23.01.2017

18.18 – Hakkan linuxiga jälle tegelema.

Paneme folderitd kõik nähtavaks. Kustutasime ja shrinkisime foldereid. Tahaks suure restardi teha. Uurisin, et võimalik linux mittejärjesikustela partitsioonidele ühtlaselt installid .

Sisuliselt 2 võimalus

1) Kaotame kuidagi partitisooonid
2) Installime Linuxi mitmela partitsiooniless

01.02.2017

1. Kasutasin sellist programmin nagu EaseUS partitsioneerimiseks
2. Installisin Ubuntu USB-le http://www.howtogeek.com/howto/linux/create-a-bootable-ubuntu-usb-flash-drive-the-easy-way/
3. https://courses.cs.ut.ee/2016/os/fall/Main/Praktikum2 Hakkan selle järgi installima ubuntut
    a. Enne tutvun kindlamalt partitsioonidega

Partitsioonid

https://www.bleepingcomputer.com/tutorials/understanding-hard-disk-partitions/ - hea tutorial

In the current IBM PC architecture, there is a partition table in the drive's Master Boot Record (section of the hard drive that contains the commands necessary to start the operating system), This partition table is then **further split into 4 partition table entries**, with each entries corresponding to a partition. This partition table is then further split into 4 partition table entries, with each entries corresponding to a partition. To overcome this restriction, system developers decided to add a new type of partition called the **extended partition**. By replacing one of the four primary partitions with an extended partition, you can then make an additional **24 logical partitions** within the extended one.

Partition Table

Primary Partition #1

Primary Partition #2

Primary Partition #3

Primary Partition #4 (Extended Partition)

    Logical Partition #1

    Logical Partition #1

Each hard drive also has one of its possible 4 partitions flagged as an **active partition.** The active partition is a special flag assigned to only one partition on a hard drive that the Master Boot Record (MBR) uses to boot your computer into an operating system

As only one partition may be set as the active partition, you may be wondering how people can have multiple operating systems installed on different partitions, and yet still be able to use them all. This is accomplished by installing a boot loader in the active partition. When the computer starts, it will read the MBR and determine the partition that is flagged as active. This partition is the one that contains the boot loader

When the operating system boots off of this partition the boot loader will start and allow you to choose which operating systems you would like to boot from

 16:48 Mounting a fileSystem in Linux

https://www.bleepingcomputer.com/tutorials/introduction-to-mounting-filesystems-in-linux/

In order to access a filesystem in Linux you first need to mount it. Mounting a filesystem simply means making the particular filesystem accessible at a certain point in the Linux directory tree. When mounting a filesystem it does not matter if the filesystem is a hard disk partition, CD-ROM, floppy, or USB storage device.

You simply need to know the device name associated with the particular storage device and a directory you would like to mount it to.

. Having the ability to mount a new storage device at any point in the directory is very advantageous. For example, lets say that you have a web site stored in /usr/local/website. The web site has become very popular and you are running out of space on your 36 GB hard drive. You can simply go out and purchase a new 73 GB hard drive, install it in the computer, and then mount that entire drive as /usr/local/. Now your /usr/local mount point has a total hard drive space of 73 GB, and you can free up the old hard drive by copying everything from the old /usr/local to the new one. As you can see, adding more hard drive space to a computer, while still keeping the same exact directory structure, is now very easy.

16:54 Nüüd edasi tulevad käsud, seega vaatan, et saaksin usb pealt bootida.

http://www.isunshare.com/windows-password/four-methods-to-access-uefi-bios-setup.html

17.09 Sain käima UBUNTU TRY osa. Proovin nüüd ubuntu partitsionides aru saada.

https://help.ubuntu.com/community/PartitioningSchemes MINGI GPT jama teksti alguses. Uurin mis schemen mul on. **Partition Scheme** keyword

17.24 Installin Ubuntut eelnevalt mainitud tutoriali järgi.

Seal õpetatakse, et kui insallima hakkan siis

Boot jaoks 32 Swap jaoks 4 ja Home ülejäänud

17.29 – 15 mintsa paus.

02.02.2017

9.44

Newer computers come with a replacement firmware for the old BIOS system called UEFI (Unified Extensible Firmware interface), and GPT is a part of the UEFI standard. If you bought a Windows 8 computer, it's most certainly installed on a GPT partitioning scheme. If you're already running a Linux distribution or attempting to install one on a recent Windows 7/8 computer, the easiest method of finding out what partitioning standard is used is to launch a shell terminal and type **sudo fdisk -l** (leave out the *sudo* if the distribution does not use it).

sudo fdisk -l

Mul SDA -l gpt

File Systems: Before a disk partition can be used to store data, it must first be formatted. The formatting process includes stamping it with a file system.

10.27 – Peaks natuke tegelema mõistetega ja läbi mõtlemiseg..

GPT MBR

Partitsiioonid linuxis

LVM linuxis

Linuxi installimine on võimalik ilma LVM-ita Võimalik koos LVM-iga.

Linuxil on mount pointid . Alguseks oluline installida boot, home ja swap.
Failisüssteem kas ext 4 või bdfs

Põhiküsimused

1. Kas kasutada LVM

Mis on grub ja mis on bootloadeR This partition is the one that contains the boot
loader

When the operating system boots off of this partition the boot loader will start and
allow you to choose which operating systems you would like to boot from

2.
3.

10.52

1. 1. Kas kasutada LVM – EI
2. Mis on EFI, KAS MUL ON VAJA BOOTLOADERI JAOKS MIDAGI ERALDI
   a. Kas on efit vaja And if you're attempting to set up a dual-boot system
      between Windows 10 and Ubuntu 15.10 on a single hard drive, you'll
      be creating just three partitions – each mounted at / (root), /home
      and swap. That means creating an EFI System Partition (ESP), a root,
      /home and swap partitions. The /home partition is optional but
      recommended. => Seega minul ei ole vaja

http://www.tecmint.com/ubuntu-16-04-installation-guide/#C3

http://linuxbsdos.com/2015/10/30/gpt-and-mbr-manual-disk-partitioning-guide-for-ubuntu-15-10/

Boot loader i

http://askubuntu.com/questions/326662/which-partition-to-select-as-device-for-boot-loader-in-an-efi-mode-install

https://ubuntuforums.org/showthread.php?t=2309806

Ubutntu installitud valisin efi bootloaderiks. Praefu peab bioses manuaalselt järjekorda vahetama.

04.02.2017

Apt-get update

Apt-get install ssh

11.09

SSH -ga tegelen .

Praegu ühendatud läbi kahe arvuti. Kuidas faile transportida?

11:33 Avasin tutoriali 1

11.57 -Definition of hypervisor.

On 2-te tüüpi.

Type-1, native or bare-metal hypervisors

These hypervisors run directly on the host's hardware to control the hardware and to manage guest operating systems. For this reason, they are sometimes called bare metal hypervisors. The first hypervisors, which IBM developed in the 1960s, were native hypervisors.[4] These included the test software SIMMON and the CP/CMS operating system (the predecessor of IBM's z/VM). Modern equivalents include Xen, Oracle VM Server for SPARC, Oracle VM Server for x86, the Citrix XenServer, Microsoft Hyper-V and VMware ESX/ESXi.

Type-2 or hosted hypervisors

These hypervisors run on a conventional operating system just as other computer programs do. A guest operating system runs as a process on the host. Type-2 hypervisors abstract guest operating systems from the host operating system. VMware Workstation, VMware Player, VirtualBox, Parallels Desktop for Mac and QEMU are examples of type-2 hypervisors.

https://en.wikipedia.org/wiki/Hypervisor

https://en.wikipedia.org/wiki/Xen

11.37 – Vahetasin parooli. Esimene ühendus tehtud

Tahaks rohkem teada kuidas ühendused toimivad.

1) Try to connect in LAN with gateguard

Exactly like in tutorial with putty .

1

Gatekeeper.cs.hku.hk Port: 22

anamm

AedrMDwH

2

ssh student@202.45.128.135

3

ssh student@student13 tundis ilusti ära

2) Proovin ühenda võrgus ilma gateguardita. - EI

202.45.128.135

student@202.45.128.135

3) Proovin ühendada võrgust väljas ilma gateguardita, ilma vpn-ita - EI
4) Proovin ühendada võrgust väljas gateguardita ja vpn-iga   - SOBIB

ssh student@202.45.128.135

⇨ Ilmselt on vaja olla CS department vpn-is

5) Proovin ühendada võrgust väljas gateguardi ja vpn-iga SOBI
6) Proovin ühendada VÕRGUST VÄLJAS ilma vpnita gateguardiga SOBIB

Täpselt sama, kui kõik muu

13.28 FAILIDE LIIGUTAMINE TOIMUB VAID LÄBI VPN-I SOBIB.

16.27

See teine vahelüli võrgus

- NAT https://en.wikipedia.org/wiki/Network_address_translation
  http://whatismyipaddress.com/nat

Network address translation (NAT) is a method of remapping one IP address space into another by modifying network address information in Internet Protocol (IP) datagram packet headers coc server

- DHCP server https://technet.microsoft.com/en-us/library/dd145320(v=ws.10).aspx Automatic ip address mappin
- TFTP server FILE RELATED

17.04

Part I Specification
sudo lshw -short

lshw(Hardware Lister): a small tool to provide detailed information on the hardware configuration of the machine

H/W path	Device	Class	Description
		system	HP Compaq Elite 8300 SFF (QV996AV)
/0		bus	3397
/0/0		memory	64KiB BIOS
/0/4		memory	256KiB L1 cache
/0/5		memory	1MiB L2 cache
/0/6		memory	6MiB L3 cache
/0/7		**memory**	**16GiB System Memory**
/0/7/0		memory	DIMM [empty]

/0/7/1   memory   8GiB DIMM DDR3 Synchronous 1600 MHz (0.6 ns)

http://www.pcmag.com/article2/0,2817,2400801,00.asp

/0/7/2   memory   DIMM [empty]

/0/7/3   memory   8GiB DIMM DDR3 Synchronous 1600 MHz (0.6 ns)

/0/10   processor   Intel(R) Core(TM) i5-3570 CPU @ 3.40GHz

https://arstechnica.com/gadgets/2016/02/pentium-core-i5-core-i7-making-sense-of-intels-convoluted-cpu-lineup/

These are all quad-core CPUs without **Hyperthreading,**

/0/100   bridge   Xeon E3-1200 v2/3rd Gen Core processor DRAM Controller

/0/100/2   display   Xeon E3-1200 v2/3rd Gen Core processor Graphics Controller

/0/100/14   bus   7 Series/C210 Series Chipset Family USB xHCI Host Controller

/0/100/19   eth0   network   82579LM Gigabit Network Connection

/0/100/1a   bus   7 Series/C210 Series Chipset Family USB Enhanced Host Controller #2

/0/100/1b   multimedia 7 Series/C210 Series Chipset Family High Definition Audio Controller

/0/100/1d   bus   7 Series/C210 Series Chipset Family USB Enhanced Host Controller #1

/0/100/1e   bridge   82801 PCI Bridge

/0/100/1f   bridge   Q77 Express Chipset LPC Controller

/0/100/1f.2   storage   7 Series/C210 Series Chipset Family 6-port SATA Controller [AHCI mode]

/0/100/1f.3   bus 7 Series/C210 Series Chipset Family SMBus Controller

/0/1   scsi0 storage

/0/1/0.0.0 /dev/sda disk 500GB ST500DM002-1BD14

/0/1/0.0.0/1 /dev/sda1 volume 449GiB EXT4 volume

/0/1/0.0.0/2 /dev/sda2 volume 15GiB Extended partition

/0/1/0.0.0/2/5 /dev/sda5 volume 15GiB Linux swap / Solaris partition

/0/2   scsi2 storage

 /0/2/0.0.0 /dev/cdrom disk DVD-RAM GH80N

QUESTION – DOES IT HAVE EFI OR MBR – Probably mbr, because we use an extenden partition

cat /proc/cpuinfo

cat /proc/cpuinfo | grep processor | wc -l

You need to use Proc (/proc) file system provides information about CPU and their speed which is a pseudo-filesystem https://www.cyberciti.biz/faq/linux-display-cpu-information-number-of-cpus-and-their-speed/

lcpuhy it shows multiple cpu-s in this file

http://unix.stackexchange.com/questions/146051/number-of-processors-in-proc-cpuinfo

The words "CPU", "processor" and "core" are used in somewhat confusing ways. They refer to the processor architecture.

 A core is the smallest independent unit that implements a general-purpose processor;

A processor is an assemblage of cores (on some ARM systems, a processor is an assemblage of clusters which themselves are assemblages of cores).

A chip can contain one or more processors (x86 chips contain a single processor, in this sense of the word *processor*).

Hyperthreading means that some parts of a core are duplicated. A core with hyperthreading is sometimes presented as an assemblage of two "virtual cores"

In software manuals, terms *CPU* and *processor* are used to mean any one piece of hardware that executes program code. In hardware terms, this means one core, or one virtual core with hyperthreading.

Hardware configuration in STUDENT 13

CPU model: Intel(R) Core(TM) i5-3570 CPU @ 3.40GHz

RAM size: memory   16GiB System Memory

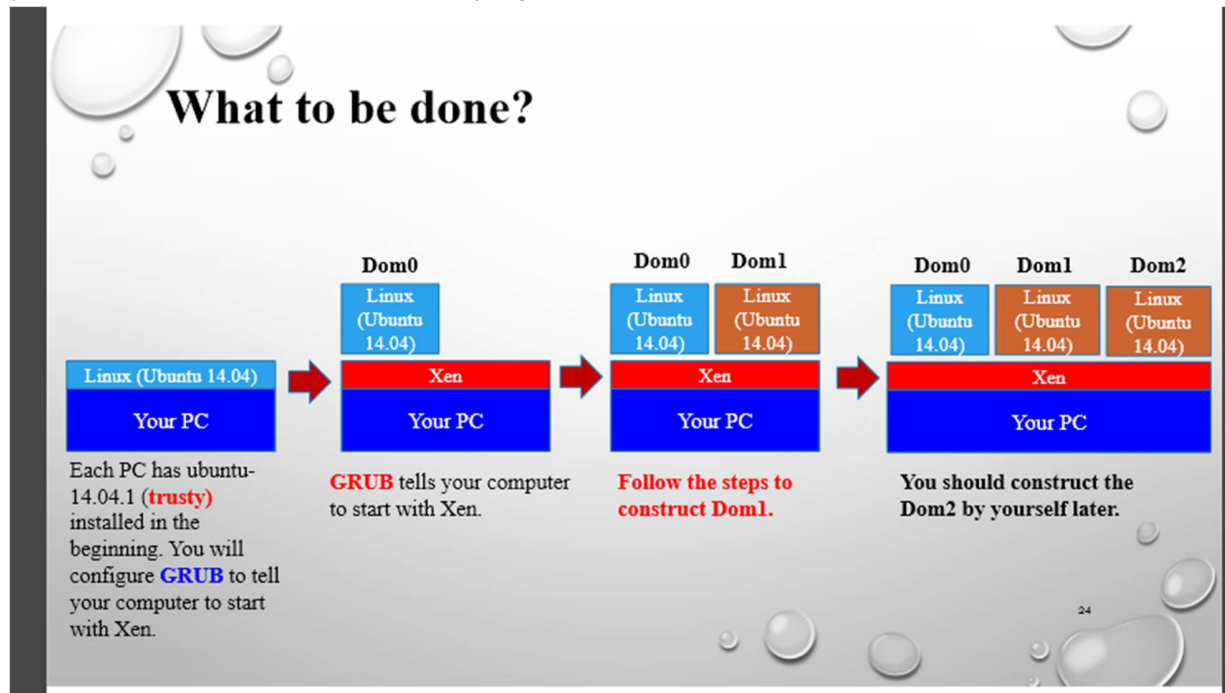**# of cores : 4** No hyperthreading

XEN installation
18:38

1. Started with the slides
2. Started with the video in moodle
3. Went on to tutorial
   https://wiki.xenproject.org/wiki/Xen_Project_Beginners_Guide

18.44

punane nool siin slaidil tähistab järgmisele sammule liikumist.

05.02.2015

Network switch https://en.wikipedia.org/wiki/Network_switch Does stuff in OSI layer 2

https://www.quora.com/What-is-the-difference-between-HTTP-protocol-and-TCP-protocol

19.05

sudo apt-get update

sudo apt-get upgrade

xen

sudo apt-get install xen-hypervisor-amd64

vim

sudo apt-get install vim

READING ABOUT BOOTLOADER

https://en.wikipedia.org/wiki/GNU_GRUB

http://www.howtogeek.com/196655/how-to-configure-the-grub2-boot-loaders-settings/

Distribution info

lsb_release -a

Distributor ID: Ubuntu

Description: Ubuntu 14.04.5 LTS

Release: 14.04

Codename: trusty

GRUB settings

sudo vim /etc/default/grub

https://www.gnu.org/software/grub/manual/html_node/Simple-configuration.html

GRUB_CMDLINE_LINUX'

 Command-line arguments to add to menu entries for the Linux kernel.

Do a copy first

sudo cp /etc/default/grub cp_grubfile

sudo

http://www.openvim.com/

Väike vim tuts

sudo update-grub

Network configuration

virtual switch – like a physical switch, but virtual - Dom0(student10) that takes packets from the virtual machines (Dom1 & Dom2) and forwards them onto the

physical network so they can see the Internet and other machines on your network.

We use linux bridge for this

https://wiki.xenproject.org/wiki/Xen_Project_Beginners_Guide

Uurin, mis varasemalt valesti tehti.

0:07.36 xen install videost

18.10

sudo apt-get install bridge-utils

sudo vim /etc/network/interfaces

sudo reboot

Checkpoint 3: I am creating a new VM! Choice. Creating your first VM (up to Xen slides p.44).

sudo xl list

ifconfig

/etc/network/interfaces fail lähemalt

https://wiki.xenproject.org/wiki/Xen_Project_Beginners_Guide#Setup_Linux_Bridge_for_guest_networking

eth0 is the first interface

A *network interface* is the point of interconnection between a computer and a private or public network. A network interface is generally a network interface card (NIC), but does not have to have a physical form. Instead, the network interface can be implemented in software. For example, the loopback interface (127.0.0.1 for IPv4 and ::1 for IPv6) is not a physical device but a piece of software simulating a network interface. The loopback interface is commonly used in test environments.

http://www.computerhope.com/unix/uifconfi.htm

- **eth0** is the first ethernet interface. (Additional ethernet interfaces would be named **eth1**, **eth2**, etc.) This type of interface is usually a [NIC](NIC) connected to the network by a category 5 cable.
- **lo** is the loopback interface. This is a special network interface that the system uses to communicate with itself.
- **wlan0** is the name of the first wireless network interface on the system. Additional wireless interfaces would be named **wlan1**, **wlan2**, etc.

dom0 is a host

sudo apt-get install xen-utils-4.4 xenwatchxen-toolsxen-utils-common xenstore-utilsvirtinstvirt-viewer virt-manager

// Tahtsin endale teha koopiad enda arvutisse nendest failidest aga järelikult ei saa

https://winscp.net/eng/docs/guide_tunnel , sest mitmne ühendus ☹ ssh hop, ssh tunneling

0.17.21 – Järgnev tegevus algab siit

Xen daemon configuration file

sudo vim /etc/xen/xend-config.sxp

 In multitasking computer operating systems, a **daemon** (/ˈdiːmən/ or /ˈdeɪmən/) is a computer program that runs as a background process, rather than being under the direct control of an interactive user.

CLI – command line interface

Communication within a host
https://troydhanson.github.io/network/Unix_domain_sockets.html

xend-unix-server: A boolean value that tells xend whether or not to start the unixdomain socket management server. This is required for the CLI tools to operate. Default isyes.

UNIX **domain sockets** are a method by which processes on the same host can communicate. Communication is bidirectional with stream sockets and unidirectional with datagram sockets.

fd = socket(AF_UNIX, SOCK_STREAM, 0);

Identity

Instead of identifying a server by an IP address and port, a UNIX domain socket is known by a pathname. Obviously the client and server have to agree on the pathname for them to find each other. The server binds the pathname to the socket:

sudo mkdir /home/xen/

sudo chmod 777 -R /home/xen

sudo ln -s /usr/lib/xen-4.4 /usr/lib/xen

07.02.2017

0.19.05 – Järgnev tegevus algab siit

Set Default Virtual Machine Configuration.

/etc/xen-tools/xen-tools.conf: This file contains the default values that are used by thexen-createimagescript unless you specify other values on the command line.

sudo vim /etc/xen-tools/xen-tools.conf

vahetasin väga palju settinguid

gateway   10.42.0.1 https://www.lifewire.com/definition-of-gateway-817891 Tavavõrkudes rooter

netmask 255.255.254.0 http://www.computerhope.com/jargon/n/netmask.htm To divide the network into available hosts

        network address from host address - https://en.wikipedia.org/wiki/Host_address

        ip address https://en.wikipedia.org/wiki/IP_address -network or host adressing , location

        Ip addres allocation - with DHCP https://en.wikipedia.org/wiki/Dynamic_Host_Configuration_Protocol

https://support.microsoft.com/en-hk/help/164015/understanding-tcp-ip-addressing-and-subnetting-basics Üllatavalt hea kirjeldus eelnevast teemast

adressing

broadcast 10.42.0.1 A **broadcast address** is a logical **address** at which all devices connected to a multiple-access communications network are enabled to receive datagrams. A message sent to a **broadcast address** is typically received by all network-attached hosts, rather than by a specific host. https://en.wikipedia.org/wiki/Broadcast_address

nameserver **Nameserver** is a server on the internet specialized in handling queries regarding the location of a domain name's various services. **Nameservers** are a fundamental part of the Domain Name System (DNS). They allow using domains instead of IP addresses.

https://www.namecheap.com/support/knowledgebase/article.aspx/9434/32/using-default-nameservers-vs-hosting-nameservers

12:59

sudo xen-create-image --hostname=student13-x1

Check the log in real time

sudo tail -f /var/log/xen-tools/student13-x1.log

installation completed

Logfile produced at:

/var/log/xen-tools/student13-x1.log

Installation Summary

Hostname   : student13-x1

Distribution   : trusty

MAC Address   : 00:16:3E:22:B1:1D

IP Address(es) : dynamic

RSA Fingerprint : 16:e5:e8:92:ee:e7:6a:c2:a1:76:48:b3:54:60:4e:e6

Root Password   : 8tmPZCJd

configurin vm

sudo vim /etc/xen/student13-x1.cfg changing the mac adress

start vm

sudo xl create /etc/xen/student13-x1.cfg -c

VM IS UP

sudo xl console student13-x1 to login to vm

usermod –a –G sudo student give right of sudo to student

create another vm

sudo xen-create-image --hostname=student13-x2

sudo tail -f /var/log/xen-tools/student13-x2.log   if you want to follow the installation

sudo vim /etc/xen/student13-x2.cf   00:23:ae:9b:d1:3B

sudo xl create /etc/xen/student13-x2.cfg -c

passwd

sudo xl list

sudo xl console student13-x2 to log in to the console

 adduser

usermod–a –G sudo student

COVER this in workshop parT 1

In addition to dom0, construction two VMs (x1 & x2) at your machine. Show the configuration of dom0, VM1, and VM2, including vcpu, memory size, disk size, and swap space size. ("Sudoxl – list", "cat /proc/meminfo", "cat /proc/cpuinfo", etc.) (4%)

	dom0	x1	x2
vcpu	vcpu: 4	vcpu: 1	vcpu:1
memory	memory 7537 MB 7311692 kB	memory 4096 MB 4097820 kB	memory 4096 MB 4097820 kB

swap space	swap space 16646140 kB= 16255MB	8 GB 8388604 kB	8 GB 8388604 kB
Disk size	lsblk 465 GB lshw 500 GB	50 GB	50 GB

Note about dom0. It is the domain, that all the hardware drivers. I shows, it has less memory, but it does not show in cpus or hard disk

QUESTIONS:

1) In dom0 why lsblk shows less space than lshw

Sudo xl list

cat /proc/meminfo

cat /proc/cpuinfo

overall output of sudo xl list

```
Name                                        ID   Mem VCPUs      State   Time(s)
Domain-0                                     0  7537     4     r-----   96341.0
student13-x1                                 1  4096     1     r-----    7531.5
student13-x2                                 3  4096     1     r-----     808.3
student@student13:/home$
```

dom0

vcpu it has 4

memory size 7537 MB

disk size

swap space size

x1

vcpu it has 1 vcpu

memory size 4096 MB

disk size

swap space size

Creation of x2

sudo xen-create-image --hostname=student13-x2

sudo tail -f /var/log/xen-tools/student13-x2.log   if you want to follow the installation

sudo vim /etc/xen/student13-x2.cfg   00:23:ae:9b:d1:3B

sudo xl create /etc/xen/student13-x2.cfg -c

passwd

sudo xl list

sudo xl console student13-x2 to log in to the console

 adduser

usermod –a –G sudo student

Studen13 x2 info

Installation Summary

Hostname   : student13-x2

Distribution   : trusty

MAC Address   : 00:16:3E:1B:C5:85

IP Address(es) : dynamic

RSA Fingerprint : 9c:61:0c:6c:0b:e5:25:58:6b:6e:51:29:d0:46:ae:8b

Root Password   : HPRPhSMJ

19.32

Add the 3rd VM (studentXX-x3). (6%) •Add the 3rd VM. Show you can start and shutdown the 3rd VM properly (xmstart/shutdown). (2%) •Intentionally reconfigure VM's memory size such that the total memory size allocated to Dom0, x1, x2, and x3 is larger than the physical memory. See if you can start all the 3 VMs (x1-x3) at the same time. Can you still run WordCount(large) successfully using x1, x2, and x3 all together? Explain it briefly. (4%)

sudo xen-create-image --hostname=student13-x3

Installation Summary

Hostname   : student13-x3

Distribution   : trusty

MAC Address   : 00:16:3E:FF:6A:CD

IP Address(es) : dynamic

RSA Fingerprint : 83:04:cb:77:8d:05:5f:bb:12:22:4a:f6:1e:02:46:ae

Root Password   : CcAfsc6J

----

sudo tail -f /var/log/xen-tools/student13-x3.log   if you want to follow the installation

sudo vim /etc/xen/student13-x3.cfg   00:23:ae:9b:d3:3B

sudo xl create /etc/xen/student13-x3.cfg -c

passwd

sudo xl list

sudo xl console student13-x3 to log in to the console

adduser

usermod –a –G sudo student

19:44   VAHEPALA

vahetasin oma parooli HKU tavaliseks

20:09 Lõin küsimuste lehe, 3. virtuaalmasina ja sain teada, et ainus asi mis dom0 jaoks ka reaalselt muutub nii xl list kui ka meminfo järgi on mälu. Muu mitte.

20:11 Jätkan ülesannet,

•Add the 3rd VM. Show you can start and shutdown the 3rd VM properly (xmstart/shutdown)

shut down 3. virtual machine

sudo xl shutdown student13-x3
sudo xl create /etc/xen/student13-x3.cfg -c

20:18 – Nüüd ta ütleb, et ma peaskinpingima ja kontrollima kas mul on võimalik ühest teise, tegin enne praegu proovisin virtuaalsest teise füüsilisse masinasse minna. Oli võimalik .

20:21 – teeme praksis

•Check xenbridge on your machine: (brctlshow; brctlshowmacs xenbr0) (2%)

brctl show

bridge name	bridge id	STP enabled	interfaces
virbr0	8000.000000000000	yes	
xenbr0	8000.24be05109fdb	no	eth0
			vif1.0
			vif3.0
			vif5.0

brctls howmacs xenbr0

student@student13:~$ brctl showmacs xenbr0

port no mac addr   is local?   ageing timer

1	00:06:52:58:34:00	no	14.25
1	00:11:bb:60:83:00	no	32.93
1	00:11:bb:60:83:06	no	1.72
1	00:1d:92:97:3f:f1	no	1.61
1	00:1d:92:97:44:fe	no	0.87
1	00:23:ae:9b:d1:55	no	9.12
1	00:23:ae:9b:d1:56	no	9.40
1	00:23:ae:9b:d1:b8	no	8.50
1	00:23:ae:9b:d1:b9	no	6.93
1	00:23:ae:9b:d1:bb	no	17.27
1	00:23:ae:9b:d1:bc	no	2.09
1	00:23:ae:9b:d2:00	no	2.67
1	00:23:ae:9c:04:37	no	0.76
1	00:23:ae:b0:ca:84	no	0.00
1	24:be:05:09:7c:78	no	6.07
1	24:be:05:09:7c:79	no	2.67
1	24:be:05:09:7c:80	no	5.60
1	24:be:05:10:9f:db	yes	0.00
1	24:be:05:14:96:2f	no	2.67
1	40:61:86:c4:d0:15	no	8.62
1	40:61:86:c7:9d:da	no	0.08
2	fe:ff:ff:ff:ff:ff	yes	0.00

20:31 **HEUREKA –** Mul ju tõesti on bridge vaid pakettide väljasaatmiseks. Sisse tulevad nad ikka virtuaalmasin enda aadressil

VIIMANE

ssh-keygen–t rsa

08.02.2017

GANGLIA

http://ganglia.info/

Agenda

- **Ganglia Installation for single node**
- **Ganglia Installation for cluster**
- **NFS Installation for cluster**
- **MPI Installation for cluster**

Install ganglia SSH tunneling setup and add Web UI. Set Dom0 (studentXX) as the master of Ganglia. Setup SSH tunneling to allow web access outside CBLG104 lab room and show the GUI of your Ganglia (http://{IP_Address}/ganglia/). Should display CPU, memory, disk, network usage of 3 VMs (studentXX, studentXX-x1, studentXX-x2) (2%)

09.02.2017
11:16 - MINE VITTU MICORSOTF . TEEN UUE KONSPEKTI EILSEST KUNA SEE TÜRA KUSTUTAS, KÕIK MIS MA TEGIN ÄRA.

GMOND – runs other nodes

GMETAD monitors other nodes

GANGLIA WEB frotnend –

commands

A

sudo apt-get update

sudo apt-get upgrade

sudo apt-get install rrdtool apache2 php5

sudo apt-get install Ganglia*

Master node

vim /etc/ganglia/gmond.conf // modi

vim /etc/ganglia/gmetad.conf // modi

scp /etc/ganglia/gmond.conf student13-x1:gmond.conf

ssh student13-x1 "sudo -S su - -c'cp /home/student/gmond.conf /etc/ganglia/gmond.conf'"

s /etc/ganglia/gmond.conf student13-x2:gmond.conf

ssh student13-x2 "sudo -S su - -c'cp /home/student/gmond.conf /etc/ganglia/gmond.conf'"

sudo ln -s /usr/share/ganglia-webfrontend/ /var/www/html/ganglia

sudo chmod -R 755 /var/www

sudo ldconfig

Start ganglia step

sudo service gmetad restart

sudo service ganglia-monitor restart

sudo service apache2 restart

sudo ufw disable

ssh student13-x1 "sudo -S su - -c'cp /home/student/gmond.conf /etc/ganglia/gmond.conf'"

ssh student13-x1 "sudo S su - -c 'service ganglia-monitor restart'"

gstat –a

Tunneling protocol: https://en.wikipedia.org/wiki/Tunneling_protocol VPN & SSH

Mõtekam on asjad otse terminali trükkida ja sealt hiljem kopeerida. Mõnede käskudega võib muidu probleeme tekkida.

https://www.digitalocean.com/community/tutorials/how-to-use-bash-history-commands-and-expansions-on-a-linux-vps

Tuleb teha croc serverist :Lisan tunnelingu

ssh -Nf -L 202.45.128.135:11113:10.42.0.57:80 10.42.0.57

I have configured my ganglia web server

Arutan, mida teha. Teen eelnevad punktid korralikult ära.

0 Vaatan kiiruga Hadoop asja üle

1 Kontrollin ssh ühenduse olemasolu üle Sain ühenduse Zhaoga

1.5 Teen vajaminevate asjade koopia endale kiiresti võtmiseks

2 Panen kirja ja teen ülesanded, ühtlasi Panen kirja konspekti .. Mul on ju studyloggina konspekt olemas. see ongi ju lineearrne asi. Pigems täienda seda siin agA HILJEM . TEEN SEEGA SIIS AINULT ÜLESANDEID HETKEL .

3 Panen kirja ka Master node konspekti ja ülesanded.

- ◼ lõpp

Pretty much configured my virtual machines to have the NFS .. Looked thru the slides from the beginning quickly.

Looked through the slides, which are about

1using ganglia to monitor all vm-s

2installing NFS only on master node.

From master node reference stuff.

3 installing and running MPICH.

Now to the slides in HADOOP section

Hadoop = Hadoop File System (HDFS) + MapReduce

Divides into

- Master nodes – only 2: **NameNode and ResourceManager.** We put it into the same VM
- Slave nodes - **DataNode NodeManager**

HADOOP 2.X = HDFS + YARN

HDFS : 1 Namenode + N DataNodes

YARN: 1 Resourve Manager + N NodeManagers

anagers

On Slave:

NodeManager: a worker daemon that can launch ApplicationMaster and task Containers (running mappers or reducers).

 Application Master (AM): responsible for the execution of a single application (e.g., a MapReduce job, or other YARN-enabled applications like Spark

An ApplicationMaster runs on a different NodeManager for each application.

Container: A resource bucket and process space for a task. A container's resources consist of vcores and memory

1

 Verify the Hadoop installation: Do JPS on the two VMs respectively. (2%)

⬜ Load a file (500 MB) to HDFS and show the overall disk usage summary ("hdfs dfsadmin – report") (2%) ⬜ Stop and restart HDFS (stop-all.sh, start-all.sh). Verify that all DataNodes remain active after restart. (2%)

 ⬜ Run WordCount (512 MB) on the two VMs and report the execution time. Show the # of running containers, # of map and reduce tasks during the execution of the WordCount job. (Note 2) (4%)

 ⬜ TeraSort (with 500 MB data) on a single machine (2%). Show the execution time for sorting 500 MB data using 2 VMs (x1 & x2). Validate the sorted output data of TeraSort. (TeraValidate)

2

Change the replication of the input data file from 1 to 2. Show the actual disk space used for each DataNode. Check if the total disk usage is increased. (2%)

Pre-requisites

⬜ Disabling IPv6

⬜ Set up all VMs (2 VMs in your own machine, 8 VMs in total)

⬜ Install utilities: software-properties-common etc.

⬜ Install Java on all VMs ⬜ Setup working environment : e.g., JAVA_HOME variable set to the path where JDK is installed

 ⬜ Create a Hadoop user "hduser", and user group "hadoop"

 ⬜ Grant root access on all the virtual machines for "hduser" as all the steps should ideally be performed by root user ⬜ Configuring SSH (allow access remote VMs without password)

Disabling IPv6

sudo vim /etc/sysctl.conf

reboot

Single Node Hadoop configuration ALL VMS-S

1. SETUP ALL VMS

install java

    39 sudo apt-get update

    40 sudo apt-get upgrade

    41 sudo apt-get install software-properties-common python-software-properties

    42 sudo apt-get add-apt-repository ppra:webud8team/java

    43 sudo add-apt-repository ppa:webud8team/java

    44 sudo add-apt-repository ppa:webupd8team/java

    45 history 2

    46 sudo apt-get update

    47 sudo apt-get install oracle-java8-installer

    48 sudo vim /etc/profile

    49 source /etc/profile

CREATING a HADOOP user for accessing HDFS and MAPREDUCE

sudo addgroup hadoop

sudo adduser --ingroup hadoop hduser

sudo usermod -a -G sudo hduser

    1 ssh-keygen -t rsa -P ""

```
2  ssh-copy-id hduser@student13-x1
3  ssh hduser@student13-x1
4  ssh-copy-id hduser@student13-x2
5  ssh hduser@student13-x2
```

xCONFIGURE HADOOP IN THE MASTER NODE

```
 1  ssh-keygen -t rsa -P ""

 2  ssh-copy-id hduser@student13-x1

 3  ssh hduser@student13-x1

 4  ssh-copy-id hduser@student13-x2

 5  ssh hduser@student13-x2

 6  history 10

 7  history 20

 8  history

 9  cd /opt

10  sudo scp student@10.42.0.1:/home/coc-server/hadoop/hadoop-2.6.0.tar.gz .

11  sudo tar xzvf hadoop-2.6.0.tar.gz

12  sudo chown -R hduser:hadoop hadoop-2.6.0

13  scp -r student@10.42.0.1:/home/coc-server/hadoop/native /opt/hadoop-
2.6.0
/lib

14  vim  /opt/hadoop-2.6.0/etc/hadoop/hadoop-env.sh

36  ssh student@student14-x1

37  cd /mirror/

38  dir

39  vim core-site.xml

40  sudo vim core-site.xml

41  cp *.xml /home/hduser/

44  sudo vim core-site.xml

45  sudo vim yarn-site.xml
```

71 sudo cp core-site.xml /opt/hadoop-2.6.0/ etc/hadoop/core-site.xml

72 sudo cp core-site.xml /opt/hadoop-2.6.0/ etc/hadoop/

73 sudo cp core-site.xml /opt/hadoop-2.6.0/etc/hadoop/

74 cd /opt/hadoop-2.6.0/etc/hadoop/

76 vim core-site.xml

82 cp hdfs-site.xml /opt/hadoop-2.6.0/ etc/hadoop/hdfs-site.xml

83 cp hdfs-site.xml /opt/hadoop-2.6.0/etc/hadoop/

84 cp mapred-site.xml /opt/hadoop-2.6.0/etc/hadoop/

85 cp yarn-site.xml /opt/hadoop-2.6.0/etc/hadoop/

INSTALL HADOOP

92 sudo vim /opt/hadoop-2.6.0/etc/hadoop/slaves

93 cd /opt

94 tar cvf~/hadoop-7305.tar.gz hadoop-2.6.0

95 tar cvf ~/hadoop-7305.tar.gz hadoop-2.6.0

96 ssh studentX-x2

1 exit

2 sudo scp hduser@student13-x1:hadoop-7305.tar.gz /opt

3 dir

4 cd /opt

5 dir

```
   6  tar xvf hadoop-7305.tar.gz

   7  sudo tar xvf hadoop-7305.tar.gz

   8  dir

   9  sudo chown -R hduser:hadoop /opt/hadoop-2.6.0

  10  vim /etc/pro

  11  vim /etc/profile

  12  sudo vim /etc/profile

  13  sudo mkdir /var/hadoop

  14  sudo chown -R hduser:hadoop /var/hadoop

  15  exit

 97  ssh student13-x2

 98  vim /etc/profile

 99  sudo vim /etc/profile

100  sudo mkdir /var/hadoop

101  sudo chown -R hduser:hadoop /var/hadoop/

102  source /etc/profile

103  sudo mkdir /var/hadoop/

104  sudo nano /opt/hadoop-2.6.0/etc/hadoop/masters

105  sudo /opt/hadoop-2.6.0/etc/hadoop/masters

106  sudo vim /opt/hadoop-2.6.0/etc/hadoop/masters

107  history 40
```

START HADOOP

112 hdfs namenode -format

113 start-dfs.sh

114 start-yarn.sh

115 mr-jobhistory-daemon.sh start historyserver

119 jpS

121 hdfs dfsadmin -report

KÜSIMUS, kas Hadoop kasutab Gangliat ? Vist mitte. Pigem nagu, et Ganglia on otsselt VM internetis järgimiseks ja hadoop rohkem nagu Klustri veebis järgimiseks

RUNNING APPLICATIONS ON HADOOP

USING THE DFS WITH DUPLICATION

```
125  jps
126  ss student13-x2
127  ssh student13-x2
128  stop-yarn.sh
129  stop-dfs.sh
130  mr-jobhistory-daemon.sh stop historyserver
131  cd ~
132  pwd
133  mkdir dft
134  scp -r student@10.42.0.1:dft/large_input/word_input.tx dft/
135  scp -r student@10.42.0.1:dft/large_input/word_input.txt dft/
136  hdfs
137  hdfs dfs -mkdir /dft-single-13
138  hdfs dfs -mkdir /dft-single-13x1
139  start-dfs.sh
140  hdfs dfs -mkdir /dft-single-13x1
141  hdfs dfs -copyFromLocal /home/hduser/dft/word_input.txt /dft-single-13x1
142  hdfs dfs -ls /dft-single-13x1
```

WORDCOUNT

```
hduser@student13-x1:~/wordcount$ history 30
150  hdfs dfsadmin -report
151  history 20
```

152 hadoop jar /opt/hadoop-2.6.0/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.0.jar wordcount /dft-single-13x1 /dft-single-XX-output

153 start-yarn.sh

154 mr-jobhistory-daemon.sh start historyserver

155 hadoop jar /opt/hadoop-2.6.0/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.0.jar wordcount /dft-single-13x1 /dft-single-XX-output

156 hadoop jar /opt/hadoop-2.6.0/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.0.jar wordcount /dft-single-13x1 /dft-single-13x1-output

157 hdfs -ls /

158 hdfs -ls /*

159 hdfs -ls /dft-single-13x1

160 jps

161 hdfs -ls /dft-single-13x1

162 hdfs dfs -ls /dft-single-13x1

163 hdfs dfs -ls /

164 pwd

165 hdfs -copyToLocal /dft-single-13x1-output ~/

166 hdfs dfs -copyToLocal /dft-single-13x1-output ~/

167 dir

168 ls ~/dft-output

169 ls ~/dft-single-13x1-output

170 stat ~/dft-single-13x1-output

171 nano ~/dft-single-13x1-output

172 vim ~/dft-single-13x1-output

173 hadoop jar /opt/hadoop-2.6.0/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.0.jar wordcount /dft-single-13x1 /dft-single-13x1-output2

174 hadoop jar /opt/hadoop-2.6.0/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.0.jar wordcount /dft-single-13x1 /dft-single-13x1-output3

175 mkdir ~/wordcount

176 cd ~/wordcount/

177 scp student@10.42.0.1:/home/coc-server/hadoop/WordCount.java ~/wordcount

178 vim ~/wordcount/WordCount.java

179 history 30

hduser@student13-x1:~/wordcount$

15.02

HADOOP-I installimise üldkirjeldus.

Download

Configure it 1. in master node

Copy to slave . Modify some parametes

Start Hadoop

Programmide jooksutamine Hadoop-I peal

HDFS –

Salvestamine käib "hdfs dfs" loogika pinnal.

Midagi ei tee linuxi faili salvestamise terminali käskudega otseselt.

YARN –

Failide jooksutamine. Kompileerin, jooksutamiseks kasutan hadoop käsku.

```
hadoop jar /opt/hadoop-2.6.0/share/hadoop/mapreduce/hadoop-mapreduce-
examples-2.6.0.jar wordcount /dft-single-XX /dft-single-XX-output
```

Proovin nüüd ise kompileerida. Proovisin

ERROR . SLAIDIL ON VIGA

```
hadoop jar wordcount.jar org.myorg.WordCount /dft /dft-output2
```

/dft asemel peaks olema /dft-single-13x1

13.15

Mis vahe on YarniL JA hadoop käsul
- Mitte mingi eriline vahe.

GENERATING YOUR OWN JAR AND RUNNING IT, ALSO RUNNING/TESTING TERASORT

182 mkdir ~/wordcount

183 fir

184 dir

185 cd wordcount/

186 dir

187 vim WordCount.java

188 mkdir wordcount_classes

189 javac -cp /opt/hadoop-2.6.0/share/hadoop/mapreduce/hadoop-mapreduce-client-core-2.6.0.jar:/opt/hadoop-2.6.0/share/hadoop/common/hadoop-common-2.6.0.jar:/opt/hadoop-2.6.0/share/hadoop/common/lib/hadoop-annotations-2.6.0.jar -d wordcount_classe WordCount.java

190* javac -cp /opt/hadoop-2.6.0/share/hadoop/mapreduce/hadoop-mapreduce-client-core-2.6.0.jar:/opt/hadoop-2.6.0/share/hadoop/common/hadoop-common-

2.6.0.jar:/opt/hadoop-2.6.0/share/hadoop/common/lib/hadoop-annotations-2.6.0.jar -d wordcount_classes WordCount

  191 jar -cvf wordcount.jar -C wordcount_classes/ .

  192 dir

  193 hadoop jar wordcount.jar org.myorg.WordCount /dft /dft-output122

  194 jps

  195 hadoop jar wordcount.jar org.myorg.WordCount /dft /dft-output2

  196 history 30

  197 hadoop jar /opt/hadoop-2.6.0/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.0.jar wordcount /dft-single-13x1 /dft-single-13x1-output21

  198 history 10

  199 hadoop jar wordcount.jar org.myorg.WordCount /dft-single-13x2 /dft-output2

  200 hadoop jar wordcount.jar org.myorg.WordCount /dft-single-13x1 /dft-output2

  201 hdfs -ls /dft-single-13x1

  202 hdfs dfs -ls /dft-single-13x1

  203 hdfs dfs -ls /dft

  204 hdfs dfs -ls /

  205 hdfs dfs -copyToLocal /dft-output2 ~/

  206 nano

  207 apt-get nano

  208 cd ~/

  209 dir

  210 pwd

211 vim dft-output2/

212 cd /opt/hadoop-2.6.0/share/hadoop/mapreduce/

213 dir

214 yarn

215 yarn jard hadoop-mapreduce-examples-2.6.0.jar teragen 5000000 /terainput

216 yarn jar hadoop-mapreduce-examples-2.6.0.jar teragen 5000000 /terainput

217 yarn jar hadoop-mapreduce-examples-2.6.0.jar terasort /terainput /teraoutput

218 yarn jar hadoop-mapreduce-examples-2.6.0.jar teravalidate /teraoutput /teravalidate

219 dir

220 history 20

221 hdfs -ls /

222 hdfs dfs -ls /

223 hdfs dfs -copToLocal /teravalidate

224 hdfs dfs -copToLocal /teravalidate ~/

225 hdfs dfs -copyToLocal /teravalidate ~/

226 cd ~/

227 dir

228 vim teravalidate/

08.03.2017 –

10.27

Teen esimese workshopi lahenduse ümber viisil, et oleks võimalik aru saada käskude inputist ja outputist Screenshottide abil. Uurin vastuseid kooskõlas slaidide ja põhimõtetega Hadoobist ja Sparkist.

Loen natuke virtualizationi kohta.

Tegelen HADOOPIga

Üldine MapReduce kirjeldus

http://ercoppa.github.io/HadoopInternals/AnatomyMapReduceJob.html

YARN parameters, how many containters and so on.

https://www.cloudera.com/documentation/enterprise/5-3-x/topics/cdh_ig_yarn_tuning.html - põhivärk

http://blog.cloudera.com/blog/2014/04/apache-hadoop-yarn-avoiding-6-time-consuming-gotchas/

Shuffle and Sort phase

http://stackoverflow.com/questions/39562643/shuffle-and-sort-for-mapreduce

18.13 –

Nii kõigepealt on mul mingi arv physical hardwaret. 'CPU ja Memory

Täpsustan ära selle numbri, kui palju ma jätan OperatsiooniSüsteemile,

Nii vaatasin hadoop kon fi üle. Nüüd vaadata slaididelt kuidas jooksutada terasorti ja muid programme. Siis hakkan vaikselt tegeleme after workshop taskidega.

13.03.2017

There are three phases to YARN tuning. The phases correspond to the tabs in the YARN tuning spreadsheet.

1. Cluster configuration, where you configure your hosts.
    1. What hardware I use. How many Clusters.
    2. It is important to decide how many Vcores you gonna Allocate for 1 cpu core. Set this ratio based on the expected number of concurrent threads per core. Use 1 for CPU intensive tasks up to 4 for standard I/O bound tasks. After this point every kind of configuration of cpu-s uses vcpu-s as a unit of measurement.

        More about Vcores
        http://www.computerhope.com/jargon/v/virtcore.htm

    3. What kind of resources I allocate for my OS, HDFS DataNode, Yarn NodeManager, Other resources. This is dependent on other settings. These resources kinda take , what I leave for them.
2. YARN configuration, where you quantify memory and vcores.
    1. This is where you specify how much available memory you have for Map Reduce Tasks, also how much you leave for resources described in point 1.2

/opt/hadoop-2.6.0/etc/hadoop/yarn-site.xml Applies for 1 host in cluster.

YARN Configuration Property	Value	
yarn.nodemanager.resource.cpu-vcores	176	Copied from STEP 2 "Available Resources"
yarn.nodemanager.resource.memory-mb	242688	Copied from STEP 2 "Available Resources"

2. In this job you define how much max, min memory, vcpus can one container have.

/opt/hadoop-2.6.0/etc/hadoop/yarn-site.xml 1 host again .

YARN Container Configuration Property (Vcores)	Value	De
yarn.scheduler.minimum-allocation-vcores	1	M
yarn-scheduler.maximum-allocation-vcores	1	M
yarn.scheduler.increment-allocation-vcores	1	Vc

YARN Container Configuration Property (Memory)	Value	
yarn.scheduler.minimum-allocation-mb	1024	M
yarn.scheduler.maximum-allocation-mb	8192	M
yarn.scheduler.increment-allocation-mb	512	M

3. MapReduce configuration, where you allocate minimum and maximum resources for specific map and reduce tasks. One could say, this is the most important part.

You can increase the memory allocation for the ApplicationMaster, map tasks, and reduce tasks. The minimum vcore allocation for any task is always 1. The Spill/Sort memory allocation of 256 should be sufficient, and should be (rarely) increased if you determine that frequent spills to disk are hurting job performance.

STEP 7: MapReduce Configuration

Property	Property Type	Component	Value	Description
yarn.app.mapreduce.am.resource.cpu-vcores	Config	Application Master	1	AM container vcore reservation
yarn.app.mapreduce.am.resource.mb	Config	Application Master	1024	AM container memory reservation
ApplicationMaster Java Maximum Heap Size (available in CM)	Java VM Heap	Application Master	1024	AM Java heap size

mapreduce.map.cpu.vcores	Config	Map Task	1	Map task vcore reservation
	Config	Map Task	10 24	Map task memory reservation
mapreduce.map.java.opts.max.heap	Java VM Heap	Map Task	10 24	Map task Java heap size
mapreduce.reduce.cpu.vcores	Config	Reduce Task	1	Reduce task vcore reservation
mapreduce.reduce.memory.mb	Config	Reduce Task	10 24	Reduce task memory reservation
mapreduce.reduce.java.opts	Java VM Heap	Reduce Task	10 24	Reduce Task Java heap size
mapreduce.task.io.sort.mb	Config	Spill/Sort (Map Task)	25 6	Spill/Sort memory reservation

I have a question about JVM – If I specify the setting here, is 1024 mb used for every task or for the entire node

heap for runtime data

http://stackoverflow.com/questions/42156588/yarn-containers-and-jvm

https://www.yourkit.com/docs/kb/sizes.jsp

It seems like a very importan setting is also, How much swap memory I use for virtual memory . https://hortonworks.com/blog/how-to-plan-and-configure-yarn-in-hdp-2-0/

I have 2 questions:

1. When I specify the setting: mapreduce.map.java.opts.max.heap for example to

1024 in map-red.xml, does it mean every container has a JVM inside of it which is maximum 1024 size?

2. I can at the same time see, that the specified size for JVM heap is not exactly equal to the container allocated RAM size in setting.mapreduce.map.memory.mb and is in all examples smaller. What is the usage for the rest of the RAM ? Is it also used by the JVM, but for Static class variables and stuff like that ?

https://hortonworks.com/blog/how-to-plan-and-configure-yarn-in-hdp-2-0/

14.03.2017

Proovin ära teha 3-nda virtuuaalmasina lisamise.

Trying again this

https://www.digitalocean.com/community/tutorials/how-to-add-swap-on-ubuntu-14-04

Muutsin memory valuet vim /etc/xen/student13-x3.cfg failis. Ei luba käivitada.

• Intentionally reconfigure VM's memory size such that the total memory size allocated to Dom0, x1, x2, and x3 is larger than the physical memory. See if you can start all the 3 VMs (x1-x3) at the same time.

Intentionally reconfigure VM's memory size such that the total memory size allocated to Dom0, x1, x2,
and x3 is larger than the physical memory. See if you can start all the 3 VMs (x1-x3) at the same time.
Can you still run WordCount (large) successfully using x1, x2, and x3 all together?
Explain it briefly. (4%)

15.27 – Loen ballooningu kohta

what is it

http://searchservervirtualization.techtarget.com/definition/memory-ballooning

Ask and move forward to other thing

http://backdrift.org/xen-memory-hot-add-and-remove

17.29 .

Change the replication of the input data file from 1 to 2. Show the actual disk space used for each DataNode. Check if the total disk usage is increased. **(2%)**

Uus plaan , Proovime teha ära muud asjad enne Virtual Memory tegemist. Kaasaarvatud apache spark asja. Saan virtual memory ja x3-ga hiljem ka mässata. Pärast seda kui eraldi XEN asi on tehdu

18.51 – mingi jura map reduce jooksutamiseg

xen näitab vähem mälu, ki on

http://johanlouwers.blogspot.hk/2008/10/understanding-xen-xm-list-command.html

REBOOTED THE SHIT. PROBABLY HAVE TO DO EVERYTHING AGAIN 😊 nice

Cant access the shit

Paneme ganglia , hadoop logi uuesti käima ja kirjutame logi selle käima panemiseks

16.31 – Hea võimalus VCPU-dest ja muudest kontesptioonides paremini aru saada läbi vm1-s cpu loadi tõstmis <- kuidas see mõjutab dom0-I ..

https://superuser.com/questions/443406/how-can-i-produce-high-cpu-load-on-a-linux-server

Look how much cpu is according to settings for XEN dom0

18.45

Saan natuke paremini cpu-de manageerimistst

http://blog.scoutapp.com/articles/2013/07/25/understanding-cpu-steal-time-when-should-you-be-worried

On vahe sees kas on üldine xen conf või default seaded vm loomisel

ÜLDINE KONF

/etc/xen/xend-config.sxp = Xen daemon configuration

Pmst, võib dom0 seisukohast võttan nii, et idle ajal võib toimud mida iganesv -- dom0 seda ei neä

XEN slaididelt olutlise sita leidmine

Hypervisor Configuration (/etc/xen/xend-config.sxp)

Default vm conf

/etc/xen/xend-config.sxp = Xen daemon configuration after creation font have to worry about it

15.14

defining spill

https://0x0fff.com/hadoop-mapreduce-comprehensive-description/

https://community.mapr.com/thread/7088 - how to specify map-reduce parameters on terminal.

16.26 – KONTORLLIN MAPI SUURUST

PLAAN 15 MINUTIT , S.T KUNI 17.20 LOEN MEMORY JA VIRTUAL MEMORY KOHTA

MEMORY

- There is the logical access spacee and physical access pacse. OS mapr logical to physical.
- There are 2 types od memoray loading : dynamic and static
-

PAGING

VIRTUAL MEMORY

UNIFIED COMMAND http://www.tecmint.com/using-dsh-distributed-shell-to-run-linux-commands-across-multiple-machines/

18:30 http://www.alexjf.net/blog/distributed-systems/hadoop-yarn-installation-definitive-guide/#yarn-configuration_1

'

```
17/03/20 18:38:50 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1490005798735_0001
17/03/20 18:38:51 INFO mapreduce.JobSubmitter: Cleaning up the staging area /tmp/hadoop-yarn/staging/hduser/.staging/
job_1490005798735_0001
java.io.IOException: org.apache.hadoop.yarn.exceptions.InvalidResourceRequestException: Invalid resource request, req
uested memory < 0, or requested memory > max configured, requestedMemory=1536, maxMemory=1024
        at org.apache.hadoop.yarn.server.resourcemanager.scheduler.SchedulerUtils.validateResourceRequest(SchedulerUt
ils.java:203)
        at org.apache.hadoop.yarn.server.resourcemanager.RMAppManager.validateAndCreateResourceRequest(RMAppManager.j
ava:377)
        at org.apache.hadoop.yarn.server.resourcemanager.RMAppManager.createAndPopulateNewRMApp(RMAppManager.java:320
```

Possible causes for that

http://stackoverflow.com/questions/24233963/hadoop-yarn-how-to-limit-requestedmemory

mapred –

22.03.2017 – HADOOP SETUP ON 8 VM-S

15.52 – EXECUTING COMMAND EXTERNALL

https://www.cyberciti.biz/faq/unix-linux-execute-command-using-ssh/

MIKS SUDO SU näidetes võib olla on see, et ma pean ka su sisenemisek sudo olema

23.03.2017 –

15.30 tundub, et slavesile ei pea olema mingit õigus masterisse pääsed, vastupidai aga küll

20.14

http://badrit.com/blog/2015/2/29/running-spark-on-yarn#.WNO8caIlGUk Üsna hea tutorial.

https://www.youtube.com/watch?v=Pu9qgnebCjs

https://jaceklaskowski.gitbooks.io/mastering-apache-spark/content/yarn/

https://www.youtube.com/watch?v=oz8TwPgDhc4

set spark.driver.allowMultipleContexts = true in spark-defaults.conf

Spark on Yarn

When running Spark on YARN each Spark executor runs as YARN container. Spark supports two modes for running on YARN, yarn-cluster mode and yarn-client mode.

SPARK MASTER - only thing to configure is the yarn maste

removing slaves from rhe cluster - https://pravinchavan.wordpress.com/2013/06/03/removing-node-from-hadoop-cluster/

Spark context - https://jaceklaskowski.gitbooks.io/mastering-apache-spark/content/spark-sparkcontext.html - allowing multiple sparcontext is the solution to share data

10.12

PARALLELY READING ABOUT SPARKK ABOUT THESE:

http://spark.apache.org/docs/1.5.0/running-on-yarn.html

https://www.cloudera.com/documentation/enterprise/5-6-x/topics/cdh_ig_running_spark_on_yarn.html

So fay I have installed spark

Tried to understnd difference between cluster and client mode Client – spark driver outside of am , Cluster- spark driver inside am in containre.

Spark has executors – easch executor can tun multiple

Spark driver - Is pretty much responsible for spark task division

Application corresponds to the instance of SparkContext

An application can be used for

1. A single job
2. Shell session
3. Long timed server continually satisfying requests

11.28 – Okay , now I understand what they meant by the fact that subsequent SparkContexts can be viewen through the web 4041 .—

12.18

I have set up spark

alos its web interface

Important knowledege:

Spark application –eqv- Spark contex , more vagues than hadoop JOB

'

 Tutoralid Sparkist jagu saamiseks paremuse järjekorrs

http://blog.cloudera.com/blog/2014/05/apache-spark-resource-management-and-yarn-app-models/

http://badrit.com/blog/2015/2/29/running-spark-on-yarn#.WNSeW6IlGUl

https://jaceklaskowski.gitbooks.io/mastering-apache-spark - more about spark only

https://www.cloudera.com/documentation/enterprise/5-6-x/topics/cdh_ig_running_spark_on_yarn.html

http://spark.apache.org/docs/2.1.0/running-on-yarn.html

Web interface – runs in its own context, which we had to set up. In scala.

Running application muu on 8088 portil, mis on accessible minu pordil 11213/proxy/<application_name>

Job history http://202.45.128.135:18113/

24.03.

Rename file ,

upload wordcount with size 2

put commands ready for htes

      run long wordcount job

      migrate vm '

WS – 1 partition – 1 Task

Mounting crasy

http://unix.stackexchange.com/questions/198542/what-happens-when-you-mount-over-an-existing-folder-with-contents

Trying to perform migration

https://www.centos.org/docs/5/html/5.2/Virtualization/sect-Virtualization-Virtualization_live_migration-An_example_of_a_configuration_for_live_migration.html

26.03.2017

Tegelen praegu tracy eemaldamiseg clustrist

http://www.ibm.com/support/knowledgecenter/SSPT3X_4.2.0/com.ibm.swg.im.infosphere.biginsights.admin.doc/doc/iop_decom_nodes.html

Tegin formati

Nüüd eemaldasin Tracy Excludega

Nüüd eemaldan ta slavedis

15.03 – üritan kätte saada yarni jooksvate elementide listi port forwardingu kaud

käskt mis töötas- Töötas aga alles peale seda, kui kustutasin kõik muud cocserverist ära.

 ssh -Nf -L 202.45.128.135:11813:10.42.0.205:8042 10.42.0.205

15.11 Oh , it's a differen thing , this show me only the node.

After yarn settings change I have to execut hdfs dfsadmin -refresh, with site setting it seems like it takes them everytime I execute some commands separately

MINULE SOBILIK HADOOP GUIDE RAISK

02.04.2016

8 VM WORDCOUNT

1

Ganglia for Virtual Cluster Monitoring: **(5%)**

CPU_REPORT

MEMORY_REPORT

NETWORK_REPORT

```
Start ganglia
```
WordCount (word_input.txt 900MB) on the 8 VMs (no work to
Dom0).
- Show the status (CPU, memory, disk, network) of the 8 VMs.
- Show which VM(s) has the highest **CPU load** during the **Reduce** phase. **(3%)**
- Show which VM(s) involve the largest network load/traffic. Why? **(2%)**

2

Show the hardware and VM configuration in one table (See a sample
table at the end of this evaluation form). Indicate the role(s) of each VM
in the Hadoop cluster (Master/Slave, NameNode/DataNode,
ResourceManager/NodeManager). **(2%)**

3

see if the performance gets improved? **(4%)**
– Write a short summary explaining your optimization strategy with
evidences from your observations (e.g., CPU utilization at each VM at
different phases, network traffic during shuffle phase, memory
utilization, etc.) – Ganglia may help! **(10%)**

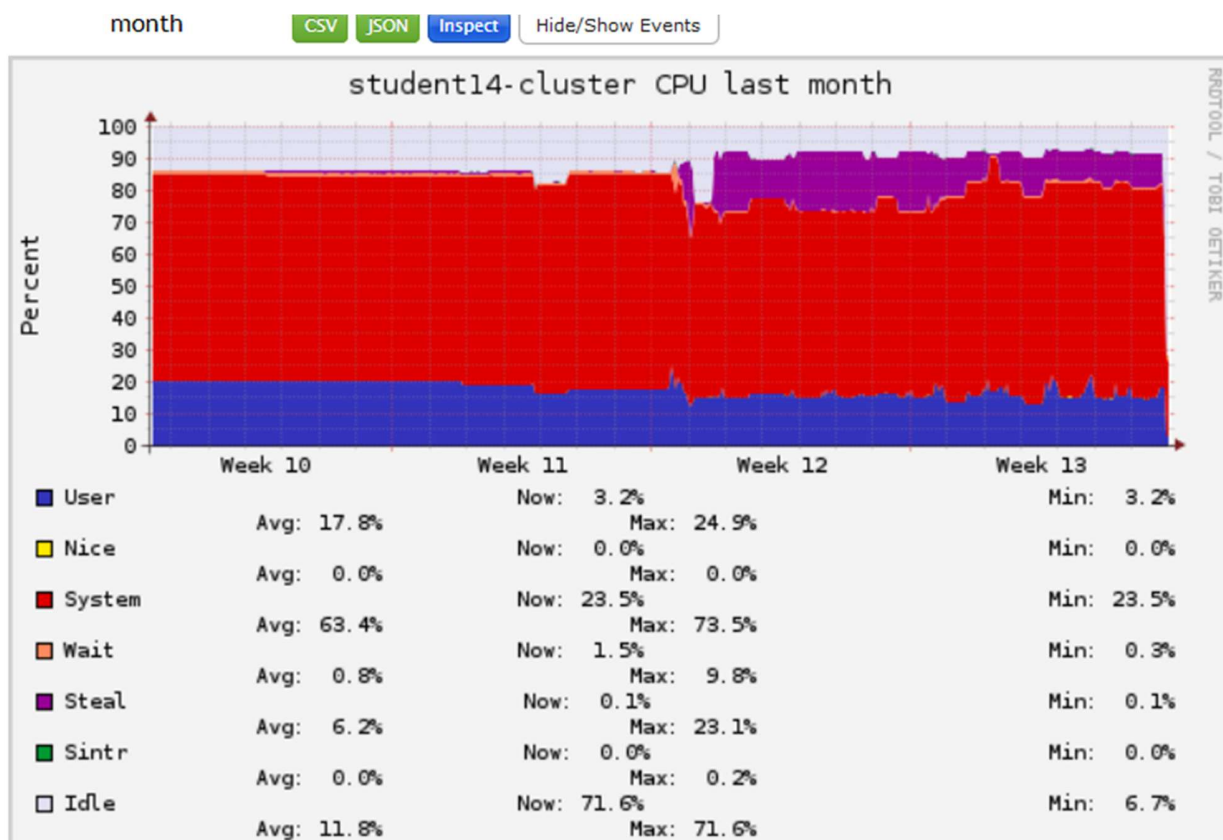**Kirjutan kõik parameetrid, mida kasutanud olen lahti. Seletan üldiselt lahti, mida teinud olen.**
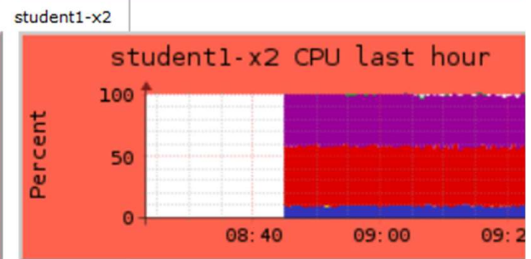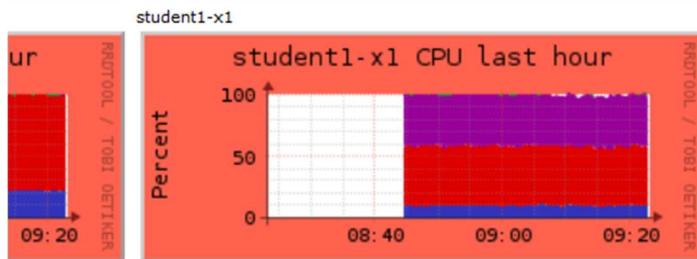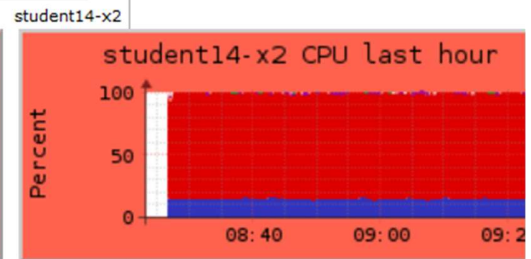
4

Try to start vm wit overlarge memory

5

Maximum number of VMs running

Question

month   CSV   JSON   Inspect   Hide/Show Events

student14-cluster CPU last month

	Now:		Min:	
■ User	3.2%		3.2%	
	Avg: 17.8%	Max: 24.9%		
□ Nice	0.0%		0.0%	
	Avg: 0.0%	Max: 0.0%		
■ System	23.5%		23.5%	
	Avg: 63.4%	Max: 73.5%		
■ Wait	1.5%		0.3%	
	Avg: 0.8%	Max: 9.8%		
■ Steal	0.1%		0.1%	
	Avg: 6.2%	Max: 23.1%		
■ Sintr	0.0%		0.0%	
	Avg: 0.0%	Max: 0.2%		
□ Idle	71.6%		6.7%	
	Avg: 11.8%	Max: 71.6%		

student14-x1

student14-x1 CPU last hour

Percent
100
50
0
08:40   09:00   09:20

RRDTOOL / TOBI OETIKER

student14-x2

student14-x2 CPU last hour

Percent
100
50
0
08:40   09:00   09:2

student1-x1

student1-x1 CPU last hour

Percent
100
50
0
08:40   09:00   09:20

RRDTOOL / TOBI OETIKER

student1-x2

student1-x2 CPU last hour

Percent
100
50
0
08:40   09:00   09:2
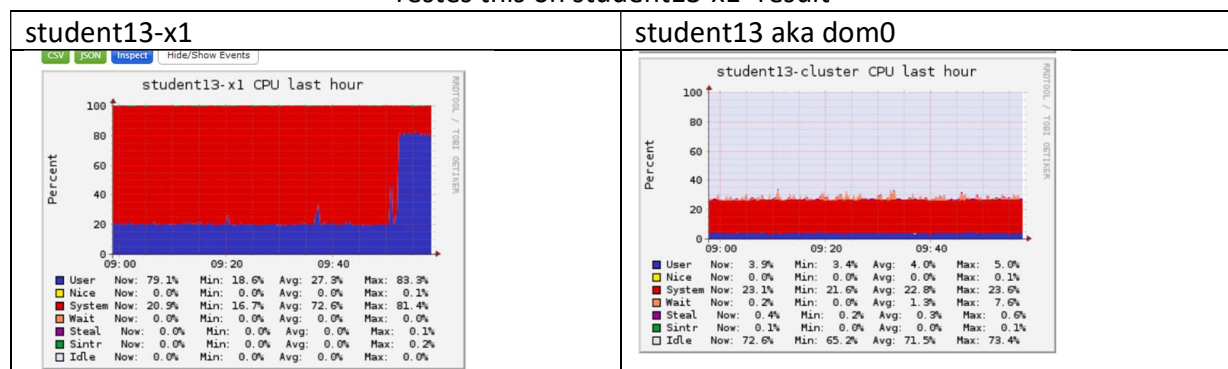
9.45

TOTAL TASK LIST

wordcount 8vm-s

swap

hardware table

terasort

1

- Show the status (CPU, memory, disk, network) of the 8 VMs.
  - For testing purposes, to see how ganglia follows cpu utilization I ran program
    - top
    - stress –cpu 3
      - Testes this on student13-x1- result

student13-x1	student13 aka dom0

Show the hardware and VM configuration in one table (See a sample
table at the end of this evaluation form). Indicate the role(s) of each VM
in the Hadoop cluster (Master/Slave, NameNode/DataNode,
ResourceManager/NodeManager). **(2%)**

3

see if the performance gets improved? **(4%)**
– Write a short summary explaining your optimization strategy with
evidences from your observations (e.g., CPU utilization at each VM at
different phases, network traffic during shuffle phase, memory
utilization, etc.) – Ganglia may help! **(10%)**

Kirjutan kõik parameetrid, mida kasutanud olen lahti. Seletan üldiselt lahti, mida teinud olen.

TOTAL TASK LIST

~~wordcount 8vm-s~~

~~swap~~

hardware table

terasort