Path: gmdzi!unido!math.fuberlin.de!ox.com!yale.edu!qt.cs.utexas.edu!cs.utexas.edu!uunet!decwrl!pa.dec.com!new s From: re...@decwrl.dec.com (Brian Reid) Newsgroups: news.groups,news.lists,news.admin Subject: USENET READERSHIP SUMMARY REPORT FOR AUG 91 Message-ID: <1991Sep1.224649.27186@PA.dec.com> Date: 1 Sep 91 22:46:49 GMT Sender: ne...@PA.dec.com (News) Organization: DEC Western Research Laboratory Lines: 392 Approved: re...@decwrl.dec.com

Xref: gmdzi news.groups:143772 news.lists:1437 news.admin:63459

USENET READERSHIP SUMMARY REPORT for Aug 91 This is the first article in a monthly posting series from the Network Measurement Project at the DEC Western Research Laboratory in Palo Alto, California.

This survey is based on a sample of data taken from various USENET sites. At the end of this message there is a short explanation of the measurement techniques and the meaning of the various statistics. The messages that follow this one show survey data sorted by various criteria.

The newsgroup volume and article counts that I post are often significantly different from the ones posted by Rick Adams, because he includes the size of a crossposted article in every group to which it is posted, whereas I charge that size only to the first-named group.

The complete set of readership data (of which this is a summary) is posted in news.lists. The software that will let your site participate in the survey is in comp.sources.d and news.admin

24.73%

22.142

9105

692

28 days

75 days

Brian Reid

OVERALL SUMMARY:		
	This	Estimated
	Sample	for entire net
Sites:	693	34000
Fraction reporting:	2.04%	100%
Users with accounts:	139017	6820000
Netreaders:	34385	1686000
Average readers per site:		50

Average readers per site: Percent of users who are netreaders: Average traffic per day (megabytes): Average traffic per day (messages): Traffic measurement interval: last Readership measurement interval: last Sites used to measure propagation:

Valid data received from these sites:

6sceng(2) 6sigma(11) abacab(6) ablnc(253) academ01(73) acheron(27) actew.oz.au(529) ads.com(196) adspp(27) agate.berkeley.edu(30) ahkcus.org(51) aidx.no(2) alba21(5) alchemy(94) alembic(13) alfred(4) alias(233) alpha.ru.ac.za(23) alsys.de(43) alv(147) alzabo(35) amcvlsi(41) amdahl(333) amun-re(112) anacom1(30) anagld(66) andrew.cmu.edu(3880) anomaly(60) apollo(2896) apple(1668) arak.winix.tih.no(0) aramis.rutgers.edu(406) ariadne.csi.forth.gr(195) arran.tcom.stc.co.uk(76) arthur.cs.purdue.edu(121) aspentec.com(207) ast(37) atari(40) atlas(30) awful(8) bclwn(19) bcsfse(187) bcstec(131) bdcsys(4) bdt(146) becker(19) bermls(6) bfmny0(1) bgsuvax(623) bitcave(18) bjs(77) blinky(13) bonnie.ics.uci.edu(985) bosco.berkeley.edu(125) brad(1) brolga(145) brspyr1(59) brubeck.uucp(16) btni(38) bu-ast(61) bu-it(249) bu-pub(772) bu.edu(112) bucsf(649) buengc(274) buengc.bu.edu(276) buengf.bu.edu(731) bugs.mty.itesm.mx(20) bute.tcom.stc.co.uk(86) bwdls58(1) cabezon(86) cadreor.uucp(12) cae780.csi.com(77) caip.rutgers.edu(421) cals01(2) camb.com(32) cambridge.apple.com(37) camco(46) carroll1(662) cartan.berkeley.edu(145) cc-server4.massey.ac.nz(1249) cc.usu.edu(1261) cck.cov.ac.uk(1142) cdin-1(23) cel(402) cerritos.edu(321) cfctech.cfc.com(43) cg-atla(342) cgl.ucsf.edu(2) chance(3) cheers(4) chyde.uwasa.fi(110) ciit85.ciit.nrc.ca(92) cis(50) cix(906) clinet(175) clmqt(164) clsib21(0) cmgm.stanford.edu(327) cmkrnl.uucp(7) cobber(41) cognos(192) com50(21) comm.wang.com(45) concurrent.co.uk(117) convex.mitek.com(172) cos(175) crash(123) crayjr(8) crcc(0) crowfix(1) cs(320) cs-col.columbia.ncr.com(57) cs.columbia.edu(650) cs.utexas.edu(2801) csadfa(88) csam(58) csustan.csustan.edu(25) cuff(50) cutler.uucp(26) cuuxb(342) cvman(21) cwsmk1(28) cybvax0(33) darth(22) datacube(119) david(4) davis(163) dayton.saic.com(58) db1(46) dbase(187) dbulm1.uucp(99) dciem(234) dcl-vitus(105) ddmp.dla.mil(12) ddsw1(131) deimos.cis.ksu.edu(191) dell(45) demon(25) desc.dla.mil(10) df.lth.se(24) digi(557) dimacs.rutgers.edu(134) disc.dla.mil(303) disuns2(289) dixie.com(27) dlvax2.datlog.co.uk(132) dmc.com(2) dorm.rutgers.edu(108) dplace(0) dptspd(29) dranet.dra.com(258) drd(5) dri500(38) druid(36) drycas.club.cc.cmu.edu(75) dsac.dla.mil(270) dsacg2.dsac.dla.mil(160) dsacg3.dsac.dla.mil(78) dsinc(8) dsndata(20) duke(316) dumbcat(10) dwerger(15) dycom(13) eab(2) earvax(48) ecl.psu.edu(646) ector.cs.purdue.edu(85) edc(30) edi386(4) eff(65) egsner(83) elan.elan.com(31) elbereth.rutgers.edu(1040) elbrun.winix.tih.no(1) eleazar.dartmouth.edu(397) electro(22) elements.rpal.com(36) elrond(10) elroy(55) encore.com(260) ensys(7) eos(86) eplunix(37) ericom.ericsson.se(29) esatst(11) esleng(11) euler.rutgers.edu(290) europe.drms.dla.mil(96) fang(35) fantasci(32) fasfax(66) fdls(55) fermat.rutgers.edu(113) fibers.uucp(3) flab(741) flatlin(7) fnnews(1) fnx(27) forsyth.uucp(125) foster(32) fpssun(388) frcs(11) freedm(5) frmug(59) frog(35) froya.winix.tih.no(1) frye.rain.com(19) fryeten.uucp(21) fts1(12) fxgrp(103) gaboon(32) gacvx2.gac.edu(42) galaxia(14) galaxy.southpower.co.nz(281) gang-of-four(75) garnet.berkeley.edu(3204) gauss.math.purdue.edu(257) gauss.rutgers.edu(282) gcc.edu(63) geac(69) geovision(82) gerbil.ladc.bull.com(38) giga(104) gistdev(48) glibli(10) goofy(131) gouldnl(65) grad1.cis.upenn.edu(386) grasp1(20) grebyn(34) grian(12) gtenmc(23) gtisqr(25) gypsum.berkeley.edu(175) gyre.cs.umd.edu(55) hackes(863) hagbard(154) hardees.rutgers.edu(218) harris.cis.ksu.edu(57) harrnl(13) harvisr(791) hawkmoon(13) hayes.ims.alaska.edu(54) hbiso(99) hccwmv2(10) hcsrnd(110) hemuli.tik.vtt.fi(72) henson(641) herkules.sssab.se(14) hern.stonemarche.org(12) hexagon(45) hilbert.rutgers.edu(116) hippo.ru.ac.za(185) hitra.winix.tih.no(1) hobbes.physics.uiowa.edu(8) holly.stl.stc.co.uk(51) honey(509) hoptoad(65) hp835.mitek.com(168) hpmclc(44) hpnmdla(435) hriso(77) hsi(107) hudson.stanford.edu(72) hugpar(2) humpty.edb.tih.no(20) hybrid(7) iam.unibe.ch(80)

icot.or.jp(202) ics(15) ics.uci.edu(343) iecc(8) ifi.uio.no(2516) igor(127) iitmax(889) inco(43) indigo.stl.stc.co.uk(125) inland.com(275) innovative.com(4) inttech.com(18) investor(9) iowasp.physics.uiowa.edu(128) ipc.uucp(66) iris.claremont.edu(5) isaak(56) islabs(12) ispi(7) iuvax(186) ixi(37) izar(242) jabba(43) janus.mtroyal.ab.ca(452) jato(97) jessica.stanford.edu(312) jetson(2) jfwhome.funhouse.com(5) jjmhome(41) jodokus(17) johnny5(2) joiner.com(69) jonlab(10) jrdzzz.jrd.dec.com(127) junkyard(23) ka3ovk(65) kannel(420) kericks(11) khijol(36) kielo(2790) kirk(17) kithrup(12) kjell.uucp(14) koala(18) kralizec(24) kscorp(36) kudosl.winix.tih.no(6) kurz-ai(110) lab001(25) ladcgw(15) lance(5) larry(11) latour(28) laverne(149) lehi3b15(9) leland(5404) leland.stanford.edu(5645) lethe(40) linac(18) lindy.stanford.edu(525) 1l1a(1) loglule(29) lokkur(12) longs.lance.colostate.edu(3729) loria.crin.fr(380) lpi(38) lsicom2(6) lsuc(8) lta(11) lufu(18) lysator(149) lyxys(7) m2c.m2c.org(137) m2xenix.psg.com(110) mach2.wlu.ca(52) madnix(107) maple(11) maple.stl.stc.co.uk(48) mariela(0) marque(90) math(123) math.berkeley.edu(19) math.bu.edu(181) math.ksu.edu(51) math.rutgers.edu(487) mathrt0.math.chalmers.se(201) maths.su.oz.au(251) matilda(154) matrox(28) matrx(22) mcclb0.med.nyu.edu(207) mccuts(85) mcdchg(31) mcgp1(51) mcmi(28) mdtv.uucp(11) meaddata(474) med.wcc.govt.nz(891) mediatex.uucp(13) memstvx1.memst.edu(897) metasoft(49) metpage.mps.com(116) mgweed(129) mh.nl(50) mica.berkeley.edu(81) miclon(17) mikebat(11) mikrosys.forut.no(4) mikrosys.mikrosys.no(8) mimsy.umd.edu(460) mips.mitek.com(134) mirror(34) mit-eddie(167) mnetor(141) monu6.cc.monash.edu.au(3260) monymsys(3) moon(11) moria(1) morrow.stanford.edu(13) moscom(59) mri(125) msor0(200) mtecv2(431) mtek(3) mthvax.cs.miami.edu(70) mti(47) mtxinu(55) mudos(10) muhthr.cdad.hlo.dec.com(68) mulberry(31) munsell(144) n8ino(3) nanovx(20) natinst(62) natinst.com(67) nbc1(31) nbires(18) ncoast(386) ncrcae.columbia.ncr.com(212) neon.stanford.edu(848) netagw(8) netcom(298) news.cis.ohio-state.edu(1605) newton.rutgers.edu(86) nightowl(8) nmrdc1(59) nms(26) nocusuhs(15) nosun.west.sun.com(35) nplpsg(37) nrlvx1.nrl.navy.mil(246) nshore(5) ntpal(47) nttta(35) nucleus(65) nyama(11) obdient(38) obh(11) odysseus.sssab.se(0) ogicse(1437) oliveb(16) olivee(75) olivej(122) oliven(61) olivey(110) omni(23) ondine(20) onion(12) ontek(18) orac(0) orfeo(55) orion.arc.nasa.gov(195) orkdal.winix.tih.no(2) osdvo.winix.tih.no(6) oucsace.cs.ohiou.edu(173) ovaltine.berkeley.edu(192) oxy(1174) p4tustin(15) pa621a.inland.com(174) pandora.matrox.com(150) paris.ics.uci.edu(642) park(115) park.bu.edu(115) parsely(51) pbhya(174) pbhyb(180) pbhyc(254) pbhyd(202) pbhye(187) pbhyf(216) pbhyg(283) peanut(3) pedagog.no(1) pedagog.syh.winix.tih.no(2) penguin(7) pentagon-ai(25) peregrine(135) phantom(51) philpav(7) phobos.cis.ksu.edu(396) phoenix.princeton.edu(3428) piaggio(105) pitt(147) plains(761) planck(33) pmafire(30) polari(247) practic(21) presto.ig.com(113) pride386(10) prism1(18) proeast.propress.com(201) prowest.propress.com(7) psuvax1(410) ptcburp(8) ptsfa(184) pttdis(19) pute(18) putz.uucp(9) pyramid(188) pzbaum(4) qiclab.scn.rain.com(13) qmw-cs(336) qmw-dcs(332) quack(14) quad1(35) quagga.ru.ac.za(37) questar.questar.mn.org(158) quick(5) quorum.com(8) radio.astro(64) rancho(9) rducky(7) redpoll(2) remus.rutgers.edu(1008) rena(4) rescon(6) research.canon.oz.au(42) retix.retix.com(153) rhi.hi.is(2672) ria.ccs.uwo.ca(110) rimfaxe.sssab.se(5) rivm.nl(747) robohack(34) rochester(243) roselin(540) rosevax(828) rtweyco(1) ruacad(375) rucs(25) rucs2(143) rulcvx(5) rutcor.rutgers.edu(66) rutgers.rutgers.edu(72) rwmke(3) sacred(122) sactoh0.sac.ca.us(198)

6/7/13

```
sage(16) samsung(4) samt19(29) sandoz(19) satchmo(31)
saul.cis.upenn.edu(66) schoenfinkel(5) sci.ccny.cuny.edu(244)
scicom(47) scifi(31) sclcig(24) sdd550(22) sdf(73) sdrc(77) sed00(75)
seer(37) seismo.css.gov(262) sequoia(122) sgfb(349) sh8810(265)
shade(1) shapetc(124) shark.edb.tih.no(1) shelby(8)
shelby.stanford.edu(9) shiva.com(70) sirius.ucs.adelaide.edu.au(78)
sis(30) sixnine(2) sjh.org(73) sjuphil(636) slcs.slb.com(162) slimer(5)
sluger(40) sma2(11) smith(4) sol.ctr.columbia.edu(274) sooner(5)
sparky(28) spatial.com(67) spcc.uucp(100) sq(63)
srchtec.searchtech.com(34) stade(8) stan(421)
starfish.convergent.com(53) stat.appstate.edu(1158)
storax.stl.stc.co.uk(12) stordlh.winix.tih.no(1) student(80) sugar(134)
sunburn.stanford.edu(168) suned1(347) surf(39) swatsun(87) swlabs(1)
tabbs(15) tac.fi(828) tahoe.cs.unr.edu(201) tahoe.unr.edu(159)
talos(27) tamsun.tamu.edu(58) tcomeng(56) tct(21) td2cad(147)
techbook(164) tellab5(101) telops.ma02.bull.com(3) teti(33)
theporch(25) thinc(3) think(729) tkou02.enet.dec.com(11) tol-ed.com(39)
torag(36) tove.cs.umd.edu(245) tricord(55) trident.arc.nasa.gov(16)
trux(200) ttidca(527) ttsi(41) tucos.utu.fi(77) tukki(1023)
tut.cis.ohio-state.edu(1562) twg(13) twinsun(35) txsil(24) uafhp(1810)
ucbarpa.berkeley.edu(450) ucselx(316) ug.cs.dal.ca(301)
uhura.cc.rochester.edu(349) uk96sup(47) ukma(601) ukpoit(12)
ulla.und.slu.se(2) ulmus.sdi.slu.se(6) umd5.umd.edu(132) unicom(416)
unify(66) uniwa(135) unocal(189) ursa(147) uu.psi.com(204) uunet(91)
ux0(46) valinor(7) valinor.mythical.com(8) vax5.cit.cornell.edu(1450)
vela.acs.oakland.edu(1850) verifone.com(1003) vicstoy(16) victrola(4)
victrola.halcyon.wa.com(4) videovax.tv.tek.com(2) vidiot(1)
violet.berkeley.edu(1650) viper(1) viusys(179) voodoo(22) vsi1(76)
vtsai2.sai.vtt.fi(43) vtserf(56) wang(111) warwick(3819)
water.berkeley.edu(171) wattres(10) wb3ffv(68) weaver.berkeley.edu(301)
weitek(141) well(5754) wet(136) wgate(24) whscdp.uucp(36)
widow.berkeley.edu(272) wilbur(113) wiley(191) winixus(7) winsoft(6)
wizvax(19) wjvax(49) wolf.berkeley.edu(248) wolves(33) woodowl(33)
wrdis01(2234) wrglex(34) wrglex.uucp(109) wright(474) wshb(43) wsl(19)
wsrcc.com(2) wybbs(150) xenitec(33) xenon(432) xmws.uucp(8) xray(53)
xylogics.com(52) yetti(38) yorkohm(12) yuban.berkeley.edu(118)
zap.mlb.semi.harris.com(151) zaphod(129) zardoz.club.cc.cmu.edu(12)
zeus(29) zinn(27) zorch(36)
```

EXPLANATION OF THE MEASUREMENTS AND STATISTICS

Survey data is taken by having one person at each site run a program called "arbitron", which looks at the news or notes files and determines the newsgroups that the user has read within a recent interval. To "read" a newsgroup means to have been presented with the opportunity to look at at least one message in it. Going through a newsgroup with the "n" key counts as reading it. For a news site, "user X reads group Y" means that user X's .newsrc file has marked at least one unexpired message in Y. If there is no traffic in a newsgroup for the measurement period, then the survey will show that nobody reads the group. For a notes site, "user X reads group Y" means that user 14 days. The "14 days" interval for notesfiles corresponds to "unexpired" for news.

The "arbitron" program is periodically posted to comp.sources.d, or is available from me (decwrl!reid). The notesfiles version of the program should be available through standard notesfiles software distribution channels as well.

SITES SURVEYED IN THIS SAMPLE

"This Sample" means the set of sites that have sent in an arbitron report within the past "Readership measurement interval" days. In every case the most recent report from each site is used. At the moment, some of the readership reports are several months old. In future postings those reports will have expired and will not be included.

The number in parentheses after the site name is the number of users that the site reported. A value of (0) usually means that the software has been configured to use the wrong technique for counting users at that site; a report showing 0 users but 6 readers of rec.humor.funny is statistically meaningful.

One might argue that the sample is self-selected, and thereby be biased. It does in fact have a certain self-selection factor in it, because we only get data from sites at which someone participates in the survey. However, we do not require the participation of every user at a site, only one user. The survey program returns data for every user on the system on which it was run. Since there are an average of 30 people per site reading news, there is a certain amount of randomness introduced that way. Of course, the sample is biased in favor of large sites (they are more likely to have a user willing to run the survey program) and software-development-oriented sites (more likely to have a user *able* to run the survey program).

NETWORK SIZE

I determine the network size by looking at the set of sites that are mentioned in the Path lines of news articles arriving at decwrl. This number is consistently higher than the number of sites that posted a message (as measured and posted from uunet) because it includes passive sites that are on the paths between posting sites and decwrl. Each month I store the names of the hosts that are named that month, and for this report I used the past 10 months worth of data.

There are 33831 different sites in the Path lines of articles that arrived at decwrl in the last 10 months. There are 11277 different sites in the comp.mail.maps data, but comp.mail.maps tends to include only one or two machines for each organization, leaving the rest unmentioned. Also a large number of sites participate in USENET without participating in UUCP.

I believe that 34000 is the best estimate for the size of USENET. Because it is actually a measurement of the number of sites that have posted a message or that are on the path to a site that has posted a message, it will be slightly smaller than the number of sites that actually read netnews. Any site that believes it is not being counted can just ensure that it posts at least one message a year, so that it will be counted.

NUMBER OF USERS

The number of users at each site is determined in a site-specific fashion. Sometimes it is done by counting the number of user accounts that have shells and login directories. Sometimes it is done by counting the number of people who have logged in to the machine in some interval. Sometimes other techniques are used. This number is probably not very accurate--certainly not more accurate than to within a factor of two.

ESTIMATED TOTAL NUMBER OF PEOPLE WHO READ THIS GROUP, WORLDWIDE

There are two sources of error in this number. The number is computed by multiplying the number of people in the sample who actually read the group by the ratio of estimated network size to sample size. The estimated total can therefore be biased by errors in the network size estimate (see above) and also by errors in the determination of whether or not someone reads a group. Assuming that "reading a group" is roughly the same as "thumbing through a magazine", in that you don't necessarily have to read anything, but you have to browse through it and see what is there, then the measurement error will come primarily from inability to locate .newsrc files, which can either be protected or moved out of root directories. There is no way of measuring the effect on the measurements from unlocated .newsrc files, but it is not likely to be more than a few percent of the total news readers.

PROPAGATION: HOW MANY SITES RECEIVE THIS GROUP AT ALL

This number is the percent of the sites that are even receiving this newsgroup. The information necessary to compute propagation was not generated by early versions of the arbitron program, so the "basis" (number of sites) used to generate the Propagation figure is smaller than the "Sites in this sample" figure. A site's data will be used to compute propagation if either (a) it reports zero readers for at least one group, or (b) it is using an arbitron with an explicit version number that is high enough.

MESSAGES PER MONTH AND KILOBYTES PER MONTH

Traffic is measured at decwrl, in Palo Alto, California. Any message that has arrived at decwrl within the last "Traffic measurement interval" days is counted, regardless of when it was posted. Monthly rates are computed by taking the total traffic, dividing by the number of days in the traffic measurement interval, and multiplying by 30. Decwrl runs 2.10.3 news, which does not store the "Date-Received", "Relay-version" or "Posting-version" header lines; the amount of space occupied at your site might be higher, and the number of bytes transmitted between machines is probably higher. By definition this number is correct, because it is an exact measurement, but it may differ from the traffic at your site by as much as 15% due to timing differences and news version differences. Timing differences will be random, but will average out in the long run. News version differences will cause a systematic error that is additively uniform across all newsgroups, and which therefore does not significantly affect ratios.

If a message is crossposted to several groups simultaneously, it is charged only to the first-named group in the list. Note that this differs from the statistics posted from uunet every 2 weeks: the uunet data charge a message equally to every group that it is crossposted to.

CROSSPOSTING PERCENTAGE: WHAT FRACTION OF THE ARTICLES ARE CROSSPOSTED

"Crossposting" means to post the same article simultaneously in more than one newsgroup. In genuine "news" systems crossposting is implemented with Unix links and does not increase the storage or transmisison cost, though in some other systems crossposted articles are unbundled and must be stored and 6/7/13

transmitted separately.

The "crossposting percentage" is the percentage of the articles in this group that are crossposted to at least one other group. If every article in this group is crossposted, the percentage will be 100%; if none is crossposted, then the percentage will be 0%. The crossposting percentage figure does not take the size of the article into account, only the number of articles. Crossposting a 50,000-byte article or a 50-byte article both cause the same tally.

COST RATIO: DOLLARS PER MONTH PER READER

The most controversial field in the survey report is the "\$US per month per reader". It is the estimated number of dollars that are being spent on behalf of each reader, worldwide, on telephone and computer costs to transmit this newsgroup. The rate of \$.0025 per kilobyte is the same value used in the UUNET statistics reported biweekly. It is based on discussions among system administrators about the true cost of news transmission.

The cost ratio is computed as follows:

```
$US/month/reader = ($USPerMonthPerSite * numberOfSites) / numberOfReaders
$USPerMonthPersite = KBytesTrafficPerMonth * $USPerKByte * Propagation factor
$USPerKByte = 0.0025
```

Combining all these gives

```
$USPerMonthPersite =
    KBytesTrafficPerMonth * 0.0025
    = KBytesTrafficPerMonth / 400
```

Therefore:

```
$US/month/reader =
    (KBytesTrafficPerMonth * numberOfSites) / (400 * numberOfReaders)
```

The accuracy of this number is in fact better than the accuracy of the participation ratio, because the source of error--the network size estimate--is present both in the numerator and the denominator, and therefore cancels out. The primary source of bias in this number comes from the bias in

the "estimated number of readers, worldwide", which is described above. Treat this value as being accurate to within about 25%.

SITE PARTICIPATION

I would like to receive data from every site on USENET. The arbitron programs (posted comp.sources.d along with this report) work on news 2.9, 2.10.[1-3], 2.11, and on many versions of notesfiles.

Brian Reid
DEC Western Research Laboratory, Palo Alto CA
re...@decwrl.DEC.COM
{uunet,ucbvax,sun,pyramid,att}!decwrl!reid

6/7/13