TUSK Server Architecture

TUSK Server Installations (OLD as of 7/13)

TUSK installs can range widely from a small, single server instance to a full blown, multi data center redundant setup. Here are three different configurations and how they relate. Most TUSK installations run on some form of RPM based Linux, either RedHat, Fedora or CentOS. Tufts University runs its installation of TUSK on Sun Solaris SPARC. These proposed configurations will work running either OS.

TUSK can be considered a LAMP application, as it uses Apache, MySQL and Perl on Linux or Solaris platforms. Currently we are limited to using Apache 1.3, mod_perl 1.3 and MySQL 4.1. Work is being done to move to MySQL 5.0. Once that is finished Apache and mod_perl will be addressed.

Small Installation

This is the simplest of TUSK installations. The Apache web server as well as the MySQL database run on the same system. http://demo.tusk.tufts.edu is such a type of installation which is hosted on a Red Hat Enterprise Linux virtual machine, with 3 core of 3 GHz CPU, 4 GB RAM, and 10 GB disk space with 300 Gig attached storage. It relies on a separate MySQL database server, also on a Red Hat Enterprise Linux virtual machine with 2 cores of 3 GHz CPU, 4 GB of RAM, and 10 disk space with a small database.

In the diagram, the backup server runs 2 copies of MySQL, one to service the test environment, the second is a MySQL replication server of the production data. rsync processes are run to copy over the TUSK content to the remote machine. In an event of a catastrophic failure of the production server, the backup server can assume the identity of the production server and bring the service back up.

The PPT2PDF DOC2HTML conversion PC is an optional part of the infrastructure. It is just a Windows XP box running Office 2007 plus some special glue software. If a site implements it, then when uploading PPT files, the author has the choice of just having the students download the PPT, just viewing PDF of each slide or both. With out it, then the only choice is for the students to download the PPT. The DOC2HTML will auto convert any Office doc file into a native TUSK HTML file instead of treating it as only a downloadable file.

Other TUSK schools run their installations on hardware like this:

> *Systems(s): 2*
> *Processors per system: (2) Intel Quad Core 2GHz*
> *Total Cores per system: 8*
> *Memory: 32GB*
> *Hard Drives(s): (2) 500GB SATA, Raid1*
> *NICs: (2) GigE ports*
>
> *Each of the above systems will be configured to run KVM virtualization software:*
> *One of the above systems will be configured with 2 KVM Guest hosts*
> *1 KVM Guest Host for TUSK*
> *1 KVM Guest Host for DB*
>
> *Example TUSK Guest Host:*
> *Processor(s): 2 cores*
> *Memory: 4GB*
> *Disk Configuration:*
> */ 10GB*
> */var/lib/mysql 40GB*
> */data 100GB*
>
> *Example DB Guest Host:*
> *Processor(s): 2 cores*
> *Memory: 4GB*
> *Disk Configuration:*
> */ 10GB*
> */var/lib/mysql 40GB*
> */data 100GB*
>
> *The second server will also run KVM virtualization software and serve as a backup for the first KVM server*
>
> *An Example of Server System(s) follow:*
>
> *Dell Power Edge R610:*
> *Processors: (2) Intel Quad Core E5620 2.4GHz*
> *Memory: 32 GB 2R PC3-8500 RDIMM(6x2GB)*
> *Hard Drives(s): (2) 146GB 15K SAS 2.5in*
> *NICs: (1) Dual Port 5709 Gigabit Ethernet PCIe Network Interface Card*
>
> *Additional Hardware:*
> *UPS: APC 3000VA Smart UPS SMT3000*
> *NAS: IOMEGA StoreCentre IX12 300R-8TB*
>
> *NOTE: The 100 plus GB partition is typically a NFS or NAS server mounted on /data*

Capacity

500 Active Users
7,500 Pieces of content

Medium Installation

A Medium TUSK installation consists of at least 2 Apache Web servers, 2 dedicated MySQL database servers (master and slave), a external LDAP/AD authentication server and a NFS file server. TUSK content is either stored in the MySQL database or on the NFS file server, so both web servers have access to the content at the same time. Traffic can be directed to one of the web servers via DNS Round Robin entries or a Load Balancer in the Data Center. An installation may want to have the Apache Web servers be multi homed where the web traffic is over the public network and the NFS/MySQL /LDAP traffic on a private network.

Large Installation

The Large TUSK installation is fully redundant between 2 Data Centers with a global Load Balancer, LDAP/AD replicated servers and MySQL in a Clustered setup. The NFS server is replaced by a SAN that is replicated as well.

It also contains a Cluster of Helix Streaming servers. If a site has quite a bit of Audio and Video content to stream to users, a streaming server may be useful. All other installations use the Apache web server to "stream" the Audio and Video content. TUSK can work well with either setup.

The UMLS indexing server now takes place of one of the Apache servers to run the indexing crontab job as well as accepting SCP connections from the PPT2PDF and DOC2HTML Windows PCs. It needs access to the shared SAN space and direct access to the MySQL database.

Tufts' TUSK installation is currently somewhere between a Medium and Large installation, with our long term goal to be fully redundant. Currently we have a NAS that has some failover capability to the 2nd Data Center, still running MySQL in a master slave configuration, one Helix Streaming Server and have a Data Center only Load Balancer.

Capacity

4,000+ Active Users
800,000+ Pieces of content

File	Modified
PDF File TUSK-Installation-S.pdf	Aug 26, 2009 by Jim Pirzyk
PDF File TUSK-Installation-M.pdf	Aug 26, 2009 by Jim Pirzyk
PDF File TUSK-Installation-L.pdf	Aug 26, 2009 by Jim Pirzyk

Download All