

# Log everything

## ...und dann?

# print Dumper(\%speaker);

- Sebastian Willing
  - CPAN: sewi
  - Blog: [www.pal-blog.de](http://www.pal-blog.de)
- Jahrgang 1979
- Erste Programmierversuche ~1988
- Perl seit 1997
- Padre, the Perl IDE
- Seit 2011 bei **spark:5** (und Vorgängern)
  - Perl, MovableType, eCommerce, Webdesign, u.v.m.
  - [www.spark5.de](http://www.spark5.de)
  - [sebastian.willing@spark5.de](mailto:sebastian.willing@spark5.de)
- Privat: 2 Kinder, Buchautor, Hobbypilot

# Ausgangssituation

Cron-Mails

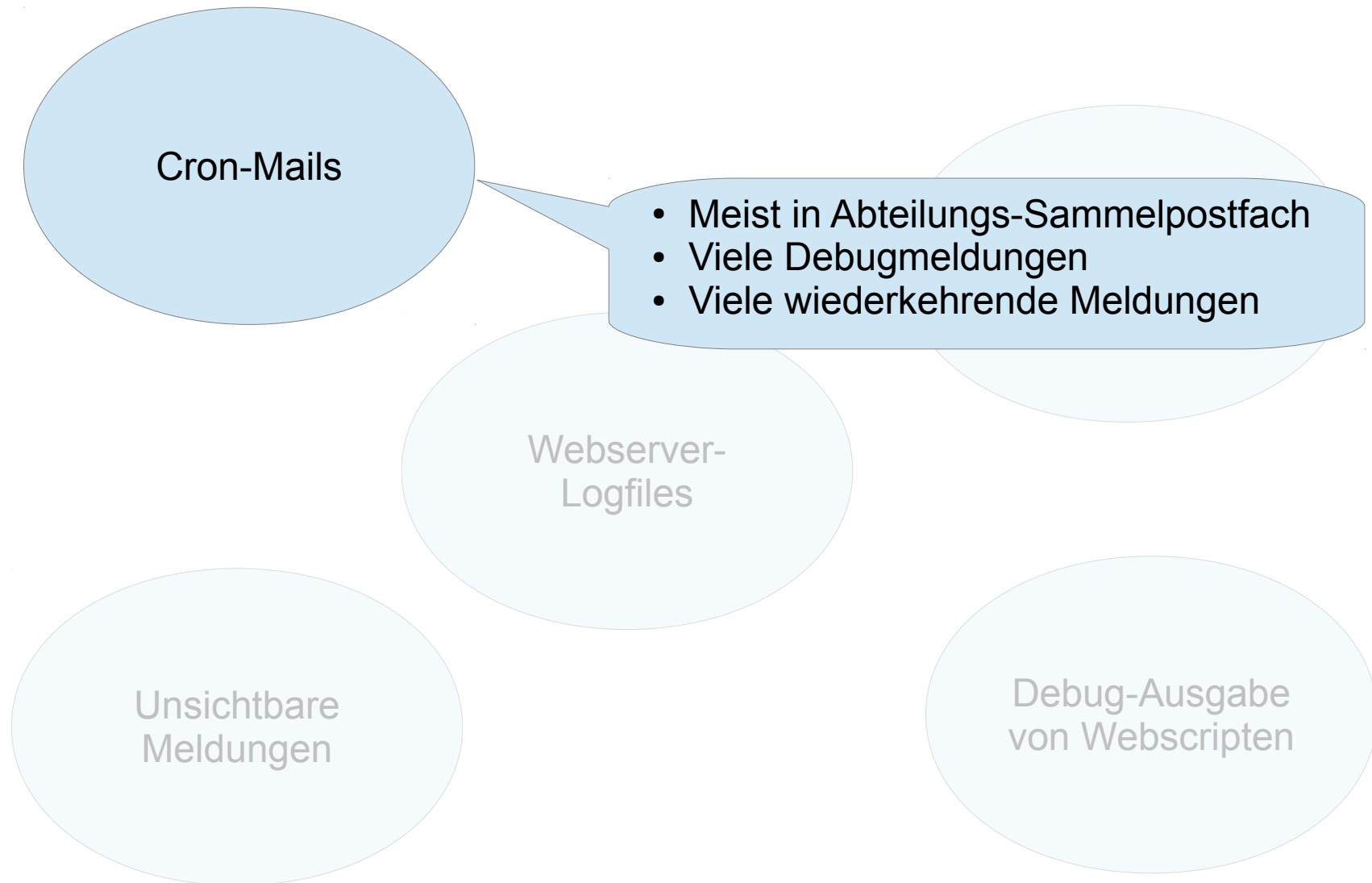
Mails an einzelne  
Entwickler

Webserver-  
Logfiles

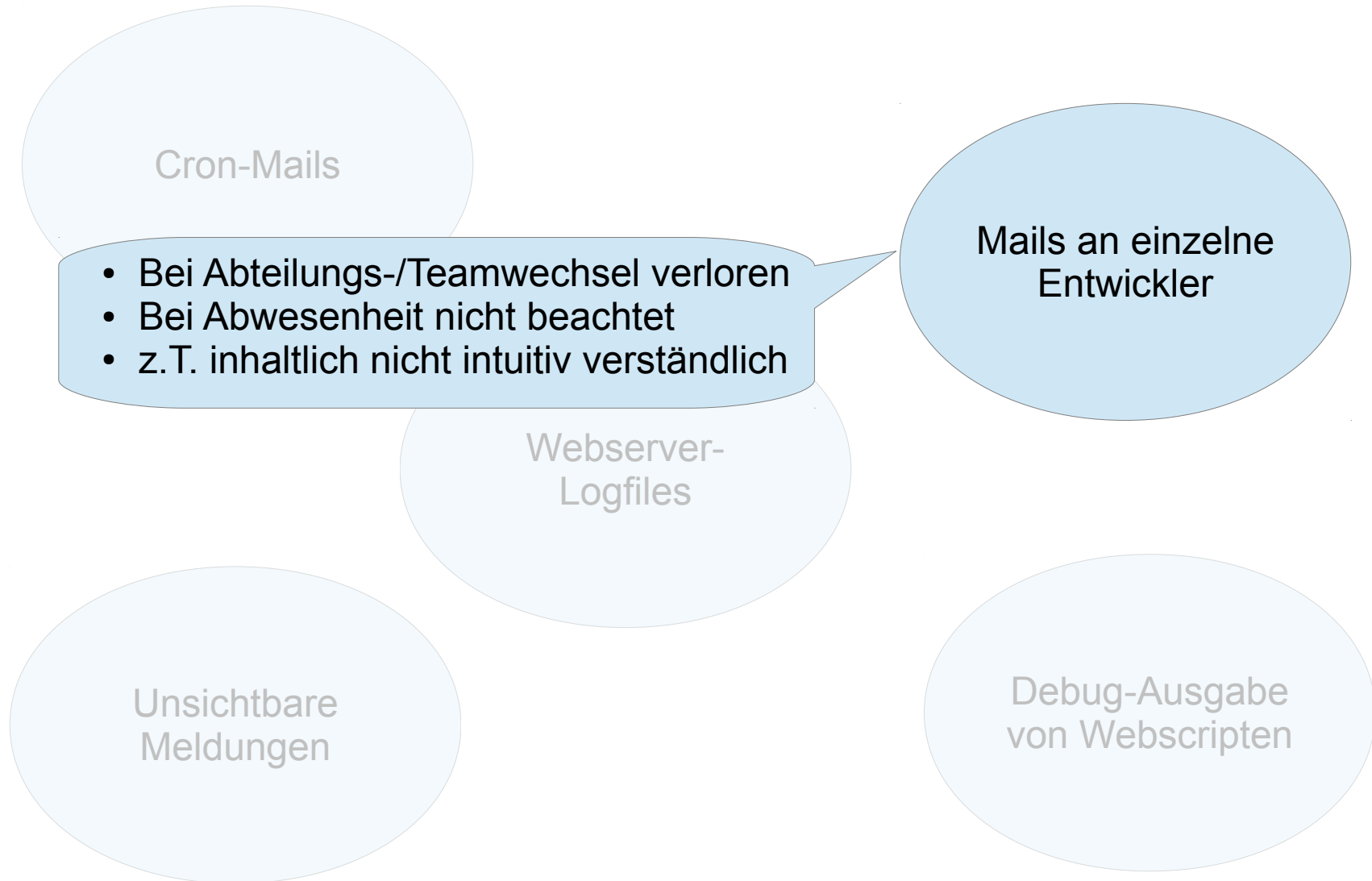
Unsichtbare  
Meldungen

Debug-Ausgabe  
von Webscripts

# Ausgangssituation



# Ausgangssituation



# Ausgangssituation

- Werden, wenn überhaupt, nur bei Probleme beachtet
- Zusammenhang mit zusätzlichen Daten (z.B. POST) kaum herstellbar

Mails an einzelne  
Entwickler

Webserver-  
Logfiles

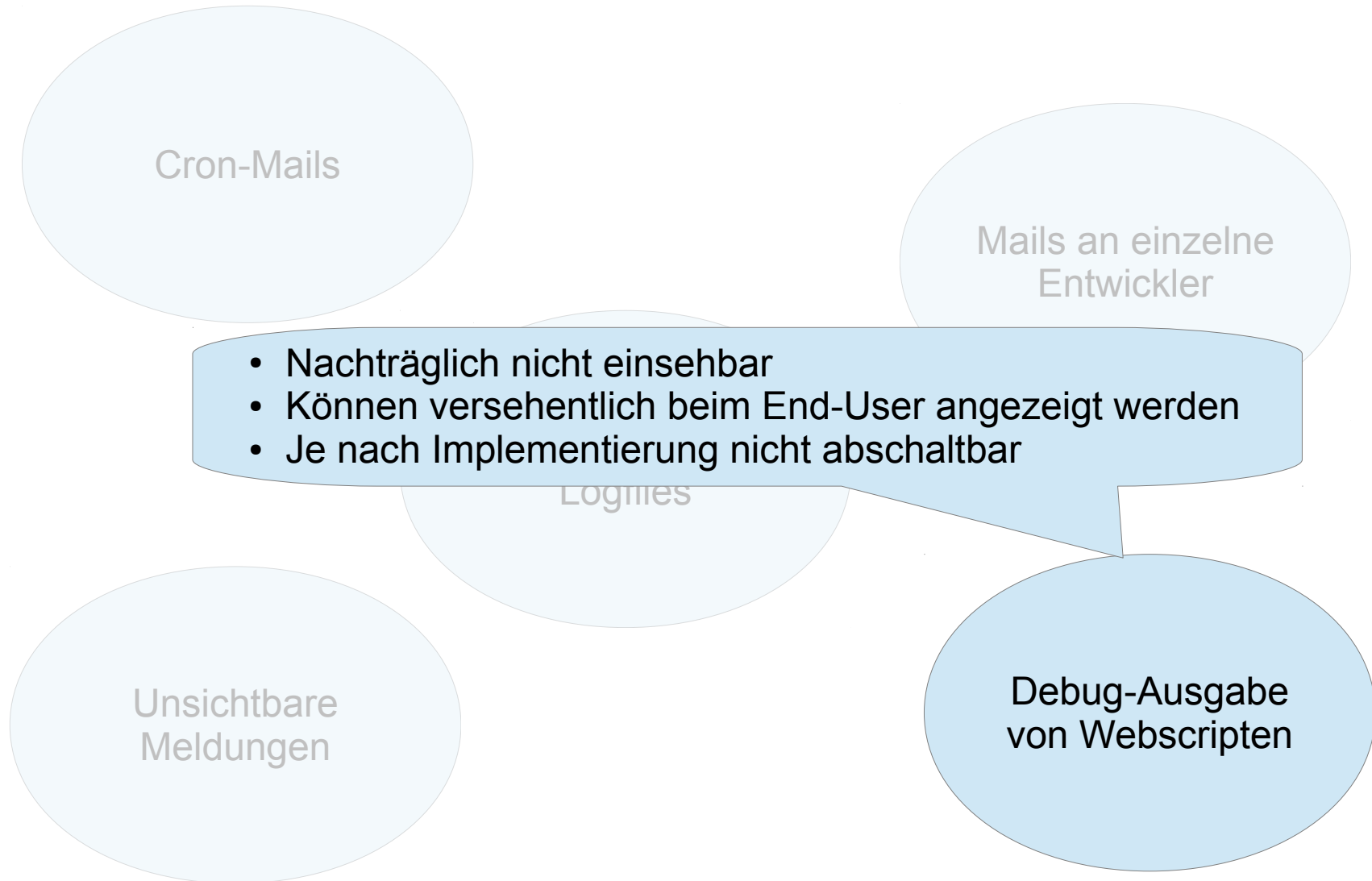
Unsichtbare  
Meldungen

Debug-Ausgabe  
von Webscripts

# Ausgangssituation



# Ausgangssituation





# Die Lösung: TechLog

Generierung

Transport

Datenbank

Klassifizierung

Erledigung

# Generierung der Meldungen

- *warn* und *die* via `$SIG{__WARN__}` und `$SIG{__DIE__}`
  - Vorteil: Bekannt und leicht zu implementieren
  - Nachteil: Kein Platz für projektspezifische Erweiterungen
- *tw* (trigger warning) und *te* (trigger error) als Ersatz für *warn* und *die*
  - Bereits teilweise im Projekt eingesetzt
  - Fehlerausgabe für Web-Scripte bereits implementiert
  - Nachteil: Benötigt `use Project::ErrorHandlers` und benutzt `Exporter`
- Aufruf:

```
tw(„Message text“,  
    report => „additional long text“,          # optional  
    dump  => \%any_ref_for_Data_Dumper,      # optional  
);
```
- Hilfsfunktion `t_info('some debug message like $var=',$var);`
  - Zwischenspeicher für Debug-Informationen
  - Daten werden nur bei folgendem *tw* oder *te* Aufruf verwendet
  - Speicher begrenzt auf die letzten 25 *t\_info* Aufrufe

Generierung

Transport

Datenbank

Klassifizierung

Erledigung

# Beispielmeldung

Error running Project::Gearman::OrduAPI::detect: garbage after JSON object, at character offset 4 (before "Can't connect to api...") at (eval 928) line 163.

Text

Received 352 bytes of data.

report => '...'

```
$VAR1 = {  
  'HTTP_USER_AGENT' => 'MQQBrowser/2.8/Adr (Linux; U; 2.3.5; zh-cn; T9300A Build/MocorDroid;480*854)',  
  'HTTP_X_FORWARDED_FOR' => '123.34.67.89',  
  'IDENT' => '410.1393258093.92789.2777074013',  
  'QUERY_STRING' => 'user_id=d30619bf44111c4ffb18c735758712b4&sub_id=CB9776',  
  'REDIRECT_STATUS' => '200',  
  'REMOTE_ADDR' => '172.19.49.21',  
  'REMOTE_PORT' => '34418',  
  'REQUEST_METHOD' => 'GET',  
};
```

dump => {...}

Gearman-Job-Daten

```
Job-Info:$VAR1 = {  
  'arg' => {  
    'caller' => '/srv/www/cgi-bin/Project/Web/User.pm:137',  
    'context' => 'cn',  
    'data' => {  
      'HTTP_USER_AGENT' => 'MQQBrowser/2.8/Adr (Linux; U; 2.3.5; zh-cn; T9300A  
Build/MocorDroid;480*854)',  
      [...] }  
    },  
  'job' => 'OrduAPI::detect'  
};
```

Generierung

Transport

Datenbank

Klassifizierung

Erledigung

# Beispielmeldung

Generierung

Stacktrace:

0. /srv/www/cgi-bin/Project/Gearman/JobMaster.pm - call of Project::ErrorHandlers::tw in  
Project::Gearman::JobMaster , line 674

1. /usr/share/perl5/Gearman/Worker.pm - call of Project::Gearman::JobMaster:: ANON in Gearman::Worker ,  
line 240

Stacktrace

Land/Projekt

Server

Datum/Uhrzeit  
(Land/Projekt-Zeitzone, UTC, Deutschland)

Transport

Context: 'cn' / Server: vs-gearmanw-a-03

Time: Tue Feb 25 00:08:33 2014 (local) / Mon Feb 24 16:08:33 2014 (UTC) / 2014-02-24T17:08:33 (Germany)

Assumed last SQL: INSERT INTO USER\_LOGIN\_LOG (user\_id, login\_dt, ip) VALUES (?, ?, ?)

Last DBI error: No DBI error message found

!\$: No such device or address

\$@:

DBI->err und DBI->errmsg

Letztes SQL-Query  
(fast immer richtig)

Datenbank

Evtl. hilfreiche Perl-Standardvariablen

Latest HTTP arguments:

-----

Bei Web-Scripten: GET und Post Daten

Klassifizierung

Content of @ARGV:

-----

1. gearman\_slow\_jobs

2. gearman\_fast\_jobs

Erledigung

Content of %ENV:

-----

HOME => /srv/www/cgi-bin

LANG => en\_US.UTF-8

PATH => /bin:/sbin:/usr/bin:/usr/sbin

# Beispielmeldung

Process status:

-----

```
USER    PID %CPU %MEM    VSZ   RSS TTY      STAT START   TIME COMMAND
root    429  1.0  0.0 53372 10048 ?        S   Feb24  0:07 /usr/bin/perl /var/www/project/bin/gearman_workers.sh
www-data 554  4.7  0.2 322760 172772 ?       Ss  Feb24  0:31 \_ /var/www/project/bin/gearman_worker.pl
OrduAPI::detect H:gearmand-01:384332148
```

``ps fuww $parent_pid $$``

t\_info stack:

-----

```
1393258093 /var/www/project/cgi-bin/Project/Environment.pm:82@cn: cn from /srv/www/project/bin/Project/JobMaster.pm:633
1393258092 /srv/www/project/cgi-bin/Project/JobMaster.pm:594@(undef): Start job OrduAPI::detect
```

Die letzten 25 t\_info Debugmeldungen

t\_info Parameter

Timestamp  
(relative Zeit ist oft  
interessanter als absolute)

Land

Caller von t\_info

Generierung

Transport

Datenbank

Klassifizierung

Erledigung

# Andere Quellen

Generierung

- MySQL
  - Alle Fehlermeldungen
  - Viele Warnungen
  - DBI::st Überlagerung
- Alle Perl-Warnings via `$SIG{__WARN__}`
- Syslog-Meldungen (gefiltert)
- Webserver
  - FastCGI Fehlermeldungen
  - Error-Log (zeitweise)
- Projekt-Monitoringsystem
- Verwandte Projekte

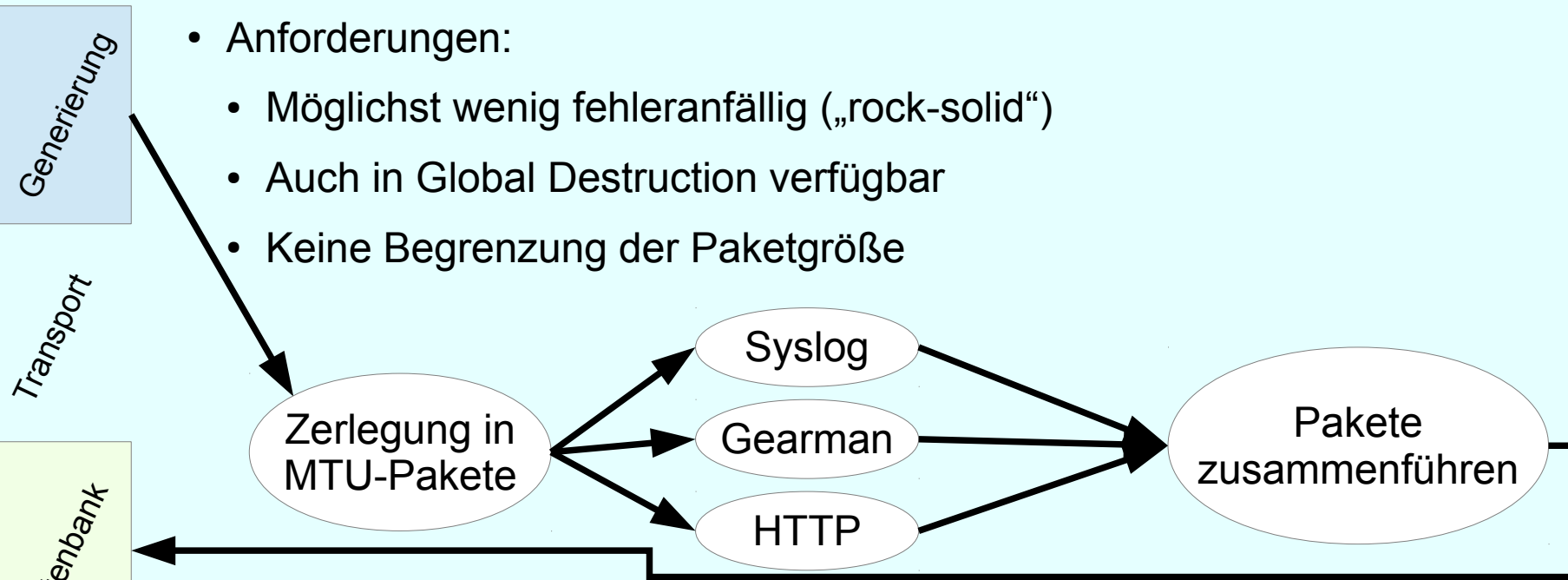
Transport

Datenbank

Klassifizierung

Erledigung

# Transport



- Anforderungen:
  - Möglichst wenig fehleranfällig („rock-solid“)
  - Auch in Global Destruction verfügbar
  - Keine Begrenzung der Paketgröße

- Lösung:
  - Modularer Transport über verschiedene Wege
    - Syslog
    - Gearman-Jobs (beste Ergebnisse = bevorzugter Transportweg)
    - HTTP
    - UDP (nicht im Praxiseinsatz)
  - Individuelle MTUs je Transportweg

# Datenbank

Generierung

Transport

Datenbank

Klassifizierung

Erledigung

- Unterschiedliche Anforderungen für Development und Livesystem
  - Development
    - Hohe Fluktuation = kurze Lebenszeit der Einzelmeldungen
    - Möglichst ein Datenspeicher pro Entwickler, auch bei ansonsten geteilten Entwicklungsressourcen
  - Livesystem
    - Hohe DB-Last durch viele und z.T. sehr große Meldungen
    - Unterschiedlich lange Vorhaltezeiten („same day“ bis zu „end-of-ticket“)
- Lösung
  - Modulare Speicherlösung (weitgehend ORM)
  - Wegen Performanceoptimierung selbstentwickelt und mit eingeschränkten Funktionen
  - SQLite (Development)
  - MySQL (Livesystem, z.T. an der Leistungsgrenze)
  - MongoDB (wünschenswert für Livesystem)



# Klassifizierung

- Gleiche Meldungen werden zu „Events“ gruppiert
  - Ziel: Eine Ursache = ein Event
- Beispielmeldungen:
  - Use of uninitialized value in numeric comparison (<=>) at some/path/lib/Project/Backoffice/Statistics.pm line 142.
  - Argument "\x{33}\x{38}..." isn't numeric in numeric ne (!=) at some/path/lib/Project/Cntl/Web.pm line 68.
  - Query execution failed: Duplicate entry '9748541' for key 'PRIMARY' for USER\_ANSWER
  - Invalid TIMESTAMP value in column ? at row ? for DELAYED\_JOBS (1299)

Generierung

Transport

Datenbank

Klassifizierung

Erledigung

# Automatische Klassifizierung

- „Auto-Unique-ID“
  - Text wird um variable Werte bereinigt
  - Datum/Uhrzeit  $s/\wedge b(?:Mon|...|Sun) (?:Jan|...|Dec)\s*d+\s*d+:\s*d+:\s*d+\wedge d\{4}\wedge b//g$
  - Email-Adressen
  - Ref's  $s/(HASH|ARRAY|SCALAR|CODE)\(0x[\wedge da-f]+\)/\$1\(\wedge\)/g$
  - Eval-Nr.  $s/\wedge (eval \wedge d+\wedge)\wedge (eval\wedge)/g$
  - Filehandle-Zeilenummern  $s/, \wedge <|\$?\wedge w+> line \wedge d+\wedge .\wedge .\wedge /$
  - Alle Leerzeichen, Zeilenumbrüche, Tabs  $s/\wedge s+//g$
  - Ziel: Um Variablen bereinigter Meldungstext
- Beispiel:
  - Query execution failed: Duplicate entry '9748541' for key 'PRIMARY' for USER\_ANSWER
  - => Queryexecutionfailed:Duplicateentry"forkey'PRIMARY'forUSER\_ANSWER

Generierung

Transport

Datenbank

Klassifizierung

Erledigung

# Manuelle Klassifizierung

- Manuelle „Filter rules“

TechLog Events >> Query execution failed: Duplicate entry '9748541' for key 'PRIMARY' for USER\_ANSWER >> Rules >>

## TechLog Event 166016 rule

### Edit rule

Priority:

Title: Match duplicate keys in all answer tables

Condition:

(\$msg, \$rule) = @\_;

```
$msg->text =~ /Duplicate entry '.*' for key 'PRIMARY' for (?:USER|VISITOR)_ANSWER/;
```

Expected return values:

false Rule didn't match, try next rule

true Rule matched, apply action

Sample message:

\$msg->id XiCWVeqZUWingame

\$msg->datetime 2014-03-10 14:34:20

\$msg->context br

\$msg->priority 4

\$msg->source\_file /some/path/cgi-bin/script.pl

\$msg->source\_line 2656

\$msg->text (beginning) Query execution failed: Duplicate entry '9748541' for key 'PRIMARY' for |

\$msg->auto\_unique Queryexecutionfailed:Duplicateentry''forkey'PRIMARY'forUSER\_ANSWER

Action on matched condition: No action

Notes: No action

Use this event

Skip this event

Test

Cancel

Menschenlesbarer Titel

Regel als Perl-Source  
\$msg = aktuelle Nachricht

Return value:  
true=match / false=non-match

Ansicht der wichtigsten  
Eigenschaften von \$msg

Aktion bei „match“

Notizen

Vor dem Speichern: Syntax- und Match-Test

Generierung

Transport

Datenbank

Klassifizierung

Erledigung

# Routineaufgaben automatisieren

- Konfigurierbare Actions

TechLog Events >> Long running frontend request found. >> Actions >> delete old  
**TechLog Event 13875 action 133029**

Title: delete old

Timeout: 2038-01-01 23:59:59 (JJJJ-MM-DD hh:mm:ss)

Action: Delete read messages after 7 days

Notes:

- No action
- Mark messages as read
- Delete all messages
- Send email notification
- Add note to task
- Delete read messages after 1 days
- Delete read messages after 7 days
- Delete read messages after 14 days**
- Delete read messages after 30 days
- Delete read messages after 90 days
- Keep no more than 100 messages
- Keep no more than 1000 messages

Save Cancel

Menschenlesbarer Titel

Action wird nach Timeout automatisch gelöscht

Nicht mehr zur manuellen Bearbeitung anzeigen z.B. solange das Ticket noch offen ist, wird auch automatisch erstellt

Nachrichten ignorieren  
Nicht ungefährlich!

Volumenbegrenzungen

Zusätzliche Benachrichtigungen via Email oder Ticket

Generierung  
Transport  
Datenbank  
Klassifizierung  
Erledigung

# Meldungen abarbeiten

- Übersichtsseite optimiert für schnelle Ab...

Standardansicht:  
Events mit unbearbeiteten Nachrichten

Unread events | All events | All rules | Statistics

Site reload in 56 seconds. [reload now](#)

Actions for selected events: [Mark as read](#) [Mark as read for today](#) [Delete events & reassign messages!](#) [Assign to task #](#)

! Applies only to events without rules, actions and taskid.  
[ALT + CLICK] on a tablerow checks this entry.  
[SHIFT + CLICK] on a tablerow and again on another one to select all in between.

2014-03-10: 46372 messages  
2014-03-09: 48550 messages  
2014-03-08: 159661 messages

TaskID:  [Search](#)

Newest message

<input type="checkbox"/>	166020	cgi-bin/backoffice/statistics/statistics.pl failed: Script Runtime Error: Missing comment! The comment-field is required! at (eval 3901)		2014-03-10 16:00:03	<a href="#">+</a> new	2
<input type="checkbox"/>	107326	Norway - GWS: DOI - Worst current DOI percentage on one website			<a href="#">+</a>	
<input type="checkbox"/>	108285	User manipulated URL:		2014-03-10 14:07:36	<a href="#">+</a> new	26
<input type="checkbox"/>	165731	Q-die: Can't locate AnyEvent/AIO.pm in @INC (@INC contains: ../lib lib /etc/perl /usr/local/lib/perl/5.14.2 /usr/local/share/perl/5.14.2 /usr/lib/perl5 /usr/share/perl5 /usr/lib/perl/5.14 /usr/share/perl/5.14 /usr/local/lib/site_perl .) at /usr/lib/perl5/		2014-03-10 10:51:38	<a href="#">+</a> new	2
<input type="checkbox"/>	164936	Data truncated for column ? at row ? for MAIL_SPAM_TESTS (1265)		2014-03-10 07:52:42	<a href="#">+</a> new	1
<input type="checkbox"/>	165573	Use of uninitialized value \$compressed in concatenation (.) or string at		2014-03-10 05:19:37	<a href="#">+</a> new	2

Automatische Aktualisierung  
Stoppt sobald Nachrichten  
markiert werden

Kurzstatistik der  
letzten Tage

Kurzbedienungsanleitung  
für Gelegenheitsnutzer

Generierung

Transport

Daten

Klassifizierung

Erledigung

# Meldungen abarbeiten

- Übersichtsseite optimiert für schnelle Abarbeitung

Generierung  
Transport  
Datenbank  
Klassifizierung  
Erledigung

Unread events | All events | All rules | Statistics

Site reload in 56 seconds. [reload now](#)

Actions for selected events: [Mark as read](#) [Mark as read for today](#) [Delete events & reassign messages<sup>1</sup>](#) [Assign to task #](#)

Event-ID

Titel (i.d.R. Aus der ersten Nachricht generiert)

Letztes Vorkommen absteigend sortiert

+ New event Search events:  TaskID:  [Search](#)

all	ID	Title	Newest message	TaskID	Unread
<input type="checkbox"/>	166018	MARS Project Overview (166018)	2014-03-10 16:00:03	+ new	2
<input type="checkbox"/>	166020	cgi-bin/backoffice... Missing comment! The comment-field is required! at (eval 3901) line 670	2014-03-10 15:58:00	+ new	2
<input type="checkbox"/>			15:55:06	100263 +	2
<input type="checkbox"/>			15:55:06	+ new	26
<input type="checkbox"/>	165731	Q-die: Can't locate AnyEvent/AIO.pm in @INC (@INC contains: ../lib lib /etc/perl /usr/local/lib/perl/5.14.2 /usr/local/share/perl/5.14.2 /usr/lib/perl5 /	2014-03-10 10:51:38	+ new	2
<input type="checkbox"/>	164936	Data truncated for column ? at row ? for MAIL_SPAM_TESTS (1265)	2014-03-10 07:52:42	+ new	1
<input type="checkbox"/>	165573	Use of uninitialized value \$compressed in concatenation (.) or string at	2014-03-10 05:19:37	+ new	2

1-Klick-Ticketerstellung

Ticket-ID, Hintergrundfarbe=Status (wie im Ticketsystem)  
Icons für „Ticket wiedereröffnen“ und „Neues Ticket öffnen“

Anzahl der unbearbeiteten Nachrichten zu diesem Event



# Meldungen abarbeiten

- Übersichtliche Abarbeitung

Heute nicht mehr anzeigen  
erstellt „Mark messages  
as read“ mit Timeout 23:59:59

Site reload in 56 seconds. [reload now](#)

Actions for selected events: [Mark as read](#) [Mark as read for today](#) [Delete events & reassign messages<sup>1</sup>](#) [Assign to task #](#)

<sup>1</sup> Applies only to events with titles, actions and taskid.  
<sup>2</sup> [ALT + CLICK] on a table row links this entry.  
<sup>3</sup> [SHIFT + CLICK] on a table row and again on another one to select a range between.

2014-03-10: 46372 messages  
 2014-03-09: 48550 messages  
 2014-03-08: 159661 messages

TaskID:  [Search](#)

			Newest message	TaskID	Unread
<input type="checkbox"/>	166018				2
<input type="checkbox"/>	1660				2
<input type="checkbox"/>	1073				2
<input type="checkbox"/>	108285	User manipulated URL:	2014-03-10 14:07:36	+ new	26
<input type="checkbox"/>	165731	Q-die: Can't locate AnyEvent/AIO.pm in @INC (@INC contains: ../lib lib /etc/perl /usr/local/lib/perl/5.14.2 /usr/local/share/perl/5.14.2 /usr/lib/perl5 /usr/share/perl5 /usr/lib/perl/5.14 /usr/share/perl/5.14 /usr/local/lib/site_perl .) at /usr/lib/perl5/	2014-03-10 10:51:38	+ new	2
<input type="checkbox"/>	164936	Data truncated for column ? at row ? for MAIL_SPAM_TESTS (1265)	2014-03-10 07:52:42	+ new	1
<input type="checkbox"/>	165573	Use of uninitialized value \$compressed in concatenation (.) or string at	2014-03-10 05:19:37	+ new	2

Generierung

Support

Datenbank

Klassifizierung

Erledigung

Ausblenden  
(als „bearbeitet“ (read) markieren)

Manuelle Regel wurde erstellt:  
Event löschen und Nachrichten  
neu klassifizieren

Bekannter Ticket-ID  
zuweisen

# Event-Ansicht

Generierung  
 Transport  
 Datenbank  
 Klassifizierung  
 Erledigung

Unread events | All events | All rules | Statistics

TechLog Events >> Long running frontend request found.

TechLog

pk	title	created	taskid
13875	Long running frontend request found.	Thu Jan 10 11:22:43 2013	102077

Neue Nachrichten der letzten 30 Tage

10.03.	09.03.	08.03.	07.03.	06.03.	05.03.	04.03.	03.03.	02.03.
347	1	319	214	359	343	312	431	432

Alle Nachrichten aller Events dieses Tickets

Mark all messages as read (today) | Show all messages for task 102077 | Filter context: (All)

Date	Type	Context	Source
2014-03-10 15:21:16	unknown	au	...BAL/Profiling.pm: 107 Long running frontend request found (#86884)
2014-03-10 15:21:16	unknown	de	...BAL/Profiling.pm: 107 Long running frontend request found (#86884)
2014-03-10 15:21:14	unknown	us	...BAL/Profiling.pm: 107 Long running frontend request found (#86884)

Event-Daten

Filter & Actions

Aktionen

Liste der Einzelnachrichten



# Ticketsystem

Generierung

Transport

Datenbank

Klassifizierung

Erledigung

(0567014)  
**Notiz**  
**Cron\_user** (Administrator)  
2014-02-27 15:03

Bearbeiten   Löschen  
Wichtig machen  
Sticky machen  
Zuordnen zu **Cron\_user**

1 new messages in [https://techlog.internal/cgi-bin/techlog\\_event.pl?todo=show\\_event&pk=157201](https://techlog.internal/cgi-bin/techlog_event.pl?todo=show_event&pk=157201) since 2014-01-27 11:57:55

(0567015)  
**Notiz**  
**Cron\_user** (Administrator)  
2014-02-27 15:03

Bearbeiten   Löschen  
Wichtig machen  
Sticky machen  
Zuordnen zu **Cron\_user**

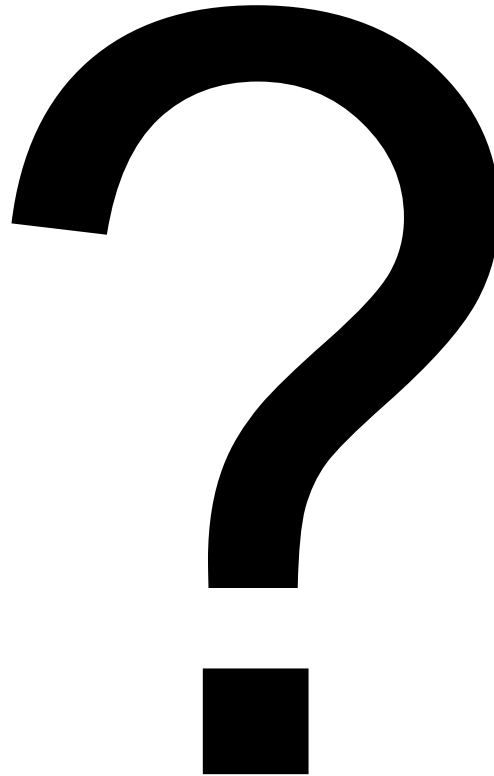
3 new messages in [https://techlog.internal/cgi-bin/techlog\\_event.pl?todo=show\\_event&pk=123362](https://techlog.internal/cgi-bin/techlog_event.pl?todo=show_event&pk=123362) since 2014-01-27 11:57:55

(0569167)  
**Notiz**  
**Cron\_user** (Administrator)  
2014-03-10 09:46

Bearbeiten   Löschen  
Wichtig machen  
Sticky machen  
Zuordnen zu **Cron\_user**

130 new messages in [https://techlog.internal/cgi-bin/techlog\\_event.pl?todo=show\\_event&pk=123362](https://techlog.internal/cgi-bin/techlog_event.pl?todo=show_event&pk=123362) since 2014-02-27 15:03:07

# Fragen



# Vielen Dank fürs zuhören

<http://www.pal-blog.de>  
[sebastian.willing@spark5.de](mailto:sebastian.willing@spark5.de)  
<http://www.spark5.de>

[https://twitter.com/PAL\\_Blog](https://twitter.com/PAL_Blog)

<https://www.facebook.com/sebastian.willing>

<https://plus.google.com/u/0/+SebastianWilling>