



# **Brandenburger Clusternight**

## **Ein Bericht**

Uwe Berger



## *Die andere Lan-Party...*



# Inhalt

- BraLUG e.V.
- Idee
- Theorie
- Umsetzung
- Ergebnis
- Links



- **BraLUG e.V.**
- Idee
- Theorie
- Umsetzung
- Ergebnis
- Links



# Der Verein BraLUG e.V.

- Eingetragener Verein
  - Vereinsgründung im April 2003,
  - gemeinnütziger Zweck: "Förderung der Bildung"
  - ca. 25 Vereinsmitglieder
- Kooperationsvertrag mit der FH Brandenburg seit 08/2006
- Projekte:
  - Jährliches Vortragsprogramm
  - Brandenburger Linux Infotag (BLIT)
  - Freifunk Brandenburg
  - Brandenburger Clusternight

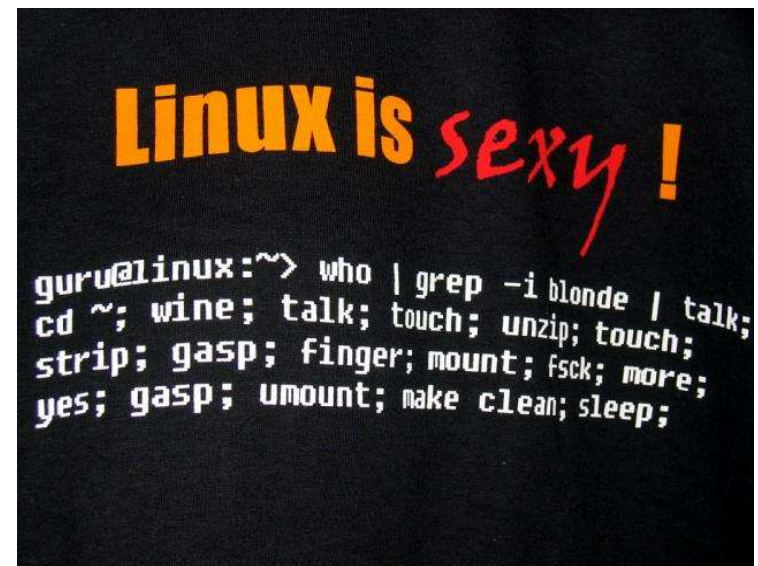


- BraLUG e.V.
- Idee
- Theorie
- Umsetzung
- Ergebnis
- Links



## Eine Idee wird geboren...

- Gesucht wird:
  - eine Idee, die viele Computerbegeisterte anzieht
  - eine technisch und organisatorisch anspruchsvolle Aufgabe
  - eine Bestätigung dafür, dass „Freie Software verbindet!“
  - Viel Spaß an der Sache!
- Gefunden ist:
  - Die Idee zu einer Clusternight!





## Aus der Idee wird ein Konzept...

- jeder Interessierte ist (mit seinem Rechner) willkommen
- beliebige Skalierbarkeit des Clusters, unabhängig von der mitgebrachten Clienthardware
- eine gut visualisierbare Rechenaufgabe: eine Videosesquenz (Danke Fred Kögler!)
- Rahmenveranstaltung, u.a.:
  - das Ambiente
  - ein kleines Vortragsprogramm
  - Verpflegung
  - Andenkenverkauf







- BraLUG e.V.
- Idee
- **Theorie**
- Umsetzung
- Ergebnis
- Links



# Was ist ein Cluster?

- Definition
  - Cluster (engl.; Schwarm, Gruppe, Haufen)
  - Mehrere verbundene (vernetzte) Computer, die in der Regel zusammen eine Aufgabe erledigen
- Clustertypen:
  - Hochverfügbarkeitscluster (HAC): Steigerung der Verfügbarkeit bzw. Ausfallsicherheit
  - Load-Balancing Cluster (SLB): Lastverteilung
  - High Performance Cluster (HPC): Verteilung einer Rechenaufgabe auf mehrere Rechner (parallele Verarbeitung)



# Clusterhardware (HPC)

- Top 500: <http://www.top500.org/>
  - Hochparallele (und spezialisierte) Supercomputer (spezielle Hardware, Speicherverwaltung, Netzwerk, OS usw.)
  - Computercluster (COTS -> commercial off-the-shelf -> kommerzielle Produkte aus dem Regal)
  - Der Schnellste: IBM BlueGene/L; 280,6 Tflop/s; 131072 CPUs; USA
  - Schnellster Deutsche: „JUBL“; IBM BlueGene Solution; 37,33 Tflop/s; 16384 CPUs; Forschungszentrum Jülich; Platz 13
- Green500-Rangliste: <http://www.green500.org>
  - Verhältnis Rechenleistung zu Energieverbrauch (Flops/Watt)



# Clustersoftware für Linux

- Betriebssystemerweiterung (spezieller Kernel)
  - Die Verteilung der Prozesse wird vom Kernel (Scheduler) übernommen
  - Mosix, openMosix
  - ClusterKnoppix (basiert auf openMosix)
- Auf Kommunikationsbibliotheken basierend:
  - Die Anwendung selbst übernimmt die Aufgabenverteilung.
  - Beowolf: Rahmenbezeichnung/-Projekt für Linux-Clusters deren Clients via TCP/IP kommunizieren
    - PVM – Parallel Virtual Machine
    - MPI – Message Passing Interface



# Aufgabenverteilung im HPC-Cluster (1)

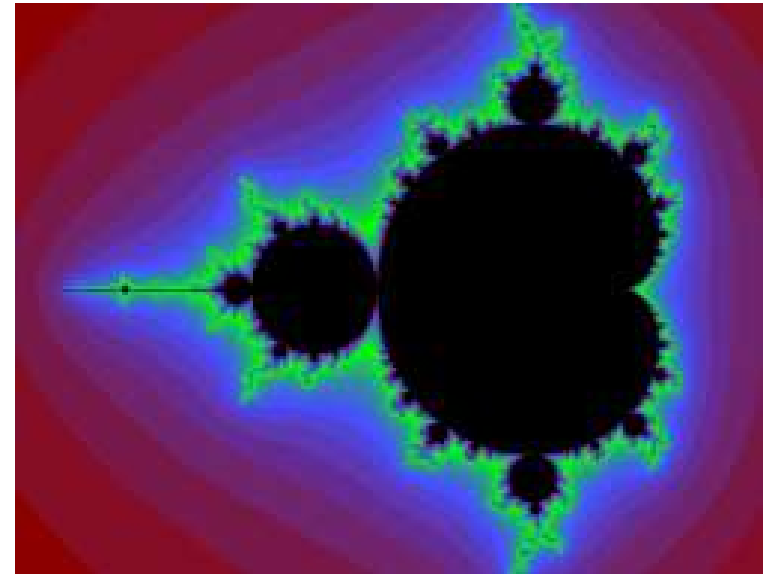
- Was muß erfüllt sein?
  - es muß eine aufteilbare Aufgabe sein
  - Teilergebnisse sollten (möglichst) nicht voneinander abhängen
- Beispiele:
  - paralleles Übersetzen von Programmen
  - Schach
  - chaotische Systeme, Simulationen
  - Zahlenfolgen (z.B. Pi)
  - Bild-/ Ton-/Video-Berechnungen
  - uvm.



# Aufgabenverteilung im HPC-Cluster (1)

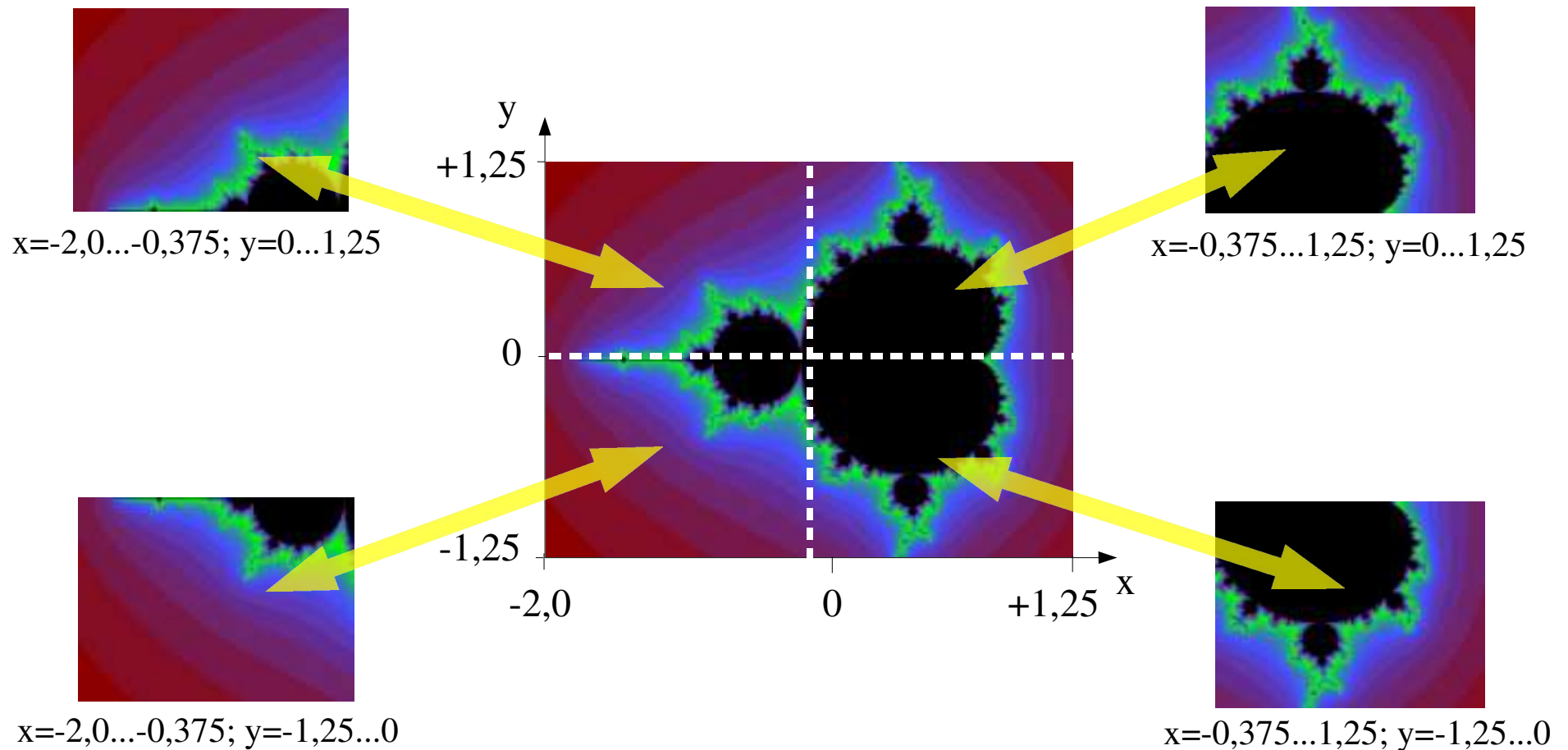
- Beispiel „Mandelbrot“-Menge:

```
for(y=-1.25; y<1.25; y+=0.005) {  
  for(x=-2.0; x<1.25; x+=0.005) {  
    az = bz = z = 0;  
    do {  
      az = sqr(az)-sqr(bz)+x;  
      bz = az1*bz*2+y;  
      az1 = az;  
      g = sqr(az)+sqr(bz);  
      z = z+1;  
    } while(g<4 && z<=75);  
    if(z<75) aa_putpixel(c, (int)(x*200+400), (int)(y*200+240), z);  
  }  
}
```



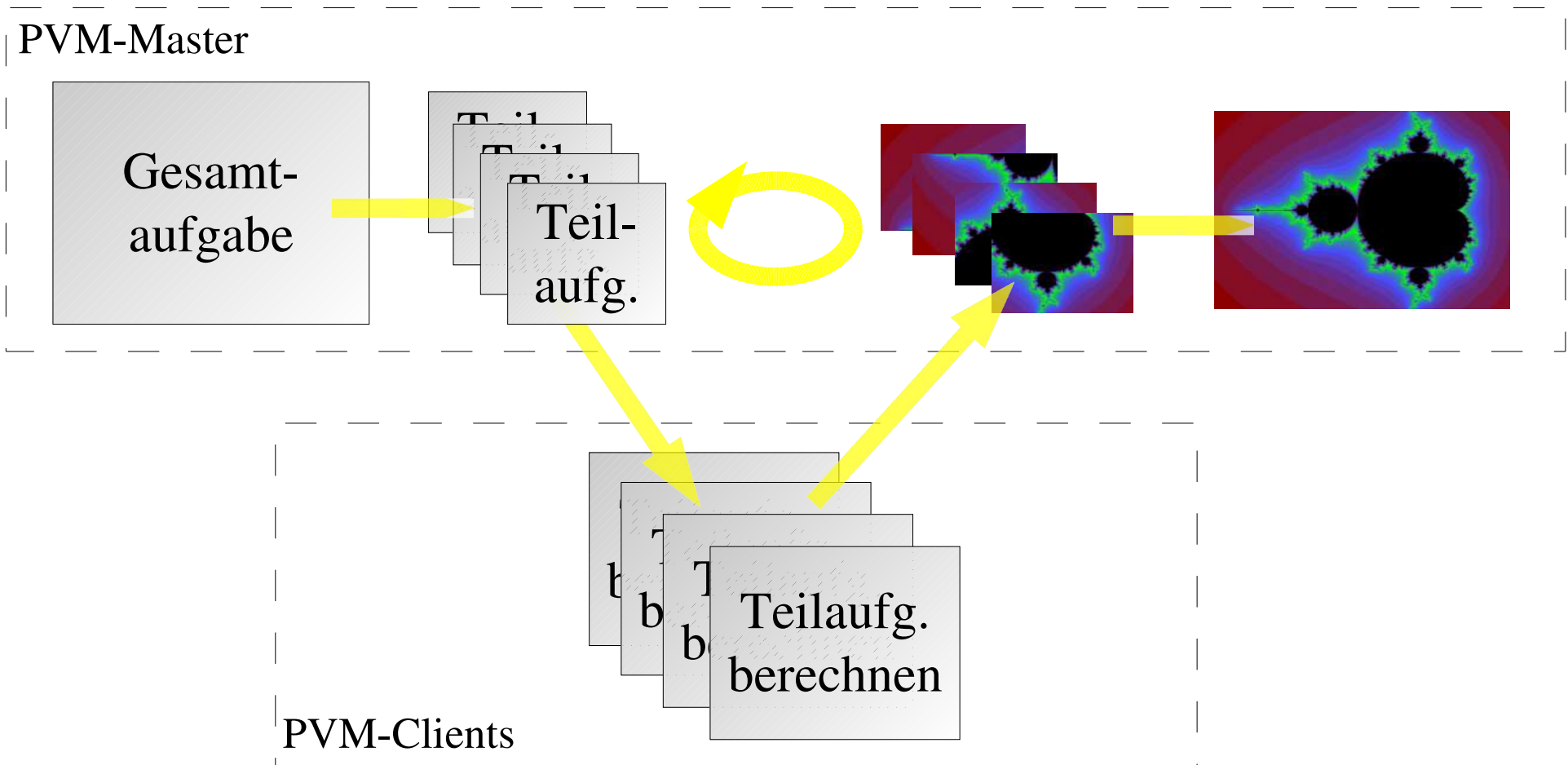


# Aufgabenverteilung im HPC-Cluster (2)





# Aufgabenverteilung im HPC-Cluster (3)



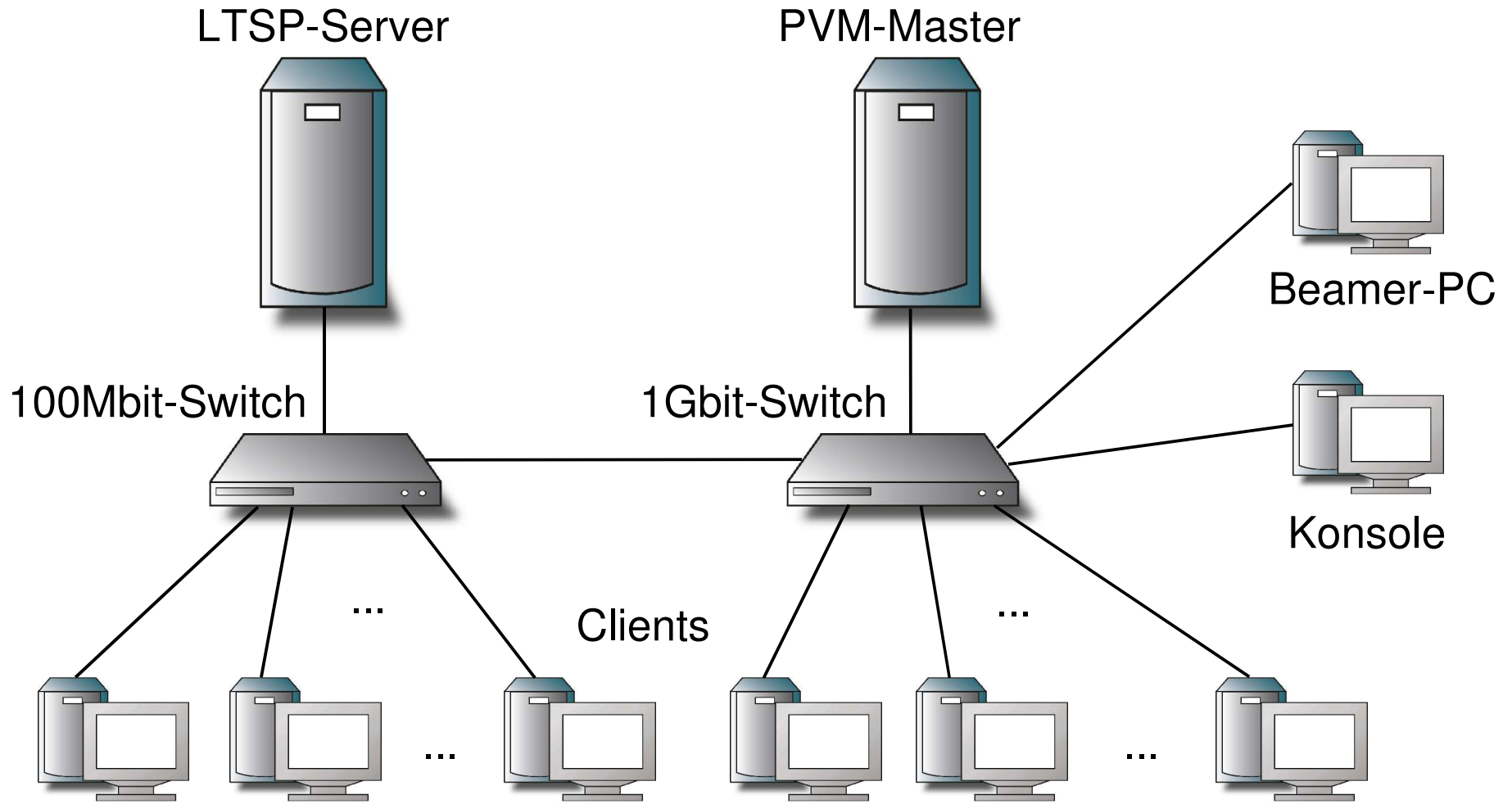




- BraLUG e.V.
- Idee
- Theorie
- **Umsetzung**
- Ergebnis
- Links



# Topologie (1)





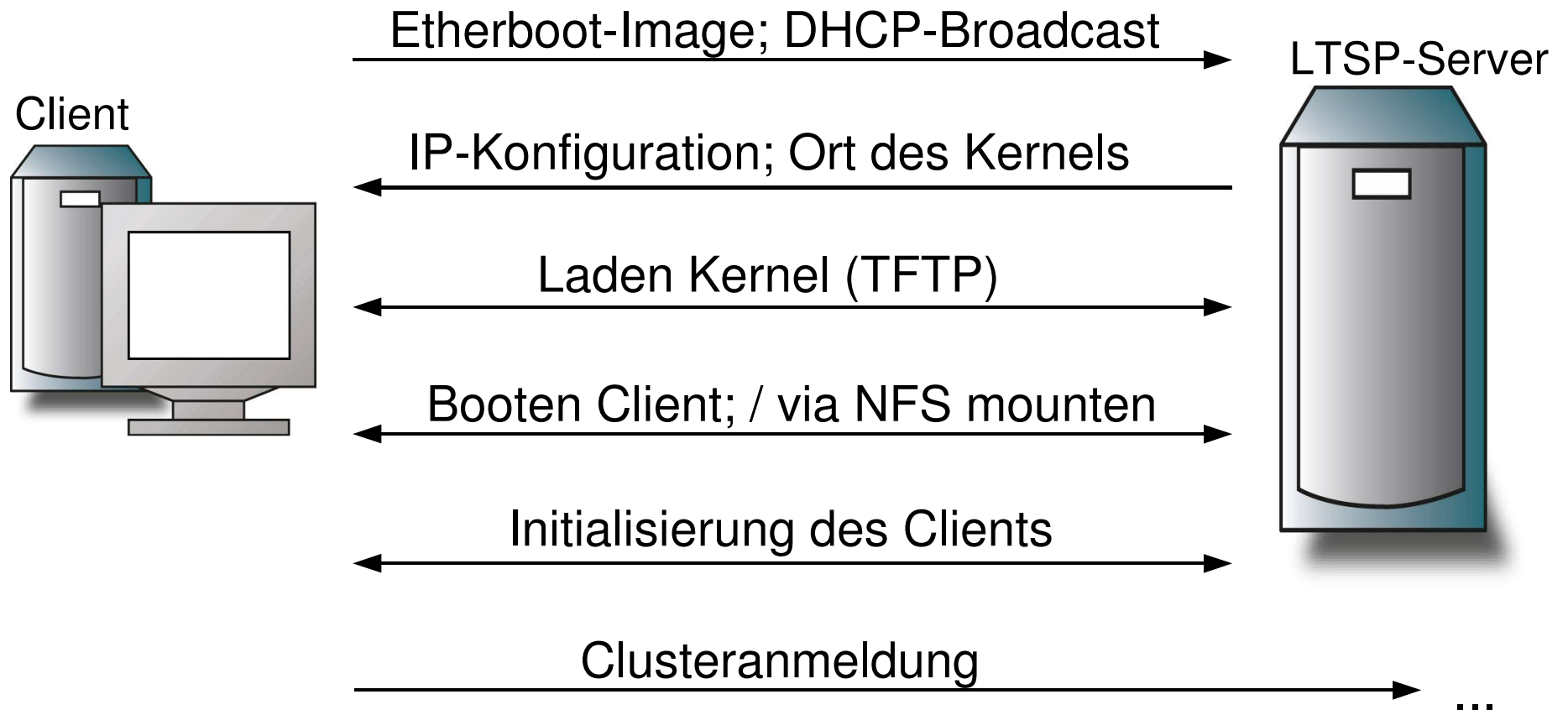
## Topologie (2)

- Software auf dem LTSP-Server:
  - Linux
  - LTSP-Paket (modifiziert, u.a. automat. Rechnernamenvergabe)
  - Netzwerkdienste (DHCP, DNS, NFS, TFTP)
- Software auf dem PVM-Master:
  - Linux, sshd
  - PVM, PVMPOV
- Clients:
  - Etherbootimage (ROM oder Diskette) bzw. PXE (Preboot Execution Environment)



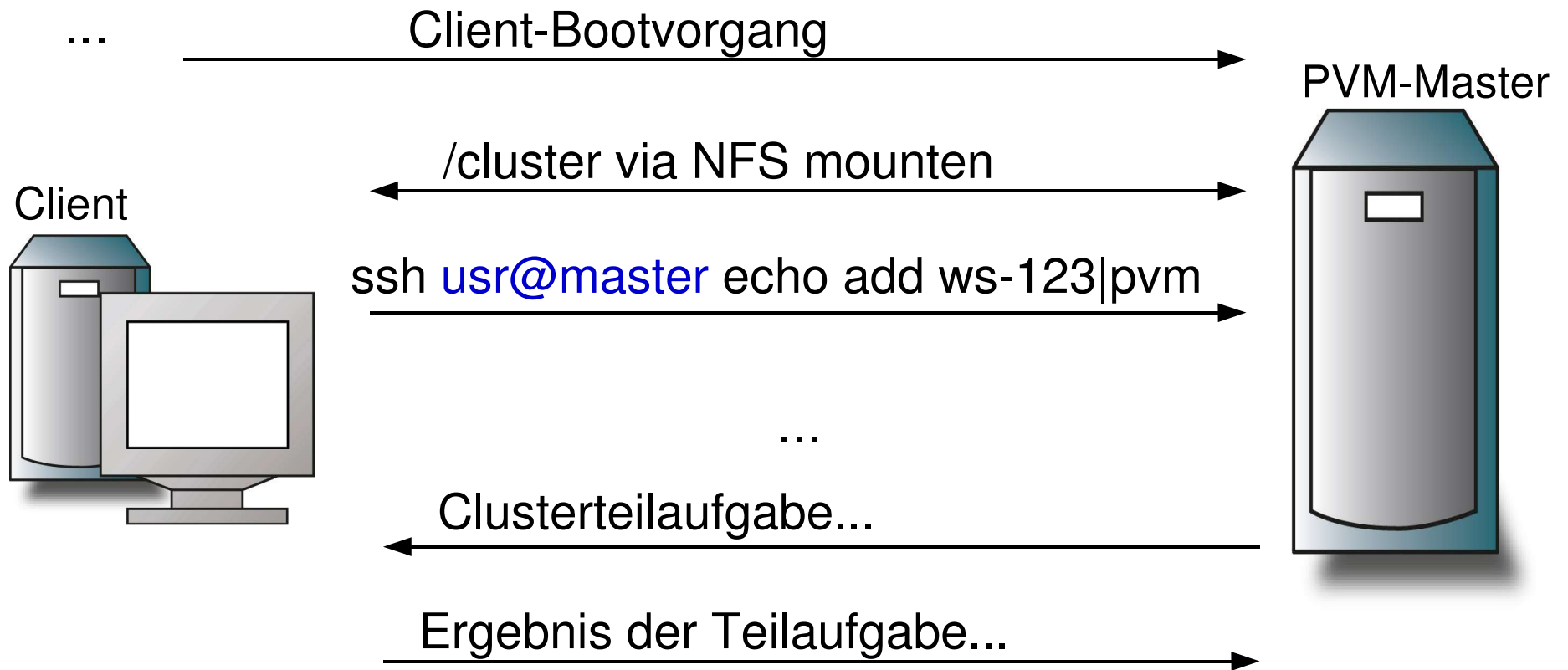


# Bootkonzept (Clientstart)





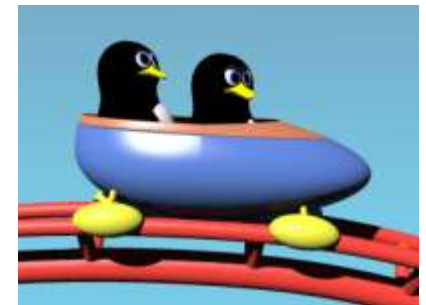
# Bootkonzept (Clusteranmeldung)





# POVRAY

- Freie Software
- POV: Persistence of Vision - Nachleuchten eines Bildes
- RAY: Ray to trace – Strahl zurückverfolgen
- Berechnet Bilder unter Berücksichtigung der elementaren Naturphänomene (Lichtquelle, Texturen, optische Brechung, Schatten etc.)
- Input: Szenenbeschreibung in Textform (Scene Description Language (SDL))
- PVMPOV:
  - Erweiterung von POV-Ray für PVM-Cluster
  - leider nicht ganz „brandaktuell“





# Clusterlog-Dateien

- cluster\_info (Autor: Markus Dahms; BraLUG)
  - „ortet“ jeden Client im PVM-Cluster und sammelt Informationen über diese ein (Rechnername, Prozessortyp, RAM, rudimentärer Benchmark etc.)
- „BraLUG“-Modifikation von PVMPOV:
  - Protokollierung der von PVMPOV an die Clients verteilten Rechenpakete
  - Protokollierung von Clientname, Framenummer, Frameausschnitt, Sende-/Empfangszeitpunkt des Paketes relativ zum Start der Berechnung des Filmes/Posters
  - Ausgabe in formatierter Textdatei

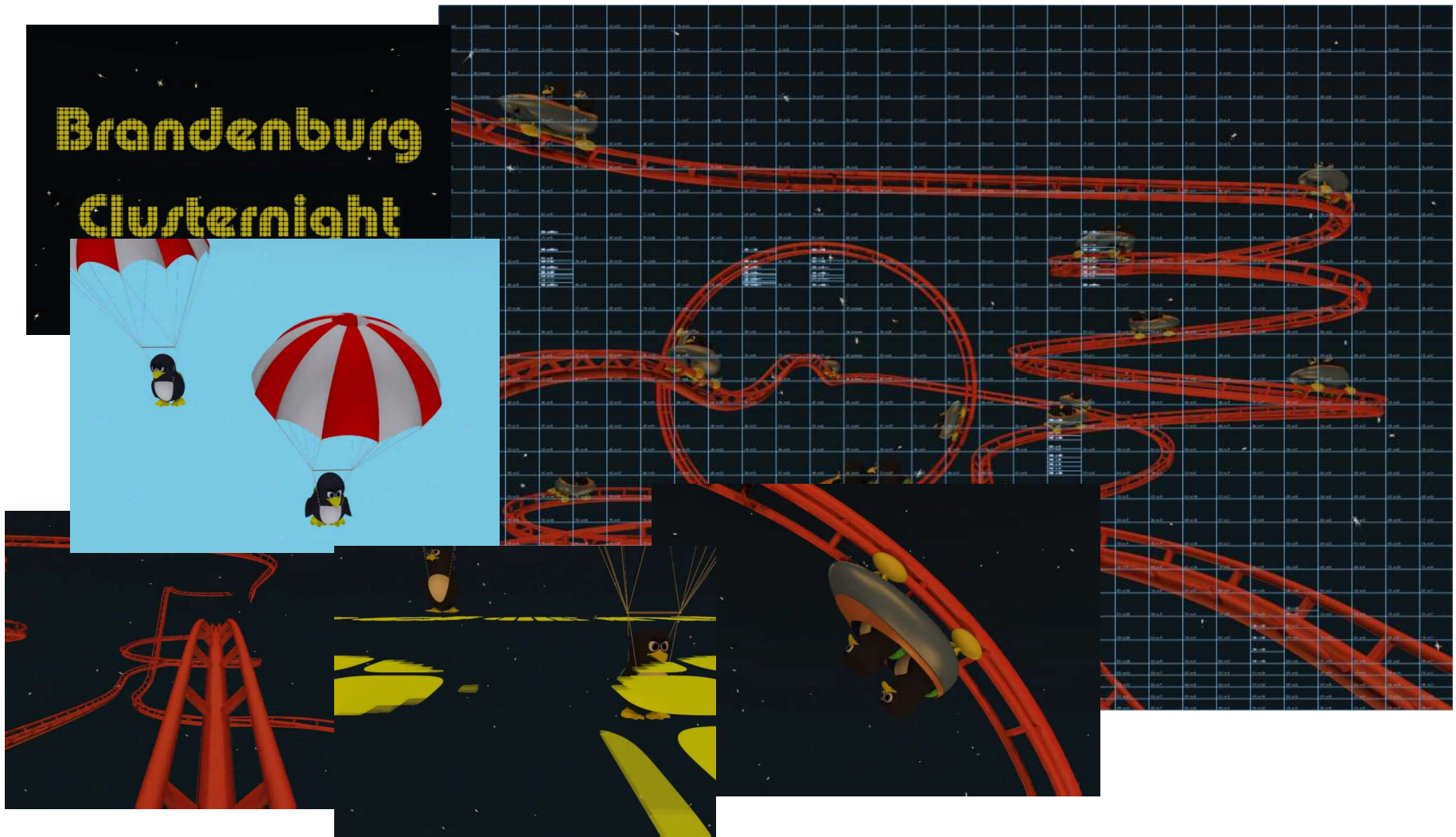




- BraLUG e.V.
- Idee
- Theorie
- Umsetzung
- **Ergebnis**
- Links



# Eine Videosequenz, ein Poster...





## Ein paar Zahlen...

- Cluster-Clients:
  - Bis zu 45 verschiedene Rechner
    - Stärkster Rechner: Pentium 4; 3,20GHz
    - Langsamster Rechner: VIA C3; 533Mhz
- Anzahl berechneter Bilder:
  - Video: 2500 Frames; 720x540 Pixel pro Bild; 3,2 Mio. Pakete
  - Poster: 9600x7200 Pixel; 839 Pakete
- Rechenzeiten:
  - Video: durchschnittlich ca. 40 Rechner; 11855s (ca. 3h 17min)
  - Poster: 40 Rechner; 2138s (ca. 36min)



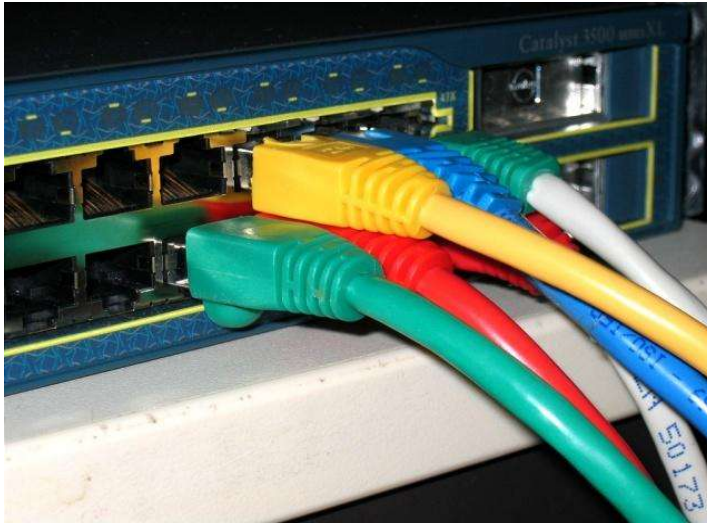
# Ein kleiner Dokumentarfilm...





## Probleme?

- Netzwerk: teilweise chaotische und nicht durchdachte Skalierung
- „utopische“, nicht unterstützte Netzwerkkarten (Stand 2005!)
- Fehlendes Feintuning des Clusters (Aufgabenverteilung je nach Leistung des jeweiligen Clients)





# Der unendlich große Cluster?

- Frage:
  - Ist der BraLUG-Cluster unendlich erweiterbar?
- Antwort:
  - Rein technisch gesehen (mit entsprechendem Aufwand), Ja!
  - Nutzbringend, Nein, denn:
    - Flaschenhals Netzwerk
    - Die unendlich teilbare Aufgabe?
  - Es ist das Optimum zwischen Rechenzeit/Datenvolumen der Teilaufgabe und dem dazu notwendigen Overhead zu finden!



- BraLUG e.V.
- Idee
- Theorie
- Umsetzung
- Ergebnis
- **Links**



## Weiterführende Informationen

- <http://www.bralug.de>
- <http://www.freifunk-brb.de>
- <http://www.linuxinfotag.brb.de>
- <http://www.bralug.de/clusternight>
- [http://bralug.de/clusternight/dokumente/film\\_bla\\_bla.pdf](http://bralug.de/clusternight/dokumente/film_bla_bla.pdf)
- [http://www.csm.ornl.gov/pvm/pvm\\_home.html](http://www.csm.ornl.gov/pvm/pvm_home.html)
- <http://www.ltsp.org>
- <http://www.povray.org>
- <http://pvmpov.sourceforge.net/>





***Gesucht wird eine neue Rechenaufgabe!***

