

Platformă web pentru clipuri video

<http://highscalability.com/youtube-architecture>

http://ro.wikipedia.org/wiki/Analiza_SWOT

<http://ro.wikipedia.org/wiki/YouTube>

<http://ro.wikipedia.org/wiki/Vimeo>

<http://builtwith.com/youtube.com>

Detalii privind soluția propusă

Analiza SWOT

	Benefice antingerii obiectivelor	Pun în pericol antingerea obiectivelor
Sursă internă (organizația)	Puncte tari	Puncte slabe
Sursă externă (mediul extern)	Oportunități	Amenințări

Analiza SWOT

Analiza SWOT este o metodă folosită în Managementul Proiectelor Software pentru a ajuta la proiectarea unei viziuni de ansamblu asupra produsului. Funcționează ca o radiografie a proiectului din punct de vedere managerial și evoluează în același timp factorii de influență interni și externi ai organizației, precum și poziția acesteia pe piață sau în raport cu ceilalți competitori cu scopul de a pune în lumină punctele tare și cele slabe ale companiei, în relație cu oportunitățile și amenințările existente la un moment dat pe piață.

Ținând cont de descrierea analizei menționate mai sus, urmează să le exemplificăm pe baza Platformei web pentru clipuri video.

Structura va reprezenta și inițialele din cadrul Analizei SWOT:

S – Puncte tari

- La ce suntem cei mai buni?

Echipa fiind compusă din absolvenți de Informatică, partea pentru care am studiat și ne-am aprofundat cunoștințele sunt cele ce țin de software. Toate eforturile se vor concentra pe partea de dezvoltarea produsului, implementarea feature-urilor și rezolvarea eventualelor bug-uri.

- Ce aptitudini specifice are forța de muncă de care dispunem?

Aptitudinile pe care ne bazăm sunt cele de implementat ideea aplicației, de rezolvarea problemelor care pot apărea în acest demers, realizarea documentațiilor și a paginilor de wiki despre produsele noastre, cât și promovarea platformei prin membrii ce dețin acest rol în cadrul echipei.

- De ce avantaje dispunem pentru a atrage personal de calitate?

Dispunem de un mediu de lucru tânăr și dornic de afirmare și de lansare pe piață a unui produs care să revoluționeze nișa de public căreia ne adresăm.

- Ce experiență deține echipa de proiect din proiecte similare?

Experiența membrilor se rezumă la proiectele individuale și cele realizate în echipă din anii facultății, cât și a lucrului la aplicații individuale, contribuții la proiecte open-source, dar și experiența de la locurile de muncă precedente și a cunoștințelor acumulate în acel mediu.

- De ce resurse financiare dispunem?

Dispunem de un capital ce ne asigură buna desfășurare a activităților într-un cadru comun, închiriat în cadrul TechHub Bucharest, pe o perioadă de 6 luni.

- Ce tehnologie folosim?

Frontend - HTML5, CSS3, Javascript, jQuery, Coffeescript, AngularJS, Twitter Bootstrap.

Backend - Scala, Java, Python, C++, C.

Platforme folosite: Apache - Hadoop, Lucene, Spark -, Linux - Ubuntu, MySQL, lighttpd pentru video related, Python, psyco - a dynamic python to C compiler

Web Servers:

- Amazon AWS: mașinile din cloud vor avea instalate Ubuntu
- NetScaler folosit ca Load Balancer, și pentru caching static content
- Request-urile sunt route-ate pentru a fi procesate de către un server Python

- Server-ele comunică cu diverse sisteme de baze de date pentru a primi toate datele, urmând apoi trimiterea lor către html
- Pentru activități CPU-intensive precum criptarea datelor, se folosesc extensii scrise în limbajul C
- Conține porțiuni HTML scump de generat - în materie de timp - cu latency ridicat, pregenerate în cache
- Memcache și Redis pentru a salva conținut în cache
- Folosesc server-ele lighttpd pentru procesare video
- Conținutul cel mai popular este mutat într-un CDN - content delivery network. CDN-ul multiplică conținutul în mai multe locuri pe server-e.

Databases:

- MySQL pentru stocare metadata, precum user-i, tag-uri, descriptions
- Cassandra
- Git pentru version control
- GitHub pentru hosting codebase-ului
- GitHub Issue Tracker pentru evidența issue-urilor
- Search feature is implemented using Elasticsearch platform.

W – Puncte slabe

- La ce suntem cei mai slabi?

La partea de configurat server-e ne așteptăm să nu fim așa fluenți, niciun membru neavând de-a face prea mult cu partea administrativă a unui sistem până acum.

- Ce fel de instruire le lipsește angajaților noștri?

Le lipsește instruirea referitoare la configurarea mediului de versionare al surselor la nivel de proiect, script-urile referitoare la ușurința dezvoltării și configurarea server-elor.

Toate aceste lipsuri se vor rezolva în timp, pe parcursul întâlnirii problemelor.

- Care e poziția noastră financiară?

Poziția noastră financiară este suficientă pentru a acoperi majoritatea necesităților pentru următoarele 6 luni.

- Este disponibilă o estimare solidă a costurilor?

Până în momentul de față nu este disponibilă o estimare solidă a costurilor, dar cu siguranță beneficiază de toată atenția și va fi tratată în perioada următoare.

- A alocat compania un buget suficient pentru a acoperi anumite cheltuieli neprevăzute?

Da, este alocat un buget destinat pentru acoperirea cheltuielilor neprevăzute ce pot apărea.

- E nevoie ca anumite părți din proiect să fie externalizate?

Da, unele părți din proiect vor fi externalizate, deoarece luăm în considerare și achiziționarea unor pachete pentru promovarea în mediile online a produsului nostru, precum

și planul de închiriere a unor mașini în cloud-ul celor de la Amazon Web Services - AWS, în această situație nu deținem accesul fizic la mașinile puse la dispoziție de AWS, ci doar legătura virtuală cu acestea prin intermediul terminalelor, ținând de noi configurarea lor cu tot ce este necesar pentru deploy-ul aplicației.

- Ce nu facem bine?

Nu avem implementate încă toate feature-urile.

- Ce ar trebui evitat pentru a nu repeta greșelile din trecut?

O mai bună documentație în privința unui workflow eficient și a metodelor Agile de rand înalt, demonstrate practic.

- Care sunt dezavantajele proiectului?

Că nu este singurul produs de acest gen de pe piață și trebuie să apară cu ceva nou pentru a atrage utilizatorii.

O – Oportunități

- Ce schimbări ale mediului extern putem exploata?

Putem exploata tendința continuă a publicului de a încerca aplicații noi în privința share-ing-ului de conținut video și a feature-urilor diverse, mai ales care să facă distincția față de ceea ce se oferă acum pe piață.

- La ce tehnologie nouă am putea avea acces?

Am putea să luăm în considerare și alternative la metodele alese și folosite și în urmă cu mai mulți ani, cum ar fi:

- elementele de programare funcțională
- object oriented databases
- document oriented databases
- graph databases
- Care sunt direcțiile strategice majore ale afacerii:

- Consolidare / Diversificare ?

- Specializare / Generalizare ?

- Unde se poate identifica, sau cum se poate crea, un avantaj concurențial?

Un avantaj concurențial se poate crea prin introducerea de feature-e care nu se găsesc pe piață la momentul lansării, tocmai pentru a reprezenta o noutate pentru potențialii utilizatori ai platformei.

T – Amenințări

- Există deja pe piață o competiție bine încheată?

Da, pe piață există deja o competiție bine încheiată. YouTube, Vimeo, Dailymotion, Vine sunt doar câteva dintre competitorii actuali care au parte de numeroși utilizatori.

- Ce ar putea face concurența în detrimentul nostru?

Concurența dispune în primul rând de fonduri mult mai mari, precum și forța de muncă mai specializată. Unele dintre celelalte platforme sunt părți din companii mai mari, și dețin și avantajul de a putea “transfera” cu ușurință specialiști de la un proiect mare la altul, fără o pierdere însemnată de timp și, implicit, cu o eficacitate ridicată pe mai multe planuri.

- Ce schimbări ale normelor sociale, ale profilurilor populației și a stilurilor de viață ar putea fi o amenințare pentru noi:

- Schimbările tehnologice?

Da. Pot apărea tehnologii semnificativ mai eficiente pentru produsele asemănătoare cu al platformei prezentate, care ar trebui luate în considerare pentru a fi portate. De asemenea, unele tehnologii folosite pot să nu mai aibă parte de suport, ceea ce le-ar face mai puțin indicate pentru continuarea folosirii lor. Trebuie analizată foarte bine oferta tehnologiilor, să fie credibile și să existe o oarecare siguranță, să pară promițătoare, dacă nu chiar cu rezultate anterioare din industrie pentru buna lor funcționare și utilizare la scară largă.

- Schimbări ale cererii pentru anumite tipuri de servicii, probabil legate de influența deosebită a Internetului?

Nu ne așteptăm la astfel de schimbări.

- E dificilă înlocuirea personalului cu experiență?

Da, pe parcurs ce proiectul crește în anvergură, este nevoie de un timp ceva mai lung pentru obișnuirea noilor angajați cu dezvoltarea platformei.

- A fost noua tehnologie testată corespunzător?

Produsul va avea nevoie de testări masive, de asemenea existând și sisteme speciale prin care trebuie să treacă fiecare modificare adusă codebase-ului.

Cine mai face același lucru?

În cadrul acestui articol prezentăm niște platforme similare care produc același conținut cu platforma prezentată în cadrul acestui document.

YouTube

Este un Site Web unde utilizatorii pot încărca și vizualiza videoclipuri. A fost creat în februarie 2005 de trei foști angajați ai PayPal. Platforma a fost achiziționată în noiembrie 2006 de Google. Utilizatorii neînregistrați pot vizualiza toate videoclipurile de pe site, iar cei înregistrați pot și încărca videoclipuri, pot introduce comentarii și pot da note filmulețelor. Un feature ce îl distinge într-o oarecare măsură față de competitori este posibilitatea de Live Video.

Vimeo

Este un site web ce oferă servicii de video hosting, unde utilizatorii pot încărca, distribui și viziona clipuri video.

Dailymotion

Este un site web unde utilizatorii pot încărca și vizualiza videoclipuri. A fost înființat în 2005 de către doi francezi. Se consideră că este principalul concurent al YouTube.

Comparația cu alte soluții asemănătoare

Diferențele între platforma prezentată și soluțiile asemănătoare prezentate mai sus nu sunt așa de multe. YouTube, fiind un serviciu cu trafic foarte ridicat, este foarte bine scalat pentru a asigura un răspuns rapid și fără întârzieri semnificative pentru utilizatori, de multe ori diferența practic făcându-se doar la nivelul vitezei cu care circula datele de la provider-ul de internet. În termeni de servicii oferite, părțile comune evidente pentru utilizatori sunt vizionarea și upload-ul de clipuri video, comentarea lor, posibilitatea de share-ing, rating-ul. După cum s-a menționat și într-un punct din analiza SWOT, un punct care-l face aparte pentru YouTube este posibilitatea de a transmite Live Video, ceea ce-l face varianta preferată în situațiile în care se urmărește transmiterea de evenimente sau emisiuni live stream.

Tehnologii folosite

Frontend - HTML5, CSS3, Javascript, jQuery, Coffeescript, AngularJS, Twitter Bootstrap.

Backend - Scala, Java, Python, C++, C.

Platforme folosite: Apache - Hadoop, Lucene, Spark -, Linux - Ubuntu, MySQL, lighttpd pentru video related, Python, psyco - a dynamic python to C compiler

Web Servers:

- Amazon AWS: mașinile din cloud vor avea instalate Ubuntu
- NetScaler folosit ca Load Balancer, și pentru caching static content
- Request-urile sunt route-ate pentru a fi procesate de către un server Python
- Server-ele comunică cu diverse sisteme de baze de date pentru a primi toate datele, urmând apoi trimiterea lor către html
- Pentru activități CPU-intensive precum criptarea datelor, se folosesc extensii scrise în limbajul C
- Conține porțiuni HTML scump de generat - în materie de timp - cu latency ridicat, pregenerate în cache
- Memcache și Redis pentru a salva content în cache
- Folosesc server-ele lighttpd pentru procesare video
- Content-ul cel mai popular este mutat într-un CDN - content delivery network. CDN-uri multiplică content-ul în mai multe locuri pe server-e.

Databases:

- MySQL pentru stocare metadata, precum user-i, tag-uri, descriptions
- Cassandra
- Git pentru version control
- GitHub pentru hosting codebase-ului
- GitHub Issue Tracker pentru evidența issue-urilor

Search feature este implementat folosind Elasticsearch platform.

Pentru partea de advertising a website-ului se folosește DoubleClick.net și Google Adsense.

Pentru Analytics and Tracking s-a optat pentru DoubleClickFloodlight, Adometry, MediaMind, Safe Count.

Unitățile de lucru pentru fiecare developer vor avea, la alegere, fie Ubuntu, fie Mac OS X. Pentru fiecare alegere există script-uri pentru a avea toate dependențele instalate și configurate.

Riscuri posibile: identificare, analiza calitativa/cantitativă, planuri de răspuns

În această secțiune urmează să prezentăm o serie de riscuri peste care este posibil să dăm pe parcursul dezvoltării produsului:

- insuficiența fondurilor pentru buna desfășurare a activității
- folosirea de tehnologii outdated sau care nu mai au parte de suport

- documentarea insuficientă în privința alegerii variantei cele mai bune pentru o anumită tehnologie
- nerespectarea deadline-urilor stabilite inițial
- pregătirea insuficientă pe anumite arii de expertiză
- promovarea inefficientă a produsului
- lipsa interesului din partea utilizatorilor, clienții nu mai revin să folosească platforma după o anumită perioadă de timp

Pentru toate acestea, avem la îndemână unele soluții ce ar putea fi puse în aplicare.

Insuficiența fondurilor ar putea fi anulată sau măcar știută din timp dacă s-ar face periodic bilanțuri și previziuni pentru perioadele ce urmează să vină și a costurilor pe care le percep acestea. Documentarea serioasă este necesară pentru a exclude acele tehnologii care nu prezintă suficientă credibilitate pentru developer-i, este crucial să ofere siguranță ca nu se va renunța la îmbunătățirea acelei tehnologii, altfel atrage după sine costuri ridicate pentru portarea altora. Nerespectarea deadline-urilor poate interfera cu multe alte probleme, și este nevoie de planificări și actualizarea ETA-urilor - Estimated Time of Arrival - cât mai des posibil. Din moment ce se dă acordul din partea celor în cauză asupra folosirii unei anumite tehnologii, pentru persoanele ce urmează să o implementeze și să o folosească trebuie să urmeze o perioadă de documentare intensă și de asumare a cunoștințelor pentru a folosi tot ce oferă acestea, este benefic și ca persoanele care au mai avut de-a face cu acele produse și posedă o cunoștință în cauză mai ridicată față de ceilalți să le prezinte și colegilor, prin așa numitele training-uri interne pentru diferite tehnologii, platforme, produse, best-practices, etc. Astfel se asigură un continuu schimb de experiență ce își lasă amprenta adânc asupra dezvoltării. Persoanele responsabile de promovarea produsului în diferite arii de interes - media, site-uri de socializare, etc - trebuie să aibă cunoștințe pentru a eficientiza acest aspect, crucial pentru o deschidere cât mai largă pentru publicul țintă, evident, activități ce implică și direcționarea de fonduri pentru sponsorizări de diverse evenimente, secțiuni publicitare, dar și pentru “cumpărarea” de timp la diverse evenimente în cadrul cărora să se prezinte produsul, echipa, scopul, etc. cu accent pe interesele persoanelor din public cu privire la produsul prezentat.

Part II

4. Identify project products and activities

4.1 Identify and describe project products

Platforma prezentată este reprezentată de un site web în care utilizatorii pot vedea, upload-a conținut video, precum și să comenteze la clipuri, să le share-uiască sau să le dea note. Dacă nu se optează pentru crearea unui cont, i se permite doar să vizualizeze videoclipurile. Dacă a optat să-și creeze cont, din momentul în care s-a log-at, are acces și la toate celelalte funcționalități.

!!!!!!!!!!!!!!!!!!!!

PRINTSCREEN cu youtube homepage, și log-at și nelog-at

!!!!!!!!!!!!!!!!!!!!

În partea de sus se poate observa search-box-ul, apoi în dreapta username-ul utilizatorului. În partea centrală se observă clipurile cele mai populare. Dacă utilizatorul este log-at, în partea centrală îi vor apărea clipurile de la user-ii la care a dat în trecut “Subscribe”, iar în partea stângă apar informații cu privire la istoricul vizualizărilor, listele create și personalizate de-a lungul timpului, precum și lista cu paginile la care a dat “Subscribe”. Scroll-ând pe pagina principală se pot observa clipurile care îndeplinesc cerințele preferate, în particular, clipurile postate de paginile la care a dat “Subscribe”.

!!!!!!!!!!!!!!!!!!!!

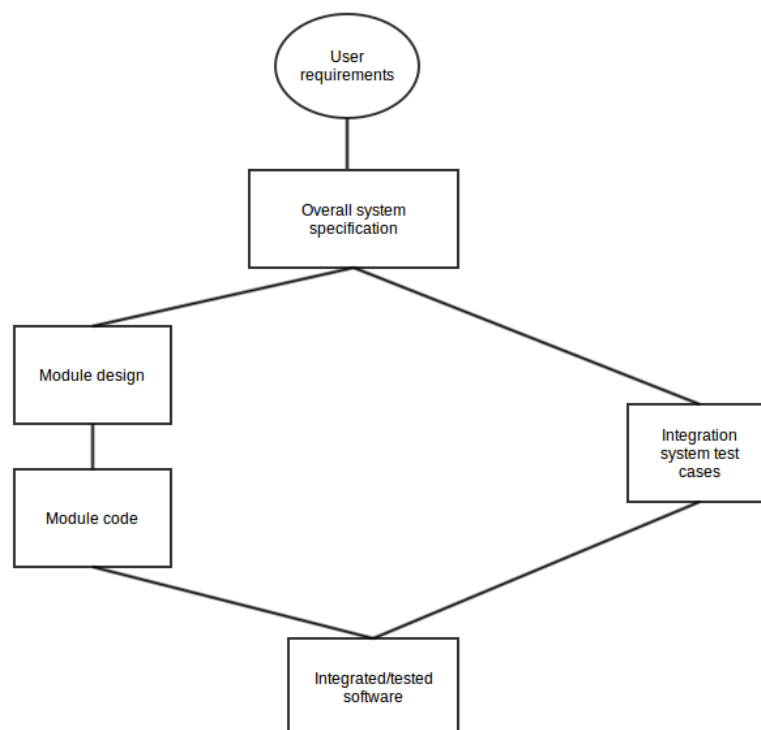
- PRINTSCREEN cu youtube video playing, description, comments section
- de menționat quality criteria la fiecare, ca e ok, etc

!!!!!!!!!!!!!!!!!!!!

În partea principală se observă clipul video, iar în partea dreapta lista de sugestii cu clipuri asemănătoare, pe anumite criterii observate de-a lungul timpului pentru fiecare user în parte, clipurile ce se consideră a fi cele mai probabile să dorească să le vizioneze. Imediat sub clip sunt menționate numele user-ului care a postat înregistrarea video, butonul de “Subscribe”, numărul de subscribe-eri, butoanele pentru adăugare în cadrul unei liste create anterior de user, de Share, precum și alte opțiuni. De asemenea, apare și numărul de vizualizări, numărul de like-uri și dislike-uri. În continuare este secțiunea cu informații despre data la care a fost postat, o scurtă descriere, după care urmează secțiunea de comentarii la care poate contribui orice user log-at, cu excepția acelor clipuri pentru care utilizatorul a dezactivat posibilitatea de a primi comentarii.

4.2 Document generic product flows

Unele produse necesită unul sau mai multe alte produse să existe înainte ca acestea să poată fi create. De exemplu, design-ul unui program trebuie să fie creat înainte ca programul să fie implementat, să existe o măsură cu privire la rezultatul final dorit. Aceste relații pot fi exprimate prin Product Flow Diagram.



Product Flow Diagram

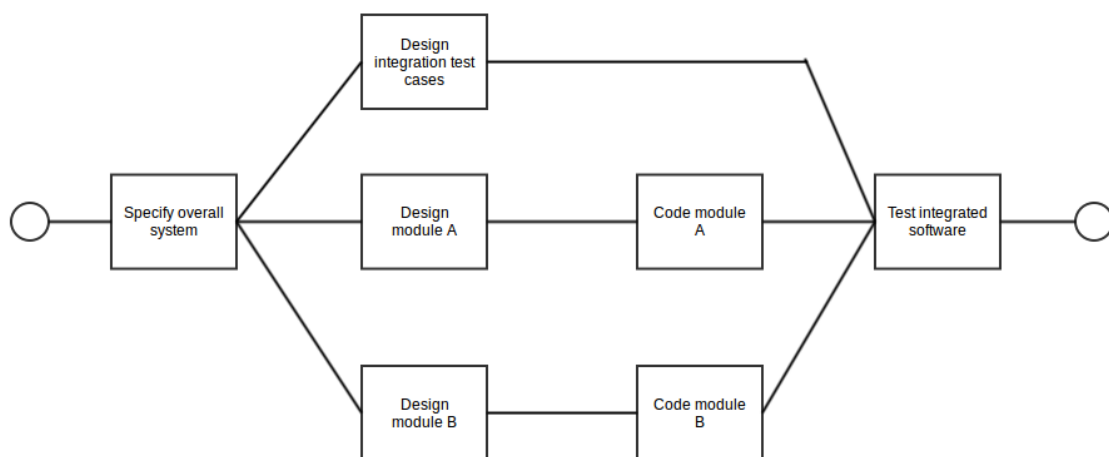
În partea de User requirements trebuie să fie clar stabilite cerințele utilizatorilor și ce urmăresc aceștia de la aplicația prezentată. Ultimul pas din schema de mai sus reprezintă efectiv implementarea ideii și testarea riguroasă a soluției.

4.3 Recognize product instances

Instanța finală a produsului va fi reprezentată de website-ul ce va cuprinde toate funcționalitățile descrise mai sus la capitolele anterioare.

4.4 Produce ideal activity network

Pentru a crea un produs din altele, trebuie să aibă loc mai multe activități. Prin identificarea acestor activități, putem crea o rețea de activități care arată task-urile ce trebuiesc duse la bun sfârșit.



Un exemplu de activity network

Considerând diagrama de mai sus la modul general, o putem particulariza pentru produsul prezentat sub forma următoare:

- Specify overall system: prezentarea generală a sistemului, la fel cum a fost surprinsă și în capitolele precedente

- Design integration test cases: pentru această etapă este necesară elaborarea de script-uri care să testeze în diferite moduri platforme - inclusiv White Box Testing și Black Box Testing, dar și modificările asupra codebase-ului pe parcursul timpului, cu diverse tool-uri specifice: FindBugs, Java Pathfinder, Coverity, Sparse, Checkstyle, etc.
- Design module A, B, ...: aceste părți necesită a fi completate pe cât posibil încă de înainte de a fi începută partea efectivă de implementare. Este nevoie de stabilirea clară a aspectului fiecărei părți ce va fi prezentată utilizatorului.
- Code module A, B, ...: aceste părți reprezintă secțiunea cea mai solicitantă, întrucât se urmărește scriere de cod eficient și cu cât mai puține bug-uri și efecte colaterale neașteptate.
- Test integrated software: este strâns legată de partea de implementare, pentru că fiecare modificare adusă sistemului trebuie să treacă de un code review, apoi o serie de teste automatizate, dar și unele efectuate manual. Doar după toți acești pași modificările sunt trimise către codebase-ul central și ajung astfel să fie vizibile modificările la nivelul aplicației live.

4.5 Modify ideal to take into account need for stages and checkpoints

Conform descrierii anterioare a planificării activităților, se ajunge la formularea unui plan care va minimiza timpul total necesar - elapsed time - pentru proiect. Se presupune că o activitate va începe imediat ce cele precedente de care depinde au fost terminate. Dar, totuși, poate fi nevoie de divizarea proiectului în etape și introducerea de checkpoint-uri. Acestea trebuie aranjate în așa fel încât să nu se ajungă la un număr ridicat de task-uri care depinde unele de altele, să se evite situația în care se așteaptă mult timp pentru ca un anumit task să fie dus la capăt.

Ținând cont de aceasta, fiecare task care este prezentat pentru a fi rezolvat trebuie să aibă niste checkpoint-uri care vor fi update-ate periodic cu câte un status, pentru a se putea aprecia ce se poate desfășura în același timp, să nu interfereze sau să fie muncă în zadar. Fiecare membru/echipă care este responsabil/responsabilă de acel task trebuie să aibă răspunderea de update-area checkpoint-urilor și a status-ului. Astfel ceilalți membri implicați

în alte task-uri care depind de activitatea desfășurată în acel moment au un plus de informații pentru a-și desfășura activitatea cât mai eficient.

5. Estimate effort for each activity

5.1 Carry out bottom-up estimates (i.e. estimate efforts for each activity, then add them to obtain the effort for the whole project - how many person-months (or days/weeks) are necessary to complete the project)

Partea de stabilire a detaliilor referitoare la design-ul aplicației, precum și organizarea backend-ului ar avea estimated time de o lună.

Pentru partea de configurare a server-elor de pe platforma Amazon AWS, estimated time ar fi de 7 zile, pentru a fi instalate toate mediile și dependențele necesare. În acest timp este inclus și timpul necesar pentru configurarea inițială a Hadoop-ului pe cluster-e, MySQL, Memcache, Redis, Cassandra, Git.

În maxim 4 ore ar trebui să fie instalată orice varianta de OS aleasă, Ubuntu sau Mac OS X, alături de rularea script-urilor de configurare a mediilor de lucru necesare - dependențe, etc.

Implementarea frontend are stabilit inițial un ET - estimated time - 3 săptămâni.

Pentru backend cu tot ce ține - partea de upload clipuri video, adăugare feature de comentarii, afișarea numărului de vizualizări, clasificare după nota generală dată de user-i, afișarea de clipuri asemănătoare în partea dreaptă, History list pentru fiecare utilizator, ar avea un ET de 4 luni.

În total, ET ar fi în jur de 6 luni.

Trebuie menționat că toate aceste elaborări se referă doar la prima variantă a aplicației, așa numita versiune beta. După aceea urmând mentenanța, adăugarea de feature-e noi, rezolvarea bug-urilor pentru buna funcționare a platformei, care se vor întinde în timp.