

## Webbanalys

Hemsidan  
Webbanalys artiklar

### Utvärdering av logganalysverktygen Awstats och Webalizer

*Awstats och Webalizer är två gratisverktyg för logganalys som är relativt populära bland webbmasters och webbredaktörer i Sverige. Men kan de beräkna trafiken rätt överhuvudtaget vad gäller sidvisningsar och unika besökare?*

#### Förutsättningar

Som indata användes Apache loggfiler från www.Webbanalys.se (Combined format inkl. cookie data), där följande data användes av valideringsverktyget:

- sidvisningar (samtliga GET eller POST anrop till sidor av typen .html)
- endast lyckade 200 eller 304 (cache) träffar
- cookie data (IP nummer + slumpnummer sammansatt)

I loggfilen fanns det 3306 korrekta anrop till olika webbsidor enligt ovan vilket kom från 1027 unika besökare (baserat på cookie data), vilket rapporterades av ett valideringsverktyg samt verifierades visuellt i en fileditor. Dessutom fanns både IP nummer information och cookie data då webbservern sänder ut 1:a parts cookies till varje besökare av webbplatsen. Allt detta gav förutsättningarna för att det skulle vara möjligt att korrekt beräkna antal sidvisningar och unika besökare.

#### Filtrering av robotar

Eftersom valideringsverktyget per automatik exkluderade sökrobotar och liknande kom antalet sidvisningar som den presenterade att bli betydligt lägre än vad de två testade verktygen visade. Awstats har betydligt bättre möjligheter än Webalizer att kunna filtrera bort irrelevanta anrop, och är i detta hänseende det bättre valet. Dock kräver Awstats att man kontinuerligt håller ett öga på tabellen med signaturer för webbläsare för att kunna hålla listan på robotsignaturer uppdaterad. Filtrering av robotar sker inte per automatik i Webalizer, och den kräver mer arbete för att åstadkomma detta. I Webalizer sker detta med hjälp av *hide* optionen, vilket kräver ett fortlöpande underhåll till skillnad mot Awstats som vid uppdatering av programmet kommer med senaste robotlistan.

#### Sidvisningar

Genom att endast inkludera anrop på sidor vilka levererats av webbservern (status 200) eller påträffats i cachen (status 304) erhöles värden på faktiska sidvisningar. Både Awstats och Webalizer gör sina beräkningar av sidvisningar på detta sätt.

#### Unika besökare

Det är vid beräkning av unika besökare som gratisverktygen visar upp en annan svaghet, och dessa två verktyg har valt olika sätt att försöka få fram en relevant siffra. Awstats beräknar unikheten på anrop efter webbsidor och ignorerar helt träffar (hits) på andra objekt som webbservern sänder ut. Genom att göra beräkningen på så sätt minskas felräkningen, men man anger helt riktigt i sin FAQ att detta inte ger en korrekt siffra eftersom besökare bakom en proxy kan få sidor anropade via fler proxy servrar som kör på olika IP nummer.

Webalizer anger i sin README fil att anrop kommer ifrån unika *sites* vilket likställs med ett IP nummer eller namn (vid namnuppslagning av IP numret), vilket inte bara är en felaktig benämning utan även rätt vilseledande. Även här anges att den siffra som erhålls inte ger en korrekt beräkning av antalet unika besökare.

Beräkning av unika besökare när en loggfil är datakällan kräver att man kan särskilja besökare bakom brandväggar och proxyservrar, och ett sätt är att använda sig av en cookie som den egna webbservern sänder ut. Men för Awstats och Webalizer spelar det ingen roll, de saknar helt funktionaliteten att kunna ta vara på cookie information. Att beräkna unika besökare på enbart IP nummer ger direkt felaktiga siffror och bör undvikas!

#### Summering

Båda verktygen ger rena glädjesiffrorna på sidvisningar om inte omfattande konfiguration av dem sker innan analys av loggdata. Dessutom kommer de trots detta att visa besöksiffror vilka inte var korrekta, på loggdata filtrerat av valideringsverktyget lyckades de båda fortfarande visa felaktig information:

Valideringsverktyget - 1027 unika besökare  
Awstats - 925 unika besökare  
Webalizer - 1099 unika besökare

Däremot blev siffrorna korrekta vad gäller sidvisningar när det filtrerade underlaget användes, dock tappade Awstats bort 1 sidvisning vilket nog är kopplat till beräkningskoden där sista sidan inte räknas.

#### För korrekta siffror krävs ett kompetent verktyg

Men sidvisningar som syns i loggfilen är långt ifrån allt som faktiskt har visats, speciellt med tanke på att innehåll kan levereras via mellanlager i proxy servrar och liknande system. Under samma tidsperiod mättes webbtrafiken med HBX Analytics som rapporterade hela 14% mer sidvisningar! HBX använder sig av javascript taggar på webbsidor vilka anropar en insamlingsserver varje gång sidan laddas i en webbläsare. I detta fall visade det sig att på en så liten datamängd som nu användes så var felavvikelsen anmärkningsvärd, om man är seriös med de mätvärden man presenterar måste filtrering och validering ske.

Vill man dessutom ha mer verkliga sidvisningar bör man överväga att använda sig av en klientinsamlad lösning (HBX Analytics) som med JavaScript samlar in besöksinformation. För rättvisande siffror på unika besökare måste cookies användas. Sist men inte minst är det viktigt att ta bort bruset i det data som mätningen utförs på för att säkerställa att relevant information presenteras.

Har du inte möjlighet/budget att skaffa dig ett riktigt webbanalysystem så rekommenderas Awstats trots att den visar lite lägre antal unika besökare eftersom den är betydligt mer konfigurerbar och har en mer komplett och lättläst rapport samt bättre konfigurations möjligheter. Men det kräver då att det data som används antingen är förbehandlad alternativt ett omfattande underhåll av Awstats konfigurationen fortlöpande.

I testet användes:

**WebSideStory® HBX™ Analytics version 2.6**

**Webalizer version 2.01**

**Awstats (Advanced Web Statistics) version 6.4 [build 1.814]**

««



Analysera dina sidvisningar med HBX Analytics!