

Scalable web solutions: Use case: Regulatory reform in Vietnam on eZ Publish CMS



Ivo Lukač

www.netgen.hr feat. www.infomar.hr

Regulatory reform in Vietnam

- Finding procedures fast for public users
- Project by Vietnam government with foreign advisers, consultant & know-how
- Sense Consulting, Zagreb part of international team
- Netgen responsibilities:
 - Information architecture
 - Implementation
- Infomar responsibilities:
 - Infrastructure architecture
 - Installation, security, performance



Why scalability is important?



- 88 million people
- 21 million Internet users

Starting point (1)

Main parameters

- Amount of information
- Structure of information
- Number of visitors and editors
- Availability
- Response time
- Security
- Extra functionalities:
 - Search....

Starting point in use case

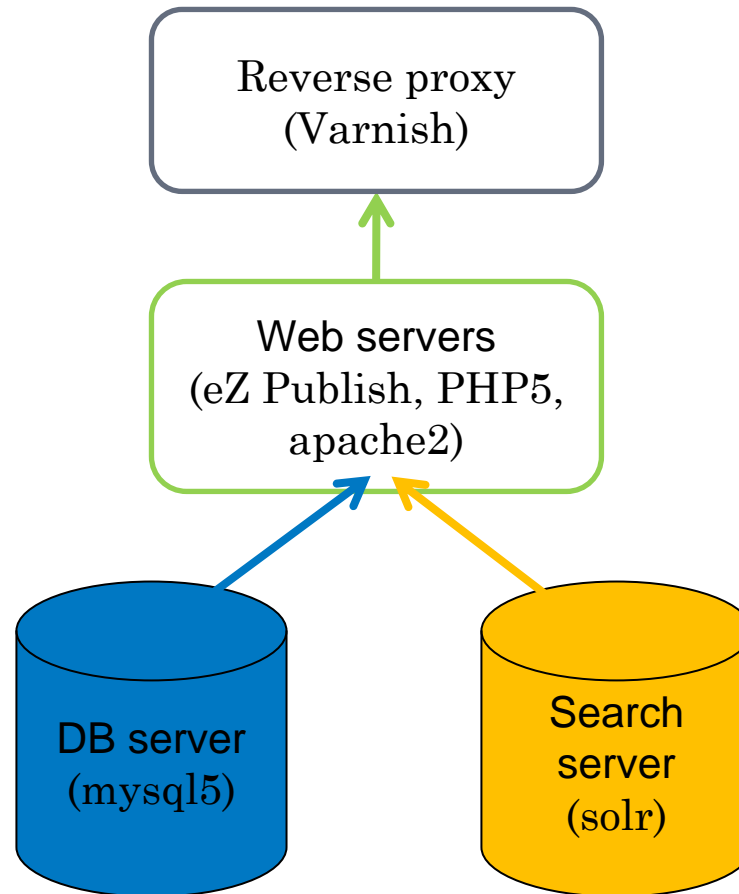
Main parameters

- 125 000 dossiers with additional data, 300 000 objects overall
- Tree structure based on task forces, lot of object relations
- Concurrent 80 editors in peaks
- 200GB of data & cache on disk, 4 million files
- 40GB data in database
- Usefull and fast search

Infrastructure

- Number of servers per task
- Data storage
- Services (apache, mysql, varnish, eZ)
- Clustering
- Redundance
- Loadbalancing
- Easy monitoring

Architecture overview (basic)



Infrastructure in use case

- 8 servers on disposition
- SAN iscsi + Oracle Cluster FS2
- 2 varnish, 3 web servers, 2 database servers, 1 search server
- Redundant varnish in failover mode
- Redundant web server monitored and balanced by varnish
- Mysql master – slave replication



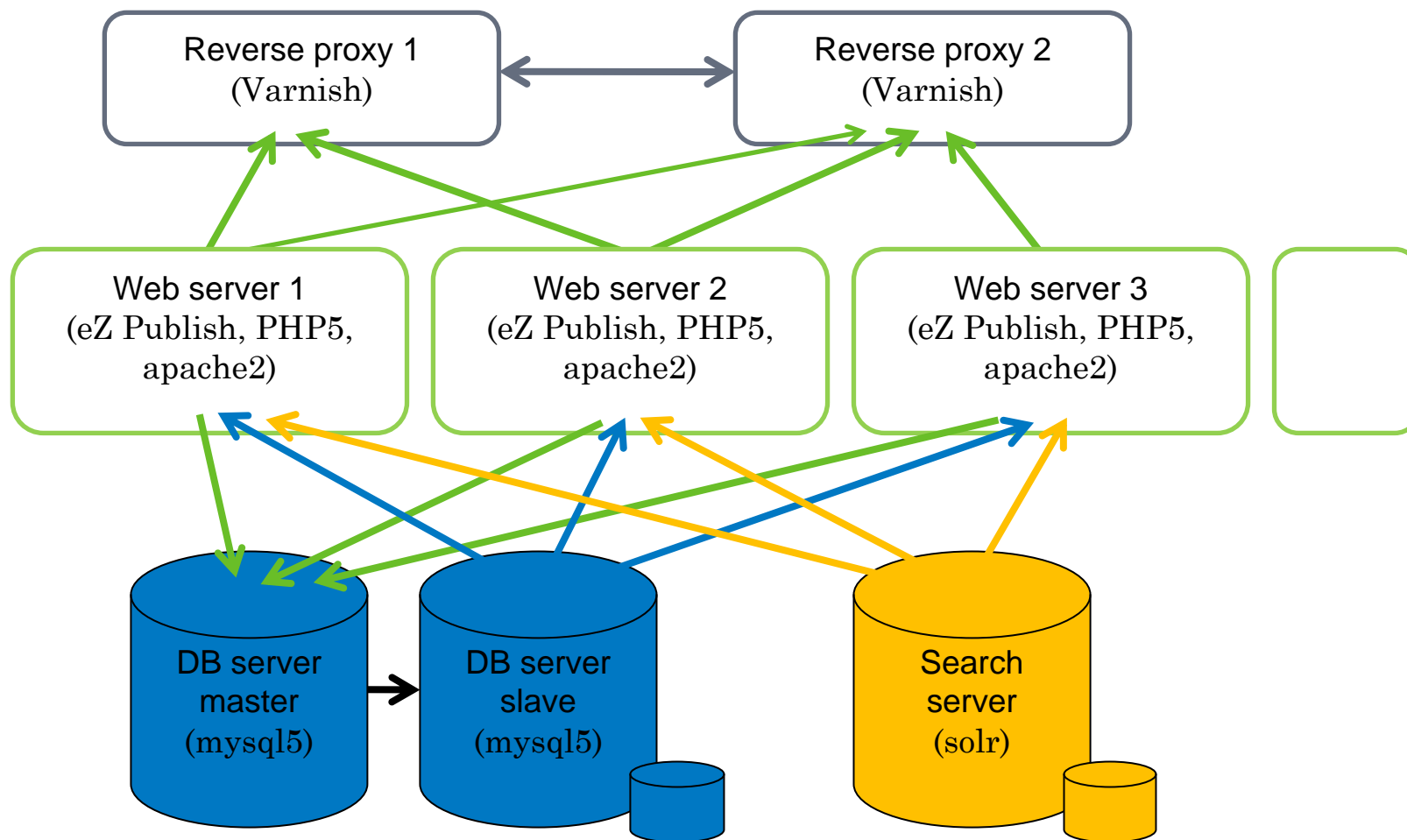
System considerations (1)

- iSCSI SAN:
 - Cheaper option,
 - Middle range performance ~1000 IOPS max, realno 500 IOPS (RAID5 10krpm SAS)
 - RAID1 – for MySQL
 - RAID5 – for files
 - Scalability possibilities: using more chasis, flashcopies, more controlers, more disks, transfrmation to RAID10

System considerations (2)

- Mimimized writing (in application)
 - To reduce cluster-wise flock() calls, use Memcached or similar
- Minimize HTTP calls to backend servers
 - Use Varnish HTTP reverse proxy
 - Maximize CMS caching
 - Consider authenticated users and sessions (ESI)
 - Varnish Cache ratio should be 80% at least

Architecture overview (current)



System implementation (1)

- HTTP reverse proxy-cachevi (Varnish):
 - DNS RR + HA IP cluster: automatic IP migration
 - Random load-balancing with is_alive
 - anti-DoS
- iSCSI (Open-iSCSI) on SAN:
 - IBM DS3300 as entry-level SAN iSCSI model
 - iSCSI + multipathing + OCFS2(Web) / Ext3 (SQL)
 - 500-600IOPS

System implementation (2)

- Web servers:
 - cluster-aware OCFS2 with no additional cluster services
 - Independent! Easier maintenance, easier scalability
- SQL servers:
 - Ext3
 - Problems with replication
 - Not so easy to scale, but not the bottleneck
 - For future: multimaster, SQL proxy, IPVS, etc.

Security, monitoring, redundancy

- Monitoring:
 - Checksums, logs: OSSEC HIDS
 - Service testing: Monit
 - Monitoring & reporting: ZenOSS
 - Authentication: Fail2Ban
- Redundancy:
 - 2x iSCSI ports on servers (multipath)
 - 2x LAN ports on servers (bonding)
 - 2x switches in cluster mode
 - 4x iSCSI ports on SAN (multipath, 2x controler)
 - 2x mgmt ports on SAN

Application implementation

- Information structure in eZ Publish
 - Hierarchy, relations, attributes
- Interface in regard to different types of caching (view, block, static, ESI)
- Search (ezfind / lucene / solr)



Implementation in use case (1)

- Based on tree structure with objects of different classes (class -> set of attributes)
- Relations between dossier objects heavily used
- Caching:
 - Template cache
 - Cache blocks
 - View cache
 - Static cache
 - ESI (in future)

Implementation in use case (2)

- Ajax based functions
 - Search suggestions
 - Tree browse (not used yet)
- Search – eZFind extension:
 - Based on Solr/Lucene
 - Results ranking based on relevance
 - Filtering
 - Faceting (not used yet)
 - Extremely fast

CƠ SỞ DỮ LIỆU QUỐC GIA VỀ THỦ TỤC HÀNH CHÍNH
Đăng nhập | Đăng ký

Trang đầu
Tìm kiếm TTCH
Tìm kiếm VB
Cơ quan
Tin tức
05/02/2010

Lời giới thiệu

Cơ sở dữ liệu quốc gia về thủ tục hành chính là hệ thống thông tin về thủ tục hành chính và các văn bản pháp luật quy định về thủ tục hành chính liên quan đến cá nhân và tổ chức được thiết lập trên cơ sở các Quyết định công bố thủ tục hành chính hoặc thủ tục giải quyết công việc của 24 bộ, ngành và 63 địa phương. Mục tiêu chính của cơ sở dữ liệu nhằm cung cấp một địa điểm duy nhất để người sử dụng có thể tìm kiếm các thủ tục hành chính quan tâm. Cơ sở dữ liệu sẽ tăng cường khả năng tiếp cận thông tin về các quy định, tăng tính minh bạch của hệ thống thể chế và thiết lập một cơ sở lịch sử về hệ thống thủ tục hành chính.

Cơ sở dữ liệu này sẽ tiếp tục được hoàn thiện trong các giai đoạn rà soát và giai đoạn thực thi các kiến nghị của đề án 30. Tổ công tác chuyên trách cải cách thủ tục hành chính của Thủ tướng Chính phủ kêu gọi mọi cá nhân và tổ chức tham gia đóng góp ý kiến đối với các quy định về thủ tục hành chính nhằm góp phần hoàn thiện cơ sở dữ liệu quốc gia về thủ tục hành chính.

TTCH

VB

Tìm kiếm thủ tục hành chính trợ giúp ?

Nhập tên thủ tục hành chính hoặc từ khóa liên quan đến thủ tục hành chính vào cửa sổ tìm kiếm dưới đây

Tìm kiếm

[tìm kiếm nâng cao >](#)

h

vào cửa sổ tìm kiếm

Tìm kiếm

[tìm kiếm nâng cao >](#)

Tham gia

Nhận xét

Liên hệ

TÌM KIẾM NHIỀU NHẤT

- [→ Đăng ký mã số thuế cá nhân](#)
- [→ ô tô](#)
- [→ cấp sổ đỏ](#)
- [→ mã số thuế](#)
- [→ sổ đỏ](#)

NHẬN XÉT NHIỀU NHẤT

- [→ Cấp giấy phép kinh doanh lữ h...](#)
- [→ Cấp Giấy phép thành lập Văn p...](#)
- [→ Cấp lại Giấy phép thành lập V...](#)
- [→ Cấp lại Giấy phép thành lập V...](#)
- [→ Gia hạn Giấy phép thành lập V...](#)

BỔ SUNG GẮN ĐÃY NHẤT

- [→ Đăng ký thỏa ước lao động tập...](#)
- [→ Đăng ký điều chỉnh dự án đầu...](#)
- [→ Chấm dứt hoạt động Chi nhánh/...](#)
- [→ Giải thể doanh nghiệp đối với...](#)
- [→ Đề nghị cấp lại Giấy chứng nh...](#)

LIÊN KẾT

- [→ thutuchanhchinh.vn](#)

THAM GIA
nhận xét hồ sơ >
liên hệ >

Các liên kết
TTCH mới bổ sung >

Cơ sở dữ liệu quốc gia về TTCH Sơ đồ trang tin >

<http://csdl.thutuchanhchinh.vn>

Bad experiences (1)

- MySQL 5.0:
 - usual SBR replication #FAIL
 - Non consistent replication. Rollback?
- MySQL 5.1:
 - OOPS on master server (and reboot): slave server cannot continue replication from impossible position...
 - Need to monitor replication trough ZenOSS
- Debian Lenny kernels:
 - OOPS can happen
 - Unstable kernels moduls: GFS, GFS2

Bad experiences (2)

- Debian RHCS:
 - Old packages2
 - No active Debian maintainer
 - won't fix bugs
 - Hard to maintain (umount problems, etc.) uzrokuje razne popratne pojave: nemogućnost umounta GFS, čistog reboota stroja...
- GFS i GFS2:
 - Slow, unstable, bad scalability
 - GFS2: OOPS on all nodes after 5 hours of testing
 - GFS: decreasing performance over time

Advices

- Use OCFS2 – negligible problems comparing to GFS
- Optimize caching on all levels
- Minimize writes to database and disk
- Use ESI to utilize Varnish for authenticated users
- Use Solr for reading and searching
- Make AJAX features

Questions?



Ivo Lukač

www.netgen.hr feat. www.infomar.hr