

OpenPOWER™

OpenPOWER – Vorteile und Anwendungsgebiete der offenen POWER Architektur

Webinar vom 25.01.2017

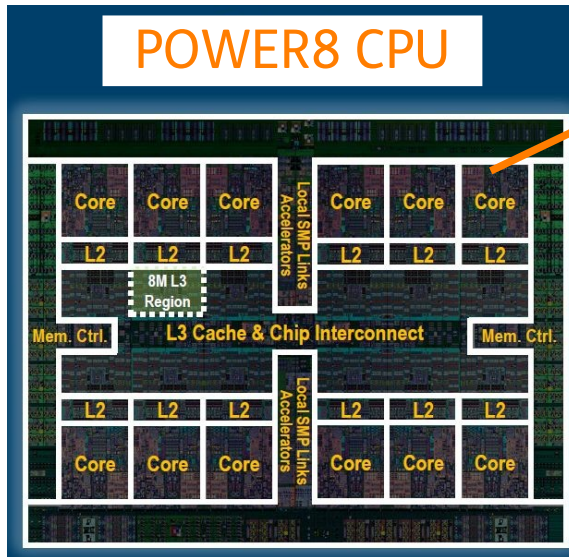
Werner Fischer, Thomas-Krenn.AG

**THOMAS
KRENN®**

Agenda:

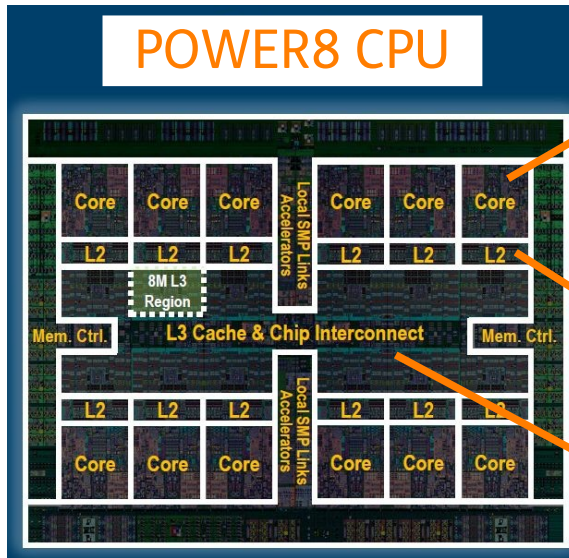
- POWER Architektur
- OpenPOWER Firmware
- RAS-Features: Reliability, Availability, Serviceability
- Datenbanken
- Virtualisierung
- OpenPOWER von Thomas-Krenn

Performance
Optimization
With
Enhanced
RISC



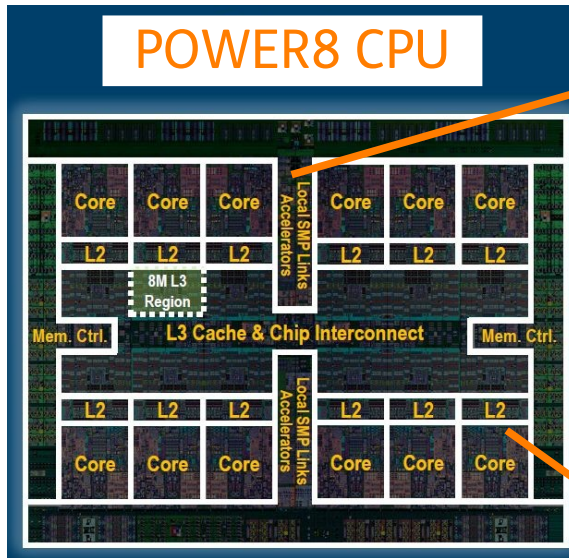
- Multi-Core:
 - 8 Cores
 - 10 Cores
 - 12 Cores
- 8x SMT pro Core (Simultaneous Multi-Threading)
- 64, 80 oder 96 Threads

POWERful Caches



- L1 Cache:
 - 64KB Data / Core
 - 32KB Instr. / Core
- L2 Cache: 512KB / Core
- L3 Cache: 8MB / Core
- L4 Cache in den Memory Buffers

POWERful Accelerators



Pro Chip:

- Symmetric Crypto
- Compression Engine
- Random Number Generator

Pro Core:

- Symmetric Crypto
- Cyclic Redundancy Check

Crypto Accelerators

https://github.com/openssl/openssl/blob/OpenSSL_1_0_2/CHANGES#L67

```
67 *) Initial support for PowerISA 2.0.7, first implemented in POWER8.  
68 This covers AES, SHA256/512 and GHASH. "Initial" means that most  
69 common cases are optimized and there still is room for further  
70 improvements. Vector Permutation AES for AltiVec is also added.  
71 [Andy Polyakov]
```

OpenSSL >= 1.0.2 ab

- SLES 12 SP2
- RHEL 7.4¹
- Ubuntu 15.10
- Debian 9.0²

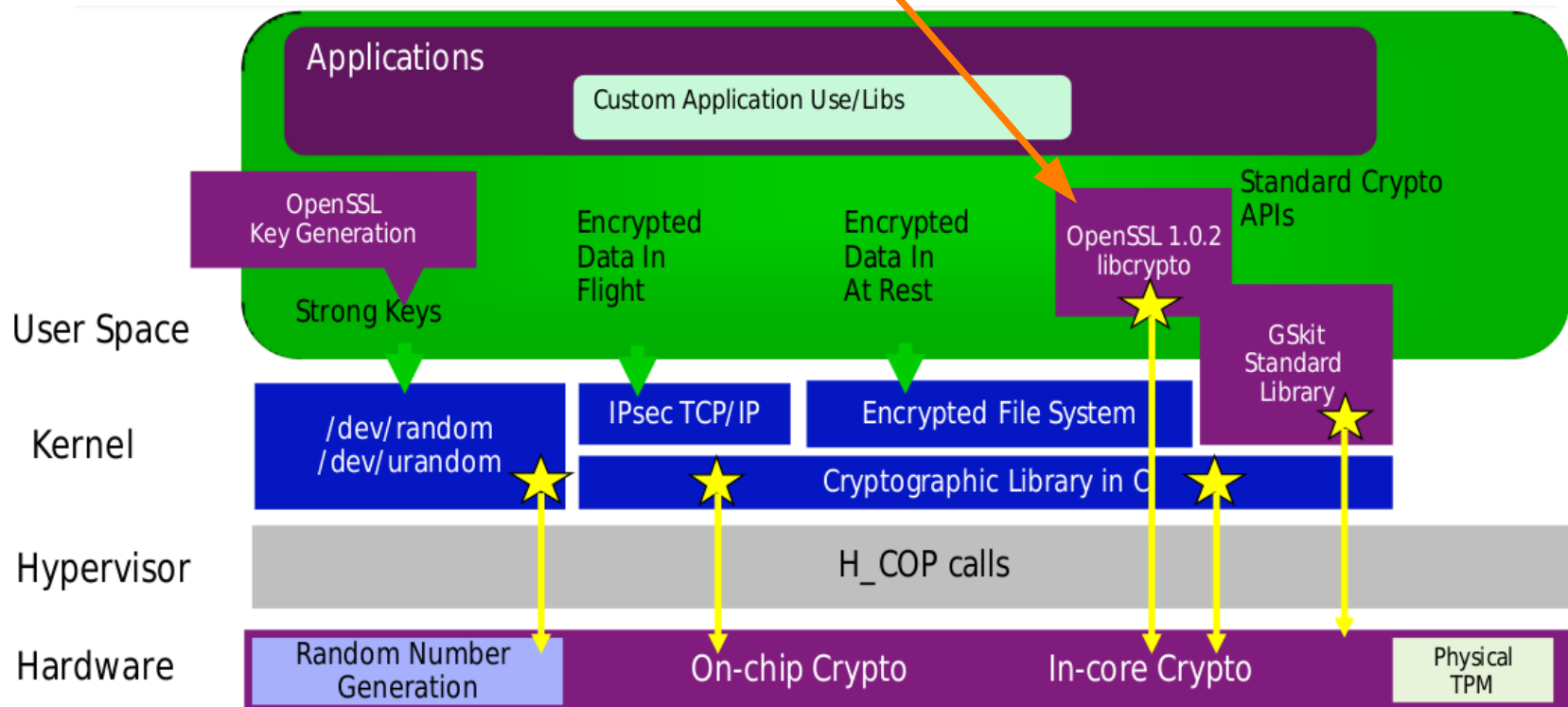
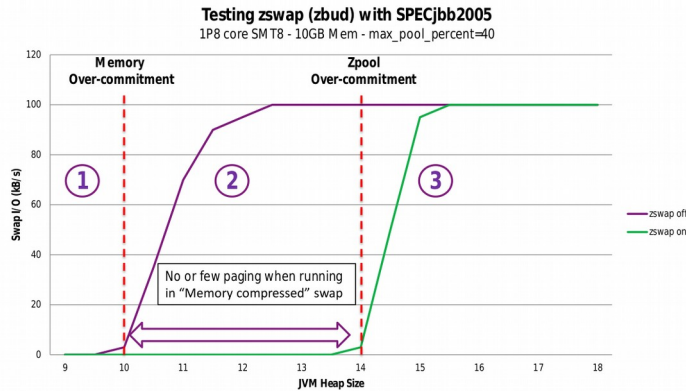


Bild: <http://de.slideshare.net/sebastienchabrolles/enabling-power-8-advanced-features-on-linux>

1) https://bugzilla.redhat.com/show_bug.cgi?id=1276310

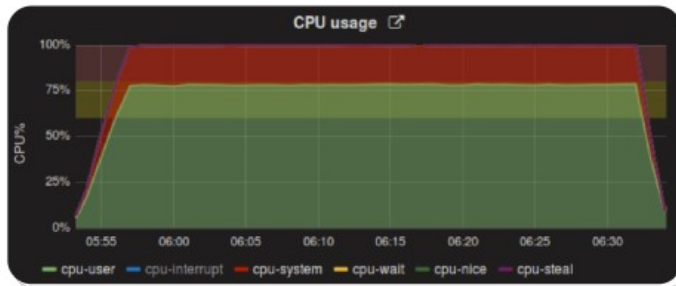
2) <https://packages.debian.org/stretch/openssl>



zswap ab Kernel 3.11¹

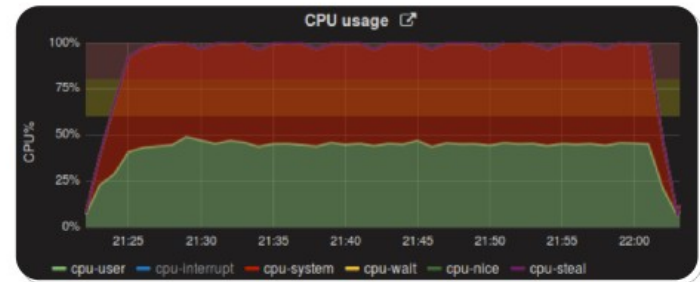
POWER HW Support (NX-842) ab Kernel 4.2²

Zswap HW compression 842



- 10GB RAM , 14GB Java Heap Size
- 25% of System CPU (overhead) due to memory page scanning.
- Compression offloaded to NX 842

Zswap Soft. Compression LZ0



- 10GB RAM , 14GB Java Heap Size
- 50% of system CPU (overhead) due to memory page scanning **and compression**

Bilder: <http://de.slideshare.net/sebastienchabrolles/enabling-power-8-advanced-features-on-linux>

1) https://www.thomas-krenn.com/de/wiki/Linux_Kernel_Versionen#Kernel_3.11

2) https://www.thomas-krenn.com/de/wiki/Linux_Kernel_Versionen#Kernel_4.2

CPU Vergleich

	POWER8	Xeon 2600 v4
Cores	8 / 10 / 12	4 ... 22
Multithreading	8x SMT	2x HT
Threads	64 / 80 / 96	8 ... 44
L1 Cache (Instruktionen)	32 KB	32 KB
L1 Cache (Daten)	64 KB	32 KB
L2 Cache	512 KB je Core (4 / 5 / 6 MB)	256 KB je Core (1 ... 5,5 MB)
L3 Cache	8 MB je Core (64 / 80 / 96 MB)	2,5 MB je Core (10 ... 55 MB)
L4 Cache	(Centaur Chips)	
IDLE Energie Optimierung	nein	ja

Current Members



Platinum Level



Silver Level



...



...

Agenda:

- POWER Architektur
- OpenPOWER Firmware
- RAS-Features: Reliability, Availability, Serviceability
- Datenbanken
- Virtualisierung
- OpenPOWER von Thomas-Krenn

BIOS? UEFI?

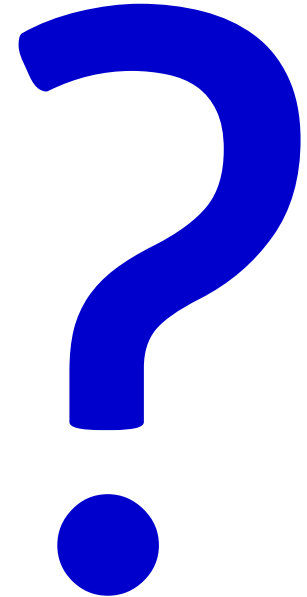
Aptio Setup Utility - Copyright (C) 2012 American Megatrends, Inc.

Main Advanced Event Logs IPMI Boot Security Exit

System Date	[Mon 09/15/2014]	Set the Date. Use Tab to switch between Date elements.
System Time	[14:11:50]	
Supermicro X9SCL(+)/X9SCM		
Version	2.0c	
Build Date	10/17/2013	
Processor		
Intel(R) Xeon(R) CPU E3-1220 V2 @ 3.10GHz		
Speed	3100 MHz	
Physical Count	1	
Logical Count	4	
System Memory		
Size	8192 MB	

++: Select Screen
↑↓: Select Item
Enter: Select
+/-: Change Opt.
F1: General Help
F2: Previous Values
F3: Optimized Defaults
F4: Save & Exit
ESC: Exit

Version 2.15.1234. Copyright (C) 2012 American Megatrends, Inc.




OpenPOWER Firmware

GitHub.com/open-power

OpenPOWER Firmware

GitHub, Inc. (US) | <https://github.com/open-power> Search

Personal Open source Business Explore Pricing Blog Support This organization Search Sign in Sign up

 **open-power**
<http://openpowerfoundation.org>

Repositories People 15


Filters Find a repository...

serverwiz Java ★ 1 📄 11
Updated 12 hours ago

common-mrw-xml ★ 0 📄 0
Updated 20 hours ago

op-build Makefile ★ 22 📄 85

People 15 >

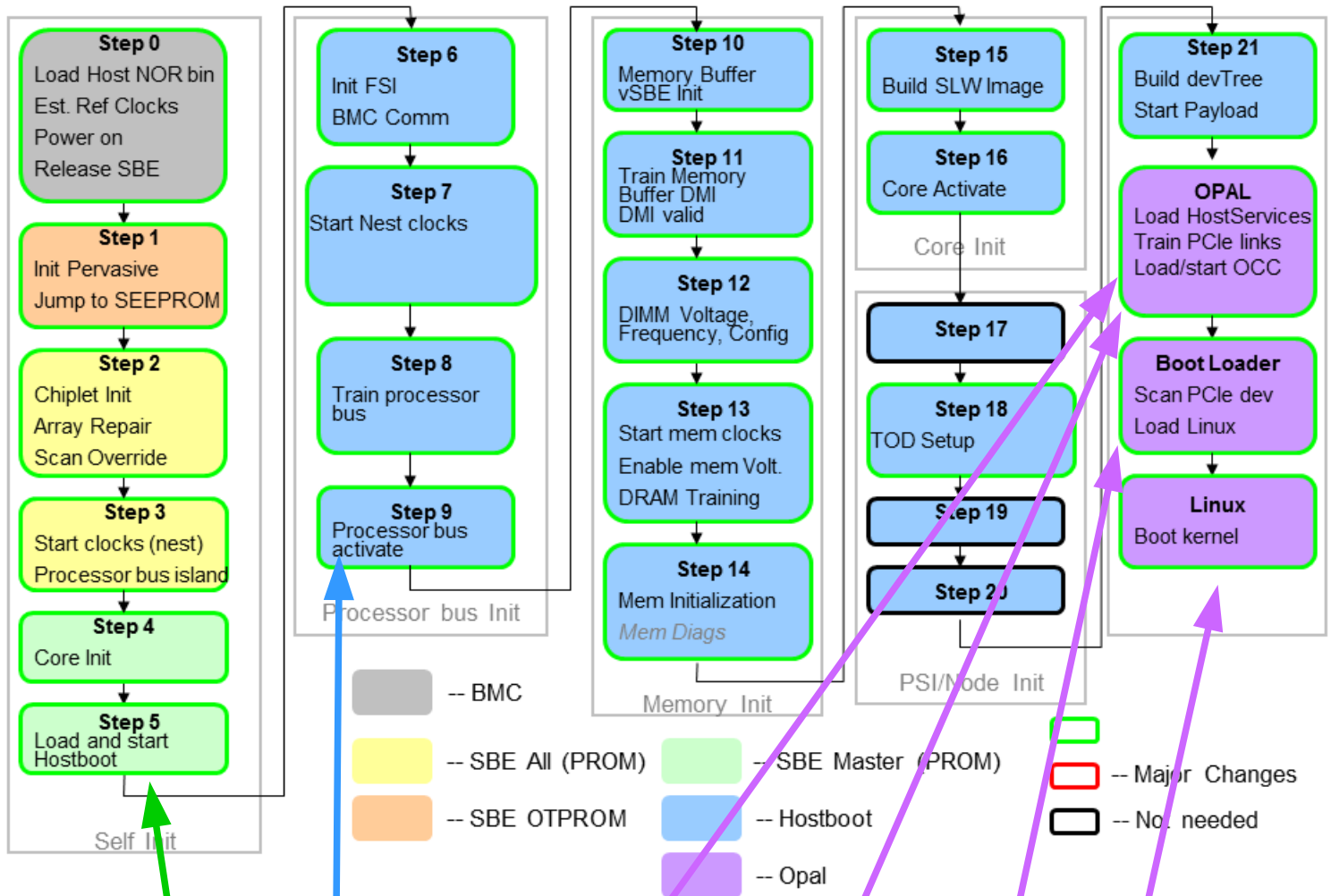


OpenPOWER Boot Prozess

- Self Boot Engine (SBE, ISTEPs 1-4)
- HostBoot (ISTEP 5-21), z.B. ECC leeren
 - microkernel, has userspace
 - CPU bus init, memory init, core init,
- On Chip Controller (OCC) – hard hw limits
 - PowerPC 405 core, hat eigenes Realtime-OS
- SkiBoot (OPAL)
- Linux / Petitboot (Bootloader)
- Betriebssystem (Linux)

(ISTEP = IPL Step)
(IPL = Initial Program Load)

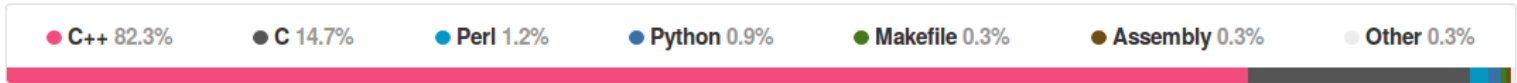
OpenPOWER Boot Prozess



- Self Boot Engine (SBE)
- HostBoot
- On Chip Controller (OCC)
- SkiBoot (OPAL)
- Linux / Petitboot (Bootloader)
- Betriebssystem (Linux)

OpenPOWER Boot Prozess

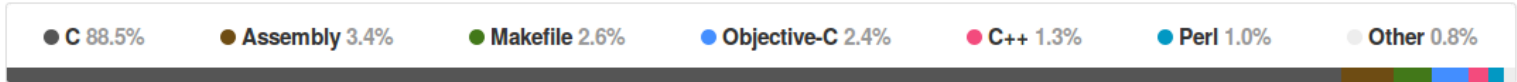
Self Boot Engine (SBE)



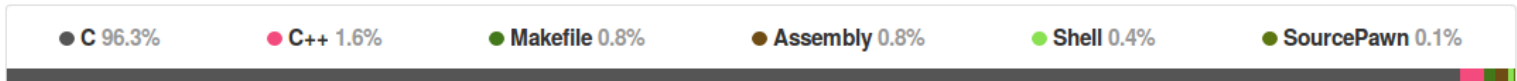
Hostboot



On Chip Controller (OCC)



Skiboot



Linux



Petitboot



TYAN TN71-BP012 PNOR V1.01

```
[Disk: sda2 / 3a71c024-b5ea-4031-98df-1c99beabbelc]
  Ubuntu, with Linux 4.4.0-31-generic (recovery mode)
  Ubuntu, with Linux 4.4.0-31-generic
  Ubuntu, with Linux 4.4.0-36-generic (recovery mode)
  Ubuntu, with Linux 4.4.0-36-generic
  Ubuntu
[Disk: sdb2 / 538e8e0d-84cb-4522-86c3-f2d7431923f6]
  Red Hat Enterprise Linux Server (0-rescue-ab029427b3fb4b87b79d81fba04c510c)
  Red Hat Enterprise Linux Server (3.10.0-327.el7.ppc64) 7.2 (Maipo)

System information
System configuration
Rescan devices
Retrieve config from URL
*Exit to shell
```

Enter=accept, e=edit, n=new, x=exit, h=help

Agenda:

- POWER Architektur
- OpenPOWER Firmware
- RAS-Features: Reliability, Availability, Serviceability
- Datenbanken
- Virtualisierung
- OpenPOWER von Thomas-Krenn

OpenPOWER RAS

An aerial photograph of a city skyline, likely Atlanta, Georgia, showing a dense urban area with numerous skyscrapers and a large railway yard in the center. The sky is clear and blue.

Reliability
Availability
Serviceability

„Everything has checksums“

- Adress- und Daten-Pfade der CPU sind durch Paritäts-Mechanismen oder Fehlerkorrekturverfahren (Error Correcting Codes, ECC) geschützt.
- Beispiel: ECC Fehler auf einer L2-Cache-Line
→ Deaktivierung dieser Line

„Processor Instruction Retry“

- Sorgt dafür, dass bei einem Soft-Fehler innerhalb eines Prozessorkerns der Prozessor die fehlgeschlagene Operation wiederholt.

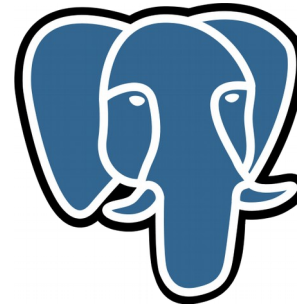
„Enhanced PCIe error handling“

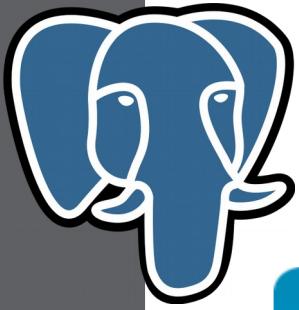
- Ausfall/Crash einer PCIe Karte → Recovery Mechanismus wird gestartet
- Re-Init zur Laufzeit
- Kein Server-Reboot erforderlich

Agenda:

- POWER Architektur
- OpenPOWER Firmware
- RAS-Features: Reliability, Availability, Serviceability
- Datenbanken
- Virtualisierung
- OpenPOWER von Thomas-Krenn

- SMT8 (8 symm. Threads pro Core)
 - Ermöglicht mehr parallele Transaktionen, erhöht Durchsatz
- Große Caches
 - Größerer „Arbeitsbereich“ am Chip, bevor Daten in RAM geschrieben werden → minimiert Latenz
- Optimierungen für die POWER-Plattform





The world's most advanced
open source database.

[Home](#) | [About](#) | [Download](#) | [Documentation](#) | [Community](#) | [Developers](#) | [Support](#) | [Your account](#)

- » [About](#)
- » [Advantages](#)
- » [Feature Matrix](#)
- » [Awards](#)
- » [Donate](#)
- » [Case Studies](#)
- » [Quotes](#)
- » [Featured Users](#)
- » [History](#)
- » [Sponsors](#)
- » [Servers](#)
- » [Latest news](#)
- » [Upcoming events](#)
- » [Press](#)
- » [Licence](#)

ppc64el packages now available on apt.postgresql.org

Posted on **2016-09-30**

PostgreSQL's repository for Debian and Ubuntu packages, apt.postgresql.org, has been extended to provide pre-built binary packages for a third architecture. ppc64el, the little-endian incarnation of IBM's POWER architecture, is joining the existing amd64 (64-bit x86 Intel) and i386 (32-bit x86 Intel) architectures.

The PostgreSQL 9.6 release is already included in the repository.

Just as on the existing architectures, all packages will be supported for all PostgreSQL major releases - which is currently all releases from 9.1 to 9.6. With only a few exceptions, all packages previously included in the repository have been compiled for ppc64el. [1]

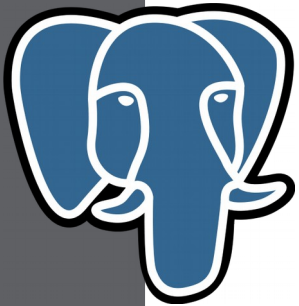
Supported Debian distributions are Debian 8 (Jessie) and unstable (Sid); supported Ubuntu releases are 14.04 (Trusty) and 16.04 (Xenial).

This initiative has been led by [2ndQuadrant](#) in cooperation with [IBM](#) Italy. Thanks must also be given to [credativ](#) for helping with the setup of the build infrastructure for the ppc64el architecture.

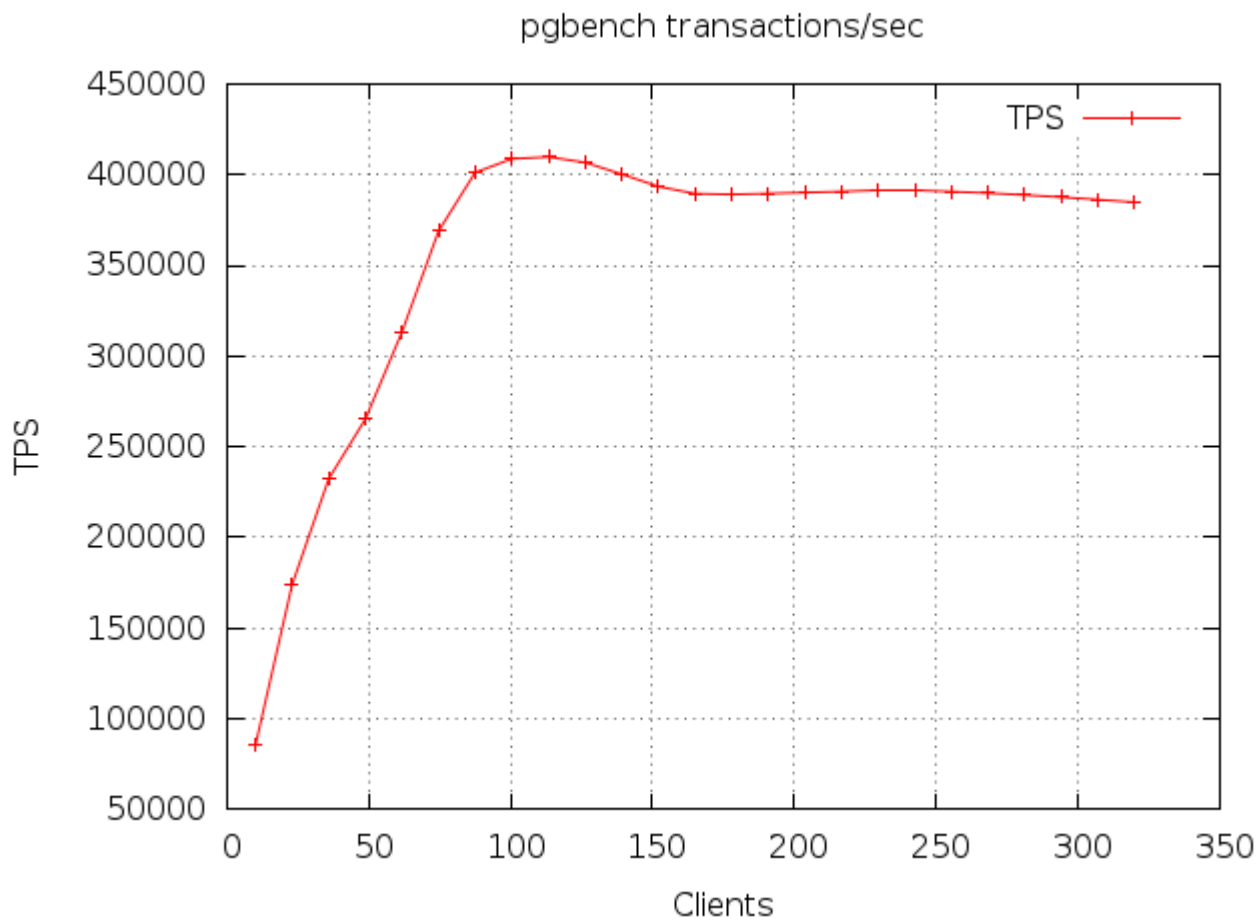
The ppc64el build host is provided by [IBM Power Systems Linux Center - Montpellier](#).

Kind Regards, Marco / Debian PostgreSQL Maintainers

[1] pg-partman is not available for 9.6 yet; psqlodbc and libpqtypes are not available on Trusty.



DB: PostgreSQL



Test Bernd Helmle (Technischer Leiter Datenbanken, creativ GmbH)
PostgreSQL 9.6.1 on powerpc64le, auf Thomas-Krenn OpenPOWER System (8-Core CPU, 64 Threads)
Linux debian 3.16.0-4-powerpc64le (Debian 8.6)

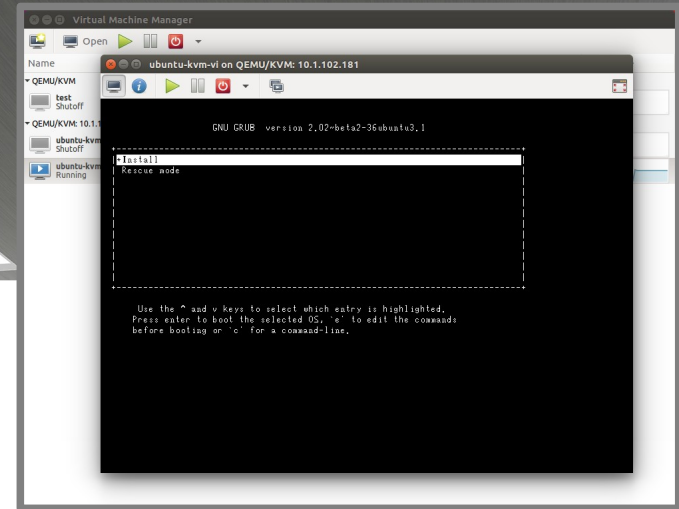
Agenda:

- POWER Architektur
- OpenPOWER Firmware
- RAS-Features: Reliability, Availability, Serviceability
- Datenbanken
- Virtualisierung
- OpenPOWER von Thomas-Krenn

- CPU: „zero measurable overhead“
- Memory: „zero measurable overhead“
- I/O kann limitierend wirken, PCIe Passthrough ist jedoch möglich¹

1) https://www.raptorengineering.com/TALOS/op_qemu_vfio.php

- KVM Virtualisierung
 - wie bei „normalen“ x86 Linux
- Netzwerk-Bridge
 - bnx2x Treiber
 - Linux Kernel 4.8 oder gepatchter Kernel 4.4 erforderlich
- Wiki-Artikel: KVM mit OpenPOWER¹



1) https://www.thomas-krenn.com/de/wiki/KVM_mit_OpenPOWER

Agenda:

- POWER Architektur
- OpenPOWER Firmware
- RAS-Features: Reliability, Availability, Serviceability
- Datenbanken
- Virtualisierung
- OpenPOWER von Thomas-Krenn

- POWER8 8-Core 3.325 GHz
- RAM: 8 x R-DDR3L 16GB = 128 GB
- 2 x 4 TB SATA
- 12 x Front-HDDs/SSDs (optional)
- NIC: 4 x 10 GBit



POWER von Thomas-Krenn



Open Source FW
(ohne BMC)



RAS Funktionen von
POWER8 / OPAL



Optimierte DB:
PostgreSQL, MariaDB



Hohe Speicher
Bandbreite



Für Big Data,
Hadoop, ...





wfischer@thomas-krenn.com