

IPOP – IP over P2P

Virtual Networking for Grid Computing

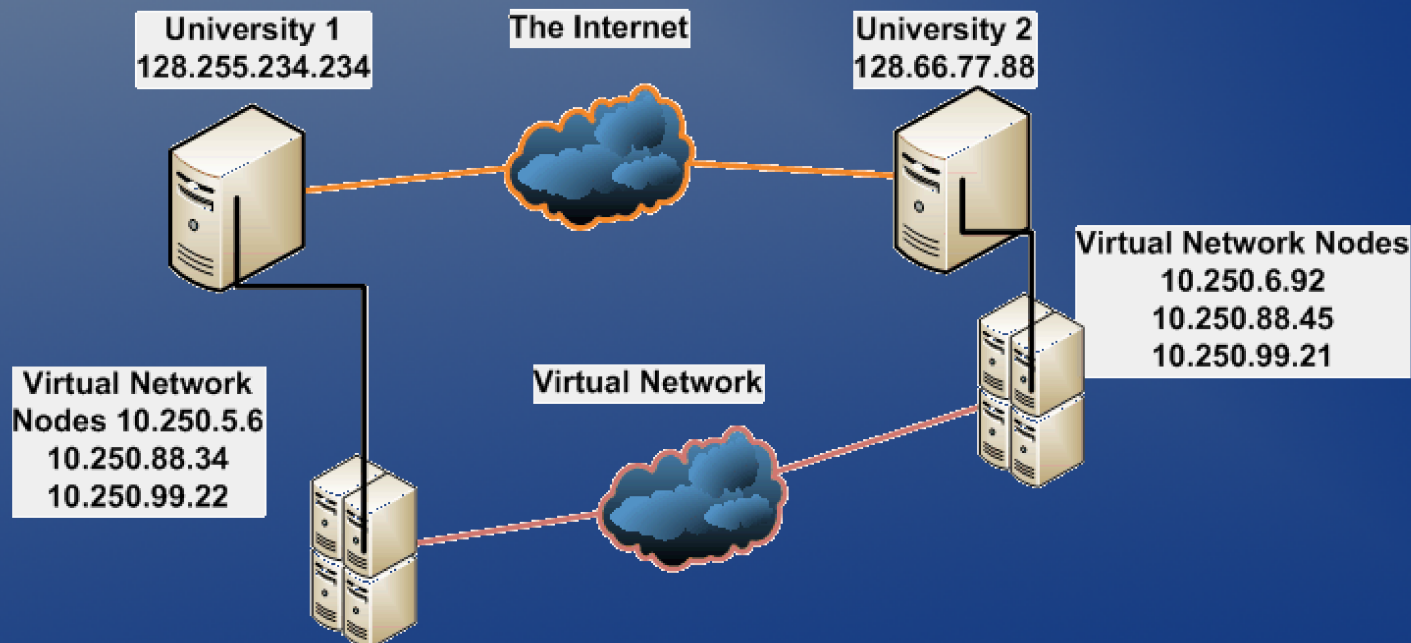
David Wolinsky
ACIS P2P Group
University of Florida

Outline

- Virtual Networking + Grid Computing
- Making It Easy
- Deploying in Clusters and Workstations
- Demonstration
- Performance
- Moving Forward...

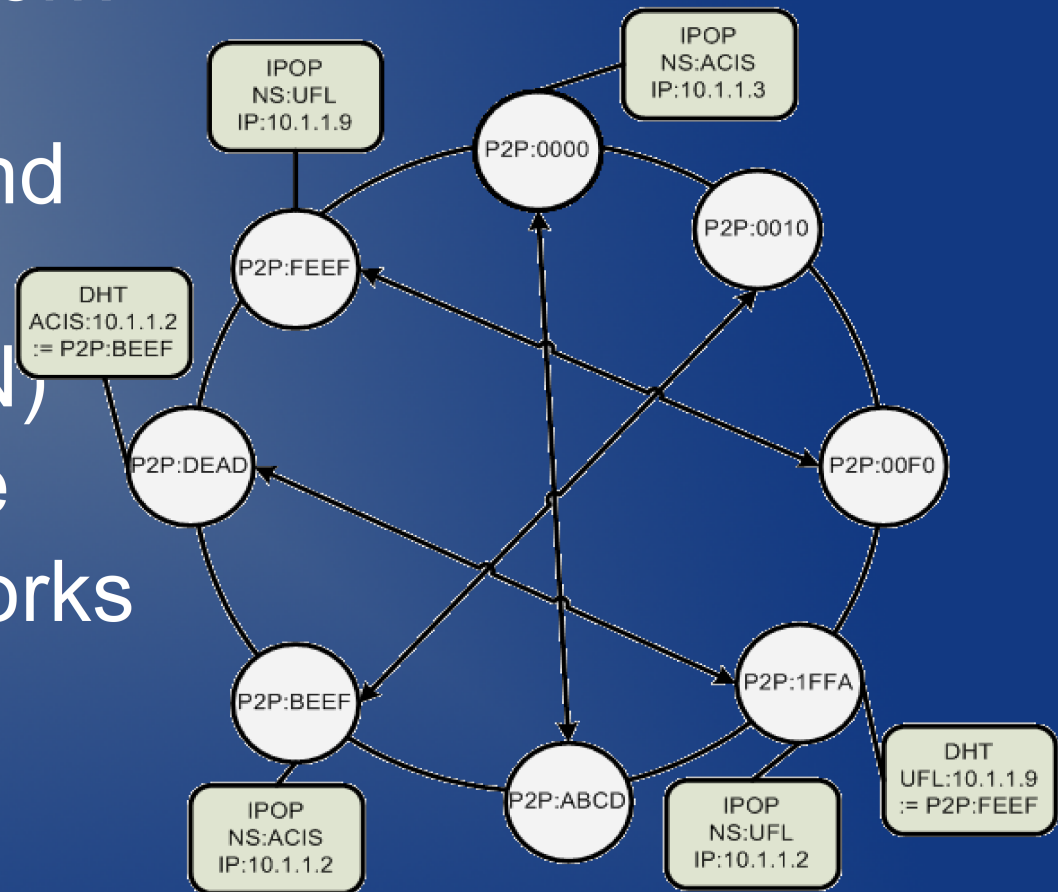
Virtual Networking in the Grid

- Unified layer 3 (IP) network for all machines
- Cross-site communication without a middleware broker



IPOP Architecture

- Structured P2P Network Overlay
- Provides tunneling and direct shortcuts
- NAT Traversal (STUN)
- Distributed data store
- Multiple Virtual Networks Per Overlay



Unique Features of IPOP

- We provide the infrastructure for you!
- How to connect:
 - Specify a network to join
 - Start IPOP and you are dynamically given an address in that network
 - You can find all resources by having a central manager or using multicast
- IPOP takes care of configuring routes, address allocation, and DNS for you!
- Support for individual workstations and clusters

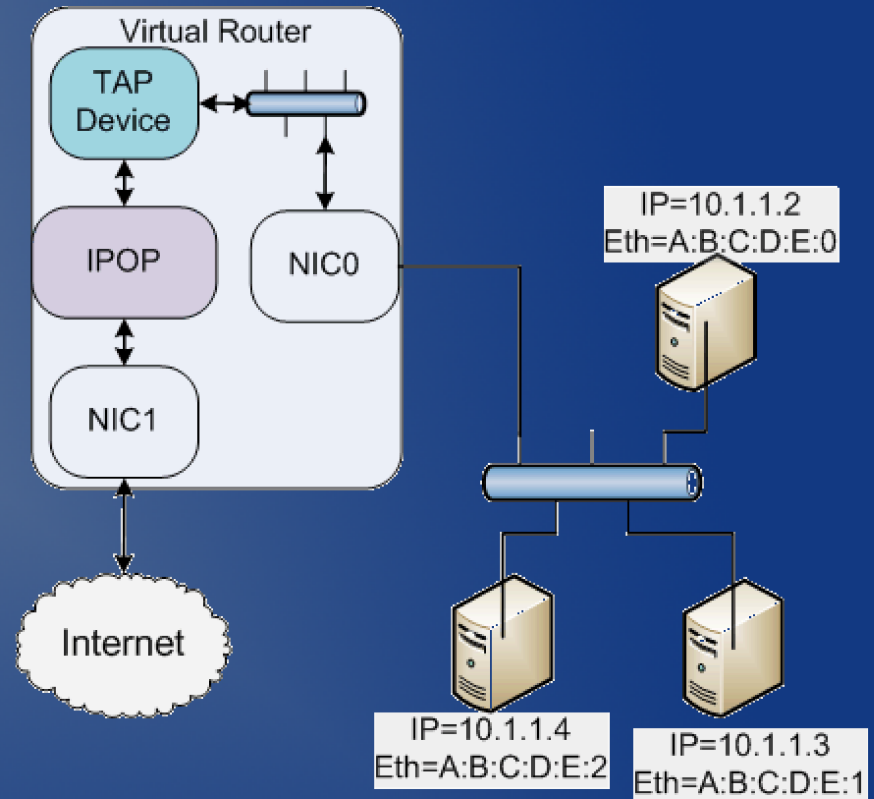
Setup in a Workstation

- Each machine has IPOP running locally
- Machine has IPOP and “Internet” connectivity
- Low latency



Setup in a Cluster

- Single IPOP instance for entire cluster
- Machines have connectivity over IPOP
- May have “Internet” connectivity if there is an “Internet” router



- Limited to no resource configuration
- Reduced virtualization overhead

```
griduser@C240195038: /home/griduser
griduser@C240195038:~$ /sbin/ifconfig
eth0      Link encap:Ethernet  HWaddr 08:00:27:B4:FF:B4
          inet addr:10.0.2.15  Bcast:10.0.2.255  Mask:255.255.255.0
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:4986 errors:0 dropped:0 overruns:0 frame:0
          TX packets:5462 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:1260249 (1.2 MiB)  TX bytes:1007914 (984.2 KiB)
          Interrupt:11 Base address:0xc020

lo        Link encap:Local Loopback
          inet addr:127.0.0.1  Mask:255.0.0.0
          UP LOOPBACK RUNNING  MTU:16436  Metric:1
          RX packets:185 errors:0 dropped:0 overruns:0 frame:0
          TX packets:185 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:23496 (22.9 KiB)  TX bytes:23496 (22.9 KiB)

tapipop   Link encap:Ethernet  HWaddr 00:FF:09:D2:21:76
          inet addr:242.240.195.38  Bcast:242.255.255.255  Mask:255.0.0.0
          UP BROADCAST RUNNING MULTICAST  MTU:1200  Metric:1
          RX packets:512 errors:0 dropped:0 overruns:0 frame:0
          TX packets:486 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:500
          RX bytes:405269 (395.7 KiB)  TX bytes:50157 (48.9 KiB)

griduser@C240195038:~$ hostname -f
C240195038.ipop
griduser@C240195038:~$
```

Taskbar: debian, xmessage, griduser@C240..., griduser@C240195..., 10:23:33 AM

ent/
ig/
nf
fil
us.
anc
READ
bu
anc
ce i
Gt
ffi
9-0
fco

```

griduser@C240195038: /home/griduser

slot10C185000234.i LINUX INTEL Unclaimed Idle 0.000 1485 0+03:15:05
slot20C185000234.i LINUX INTEL Unclaimed Idle 0.000 1485 17+11:44:48
slot10C185025227.i LINUX INTEL Unclaimed Idle 0.000 620 0+00:55:05
slot20C185025227.i LINUX INTEL Unclaimed Idle 0.000 620 1+08:41:03
C186098058.ipop LINUX INTEL Unclaimed Idle 0.000 2971 0+00:55:04
slot10C186185058.i LINUX INTEL Unclaimed Idle 0.000 1485 0+00:50:04
slot20C186185058.i LINUX INTEL Unclaimed Idle 0.000 1485 6+08:40:33
slot10C191158136.i LINUX INTEL Unclaimed Idle 0.000 1775 0+02:15:05
slot20C191158136.i LINUX INTEL Unclaimed Idle 0.000 1775 0+18:07:09
slot10C193080211.i LINUX INTEL Unclaimed Idle 0.000 620 17+11:10:58
slot20C193080211.i LINUX INTEL Unclaimed Idle 0.000 620 0+00:10:05
C194165156.ipop LINUX INTEL Unclaimed Idle 0.000 3225 0+03:25:05
slot10C195096083.i LINUX INTEL Unclaimed Idle 0.000 1485 0+03:40:04
slot20C195096083.i LINUX INTEL Unclaimed Idle 0.000 1485 3+19:27:53
slot10C206214018.i LINUX INTEL Unclaimed Idle 0.000 1775 0+14:35:52
slot20C206214018.i LINUX INTEL Unclaimed Idle 0.000 1775 0+02:45:06
slot10C211145230.i LINUX INTEL Unclaimed Idle 0.000 1775 0+17:56:31
slot20C211145230.i LINUX INTEL Unclaimed Idle 0.000 1775 0+02:25:06
slot10C223073125.i LINUX INTEL Unclaimed Idle 0.000 620 0+00:50:04
slot20C223073125.i LINUX INTEL Unclaimed Idle 0.000 620 3+08:37:05
slot10C224255233.i LINUX INTEL Unclaimed Idle 0.000 1775 0+02:20:04
slot20C224255233.i LINUX INTEL Unclaimed Idle 0.000 1775 0+18:08:52
slot10C228022109.i LINUX INTEL Unclaimed Idle 0.030 1485 0+00:25:04
slot20C228022109.i LINUX INTEL Unclaimed Idle 0.000 1485 9+00:37:36
slot10C232105165.i LINUX INTEL Unclaimed Idle 0.000 1775 0+02:30:04
slot20C232105165.i LINUX INTEL Unclaimed Idle 0.000 1775 0+14:35:25
slot10C235052143.i LINUX INTEL Unclaimed Idle 0.000 1485 0+00:45:04
slot20C235052143.i LINUX INTEL Unclaimed Idle 0.000 1485 7+16:51:09
slot10C235252250.i LINUX INTEL Unclaimed Idle 0.000 1485 4+08:51:42
slot20C235252250.i LINUX INTEL Unclaimed Idle 0.000 1485 0+00:45:05
C240195038.ipop LINUX INTEL Owner Idle 0.320 249 0+00:05:12
slot10C245091047.i LINUX INTEL Unclaimed Idle 0.000 1485 0+02:00:04
slot20C245091047.i LINUX INTEL Unclaimed Idle 0.000 1485 9+09:59:04
slot10C254063065.i LINUX INTEL Unclaimed Idle 0.000 1775 0+00:40:07
slot20C254063065.i LINUX INTEL Unclaimed Idle 0.000 1775 0+16:07:10

Total Owner Claimed Unclaimed Matched Preempting Backfill
INTEL/LINUX 136 2 1 133 0 0 0
Total 136 2 1 133 0 0 0
griduser@C240195038:~$

```

debian
0
X xmessage
griduser@C240...
griduser@C240195...
10:24:03 AM


```

griduser@C240195038: /home/griduser
griduser@C240195038:~$ ping C174169130
PING C174169130.ipop (242.174.169.130) 56(84) bytes of data.
64 bytes from C174169130.ipop (242.174.169.130): icmp_seq=2 ttl=64 time=581 ms
64 bytes from C174169130.ipop (242.174.169.130): icmp_seq=3 ttl=64 time=439 ms
64 bytes from C174169130.ipop (242.174.169.130): icmp_seq=4 ttl=64 time=439 ms
64 bytes from C174169130.ipop (242.174.169.130): icmp_seq=5 ttl=64 time=581 ms
64 bytes from C174169130.ipop (242.174.169.130): icmp_seq=6 ttl=64 time=529 ms
64 bytes from C174169130.ipop (242.174.169.130): icmp_seq=7 ttl=64 time=422 ms
64 bytes from C174169130.ipop (242.174.169.130): icmp_seq=8 ttl=64 time=506 ms
64 bytes from C174169130.ipop (242.174.169.130): icmp_seq=9 ttl=64 time=548 ms
64 bytes from C174169130.ipop (242.174.169.130): icmp_seq=10 ttl=64 time=425 ms
64 bytes from C174169130.ipop (242.174.169.130): icmp_seq=11 ttl=64 time=479 ms
64 bytes from C174169130.ipop (242.174.169.130): icmp_seq=12 ttl=64 time=554 ms
64 bytes from C174169130.ipop (242.174.169.130): icmp_seq=13 ttl=64 time=47.5 ms
64 bytes from C174169130.ipop (242.174.169.130): icmp_seq=14 ttl=64 time=39.3 ms
64 bytes from C174169130.ipop (242.174.169.130): icmp_seq=15 ttl=64 time=37.9 ms
64 bytes from C174169130.ipop (242.174.169.130): icmp_seq=16 ttl=64 time=34.2 ms
64 bytes from C174169130.ipop (242.174.169.130): icmp_seq=17 ttl=64 time=39.7 ms
64 bytes from C174169130.ipop (242.174.169.130): icmp_seq=18 ttl=64 time=50.5 ms
^C
--- C174169130.ipop ping statistics ---
18 packets transmitted, 17 received, 5% packet loss, time 17215ms
rtt min/avg/max/mdev = 34.278/338.699/581.804/224.589 ms
griduser@C240195038:~$

```

| | | | | | | | |
|--------------------|-------|-------|-----------|------|-------|------|-------------|
| slot20C160231102.i | LINUX | INTEL | Unclaimed | Idle | 0.000 | 886 | 0+01:30:06 |
| slot10C174093235.i | LINUX | INTEL | Unclaimed | Idle | 0.000 | 1485 | 0+03:45:04 |
| slot20C174093235.i | LINUX | INTEL | Unclaimed | Idle | 0.000 | 1485 | 3+19:34:58 |
| slot10C174136132.i | LINUX | INTEL | Unclaimed | Idle | 0.000 | 1485 | 14+00:18:29 |
| slot20C174136132.i | LINUX | INTEL | Unclaimed | Idle | 0.000 | 1485 | 0+00:05:05 |
| slot10C174169130.i | LINUX | INTEL | Unclaimed | Idle | 0.000 | 1775 | 0+15:11:57 |
| slot20C174169130.i | LINUX | INTEL | Unclaimed | Idle | 0.020 | 1775 | 0+01:35:05 |
| slot10C178183202.i | LINUX | INTEL | Unclaimed | Idle | 0.000 | 1485 | 0+02:05:04 |
| slot20C178183202.i | LINUX | INTEL | Unclaimed | Idle | 0.000 | 1485 | 12+14:01:26 |
| slot10C183190251.i | LINUX | INTEL | Unclaimed | Idle | 0.000 | 1485 | 0+03:05:05 |
| slot20C183190251.i | LINUX | INTEL | Unclaimed | Idle | 0.000 | 1485 | 17+13:07:42 |

0

xmessage

griduser@C240195...

griduser@C240...

10:27:35 AM

Documentation

- <http://www.grid-appliance.org/documentation/BasicNodeHowTo.html>
- <http://www.grid-appliance.org/documentation/>

Below is a sample DHCPServerConfig in XML format:

```
<DHCPServerConfig>
  <Namespace>Generic</Namespace>
  <netmask>255.255.0.0</netmask>
  <pool>
    <lower>10.250.0.0</lower>
    <upper>10.250.255.255</upper>
  </pool>
  <DHCPReservedIPs>
    <DHCPReservedIP>
      <ip>0.0.0.1</ip>
      <mask>0.0.0.255</mask>
    </DHCPReservedIP>
  </DHCPReservedIPs>
  <leasetime>3600</leasetime>
</DHCPServerConfig>
```

Ipop requires an **IpopConfig** xml file in addition to the NodeConfig xml file. Below is a sample config followed by a description of the fields.

```
<IpopConfig>
  <IpopNamespace>Generic</IpopNamespace>
  <VirtualNetworkDevice>tap0</VirtualNetworkDevice>
  <AddressData>
    <Hostname>Example</Hostname>
  </AddressData>
  <EnableMulticast>true</EnableMulticast>
</IpopConfig>
```

IpopNamespace must match the namespace specified in the

Now line by line...

Performance

| | Bandwidth (mbps) | Latency (ms) |
|------------------|---------------------|-----------------|
| Host | 940 | .3 |
| IPOP Workstation | 205 | .6 |
| IPOP Router | 220 | .8 |

Conclusions

- Coming up next:
 - Security is available working on improving deployment
 - Automated tools for creating network configurations
 - Working on support for more operating systems
- Going on now:
 - Used in real systems:
 - PlanetLab - 500 boot strap nodes
 - Archer – 250 active Grid Appliances
 - <http://www.grid-appliance.org>
 - <http://www.ipop-project.org>
 - Completely free and Open Source (GPLv2)
 - Product of ACIS P2P Team
- Questions?