



Apache Bigtop

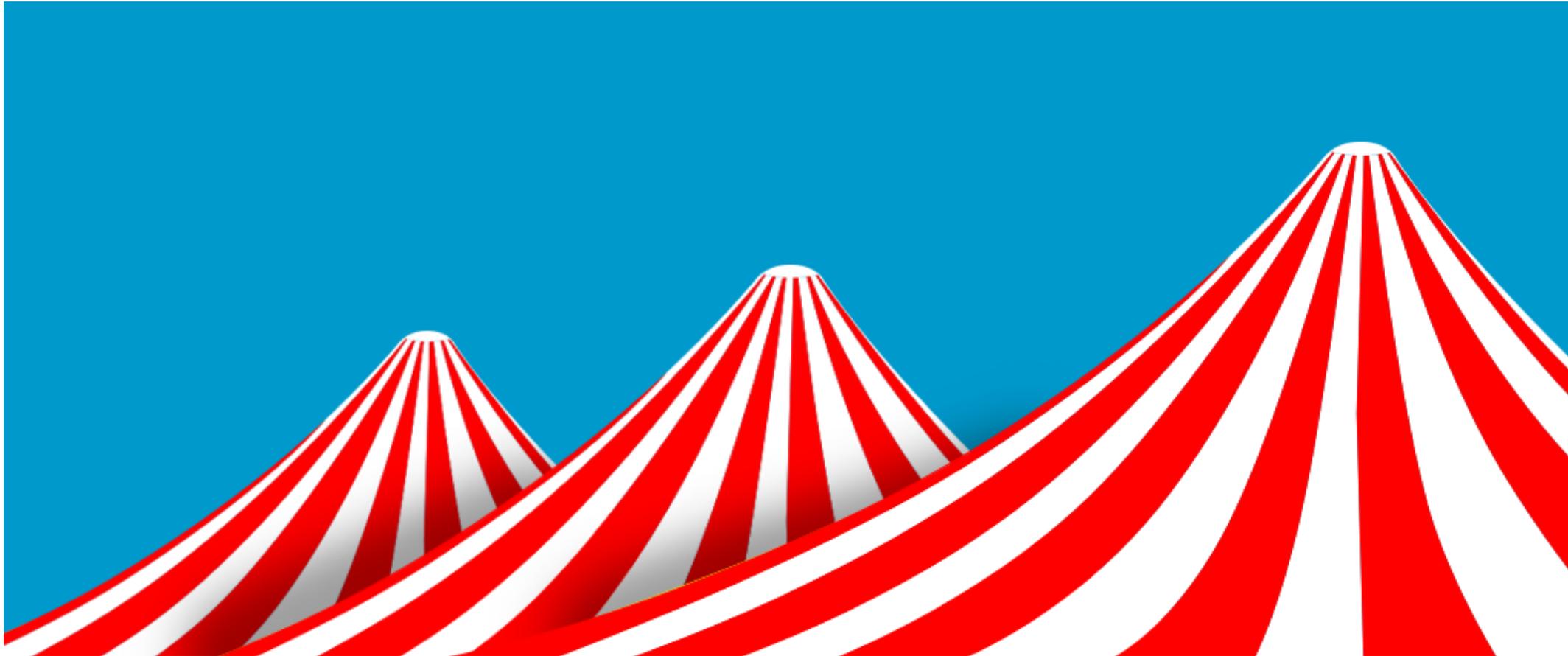
Dr. Olaf Flebbe

science + computing ag

IT-Dienstleistungen und Software für anspruchsvolle Rechnernetze

Tübingen | München | Berlin | Düsseldorf

Apache Bigtop



- Warum Bigdata ?
- Was hat das für Konsequenzen?
- Was heisst da Hadoop Ökosystem?
- Wer sind die Market Player?

Warum Big Data?

- Relationale Datenbanken kommen an Ihre Grenzen.
- Beispiel Index verwalten:
 - Festplatten Transaktionen (Durchs Filesystem): Seek/Sec: 80
 - Index aufbauen: Seek, Lesen, Schreiben: 1 Seek
 - 80 Datensätze pro Festplatte / sec. (7 Mio/Tag)
- Scale up wird schwierig: Ausweg Partitionierung, Sharding -> ACID wird immer aufwändiger zu implementieren
- Sehr aufwändig! -> Aufgeben der ACID Eigenschaften & Scale Out

Warum Bigdata

- Relationale Datenbanken kommen an Ihre Grenzen.
- Daten müssen an die Query Engine geschickt werden:
- Festplatte @ 1 TB Brutto, 24 Datenplatten => effektiv 20 TB lokal machbar
- SAN mit mehr Platten -> Bandbreiten Probleme

- Daten werden auf einem Cluster mit Commodity Hardware auf simplen Festplatten gelagert. (Data nodes genannt)
"Skalierbarkeit"
- Verteilte Algorithmen wie zum Beispiel Map/Reduce auf jedem Datanode (d.h. die Algorithmen kommen zu den Daten und nicht umgekehrt) **"Datenlokalität"**
- Durch Redundanz (mehrfaches Abspeichern) wird die Datensicherheit gewährleistet. **"Ausfallsicherheit"**

- HDFS ist ein Dateisystem im Userspace: HDFS
- HDFS: Metadatenervices "namenodes" und „datanodes“ als Datenspeicher
- HDFS Metadatenervice "namenode": Welcher Datensatz ist vorhanden, welche Permissions, wo liegen die Datei Fragmente.
- Für High Availability zusätzlich: Journalnode und ZKFC

Hadoop HDFS Befehle

- `hdfs dfs -ls <path>`
- `hdfs dfs -cp, -find, -mkdir`

- (oder `hadoop fs -ls ...`)

- Ein/Auslagern
- `hdfs dfs -copyFromLocal` , `hdfs dfs -copyToLocal`

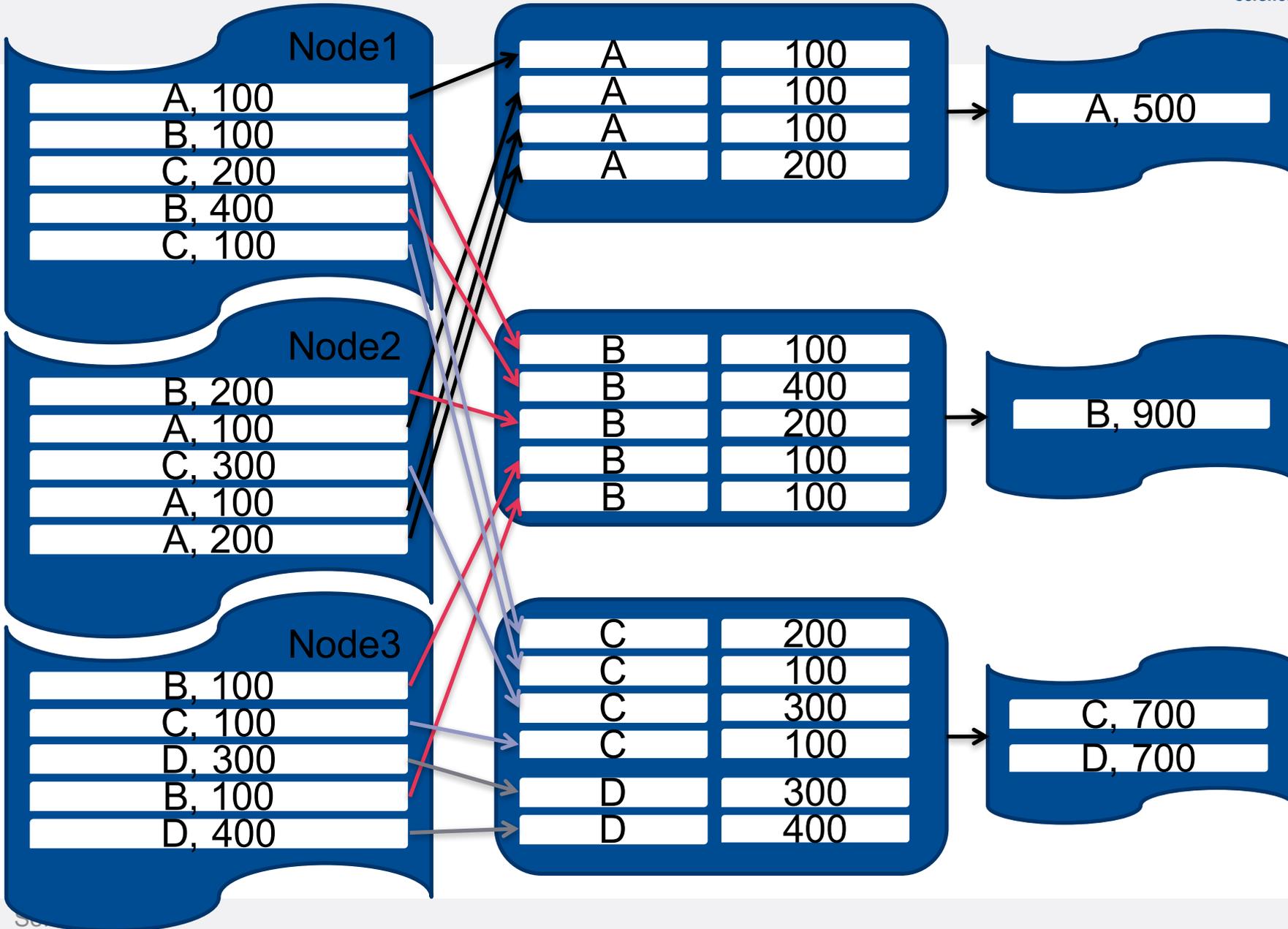
Hadoop: MapReduce

- Map Reduce ist ein Algorithmus Blaupause
 - Scatter/Gather auf Daten, gähn
- Beispiel Sortierte Summe eines Datenfiles mit zwei Spalten
 - SQL: "select eins, sum(zwei) from data group by eins"
 - über alle paare(eins,zwei)
sum[eins] += zwei

Beispiel: Sortierte Summe

Map

Reduce



Apache Bigtop

- Das „Debian“ der Big Data Distributionen
 - Enthält sehr viele verschiedene Apache Projekt und darüber hinaus.
 - Build Umgebung (Docker basiert)
 - Convenience Artifacts: Repositories mit Paketen
 - Support für Power8 (ppc64le) und beginnend ARM64 (aarch64)
 - Puppet Deployment Rezepte
 - Eine Smoke Test Umgebung
 - Data Generatoren
- 2015: Committer, PMC
- 2016 V.P., Apache Bigtop

- Yet Another Resource Negotiator
 - Resource Manager (RM) Zuweisung von Jobs auf Nodes
 - Datenlokalität
 - Verfügbaren Memory
 - Auslastung
 - Retries auf anderen Nodes
 - Application Manager (AM): Plugin Konzept
 - Dynamische Job Generierung und Platzierung
 - Kann im HDFS plaziert sein (Apache Tez)
 - Node Manager:
 - Started die Jobs (Im Linux Container)

Was ist Apache Bigtop

zookeeper

hadoop

ignite-hadoop

hbase

pig

hive

tez

sqoop

sqoop2

oozie

mahout

flume

giraph

hue

datafu

solr

crunch

kafka

ycsb

kite

hama

zeppelin

tajo

apex

qfs

spark

flink

phoenix

bigtop-groovy

bigtop-utils

bigtop-jsvc

bigtop-tomcat

alluxio

- Wurzeln aus Yahoo & Cloudera
- 2011: Cloudera 4 repo re-engineering als Open Source Projekt Bigtop
- Seit 2012 Apache Software Foundation TLP
- Release 1.0 in 2015

Wie kam ich zu Apache Bigtop

- Brauchte einen Open Source Stack d.h. incl. Build
 - Brauchte eine Deploy Mechanismus um die Pakete im Cluster zu installieren
 - Brauchte Kerberisierte Umgebung
 - Brauchte überhaupt eine Automatisierung
-
- Compile Toolchain Reproduzierbar,
 - Puppet Toolchain zum Aufsetzen der Compile Rechner

Aber ..

- Es gab ubuntu, aber kein Debian.
- Debian Jessie kam gerade heraus.
- Compile Toolchain war nicht reproduzierbar
- Ich hab das Git repository nicht gefunden
- Es gab make und gradle. Ein Teil der CI skripte lief mit make der andere mit gradle

Eins nach dem Anderen

- Konnte die Infrastruktur Probleme eine nach dem Anderen lösen und die Committer waren auch sehr dankbar, das diese gelöst wurden
- Anderer Contributor Evans Ye, brauchten offensichtlich auch eine Deploy Umgebung
- Puppet Module gegenseitig reviewed und die anderen Committer vor uns hergetrieben

Einordnung von Bigtop

- Kommerzielle Distributionen
 - Free to use, Support and added Value \$\$
 - Cloudera
 - Hortonworks (ex Yahoo)
 - MapR
 - Pivotal, Teradata, IBM, Trend Micro
- Freie Distribution
 - Apache Bigtop
- Cloud Angebote
 - Amazon Elastic MapReduce
 - Google DataProc
 - Azure

cloudera[®]
Ask Bigger Questions



Cloudera

- Die erste Hadoop Distribution
- CDH 5 sieht noch sehr nach Apache Bigtop aus

Hortonworks

- Enthält bigtop_* Pakete.
- Mir unklar wie der Zusammenhang
- Ist Mitglied der ODPi

Pivotal, IBM, Teradata, ... ODPi

- ... hatten alle mal eigene Distributionen
- Open Data Platform Initiative (odpi.org) Mitglieder haben sie ihr Distros weggeworfen und verwenden jetzt... Hortonworks
- "ODPi Kompatibel." .. heisst Apache Bigtop Kompatibel, API mässig.
- Gold Sponsor Apache Foundation.
- Die Standards basieren auf ... Apache Bigtop

Amazons Elastic Mapreduce

- Amazons Managed Hadoop
- Amazon EMR Entwickler contributen zu Apache Bigtop:
 - Pakete und Puppet Manifeste

Google Dataproc

- Baut offensichtlich auf Apache Bigtop 1.1.0 für Debian auf
 - Mein Ding ;-)
- Mehr als die Hälfte der Pakete ist Release 1.1.0

Orchestrierung:

- Ubuntu juju:
 - Juju is an application and service modelling tool that enables you to quickly model, configure, deploy and manage applications in the cloud with only a few commands. Use it to deploy hundreds of pre-configured services, OpenStack or your own code to any public or private cloud.
 - Apache Bigtop Support geht gerade upstream.

- Gradle, jetzt auch bekannt von Android
- Übernimmt manche Gedanken von Maven (Convention over Coding)
- Eigene DSL aufbauend auf groovy, damit muss nicht alles in extra plugins ausgelagert werden

Bigtop Compile

- Es wird der gradle wrapper unterstützt. D.h. kein gradle Installation nötig.
- JDK 7 (oder JDK 8) wird benötigt.
- `./gradlew`
- BTW: Hadoop hat ein paar Leichen im Keller, weswegen es nicht richtig mit JDK8 compiliert.



```

: aass..
=XS22nai,>_ . .
=n-- +!!!"^^--
.vX>
.)e<o;.
.v2^-{S> .
..<de-..;)Sa,
.._aoX}:==>=-?Xo>, . .
. . _aaZe!^=><i=s+s;~*XXos,,_ . . .
.....=iisaa oXXZ Y!"- . v (=d=:nc-ls,--?SX#Xouass,,____.:. . . . .
=XXoXXXXXXXXXXXXXZUZXX21?!"^-.._au*^=u2` jX>.*a>,-"!Y1XSSX##Z ZX XXXX XXXX XxoX Xc
.{XXXXXX2*?!!"!^^---- .._aa2!^- =dX( .+XXc. -!lnas,,-----^"!!"!!!?YSX XXXX 2+
-"YSXXXo=. . _=sssaaavl!!- . _aXxe` )SXo>. --"?Yooouass_s,, _vXXXX2}-
-{XXZoai%*XXSSSX>.. . <uXXX2~ . {XXXXs,. . =dXXXXZ21ii%uXXXXe-
.<XXXXX%- <XXXXX1|==%vdXXXXXo;:. . vXXXXXXos =i|*XXXXX> -<XXXXX`
=SXXXZc . .nXXX2> ----=2XXX2^~"||}|"---{ZX XX1-- .:XXXXo; . )XXXX2`
.=XXXXZc nXXXX> =XXXxe.._s=>_...)XX XX1 . .:SXXXo; .)XXXX2..
<SXXXc . .nXXXS> =XXXosummmBmma,)ZX XX1 :XXXX2; )XXXX.
.<XXXXX( :nXXXS; <XXXXXm#mmmmWmmmmmoZX XX1 . .3XXXo;. .)XXXXX;
.nXXXX;. :XXXX; .=XXXXXmmmmBmmmmWmB#XXX XX1 . .nXXXX> :XXXXc .
.=oZXXXe; .<XXXX2` .)XXXXZmBmBmWmmmmW#2XX XX1 . .vXXXXc vXXXXo;
+Y3S2Zx _...vXXXXe . .)ZXXXZmmWmBmmBBm#XXX Xxo. . {XXXXz:._vSS2Y1=
. ---+""*!?!*Y1s|===uXSS SXZU XUXU XUXU XS2X X2n|_=_|| %Y*?!" "^^----

```

```

.o.
.888.
.8"888. oo.ooooo. .oooo. .ooooo. 888 .oo. .ooooo. 888 888oooo .ooooo. o888oo .ooooo. oo.ooooo.
.8' `888. 888' `88b`P )88b d88' `Y8 888P"Y88b d88' `88b 888oooo888' `888 888' `88b 888 d88' `88b 888' `88b
.88ooo8888. 888 888 .oP"888 888 888 888 888ooo888 888 `88b 888 888 888 888 888 888 888 888 888 888
.8' `888. 888 888d8( 888 888 .o8 888 888 888 .o 888 .88P 888 `88bod8P' 888 .888 888 888 888 888
o88o o8888o 888bod8P' `Y888""8o`Y8 bod8P' o8 88o o888o`Y8bod8P' o888bod8P' o888o`8ooooo. "888" `Y8bod8P' 888bod8P'
888 888 d" YD 888
o888o "Y88888P' o888o

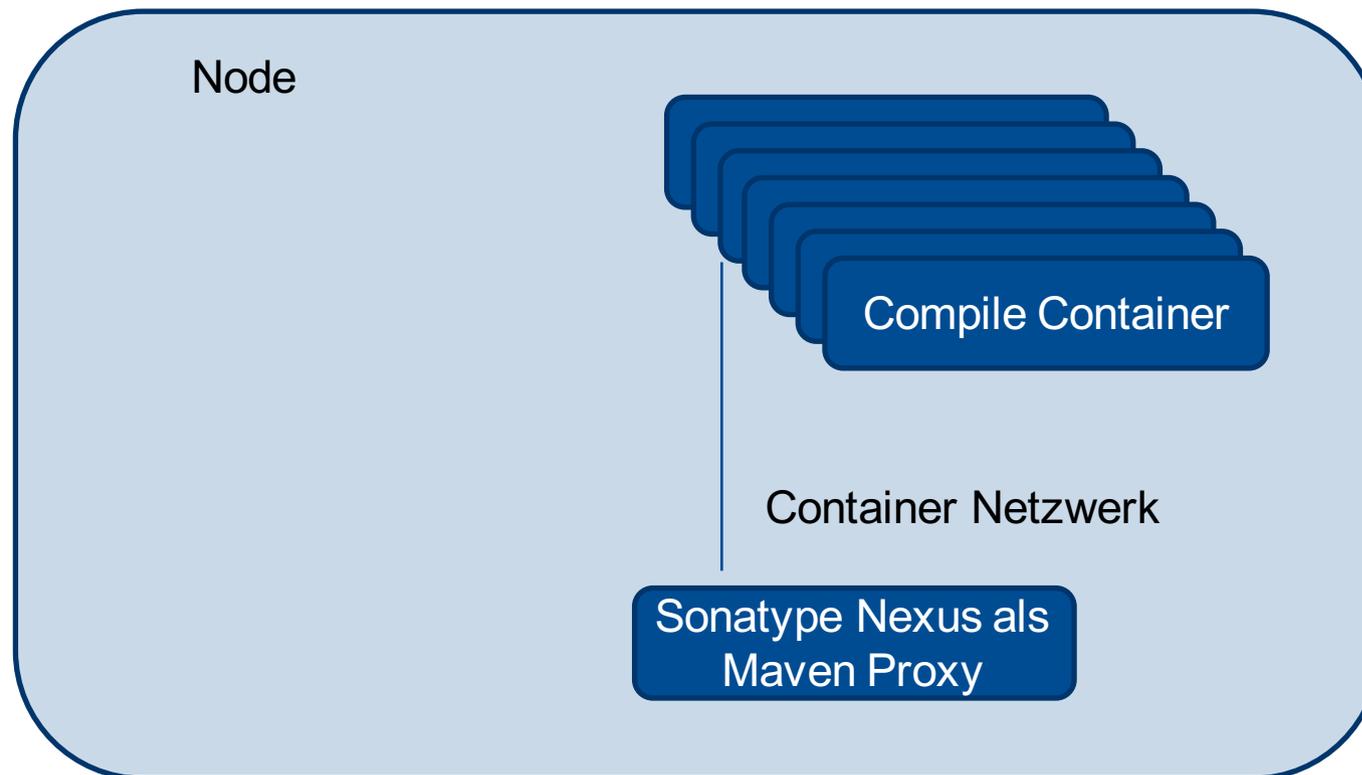
```

RPM und DEB Packages

- Alles unter bigtop-packages
- Steuerung durch ein Bill of Material BOM file bigtop.bom
- Ein einziges build script (do_component_build) pro Paket
- Wird von debuild als auch von rpmbuild aufgerufen + Metadata.
- Erzeugt Repositories
- Die Spezifischen Dateien sind in RPM bzw. DEB
- Verlinkt über Releases

- Docker für die Compile Umgebung (Erzeugung in docker)
- Jenkins Builds für Docker Matrix
- <https://ci.bigtop.apache.org/job/Bigtop-trunk-packages>
- Sonatype Nexus für wiederholte Downloads

Setup CI System



- PPC64LE Compiler an Bare metal, später Docker
- Probleme mit den Java Bibliotheken aus maven. Einige enthalten shared libraries die per Hand in die Jar eingefummelt sind

Deployment

- Puppet module bigtop-deploy/
- Gradlew deploy
- Gerade Umstellung von vagrant auf docker-compose



Vielen Dank für Ihre Aufmerksamkeit.