



Distributed Internet Archiving

Unleash the power of existing systems

DIAP – Distributed Internet Archive Protocol

Talk by author of DIAP

Damian L. Brasher 4th April 2008

Hampshire Lug Group



Hampshire Linux User Group

<http://www.hants.lug.org.uk>



Distributed Internet Archiving

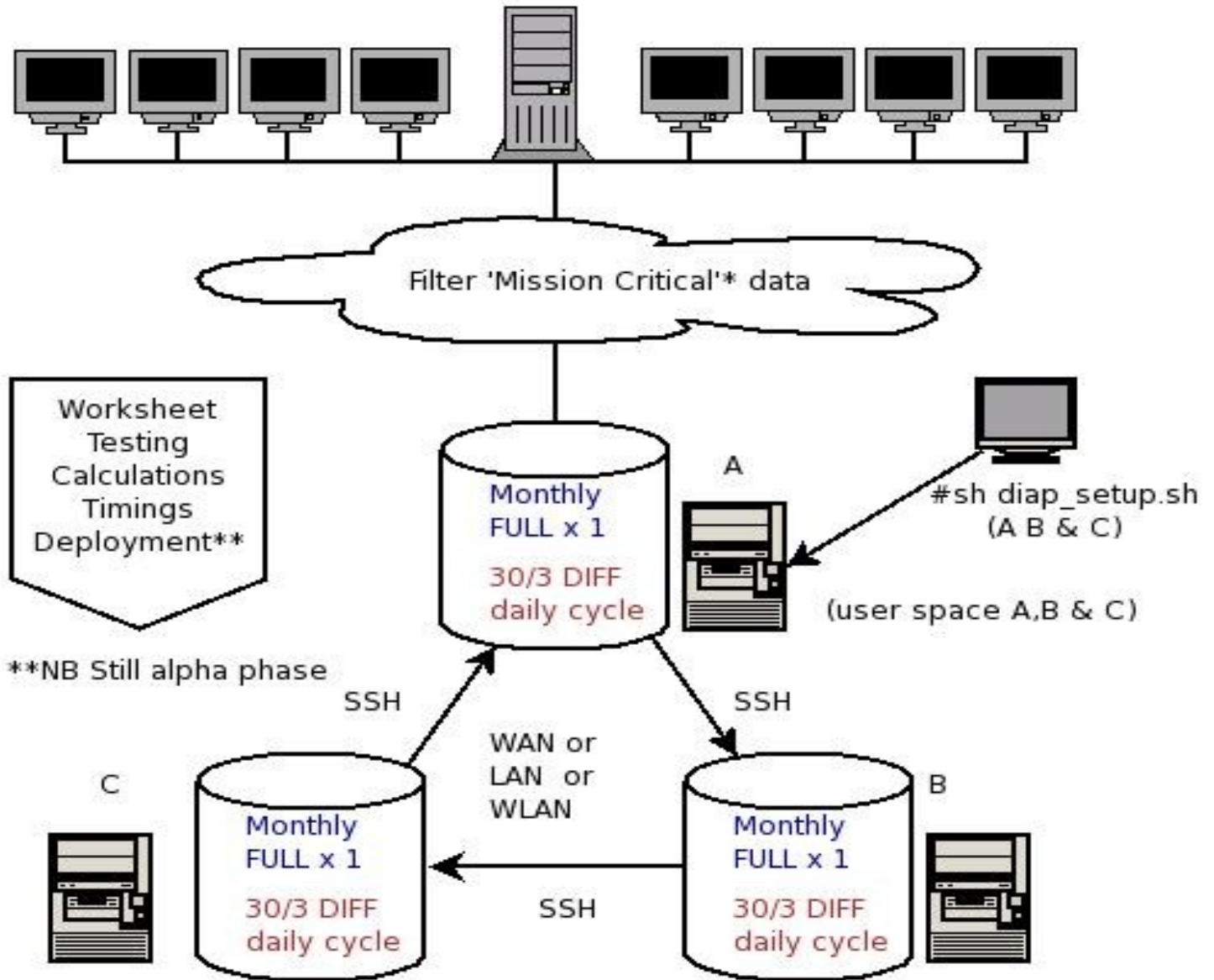
Unleash the power of existing systems

- Distributed Internet Archiving Protocol (Program)
- Began 3 1/2 years ago, currently developed in free time
- I worked for a multi site charity where resources where sparse

- Applied for UK IP Property in the form of a TM
- Alpha development – Open Source – GPL Bash Toolkit – working prototype

- Emerging technology, DVTL Distributed Virtual Tape Library
- Thinking about the years ahead - energy conservation

- Create an **area** of very secure, managed and (almost) **free** archive space with unused bandwidth & disk space



**NB Still alpha phase

* Mission Critical is organisation subjective, depending on business requirements different data can be classified with higher or lower vitality vlaues. A staff payroll or contacts database could fall into this catagory where as a video of the staff party does not (perhaps).



Distributed Internet Archiving

Unleash the power of existing systems

Scripts - Prototype and development tool

- Allow the Admin to build 3 node DIAP system on Linux servers
- Data is 'dropped off' at delta ad0 - over time, trickle between 3 servers
- Different offices or in different parts of the world
- Script creates a user account filled with the necessary configuration using expect, ssh and other native tools
- Run from laptop or workstation:-

usage: `diap_setup.sh [-k accounts and keys] [-s setup dir and cron] [-u undo]`

- A way to reduce the use of Tape Drives?
- DPA compliance



```
diap@test1:~  
File Edit View Terminal Tabs Help  
diap@test1:~ x diap@test2:~ x diap@test3:~ x  
[diap@test1 ~]$ ls  
ad0 ad2 ad4 ad6 ad8 dump key_copy_a.sh key_from_c log_b  
ad1 ad3 ad5 ad7 aFull expect key_from_b log_a tab_a.sh  
[diap@test1 ~]$ cat dump  
-rwxr-x--- 1 diap users 156894 2008-04-01 21:15 expect  
-rwxr-x--- 1 diap users 1065 2008-04-01 21:15 key_copy_a.sh  
-rw-r--r-- 1 diap users 409 2008-04-01 21:15 key_from_b  
-rw-r--r-- 1 diap users 409 2008-04-01 21:16 key_from_c  
-rw-r--r-- 1 diap users 121543 2008-04-04 16:05 log_a  
-rw-r--r-- 1 diap users 11544 2008-04-04 16:00 log_b  
-rwxr-x--x 1 diap users 2333 2008-04-01 21:16 tab_a.sh  
ad0: total 4 -rw-r--r-- 1 diap users 0 2008-04-04 15:40 ad0test  
ad1: total 4 -rw-r--r-- 1 diap users 0 2008-04-04 12:40 ad0test  
ad2: total 4 -rw-r--r-- 1 diap users 0 2008-04-04 09:40 ad0test  
ad3: total 4 -rw-r--r-- 1 diap users 0 2008-04-04 06:40 ad0test  
ad4: total 4 -rw-r--r-- 1 diap users 0 2008-04-04 03:40 ad0test  
ad5: total 4 -rw-r--r-- 1 diap users 0 2008-04-04 00:40 ad0test  
ad6: total 4 -rw-r--r-- 1 diap users 0 2008-04-03 21:40 ad0test  
ad7: total 4 -rw-r--r-- 1 diap users 0 2008-04-03 19:40 ad0test  
ad8: total 0  
aFull: total 4 -rw-r--r-- 1 diap users 0 2008-04-04 15:40 aFulltest  
[diap@test1 ~]$
```

www.diap.org.uk



Distributed Internet Archiving

Unleash the power of existing systems

Project site www.diap.org.uk

Freshmeat (also just registered with SourceForge)

<http://freshmeat.net/projects/distributedinternetarchiveprotocol/>

Internet Draft written for the IETF

<http://www.ietf.org/internet-drafts/draft-brasher-diap-00.txt>

Developers list <http://soton.mpeforth.com/mailman/listinfo/diap-dev>

Announce list <http://soton.mpeforth.com/mailman/listinfo/interlinux>



Hampshire Linux User Group

<http://www.hants.lug.org.uk>