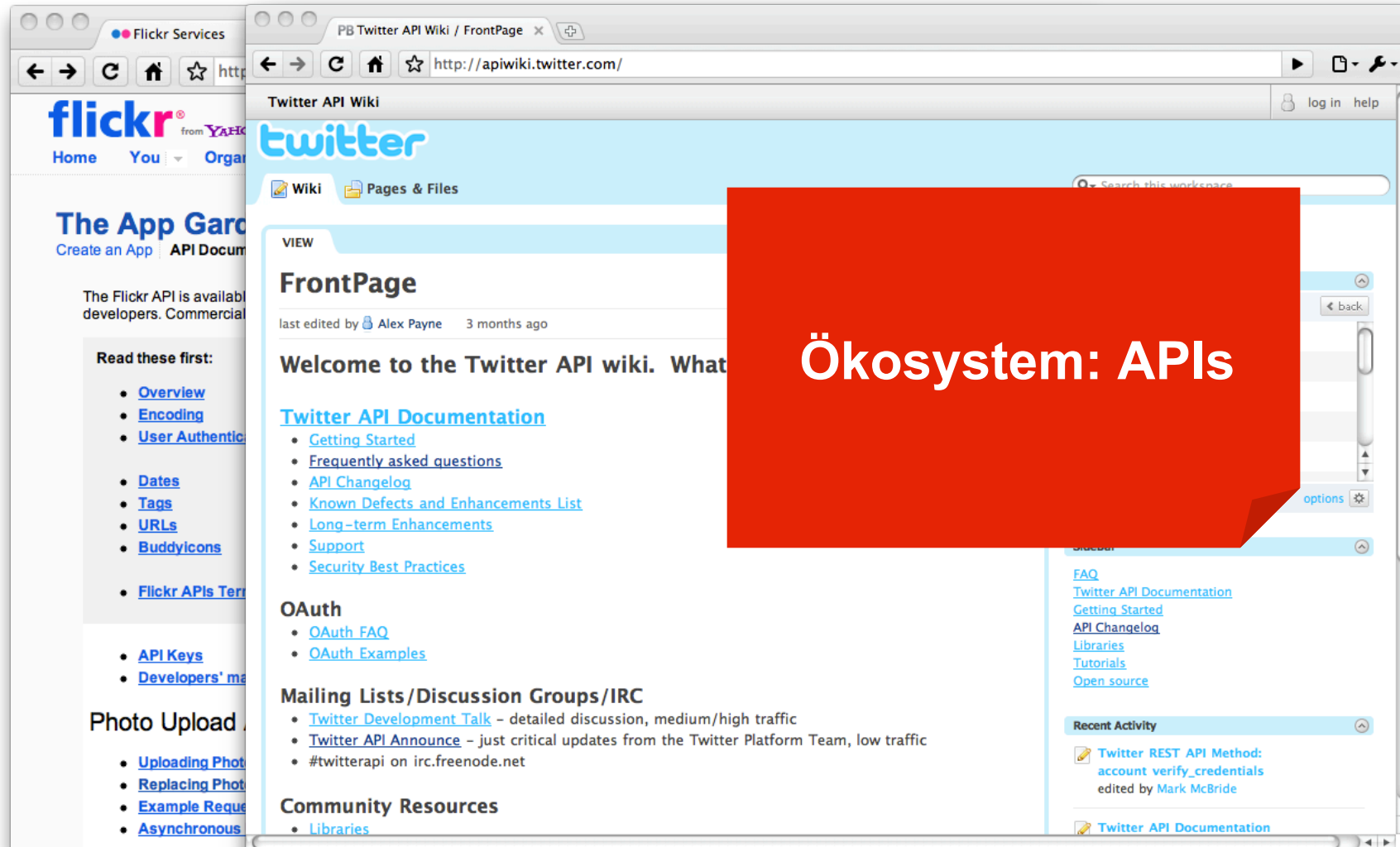
* Percent Daily Values are based on a *light* tech diet.

Was ist neu?



Präsentation /
Austausch ist
(gefühl) synchron

Was ist neu?

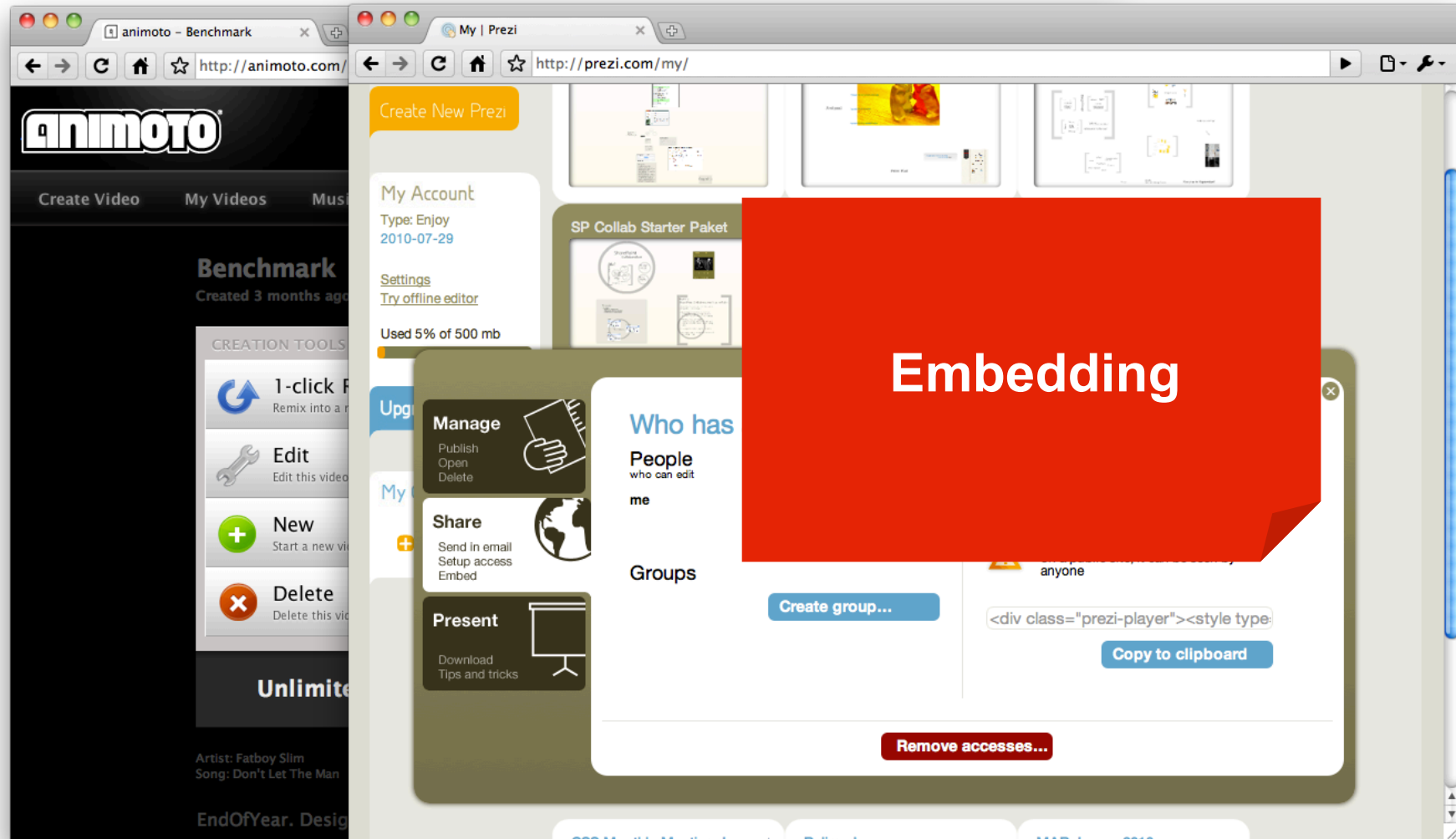


Was ist neu?

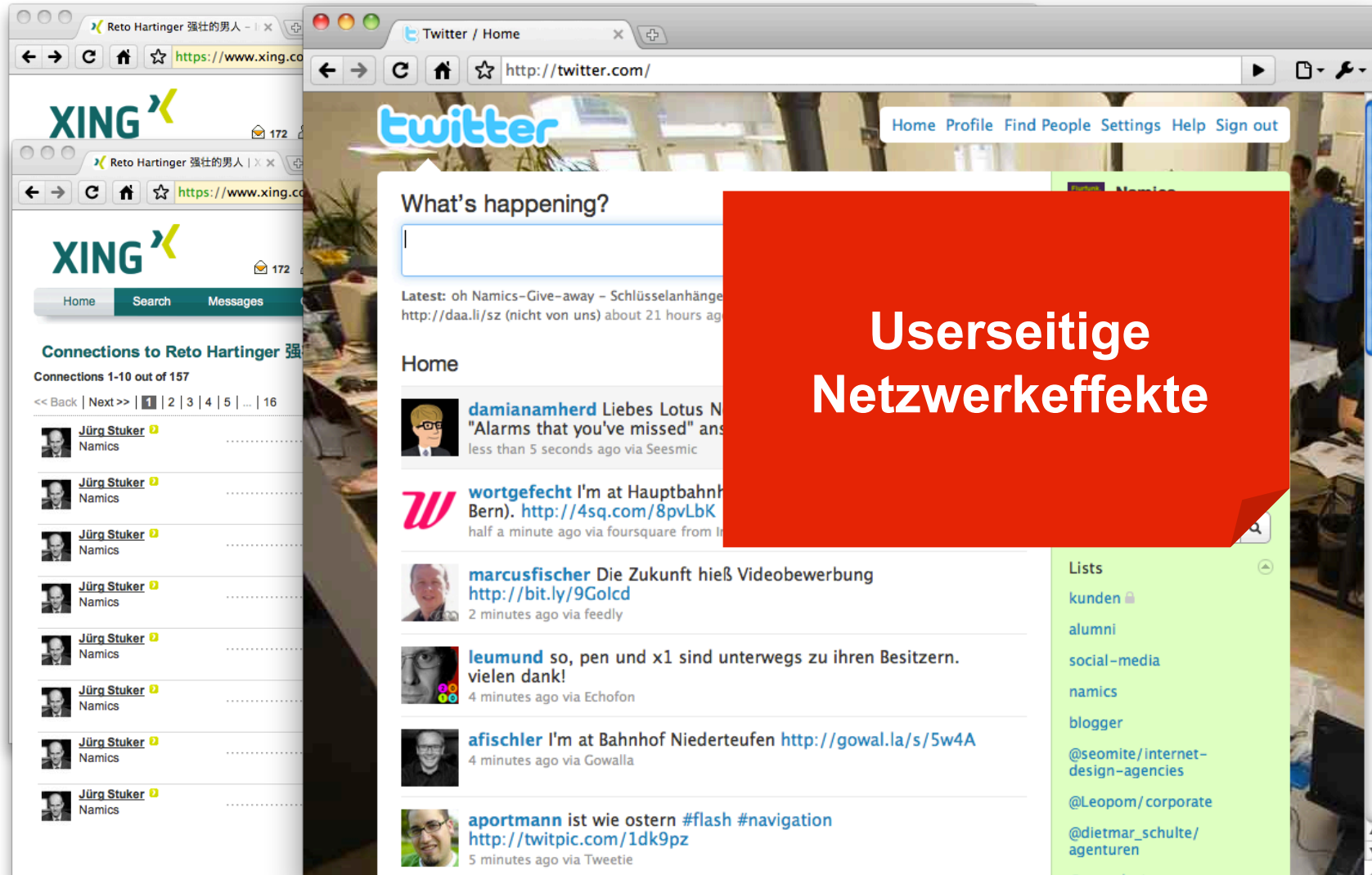
The image shows two screenshots of application authorization windows. The top screenshot is a Twitter dialog for an application named 'Yfrog' by ImageShack. It asks for permission to access and update user data. Below the text are 'Deny' and 'Allow' buttons. The bottom screenshot is a detailed 'Allow Access?' dialog for 'Akazoo Backgammon'. It explains that access will allow the application to pull profile information, photos, and friends' info. It includes a small image of dice, a 4.5-star rating, and a description of the game. At the bottom, there are 'Allow' and 'Leave Application' buttons.

Ökosystem: Anwendungs-integration

Was ist neu?

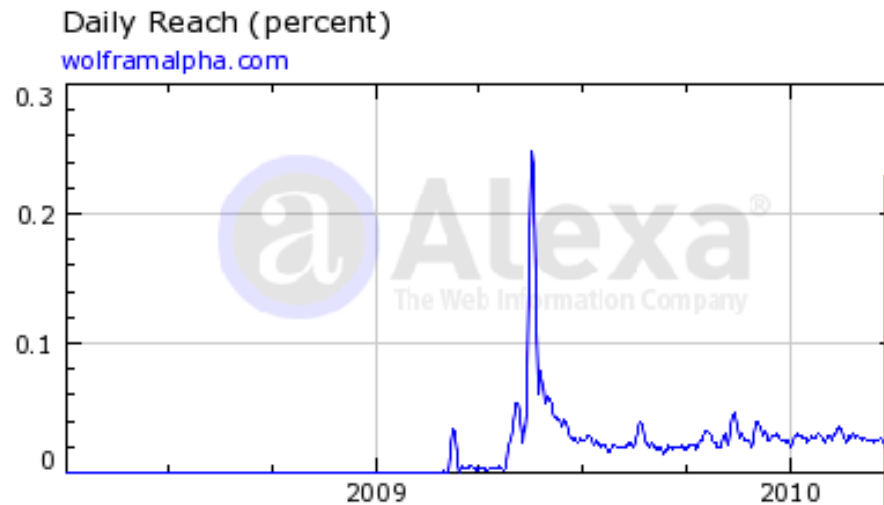


Was ist neu?



Real-Time Web. Technische Standards. Herausforderungen. **Namics.**

Was ist neu?



Animoto and Amazon EC2
Amazon EC2 easily scaled

**Netzwerkeffekte
erzeugen Lastspitzen**



Was ist neu?



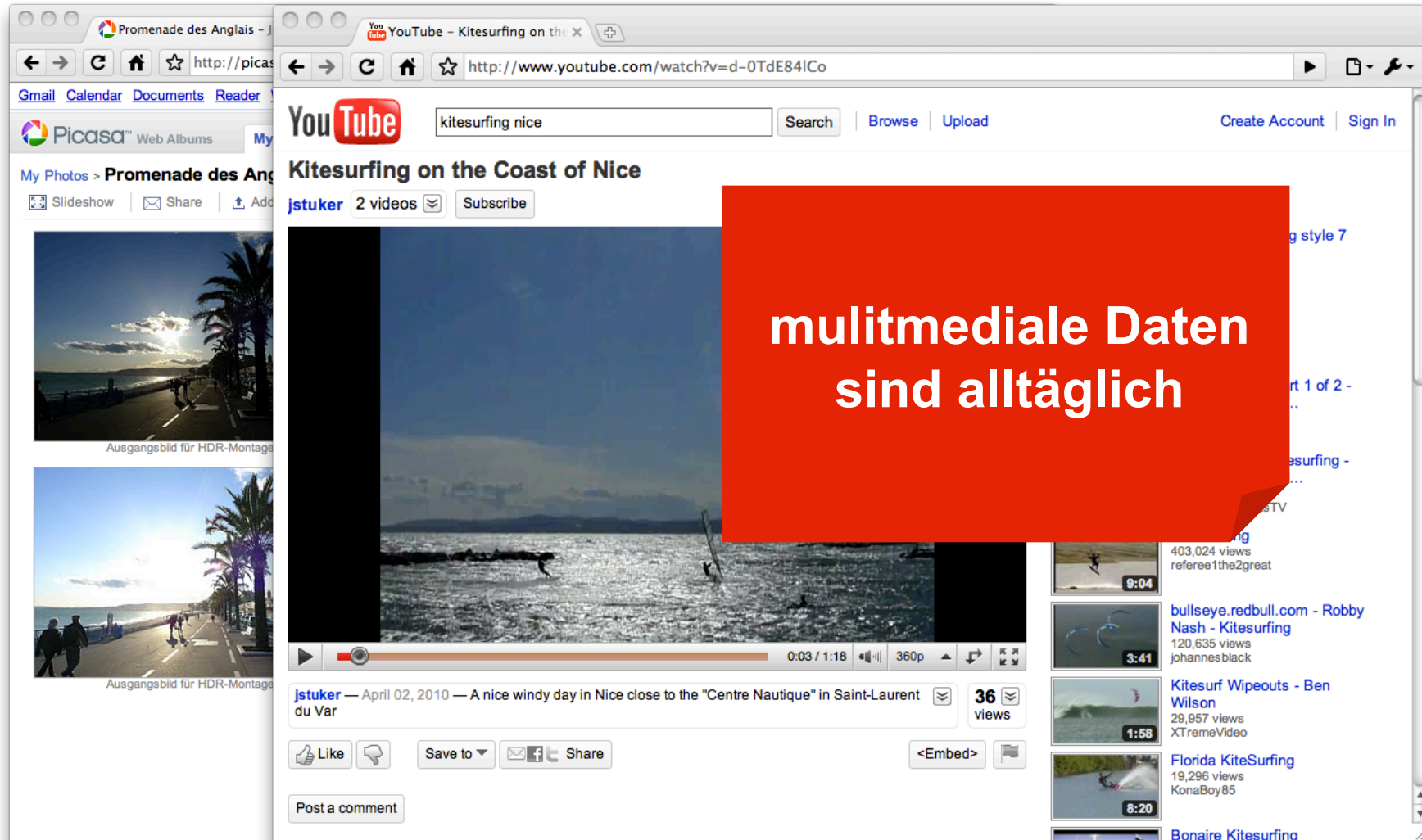
Real-Time Web. Technische Standards. Herausforderungen. **Namics.**

Was ist neu?

The image shows a screenshot of a social media feed with three posts. The top post is by Bernhard Seefeld, posted at 7:10 pm, with the text "Ipad line just 5min. Barely long enough to get a free bottle of hipster water." The middle post is by Jan Stettler, dated Mar 31, with the text "ZbW" and two photos of a hallway and a field. The bottom post is by Dorian Selz, with the text "Dem Himmel ganz nah - Das neue Büro von @". A red callout box with white text is overlaid on the middle and bottom posts, reading "Aggregation (und Republikation / Autopost) von Daten".

Aggregation (und Republikation / Autopost) von Daten

Was ist neu?



Real-Time Web. Technische Standards. Herausforderungen. **Namics.**

Was ist neu?

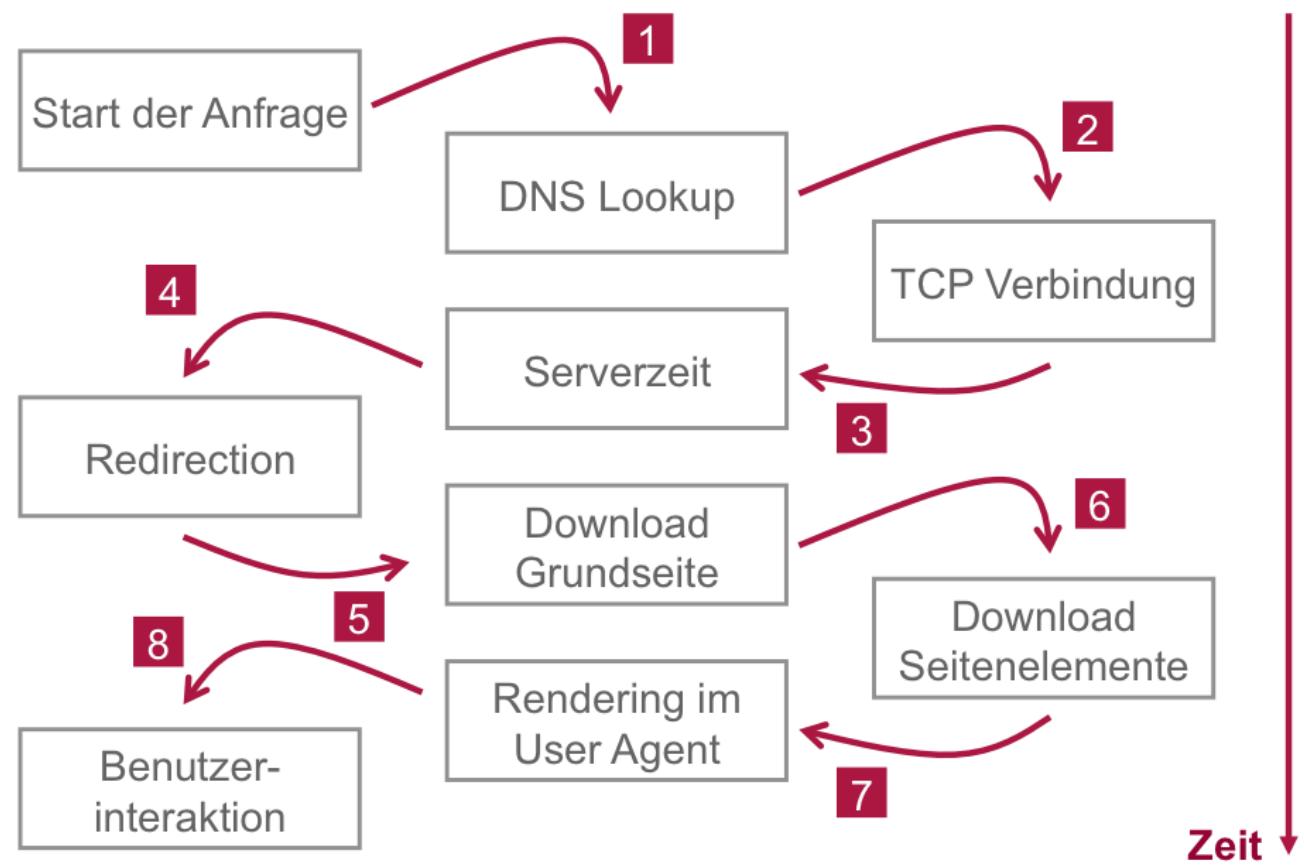


Technische Implikationen?

- **Präsentation / Austausch ist (gefühl) synchron**
- **Ökosystem**
 - APIs
 - Anwendungs-integration
- **Embedding**
- **Userseitige Netzwerkeffekte**
- **Netzwerkeffekte erzeugen Lastspitzen**
- **multimediale Daten sind alltäglich**
- **Aggregation (und Republikation / Autopost) von Daten**
- **Zeit als primäres Rangierungsprinzip**

Client Performance.

Ein Webrequest (Detail, Zeitsicht)

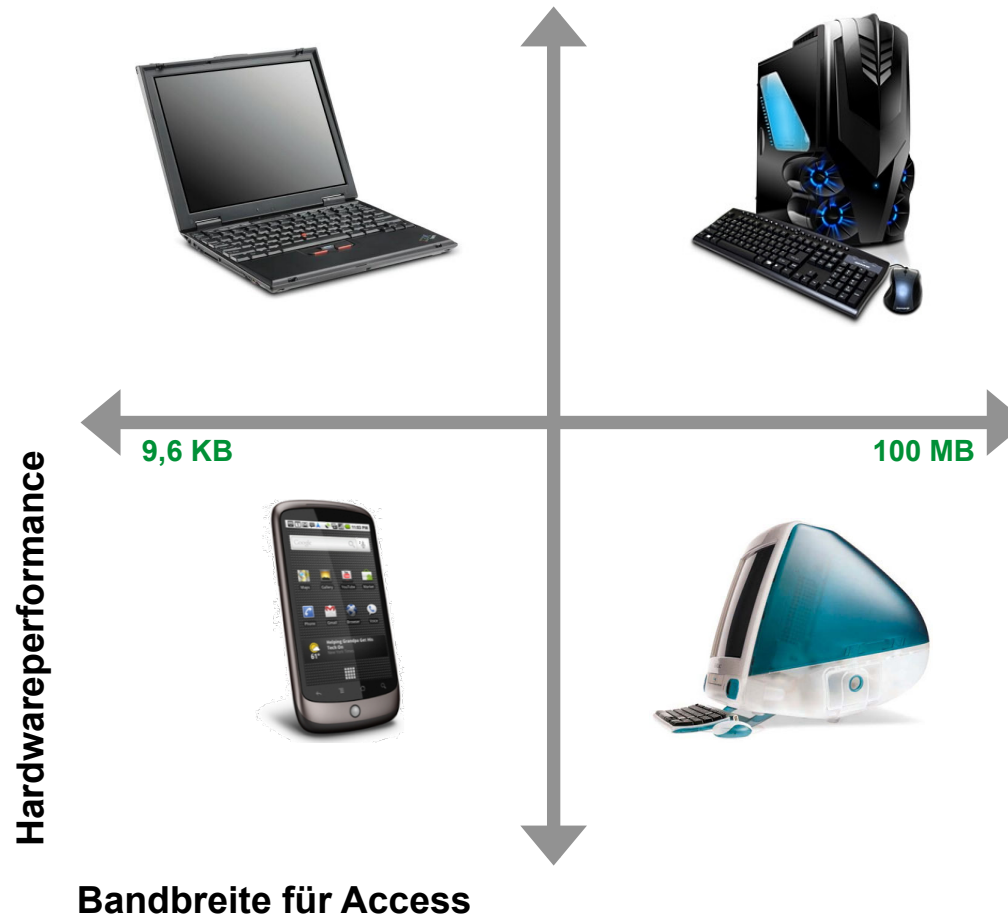


Perform
IEX 2001

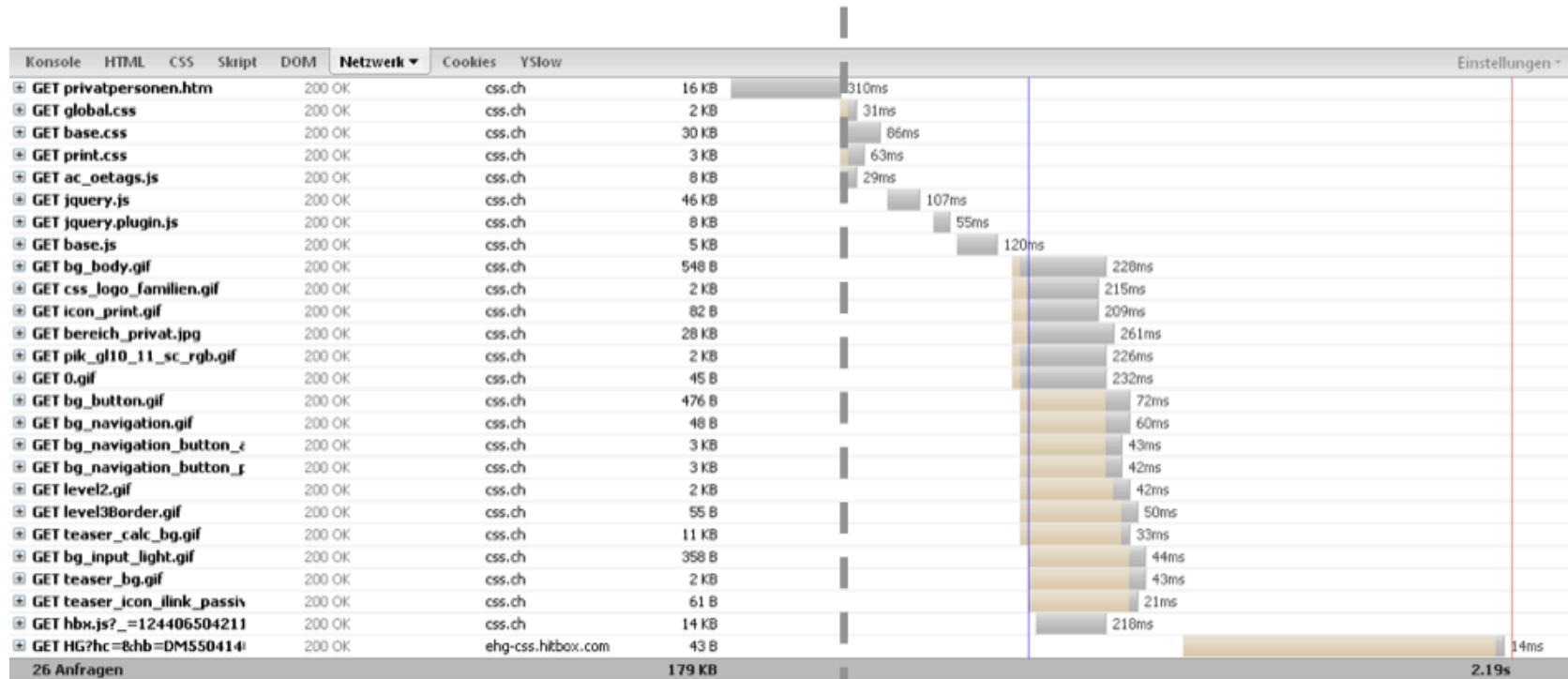
Montag, 5.
Jürg Stuke

Frankfurt, Genf,

Die Vielfalt war noch nie so gross



Die neue Problemzone



Und dann noch...



Hey IT!

The internet
thinks we
are idiots!



Please upgrade IE6 To a recent/decent browser!
hey-it.com

download .pdf (287kb)

Hey IT!

We wanna
party like
it's 2009!



Please upgrade IE6 To a recent/decent browser!
hey-it.com

download .pdf (287kb)

Hey IT!

I can haz
upgrade?
Kthnxbye!



Please upgrade IE6 To a recent/decent browser!
hey-it.com

download .pdf (288kb)

Hey IT!

Which
browser are
you using?



Please upgrade IE6 To a recent/decent browser!
hey-it.com

download .pdf (288kb)

Hey IT!

Even Bill
Gates has
upgraded!



Please upgrade IE6 To a recent/decent browser!
hey-it.com

download .pdf (288kb)

Hey IT!

Are you
sleeping?
WAKE UP!



Please upgrade IE6 To a recent/decent browser!
hey-it.com

download .pdf (287kb)

Hey IT!

2001 is
8 years ago!



Please upgrade IE6 To a recent/decent browser!
hey-it.com

download .pdf (285kb)

Hey IT!

Wanna get
laid?
Fuck IE6!



Please upgrade IE6 To a recent/decent browser!
hey-it.com

download .pdf (286kb)

Geschwindigkeit: Tipps

- **Yahoo: Best Practices for Speeding Up Your Web Site**
(<http://developer.yahoo.com/performance/rules.html>)
 - Page Speed: <http://code.google.com/speed/page-speed/>
 - YSlow: <http://developer.yahoo.com/yslow/>
- **Client Code Performance!**
 - clientseitige Tests
- **Explizite Vereinbarung (aber: Realität zählt)**
 - Ladezeit (80% aller Requests)
 - Darstellungszeit (definiertes Umfeld)
 - minimale Bandbreiten

APIs.

Nett und sinnvoll: Schnittstellen anbieten!

→ **So einfach wie nur irgendwie möglich**

- selbstbeschreibend
- zustandslos
- Client bestimmt Ausgabeformat

→ **Web-Stack**

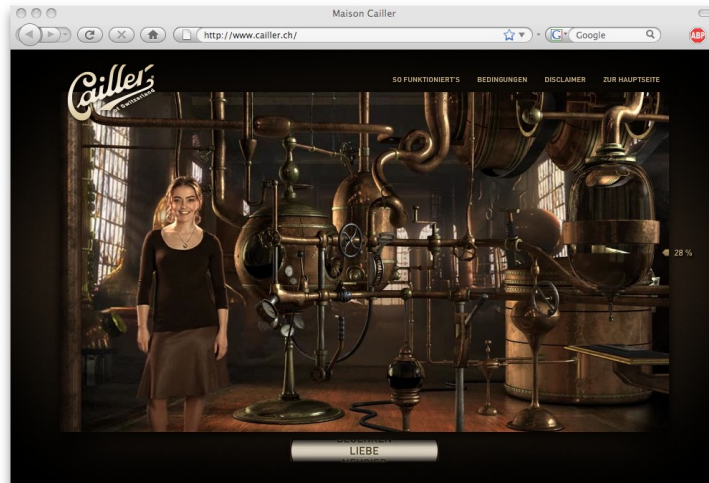
- HTTP(s): GET / POST (/PUT / DELETE)
- URI
- XML, JSON
- cachable

→ **<http://www.ics.uci.edu/~fielding/pubs/dissertation/top.htm>**

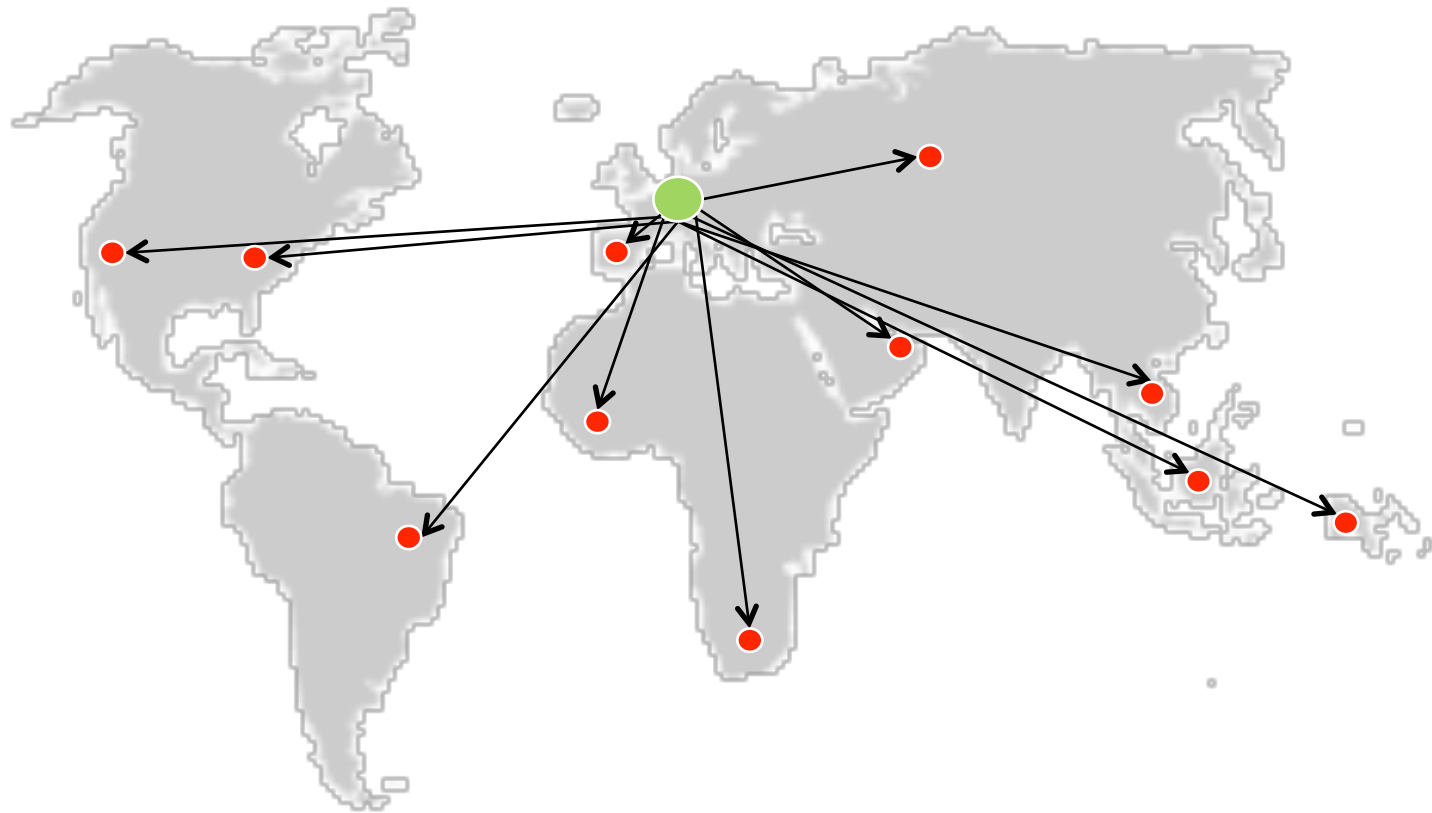
Content Distribution (Networks).

Der folgende Dreisatz

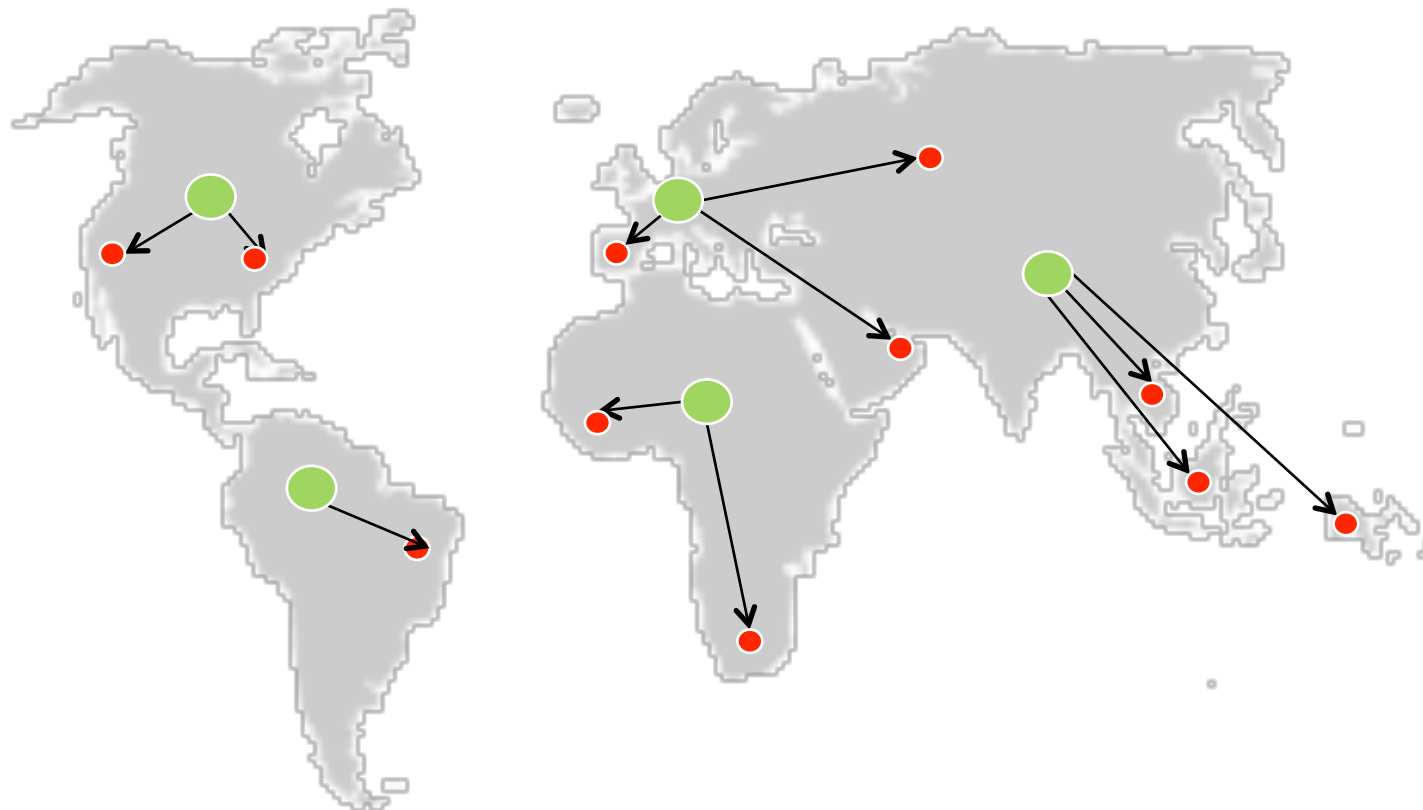
→ (Grösse Mediadatei / ø Abspielzeit) * ø Zuschauer



Meist so...



Besser wäre aber so...



Lösungsraum



→ Fragen

- Branding / Sichtbarkeit des Outsourcers
- Vertrags- / Compliancebestimmungen
- Kosten für Durchführung / Kosten für Auf- und Abbau
- (und weitere, irrationale Argumente)

CDN: Tipps

→ **Auslagern von Mediendateien!**

- bessere Infrastruktur
- mehrere Hostnamen (Sockets)

→ **(Spezialfall) Auslagern zentraler Bibliotheken**

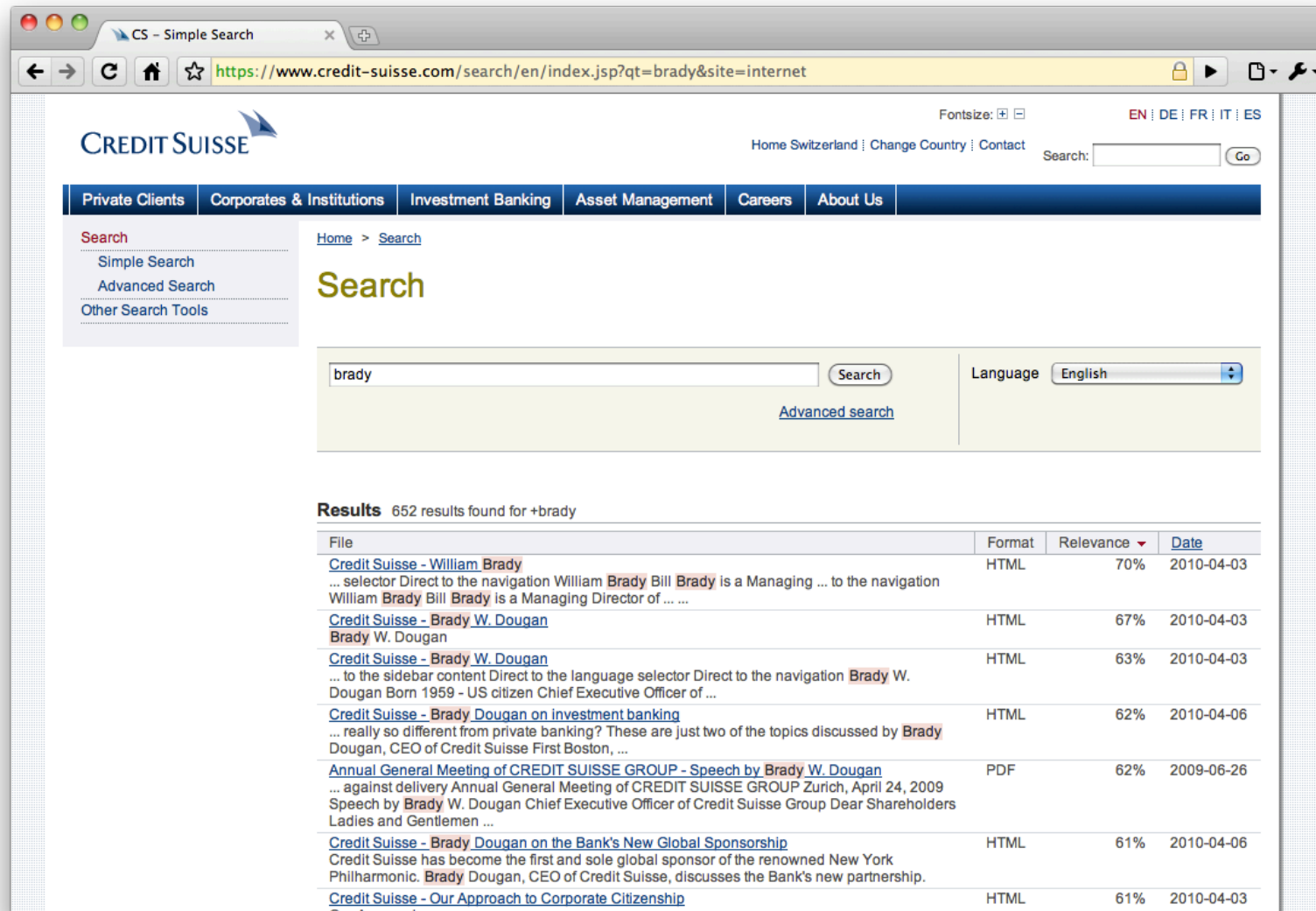
- z.B. <http://code.google.com/apis/ajaxlibs/>

→ **Mein Favorit für Content Delivery**

- Amazon S3 <http://aws.amazon.com/s3/>
- (mit oder ohne) Amazon Cloudfront
<http://aws.amazon.com/cloudfront/>

Timelines.

So eher nicht



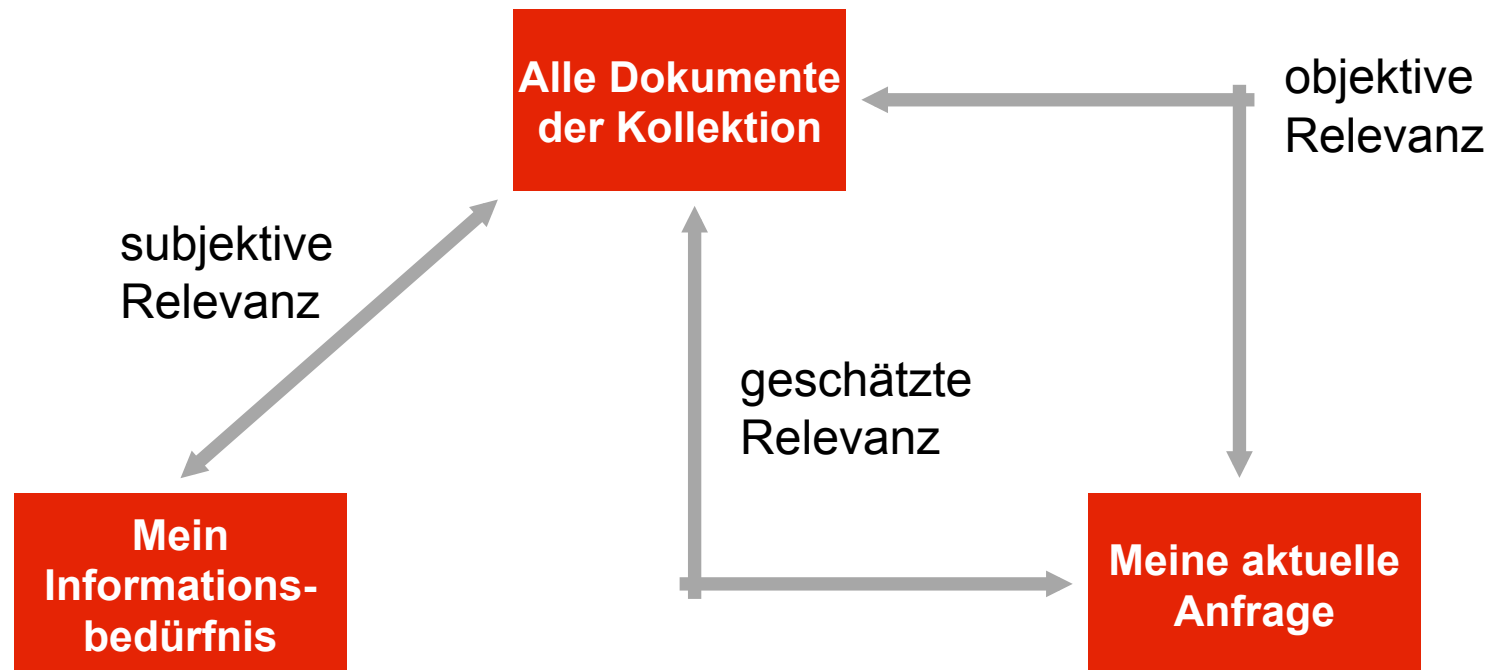
So, je nachdem

The screenshot shows the Amazon.com website with a search for "long tail" in the Books department. The search results are sorted by "Bestselling" and show two books:

- 1. Tribes: We Need You to Lead Us** by Seth Godin (Hardcover - Oct. 16, 2008)
Buy new: ~~\$40.05~~ **\$13.57**
50 new from \$10.67 19 used from \$10.52
Get it by **Thursday, Apr. 8** if you order in the next **23 hours** and choose one-day shipping.
Eligible for **FREE** Super Saver Shipping.
★★★★☆ (178)
Other Editions: Kindle Edition, Audio CD, Audio Download, Audio CD; See All.
Excerpt - page 37: "... word of mouth and viruses and outsourcing and the **long tail** and the other factors involved in social media mean that ..."
Surprise me! See a random page in this book.
- 2. Groundswell: Winning in a World Transformed by Social Technologies** by Charlene Li and Josh Bernoff (Hardcover - Apr. 21, 2008)
Buy new: ~~\$29.95~~ **\$19.77**
66 new from \$15.92 59 used from \$12.94
Get it by **Thursday, Apr. 8** if you order in the next **23 hours** and choose one-day shipping.
Eligible for **FREE** Super Saver Shipping.
★★★★☆ (105)
Other Editions: Kindle Edition, Hardcover, Audio Download, Audio CD
Excerpt - page 11: "... com. As Chris Anderson, author of The **Long Tail**,26 has pointed out, shelf space creates far less power when ..."
Surprise me! See a random page in this book.

Relevanzgewichtung

- **Relevant** heisst für mich „hier und jetzt“ wichtig
- **Doch wie weiss die Suchmaschine was „hier und jetzt“ ist (“der Kontext”)**



Welche ist nun die beste Reihenfolge?

The image shows a screenshot of a Twitter search results page. At the top, the Twitter logo is repeated three times. A search bar contains the text 'ipad OR #ipad' and a search button. To the right of the search bar are links for 'Advanced Search', 'Feed for this query', and 'Tweet these results'. Below the search bar, there are three columns of search results. The first column shows results for 'ipad OR #ipad' with tweets from users like steviewoz, sethmeyers21, timbray, ShowPatrol, PortlandORrs, and guillembach. The second column shows '617 more results since' followed by similar tweets. The third column shows results for 'ipad OR #ipad' with tweets from me1000, TariAkpodiete, bevysmith, mweisman, kayte71, MacMagazine, and lifeonprint. On the right side of the page, there are sections for 'Trending topics' (including #MyWorld2, #nowplaying, #HappyEaster, #PartOfTheProblem, #ChuckNorris, #DoctorWho, #alwayskeepthefait, #DoctorWho, #drwho, #FF) and 'Nifty queries' (including cool filter:links, 'is down', 'movie :)', 'happy hour' near:SF, #haiku).

Everything runs from memory.

Everything runs from memory

Links

→ **Twitter, an Evolving Architecture**

- <http://www.infoq.com/news/2009/06/Twitter-Architecture>

→ **Improving Running Components at Twitter**

- <http://blog.evanweaver.com/articles/2009/03/13/qcon-presentation/>

→ **Memcached: Distributed Memory Object Caching System**

- <http://memcached.org/>

Scale Out.

„Echte“ Skalierbarkeit?



scale up



scale out

Ein paar Realitäten

→ **Bei tausenden von Knoten**

- Disk, Memory und andere I/O-Fehler
- Ausfall einzelner Rechner
- Netzwerkprobleme / -partitionierung
- etc.

→ **Transaktionen müssen synchronisiert werden**

→ **Riesige Datensets müssen partitioniert werden**

Folgen

- **Neue Algorithmen**
- **Partitionierung (Separations of Concerns)**
 - simple Schnittstellen
- **Keine (möglichst wenig) zentralen Daten**
 - Zustand(slos)
 - transaktionale Informationen
- **Lose Koppelung**
 - Queueing

ACID versus BASE

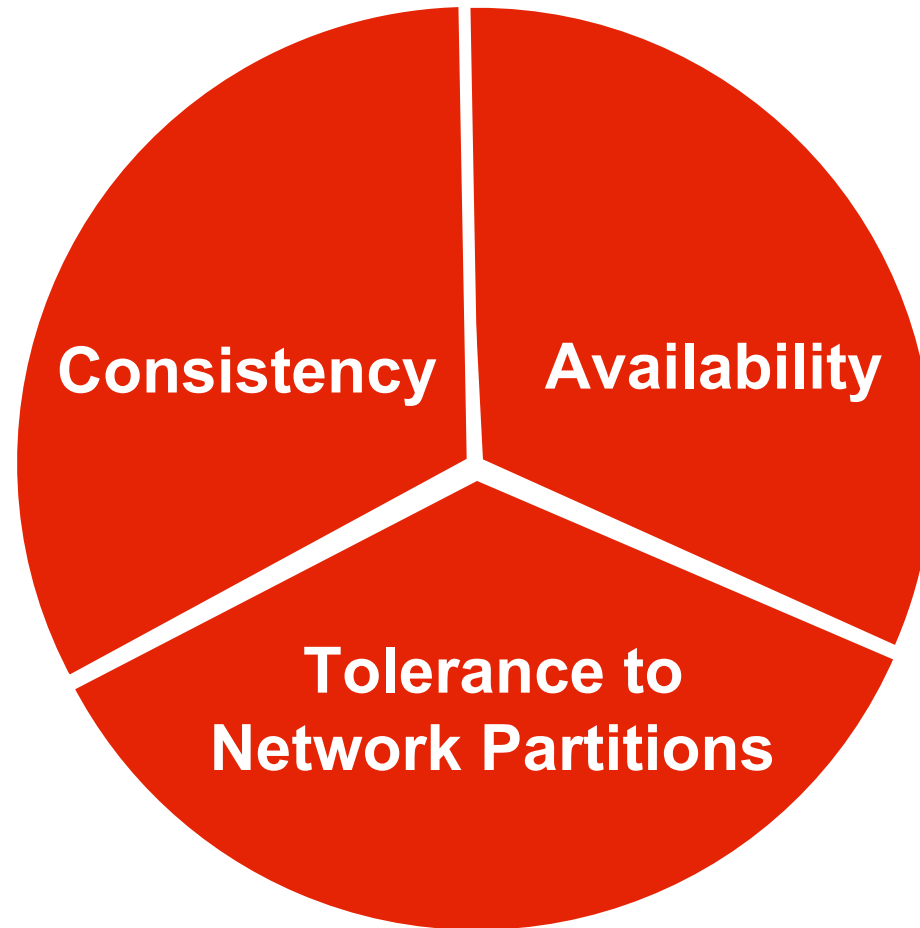
→ **ACID (Atomicity, Consistency, Isolation, Durability)**

- Strong consistency
- Isolation
- Focus in “commit”
- Nested transactions
- Availability?
- Conservative (pessimistic)
- Difficult evolution (e.g. schema)

→ **BASE (Basically Available, Soft-state, Eventual consistency)**

- Weak consistency (stale OK)
- Availability first
- Best effort
- Approximative answers OK
- Aggressive (optimistic)
- Simpler!
- Faster
- Easier evolution

Das CAP Theorem („pick two“)



Quelle: Armando Fox and Eric A. Brewer. Harvest, Yield, and Scalable Tolerant Systems. <http://tinyurl.com/cc5648>

Links (1 von 2)

- **Eric Brewer. Towards Robust Distributed Systems. 2000**
 - <http://www.cs.berkeley.edu/~brewer/cs262b-2004/PODC-keynote.pdf>

- **Werner Vogels. Eventually consistent. Communications of the ACM. 2009**
 - <http://portal.acm.org/citation.cfm?id=1435432>

- **Sanjay Ghemawat, Howard Gobioff, Shun-Tak Leung. The Google File System. 2003**
 - <http://labs.google.com/papers/gfs.html>

- **Jeffrey Dean, Sanjay Ghemawat. MapReduce: Simplified Data Processing on Large Clusters. 2004**
 - <http://labs.google.com/papers/mapreduce.html>

Links (2 von 2)

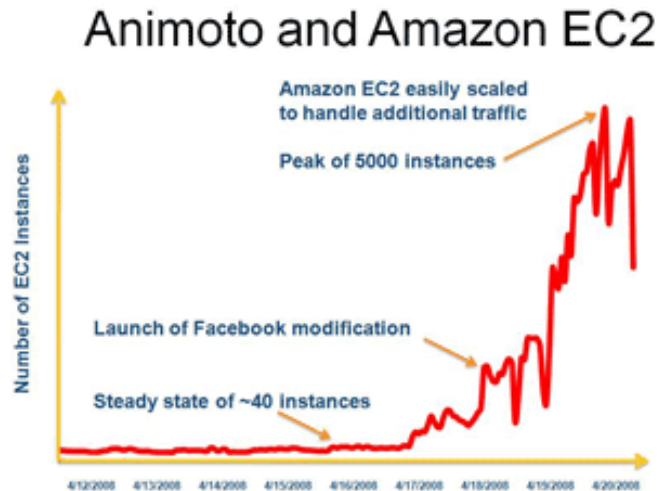
- **Fay Chang, Jeffrey Dean, Sanjay Ghemawat et al. Bigtable: A Distributed Storage System for Structured Data. 2006**
 - <http://labs.google.com/papers/bigtable.html>
- **Jinesh Varia. Architecting for the Cloud: Best Practices. AWS Whitepaper. 2010**
 - http://jinesh.varia.in/?page_id=10
- **Jinesh Varia. Cloud Architectures. 2008**
 - http://jinesh.varia.in/?page_id=10
- **Werner Vogels et al. Dynamo: Amazon's Highly Available Key-value Store. 2007**
 - http://www.allthingsdistributed.com/2007/10/amazons_dynamo.html

Cloud Deployment.

Cloud Deployment: Tipps

→ **Bei Patrice gut zuhören**

→ **Elastizität**



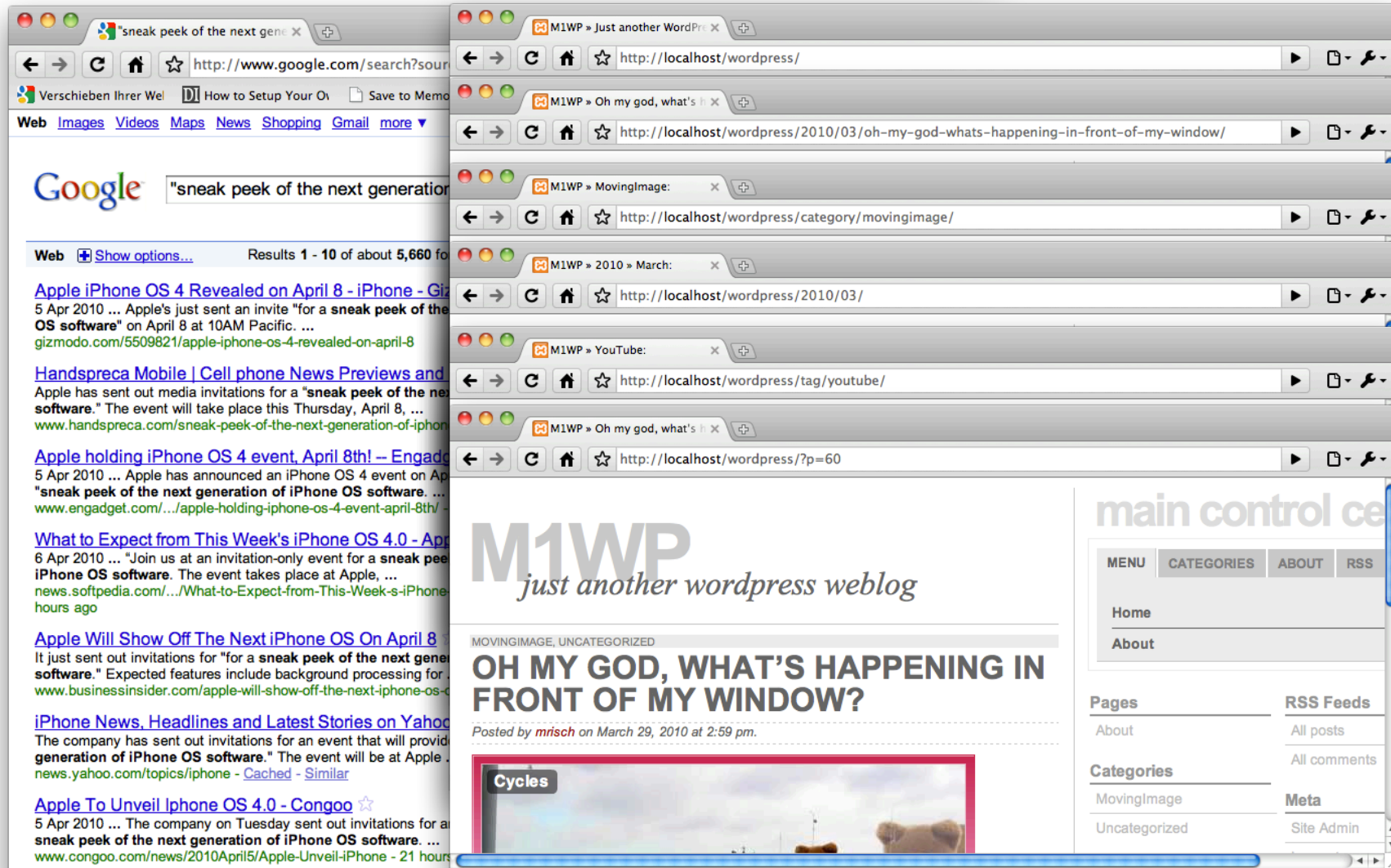
→ **Scalr: <http://aws.typepad.com/aws/2008/04/scalr-.html>**

Deduplikation.

Problem

→ **Derselbe Inhalt an hunderten von Orten?**

Etwa so?



Deduplikation: Tipps

→ **Prio 1: Canonical URL festlegen und angeben**

- <http://googlewebmastercentral.blogspot.com/2009/02/specify-your-canonical.html>

→ **Prio 2: Redirection HTTP STATUS 301**

→ **„Den Rest“ für die Crawler aussperren**

- `<META NAME="ROBOTS" CONTENT="NOINDEX, FOLLOW">`
- `/robots.txt`
- <http://www.robotstxt.org>

→ **Test über Ego-Queries in den Suchmaschinen**

Syndikation.

Ausgangspunkt

→ **Inhalte können**

- als Feeds (RSS / Atom)
- mit zustandslosen URLs (REST)

ausgerufen werden

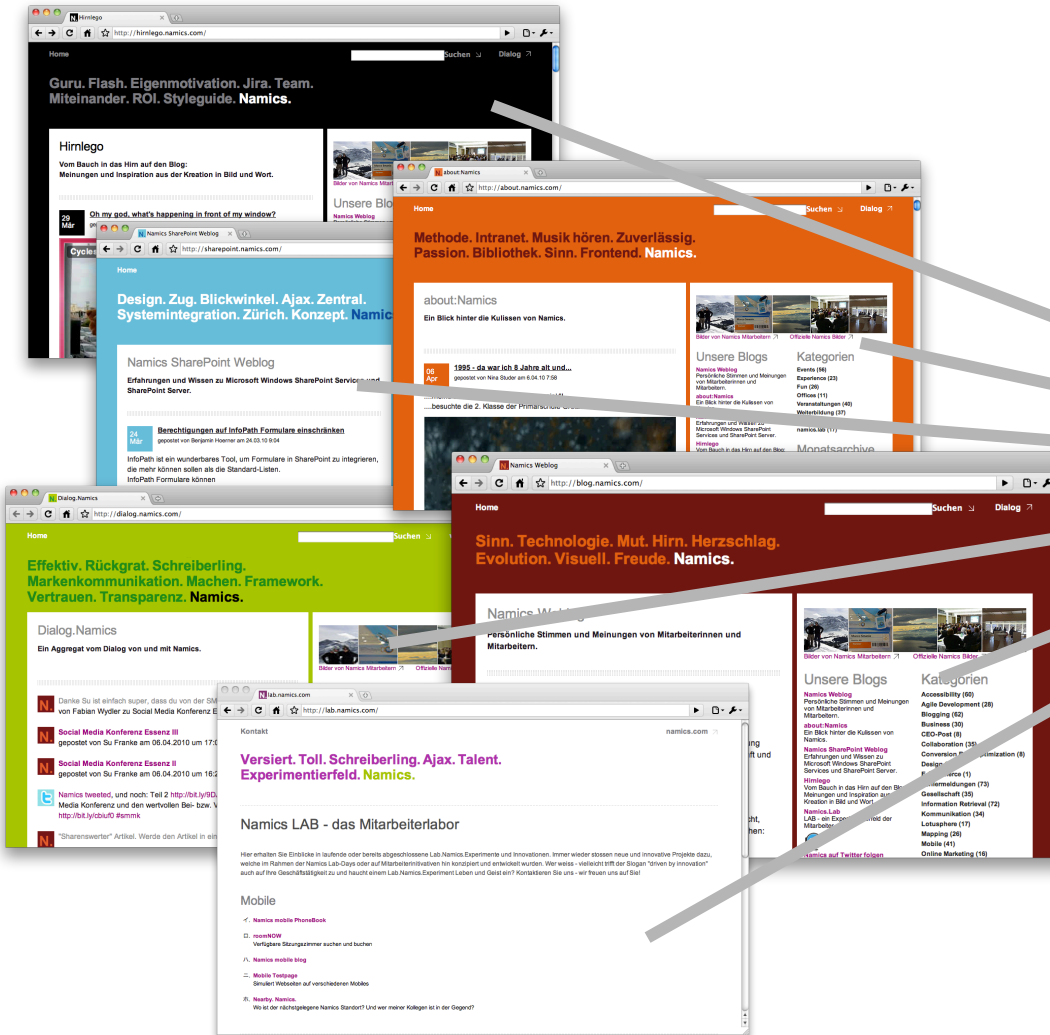
→ **Wie werde Änderungen populiert?**

- Poll
- Push
- Publish/Subscribe

Poll: „Hast Du Dich geändert?“



Push: „Ich habe mich geändert !“



XML-RPC POST
REST GET

Publish/Subscribe

→ **Paradigma aus der „alten Welt“**

- Message Queue Paradigma

→ **zentraler Hub**

- verwaltet Abos von Konsumenten („Subscriber“)
- empfängt Updates von Produzenten („Publishers“) und leitet dieser weiter

→ **Ziele**

- (in der Nähe) von Real-Time
- effizient

→ **„Alles drin“: PubSubHubbub**

Syndikation: Tipps

- **Es den Konsumenten einfach machen!**
- **Rechtlicher Rahmen**
 - <http://creativecommons.org/>
- **Autodiscovery unterstützen**
 - <http://www.rssboard.org/rss-autodiscovery>
 - <http://intertwingly.net/wiki/pie/AutoDiscovery>
- **(zumindest) Update Ping an “die Wichtigen”**
 - z.B. http://www.google.com/help/blogsearch/pinging_API.html
- **PubSubHubbub Protocol anschauen**
 - <http://code.google.com/p/pubsubhubbub/>

`<schon-fast-am-ende />.`

Technische Implikationen?

- **Präsentation / Austausch ist (gefühl) synchron**
- **Ökosystem**
 - APIs
 - Anwendungs-integration
- **Embedding**
- **Userseitige Netzwerkeffekte**
- **Netzwerkeffekte erzeugen Lastspitzen**
- **multimediale Daten sind alltäglich**
- **Aggregation (und Republikation / Autopost) von Daten**
- **Zeit als primäres Rangierungsprinzip**

Technische Implikationen!

- Schnittstellen anbieten (und nutzen)!
- Effiziente Benachrichtigung über Plattformgrenzen
- Embedding
- Skalierbare und elastische Architektur
- Cachability und Auslagern von Daten
- Zeit als primäres Rangierungsprinzip

Download:
<http://blog.namics.com/vortraege/>

@jstuker
juerg.stuker@namics.com